2D full waveform inversion of shallow seismic Rayleigh waves

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN
von der Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

Dipl.-Geophys. Lisa Groos
aus Aalen

Tag der mündlichen Prüfung: 8. November 2013
Referent: Prof. Dr. Thomas Bohlen
Korreferent: Prof. Dr. Wolfgang Friederich

Contents

Chapter 1

Introduction

The inversion of shallow seismic Rayleigh waves is attractive for shallow geotechnical site investigations. Shallow seismic Rayleigh waves can be easily excited by a vertical hammer blow on the surface and recorded with vertical component or three-component geophones along a profile. Rayleigh waves have a high sensitivity to the S-wave velocity. Therefore, the analysis of shallow seismic Rayleigh waves excited by hammer blows can reveal information about the S-wave velocity structure in the first 10 m to 15 m of the subsurface. Rayleigh waves have a high signal-to-noise ratio compared to body waves since the wavefields excited by a hammer blow are dominated by surface waves. Furthermore, Rayleigh waves can be used to investigate sites with low-velocity zones which is impossible by an analysis of refracted body waves.

Established methods for the inversion of surface waves are the inversion of dispersion curves (e. g. Wathelet et al., 2004 or review article of Socco et al., 2010) or Fourier-Bessel expansion coefficients (Forbriger, 2001, 2003a,b). These methods assume 1D subsurface structures where the material properties vary only with depth. However, this assumption is not satisfied in many applications of practical relevance. To investigate sites where the subsurface properties change also along the profile approaches are applied where the recorded data sets are divided into subsets. Afterwards, a 1D subsurface model is derived from each subset. Finally, the 1D models are connected to a pseudo 2D subsurface model for interpretation (e. g. Bohlen et al., 2004). However, the lateral resolution of the obtained 2D subsurface models is limited because the profile length of the analyzed subsets cannot be chosen arbitrarily small. The length of the profiles determine the lowest frequency for which phase velocities can be estimated and therefore the maximum depth of the derived 1D models. O'Neill et al. (2008) for example discuss this problem for a subsurface structure containing a soft sinkhole. They conclude that the lateral extent of the sinkhole must be larger than half of the spread lengths of the analyzed subsets to allow a reliable interpretation. Furthermore, these concepts break down for data sets where no continuous dispersion curves can be extracted which is the case for strong lateral variations or 3D subsurface structures (see e. g. data set Lerchenberg in Forbriger, 2001).

To invert also for strong 2D and 3D structures full waveform inversion (FWI) must be applied. Many structures of practical relevance are 3D structures rather than 2D structures. Therefore, 3D FWI is required in a long-term view. However, 3D elastic FWI is still computationally too expensive for frequently repeat runs due to the huge number of forward problems that must be solved during FWI numerically (Butzer et al., 2013). Therefore, 2D FWI is applied in this thesis.

FWI has the potential to reconstruct high-resolution images of the subsurface by using the full information content of the recorded waveforms. There are different approaches for FWI discussed in literature. Waveform sensitivity kernels can be calculated by applying the first-order Born approximation which describe changes in the seismograms due to changes in the subsurface parameters (Schumacher & Friederich, 2013; Auras et al., 2013; Chen et al., 2007). These waveform sensitivity kernels can be used to solve the inverse problem. Another wide-spread method is FWI using the adjoint approach which was introduced into seismics by Tarantola (1984) for acoustic wave propagation. Later it was generalized to elastic (Mora, 1987) and viscoelastic wave propagation (Tarantola, 1988). FWI using the adjoint approach directly provides the gradient of the misfit function without explicitly calculating the waveform sensitivity kernels. Therefore, only gradient-based methods can be used to solve the optimization problem which have a slower convergence than optimization methods that use the second derivative of the misfit function provided by the kernels. However, the numerical implementation of the adjoint approach is very efficient. The suitability of the FWI approaches depends on the specific problem. For data sets with a large number of receivers and a small number of sources the adjoint approach is preferable with respect to computation time (Chen et al., 2007). It is therefore mainly applied in active source seismics. A third approach is the application of global search algorithms in FWI. However, this approach can be only applied to problems with small parameter spaces (Auras et al., 2013). It is therefore more suited for material testing using ultrasonic measurements. Mosca et al. (2013) also successfully applied FWI using a global search algorithm to a shallow seismic field data set. The obtained model could be used as initial model in a subsequent FWI using a local optimization approach. An advantage of global search strategies is that they implicitly provide a resolution analysis since they search for optimal models in the whole parameter space.

FWI using the adjoint approach was first formulated in time-domain (e. g. Tarantola, 1984; Mora, 1987) and later in frequency-domain (e. g. Pratt & Worthington, 1990; Pratt, 1990, 1999). Calculation of the gradient of the misfit function in the frequency-domain is more efficient concerning memory consumption because the gradient is calculated only for a few discrete frequencies or even only one discrete frequency. However, the frequencies must be chosen carefully to ensure convergence to the global minimum (Sirgue & Pratt, 2004). With respect to surface waves Romdhane et al. (2011) report that they have to invert at least three discrete frequencies simultaneously to reconstruct the subsurface structure. They perform a synthetic study where they apply 2D elastic FWI in the frequency-domain to Rayleigh waves to image the shallow structure.

I apply a 2D FWI code developed by Köhn (2011). It uses the time-domain adjoint method. A time-domain finite-difference scheme is implemented as forward solver (Virieux, 1986; Levander, 1988; Bohlen, 2002). A collection of other forward solvers which are used in FWI is given by Virieux & Operto (2009) and Fichtner (2011).

There are already a few applications of FWI to shallow seismic surface waves. Romdhane et al. (2011) performed a synthetic study where they successfully inverted Rayleigh waves in a complex geological setting with topography. They apply a 2D elastic FWI in the frequency-domain. They use elastic forward modeling and assume an a priori known source wavelet in their tests.

Tran & McVay (2012) show a synthetic as well as a field data application where they invert surface waves up to 20 Hz to image the subsurface structure down to 20 m depth. They apply elastic FWI in the time-domain including elastic forward modeling. In their synthetic tests they successfully reconstruct low and high velocity layers in the 2D S-wave velocity model for both, noise-free and noisy data. Furthermore, they resolve a low velocity layer in the field data example. The S-wave velocity model obtained in the field data application is compared to a standard penetration test (SPT) analysis. The results appear to be consistent. Bretaudeau et al. (2013) apply 2D elastic FWI on a layered medium with an inclusion for synthetic data and laboratory data which are obtained by small-scale physical modeling. The used inversion scheme is formulated in the frequency-domain and considers anelastic wave propagation. Artefacts occur in the inversion result of the data recorded in the laboratory due to inaccuracies in the estimated source wavelets and singularities of the gradient close to the sources. Nevertheless, the inclusion is reconstructed.

These first successful applications show the potential of FWI applied to Rayleigh waves especially to image 2D structures. However, application of FWI to shallow seismic Rayleigh waves is in an early stage. Published studies use different forward solvers, optimization methods, and preprocessing techniques which account for the effect of the unknown source wavelet or the difference in 3D and 2D spreading between recorded data and synthetics. No common concept, especially for the inversion of field data, evolved yet.

The main objective of this work is the application of 2D FWI to shallow seismic field data. Main challenges in the inversion of field data are the 3D/2D transformation that must be applied to the field data prior to the application of 2D FWI. Furthermore, the unknown source wavelet must be estimated and the effects of anelastic damping must be considered. In my studies I partly address these problems by tests with synthetic data. Finally, I apply FWI to a field data set.

This thesis is organized as follows.

Chapter 2 discusses the fundamentals of 2D FWI using the adjoint approach. The gradient of the misfit function which is used for solving the inverse problem in the later chapters is derived considering viscoelastic wave propagation. Furthermore, some details about the numerical implementation are described. Finally, appropriate 3D/2D transformation techniques for shallow seismic Rayleigh waves are introduced. Field data are in general acquired with point sources. In 2D FWI a 2D forward solver is applied which uses implicitly line sources. Therefore, the recorded point-source seismograms must be transformed to corresponding line-source seismograms prior to the application of 2D FWI.

FWI aims to find a subsurface model that predicts the full recorded seismograms. Therefore, the forward problem is solved many times and the synthetic data are compared to the field data. An improvement of the reconstructed subsurface model is in general achieved when the data misfit between the synthetic data and the field data decreases. Therefore, a crucial prerequisite for a successful FWI is an appropriate forward solver. The simulated wave propagation must reproduce all significant physical effects that are contained in the field data. Therefore, I perform benchmark tests of the forward solver implemented in the 2D FWI code. A finite-difference time-

domain scheme is used for forward modeling. This scheme and the results of the benchmark tests are presented in Chapter 3.

Chapter 4 deals with synthetic tests where FWI is applied to synthetic data which are used as pseudo-observed data in the inversion. These tests are called reconstruction tests because the true subsurface model is known and the reconstruction of this model by FWI can be directly evaluated. The effects of anelastic damping as well as of the influence of the P-wave velocity model on the reconstruction of the S-wave velocity model are studied by means of reconstruction tests. Both, monoparameter and multiparameter tests are performed.

The application of FWI to field data is discussed in Chapter 5. The analyzed data set was acquired on a predominantly depth dependent structure in unconsolidated sediments. A 1D subsurface model is derived from the field data set by a joint inversion of Fourier-Bessel expansion coefficients and first arrival P-wave travel times for the purpose of comparison. The preprocessing applied to the field data prior to the 2D FWI is described as well as the 2D subsurface model which is obtained by 2D FWI. Finally, the 1D and 2D S-wave velocity models are compared.

Finally, Chapter 6 summarizes the main results of the work.

Chapter 2

Basic principles

2.1 Full waveform inversion using the adjoint approach

Full waveform inversion (FWI) aims to use the full information content of the seismic waveforms to reconstruct high-resolution images of the subsurface. In my work I applied an FWI based on the adjoint method. This method was firstly introduced into seismics by Tarantola (1984) for the acoustic case and was later formulated for more general cases (e. g., elastic case by Mora, 1987 and viscoelastic case by Tarantola, 1988). During an FWI a previously defined misfit function is minimized. With the adjoint method the gradient of the misfit function can be calculated numerically very efficiently. The optimization problem is then solved using a steepest descent or conjugate gradient method. The idea is to update the model parameters in the opposite direction of the gradient of the misfit function to reduce the misfit. As the calculation of the gradient of the misfit function is one of the basics of the FWI method applied in this theses, it is discussed in detail in this section.

2.1.1 Wave equation

The wave propagation in a volume V can be described by (Aki & Richards, 2002; Tarantola, 1988)

$$\hat{L}_{kl}(\vec{x})\, u_l(\vec{x},t) \;=\; f_k(\vec{x},t) \tag{2.1}$$

with the operator

$$\hat{L}_{kl}(\vec{x}) \;=\; \rho(\vec{x})\delta_{kl}\frac{\partial^2}{\partial t^2} - \frac{1}{2}\frac{\partial}{\partial x_j}\Psi_{jkml}(\vec{x},t) * \frac{\partial}{\partial x_m} - \frac{1}{2}\frac{\partial}{\partial x_j}\Psi_{jkln}(\vec{x},t) * \frac{\partial}{\partial x_n} \tag{2.2}$$

where $\vec{x} \in V$ and $t \in [t_0, \infty)$. u_l is the l-component of the particle displacement vector, ρ the density and f_k the k-component of an external force density. Ψ is a tensor of rank four that contains the time derivatives of the relaxation functions which describe the rheology of the medium. They are also named rate of relaxation functions (Tarantola, 1988). $*$ denotes a convolution. δ_{ij} is the Kronecker symbol with $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$. I assume that the material properties do not change with time. Therefore, the wave equation is time invariant. (Ψ_{jkml} is also assumed to be time invariant. The argument t occurs due to the convolution.) I use the Einstein summation convention which implies the summation over repeated indices within a product. In some later equations the summations of the convention are partly written explicitly to emphasize their physical meaning.

Due to the symmetry of the strain tensor and the following symmetry $\Psi_{jkln} = \Psi_{jknl}$ Equation (2.2) can be written as

$$\hat{L}_{kl}(\vec{x}) \;=\; \rho(\vec{x})\delta_{kl}\frac{\partial^2}{\partial t^2} - \frac{\partial}{\partial x_j}\Psi_{jkml}(\vec{x},t) * \frac{\partial}{\partial x_m} \qquad \text{with} \qquad \left[\hat{L}_{kl}\right] = 1\frac{\text{kg}}{\text{m}^3\text{s}^2}. \tag{2.3}$$

Due to the convolution in the differential operator in Equation (2.3) also Equation (2.1) implicitly involves a convolution.

To uniquely define the problem given in Equation (2.1) initial and boundary conditions must be defined. I assume the initial conditions

$$u_l(\vec{x},t) = 0 \quad \text{for} \quad t \leq t_0 \tag{2.4a}$$

$$\frac{\partial u_l(\vec{x},t)}{\partial t} = 0 \quad \text{for} \quad t \leq t_0. \tag{2.4b}$$

As boundary condition one can for example use a free surface condition (Tarantola, 1988) with

$$n_j(\vec{\xi})\; \Psi_{jkml}(\vec{\xi},t) * \frac{\partial u_l}{\partial x_m}(\vec{\xi},t) = 0 \quad \text{for} \quad \vec{\xi} \in S \tag{2.5}$$

where S is the surface of the volume V, $\vec{\xi}$ is a point on this surface and n_j is the outward pointing unit normal vector of the surface S.

Elastic material

For a purely elastic medium it is (Tarantola, 1988)

$$\Psi_{jkml}(\vec{x},t) \;=\; M_{jkml}(\vec{x})\delta(t) \tag{2.6}$$

with the tensor M_{jkml} of the elastic constants and the delta function $\delta(t)$. For this reason the differential operator of Equation (2.3) simplifies to

$$\hat{L}_{kl}(\vec{x}) \;=\; \rho(\vec{x})\delta_{kl}\frac{\partial^2}{\partial t^2} - \frac{\partial}{\partial x_j}M_{jkml}(\vec{x})\frac{\partial}{\partial x_m}. \tag{2.7}$$

For a homogeneous isotropic medium the elements of the tensor of elastic constants are given by

$$M_{jkml} = \delta_{jk}\delta_{ml}\lambda + (\delta_{jm}\delta_{kl} + \delta_{jl}\delta_{km})\mu \qquad \text{with} \qquad \left[M_{jkml}\right] = [\lambda] = [\mu] = 1\frac{\text{kg}}{\text{m s}^2} = 1\text{Pa} \tag{2.8}$$

and with the Lamé parameters λ and μ.

Viscoelastic material - Generalized standard linear solid

The 2D FWI code DENISE which I applied uses a generalized standard linear solid (GSLS) as rheological model. Therefore, I focus on the description of this rheological model. In this section I will only review the equations needed for the following derivation of the gradient of the

misfit function. A more detailed overview about a GSLS is given in Appendix A. Considering an isotropic and isothermal medium with a rheology described by a GSLS the tensor elements Ψ_{jkml} are expressed by

$$\Psi_{jkml} = \delta_{jk}\delta_{ml}\left(\dot{\psi}_p(t) - 2\dot{\psi}_s(t)\right) + \left(\delta_{jm}\delta_{kl} + \delta_{jl}\delta_{km}\right)\dot{\psi}_s(t), \tag{2.9}$$

where $\dot{\psi} = \partial\psi/\partial t$ is a time derivative and

$$\psi_p(t) = \pi_r\left[1 + \sum_{l=1}^{L}\left(\frac{\tau_{\varepsilon l}^p}{\tau_{\sigma l}} - 1\right)e^{-t/\tau_{\sigma l}}\right]H(t) = \hat{\psi}_p(t)H(t) \tag{2.10a}$$

$$\psi_s(t) = \mu_r\left[1 + \sum_{l=1}^{L}\left(\frac{\tau_{\varepsilon l}^s}{\tau_{\sigma l}} - 1\right)e^{-t/\tau_{\sigma l}}\right]H(t) = \hat{\psi}_s(t)H(t) \tag{2.10b}$$

are the relaxation functions of the L relaxation mechanisms for P and S-waves, respectively (Liu et al., 1976; Carcione et al., 1988; Robertsson et al., 1994). The relaxation functions describe the evolution of the stress caused by a unit step function in the strain (Emmerich & Korn, 1987). π_r is the relaxed P-wave modulus (modulus for compressional waves) and μ_r the relaxed shear modulus. The relaxed moduli are related to the material properties in case of very low frequencies ($\omega \to 0$ where ω is the angular frequency, see Appendix A.2.2). The P-wave modulus is $\pi = \lambda + 2\mu$. $\tau_{\varepsilon l}^p$ and $\tau_{\varepsilon l}^s$ are the retardation times of the strain of the l-th relaxation mechanism for P and S-waves and $\tau_{\sigma l}$ the corresponding relaxation times for the stress.

As a further simplification the τ method is used. It was suggested by Blanch et al. (1995) to reduce the number of variables that describe a GSLS. The L retardation times $\tau_{\varepsilon l}^p$ and $\tau_{\varepsilon l}^s$, respectively, are replaced by dimensionless variables τ^p and τ^s with

$$\tau^p = \frac{\tau_{\varepsilon l}^p}{\tau_{\sigma l}} - 1 \tag{2.11a}$$

$$\tau^s = \frac{\tau_{\varepsilon l}^s}{\tau_{\sigma l}} - 1. \tag{2.11b}$$

Generally, the τ method is used to approximate a frequency independent quality factor Q in a limited frequency interval (Blanch et al., 1995). The relaxation times or relaxation frequencies $f_l = 1/(2\pi\tau_{\sigma l})$, respectively, determine, in which frequency interval this approximation is valid. Generally, this frequency interval is equal for P and S-waves (Bohlen, 1998). Therefore, the same relaxation times are used in ψ_p and ψ_s in Equation (2.10). The values of τ^p and τ^s are related to the value of the quality factors for P and S-waves (see Appendix A.2.4).

In the limit of high frequencies ($\omega \to \infty$) the material properties are described by the unrelaxed moduli (see Appendix A.2.2)

$$\pi_u = \pi_r\left(1 + \sum_{l=1}^{L}\left(\frac{\tau_{\varepsilon l}^p}{\tau_{\sigma l}} - 1\right)\right) = \pi_r\left(1 + L\tau^p\right) = \hat{\psi}_p(0) \tag{2.12a}$$

$$\mu_u = \mu_r\left(1 + \sum_{l=1}^{L}\left(\frac{\tau_{\varepsilon l}^s}{\tau_{\sigma l}} - 1\right)\right) = \mu_r\left(1 + L\tau^s\right) = \hat{\psi}_s(0). \tag{2.12b}$$

Applying a time derivative to Equation (2.10) and inserting Equation (2.11) leads to (Carcione et al., 1988)

$$\dot{\psi}_p(t) = \hat{\psi}_p(t)\delta(t) - \Phi_p H(t) \tag{2.13a}$$
$$\dot{\psi}_s(t) = \hat{\psi}_s(t)\delta(t) - \Phi_s H(t), \tag{2.13b}$$

with the response functions of the medium

$$\Phi_p(t) = \sum_{l=1}^{L} \varphi_{p_l}(t) \tag{2.14a}$$
$$\Phi_s(t) = \sum_{l=1}^{L} \varphi_{s_l}(t) \tag{2.14b}$$

with

$$\varphi_{p_l}(t) = \frac{\pi_r}{\tau_{\sigma l}} \tau^p e^{-t/\tau_{\sigma l}} = \frac{\pi_u}{\tau_{\sigma l}} \frac{\tau^p}{1 + L\tau^p} e^{-t/\tau_{\sigma l}} \tag{2.15a}$$
$$\varphi_{s_l}(t) = \frac{\mu_r}{\tau_{\sigma l}} \tau^s e^{-t/\tau_{\sigma l}} = \frac{\mu_u}{\tau_{\sigma l}} \frac{\tau^s}{1 + L\tau^s} e^{-t/\tau_{\sigma l}}. \tag{2.15b}$$

Using the τ method the anelastic properties of a medium can be described by the unrelaxed moduli π_u and μ_u, the τ values for P and S-waves and the relaxation times $\tau_{\sigma l}$.

2.1.2 Green's function

The Green's function $G_{ln}(\vec{x}, t; \vec{x}', t')$ is defined as the solution of the differential equation

$$\hat{L}_{kl}(\vec{x}) G_{ln}(\vec{x}, t; \vec{x}', t') = \delta_{kn} \delta^3(\vec{x} - \vec{x}') \delta(t - t') \tag{2.16}$$

with a point source at location $\vec{x}'$ and time t' in direction of the n-th component. $\delta^3(\vec{x})$ and $\delta(t)$ are Dirac's delta functions with physical units of $\frac{1}{\mathrm{m}^3}$ and $\frac{1}{\mathrm{s}}$, respectively. The physical unit of the Green's function is $[G_{ln}] = 1\frac{\mathrm{m}}{\mathrm{N \cdot s}}$. The particle displacement and solution of Equation (2.1) for an arbitrary force density $f_n(\vec{x}', t')$ in $\frac{\mathrm{N}}{\mathrm{m}^3}$ is

$$u_l(\vec{x}, t) = \int_V \int_{-\infty}^{\infty} G_{ln}(\vec{x}, t; \vec{x}', t') \, f_n(\vec{x}', t') \, \mathrm{d}t' \, \mathrm{d}^3\vec{x}' \qquad \text{with} \quad [u_l] = \mathrm{m}. \tag{2.17}$$

Due to the initial conditions defined in Equation (2.4) the Green's function is causal ($G_{ln}(\vec{x}, t; \vec{x}', t') = 0$ for $t < t'$). If we additionally assume $f_n(\vec{x}', t') = 0$ for $t' < 0$ the integration limits in Equation (2.17) can be adjusted to

$$u_l(\vec{x}, t) = \int_V \int_0^t G_{ln}(\vec{x}, t; \vec{x}', t') \, f_n(\vec{x}', t') \, \mathrm{d}t' \, \mathrm{d}^3\vec{x}'. \tag{2.18}$$

Furthermore, it is (Tarantola, 1988)

$$G_{ln}(\vec{x}, t; \vec{x}', t') = G_{ln}(\vec{x}, t - t'; \vec{x}', 0) \tag{2.19}$$

since I assume that the material properties do not change with time. Additionally, the source-receiver reciprocity

$$G_{ln}(\vec{x}, t - t'; \vec{x}', 0) = G_{nl}(\vec{x}', t - t'; \vec{x}, 0) \tag{2.20}$$

becomes valid (e. g. Tarantola, 1988; Carcione, 2001).

2.1.3 Partial derivatives

In the following a reference model is considered which is defined by the vector of model parameters $\vec{m}_0$. The wave propagation in this reference model is described by the differential operator $\hat{L}_{kl}^{(0)}$ of Equation (2.1) and the Green's function solving the differential equation is $G_{ln}^{(0)}$. To apply scattering theory the material properties of the reference medium are perturbed by $\delta\vec{m}$ which leads to a new model

$$\vec{m} = \vec{m}_0 + \delta\vec{m}. \tag{2.21}$$

The differential operator depends on the model parameters. Due to the change in the material properties the differential operator

$$\hat{L}_{kl}(\vec{x},\vec{m}) = \hat{L}_{kl}^{(0)}(\vec{x},\vec{m}_0) + \hat{L}_{kl}^{(1)}(\vec{x},\delta\vec{m}) \tag{2.22}$$

changes and deviates by $\hat{L}_{kl}^{(1)}$ from the differential operator of the reference model. In addition, the Green's function changes too and can be described by

$$G_{ln}(\vec{x},t-t';\vec{x}',0) = G_{ln}^{(0)}(\vec{x},t-t';\vec{x}',0) + G_{ln}^{(1)}(\vec{x},t-t';\vec{x}',0). \tag{2.23}$$

The partial derivatives of the synthetic data in the perturbed model with respect to the model parameters are (Forbriger, 2009)

$$\begin{aligned}\frac{\partial}{\partial m_k}\bigg|_{\vec{m}=\vec{m}_0} G_{ln}(\vec{x},t-t';\vec{x}',0) = \frac{\partial}{\partial\,\delta m_k}\bigg|_{\delta\vec{m}=0} G_{ln}(\vec{x},t-t';\vec{x}',0) = \\ -\int_V\int_{-\infty}^{\infty} G_{lm}^{(0)}(\vec{x},t-t'';\vec{x}'',0)\,\frac{\partial}{\partial\,\delta m_k}\bigg|_{\delta\vec{m}=0} \hat{L}_{mi}^{(1)}(\vec{x}'',\delta\vec{m})\,G_{in}^{(0)}(\vec{x}'',t''-t';\vec{x}',0)\,\mathrm{d}t''\,\mathrm{d}^3\vec{x}'',\end{aligned} \tag{2.24}$$

where m_k is the k-th model parameter and $\partial/\partial m_k = \partial/\partial\,\delta m_k$ is valid due to Equation (2.21). In a physical interpretation the Green's function $G_{in}^{(0)}(\vec{x}'',t''-t';\vec{x}',0)$ describes the propagation of the wavefield in the reference medium from a source at $\vec{x}'$ and time t' to a scatterer in the medium at $\vec{x}''$. The scattering of the wavefield is described by the perturbed differential operator $\hat{L}_{mi}^{(1)}$. Finally, the Green's function $G_{lm}^{(0)}(\vec{x},t-t'';\vec{x}'',0)$ describes the propagation from the scatterer to a point $\vec{x}$ in the medium.

2.1.4 Gradient of the misfit function

For the derivation of the gradient of the misfit function I closely follow the work of Tromp et al. (2005) as well as Bozdağ et al. (2011) which discuss the derivation for the elastic case. However, I consider in the following viscoelastic wave propagation.

As defined in Equation (2.21) the vector of the model parameters in the reference model is denoted by $\vec{m}_0$ and the perturbation of the model parameters is denoted by $\delta\vec{m}$. The **observed particle displacement** of the j-th component at point $\vec{x}_r$ is $d_j(\vec{x}_r,t)$ and the corresponding **synthetic particle displacement** is $s_j(\vec{x}_r,t,\vec{m})$. The particle displacement field in the reference model is $s_{0_j}(\vec{x}_r,t,\vec{m}_0)$ which means that it is $s_j(\vec{x}_r,t,\delta\vec{m}=0) = s_{0_j}(\vec{x}_r,t,\vec{m}_0)$.

The misfit is defined by

$$E = \sum_{r=1}^{n_r} \sum_{j=1}^{n_c} \int_0^T g\left[d_j(\vec{x}_r,t), s_j(\vec{x}_r,t,\vec{m})\right] \mathrm{d}t \tag{2.25}$$

where g quantifies the discrepancy between the observed and the synthetic data at receiver r at point $\vec{x}_r$. n_r is the number of receivers, n_c is the number of components, and T is the recording time. For convenience, I assume a data set with only one source. If a data set with more sources is considered the misfit definition in Equation (2.25) includes an additional sum over the sources. The physical unit of the misfit is $[E] = [g]$ s.

The negative gradient of the misfit function points into the steepest descent direction in the model space and provides the direction in which the misfit decreases strongest. To minimize the misfit function the model update $\Delta\vec{m}$ in a steepest descent method is calculated by

$$\Delta m_k = -\alpha \left.\frac{\partial E}{\partial m_k}\right|_{\vec{m}=\vec{m}_0} \tag{2.26}$$

with a factor α that determines the step length. The factor α must be scaled appropriately to provide the correct physical units. This is discussed in more detail in Section 2.3.2. Equation (2.25) together with the chain rule yields the gradient of the misfit function

$$\left.\frac{\partial E}{\partial m_k}\right|_{\vec{m}=\vec{m}_0} = \sum_{r=1}^{n_r} \sum_{j=1}^{n_c} \int_{t=0}^{T} \partial_{s_j} g \left.\frac{\partial s_j(\vec{x}_r,t,\vec{m})}{\partial m_k}\,\mathrm{d}t\right|_{\vec{m}=\vec{m}_0} \tag{2.27}$$

where $\partial_{s_j} g$ denotes the Fréchet derivative following the work of Bozdağ et al. (2011).

Using Equation (2.17) together with Equation (2.24) the partial derivatives of the synthetic seismograms with respect to the model parameters can be expressed by (see Appendix B.1 for details)

$$\left.\frac{\partial}{\partial m_k}\right|_{\vec{m}=\vec{m}_0} s_j(\vec{x}_r,t) = -\int_V \int_0^t G^{(0)}_{jm}(\vec{x}_r,t-t'';\vec{x}'',0) \left.\frac{\partial}{\partial \delta m_k}\right|_{\delta\vec{m}=0} \hat{L}^{(1)}_{mi}(\vec{x}'',\delta\vec{m}) s_{0_i}(\vec{x}'',t'')\,\mathrm{d}t''\mathrm{d}^3\vec{x}''. \tag{2.28}$$

Inserting these partial derivatives into Equation (2.27) yields

$$\begin{aligned}\left.\frac{\partial E}{\partial m_k}\right|_{\vec{m}=\vec{m}_0} = &-\sum_{r=1}^{n_r}\sum_{j=1}^{n_c}\int_{t=0}^{T} \partial_{s_j} g(\vec{x}_r,t,\vec{m})\big|_{\vec{m}=\vec{m}_0} \\ &\times \int_V \int_{t''=0}^{t} G^{(0)}_{jm}(\vec{x}_r,t-t'';\vec{x}'',0) \left.\frac{\partial}{\partial \delta m_k}\right|_{\delta\vec{m}=0} \hat{L}^{(1)}_{mi}(\vec{x}'',\delta\vec{m}) s_{0_i}(\vec{x}'',t'')\,\mathrm{d}t''\mathrm{d}^3\vec{x}''\mathrm{d}t.\end{aligned} \tag{2.29}$$

In the next step the order of time integrals is interchanged which leads to a change in the integration limits. This is sketched in Figure 2.1. For the gradient of the misfit function one obtains

$$\begin{aligned}\left.\frac{\partial E}{\partial m_k}\right|_{\vec{m}=\vec{m}_0} = -\int_V \int_{t''=0}^{T} &\underbrace{\sum_{r=1}^{n_r}\sum_{j=1}^{n_c}\int_{t=t''}^{T} G^{(0)}_{jm}(\vec{x}_r,t-t'';\vec{x}'',0)\,\partial_{s_j} g(\vec{x}_r,t,\vec{m})\big|_{\vec{m}=\vec{m}_0}\,\mathrm{d}t}_{\Phi_m(\vec{x}'',t'')} \\ &\times \left.\frac{\partial}{\partial \delta m_k}\right|_{\delta\vec{m}=0} \hat{L}^{(1)}_{mi}(\vec{x}'',\delta\vec{m}) s_{0_i}(\vec{x}'',t'')\,\mathrm{d}t''\mathrm{d}^3\vec{x}''\end{aligned} \tag{2.30}$$

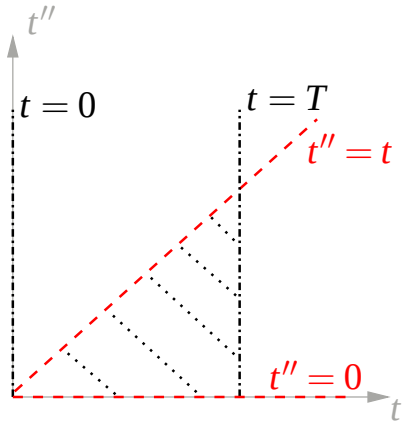


Figure 2.1: Sketch to demonstrate the determination of integral limits after a change of the integration order.

where the wavefield $\Phi_m(\vec{x}'', t'')$ is discussed in more detail in the following. Using the source-receiver reciprocity (Equation 2.20) it can be expressed by

$$\Phi_m(\vec{x}'', t'') = \sum_{r=1}^{n_r} \sum_{j=1}^{n_c} \int_{t=t''}^{T} G_{mj}^{(0)}(\vec{x}'', t - t''; \vec{x}_r, 0)\, \partial_{s_j} g(\vec{x}_r, t, \vec{m})\big|_{\vec{m}=\vec{m}_0}\, \mathrm{d}t. \tag{2.31}$$

The substitution $t = T - \hat{t}$ or $\hat{t} = T - t$, respectively, reverts the time axis and shifts the origin of the time axis to T. It is

$$\mathrm{d}t = (-1)\,\mathrm{d}\hat{t}$$

and the limits of the integral after the substitution are

$$\hat{t}_1 = T - t_1 = T - t'' \qquad \text{and} \qquad \hat{t}_2 = T - t_2 = 0.$$

Applying the substitution and reverting the integration limits results in

$$\Phi_m(\vec{x}'', t'') = \sum_{r=1}^{n_r} \sum_{j=1}^{n_c} \int_{\hat{t}=0}^{T-t''} G_{mj}^{(0)}(\vec{x}'', T - \hat{t} - t''; \vec{x}_r, 0)\, \partial_{s_j} g(\vec{x}_r, T - \hat{t}, \vec{m})\big|_{\vec{m}=\vec{m}_0}\, \mathrm{d}\hat{t}. \tag{2.32}$$

Now the so-called adjoint sources (Tromp et al., 2005; Bozdağ et al., 2011) can be defined by

$$f_j^\dagger(\vec{x}, \hat{t}) = \sum_{r=1}^{n_r} \partial_{s_j} g(\vec{x}_r, T - \hat{t}, \vec{m})\big|_{\vec{m}=\vec{m}_0}\, \delta(\vec{x} - \vec{x}_r) \qquad \text{with} \qquad \left[f_j^\dagger\right] = \frac{[g]}{\mathrm{m}^4} \tag{2.33}$$

which results in

$$\Phi_m(\vec{x}'', t'') = \sum_{j=1}^{n_c} \int_{\hat{t}=0}^{T-t''} \int_V G_{mj}^{(0)}(\vec{x}'', T - \hat{t} - t''; \vec{x}, 0)\; f_j^\dagger(\vec{x}, \hat{t})\; \mathrm{d}^3\vec{x}\; \mathrm{d}\hat{t}. \tag{2.34}$$

Again we can use the concept of Green's function (Equation 2.17) to define the adjoint wavefield

$$s_m^\dagger(\vec{x}'', \tau) = \sum_{j=1}^{n_c} \int_0^{\tau} \int_V G_{mj}^{(0)}(\vec{x}'', \tau - \hat{t}; \vec{x}, 0)\; f_j^\dagger(\vec{x}, \hat{t})\; \mathrm{d}^3\vec{x}\; \mathrm{d}\hat{t} \tag{2.35}$$

which is excited by the adjoint sources and which is related to the wavefield Φ by

$$\Phi_m(\vec{x}'', t'') = s_m^\dagger(\vec{x}'', T - t''). \tag{2.36}$$

Due to the summation over the receivers in the definition of the adjoint sources $f_j^\dagger$ (Equation 2.33) the adjoint wavefield can be calculated with only one forward modeling independent of the number of used receivers. I want to mention here that the adjoint wavefield is not a displacement wavefield. Its physical unit is $1\frac{\mathrm{s}^2[g]}{\mathrm{kg\ m}}$. The adjoint wavefield is also often called backpropagated wavefield which is related to the time reversal applied in this derivation. For the adjoint wavefield $\Phi_m(\vec{x}'', t'')$ the conditions of Equation (2.4) act as final conditions which results in

$$\Phi_m(\vec{x}'', t'') = 0 \quad \text{for} \quad t'' \geq T \tag{2.37a}$$

$$\frac{\partial \Phi_m(\vec{x}'', t'')}{\partial t''} = 0 \quad \text{for} \quad t'' \geq T. \tag{2.37b}$$

In practice, the adjoint wavefield is obtained by calculating the adjoint sources via Equation (2.33) and afterwards run a forward propagation where the adjoint sources act simultaneously as sources at the receiver positions.
Using Equation (2.36) in Equation (2.30) the gradient of the misfit function is

$$\frac{\partial E}{\partial m_k} = -\int_V \int_{t''=0}^{T} s_m^\dagger(\vec{x}'', T - t'') \left. \frac{\partial}{\partial \delta m_k} \right|_{\delta \vec{m}=0} \hat{L}_{mi}^{(1)}(\vec{x}'', \delta \vec{m}) s_{0_i}(\vec{x}'', t'') \, \mathrm{d}t'' \, \mathrm{d}^3 \vec{x}''. \tag{2.38}$$

In a next step the perturbation of the differential operator (see Equation 2.3) is inserted and integration by parts is applied (detailed calculation in Appendix B.2). Denoting the perturbations in the model parameters with $\delta\rho$ and $\delta\Psi$ it is

$$\begin{aligned} \frac{\partial E}{\partial m_k} = \int_V \int_{t''=0}^{T} & \frac{\partial s_m^\dagger(\vec{x}'', T - t'')}{\partial t''} \frac{\partial \delta\rho(\vec{x}'')}{\partial \delta m_k} \delta_{mi} \frac{\partial s_{0_i}(\vec{x}'', t'')}{\partial t''} \\ & - \frac{\partial s_m^\dagger(\vec{x}'', T - t'')}{\partial x_j''} \frac{\partial \delta\Psi_{jmni}(\vec{x}'', t'')}{\partial \delta m_k} * \frac{\partial s_{0_i}(\vec{x}'', t'')}{\partial x_n''} \, \mathrm{d}t'' \, \mathrm{d}^3 \vec{x}'' \end{aligned} \tag{2.39}$$

where I omitted that the partial derivative of $\hat{L}_{mi}^{(1)}(\vec{x}'', \delta \vec{m})$ must be evaluated at $\delta \vec{m} = 0$. Since the perturbation $\hat{L}_{mi}^{(1)}$ of the differential operator depends only linearly on the perturbation of the material parameters the derivatives $\partial \hat{L}_{mi}^{(1)} / \partial \delta m_k$ are independent of the perturbations of the model parameters $\delta \vec{m}$.

Elastic material

If an isotropic elastic material is assumed $\delta\Psi$ is given by Equations (2.6) and (2.8). Inserting in Equation (2.39) results in

$$\begin{aligned} \frac{\partial E}{\partial m_k} = \int_V \int_{t''=0}^{T} & \frac{\partial s_m^\dagger(\vec{x}'', T - t'')}{\partial t''} \frac{\partial \delta\rho(\vec{x}'')}{\partial \delta m_k} \delta_{mi} \frac{\partial s_{0_i}(\vec{x}'', t'')}{\partial t''} \\ & - \frac{\partial s_m^\dagger(\vec{x}'', T - t'')}{\partial x_j''} \frac{\partial \left(\delta_{jm}\delta_{ni}\delta\lambda(\vec{x}'') + (\delta_{jn}\delta_{mi} + \delta_{ji}\delta_{mn})\delta\mu(\vec{x}'') \right)}{\partial \delta m_k} \frac{\partial s_{0_i}(\vec{x}'', t'')}{\partial x_n''} \, \mathrm{d}t'' \, \mathrm{d}^3 \vec{x}''. \end{aligned} \tag{2.40}$$

The 2D FWI implementation used in this work is a grid based method because the finite-difference method is used for forward modeling. Therefore, in the elastic case the model parameters are the density ρ and the Lamé parameters λ and μ at the grid points of the inversion grid which in this work coincides with the grid used for the forward simulation. As the derivative of the differential operator with respect to a material parameter is only non-zero at the grid point of this model parameter the integration in Equation (2.40) over the volume V is limited to the corresponding grid cell. In the 2D case the gradient of the misfit function with respect to the different model parameters is

$$\frac{\partial E}{\partial \rho_k} = \int_0^T \left(\frac{\partial s_1^\dagger(\vec{x}'', T-t'')}{\partial t''} \frac{\partial s_{0_1}(\vec{x}'', t'')}{\partial t''} + \frac{\partial s_2^\dagger(\vec{x}'', T-t'')}{\partial t''} \frac{\partial s_{0_2}(\vec{x}'', t'')}{\partial t''} \right) \mathrm{d}t'' \Delta x''^3 \tag{2.41}$$

$$\frac{\partial E}{\partial \lambda_k} = -\int_0^T \left(\frac{\partial s_1^\dagger}{\partial x_1''} \frac{\partial s_{0_1}}{\partial x_1''} + \frac{\partial s_1^\dagger}{\partial x_1''} \frac{\partial s_{0_2}}{\partial x_2''} + \frac{\partial s_2^\dagger}{\partial x_2''} \frac{\partial s_{0_1}}{\partial x_1''} + \frac{\partial s_2^\dagger}{\partial x_2''} \frac{\partial s_{0_2}}{\partial x_2''} \right) \mathrm{d}t'' \Delta x''^3 \tag{2.42a}$$

$$= -\int_0^T \left(\frac{\partial s_1^\dagger(\vec{x}'', T-t'')}{\partial x_1''} + \frac{\partial s_2^\dagger(\vec{x}'', T-t'')}{\partial x_2''} \right) \cdot \left(\frac{\partial s_{0_1}(\vec{x}'', t'')}{\partial x_1''} + \frac{\partial s_{0_2}(\vec{x}'', t'')}{\partial x_2''} \right) \mathrm{d}t'' \Delta x''^3 \tag{2.42b}$$

$$\frac{\partial E}{\partial \mu_k} = -\int_0^T \left[2\frac{\partial s_1^\dagger}{\partial x_1''} \frac{\partial s_{0_1}}{\partial x_1''} + \frac{\partial s_1^\dagger}{\partial x_2''} \frac{\partial s_{0_2}}{\partial x_1''} + \frac{\partial s_1^\dagger}{\partial x_2''} \frac{\partial s_{0_1}}{\partial x_2''} + \frac{\partial s_2^\dagger}{\partial x_1''} \frac{\partial s_{0_1}}{\partial x_2''} + \frac{\partial s_2^\dagger}{\partial x_1''} \frac{\partial s_{0_2}}{\partial x_1''} + 2\frac{\partial s_2^\dagger}{\partial x_2''} \frac{\partial s_{0_2}}{\partial x_2''} \right] \mathrm{d}t'' \Delta x''^3 \tag{2.43a}$$

$$= -\int_0^T \left[\left(\frac{\partial s_1^\dagger(\vec{x}'', T-t'')}{\partial x_2''} + \frac{\partial s_2^\dagger(\vec{x}'', T-t'')}{\partial x_1''} \right) \cdot \left(\frac{\partial s_{0_1}(\vec{x}'', t'')}{\partial x_2''} + \frac{\partial s_{0_2}(\vec{x}'', t'')}{\partial x_1''} \right) \right. \tag{2.43b}$$

$$\left. + 2 \left(\frac{\partial s_1^\dagger(\vec{x}'', T-t'')}{\partial x_1''} \frac{\partial s_{0_1}(\vec{x}'', t'')}{\partial x_1''} + \frac{\partial s_2^\dagger(\vec{x}'', T-t'')}{\partial x_2''} \frac{\partial s_{0_2}(\vec{x}'', t'')}{\partial x_2''} \right) \right] \mathrm{d}t'' \Delta x''^3$$

where I partly omitted the dependence of the wavefield variables s and $s^\dagger$ on time t'' and point $\vec{x}''$ in interest of clarity. ρ_k, λ_k and μ_k denote the model parameters of a grid cell at point $\vec{x}''$. Assuming elastic wave propagation the gradient of the misfit function can be calculated by a zero-lag cross-correlation of the forward propagated wavefield and the adjoint wavefield.

Viscoelastic material - Generalized standard linear solid

If viscoelastic forward modeling is considered the elements of the tensor $\tilde{\Psi}$ for an isotropic material and a GSLS as rheological model are given by Equation (2.9) and (2.13). The medium can be parametrized by the density ρ, the unrelaxed P-wave modulus π_u, the unrelaxed shear modulus μ_u, the τ values τ^p and τ^s for P and S-waves, respectively, and the relaxation times $\tau_{\sigma l}$ (compare Section 2.1.1). Since in the τ method the relaxation times determine the frequency interval where a constant quality factor is approximated an inversion for them is not meaningful. The frequency interval is already fixed prior to the inversion due to the frequency content of the

observed data. The τ values τ^p and τ^s are closely related to the values of the quality factors Q_p and Q_s for P and S-waves. In the presented studies I do not invert for τ values or quality factors. An inversion for these parameters has most likely a high ambiguity especially in an inversion of field data where noise is present in the observed data and the amplitudes are also influenced by receiver coupling and differences in the transfer functions of the geophones. In my tests I assume a constant a priori known quality factor Q and determine a τ value and relaxation frequencies which approximate the constant Q value in the desired frequency range. The model parameters I invert for are the density and the unrelaxed moduli π_u and μ_u. Nevertheless, the viscoelastic case is more complicated than the elastic case. As stated in Equations (2.14) and (2.15) the response functions φ_{p_l} and φ_{s_l} and therefore also Φ_p and Φ_s depend on the unrelaxed moduli and contribute to the gradients of the unrelaxed moduli. For example, the gradient for the unrelaxed P-wave modulus is

$$\begin{aligned}\frac{\partial E}{\partial \pi_{u_k}} = &- \int_0^T \left(\frac{\partial s_1^\dagger(\vec{x}'', T-t'')}{\partial x_1''} + \frac{\partial s_2^\dagger(\vec{x}'', T-t'')}{\partial x_2''} \right) \\ &\times \Bigg\{ \Bigg[\underbrace{\frac{1}{1+L\tau^p} \left(1 + \sum_{l=1}^{L} \tau^p e^{-t''/\tau_{\sigma l}} \right) \delta(t'')}_{h_1(t'') = \frac{\partial \hat{\psi}_p}{\partial \pi_u} \delta(t'')} - \underbrace{\left(\sum_{l=1}^{L} \frac{1}{\tau_{\sigma l}} \frac{\tau^p}{1+L\tau^p} e^{-t''/\tau_{\sigma l}} H(t'') \right)}_{h_2(t'') = \frac{\partial \Phi_p}{\partial \pi_u} H(t'')} \Bigg] \\ &\qquad * \left(\frac{\partial s_{0_1}(\vec{x}'', t'')}{\partial x_1''} + \frac{\partial s_{0_2}(\vec{x}'', t'')}{\partial x_2''} \right) \Bigg\} \, \mathrm{d}t'' \Delta x''^3 \end{aligned} \tag{2.44}$$

where the identity $f(t) * [g_1(t) + g_2(t)] = [f(t) * g_1(t)] + [f(t) * g_2(t)]$ is used. The expression in square brackets is the derivative of the rate of relaxation function $\dot{\psi}_p$ with respect to the unrelaxed modulus π_u (Equation 2.13) which must be convolved with the forward propagated wavefield. Due to the δ-function the convolution of the first summand h_1 with the forward propagated wavefield simplifies and results in the forward propagated wavefield only. This contribution to the gradient can be easily calculated by a zero-lag cross-correlation of forward and adjoint wavefields like in the elastic case. The second summand includes a convolution which is computationally more expensive. To estimate the contributions of the two summands to the gradient the Fourier transforms of h_1 and h_2 can be compared. The Fourier transform of h_1 is 1 and the Fourier transform of h_2 is

$$H_2(\omega) = \sum_{l=1}^{L} \frac{\tau^p}{1+L\tau^p} \frac{1}{i\omega\tau_{\sigma l} + 1}. \tag{2.45}$$

Figure 2.2 displays the amplitude spectra of H_2 for two approximated quality factors of $Q \approx 10$ and $Q \approx 50$. In the dominant frequency band used in the inversions between 10 Hz and 70 Hz the amplitude is lower than 0.3 even for a rather small quality factor of Q=10. This indicates that the contribution of the second summand (including the convolution) is smaller than the contribution of the first summand and should therefore influence the direction of the gradient of the misfit function less.

In the studies presented in this thesis I approximate the gradient of the misfit function by neglecting the summands depending on the convolution of the response functions with the forward

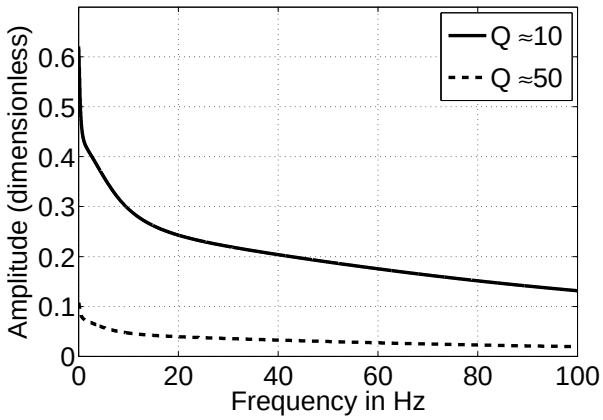


Figure 2.2: Amplitude spectra of $H_2(\omega)$ calculated with Equation (2.45) for approximated Q values of 10 and 50.

propagated wavefields. This is a critical step since it changes the gradient direction. However, as long as all significant physical effects are considered in the forward modeling the quality of a model update during an FWI can be analyzed by running a forward simulation with the new model parameters and calculating the misfit. If the approximation of the gradient is not accurate enough it can be recognized in an increase of the misfit. In the tests presented in this thesis I observe no problems with an increasing misfit due to inaccurate gradient directions. However, this observation could be model dependent.

The gradients for the unrelaxed moduli are approximated by

$$\frac{\partial E}{\partial \pi_{u_k}} = -\int_0^T \left(\frac{\partial s_1^\dagger(\vec{x}'', T-t'')}{\partial x_1''} + \frac{\partial s_2^\dagger(\vec{x}'', T-t'')}{\partial x_2''} \right) \cdot \left(\frac{\partial s_{0_1}(\vec{x}'', t'')}{\partial x_1''} + \frac{\partial s_{0_2}(\vec{x}'', t'')}{\partial x_2''} \right) \mathrm{d}t'' \Delta x''^3 \tag{2.46}$$

and

$$\begin{aligned} \frac{\partial E}{\partial \mu_{u_k}} = \int_0^T \Bigg[& -\left(\frac{\partial s_1^\dagger(\vec{x}'', T-t'')}{\partial x_2''} + \frac{\partial s_2^\dagger(\vec{x}'', T-t'')}{\partial x_1''} \right) \cdot \left(\frac{\partial s_{0_1}(\vec{x}'', t'')}{\partial x_2''} + \frac{\partial s_{0_2}(\vec{x}'', t'')}{\partial x_1''} \right) \\ & + 2 \left(\frac{\partial s_1^\dagger(\vec{x}'', T-t'')}{\partial x_1''} \frac{\partial s_{0_2}(\vec{x}'', t'')}{\partial x_2''} + \frac{\partial s_2^\dagger(\vec{x}'', T-t'')}{\partial x_2''} \frac{\partial s_{0_1}(\vec{x}'', t'')}{\partial x_1''} \right) \Bigg] \mathrm{d}t'' \Delta x''^3. \end{aligned} \tag{2.47}$$

Furthermore, the density gradient is given by

$$\frac{\partial E}{\partial \rho_k} = \int_0^T \left(\frac{\partial s_1^\dagger(\vec{x}'', T-t'')}{\partial t''} \frac{\partial s_{0_1}(\vec{x}'', t'')}{\partial t''} + \frac{\partial s_2^\dagger(\vec{x}'', T-t'')}{\partial t''} \frac{\partial s_{0_2}(\vec{x}'', t'')}{\partial t''} \right) \mathrm{d}t'' \Delta x''^3. \tag{2.48}$$

To change the parametrization from the three parameters ρ, π_u and μ_u to the parameters ρ', λ_u' and μ_u' where λ_u' and μ_u' are the unrelaxed Lamé parameters the chain rule can be applied. The

parameters are related via

$$\rho = \rho', \qquad \pi_u = \lambda_u' + 2\mu_u', \qquad \text{and} \qquad \mu_u = \mu_u'$$

and the gradients can be expressed by

$$\frac{\partial E}{\partial \rho'} = \frac{\partial E}{\partial \rho}\frac{\partial \rho}{\partial \rho'} + \frac{\partial E}{\partial \pi_u}\frac{\partial \pi_u}{\partial \rho'} + \frac{\partial E}{\partial \mu_u}\frac{\partial \mu_u}{\partial \rho'} = \frac{\partial E}{\partial \rho} \tag{2.49a}$$

$$\frac{\partial E}{\partial \lambda_u'} = \frac{\partial E}{\partial \pi_u}\frac{\partial \pi_u}{\partial \lambda_u'} + \frac{\partial E}{\partial \mu_u}\frac{\partial \mu_u}{\partial \lambda_u'} + \frac{\partial E}{\partial \rho}\frac{\partial \rho}{\partial \lambda_u'} = \frac{\partial E}{\partial \pi_u} \tag{2.49b}$$

$$\begin{aligned}\frac{\partial E}{\partial \mu_u'} &= \frac{\partial E}{\partial \pi_u}\frac{\partial \pi_u}{\partial \mu_u'} + \frac{\partial E}{\partial \mu_u}\frac{\partial \mu_u}{\partial \mu_u'} + \frac{\partial E}{\partial \rho}\frac{\partial \rho}{\partial \mu_u'} = 2\frac{\partial E}{\partial \pi_u} + \frac{\partial E}{\partial \mu_u} \\ &= -\int_0^T \left[\left(\frac{\partial s_1^\dagger}{\partial x_2''} + \frac{\partial s_2^\dagger}{\partial x_1''}\right)\cdot\left(\frac{\partial s_{0_1}}{\partial x_2''} + \frac{\partial s_{0_2}}{\partial x_1''}\right) + 2\left(\frac{\partial s_1^\dagger}{\partial x_1''}\frac{\partial s_{0_1}}{\partial x_1''} + \frac{\partial s_2^\dagger}{\partial x_2''}\frac{\partial s_{0_2}}{\partial x_2''}\right)\right] \mathrm{d}t''\Delta x''^3.\end{aligned} \tag{2.49c}$$

Using the approximation described above the gradients for density and Lamé parameters coincide with the corresponding gradients in the elastic case.

Expression of gradients in a stress-velocity forward-modeling scheme

The gradients for the Lamé parameters depend on the spatial derivatives of the forward and the adjoint wavefields. The 2D FWI code DENISE used in this study applies a finite-difference stress-velocity scheme as forward solver. Therefore it is advisable to express the spatial derivatives of the wavefields in terms of stresses since they are explicitly calculated within the stress-velocity scheme. The corresponding equations are given in Appendix B.3.

2.1.5 Gradients for P-wave velocity, S-wave velocity and density

Köhn et al. (2012) investigated the effects of different parametrization in a 2D elastic FWI and concludes that the reconstructed synthetic models are more accurate for a parametrization with density and P and S-wave velocities instead of density and Lamé parameters. They performed a synthetic study for a reflection seismic geometry and explosive sources. Although these results possibly cannot be transferred directly to surface waves I decided to use density and the seismic velocities as parametrization in my tests. The gradients for these parameters are again calculated using the chain rule. With the relations

$$\rho' = \rho \tag{2.50a}$$

$$v_p' = \sqrt{\frac{\lambda + 2\mu}{\rho}} \qquad \rightarrow \qquad \lambda = \rho'(v_p'^2 - 2v_s'^2) \tag{2.50b}$$

$$v_s' = \sqrt{\frac{\mu}{\rho}} \qquad \rightarrow \qquad \mu = \rho' v_s'^2 \tag{2.50c}$$

one obtains

$$\frac{\partial E}{\partial v'_p} = 2\rho' v'_p \frac{\partial E}{\partial \lambda} \tag{2.51a}$$

$$\frac{\partial E}{\partial v'_s} = -4\rho' v'_s \frac{\partial E}{\partial \lambda} + 2\rho' v'_s \frac{\partial E}{\partial \mu} \tag{2.51b}$$

$$\frac{\partial E}{\partial \rho'} = (v'^2_p - 2v'^2_s)\frac{\partial E}{\partial \lambda} + v'^2_s \frac{\partial E}{\partial \mu} + \frac{\partial E}{\partial \rho}. \tag{2.51c}$$

2.2 Misfit definitions and corresponding adjoint sources

2.2.1 L2 norm

A widely used misfit definition in FWI is the L2 or least-squares norm where the misfit function is defined by

$$E = \sum_{r=1}^{n_r} \sum_{j=1}^{n_c} \int_0^T \left(d_j(\vec{x}_r, t) - s_j(\vec{x}_r, t, \vec{m}) \right)^2 \mathrm{d}t. \tag{2.52}$$

The derivative of the misfit function with respect to the model parameters is

$$\frac{\partial E}{\partial m_k} = \sum_{r=1}^{n_r} \sum_{j=1}^{n_c} \int_0^T -2 \left(d_j(\vec{x}_r, t) - s_j(\vec{x}_r, t, \vec{m}) \right) \frac{\partial s_j(\vec{x}_r, t, \vec{m})}{\partial m_k} \mathrm{d}t. \tag{2.53}$$

A comparison with Equation (2.27) yields

$$\partial_{s_j} g(\vec{x}_r, t, \vec{m}) = -2 \left(d_j(\vec{x}_r, t) - s_j(\vec{x}_r, t, \vec{m}) \right) \tag{2.54}$$

for the Fréchet derivative and therefore together with Equation (2.33) the adjoint sources

$$f_j^\dagger(\vec{x}, t) = \sum_{r=1}^{n_r} 2 \left(s_{0_j}(\vec{x}, T-t, \vec{m}_0) - d_j(\vec{x}, T-t) \right) \delta(\vec{x} - \vec{x}_r). \tag{2.55}$$

When the L2 norm is used as misfit definition the residual seismograms are used as adjoint sources.

2.2.2 L2 norm of normalized wavefields or global correlation norm

Choi & Alkhalifah (2012) suggested an L2 norm of the normalized wavefields as misfit definition. In this definition each seismogram is normalized by its RMS value. Therefore, this misfit definition is not sensitive to differences in the amplitude decay with offset. Furthermore, far and near-offset traces can equally contribute to the misfit. The misfit function in time continuous form is

$$E = \sum_{r=1}^{n_r} \sum_{j=1}^{n_c} \int_0^T \left(\frac{s_j(\vec{x}_r, t, \vec{m})}{\left(\frac{1}{T} \int_0^T s_j^2(\vec{x}_r, t'', \vec{m}) \, \mathrm{d}t'' \right)^{1/2}} - \frac{d_j(\vec{x}_r, t)}{\left(\frac{1}{T} \int_0^T d_j^2(\vec{x}_r, t'') \, \mathrm{d}t'' \right)^{1/2}} \right)^2 \mathrm{d}t. \tag{2.56}$$

By application of the binomial formula this misfit function can be also formulated as (Choi & Alkhalifah, 2012)

$$E = 2\frac{n_r \cdot n_c}{\frac{1}{T}} - 2\sum_{r=1}^{n_r}\sum_{j=1}^{n_c}\int_0^T \frac{s_j(\vec{x}_r,t,\vec{m})\cdot d_j(\vec{x}_r,t)}{\left(\frac{1}{T}\int_0^T s_j^2(\vec{x}_r,t'',\vec{m})\,\mathrm{d}t''\right)^{1/2}\left(\frac{1}{T}\int_0^T d_j^2(\vec{x}_r,t'')\,\mathrm{d}t''\right)^{1/2}}\,\mathrm{d}t \tag{2.57}$$

where the second summand corresponds to a zero-lag cross-correlation between the normalized observed and the normalized synthetic data. Choi & Alkhalifah (2012) name the misfit function that consists only of the second term 'global correlation norm'. (The minus sign is included in this misfit definition so that the misfit is again minimized and not maximized.) Since the first summand of Equation (2.57) is independent of the synthetic seismograms it does not contribute to the adjoint sources. Therefore, the gradient of the misfit function formulated for the L2 norm with normalized wavefields coincides with the gradient formulated for the global correlation norm.

To derive the adjoint sources the derivative of the misfit function with respect to the model parameters is calculated (Choi & Alkhalifah, 2012). It is (for details see Appendix B.4)

$$\begin{aligned}\frac{\partial E}{\partial m_k} = \sum_{r=1}^{n_r}\sum_{j=1}^{n_c}\int_0^T -2\Bigg[&\frac{d_j(\vec{x}_r,t)}{\left(\frac{1}{T}\int_0^T s_j^2(\vec{x}_r,t'',\vec{m})\,\mathrm{d}t''\right)^{1/2}\left(\frac{1}{T}\int_0^T d_j^2(\vec{x}_r,t'')\,\mathrm{d}t''\right)^{1/2}}\\ &-\frac{\left(\frac{1}{T}\int_0^T s_j(\vec{x}_r,t',\vec{m})d_j(\vec{x}_r,t')\,\mathrm{d}t'\right)\,s_j(\vec{x}_r,t,\vec{m})}{\left(\frac{1}{T}\int_0^T s_j^2(\vec{x}_r,t'',\vec{m})\,\mathrm{d}t''\right)^{3/2}\left(\frac{1}{T}\int_0^T d_j^2(\vec{x}_r,t'')\,\mathrm{d}t''\right)^{1/2}}\Bigg]\frac{\partial s_j(\vec{x}_r,t,\vec{m})}{\partial m_k}\,\mathrm{d}t.\end{aligned} \tag{2.58}$$

In analogy to the derivation for the L2 norm (previous section) the adjoint sources can be expressed by

$$\begin{aligned}f_j^\dagger(\vec{x},t) = \sum_{r=1}^{n_r}\Bigg(&-\frac{d_j(\vec{x},T-t)}{\left(\frac{1}{T}\int_0^T s_{0_j}^2(\vec{x},t'',\vec{m}_0)\,\mathrm{d}t''\right)^{1/2}\left(\frac{1}{T}\int_0^T d_j^2(\vec{x},t'')\,\mathrm{d}t''\right)^{1/2}}\\ &+\frac{\frac{1}{T}\int_0^T s_{0_j}(\vec{x},t',\vec{m}_0)d_j(\vec{x},t')\,\mathrm{d}t'\,s_{0_j}(\vec{x},T-t,\vec{m}_0)}{\left(\frac{1}{T}\int_0^T s_{0_j}^2(\vec{x},t'',\vec{m}_0)\,\mathrm{d}t''\right)^{3/2}\left(\frac{1}{T}\int_0^T d_j^2(\vec{x},t'')\,\mathrm{d}t''\right)^{1/2}}\Bigg)\delta(\vec{x}-\vec{x}_r)\end{aligned} \tag{2.59}$$

where the factor of 2 is neglected because it only changes the overall amplitude of the adjoint sources and the gradient, respectively. However, since the gradients are scaled before the model update (see Section 2.3.2) this does not influence the results.

2.3 Numerical implementation of FWI

In my work I use the 2D FWI code DENISE which was developed by Köhn (2011). This code is available under the terms of GNU general public license under `www.gpi.kit.edu/Software.php`

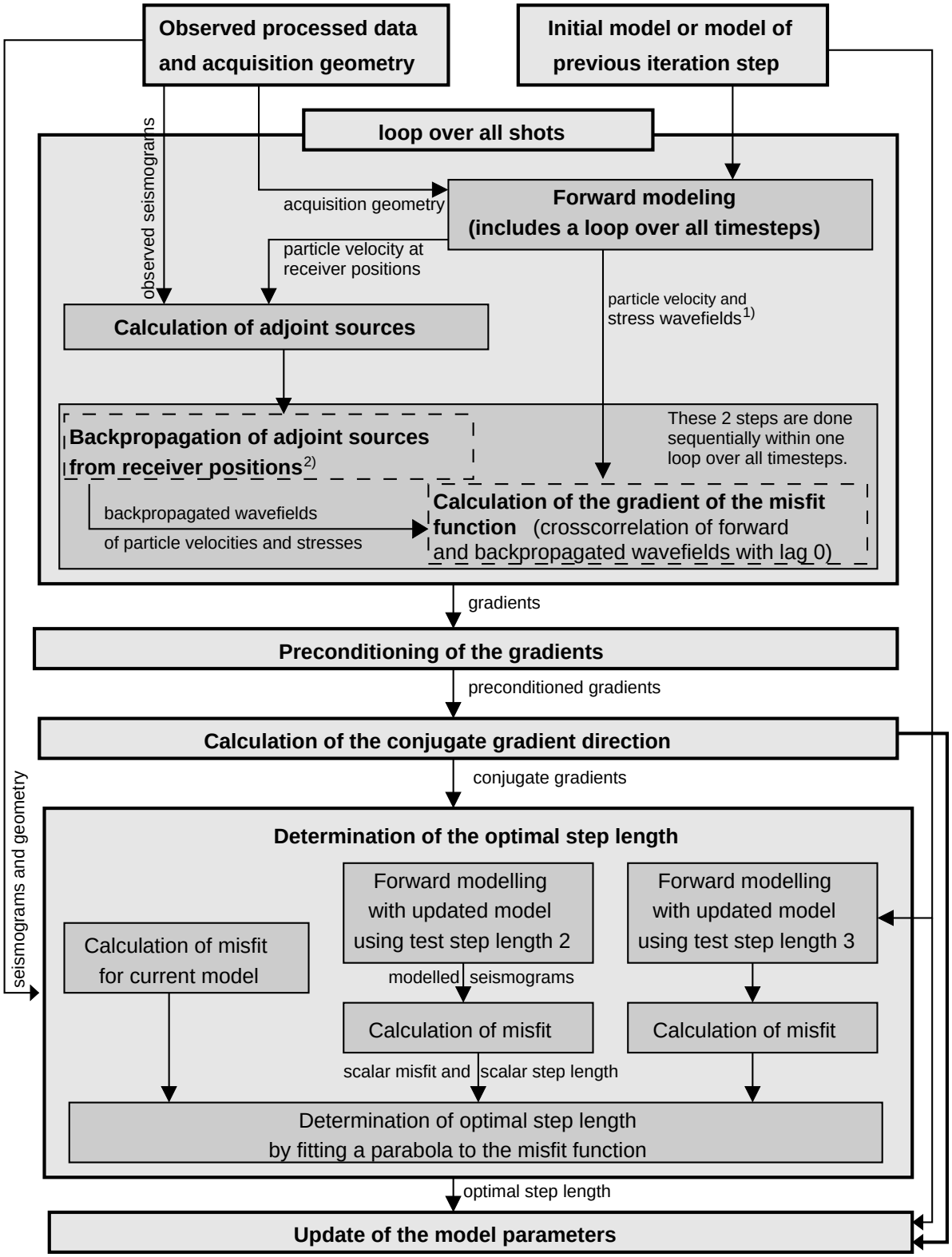
or www.opentoast.de. Figure 2.3 displays a flow chart of the main steps realized in DENISE during one iteration step. Input parameters are the observed data and the initial model or the model of the previous iteration step. First, the gradient of the misfit function is calculated within a loop over the single shots. This step includes a forward modeling on the current subsurface model where the wavefields of the stresses and the particle velocities are saved in memory. Afterwards, the adjoint sources are calculated with the observed data and the simulated seismograms at the receiver positions. Another modeling is applied where the adjoint sources are propagated from the receiver positions back into the medium. Since the gradient of the misfit function depends on zero-lag cross-correlations of the forward and the adjoint wavefields the gradient is calculated on the fly within the same loop over time steps that is used for the back-propagation. Therefore, the adjoint method is numerically very efficient. The gradient of the misfit function is divided into three parts: a gradient for the P-wave velocity, a gradient for the S-wave velocity, and a gradient for the density. These three gradients are addressed separately in DENISE in the further steps (see also Subsection 2.3.2). As forward solver a finite-difference stress-velocity scheme (Bohlen, 1998) is implemented (Chapter 3).

After the calculation of the gradients preconditioning is applied to them to avoid artefacts especially in the vicinity of the sources (see Section 2.3.1). To solve the optimization problem a conjugate gradient method is implemented in DENISE due to the higher convergence of this method in comparison to a purely steepest descent method.

Before the model update can be applied an appropriate step length must be found. Therefore, the misfit function is locally approximated by a parabola. The minimum of this parabola is used as step length. In principle, at least two additional forward modeling runs must be carried out where the models are updated with different test step lengths (for a description of the model update see also Section 2.3.2). However, in DENISE a more enhanced algorithm is implemented which performs more forward modeling runs to obtain a possibly large but still reasonable step length. For details see Köhn et al. (2013). To save computing time not all shots are used for the estimation of an optimal step length. In my work I always used three test shots, one at the left, one at the right and one in the middle of the profile. Finally, the model is updated with the optimal step length found for the current iteration step.

2.3.1 Preconditioning of gradients

As already mentioned the gradients have very high amplitudes in the vicinity of the sources. Therefore, they have to be strongly damped in these regions since otherwise the model is only updated near the sources and the inversion fails. In shallow seismics we also want to invert the S-wave velocity structure near the sources close to the surface. Hence, the gradient cannot be set to small values or even zero in the first meters in general. In an FWI of surface waves I apply a preconditioning to the gradient of each shot before summing all gradients. Thus, also a model update at the source positions is obtained in the final gradient. As preconditioning taper I use a semi-circular taper around the respective source position. An exemplary taper is shown in Figure 2.4 for a shot at x=25 m. The amplitude of the taper is zero at the source position and increases via the error function up to an amplitude of one at the boundary of the circle. In the shown example (Figure 2.4) the radius of the taper is 3.0 m.



1) The whole process is realized in one binary program. Therefore, the wavefields remain in memory for one shot.
2) Corresponds to a forward modelling with time reversed adjoint sources which act as sources at the receiver positions.

Figure 2.3: Flow chart of one iteration step as implemented in the 2D FWI code DENISE.

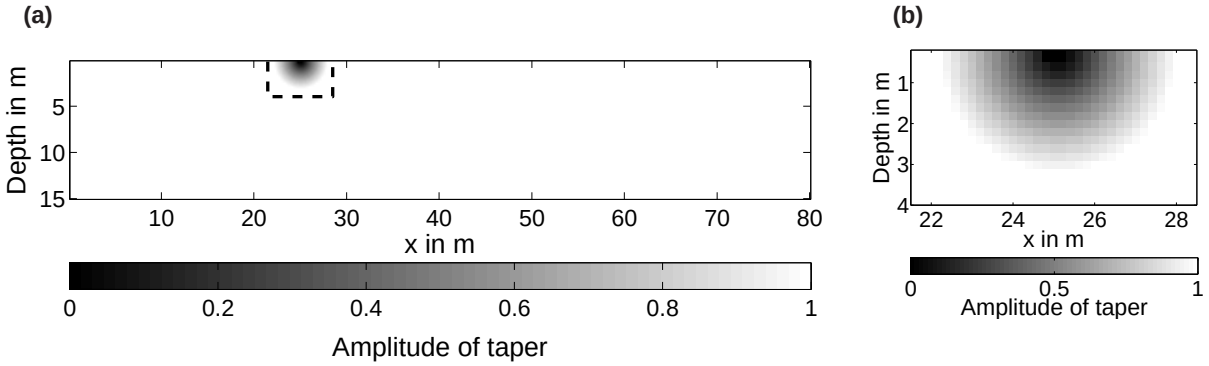


Figure 2.4: Preconditioning taper for a shot at x=25 m at the free surface. a) displays the full taper and b) displays only the part marked by the dashed line in a).

2.3.2 Scaling of the gradients

As derived in Section 2.1.4 the update $\Delta\vec{m}$ of the model parameters is calculated by

$$\Delta m_k = -\alpha \left.\frac{\partial E}{\partial m_k}\right|_{\vec{m}=\vec{m}_0} \tag{2.60}$$

where α is the step length. I parametrized the subsurface model by P-wave velocity v_p, S-wave velocity v_s and density ρ. Therefore, the whole gradient can be subdivided into three parts

$$\nabla_{\vec{m}} E = \begin{pmatrix} \nabla_{\vec{v}_p} E \\ \nabla_{\vec{v}_s} E \\ \nabla_{\vec{\rho}} E \end{pmatrix} = \begin{pmatrix} \vec{g}_{v_p} \\ \vec{g}_{v_s} \\ \vec{g}_{\rho} \end{pmatrix} \tag{2.61}$$

where the vectors $\vec{v}_p$, $\vec{v}_s$ and $\vec{\rho}$ are the P-wave velocity, S-wave velocity and density values at each grid point. In the DENISE code the update of each model parameter is decoupled by using

$$\Delta\vec{v}_p = \alpha_1 \vec{g}_{v_p} \qquad \text{with} \qquad \alpha_1 = \varepsilon \frac{\max[\vec{v}_p]}{\max[\vec{g}_{v_p}]} \tag{2.62a}$$

$$\Delta\vec{v}_s = \alpha_2 \vec{g}_{v_s} \qquad \text{with} \qquad \alpha_2 = \varepsilon \frac{\max[\vec{v}_s]}{\max[\vec{g}_{v_s}]} \tag{2.62b}$$

$$\Delta\vec{\rho} = \alpha_3 \vec{g}_{\rho} \qquad \text{with} \qquad \alpha_3 = \varepsilon \frac{\max[\vec{\rho}]}{\max[\vec{g}_{\rho}]} \tag{2.62c}$$

where $\max[\vec{a}]$ is the maximum absolute value of the elements of vector $\vec{a}$ and ε is a scalar factor. Therefore, the absolute maximum of the gradients is scaled to the maximum of the current models (Köhn, 2011). This scaling causes that the largest update in each of the three model parameters $\vec{v}_p$, $\vec{v}_s$ and $\vec{\rho}$ is ε times the maximum value of this parameter in the current model. For example, for ε=0.01 the maximum model update is 1% of the maximum P-wave velocity, S-wave velocity and density in the current model independent of the absolute values of the single gradients.

Since the scaling factors α_1, α_2 and α_3 are not equal the direction of the whole gradient $\nabla_{\vec{m}} E$ is changed by the scaling which is currently implemented in the DENISE code. However, the

direction of the gradient is also changed during preconditioning where empirically defined tapers are applied to the gradients. Therefore, I used the scaling as described above for the tests described in this thesis.

2.3.3 L2 norm of normalized wavefields

For the inversion of shallow seismic Rayleigh waves I implement the L2 norm of normalized wavefields which was already discussed in Section 2.2.2 into the DENISE code. I use this norm in the reconstruction tests which are presented in Chapter 4 as well as in the inversion of field data (Chapter 5). Therefore, I give this misfit norm additionally in discrete notation. In time discrete form the misfit function of Equation (2.56) can be expressed by

$$E = \frac{\sum_i^{n_s} \sum_r^{n_r} \sum_j^{n_c} \left| \hat{\vec{s}}_{i,r,j} - \hat{\vec{d}}_{i,r,j} \right|^2}{\sum_i^{n_s} \sum_r^{n_r} \sum_j^{n_c} \left| \hat{\vec{d}}_{i,r,j} \right|^2} = \frac{\sum_i^{n_s} \sum_r^{n_r} \sum_j^{n_c} \left| \hat{\vec{s}}_{i,r,j} - \hat{\vec{d}}_{i,r,j} \right|^2}{n_s n_r n_c} \tag{2.63}$$

where the sum over r denotes the sum over the n_r receivers and the sum over j denotes the sum over the n_c components. Additionally, I add here the sum over i which denotes the sum over the n_s sources. The vectors $\vec{s}_{i,r,j}$ and $\vec{d}_{i,r,j}$ are the synthetic and observed displacement seismograms of source i, receiver r and component j and $\hat{\vec{s}}_{i,r,j} = \vec{s}_{i,r,j}/|\vec{s}_{i,r,j}|$ and $\hat{\vec{d}}_{i,r,j} = \vec{d}_{i,r,j}/|\vec{d}_{i,r,j}|$ are the corresponding normalized displacement seismograms, respectively. In comparison to the norm suggested by Choi & Alkhalifah (2012) I normalize the misfit function. Thus, E equals one if the energy of the residual of the normalized seismograms equals the energy of the normalized observed data.

To determine the adjoint source for a receiver at $\vec{x} = \vec{x}_r$ and component j according to Equation (2.59) one first has to calculate (compare Choi & Alkhalifah, 2012)

$$\vec{f}^{\dagger}_{j,forward}(\vec{x}_r) = \frac{1}{\left|\vec{s}_{0_j}(\vec{x}_r)\right| \left|\vec{d}_j(\vec{x}_r)\right|} \left(-\vec{d}_j(\vec{x}_r) + \frac{\vec{s}_{0_j}(\vec{x}_r)}{\left|\vec{s}_{0_j}(\vec{x}_r)\right|^2} \left(\vec{s}_{0_j}(\vec{x}_r) \cdot \vec{d}_j(\vec{x}_r) \right) \right). \tag{2.64}$$

To finally obtain the adjoint source $\vec{f}^{\dagger}_j(\vec{x}_r)$ which excites the adjoint wavefield at the receiver position $\vec{x}_r$ the vector $\vec{f}^{\dagger}_{j,forward}(\vec{x}_r)$ must be flipped due to the time reversal.

2.3.4 Cycle-skipping and multiscale inversion

FWI is a non-linear inverse problem. Therefore, the misfit function contains not only a global minimum but also several local minima. If the inversion converges into the global minimum or if it gets stuck in one of the local minima strongly depends on the initial model. To achieve a convergence into the global minimum the waveforms generated with the initial model must fit the observed data within less than half a period. Otherwise, cycle-skipping will occur (e. g. Virieux & Operto, 2009) and the inversion converges towards a local minimum. In the case of cycle-skipping FWI fits cycles of the observed and synthetic data which do not correspond to each other.

To reduce the requirements for the initial model Bunks et al. (1995) suggest a multiscale approach where subsets of the observed data with different frequency content are inverted successively. At the beginning only the low frequency content of the observed data is used in the inversion since this subset is less prone to artefacts caused by cycle-skipping. After the inversion converged for the low frequencies the bandwidth of the data which are used in the inversion is increased. The multiscale approach is realized in the DENISE code by the application of a Butterworth lowpass filter to the data. The corner frequency of the Butterworth filter is increased as soon as no significant change in the misfit is observed.

2.4 3D/2D transformation

In the applied FWI code a 2D forward solver is implemented (see also Section 3.1). Therefore, line sources are used implicitly in the simulations. However, in general field data are acquired by hammer blows or explosions which can be approximated by point sources. This causes differences in the wavefields in amplitude and phase. Thus, a 3D/2D transformation must be applied to the field data prior to the inversion.

I investigate a transformation which is exact for 1D subsurface structures and is suggested by Wapenaar et al. (1992) and Amundsen & Reitan (1994). In this section I will give an overview of this transformation and I will show a numerical example where I test the accuracy of this transformation. At the end I will briefly describe further developments concerning appropriate transformations for shallow seismic wavefields.

2.4.1 Transformation using the Fourier-Bessel expansion

This transformation is suggested by Wapenaar et al. (1992) and Amundsen & Reitan (1994) and is analytically exact for subsurface structures where the material properties vary only with depth. Such subsurface structures are called one-dimensional (1D) subsurface structures in the following. Amundsen & Reitan (1994) do not give a complete derivation of the transformation. Therefore, Thomas Forbriger and I collaborated to derive the transformation. Thomas Forbriger firstly derived the expressions for the transformation of the vertical component wavefield and afterwards I derived the expressions for the horizontal component.

We assume a cylindrically symmetrical source (vertical point force or explosion) as source in the origin of the coordinate system. For a one-dimensional subsurface structure the wavefield can be expressed by an expansion with Fourier-Bessel functions and is given in cylindrical coordinates by

$$\tilde{u}_{P_z}(r,\omega) = \int_0^\infty G_z(\omega,k)\, J_0(kr)\, k \,\mathrm{d}k \tag{2.65}$$

$$\tilde{u}_{P_r}(r,\omega) = \int_0^\infty G_r(\omega,k)\, J_1(kr)\, k \,\mathrm{d}k \tag{2.66}$$

with the Fourier transform $\tilde{u}_{P_z}$ and $\tilde{u}_{P_r}$ of the vertical and radial component of the excited wavefield, respectively. The tangential component $\tilde{u}_{P_\varphi}$ is zero due to symmetry. G_z and G_r are the

expansion coefficients for the vertical and the radial component, respectively, k is the wavenumber, r the source-receiver distance or offset and ω the angular frequency. J_0 and J_1 are Bessel functions of the first kind and order zero and one.

To derive the seismograms for a line source we switch to a Cartesian coordinate system. The seismograms of a line source in a distance y to the line source can be written as a superposition of seismograms excited by an infinite number of point sources along the x-axis (Forbriger et al., 2013). The contributions of the single point sources to the vertical component wavefield at a receiver with distance y to the line source simply superimpose and the Fourier coefficients are obtained by

$$\tilde{u}_{L_z}(y,\omega) = \int_{-\infty}^{\infty} \tilde{u}_{P_z}(\sqrt{x^2+y^2},\omega)\ \mathcal{C}\ \mathrm{d}x \tag{2.67}$$

where $\mathcal{C} = 1\mathrm{m}^{-1}$. To obtain the horizontal component in y-direction of the line source wavefield the radial component point source wavefield (Equation 2.66) must be split into the Cartesian components u_{P_x} and u_{P_y} (Figure 2.5). Due to symmetry the x-component of the line-source wavefield vanishes. The Fourier coefficients of the y-component are given by

$$\tilde{u}_{L_y}(y,\omega) = \int_{-\infty}^{\infty} \tilde{u}_{P_r}(\sqrt{x^2+y^2},\omega)\ \frac{y}{\sqrt{x^2+y^2}}\ \mathcal{C}\ \mathrm{d}x. \tag{2.68}$$

Inserting Equations (2.65) and (2.66) into Equations (2.67) and (2.68) yields after some lines of calculus

$$\tilde{u}_{L_z}(y,\omega) = 2\int_0^{\infty} G_z(\omega,k)\ \cos(ky)\ \mathcal{C}\ \mathrm{d}k \tag{2.69}$$

for the vertical and

$$\tilde{u}_{L_y}(y,\omega) = 2\int_0^{\infty} G_r(\omega,k)\ \sin(ky)\ \mathcal{C}\ \mathrm{d}k \tag{2.70}$$

for the horizontal-component Fourier coefficients of the line-source seismograms (see Appendix C for detailed calculation).

Summarizing this transformation the expansion coefficients G_z and G_r have to be calculated first from recorded point-source seismograms via

$$G_z(\omega,k) = \int_0^{\infty} \tilde{u}_{P_z}(r,\omega) J_0(kr)\, r\, \mathrm{d}r \tag{2.71}$$

$$G_r(\omega,k) = \int_0^{\infty} \tilde{u}_{P_r}(r,\omega) J_1(kr)\, r\, \mathrm{d}r \tag{2.72}$$

which are the inverses of Equations (2.65) and (2.66). In practice, only spatially sampled point-source data are available and these equations must be approximated. This can be done for example by the trapezoid rule (Forbriger, 2003a, eq. 18). Afterwards, these coefficients must be used in an expansion with plane waves according to Equations (2.69) and (2.70).

2.4.2 Numerical example

To test the transformation described in the previous section numerically I simulate point-source seismograms with the reflectivity method (Fuchs & Müller, 1971) using an implementation by

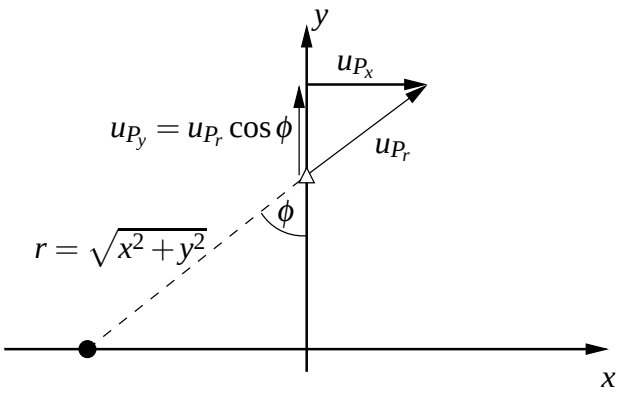


Figure 2.5: Contribution of one point source located on the x-axis to the y-component recorded by a receiver on the y-axis. The source is marked by the black circle and the receiver is marked by the white triangle. The vertical axis is perpendicular to the plane of projection.

Ungerer (1990). Furthermore, I computed line-source seismograms with the 2D finite-difference forward solver implemented in the FWI code (Bohlen, 1998; Köhn, 2011).

I use a simple model of a homogeneous layer over a homogeneous half-space. In the layer I assume a P-wave velocity of 680 m/s, an S-wave velocity of 400 m/s, and a density of 1700 kg/m^3. In the half-space the P-wave velocity is 1000 m/s, the S-wave velocity is 590 m/s, and the density is 2000 kg/m^3. The layer is 4 m thick. To model the free surface I added an air layer on top of the model. A constant quality factor of 800 for both, P and S-waves, is assumed in the calculations with the reflectivity code to simulate almost elastic wave propagation. In the finite-difference simulations a purely elastic medium is assumed.

The wavefields are excited by a vertical force at the free surface and the first half cycle of a $\sin^3$ function with a period of 64 ms is used as source time function (Equation 3.20 with T_d=32 ms). The wavefields are recorded by 72 two-component receivers with an equidistant spacing of 1 m. The smallest source-receiver distance is 1 m and the largest is 72 m.

Figure 2.6 displays a comparison of point-source seismograms (black) and line-source seismograms (red). Each of these seismograms is normalized to its maximum amplitude to emphasize phase residuals. Therefore, the difference in the amplitude decay with offset cannot be observed. Additionally, the residual seismograms between the normalized point-source and line-source seismograms are displayed in blue. The residual seismograms are not normalized. The differences in phase result in a residual wavefield whose amplitudes are almost as large as the amplitudes of the seismograms. To avoid a projection of these residuals into the reconstructed subsurface model a 3D/2D transformation must be applied to the recorded point-source data prior to the application of a 2D FWI.

Figure 2.7 displays the results after applying the 3D/2D transformation using the Fourier-Bessel expansion. The seismograms are not normalized. For offsets between approximately 5 m and 50 m the transformed point-source seismograms fit the line-source seismograms well. The dif-

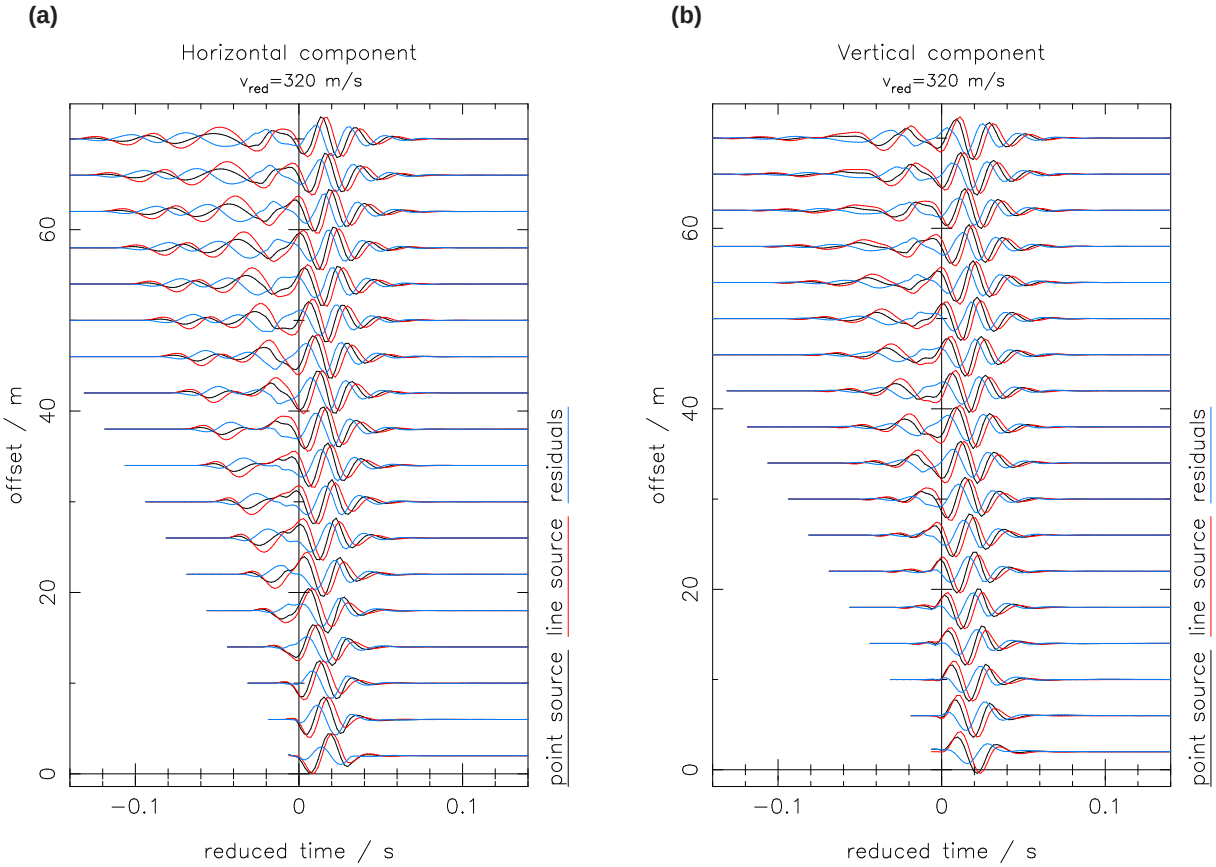


Figure 2.6: Comparison of point-source and line-source seismograms. The black seismograms are excited by a point source whereas the red seismograms are excited by a line source. Point-source as well as line-source seismograms are normalized to the maximum amplitude of each trace. The blue seismograms are the residuals between point-source and line-source seismograms. The residual seismograms are not normalized. a) displays particle velocity seismograms of the horizontal component and b) displays particle velocity seismograms of the vertical component. The time axes in both plots are reduced with a velocity of 320 m/s.

ferences in the traces with smaller offsets are a numerical artefact due to the approximation of the integrals over the wavenumber in Equations (2.69) and (2.70) by a discrete sum only up to a finite value. For larger offsets the results also deteriorate due to a cut-off that appears at constant times. This cut-off appears because point-source seismograms are only available up to a limited distance to the source (Forbriger et al., 2013).

2.4.3 Further developments

The 3D/2D transformation described in Section 2.4.1 has two disadvantages. It depends on the spatial sampling of the available point-source data and it produces strong artefacts in the simulated line-source seismograms when applied to seismograms recorded on a significantly 2D subsurface structure (Schäfer et al., 2013). Due to the assumption of a 1D subsurface structure made in the derivation of the transformation the simulated line-source seismograms do not contain effects caused by 2D or 3D structures even if such effects are present in the point-source

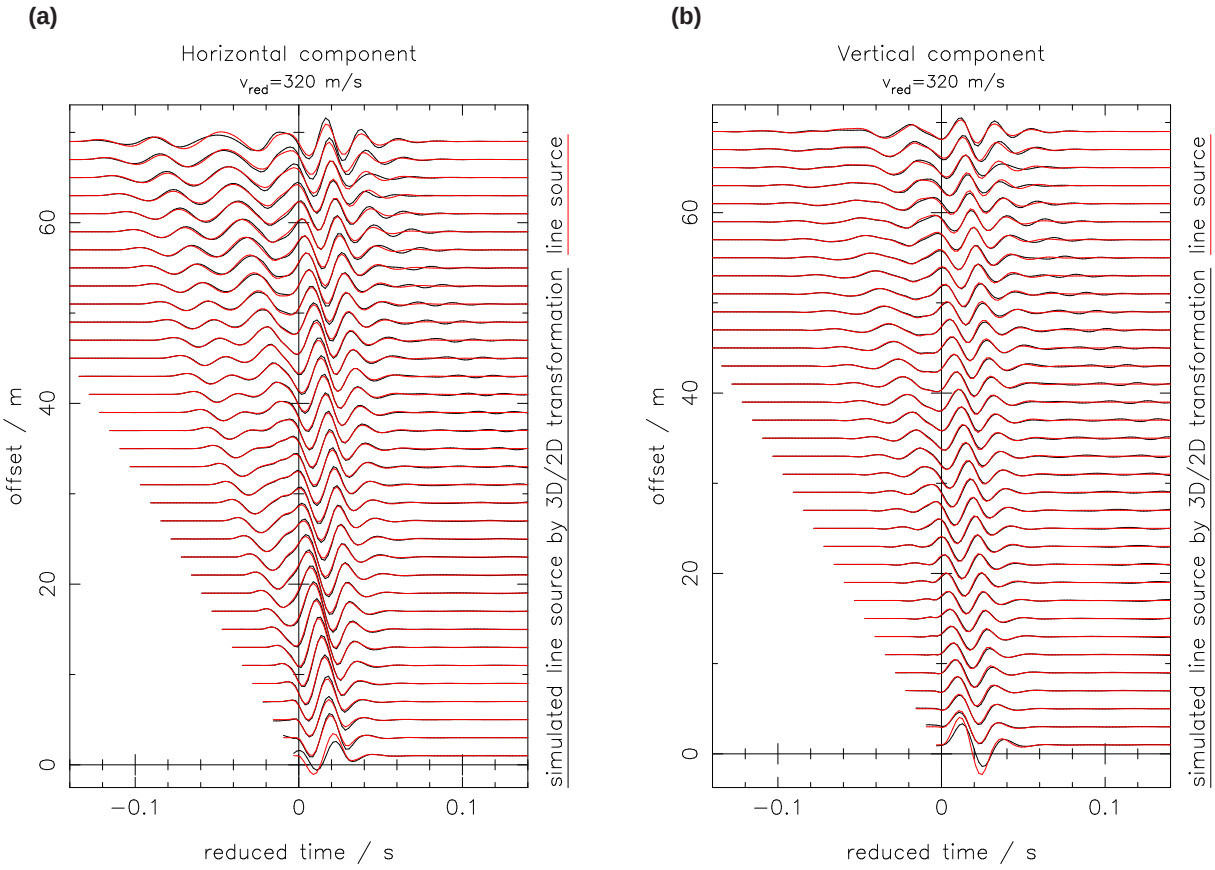


Figure 2.7: Comparison of simulated line-source seismograms by the 3D/2D transformation using the Fourier-Bessel expansion (black) and line-source seismograms calculated by a 2D forward solver (red). The seismograms are not normalized. a) displays particle velocity seismograms of the horizontal component and b) displays particle velocity seismograms of the vertical component. The time axes in both plots are reduced with a velocity of 320 m/s.

seismograms. Effects in the wavefields that are caused by 1D, 2D, and 3D subsurface structures are mixed within the transformation and the result is a line-source wavefield of a pseudo 1D structure. This behaviour is especially undesired for effects caused by 2D subsurface structures since a 2D FWI can handle these effects.

Therefore, we present in Forbriger et al. (2013) an appropriate transformation for shallow seismic wavefields which can be applied to single traces and which performs also well for seismograms recorded on a 2D structure (Schäfer et al., 2013). It is based on the acoustic wave equation. However, Forbriger et al. (2013) and Schäfer et al. (2013) confirmed by numerical tests that it is also applicable in the elastic and even viscoelastic case.

Each seismogram is convolved with $1/\sqrt{t}$ where t is the travel time. This step corresponds to a phase shift of $\pi/4$. Afterwards, a factor F_{amp} is applied to adjust the amplitudes. The factor F_{amp} can be interpreted and applied differently.

Assuming signals with a single phase velocity v_{ph} it is $F_{amp} = \sqrt{2\,r\,v_{ph}}$ where r is the source-receiver distance. Forbriger et al. (2013) call this transformation *single-velocity* transformation. Since surface waves often cover a range of different phase velocities this transformation is not suited for all data sets.

For surface waves it can be assumed that the wave propagation distance equals the source-receiver distance. Therefore, the phase velocity can be estimated by $v_{ph} = r/t$ where t is the travel time. This yields $F_{amp} = r\sqrt{2/t}$. This technique is called *direct-wave* transformation (Forbriger et al., 2013).

As third single-trace transformation Forbriger et al. (2013) and Schäfer et al. (2013) suggest to use a combination of the two previously introduced transformations to eliminate artefacts as effectively as possible. For small offsets the single velocity transformation is used and for large offsets the direct wave transformation. In between a smooth transition between both transformations is applied. This transformation is called *hybrid* transformation.

Chapter 3

Forward modeling and benchmarking

Within a full waveform inversion (FWI) several hundreds of forward modeling runs are performed to find a subsurface model that explains the recorded data within their uncertainties. The accuracy of the forward solver is a crucial point because after a model update we have to check with the next forward modeling whether the data misfit is decreased by the model update or not. Thus, the forward modeling algorithm must describe all significant physical properties of the observed data accurately. To ensure this I run benchmark tests of the forward modeling algorithm used within the FWI code.

3.1 2D finite-difference scheme

The full waveform inversion (FWI) code DENISE uses a 2D P-SV finite-difference (FD) scheme in the time domain to solve the forward problem. The implemented forward solver was developed by Bohlen (1998). It uses the stress-velocity formulation and solves the viscoelastic wave equation on a Standard Staggered Grid (SSG) (Virieux, 1986; Levander, 1988; Robertsson et al., 1994).

Coordinate system

A two-dimensional Cartesian coordinate system is used in the FWI code DENISE (Figure 3.1). The components of the vector fields (e. g. displacement field u_i or velocity field v_i) point into the direction of the corresponding coordinate axis. For example, positive amplitudes in seismograms of the y component (vertical component) correspond to a movement downwards. The origin of the coordinate system is located in the upper left corner of the model.

Interpretation of 2D simulations in a 3D space

The forward solver in the FWI code calculates the wave propagation in the $x-y$ plane, displayed in Figure 3.1, in a 2D space. However, this simulation can be also interpreted in a 3D space by assuming that the $x-y$ slice is extended to infinity in the third dimension. In this case the third dimension corresponds to the second horizontal axis which is plotted by a dashed line and named with z in Figure 3.1. The extension to infinity affects both, the subsurface structures as well as the sources. Therefore, a 2D simulation implicitly corresponds to a 3D simulation with a line source. As the observed data in an FWI are recorded in a 3D space I prefer to think in the following sections of line sources in a 3D space than of point sources in a 2D space.

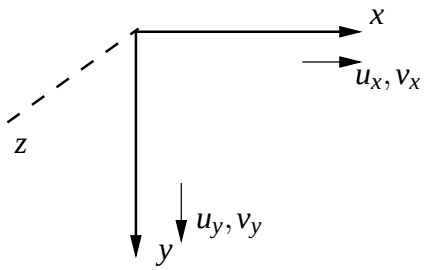


Figure 3.1: Coordinate system used within the FWI code. The origin is at the upper left corner of the model.

2D P-SV wave equation

In a 3D space the wave propagation can be described by the system of partial differential equations

$$\rho \frac{\partial^2 \vec{u}}{\partial t^2} = \frac{\partial \sigma_{ij}}{\partial x_j} + \vec{f} \tag{3.1}$$

with mass density ρ in kg/m^3, particle displacement $\vec{u}$ in m, stresses σ_{ij} in N/m^2 and volume force density $\vec{f}$ in N/m^3 together with an appropriate stress-strain relationship that describes the rheology of the subsurface medium. Elastic as well as viscoelastic wave propagation is implemented in DENISE. An elastic medium is described by the linear isotropic stress-strain relationship

$$\sigma_{ij} = \lambda \ \mathrm{div}\vec{u}\ \delta_{ij} + 2\mu\, \varepsilon_{ij}, \tag{3.2}$$

with the components of the strain tensor $\varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)$. An isothermal and isotropic viscoelastic medium can be described by

$$\sigma_{ij} = (\dot{\psi}_p - 2\dot{\psi}_s) * \varepsilon_{ll}\delta_{ij} + 2\dot{\psi}_s * \varepsilon_{ij} \tag{3.3}$$

where ψ_p and ψ_s are the relaxation functions of the medium (compare Section 2.1.1).

If a 2D subsurface structure is assumed where the properties vary only in x and y direction (see Figure 3.1) and a line source along the z-axis all derivatives with respect to z vanish due to symmetry. Equation (3.1) splits up into a system of two equations

$$\rho \frac{\partial^2 u_x}{\partial t^2} = \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} + f_x \tag{3.4a}$$

$$\rho \frac{\partial^2 u_y}{\partial t^2} = \frac{\partial \sigma_{yx}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y} + f_y \tag{3.4b}$$

describing the P-SV wave propagation and one equation

$$\rho \frac{\partial^2 u_z}{\partial t^2} = \frac{\partial \sigma_{zx}}{\partial x} + \frac{\partial \sigma_{zy}}{\partial y} + f_z \tag{3.5}$$

describing the SH wave propagation in the x-y plane. The components σ_{ij} of the stress tensor are calculated with Equation (3.2) in the elastic case and with Equation (3.3) in the viscoelastic case. In the FWI code DENISE the P-SV wave propagation is simulated by solving Equation (3.4).

Standard staggered grid

The distribution of the wavefield variables as well as the material parameters for the standard staggered grid (Virieux, 1986; Levander, 1988) are shown in Figure 3.2. The indices i and j denote the position in the Cartesian grid. Furthermore, indices i^+ and j^+ denote positions that are shifted by a half grid point in x and y-direction, respectively. The normal stresses σ_{xx} and σ_{yy} are located at full grid points. The particle velocity v_x is shifted a half grid point in x-direction and the particle velocity v_y is shifted a half grid point in vertical or y-direction. The corresponding model parameters

$$\overline{\rho}_x = \rho(i^+, j) = \frac{1}{2}\left[\rho(i,j) + \rho(i+1,j)\right] \tag{3.6a}$$

$$\overline{\rho}_y = \rho(i, j^+) = \frac{1}{2}\left[\rho(i,j) + \rho(i,j+1)\right] \tag{3.6b}$$

are calculated by the average of two neighbouring density values at the full grid points respectively. The shear stress σ_{xy} is shifted by a half grid point in x as well as in y-direction where

$$< \mu >= \mu(i^+, j^+) = 4\left[\frac{1}{\mu(i,j)} + \frac{1}{\mu(i+1,j)} + \frac{1}{\mu(i,j+1)} + \frac{1}{\mu(i+1,j+1)}\right]^{-1} \tag{3.6c}$$

is obtained by an harmonic average of the values of μ at the four neighbouring grid points.

Finite-difference discretization

For simplicity I assume in the following elastic wave propagation. Concerning viscoelastic wave propagation the reader is referred to Bohlen (1998). Since a stress-velocity scheme is implemented in DENISE Equations (3.4) are formulated with respect to particle velocities v_x and v_y by replacing $\partial u_x/\partial t$ with v_x and $\partial u_y/\partial t$ with v_y, respectively. Furthermore, the stress-strain relationship is differentiated with respect to time to obtain a relationship between the time derivative of the stresses and the particle velocities. Afterwards the partial derivatives are approximated by finite-difference operators. Taylor operators of second order are used for time derivatives. Spatial derivatives are approximated by Taylor operators of second to eighth order. Therefore, in the following equations a spatial forward operator is named by D_k^+ and a backward operator is named by D_k^- where the index k denotes the coordinate direction. For Taylor coefficients of second order this means that

$$\left.\frac{\partial a(i\Delta x)}{\partial x}\right|_{(i+1/2)\Delta x} \approx D_x^+[a(i)] = \frac{1}{\Delta x}\left[a(i+1) - a(i)\right] \tag{3.7a}$$

is the forward operator and

$$\left.\frac{\partial a(i\Delta x)}{\partial x}\right|_{(i-1/2)\Delta x} \approx D_x^-[a(i)] = \frac{1}{\Delta x}\left[a(i) - a(i-1)\right] \tag{3.7b}$$

is the corresponding backward operator. In the DENISE code the stress field is updated via

$$\sigma_{xx}^{n+}(i,j) = \sigma_{xx}^{n-}(i,j) + \Delta t\left((\lambda(i,j) + 2\mu(i,j))D_x^-\left[v_x^n(i^+,j)\right] + \lambda(i,j)D_y^-\left[v_y^n(i,j^+)\right]\right) \tag{3.8a}$$

$$\sigma_{yy}^{n+}(i,j) = \sigma_{yy}^{n-}(i,j) + \Delta t\left(\lambda(i,j)D_x^-\left[v_x^n(i^+,j)\right] + (\lambda(i,j) + 2\mu(i,j))D_y^-\left[v_y^n(i,j^+)\right]\right) \tag{3.8b}$$

$$\sigma_{xy}^{n+}(i^+,j^+) = \sigma_{xy}^{n-}(i^+,j^+) + \Delta t\ \mu(i^+,j^+)\left(D_y^+\left[v_x^n(i^+,j)\right] + D_x^+\left[v_y^n(i,j^+)\right]\right) \tag{3.8c}$$

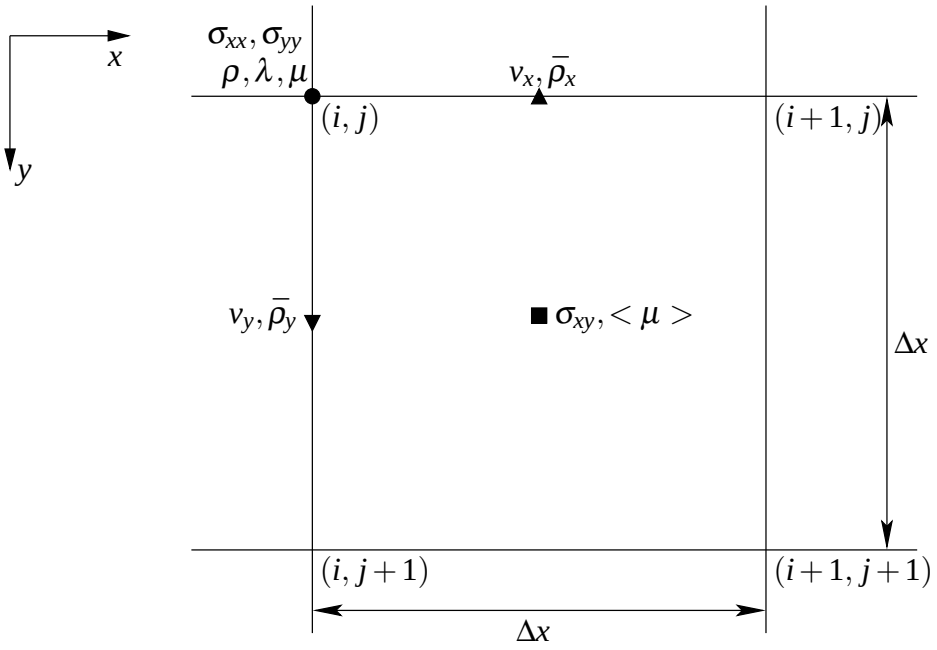


Figure 3.2: Distribution of wavefield variables and material properties on the standard staggered grid used in DENISE (Köhn, 2011, Figure 2.5). The spatial grid spacing is Δx.

and the velocity field via

$$v_x^{n+1}(i^+, j) = v_x^n(i^+, j) + \frac{\Delta t}{\rho(i^+, j)} \left(D_x^+ \left[\sigma_{xx}^{n+}(i, j) \right] + D_y^- \left[\sigma_{xy}^{n+}(i^+, j^+) \right] + f_x^{n+} \right) \tag{3.8d}$$

$$v_y^{n+1}(i, j^+) = v_y^n(i, j^+) + \frac{\Delta t}{\rho(i, j^+)} \left(D_x^- \left[\sigma_{xy}^{n+}(i^+, j^+) \right] + D_y^+ \left[\sigma_{yy}^{n+}(i, j) \right] + f_y^{n+} \right) \tag{3.8e}$$

where Δt is the temporal sampling interval and the time step is marked by the upper index n. For this index again $n+$ means $n + 1/2$ and $n-$ means $n - 1/2$.

3.2 Implementation of free surface

The implementation of the free surface is crucial to model surface waves accurately. I compare two different techniques which are shortly introduced in this section.

3.2.1 Vacuum or air layer

As discussed for example by Zahradník & Priolo (1995) or Bohlen & Saenger (2006) a free surface can be implemented implicitly in the FD scheme by adding a vacuum layer or an air layer at the top of the model. As shown in Section 3.6 the free surface in this case is located directly between the two full grid points where the upper one has the material parameters of the vacuum layer and the lower one has the material properties of the subsurface in the vicinity of the free surface. The differences between an air layer and a vacuum layer are negligible (Section 3.5). A disadvantage of this implementation of the free surface is that the simulations are only stable for second-order spatial FD operators.

3.2.2 Image technique

The image technique was firstly suggested by Levander (1988). In the staggered grid used in DENISE (Figure 3.2) the free surface is located directly at the first row of grid points of the model. σ_{yy} is set explicitly to zero at the free surface. To avoid instabilities only horizontal derivatives of the particle velocity wavefield ($\partial v_x/\partial x$) are used to calculate σ_{xx} at the free surface. To update the particle velocity fields v_x and v_y the stresses σ_{yy} and σ_{xy} are also needed above the free surface.[1] Therefore, the stresses σ_{iy} with $i = x,y$ are imaged in a way that they are odd functions with respect to the free surface. According to Robertsson (1996) the particle velocities above the free surface are set to zero. Also higher order FD operators can be used with this implementation.

3.3 Source implementation

For forward modeling it is important to ensure a correct source implementation. This means that the simulated seismograms must not depend on the temporal or spatial sampling interval as well as the order of the spatial FD operators. In this section I explain the technical details of the source implementation used in the DENISE code. The results of the corresponding benchmark tests are given in Sections 3.4 and 3.5.

Sources are considered in the wave equation by adding external volume force densities $\vec{f}$ (see Equation 3.1). In the following, explosive sources and directed force sources are discussed.

3.3.1 Explosive source

The volume force density that is caused by an explosive line source along the z-axis (Figure 3.1) can be described by a potential L via

$$\overrightarrow{f}(\vec{x},t) = \nabla L(\vec{x},t) \quad \text{with} \quad L(\vec{x},t) = M_0' s(t)\,\delta(x)\delta(y) \tag{3.9}$$

where $M_0' = \lim_{l\to\infty} \frac{M_0}{l}$ is a seismic line moment in Nm/m and l is the length of the line source. It is reasonable to define a seismic line moment in case of a line source. The definition of an absolute seismic moment M_0 which acts along the whole line source results in an infinitely small force density (Forbriger et al., 2013). $s(t)$ is the dimensionless source time function and $\delta(x)$ and $\delta(y)$ are δ-functions with physical unit 1/m.

According to Zahradník & Priolo (1995) a δ-function can be approximated by $\delta(x) \approx \frac{1}{\Delta x}$ in a finite-difference algorithm with the spatial grid spacing Δx in x-direction. This is sketched in Figure 3.3. The proposed approximation ensures that the normalization of the δ-function $\int \delta(x)\mathrm{d}x = 1$ is still valid (Zahradník & Priolo, 1995). For a source at (i,j) the discretized potential is therefore

$$L^n(i,j) = \frac{M_0'}{\Delta x^2} s^n. \tag{3.10}$$

The upper index n marks again the discrete time step. To implement the explosive source into the FD scheme the potential L can be simply added to the main stresses σ_{xx} and σ_{yy}. The

[1] How many values are needed depends on the order of the spatial FD operators.

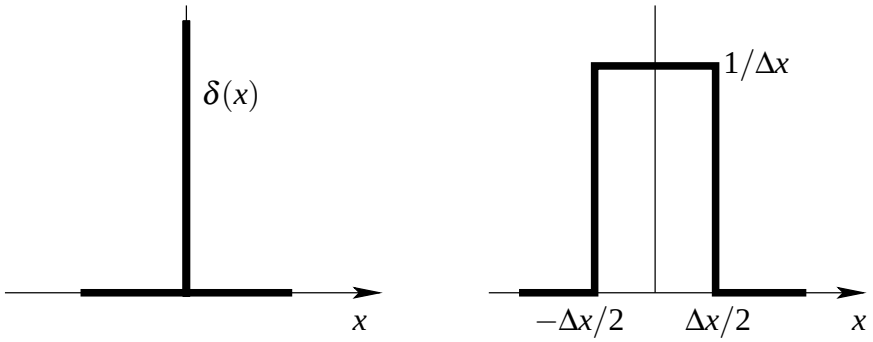


Figure 3.3: Approximation of a δ-function according to Zahradník & Priolo (1995). On the left a continuous δ-function is sketched and on the right the corresponding approximation. The grid spacing is Δx.

gradient of L is calculated automatically during the update of the velocity field (Bohlen, 1998). According to Equation (3.8d-e) the velocity field for a line source at (i, j) is calculated by

$$v_x^n(i^+, j) = v_x^{n-1}(i^+, j) + \frac{\Delta t}{\rho(i^+, j)} \left(D_x^+ \left[\overline{\sigma}_{xx}^{n-}(i, j) \right] + D_y^- \left[\sigma_{xy}^{n-}(i^+, j^+) \right] \right) \tag{3.11a}$$

$$v_y^n(i, j^+) = v_y^{n-1}(i, j^+) + \frac{\Delta t}{\rho(i, j^+)} \left(D_x^- \left[\sigma_{xy}^{n-}(i^+, j^+) \right] + D_y^+ \left[\overline{\sigma}_{yy}^{n-}(i, j) \right] \right) \tag{3.11b}$$

with $\overline{\sigma}_{xx} = \sigma_{xx} + L$ and $\overline{\sigma}_{yy} = \sigma_{yy} + L$.
To update the main stresses the time derivative of the stress-strain relationship is used. At the source point this can be written in continuous form as (personal communication with D. Köhn)

$$\frac{\partial \overline{\sigma}_{xx}}{\partial t} = (\lambda + 2\mu) \frac{\partial v_x}{\partial x} + \lambda \frac{\partial v_y}{\partial y} + \frac{\partial L}{\partial t} \tag{3.12a}$$

$$\frac{\partial \overline{\sigma}_{yy}}{\partial t} = \lambda \frac{\partial v_x}{\partial x} + (\lambda + 2\mu) \frac{\partial v_y}{\partial y} + \frac{\partial L}{\partial t} \tag{3.12b}$$

or in the discrete FD scheme (spatial indices (i, j) are omitted for clarity)

$$\overline{\sigma}_{xx}^{n+} = \overline{\sigma}_{xx}^{n-} + \Delta t \left((\lambda + 2\mu) D_x^- \left[v_x^n \right] + \lambda D_y^- \left[v_y^n \right] \right) + \Delta t \frac{L^{n+} - L^{n-}}{\Delta t} \tag{3.12c}$$

$$\overline{\sigma}_{yy}^{n+} = \overline{\sigma}_{yy}^{n-} + \Delta t \left(\lambda D_x^- \left[v_x^n \right] + (\lambda + 2\mu) D_y^- \left[v_y^n \right] \right) + \Delta t \frac{L^{n+} - L^{n-}}{\Delta t}. \tag{3.12d}$$

3.3.2 Force source

If a vertical line force source along the z-axis is considered (Figure 3.1) the force density is given by

$$\vec{f}(\vec{x}, t) = F_0' \, s(t) \, \delta(x) \delta(y) \, \hat{e}_y \tag{3.13}$$

where F_0' is a line force density in N/m and $\hat{e}_y$ is a unit vector in y-direction (vertical direction).

In analogy to the previous section the discretized form of the force density for a vertical force source at (i, j) and time step n is

$$f_y^n(i,j) = \frac{F_0'}{\Delta x^2}\, s^n \tag{3.14}$$

and can be simply applied during the update of the vertical particle velocity via

$$v_y^{n+1}(i,j^+) = v_y^n(i,j^+) + \frac{\Delta t}{\rho(i,j^+)}\left(D_x^-\left[\sigma_{xy}^{n+}(i^+,j^+)\right] + D_y^+\left[\sigma_{yy}^{n+}(i,j)\right] + \frac{F_0'}{\Delta x^2}\, s^{n+}\right). \tag{3.15}$$

A force in x-direction can be implemented analogously as long as it is not located directly at the free surface. Otherwise the scaling must be adjusted. This special case is discussed in detail in Section 3.7.

3.4 Explosion source in a homogeneous full-space

3.4.1 Analytic solution

First, I will briefly repeat the analytic solution of the wave equation for an explosive line source along the z-axis at the origin of the coordinate system displayed in Figure 3.1. Due to the symmetry of the problem it is advisable to change over to cylindrical coordinates as shown in Figure 3.4. Only radial displacements u_r are observed in this specific problem. According to Friederich (1995) the Fourier transform of the displacement field $\vec{u} = \nabla\phi$ can be expressed by the compressional potential

$$\phi(r,\omega) = \frac{M_0' S(\omega)}{4iv_p^2\rho}\, H_0^{(2)}(\frac{\omega}{v_p}r) \tag{3.16}$$

with the seismic line moment M_0' in Nm/m, the Fourier transform $S(\omega)$ of the source time function in 1/Hz, the P-wave velocity v_p in m/s, the density ρ in kg/m^3 and the zeroth-order Hankel function of the second kind $H_0^{(2)}$.

Using the relation (Gradshteyn & Ryzhik, 1980, eq. 8.472, no. 2)

$$z\frac{d}{dz}H_\nu^{(2)}(z) - \nu H_\nu^{(2)}(z) = -zH_{\nu+1}^{(2)}(z) \tag{3.17}$$

the radial displacement is given by

$$u_r(r,\omega) = \frac{\partial}{\partial r}\phi(r,\omega) = -\frac{M_0' S(\omega)}{4iv_p^2\rho}\frac{\omega}{v_p}\, H_1^{(2)}(\frac{\omega}{v_p}r). \tag{3.18}$$

The radial particle velocity

$$v_r(r,\omega) = -\frac{M_0' i\omega S(\omega)}{4iv_p^2\rho}\frac{\omega}{v_p}\, H_1^{(2)}(\frac{\omega}{v_p}r). \tag{3.19}$$

is obtained by a multiplication of Equation (3.18) with $i\omega$ which corresponds to a time derivative. For the benchmark tests I calculate the analytic solution with Matlab.

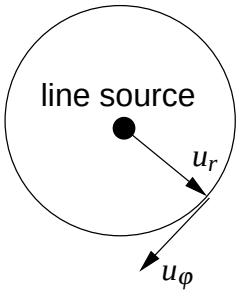


Figure 3.4: Cylindrical coordinate system used to describe the analytic solution for an explosive line source in a homogeneous full-space. The line source is marked by the black point and is located perpendicular to the plane of projection. Due to the symmetry only a radial displacement u_r is observed.

3.4.2 Benchmark model

A homogeneous full-space with a P-wave velocity of v_p=500 m/s, an S-wave velocity of v_s=300 m/s and a density of ρ=1800 kg/m^3 is used as benchmark model. The geometry of the model used in the FD simulations is sketched in Figure 3.5. The source time function $s(t)$ is given by

$$s(t) = \begin{cases} \sin^3\left(\frac{\pi t}{T_d}\right) & \text{for} \quad 0 < t < T_d \\ 0 & \text{otherwise} \end{cases} \tag{3.20}$$

with a source wavelet duration of T_d=32 ms (Figure 3.6). The maximal frequency is around 100 Hz. This corresponds to a minimum wavelength of 5 m for the P-waves. The seismic line moment is M_0'=1e+10 Nm/m.

3.4.3 Results

The benchmark tests carried out with the homogeneous full-space model are listed in Table 3.1. In this section I focus on the comparison of radial velocity seismograms rather than pressure seismograms since in the applied FWI particle displacements are inverted. Nevertheless, this benchmark test is also performed for pressure seismograms. The results are described in Appendix D.

Variation of the order of spatial FD operators

Figure 3.7 displays the results for three simulations with different orders of the spatial FD operators. No significant differences can be observed in the seismograms, neither in amplitude nor in phase. As the temporal sampling interval Δt is different for all the tests this comparison also proofs that the result does not depend on the chosen Δt.

Variation of the spatial grid spacing

Due to the geometry of the problem the x-component (radial component) which is recorded by the receivers must coincide with the particle velocities calculated by Equation (3.19). Figure 3.8

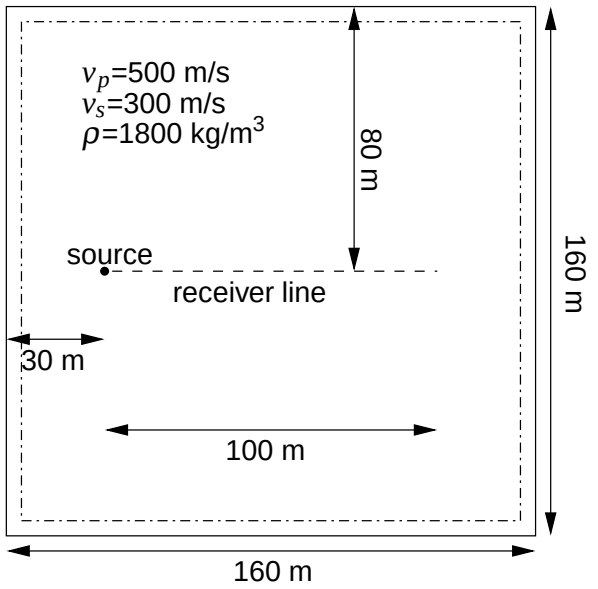


Figure 3.5: Model of a homogeneous full-space used for benchmarking. The absorbing frame (CPML layer; Komatitsch & Martin, 2007) is marked by the dash-dotted line. It has a thickness of 20 grid points.

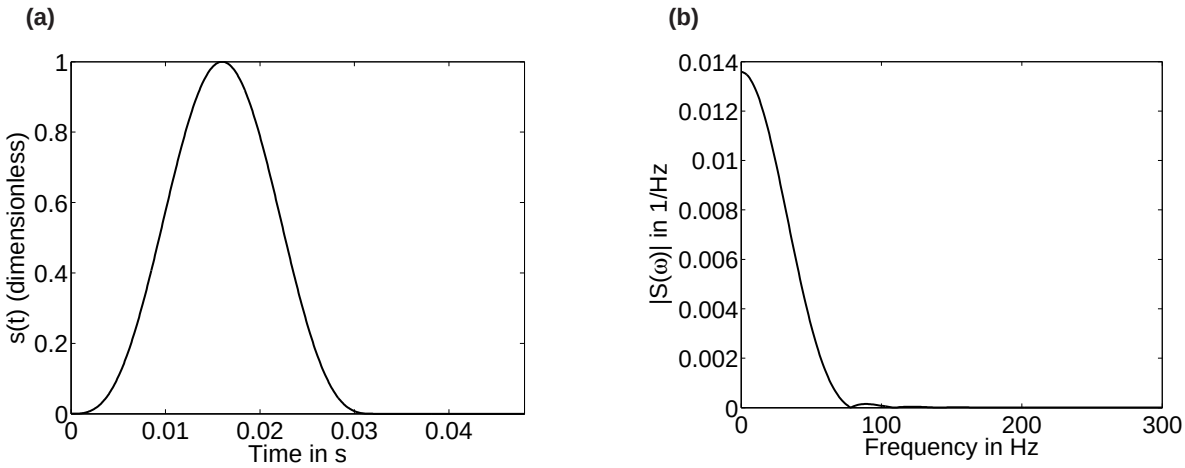


Figure 3.6: Source time function used for benchmark tests. a) displays the source wavelet in the time-domain and b) displays the amplitude spectrum.

No.	Model size (N_x x N_y)	Δx in m	Order spatial FD operator	Δt in s	Grid points per min. wavelength
1	640 x 640	0.25	2	3.0e-04	20
2	1280 x 1280	0.125	2	1.5e-04	40
3	640 x 640	0.25	4	2.9e-04	20
4	640 x 640	0.25	6	2.8e-04	20

Table 3.1: Modeling parameters used for benchmark tests with homogeneous full-space model.

displays the comparison of the analytically calculated seismograms with seismograms generated with DENISE with different spatial grid spacings (modeling no. 1 and 2 in Table 3.1). The phases for all offsets match very good. However, the amplitudes do not coincide for offsets smaller than 1 m. The amplitude differences at small offsets are predominantly caused by the staggered grid that is used in the FD code (Figure 3.2).

For an explosive source the application of a staggered grid has the following effects:

- The source is excited at a full grid point by adding the compressional potential L (see Equation (3.9)) to the main diagonal components of the stress tensor σ_{xx} and σ_{yy}. Therefore, the modeled particle velocity seismograms of the x component shown in Figure 3.8 correspond to the particle velocity of the radial component concerning the direction of the particle movement. However, the actual offsets of the seismograms generated with DENISE is $\Delta x/2$ larger. By recalculating the radial velocity seismograms with Equation (3.19) with the actual offsets of the FD simulated seismograms the amplitude misfit in the near offset traces reduces significantly (Figure 3.9).

- The grid points corresponding to the vertical particle velocity v_y are shifted by $\Delta x/2$ in y-direction in comparison to the full grid points (Figure 3.2). Therefore, they are not zero. Again Equation (3.19) can be used to calculate the analytic result. According to Figure 3.10 the offset r corresponding to the location of v_y is

$$R = \sqrt{x^2 + (\Delta x/2)^2}. \tag{3.21}$$

  The analytically calculated radial velocity $v_r(r = R)$ points along the source-receiver connection line. The corresponding y-component which is calculated by the FD code is

$$v_y(R) \;=\; v_r(R)\,\sin(\alpha) \;=\; v_r(R)\,\frac{\Delta x}{2R}. \tag{3.22}$$

  The vertical velocity seismograms are compared in Figure 3.11. All seismograms fit well concerning phase. There are some discrepancies observed in amplitudes (e. g. the seismograms at 0.5 m to 1.0 m offset for a grid spacing of 0.25 m). These differences decrease if higher orders of spatial FD operators are used. They are therefore possibly related to inaccuracies of the FD results due to grid dispersion. However, this was not analyzed in more detail. As expected, the vertical velocity seismograms are smaller in amplitude for smaller grid spacings.

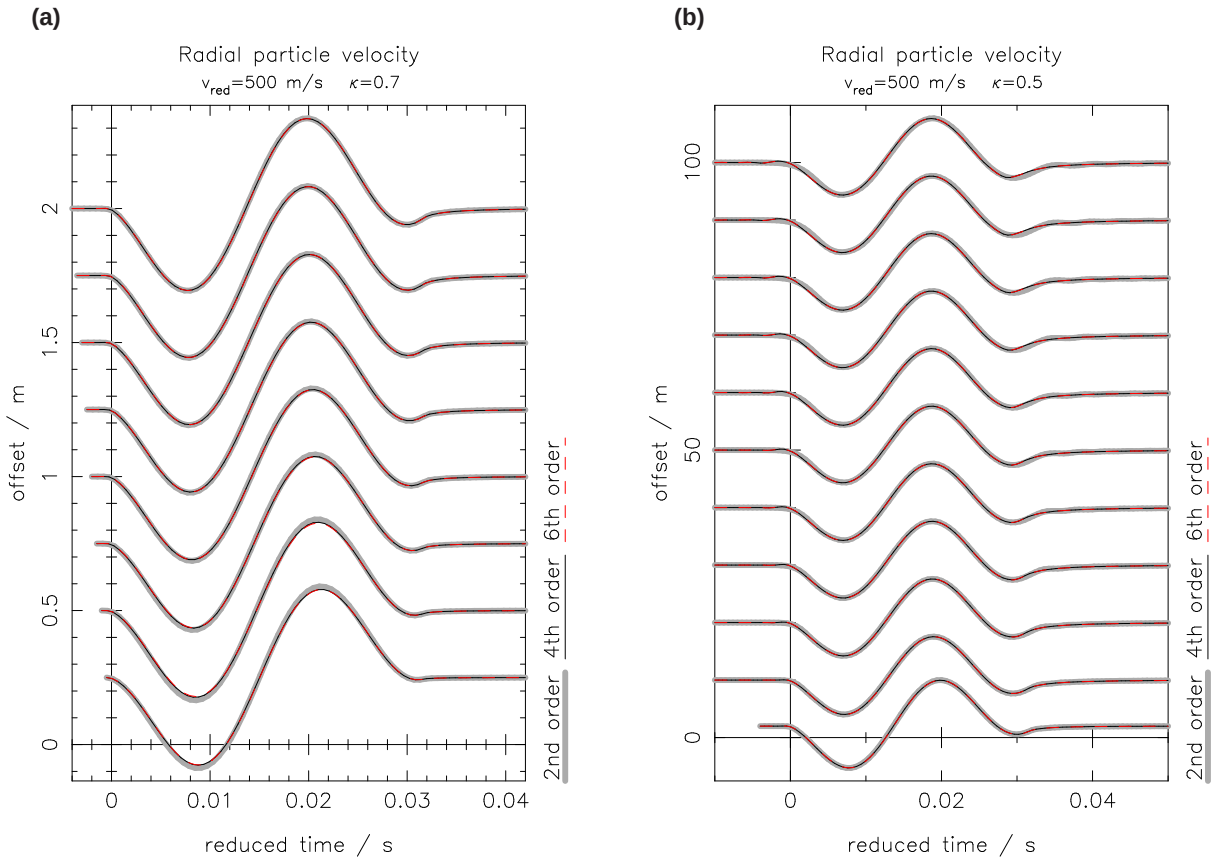


Figure 3.7: Radial particle velocity seismograms for an explosive line source in a homogeneous full-space with varying order of the spatial FD operators used in the simulations (see modelings no. 1, 3 and 4 from Table 3.1). The time axes in both plots are reduced with a velocity of 500 m/s. The seismograms are not normalized but they are multiplied by an offset dependent factor of $(r/\mathrm{m})^{\kappa}$ where κ is given in the head of each figure. a) displays seismograms with very small offsets and b) displays seismograms over the whole offset range.

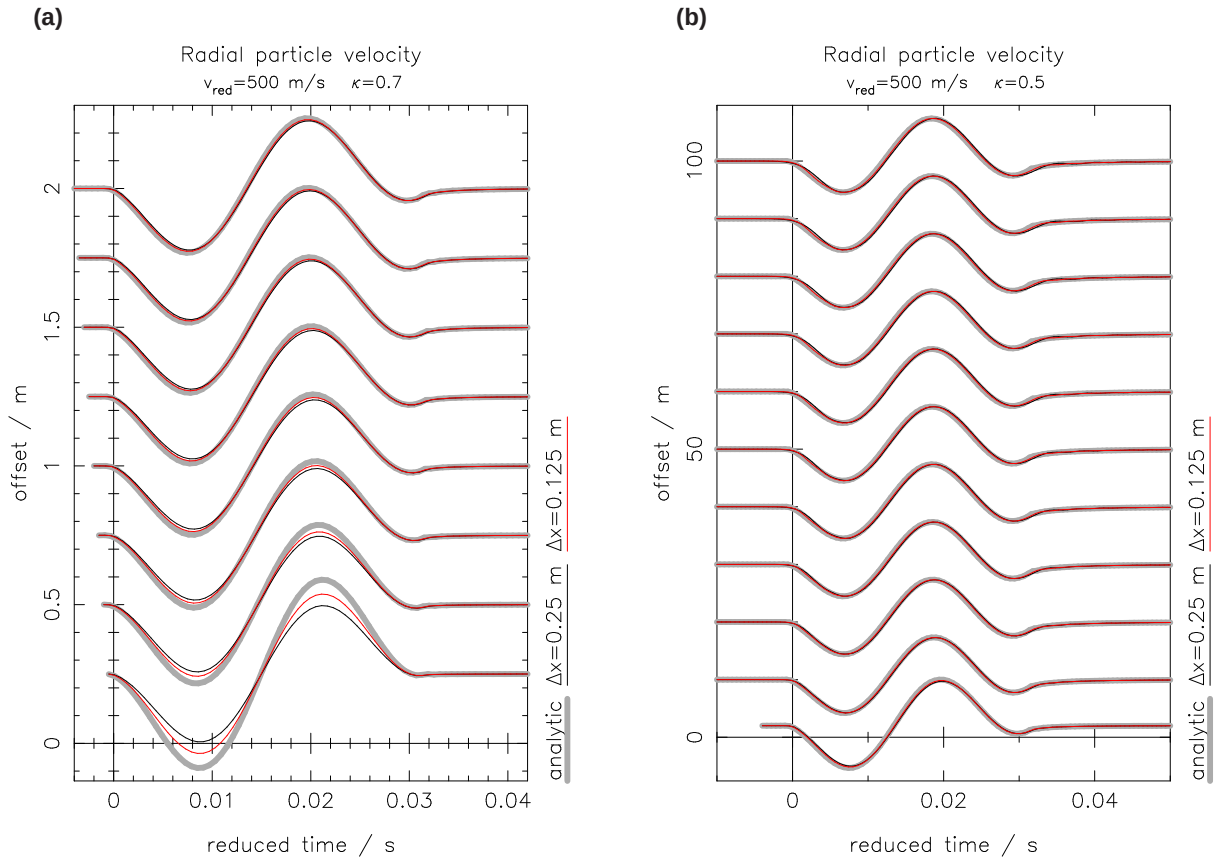


Figure 3.8: Radial particle velocity seismograms for an explosive line source in a homogeneous full-space. Comparison of analytic results with seismograms simulated with different spatial grid spacing Δx (see modelings no. 1 and 2 from Table 3.1). The time axes in both plots are reduced with a velocity of 500 m/s. The seismograms are not normalized but they are multiplied by an offset dependent factor of $(r/\mathrm{m})^{\kappa}$ where κ is given in the head of each figure. a) displays seismograms with very small offsets and b) displays seismograms over the whole offset range.

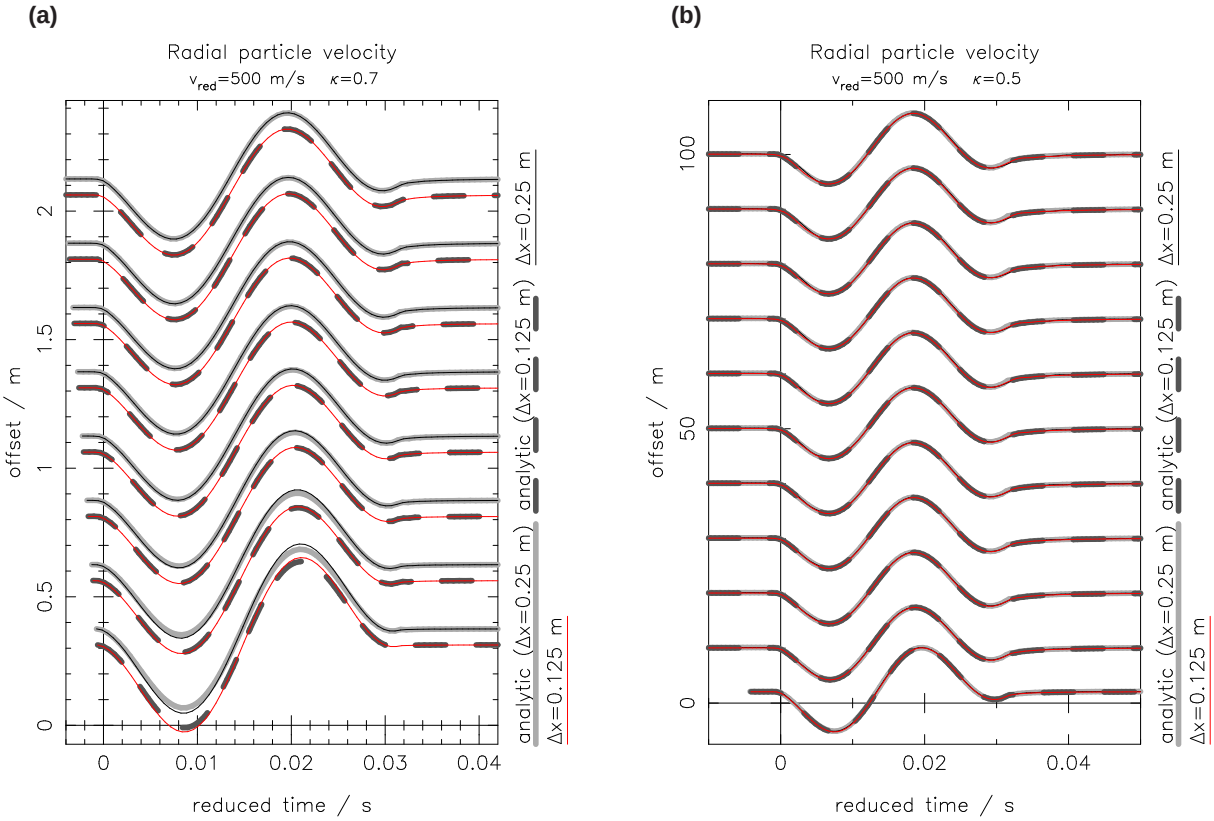


Figure 3.9: Radial particle velocity seismograms for an explosive line source in a homogeneous full-space. Comparison of analytic results with seismograms simulated with different spatial grid spacing (see modelings no. 1 and 2 from Table 3.1). The analytic seismograms are calculated at the actual offsets of the modeled seismograms and therefore consider the influence of the staggered grid. The time axes in both plots are reduced with a velocity of 500 m/s. The seismograms are not normalized but they are multiplied by an offset dependent factor of $(r/\mathrm{m})^{\kappa}$ where κ is given in the head of each figure. a) displays seismograms with very small offsets and b) displays seismograms over the whole offset range.

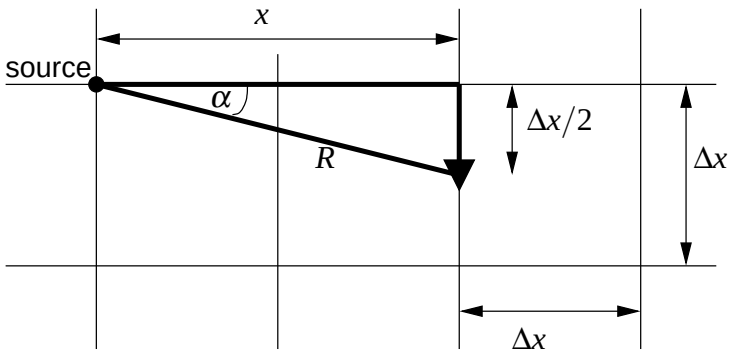


Figure 3.10: Sketch for the derivation of the analytic solution for the vertical velocity seismograms in the benchmark test of an explosive source in a homogeneous full-space. The receiver for the vertical particle velocity wavefield is marked by a black downward-pointing triangle.

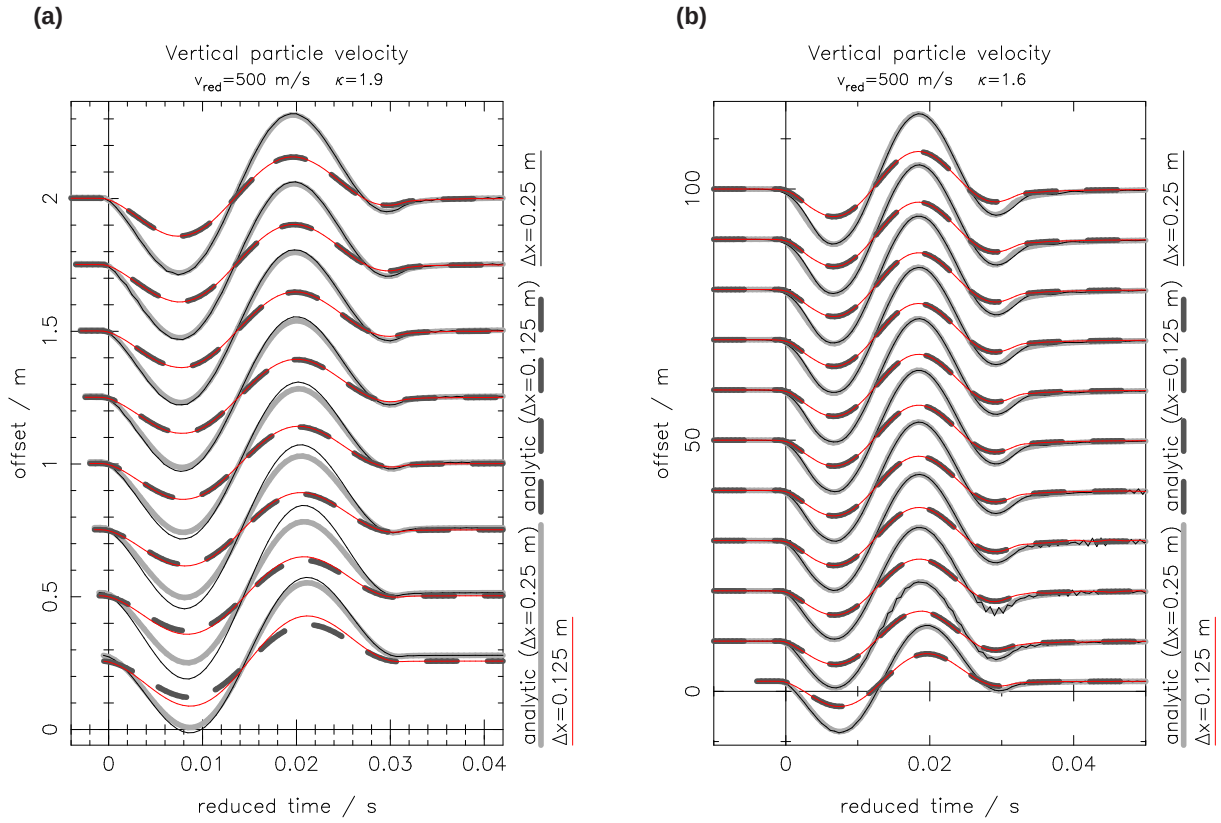


Figure 3.11: Vertical particle velocity seismograms for an explosive line source in a homogeneous full-space. Comparison of analytic results with seismograms simulated with different spatial grid spacing (see modelings no. 1 and 2 from Table 3.1). The analytic seismograms are calculated at the actual offsets of the modeled seismograms and therefore consider the influence of the staggered grid. The time axes in both plots are reduced with a velocity of 500 m/s. The seismograms are not normalized but they are multiplied by an offset dependent factor of $(r/\mathrm{m})^{\kappa}$ where κ is given in the head of each figure. a) displays seismograms with very small offsets and b) displays seismograms over the whole offset range.

3.4.4 Conclusion

The forward modeling code is successfully benchmarked for an explosive line source in a homogeneous full-space. For very small offsets the influence of the staggered grid must be taken into account to explain also amplitudes precisely. However, the influence of the staggered grid is only observable up to an offset of approximately one-fifth of the minimum wavelength and is therefore negligible in an FWI.

3.5 Vertical force at the surface of a homogeneous half-space

3.5.1 Analytic solution

A vertical line force along the z-axis is given by the force density described in Equation (3.13). The x-component of the particle velocity (Figure 3.1) at the surface of a homogeneous half-space excited by such a source can be expressed by (Ewing et al., 1957)

$$v_x(x,\omega) = F_0' S(\omega) i\omega \int_0^{+\infty} \frac{2}{\omega u} G_x(\omega,u) \sin(u\omega x) \mathrm{d}u \tag{3.23}$$

with the expansion coefficients

$$G_x(\omega,u) = \frac{\omega}{2\pi\mu} \frac{u^2(u^2-b^2+2ab)}{(u^2-b^2)^2+4u^2ab} \tag{3.24}$$

with the horizontal slowness u and the vertical components a and b of the slowness of P-waves and S-waves, respectively. The slowness components a, b, and u are linked by

$$a^2 + u^2 = 1/v_p^2 \tag{3.25a}$$

$$b^2 + u^2 = 1/v_s^2 \tag{3.25b}$$

where v_p is the P-wave velocity and v_s is the S-wave velocity.
The corresponding vertical component of the particle velocity is (Ewing et al., 1957)

$$v_z(x,\omega) = -F_0' S(\omega) i\omega \int_0^{+\infty} \frac{2}{\omega u} G_z(\omega,u) \cos(u\omega x) \mathrm{d}u \tag{3.26}$$

with the expansion coefficients

$$G_z(\omega,u) = \frac{\omega}{2\pi\mu} \frac{iua}{v_s^2(u^2-b^2)^2+4u^2ab}. \tag{3.27}$$

3.5.2 Numerical calculation of analytic solution

The analytic solution is implemented in a Matlab program in analogy to the Fortran program “lamb.f” of Forbriger (1996). The slowness integrals in Equations (3.23) and (3.26) are approximated by the trapezoidal rule. As the integration can only be done up to a finite slowness and not up to infinity a cut-off phase occurs in the calculated seismograms (Forbriger, 2003c, section 2.4). To decrease its amplitude the integrands are multiplied by a slowness dependent taper. The cut-off phase mainly disturbs the waveforms in the near-offset range.

No.	Implementation of free surface	Model size (N_x x N_y)	Δx in m	Order spatial FD operator	Δt in s	Grid points per min. wavelength
1	image technique	700 x 75	0.2	2	2.8e-04	13.75
2	image technique	1400 x 150	0.1	2	1.4e-04	27.5
3	image technique	1400 x 150	0.1	4	1.2e-04	27.5
4	air layer	700 x 100	0.2	2	2.8e-04	13.75
5	air layer	1400 x 200	0.1	2	1.4e-04	27.5
6	vacuum layer	700 x 100	0.2	2	2.8e-04	13.75
7	vacuum layer	1400 x 200	0.1	2	1.4e-04	27.5

Table 3.2: Modeling parameters used for benchmark tests with homogeneous half-space model.

3.5.3 Benchmark model

A homogeneous half-space with a P-wave velocity of v_p=500 m/s, an S-wave velocity of v_s=300 m/s and a density of ρ=1800 kg/m^3 is used as benchmark model. The non dispersive Rayleigh wave propagates in this model with a velocity of approximately 275 m/s. The free surface is implemented by three different techniques. An explicit implementation by the image technique (Section 3.2.2) and an implicit implementation by adding on the one hand a five meter thick air layer at the top of the model and on the other hand a five meter thick vacuum layer (Section 3.2.1). Within the air layer I assume v_p=331.8 m/s, v_s=0 m/s and ρ=1.3 kg/m^3 whereas the vacuum layer is approximated by v_p=v_s=0 m/s and ρ=1e-06 kg/m^3.[2] The geometry of the model is sketched in Figure 3.12. The source time function $s(t)$ is given by Equation (3.20) with a duration of T_d=32 ms (Figure 3.6). This source wavelet approximates the source time function of a hammer blow (Forbriger, 2004). The maximum frequency is around 100 Hz. This corresponds to a minimum wavelength of 2.75 m for the Rayleigh-wave. The line force is F_0'=1e+4 N/m.

3.5.4 Results

I perform several forward modeling runs which are summarized in Table 3.2. The differences of the simulated seismograms are small. The waveforms do not depend on the grid spacing Δx, the sampling interval Δt or the order of the spatial FD operators. For the comparison of the semi-analytically calculated seismograms with the modeled seismograms I focus on the different implementations of the free surface. The differences in the seismograms simulated with an air layer and a vacuum layer are insignificant. Figure 3.13 displays velocity seismograms calculated with Equations (3.23) and (3.26) and modeled seismograms calculated on the one hand with the image technique (explicit free surface, simulation no. 2 in Table 3.2) and on the other hand with an air layer (implicit free surface, simulation no. 5). The seismograms of both components match very well in amplitude and phase. The small differences especially in the near offset traces are due to the disturbance of the analytically calculated seismograms by the cut-off phase.

[2] It is not possible to set density to zero because during the update of the velocity wavefield it is divided by density and therefore numerical instabilities occur if the density is set exactly to zero. Furthermore, there is a division by the S-wave velocity v_s to calculate the averaged shear modulus $< \mu >$ in the mid of the grid cells. However, in the DENISE code it is checked whether the S-wave velocity at one of the four surrounding grid points is zero. If this is the case, the Lamé parameter $< \mu >$ in the mid of the grid cell is set to zero explicitly.

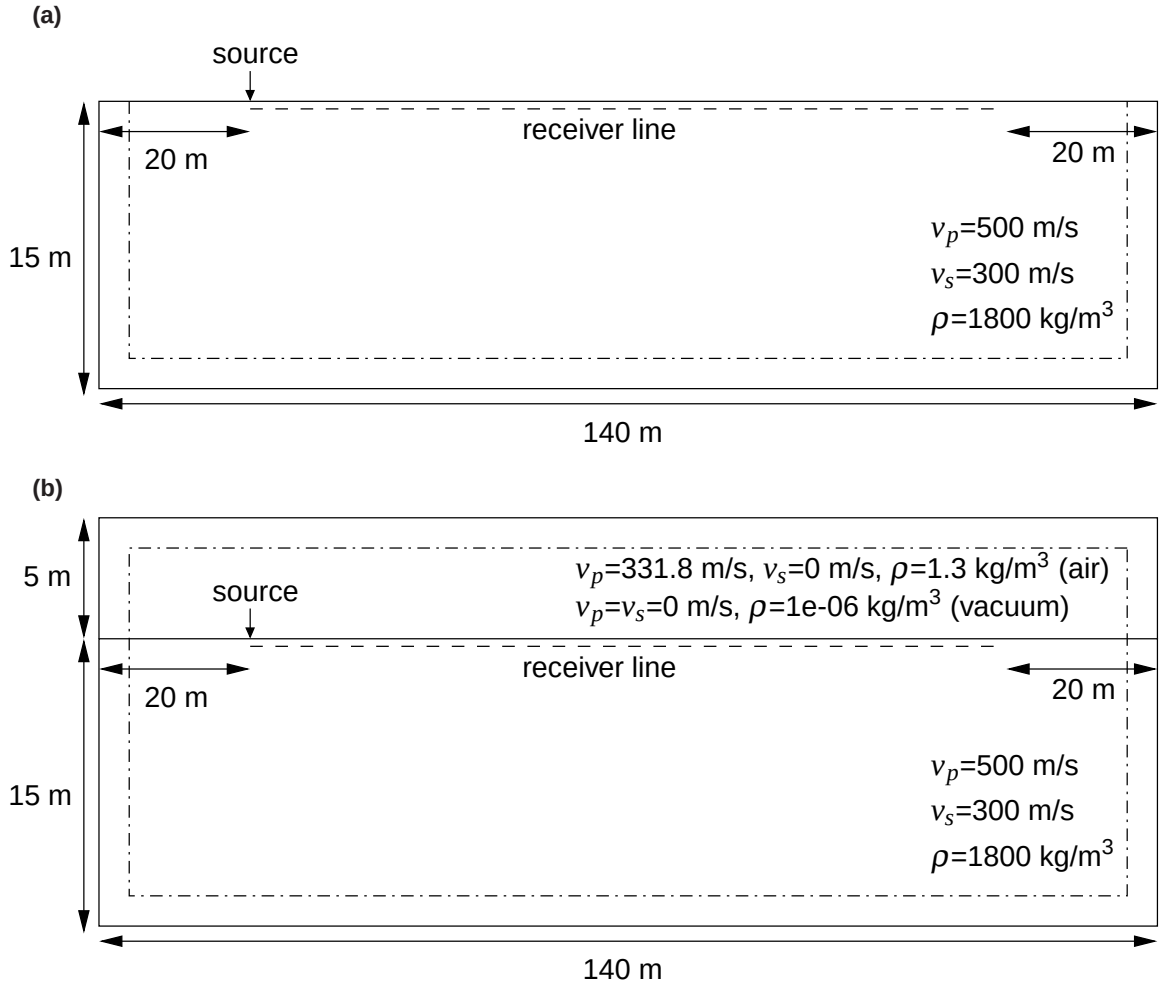


Figure 3.12: Model of a homogeneous half-space used for code benchmarking. a) displays the model used for the tests with an explicitly implemented free surface by the image technique. b) displays the model used for the tests with an implicitly implemented free surface by adding a vacuum or air layer at the top of the model. The absorbing frame (CPML layer; Komatitsch & Martin, 2007) is marked by the dash-dotted line. It has a thickness of 20 grid points.

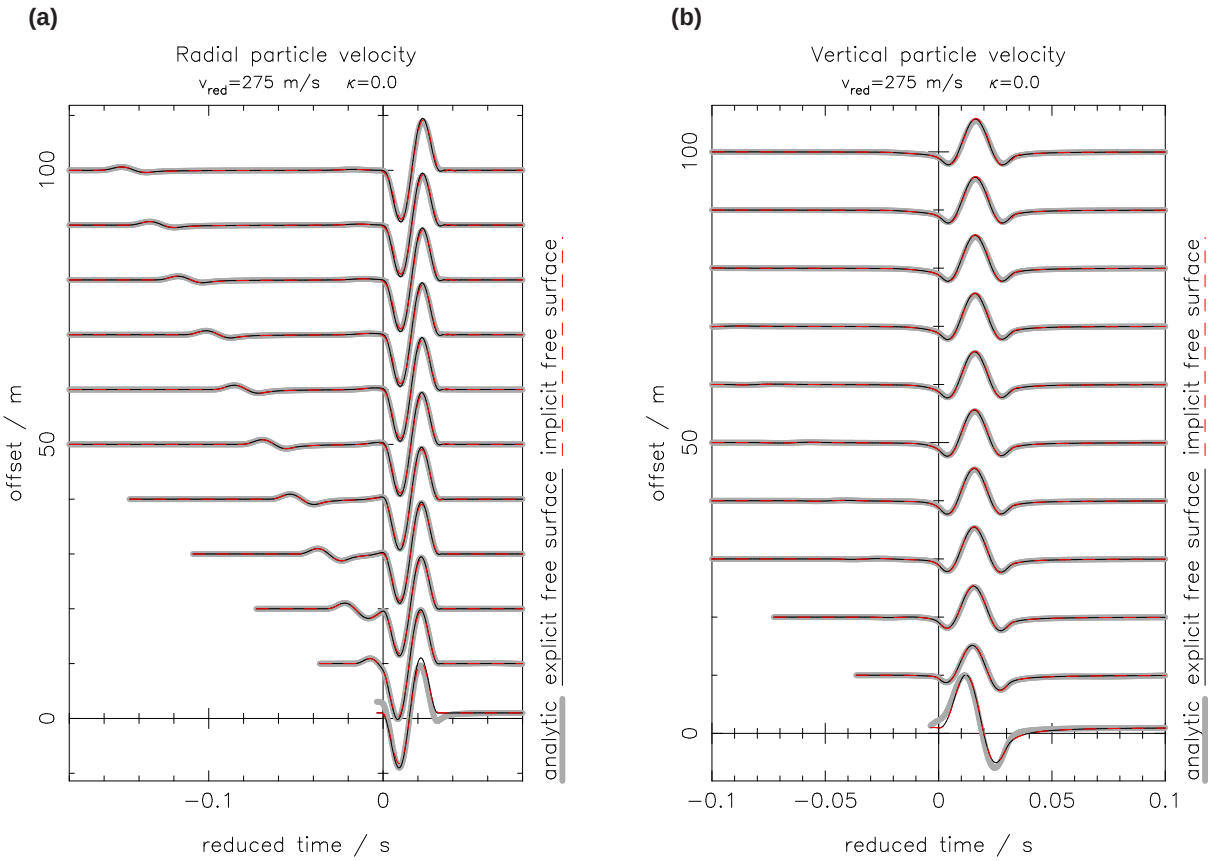


Figure 3.13: Benchmark result for a vertical line force at the surface of a homogeneous half-space. Comparison of semi-analytically calculated seismograms with seismograms simulated with different implementations of the free surface. a) displays the horizontal velocity seismograms and b) the vertical velocity seismograms. Due to the small amplitude of the P-wave in comparison to the Rayleigh wave in the vertical component the P-wave is cut off in some traces in b). The time axes in both plots are reduced with a velocity of 275 m/s. The seismograms are not normalized.

	Thickness in m	v_p in m/s	v_s in m/s	ρ in kg/m^3
Air layer (if used)	5	331.8	0	1.3
Layer	3	680	320	1700
Half-space	47	1000	590	2000

Table 3.3: Model parameters for the benchmark model of a homogeneous layer over a homogeneous half-space.

3.6 Vertical force at the surface of a homogeneous layer-over-half-space model

The benchmark tests shown in this section will give an idea about the location of the free surface and the discontinuity between the layer and the half-space for both an explicit and an implicit implementation of the free surface. In this test I compare the seismograms simulated with the DENISE code with seismograms simulated with the reflectivity method (Fuchs & Müller, 1971; Müller, 1985) in an implementation by Ungerer (1990).

3.6.1 Benchmark model

The model consists of a homogeneous layer over a homogeneous half-space. The material parameters are listed in Table 3.3. By calculating the dispersion curves for this model for a vertical force at the surface a minimum phase velocity of around 300 m/s for the fundamental mode of the Rayleigh wave is obtained. The source time function is given by Equation (3.20) with a duration of T_d=32 ms (Figure 3.6) and a line force of F_0'=1e+04 N/m is used. The minimum wavelength is approximately 3 m. The acquisition geometry used for the benchmark test is analogous to the one used for the homogeneous half-space model (Figure 3.12).

3.6.2 Results

For these benchmark tests it is important to consider the effective thicknesses of the layer in the case of an implementation of the free surface by the image technique and by an air layer. This is shown in Figure 3.14. I consider a layer of N grid points, therefore the layer thickness is $N \cdot \Delta x$. The effective layer thickness is again $N \cdot \Delta x$ if the free surface is implemented by an air or vacuum layer because the free surface as well as the discontinuity between the layer and the half-space are located between the grid rows where the material parameters change (Figure 3.14a). In contrast, if the free surface is implemented by the image technique the free surface is located directly at the first grid row (Figure 3.14b). Therefore, the effective thickness of the layer is only $(N - 0.5)\Delta x$.

In the benchmark model used in this test I assume a layer thickness of 3 m. I simulate seismograms with a spatial grid spacing of Δx=0.2 m and Δx=0.1 m. Therefore, the effective layer thicknesses in the simulations using the image technique are 2.90 m and 2.95 m. Although the maximum difference in the thickness is only 10 cm this causes a change in the dispersion of the

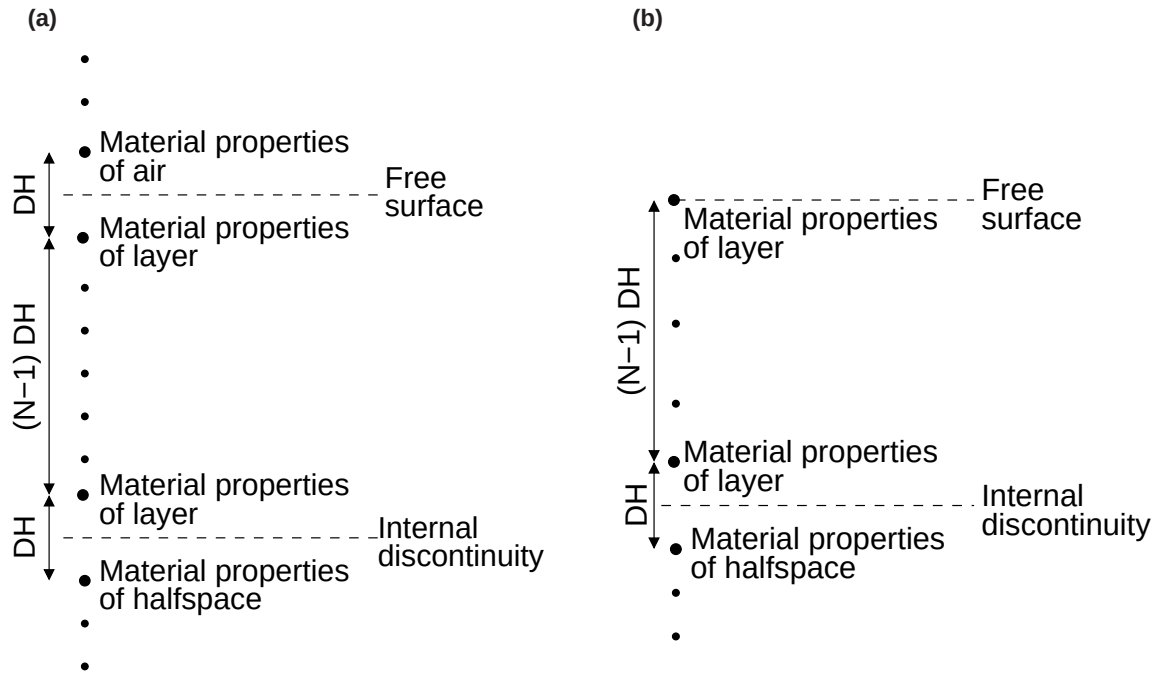


Figure 3.14: Sketch of a vertical section through the model of a layer over a half-space with a) a free surface implemented by an air layer and b) a free surface implemented by the image technique. The layer is represented by N grid rows. If the image technique is used the effective layer thickness depends on the spatial grid spacing $\Delta x = DH$ used in the simulation.

Rayleigh wave which can be observed as a significant phase shift (Figure 3.15a). In the comparison of the simulated FD seismograms with the reflectivity seismograms with corresponding effective layer thicknesses a good fit in amplitude and in phase is observed (Figure 3.15b-d). Figure 3.15 displays only vertical velocity seismograms, however, the same effects can be observed in the radial component of the velocity wavefield.

3.7 Horizontal line force at the surface of a homogeneous half-space

When FWI is applied to shallow seismic Rayleigh waves the forward propagated wavefields are excited by vertical force sources which approximate the vertical hammer blows. Nevertheless, in the tests where two-component receivers are used adjoint sources of the horizontal component (x-component) are injected during backpropagation into the FD grid. Therefore, also the implementation of a directed force source in horizontal or x-direction, respectively, must be benchmarked.

3.7.1 Benchmarking using the source-receiver reciprocity

Due to a lack of codes that are able to simulate seismograms for a horizontal line force at the free surface of a medium I apply the source-receiver reciprocity for benchmarking. Due to Equa-

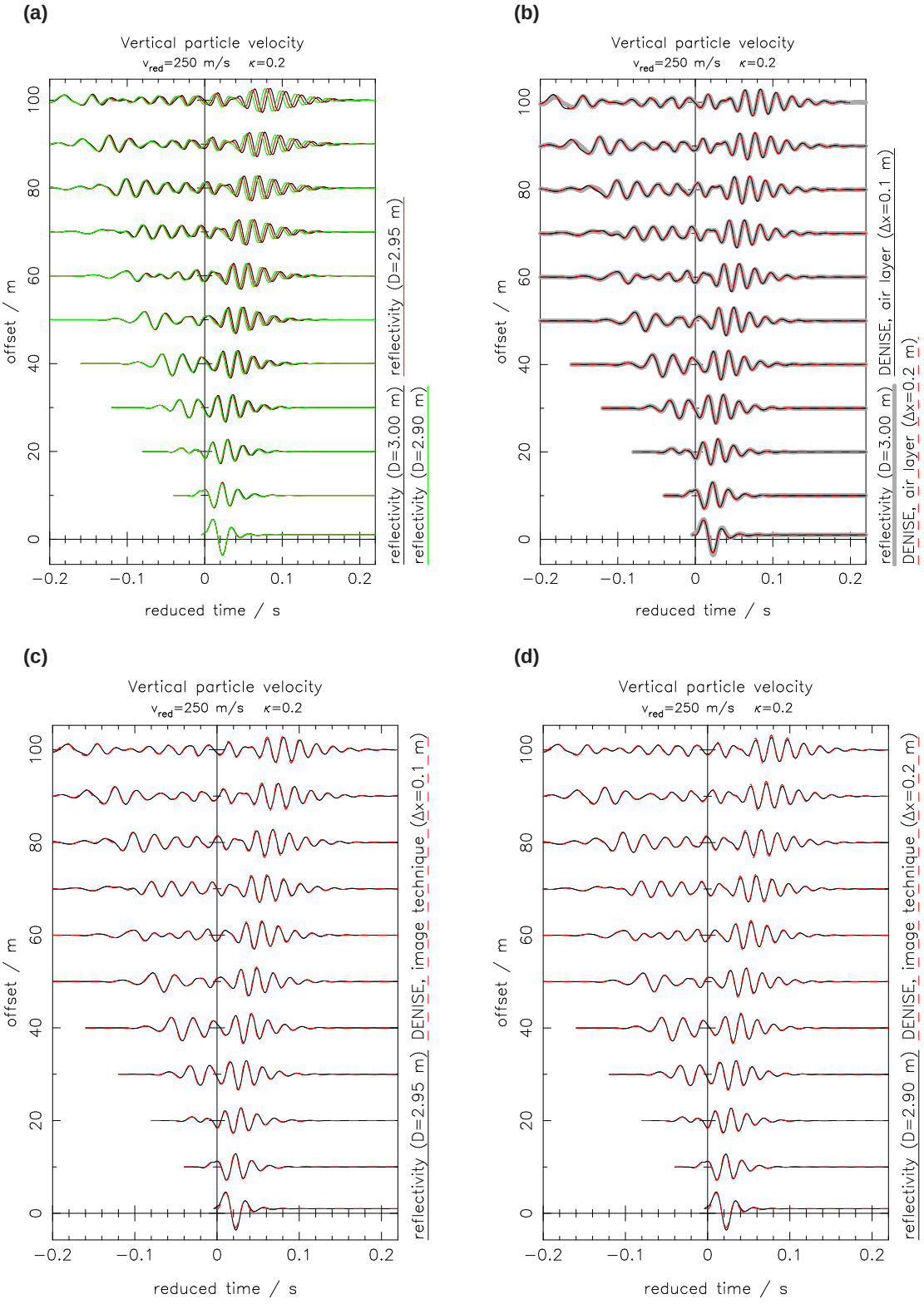


Figure 3.15: Benchmark result for a vertical line force at the surface of a model consisting of a homogeneous layer over a homogeneous half-space (Table 3.3). a) displays vertical velocity seismograms for three different layer thicknesses D calculated with the reflectivity method. b) displays a comparison of the reflectivity result and seismograms calculated with DENISE for different spatial grid spacings where the free surface is implemented by an air layer in the model. c) and d) display the reflectivity seismograms together with the corresponding FD seismograms simulated with different spacial grid spacings where the free surface is implemented by the image technique. The time axes in all plots are reduced with a velocity of 250 m/s. The seismograms are not normalized but they are multiplied by an offset dependent factor of $(r/\mathrm{m})^{\kappa}$ where κ is given in the head of each figure.

tion (2.20) it is

$$s_l(\vec{x},t) = s_n(\vec{x}',t) \tag{3.28}$$

with

$$s_l(\vec{x},t) = \int_V \int_{-\infty}^{\infty} G_{ln}(\vec{x},t-t';\vec{x}',0)\, f_n(\vec{x}',0)\mathrm{d}t'\mathrm{d}^3\vec{x}' \tag{3.29a}$$

$$s_n(\vec{x}',t) = \int_V \int_{-\infty}^{\infty} G_{nl}(\vec{x}',t-t';\vec{x},0)\, f_l(\vec{x},0)\mathrm{d}t'\mathrm{d}^3\vec{x}. \tag{3.29b}$$

Therefore, the seismograms must coincide if I run two tests where I interchange the locations of source and receiver as well as the components of them.

3.7.2 Statement of the problem

I run the benchmark test for a source-receiver pair with 55 m offset at the surface of a homogeneous elastic half-space with a P-wave velocity of v_p=500 m/s, an S-wave velocity of v_s=300 m/s and a density of ρ=1800 kg/m^3. The source time function is given by Equation (3.20) with a duration of T_d=32 ms (Figure 3.6) and a line force density of F_0'=1e+04 N/m. The free surface is implemented by the image technique (Section 3.2.2).

The first four tests are explained in Table 3.4 (Test 1-4) and compared in Figure 3.16. In these tests sources and receivers are placed directly at the free surface or half a grid point below the free surface due to the staggered grid. The results of Test 1 and Test 4 where only x and y components are compared, respectively, coincide in both, amplitude and phase. However, the seismograms of the two tests where the components are mixed (Test 2 and Test 3) differ significantly in amplitudes. Since I have already successfully benchmarked the radial and vertical velocity seismograms excited by a vertical force source (Section 3.5 and 3.6) it can be concluded that the amplitudes are too small when the medium is excited by a horizontal force source directly at the free surface.

I repeat these tests but place the source and the receiver at the second grid point. Therefore, neither the vertical force source nor the horizontal force source are at the free surface. These tests are described by Test 5 to 8 in Table 3.4 and the results are displayed in Figure 3.17. In this case the seismograms of all four tests match in both, amplitudes and phases as the seismograms excited by the horizontal forces have the correct amplitudes and are not too small by a factor of two.

The results of Tests 1 to 8 do neither depend on the spatial discretization (as long as no significant grid dispersion occurs) nor on the order of the spatial FD operators. This indicates that there exists a scaling problem when horizontal force sources are used directly at the free surface together with an implementation of the free surface by the image technique.

I also run these tests for a free surface implemented by an air layer. In this case the free surface is located between the two grid rows j and $j+1$ where the j-th grid row belongs to the air layer and the $(j+1)$-th grid row belongs to the half-space (see Section 3.6). If an air layer is used it is therefore only possible to simulate vertical force sources directly at the free surface in DENISE due to the distribution of wavefield variables within the staggered grid implemented in this code.

	Test						Reciprocal test					
	Source			Receiver			Source			Receiver		
	x_s in m	y_s in m	comp	x_r in m	y_r in m	comp	x_s in m	y_s in m	comp	x_r in m	y_r in m	comp
Test 1	10.025	0.0	x	65.025	0.0	x	65.025	0.0	x	10.025	0.0	x
Test 2	10.025	0.0	x	65.0	0.025	y	65.0	0.025	y	10.025	0.0	x
Test 3	10.0	0.025	y	65.025	0.0	x	65.025	0.0	x	10.0	0.025	y
Test 4	10.0	0.025	y	65.0	0.025	y	65.0	0.025	y	10.0	0.025	y
Test 5	10.025	0.05	x	65.025	0.05	x	65.025	0.05	x	10.025	0.05	x
Test 6	10.025	0.05	x	65.0	0.075	y	65.0	0.075	y	10.025	0.05	x
Test 7	10.0	0.075	y	65.025	0.05	x	65.025	0.05	x	10.0	0.075	y
Test 8	10.0	0.075	y	65.0	0.075	y	65.0	0.075	y	10.0	0.075	y

Table 3.4: Benchmark tests using the source-receiver reciprocity principle. All shifts of wavefield variables due to the staggered grid (Figure 3.2) are already considered. The tests are carried out with a spatial grid spacing of Δx=0.05 m. The free surface is implemented by the image technique.

For the tests I place vertical forces and receivers directly at the free surface and horizontal forces and receivers half a grid point below the free surface. The corresponding seismograms coincide and no scaling problem is observed.

3.7.3 Proposal for solution

Further investigations with a shifted staggered grid (Appendix E) show that the observed problem is not specific for horizontal force sources but for directed force sources in general that are excited directly at the free surface together with an implementation of the free surface by the image technique. The simulated seismograms are too small by a factor of two. The external force must be translated to a change in particle velocity which can be afterwards applied in the update of the particle velocity wavefield. At this point the missing factor of two originates (see Appendix E for details). I have no well-founded explanation yet what is physically wrong in the implementation of a force source directly at the free surface combined with the image technique. Therefore, I will just propose a possible explanation in the following.

If the implicit implementation of the free surface is used by adding a vacuum layer at the top of the model the grid points with the vertical particle velocity at the free surface are surrounded by full grid cells (Figure 3.18) with physically realistic parameters (e. g. σ_{yy} is zero above the free surface). Thus, the integral notation of the wave equation

$$\int_V \rho \frac{\partial v_y}{\partial t} \mathrm{d}V = \int_V \left[\frac{\partial \sigma_{xy}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y} \right] \mathrm{d}V + \int_V F_y \delta(x) \delta(y) \mathrm{d}V \tag{3.30}$$

with

$$\int_V \mathrm{d}V = \int_{x=-\Delta x/2}^{\Delta x/2} \int_{y=-\Delta x/2}^{\Delta x/2} \mathrm{d}y \mathrm{d}x \tag{3.31}$$

refers to a whole grid cell where F_y is the applied line force in N/m. As in this case a full grid cell is used the density within this volume is approximated by arithmetic averaging in the heterogeneous FD approach (Moczo et al., 2002; Moczo, 1989; Graves, 1996). Each of the δ-functions in the source term can be approximated in the FD approach by $1/\Delta x$ (Figure 3.3) and the source implementation via Equation 3.14 is correct.

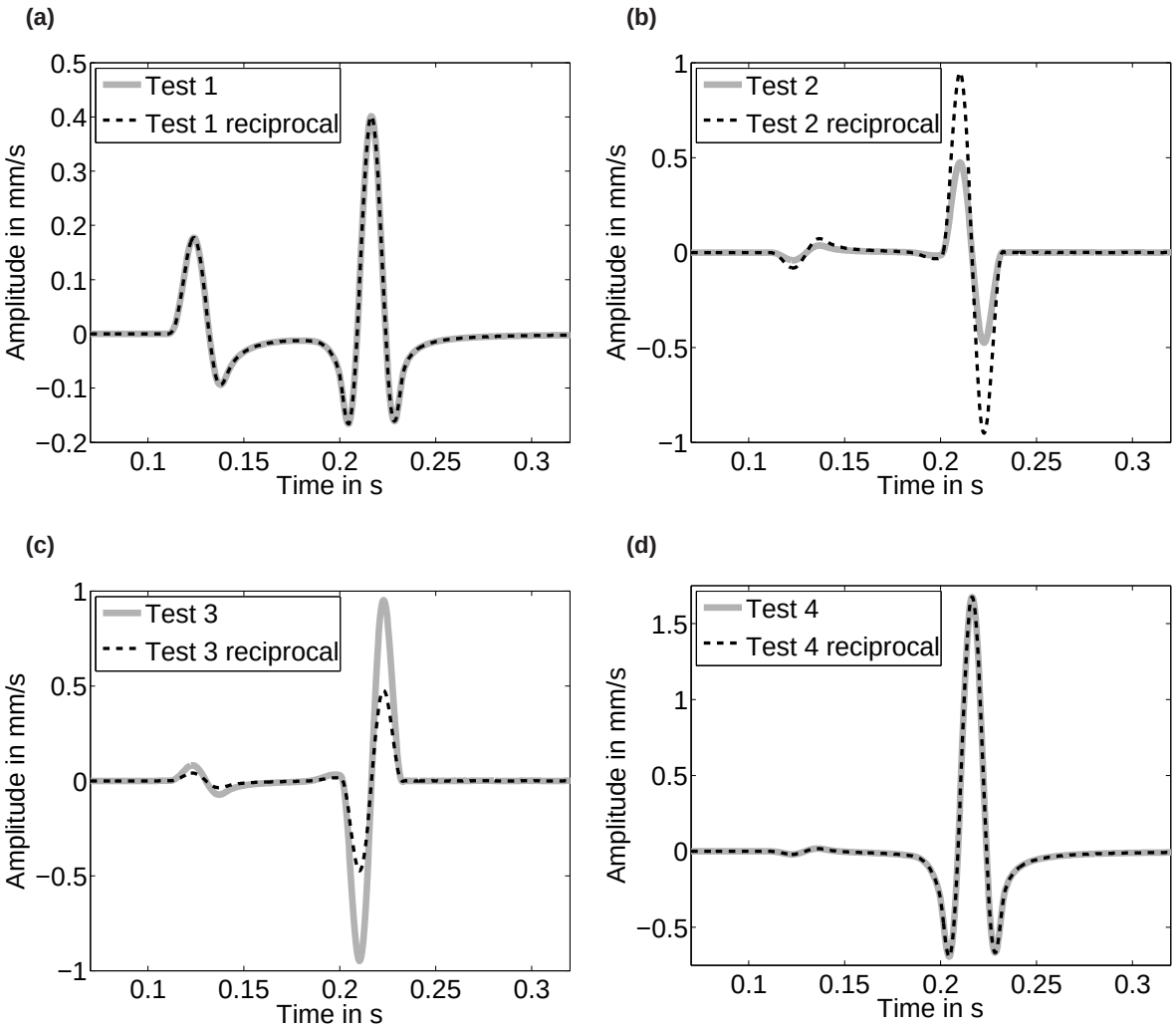


Figure 3.16: Benchmark tests 1 to 4 using source-receiver reciprocity principle. The tests are described in Table 3.4. The used source wavelet was a $\sin^3$-function with a duration of 32 ms and a maximal strength of 10 kN/m.

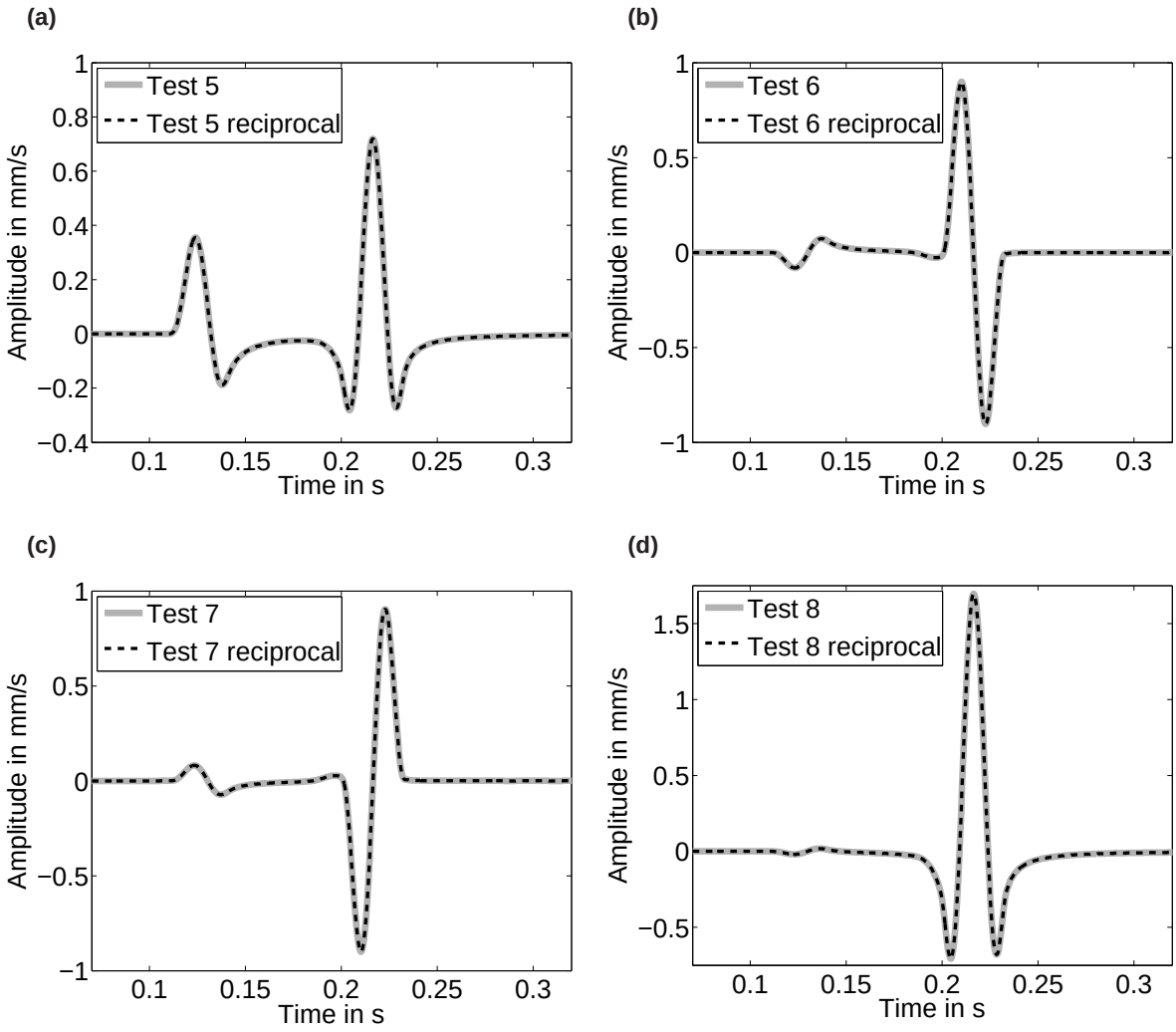


Figure 3.17: Benchmark tests 5 to 8 using source-receiver reciprocity principle. The tests are described in Table 3.4. The used source wavelet was a $\sin^3$-function with a duration of 32 ms and a maximal strength of 10 kN/m.

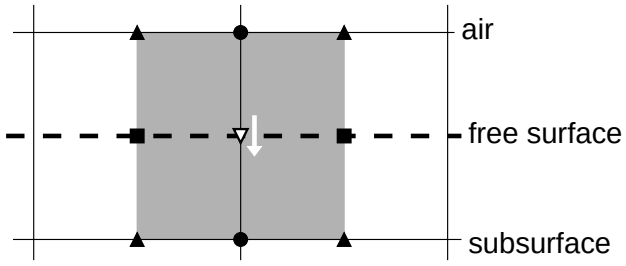


Figure 3.18: Sketch explaining appropriate source scaling for a vertical force source at the free surface where the free surface is implemented by an air layer. The grey area displays the area that approximates the δ-functions $\delta(x)$ and $\delta(y)$ in Equation (3.13). The source location is marked by the unfilled symbol. The white arrow marks the direction of the applied force source.

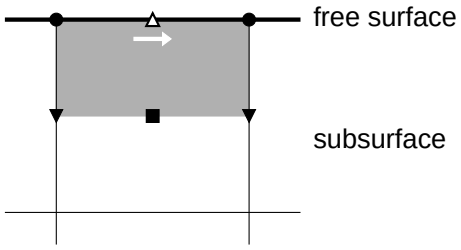


Figure 3.19: Sketch explaining appropriate source scaling for a horizontal force source at the free surface where the free surface is implemented by the image technique. The grey area displays the area that approximates the δ-functions $\delta(x)$ and $\delta(y)$ in Equation (3.13). The source location is marked by the unfilled symbol. The white arrow marks the direction of the applied force source.

However, if the image technique is used the physical grid seems to end at the free surface. To realize the image technique grid lines above the free surface are required, in fact. However, they are filled with unphysical values during the imaging of the stress field. Therefore, the wave equation used to update the particle velocities directly at the free surface must be possibly referred to only half a grid cell (Figure 3.19). In this half grid cell the density is constant which corresponds to the fact that the full density of the half-space is used. However, if only half a grid cell is considered in y-direction $\delta(y)$ should be approximated by $1/(0.5\Delta x)$. This increases the contribution to the particle velocity wavefield at the source point by the missing factor of two.

3.8 Influence of absorbing frame

In a first version the 2D FWI code DENISE uses the stress-displacement formulation for the finite-difference forward modeling. As absorbing frames a method proposed by Cerjan et al. (1985) was implemented (Köhn, 2011, chapter 2.2.3). Although, using large absorbing frames of 100 grid points (40 grid points are suggested by Cerjan et al., 1985) the damping of especially the surface waves is not satisfying using such an implementation. The absorbing frame in the stress-displacement formulation works less effective than in the stress-velocity formulation (see later comparison). Therefore, Köhn (2011) implemented a second version of the 2D FWI code DENISE which uses convolutionary perfectly matched layers, so called CPMLs (Komatitsch & Martin, 2007). They are much more efficient than the absorbing frames suggested by Cerjan et al. (1985) and can be therefore much smaller (10 to 20 grid points). This leads to a significant decrease of the model size. However, the CPMLs can be implemented more stable in a stress-velocity formulation. This caused that not only the implementation of the absorbing frame but also the FD scheme is different between the first and the second version of DENISE. In the second version the stress-velocity formulation is used.

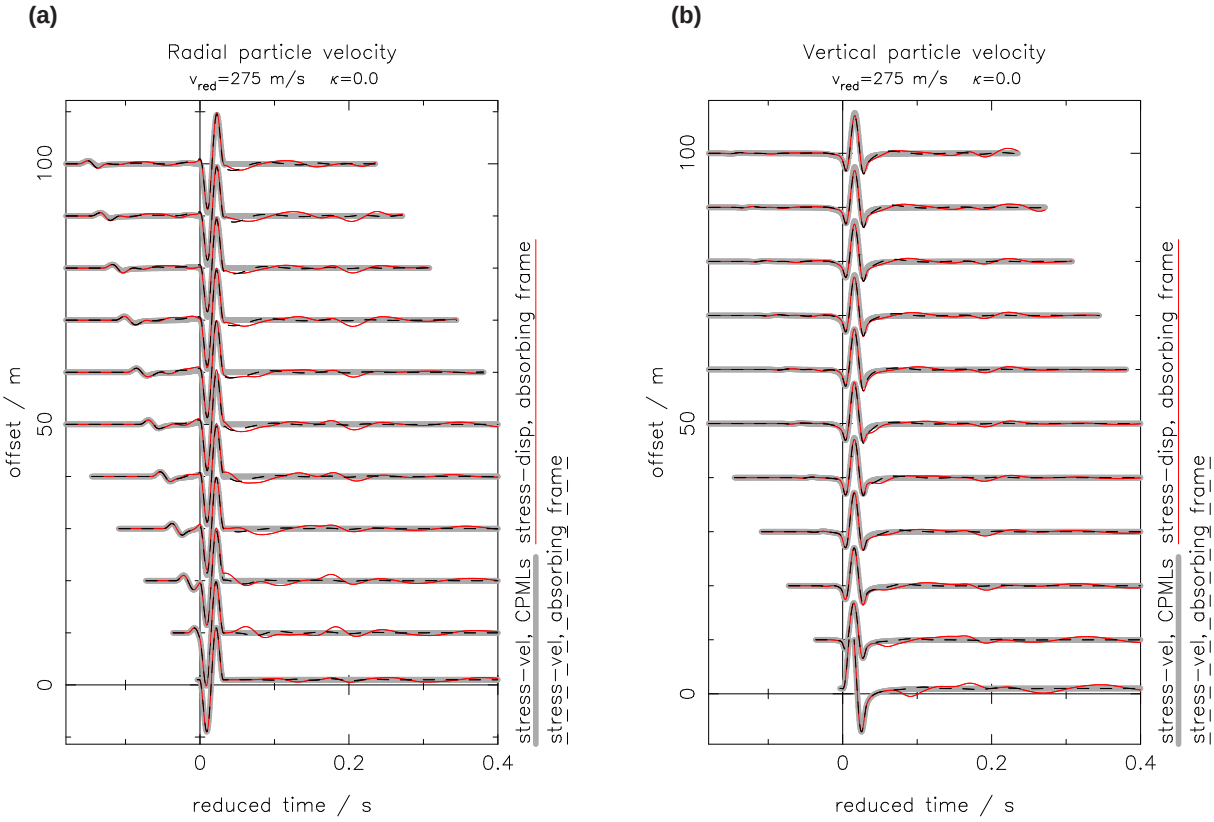


Figure 3.20: Comparison of different absorbing model frames. The thick grey seismograms are simulated with a stress-velocity formulation together with CPMLs (Komatitsch & Martin, 2007). The red seismograms are simulated with a displacement-stress formulation together with absorbing frames according to Cerjan et al. (1985) and the dashed black seismograms are simulated with a stress-velocity formulation with again absorbing frames according to Cerjan et al. (1985). The time axes in both plots are reduced with a velocity of 275 m/s. The seismograms are not normalized. a) displays horizontal particle velocity seismograms and b) vertical velocity seismograms.

Figure 3.20 displays a comparison of vertical velocity seismograms calculated for a vertical line force at the surface of a homogeneous half-space (see section 3.5 for a more detailed description of the model). The CPMLs together with the stress-velocity formulation provide the best results. Furthermore, the stress-velocity formulation with the absorbing frame suggested by Cerjan et al. (1985) suppresses the waves reflected from the boundaries more effective than in the stress-displacement formulation.

The results presented in this thesis (all other benchmark tests and FWI results) are performed with the second version of the DENISE code including CPMLs and using the stress-velocity formulation in the FD scheme.

3.9 Implementation of viscoelastic forward modeling

As will be shown in Section 4.2.1 the effects of anelastic damping are significant in shallow seismic field data. Therefore, I implement viscoelastic forward modeling in the DENISE code. It

is implemented analogously to the FD code described by Bohlen (1998). The rheology of the subsurface is described by a generalized standard linear solid (Liu et al., 1976; Carcione et al., 1988; Robertsson et al., 1994; see also Section 2.1.1 and Appendix A). A benchmark test is performed for a typical shallow seismic subsurface model. The P-wave velocity, S-wave velocity, and density structure is derived from a field data set and described in detail in Section 4.1. Furthermore, a quality factor of Q=20 for both P and S-waves is assumed. Again a vertical force source is used at the free surface and the wavelet is described by Equation (3.20) with a duration of T_d=32 ms. For the purpose of comparison a code using the reflectivity method (Fuchs & Müller, 1971; Müller, 1985) is used (implementation of Ungerer, 1990). This code uses a constant Q model where velocity dispersion due to damping can be considered via (Müller, 1985, Eq. 132 and 134)

$$c(\omega) = \Re\left[v_c(\omega)\right] \qquad \text{with} \qquad v_c(\omega) = v\left(1 + \frac{1}{\pi Q}\ln\frac{\omega}{\omega_r} + \frac{i}{2Q}\right) \tag{3.32}$$

where Q is the quality factor, v is the velocity of the medium at the reference angular frequency ω_r, v_c is a complex velocity, $c(\omega)$ is the phase velocity due to anelastic damping, and $\Re$ denotes the real part of a complex quantity.

In DENISE three simulations are performed with three, four and five relaxation mechanisms. The used relaxation frequencies and values for τ are listed in Table 3.5. They are determined according to Bohlen (1998) by solving a non-linear inverse problem. The corresponding frequency dependency of the quality factors are displayed in Figure 3.21a). The three GSLSs used in the FD simulations approximate the constant quality factor of 20 well in the frequency band between approximately 5 Hz and 70 Hz which is considered in the benchmark test. Figure 3.21b) displays a comparison of the dispersion curves caused by anelastic damping. They are very similar. Therefore, it is possible to compare the seismograms simulated with the FD scheme and the reflectivity method. Figure 3.22 displays a comparison of the seismograms. They fit each other well in phases and amplitudes. The FD results do not significantly change in dependence of the number of relaxation mechanisms. Therefore, I will use three relaxation mechanisms in the FWI tests described in the later chapters.

L	τ	$f_{\sigma l} = 1/(2\pi\tau_{\sigma_l})$ in Hz				
3	0.0966	0.52	7.67	72.68		
4	0.0813	0.24	2.41	22.67	134.64	
5	0.0720	0.15	1.05	7.78	48.61	236.15

Table 3.5: Relaxation frequencies $f_{\sigma l}$ and values of τ used for the benchmark test. The generalized standard linear solids approximate a quality factor of Q=20 in a frequency range between approximately 5 Hz and 70 Hz.

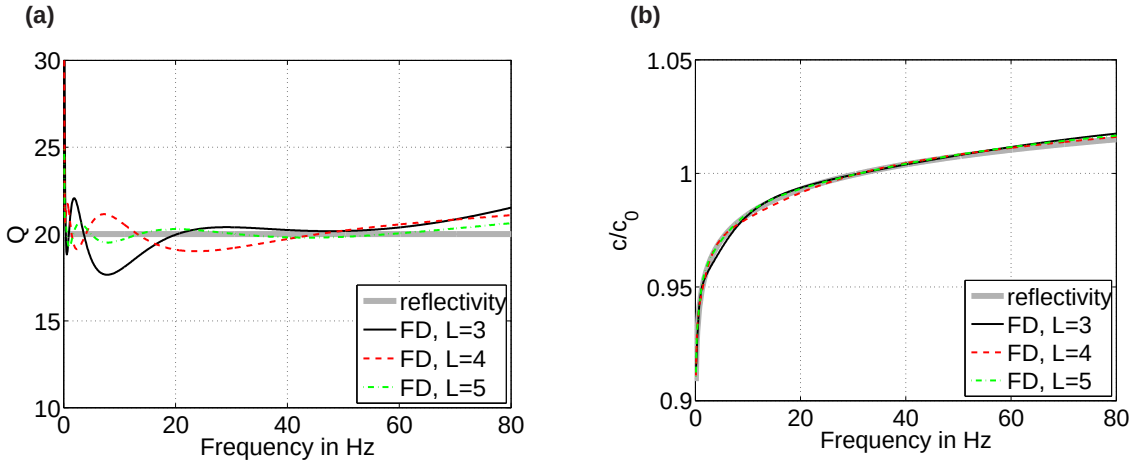


Figure 3.21: Comparison of viscoelastic parameters used for the anelastic benchmark test. a) displays the frequency dependence of the quality factors used in the FD simulations. b) displays the corresponding dispersion curves together with the dispersion curve of the model used in the simulations with the reflectivity model (calculated with Eq.(3.32).

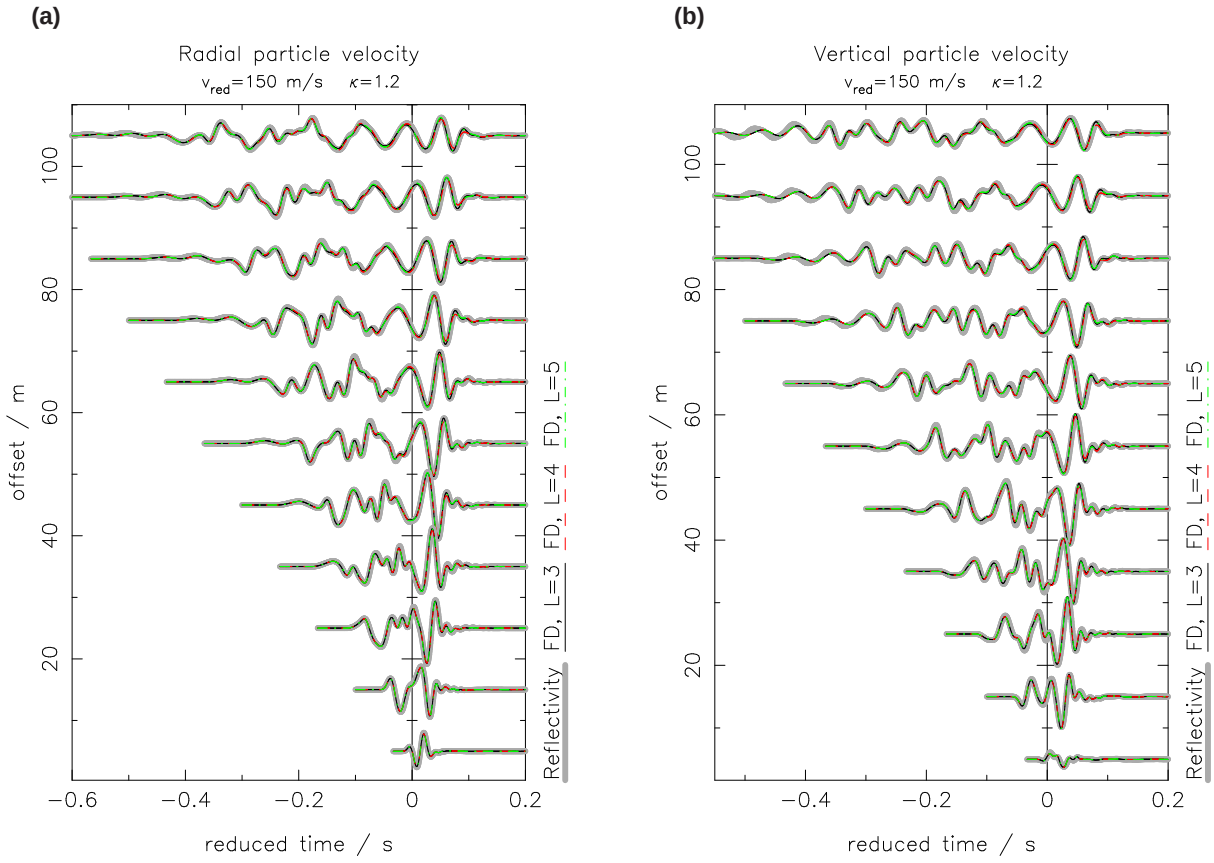


Figure 3.22: Comparison of seismograms simulated with the reflectivity method and the FD method for an anelastic benchmark model. In the FD simulations different numbers L of relaxation mechanisms are used. The time axes in both plots are reduced with a velocity of 150 m/s. The seismograms are not normalized but they are multiplied by an offset dependent factor of $(r/\mathrm{m})^{\kappa}$ where κ is given in the head of each figure. a) displays the horizontal-component and b) the vertical-component seismograms.

Chapter 4

Reconstruction tests

At first, I apply the 2D FWI to synthetic data which are used as pseudo-observed data in the inversion. These tests are called reconstruction tests in the following because the true subsurface model is known and the reconstruction of this model by FWI can be directly evaluated by a comparison of the true subsurface model and the model obtained from FWI. I investigate two topics by reconstruction tests: the influence of anelastic damping (Section 4.2) and the influence of the P-wave velocity model (Section 4.3). The true model as well as the basic inversion setup used for the reconstruction tests are identical and are described in section 4.1.

4.1 Subsurface model and inversion setup

4.1.1 True subsurface model and acquisition geometry

I use the one-dimensional model shown in Figure 4.1 for the reconstruction tests. It was derived by my colleague M. Schäfer from a field data set which was acquired on a predominantly depth dependent structure at Rheinstetten near Karlsruhe (Germany). To obtain this subsurface model he applied a joint inversion of Fourier-Bessel expansion coefficients and first arrival P-wave travel times (Forbriger, 2003a,b). The P-wave velocity model v_p (Figure 4.1a and d) consists of a layer over a half-space with a strong discontinuity at 6.3 m depth and is mainly constrained by first arrival P-wave travel times. The S-wave velocity model v_s (Figure 4.1b and e) is mainly constrained by the Fourier-Bessel expansion coefficients. It consists of a steep gradient in the topmost meter, below the v_s gradient becomes weaker. The strong gradient in the first meter of the S-wave velocity model is typical for unconsolidated sediments (Bachrach et al., 2000) and is often observed in shallow seismics. The strong discontinuity in the v_p/v_s ratio in approximately 6.3 m depth is interpreted as groundwater table because a stronger contrast in the S-wave velocity would increase the misfit which is minimized in the inversion significantly. The density model (Figure 4.1c) is not well constrained by the field data and consists of three layers with density values between 1700 kg/m^3 and 2000 kg/m^3.

As pseudo-observed data I generate viscoelastic seismograms with a quality factor of approximately 20 for both, P-waves and S-waves. I use the 2D finite-difference scheme which is described in Chapter 3 for the simulation. The same algorithm for wave propagation is used in the FWI program. The rheological model is implemented by a generalized standard linear solid

(GSLS) with three relaxation mechanisms (Robertsson et al., 1994; Bohlen, 1998). The approximation of a constant quality factor is satisfactory in the frequency band between 5 Hz and 70 Hz.

The acquisition geometry consists of eight vertical force sources with a spacing of 10 m (red stars in Figure 4.1a to c) and 63 two-component receivers (vertical and radial component) with a spacing of 1 m. The first receiver is located at x=2 m and the last one at x=70 m. I use a spatial grid spacing of Δx=20 cm. The free surface is implemented by the image technique (Section 3.2.1). Due to the standard staggered grid (Figure 3.2) the sources are located half a grid point below the free surface that is in a depth of 10 cm. The receivers are located one grid point below the free surface because of the scaling problem which occurs for sources excited directly at the free surface when the image technique is used (see Section 3.7). Thus, the radial component receivers are in a depth of 20 cm and the vertical component receivers in a depth of 30 cm (corresponds to 1.5 Δx). The smallest wavelength present in the pseudo-observed data is approximately 2.5 m. Therefore, sources and receivers can be considered as located at the free surface.

4.1.2 Inversion setup

In all reconstruction tests I use the L2 norm of the normalized wavefields as misfit function (for definition see Equation (2.63)). Furthermore, I use a multiscale inversion approach (Bunks et al., 1995) by applying frequency filtering during the inversion (Section 2.3.4). I start with a fourth-order Butterworth lowpass filter with a corner frequency at 10 Hz and increase the corner frequency up to 70 Hz in steps of 5 Hz. The bandwidth of the data is increased as soon as the relative change of the data misfit of the current iteration with respect to the next to last iteration drops below 1%. However, at least 10 iterations are performed before the criterion for the relative misfit change is activated. I use the method described in Section 2.3.1 for gradient preconditioning. The gradient of each shot is multiplied by a semi-circular taper around the respective source position with a radius of 3.0 m and an increasing amplitude from the center to the boundary of the circle. After each iteration step the models are smoothed with a 2D median filter with a filter length of 0.6 m, which is small compared to the smallest wavelength of approximately 2.5 m.

All reconstruction tests are performed on a personal computer using 4 CPUs. The number of iteration steps varies between 170 and 210 iterations and the computing time is around 20 hours on 4 CPUs for each test. Further information about the parameters used for forward modeling are listed in Table H.1 in Appendix H.

4.2 Influence of anelastic damping

The effects of viscoelastic damping can be significant in shallow seismic field data (Hatherly, 1986). Normally, quality factors Q between 10 and 50 are observed. In preparation of an inversion of field data I therefore investigate whether and to which degree anelastic damping must be considered in an FWI of shallow seismic Rayleigh waves.

There are two strategies how to consider anelastic damping in FWI. On the one hand viscoelastic forward modeling with a priori assumed attenuation can be applied in the inversion. The anelastic

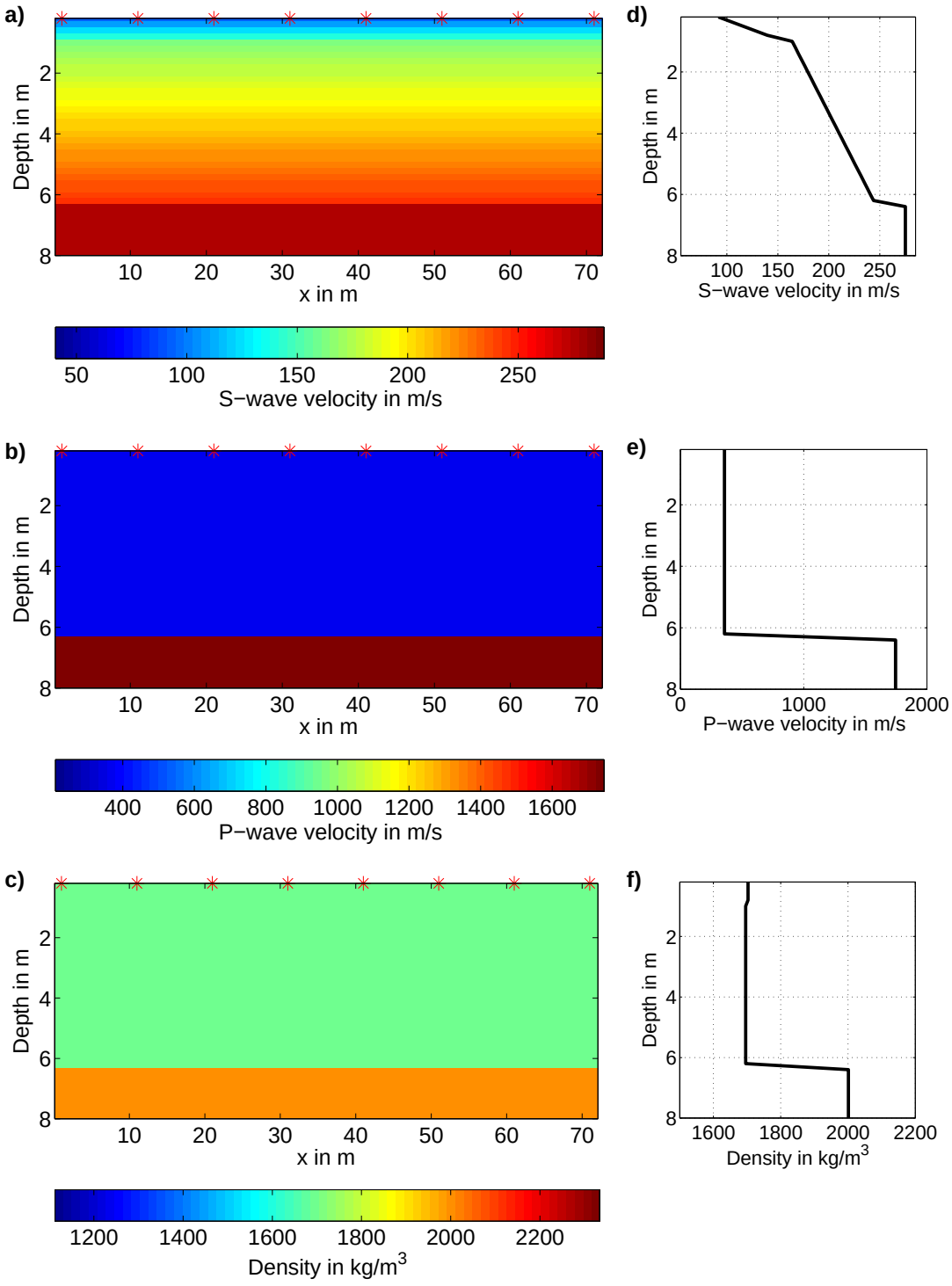


Figure 4.1: True model used in the reconstruction tests. a) displays the S-wave velocity model, b) the P-wave velocity model and c) the density model. The red stars mark the positions of eight sources which are used in the reconstruction tests. Not the whole models used for the simulations are shown. CPML boundaries as well as the lower part of the models are cut off (maximum depth of the models without CPML boundary is 11 m). The range of the colorbar for each parameter is not changed within this chapter unless it is stated in the caption of the figure. d) to f) display vertical profiles through the models shown in a) to c).

properties are used as fixed parameters in this case and one inverts only for elastic parameters (P-wave velocity, S-wave velocity, and density). This is e. g. discussed by Bretaudeau et al. (2013). They apply 2D elastic FWI on a layered medium with an inclusion for synthetic data and laboratory data which are obtained by small-scale physical modeling. They use viscoelastic forward modeling in their FWI. The viscoelastic properties of the samples in their small-scale physical model are determined by independent ultrasonic measurements. They assume them to be known well enough for a successful FWI and do not investigate the influence of inaccurately chosen viscoelastic parameters on the reconstructed model in more detail.

Beside using viscoelastic forward modeling with fixed damping in an FWI, one can also invert for dissipative properties within the FWI. In frequency-domain FWI this can be implemented with almost no extra cost by using complex moduli or complex seismic velocities, respectively (Hicks & Pratt, 2001; Malinowski et al., 2011). In time-domain FWI inversion for dissipative properties is computationally more expensive because the gradient of the misfit function does not only contain zero-lag cross-correlations of the forward and the adjoint wavefields which determine the gradient in the elastic case. Additionally, it contains full convolutions of the relaxation functions and the forward propagated wavefield (Charara et al., 2000). However, as discussed in Section 2.1.4 an inversion of dissipative properties is likely to have a high ambiguity, especially in an inversion of field data. Therefore, I consider in this thesis viscoelastic forward modeling but do not invert for dissipative properties. The dissipative properties are used as a priori known fixed parameters in the inversion (Section 2.1.4).

In the following I will show a comparison of elastically and viscoelastically simulated data with field data to demonstrate the influence of anelastic damping on field data. In addition, I will discuss the influence of a source wavelet correction (Section 4.2.1). Afterwards, I will describe the results of reconstruction tests (Section 4.2.2) which investigate the influence of dissipative properties on FWI results by comparing tests for different magnitudes of a priori assumed attenuation.

4.2.1 Comparison of simulated data with field data

Field data set and 3D/2D transformation

I use the field data set that was introduced in Section 4.1.1 for the comparison. The recorded wavefield was excited by a hammer blow which is approximately a vertical point force. To compare the field data with 2D synthetic data I transform the recorded point-source seismograms to the corresponding line-source seismograms using the direct wave transformation (Forbriger et al., 2013, see also Section 2.4.3).

Estimation of quality factor

The derived elastic subsurface model (Figure 4.1) is used to estimate an appropriate Q value for the field data set. For that reason, several forward modeling runs on this model are carried out with different quality factors. For the sake of simplicity and to ensure physical consistency of the subsurface model I assume Q_s=Q_p where Q_s is the quality factor for S-waves and Q_p

for P-waves. Tested quality factors are Q=10, 15, 20, 25, 30, 40, and 50. Furthermore, elastic forward modeling is used. After applying a source wavelet correction (see later subsection and Appendix F) the synthetic seismograms are compared with the field data. This comparison is done on the one hand qualitatively and on the other hand quantitatively by a comparison of the L2 data misfit between the recorded and the synthetic seismograms. The analyzed L2 data misfit equals the data misfit that is minimized in the estimation of the source wavelet correction filter. I apply an offset dependent weighting of $(r/\mathrm{m})^{0.6}$ in the optimization of the source wavelet correction filter so that the contribution of near and far-offset traces is approximately equal. Note that a 3D/2D transformation is already applied to the field data before. Therefore, the amplitude decay with offset is not as strong as in the original field data. I observe a minimum L2 misfit for Q_s=Q_p=25 (Figure 4.2). It is hardly possible to estimate an uncertainty for this value. The misfit increases by less than 1.5% in a range between Q_s=Q_p=20 and Q_s=Q_p=40 with respect to the misfit for Q_s=Q_p=25. However, this interval is influenced by the discrete quality factors used in the grid search as well as by the 3D/2D transformation which is applied to the field data prior to the comparison. The applied 3D/2D transformation is not exact and it can be applied either as single velocity transformation or as direct wave transformation (Forbriger et al., 2013, see also Section 2.4.3) which influences especially the amplitudes of the near-offset traces. Therefore, the given range of quality factors must be treated with caution. Nevertheless, I defined this interval to give an idea of how accurate the quality factor can be estimated from the data.

Figure 4.3a displays the frequency dependence of the quality factor which was used to approximate a constant Q value of 25 with three relaxation mechanisms of a GSLS. Figures 4.3b displays the corresponding phase velocity dispersion caused by damping. The range of the estimated uncertainties are marked by the grey area.

Influence of viscoelastic damping on the waveforms

Figure 4.4 displays the comparison of the field data (thick grey line) with elastically (red) and viscoelastically (black) modeled data with Q_s=Q_p=25. No source wavelet correction is applied to the synthetic data. I used the wavelet described by Equation (3.20) with a duration of T_d=0.0125 s for the simulation which approximates the source time function of the hammer blow (Forbriger, 2004). All wavefields are filtered with a sixth-order Butterworth lowpass filter with a corner frequency of 70 Hz. The two synthetic wavefields differ clearly. The viscoelasticity causes a distance dependent and frequency dependent damping which is not present in the elastic wavefield. The strong ringing at large offsets which is present in the elastic wavefield is caused by the steep gradient in the topmost meter of the S-wave velocity model. This ringing is strongly damped in the viscoelastic wavefield. The additional phase velocity dispersion which is caused by damping is too small (Figure 4.3b) to be observed in the viscoelastic wavefield. Note that it is not possible to observe the different amplitude decay with offset of the wavefields in Figure 4.4 because each seismogram is normalized to its maximum amplitude.

Source wavelet correction filter

The differences between the recorded data and the synthetic data (Figure 4.4) are mainly caused by differences between the true subsurface structure and the one-dimensional subsurface model that is used in the simulations. Furthermore, the discrepancies are caused by differences in the

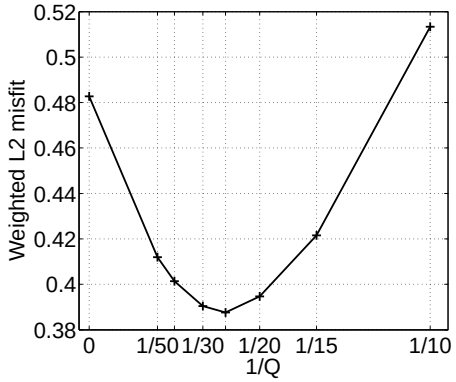


Figure 4.2: Estimation of the quality factor for the field data set. Displayed is the L2 data misfit between the field data and synthetic data generated for different quality factors Q. Prior to the calculation of the L2 data misfit a source wavelet correction is applied to the synthetic seismograms. The result for elastic forward modeling is plotted at $1/Q = 0$ since $Q \to \infty$ in the elastic case.

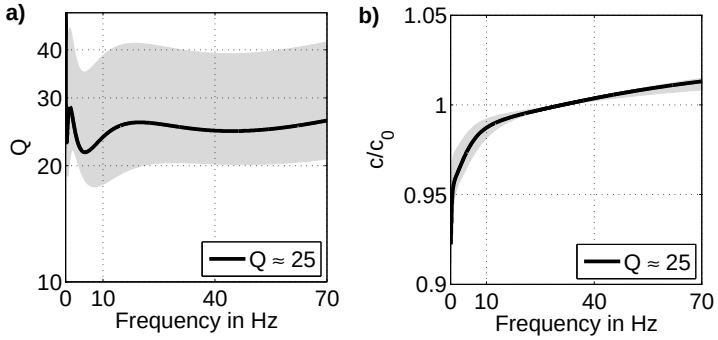


Figure 4.3: Estimated Q model assuming a generalized standard linear solid with three relaxation mechanisms. a) displays the frequency dependence of the quality factor Q and b) the corresponding velocity dispersion caused by damping for a reference frequency of 30 Hz. c is the phase velocity and c_0 is the phase velocity at the reference frequency. The grey areas mark the uncertainties in the determination of the quality factor (interval between $Q \approx 20$ and $Q \approx 40$).

source time function and the characteristics of the recording system (instrument response of geophones [eigen frequency 4.5 Hz] and digitizer). Therefore, I determine a source wavelet correction filter for the elastically as well as the viscoelastically simulated wavefield. An appropriate filter can be found by solving a damped, linear least-squares optimization problem (Appendix F). It can be formulated as stabilized deconvolution of the recorded seismograms with the simulated seismograms similar to the method proposed by Pratt (1999). In the inversion of the source wavelet correction filters I apply a weighting in the misfit which is minimized in order to infer the optimal source correction. The traces are weighted by $(r/\mathrm{m})^{0.6}$, where r is the offset. This weighting function ensures that middle and far-offset traces contribute also significantly to the L2 data misfit. I use the L-curve criterion described by Aster et al. (2013) to adjust the damping of the least-squares problems appropriately. Details on this can be found in Appendix G.

Comparison of field data and synthetic data after applying a source wavelet correction

Figure 4.5 displays the field data and the synthetic seismograms after application of the source wavelet correction. Both synthetic wavefields match the field data well but they differ in some details. The amplitude decay with offset of the fundamental mode of the Rayleigh wave is better fitted by the viscoelastic data (see time interval 0 s-0.15 s in Figure 4.5). The optimized source wavelet correction filter acts as a lowpass filter for the elastic wavefield. This can be observed in the comparison of amplitude spectra of the source wavelet correction filters determined for the elastic and the viscoelastic wavefield, respectively (Figure 4.6). The source wavelet correction filter in the elastic case eliminates the high frequencies, which are already damped in the recorded

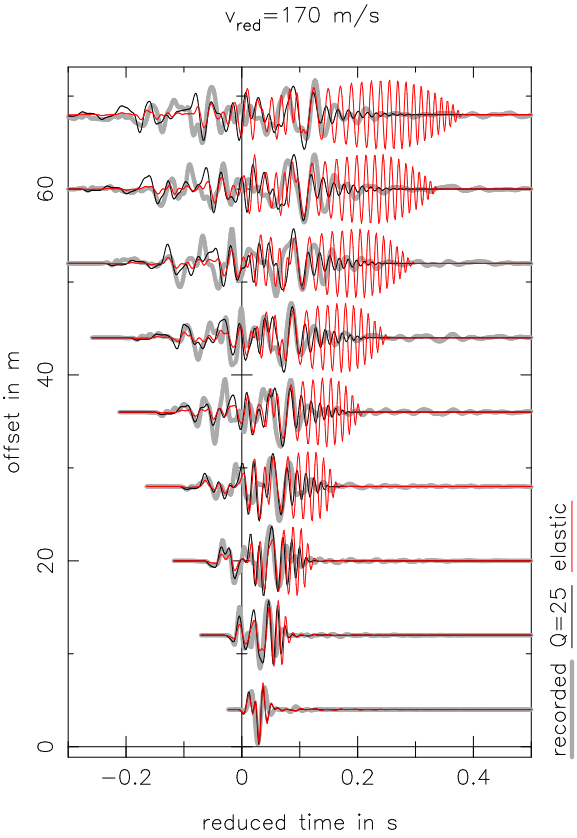


Figure 4.4: Comparison of vertical particle velocity seismograms of field data (thick grey line) and synthetic data without application of a source wavelet correction (model used for simulation see Figure 4.1). The viscoelastically modeled data are displayed in black and the elastically modeled data in red. Each trace is scaled to its maximum amplitude. All wavefields are lowpass filtered (Butterworth filter with corner frequency of 70 Hz and order 6). The field data are transformed to equivalent line-source seismograms. The time axis is reduced with a velocity of 170 m/s.

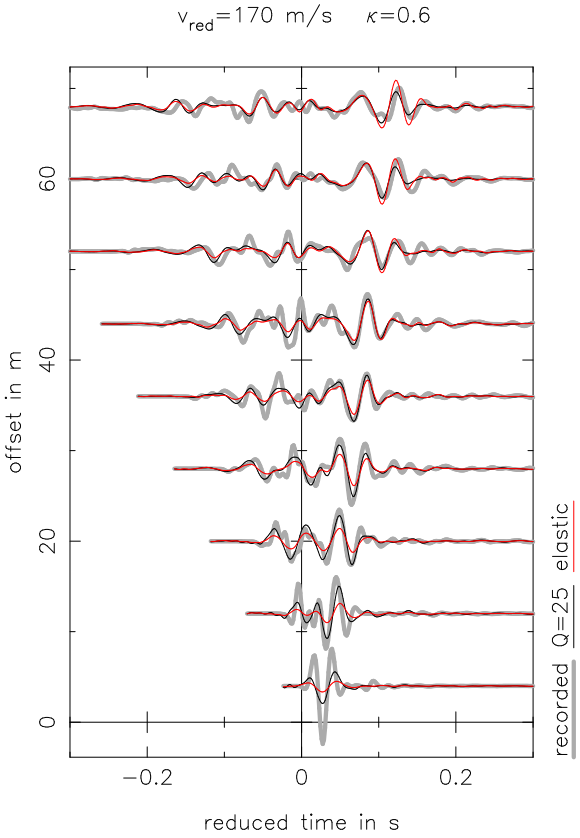


Figure 4.5: Comparison of vertical particle velocity seismograms of field data (thick grey line) and synthetic data after application of a source wavelet correction to the simulated seismograms (model used for simulation see Figure 4.1). The viscoelastically modeled data are displayed in black and the elastically modeled data in red. In case of the elastic data the source wavelet correction additionally accounts for a lowpass effect due to viscoelastic damping. The traces are multiplied by an offset dependent factor $(\frac{r}{\mathrm{m}})^{0.6}$ where r denotes offset in m. Therefore, amplitudes within traces are comparable. The field data are lowpass filtered (Butterworth filter with corner frequency of 70 Hz and order 6) and transformed to equivalent line-source seismograms. The time axis is reduced with a velocity of 170 m/s.

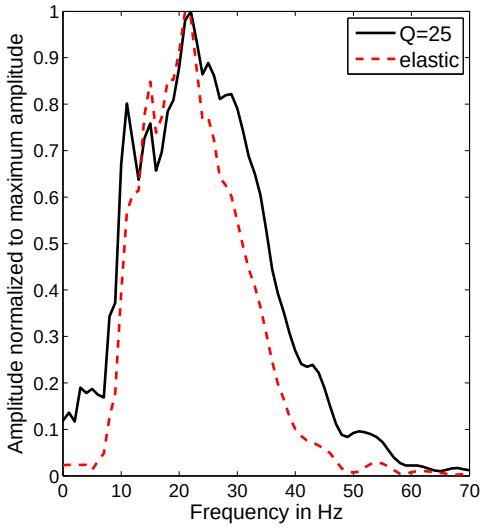


Figure 4.6: Comparison of amplitude spectra of source wavelet correction filters for viscoelastic data (black) and elastic data (dashed red) used for seismograms in Figure 4.5.

data at middle and far offsets. However, the source time function can only act as a frequency dependent filter and not as a distance dependent filter. Therefore, the high frequencies are also damped in the near-offset traces (see first trace in Figure 4.5) where the bandwidth of the elastic seismograms is too narrow compared to the bandwidth of the field data and the viscoelastically modeled data.

Concluding, relevant differences in the frequency content of elastic and viscoelastic shallow seismic wavefields can be observed. A significant fraction of these residuals can be compensated by a source wavelet correction, which is narrowing the signals' bandwidth. Nevertheless, this effect is not sufficient. Viscoelastic modeling is essential to obtain the correct amplitude decay with offset.

4.2.2 Reconstruction tests

In the second part of the study inversion tests are performed where simulated viscoelastic observations generated with Q_s=Q_p=20 are used as pseudo-observed data. First, I use purely elastic forward modeling in the FWI and viscoelastic forward modeling with a priori known Q values without applying a source wavelet correction. These tests are possible because the true source wavelet is known and can be used in the inversion. However, in an inversion of field data a source wavelet correction must be applied. Furthermore, I have shown in the previous section that there is a trade-off between viscoelastic effects and the source wavelet correction filter. Thus, I additionally test to which degree the application of a source wavelet correction can improve the FWI results.

Due to the high sensitivity of Rayleigh waves to the S-wave velocity I focus on the reconstruction of the S-wave velocity model. I apply FWI in two scenarios. In the first scenario I discuss the monoparameter inversion for S-wave velocity and I use the correct P-wave velocity model and

density model as fixed parameters in the inversion. In the second, more realistic, scenario I apply multiparameter inversion for S-wave velocity, P-wave velocity, and density. Details on the true subsurface model, the acquisition geometry, and the inversion setup used in the reconstruction tests are described in Section 4.1.

Reconstruction tests with a priori known source wavelet

Monoparameter tests

As initial S-wave velocity model a linear gradient model is used. It starts with the correct velocity of 92 m/s at the surface and increases linearly up to the half-space velocity of 275 m/s in a depth of 9 m.

To obtain a reference result I invert the pseudo-observed data with an FWI using viscoelastic modeling with the correct Q value. Figure 4.7a) displays the obtained S-wave velocity model. It is reconstructed almost perfectly. This is due to the fact that I keep the inversion problem simple by inverting only for S-wave velocity and using the correct P-wave velocity, density model, and Q value. We obtain a nearly one-dimensional subsurface structure and the vertical velocity profiles of the true model and the reconstructed model are in very good agreement. Small artefacts arise only in the vicinity of the sources. With this configuration it is possible to reconstruct all features of the model, the two shallow gradients as well as the small contrast at 6.3 m depth. Figure 4.8a) displays the final data misfit calculated with Equation 2.63. For the reference inversion result the final data fit is almost perfect as well. This is confirmed by Figure 4.9a) where two vertical displacement seismograms of the shot located at x=11 m are displayed, one in the near-offset and one in the far-offset region.

In contrast, the reconstruction of the S-wave velocity model is not possible if purely elastic modeling is used in the inversion (Figure 4.7b). Strong artefacts appear in the entire model. The final data misfit is worse compared to the reference test (Figure 4.8). Nevertheless, the artefacts in the S-wave velocity model can at least partly compensate the differences between viscoelastic observed data and elastically simulated data. This can be observed in the comparison of the far-offset trace in Figure 4.9b). Although the relative amplitudes within the trace do not match between the observed data and the simulated data all phases are explained by the reconstructed S-wave velocity model.

This result shows that viscoelastic damping cannot be neglected in an FWI of field data. Unfortunately, the correct rheology of the subsurface is not known when field data are inverted. To improve the reconstruction of the S-wave velocity model, the Q values can be estimated from the data prior to the inversion and can be used as a priori known fixed parameters in the FWI. To test for possible artefacts caused by a wrong quality factor used in the FWI I present in the following the inversion results for Q_s=Q_p=10 and Q_s=Q_p=40. I assume that the interval between these quality factors is large enough to cover the accuracy of the estimation of the quality factor from a field data set. However, this interval is derived from my experiences with the field data set described in the previous section where I use a very good subsurface model to estimate the quality factors. It must be checked for other data sets if such an accuracy of the estimated quality factors is possible in general.

I observe a periodic pattern that coincides with the source positions in the reconstructed S-wave velocity model with a quality factor of Q_s=Q_p=40 (Figure 4.7c). However, in comparison to the result obtained with elastic forward modeling the artefacts in the reconstructed S-wave velocity model as well as the final data misfit (Figure 4.8 and Figure 4.9c) are smaller. A periodic pattern is also observed in the inversion result obtained with quality factors of Q_s=Q_p=10 (Figure 4.7d). The final data misfit is larger in comparison to the result for Q_s=Q_p=40 which is mainly caused by a poor fit of the relative amplitudes for the middle and far-offset traces (Figure 4.9d). Nevertheless, the results obtained with wrong Q values show less model artefacts and lower data misfits than the result obtained by a purely elastic inversion. Therefore, FWI with viscoelastic forward modeling should be preferred for shallow seismic field data sets.

Multiparameter tests

As the correct P-wave velocity model and density model are not known in field data application we typically cannot use them as fixed models in the inversion. Therefore, I repeat the previous tests and apply multiparameter inversions where I invert for S-wave velocity, P-wave velocity and density. I use the same initial S-wave velocity model like in the monoparameter reconstruction tests. As initial P-wave velocity model and density model I use also linear gradient models. They start with the true layer values at the surface and increase linearly up to the half-space values in a depth of 9 m where they pass into a homogeneous half-space.

In all multiparameter inversion tests the reconstruction of the P-wave velocity model and the density model fails. The initial P-wave velocity model is changed only in the first three meters and the discontinuity in 6.3 m depth could not be reconstructed. The reconstructed density models are very heterogeneous and in extreme cases even unrealistic density values of below 1000 kg/m^3 and above 3900 kg/m^3 are obtained. The failure of the reconstruction of the P-wave velocity model and the density model is presumably caused by the inappropriate initial models which are used in the tests. Additionally, these model parameters are not constrained as well as the S-wave velocity by the data. Therefore, I again focus on the reconstruction of the S-wave velocity model.

Figure 4.10 displays exemplary inversion results. The reconstruction of the S-wave velocity model is worse in comparison to the monoparameter reconstruction tests (compare Figure 4.7). The small discontinuity at 6.3 m depth could not be resolved even in the reference inversion where I use the correct quality factor of Q_s=Q_p=20. The worst reconstruction is again obtained by using elastic modeling in the FWI. However, the differences between the reconstructed models in the multiparameter tests are not as large as in the monoparameter tests which can be observed in the variation of the RMS value of the relative model errors for the reconstructed models (Figure 4.11). With elastic modeling it is especially not possible to reconstruct the shallow part of the model down to 4 m depth. The reconstruction becomes worse with decreasing depth because the differences between elastic and viscoelastic data are largest at high frequencies and the shallow structure is inferred from high-frequency surface waves due to their low penetration depth. For too low quality factors of Q_s=Q_p=10 the reconstruction becomes worse in the deeper part of the model (Figure 4.10d).

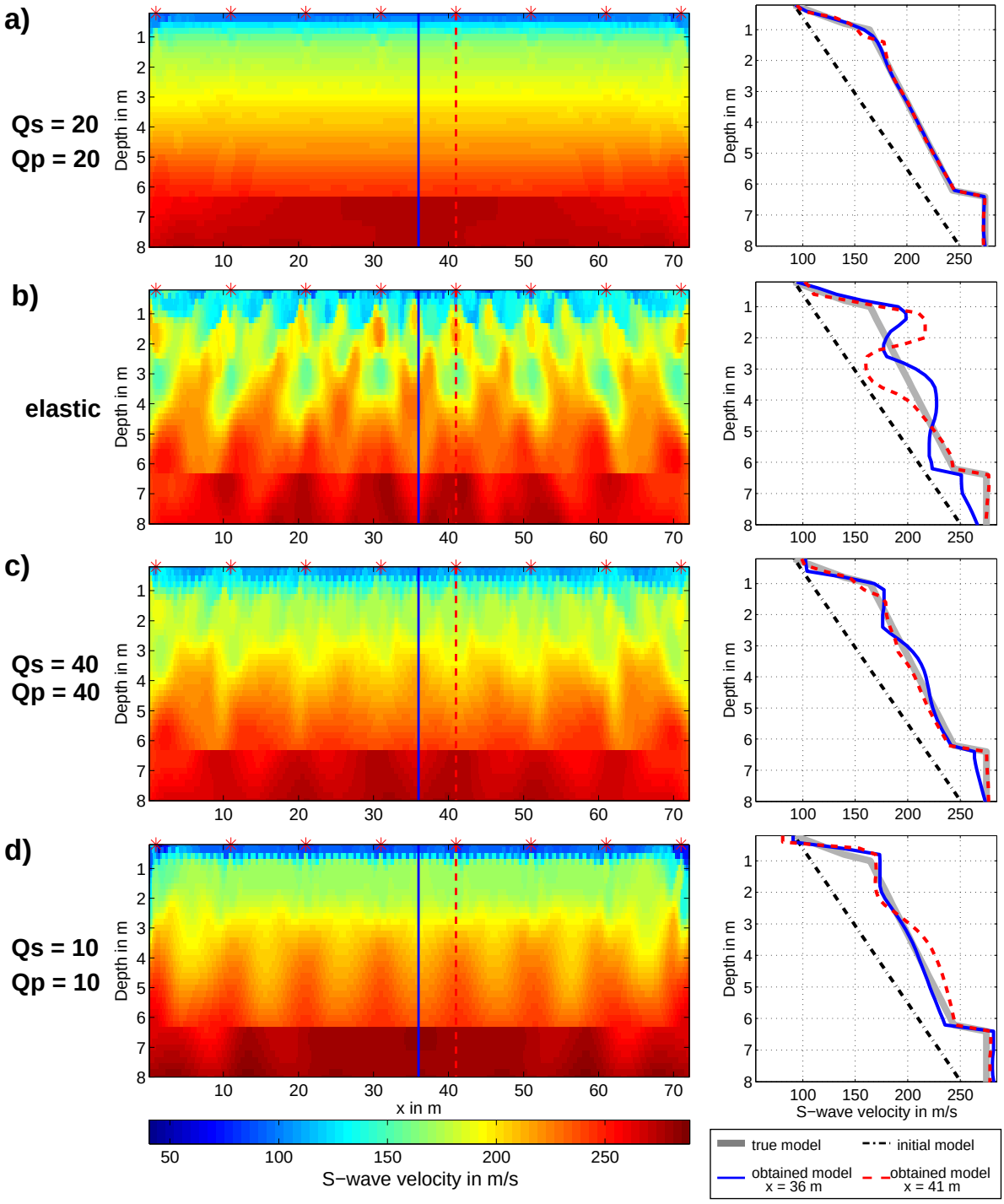


Figure 4.7: Results of monoparameter reconstruction tests without source wavelet correction. On the left, the obtained S-wave velocity models are displayed. The source positions are marked by the red stars. On the right, vertical velocity profiles at x=36 m (solid blue line) and at x=41 m (dashed red line) are shown together with the true model (thick grey line) and the initial model (dash-dotted black line). The locations of the profiles are marked by the blue and red lines in the models. CPML boundaries as well as the lower part of the models are cut off.

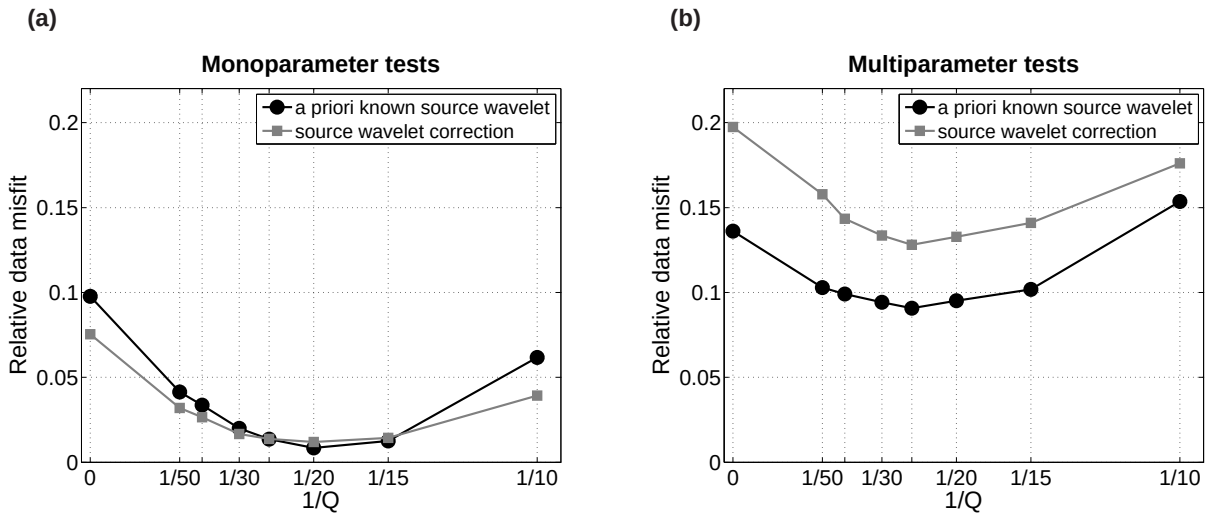


Figure 4.8: Final data misfit calculated with Equation (2.63) for the a) monoparameter reconstruction tests and b) multiparameter reconstruction tests. The black dots display the results for the tests with a priori known source wavelet and the grey squares display the results obtained for the tests with source wavelet correction. The results obtained with elastic forward modeling are displayed at $1/Q = 0$ since Q is infinite in the elastic case.

The final data fit in the multiparameter tests is worse than in the monoparameter inversion tests (Figure 4.8). Especially the P-waves and the higher modes of the Rayleigh waves can not be fitted accurately (far-offset trace in Figure 4.12) due to the failure in the reconstruction of the P-wave velocity model. The influence and handling of the P-wave velocity model in the FWI is discussed in more detail in Section 4.3. As in the monoparameter tests the best fit is obtained for approximately correct quality factors. For increasing and decreasing quality factors the data misfit increases.

Reconstruction tests with source wavelet correction

I demonstrated in the previous section that the S-wave velocity model can not be properly reconstructed when using purely elastic forward modeling in the FWI or viscoelastic forward modeling with significantly wrong Q values. However, I have shown that an optimized source wavelet correction filter can compensate a significant portion of the differences in wavefields caused by viscoelasticity. Furthermore, in an inversion of field data a source wavelet correction has to be applied. Therefore, I carry out further reconstruction tests to show how the application of a source wavelet correction effects the FWI results.

Source wavelet correction within FWI

I apply the source wavelet correction and the FWI sequentially in each frequency band of the multiscale inversion. After choosing a new frequency band an optimal source wavelet correction filter is derived for the waveforms of each shot using the current subsurface model. These correction filters are then used unaltered until a new frequency band is selected. Details on the determination of the source wavelet correction filter are given in Section 4.2.1 and in Appendix F.

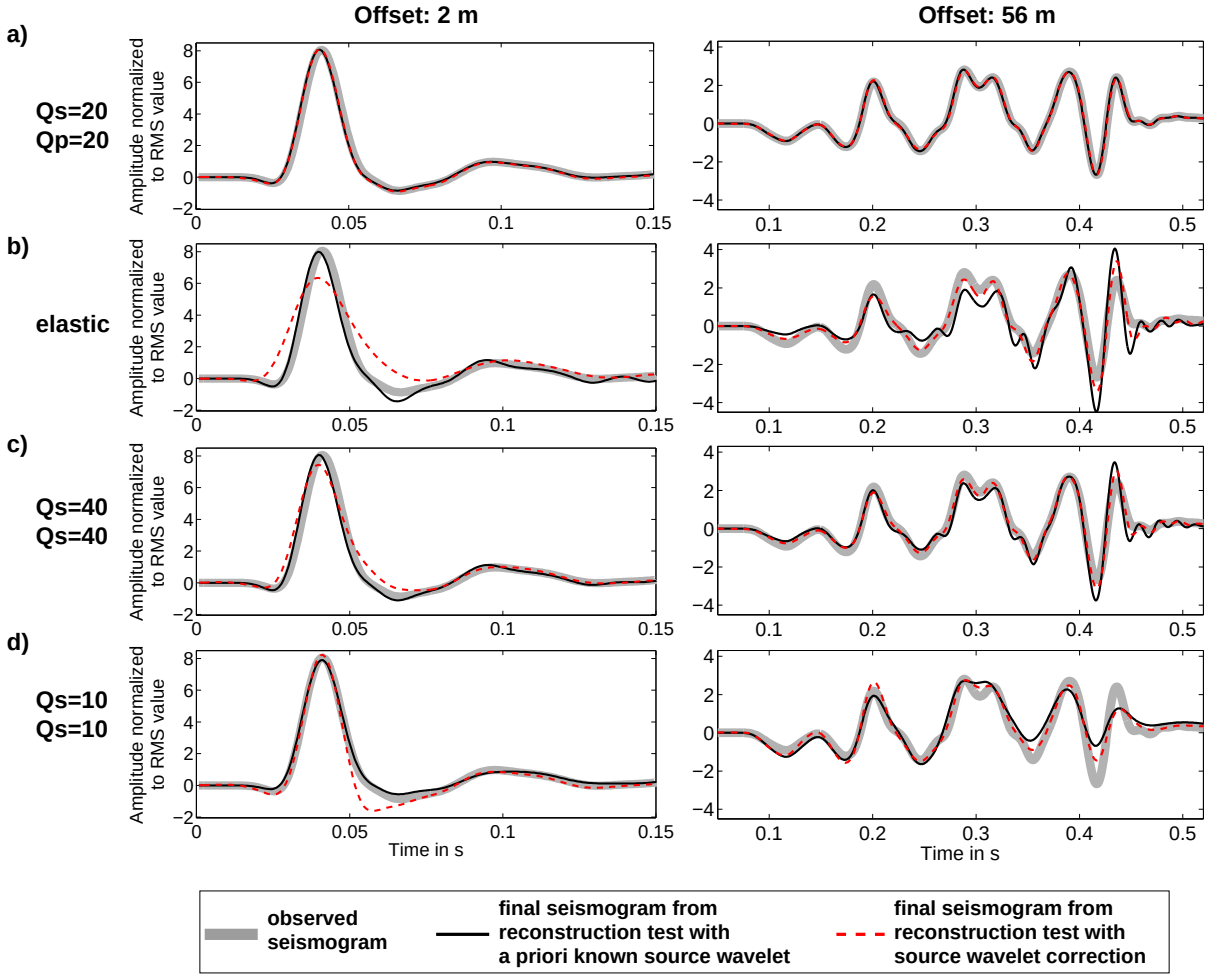


Figure 4.9: Comparison of vertical displacement seismograms for monoparameter reconstruction tests of the shot located at x=11 m for a near-offset trace at 2 m offset (left side) and a far-offset trace at 56 m offset (right side) for the reconstruction tests. All seismograms are normalized to their RMS amplitude. Pseudo-observed data are displayed by the thick grey line. Final seismograms of reconstruction tests with a priori known source wavelet are displayed in black and final seismograms of the reconstruction tests with application of a source wavelet correction are displayed by the dashed red line.

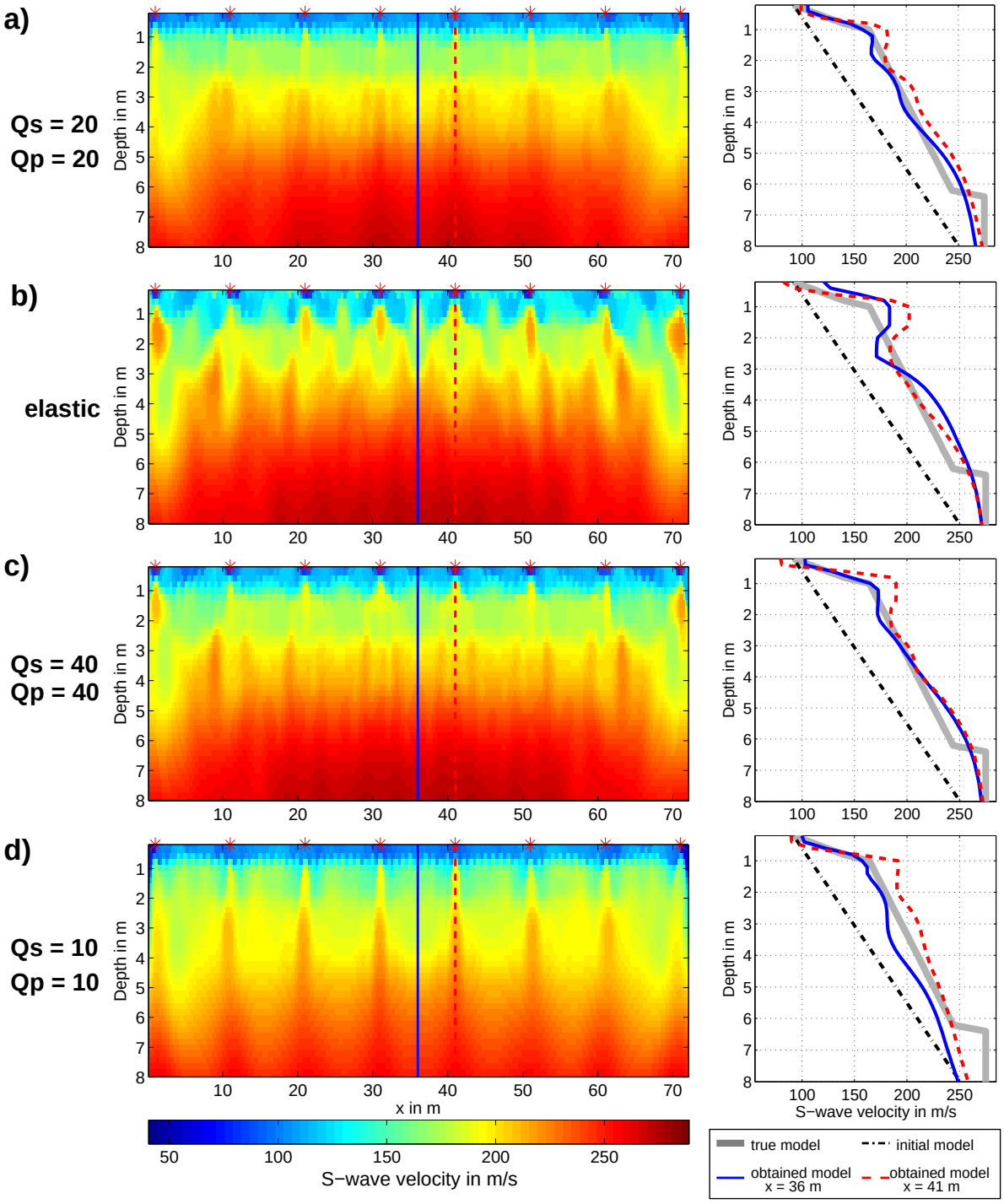


Figure 4.10: Results of multiparameter reconstruction tests without source wavelet correction. On the left, the obtained S-wave velocity models are displayed. The source positions are marked by the red stars. On the right, vertical velocity profiles at x=36 m (solid blue line) and at x=41 m (dashed red line) are shown together with the true model (thick grey line) and the initial model (dash-dotted black line). The locations of the profiles are marked by the blue and red lines in the models. CPML boundaries as well as the lower part of the models are cut off.

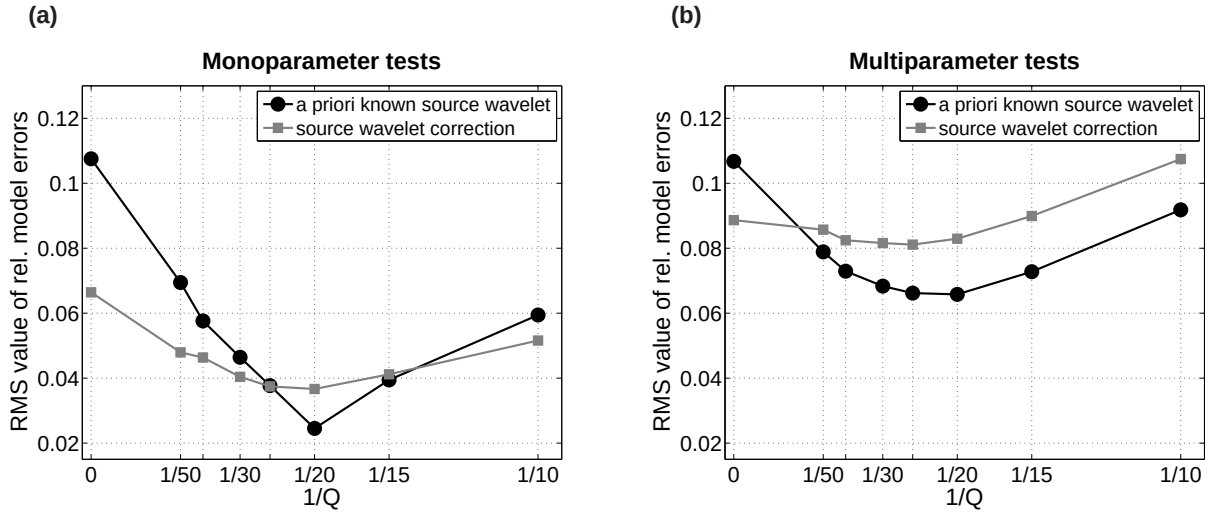


Figure 4.11: RMS value of relative model errors for S-wave velocity model for a) monoparameter reconstruction tests and b) multiparameter reconstruction tests. The black circles display the results for the tests with a priori known source wavelet and the grey squares display the results obtained for the tests with source wavelet correction. The results obtained with elastic forward modeling are displayed at $1/Q = 0$ since Q is infinite in the elastic case.

I use the same source wavelet that was already used to simulate the pseudo-observed data to model the synthetic data in the inversion. In the low frequency range up to 40 Hz I use no offset dependent weighting of the data for the determination of the source wavelet correction filter because an offset dependent weighting would support the correction of phase differences in the far-offset traces by the source wavelet correction. However, at the beginning of the inversion most of these differences are due to the subsurface structure and should not be projected into the source wavelet correction. Furthermore, for low frequencies the differences of viscoelastically and elastically modeled data are small and the lowpass effect of the source wavelet correction is therefore not needed. However, for higher frequencies and too high Q values the high-frequency ringing (Figure 4.4) in the synthetic seismograms increases. To suppress these high frequencies I use an offset dependent weighting of $(r/\mathrm{m})^{0.8}$ in the inversion of the source wavelet correction filters such that near-offset traces and far-offset traces can contribute to the misfit by the same amount.

Although the source wavelet inversion and the FWI are applied sequentially they influence each other. For example, at the beginning of the inversion the subsurface models are far away from the true models and therefore the differences in the waveforms caused by different subsurface models are at least partly projected into the source wavelet correction. This effectively increases the ambiguity of the FWI.

Monoparameter tests

Figure 4.13 displays the reconstructed S-wave velocity models that are obtained by additionally applying a source wavelet correction in the FWI. For the correct quality factors of Q_p=Q_s=20 the reconstruction of the S-wave velocity model (Figure 4.13a) as well as the fit of the observed data

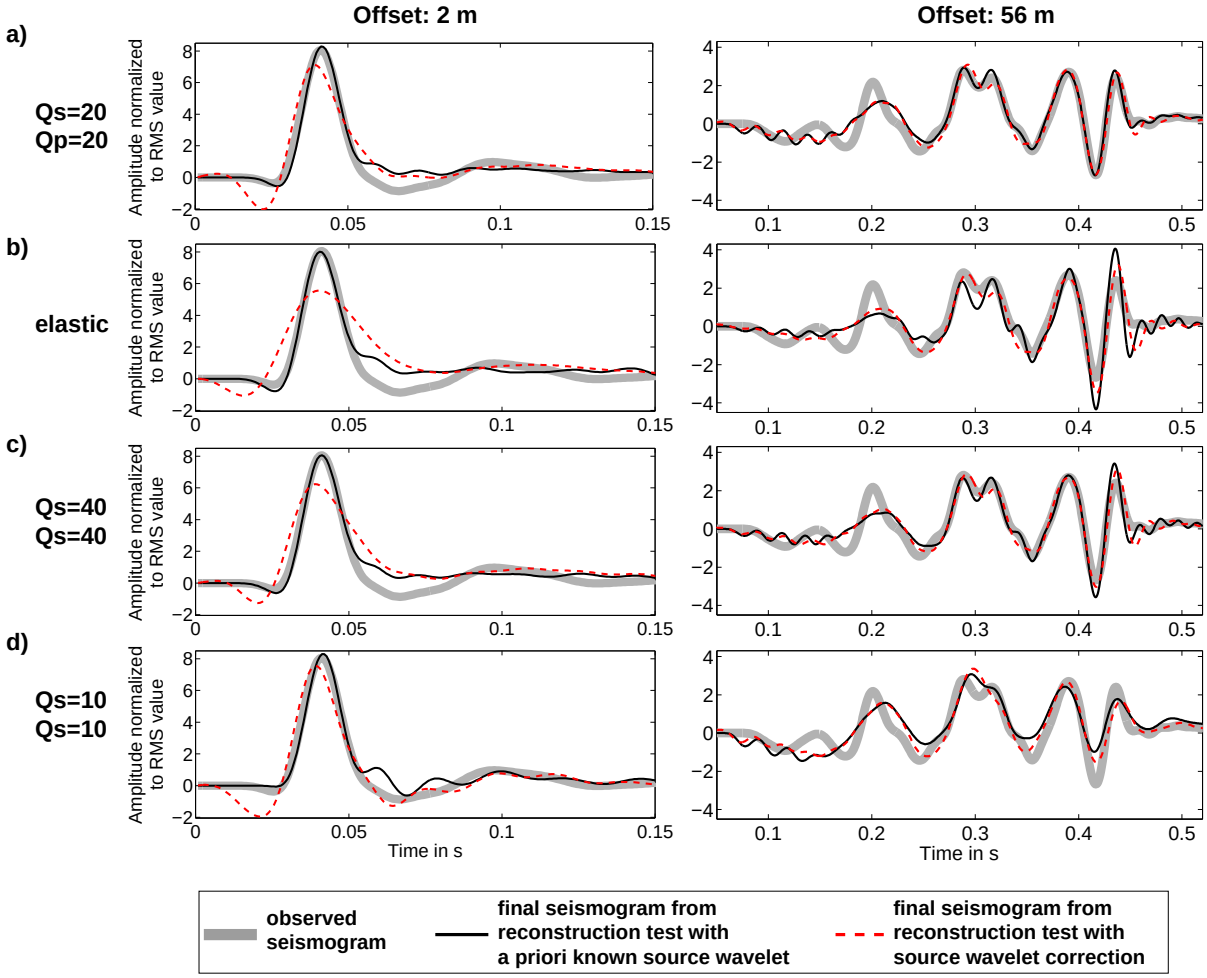


Figure 4.12: Comparison of vertical displacement seismograms for multiparameter reconstruction tests of the shot located at x=11 m for a near-offset trace at 2 m offset (left side) and a far-offset trace at 56 m offset (right side). All seismograms are normalized to their RMS amplitude. Pseudo-observed data are displayed by the thick grey line. Final seismograms of reconstruction tests with a priori known source wavelet are displayed in black and final seismograms of the reconstruction tests with application of a source wavelet correction are displayed by the dashed red line.

(Figure 4.8 and 4.9a) are again very good. In comparison to the result obtained without source wavelet correction (Figure 4.7a) the small artefacts in the vicinity of the sources decrease. In contrast, the reconstruction in larger depths slightly degrades presumably due to the higher ambiguity of the inversion when additionally applying a source wavelet correction. Figure 4.14 displays the corrected source wavelets of the shot located at x=11 m (shown in Figure 4.9) that are obtained at the end of the reconstruction tests with source wavelet correction. The thick grey line is the true source wavelet as used for the calculation of the pseudo-observed data. The source wavelet is almost not influenced by the source wavelet correction when using the correct quality factors of Q_s=Q_p=20.

If we neglect damping and use purely elastic wave propagation in the FWI the reconstructed S-wave velocity model improves significantly by applying a source wavelet correction (Figure 4.13b). The artefacts decrease in comparison to the results without source wavelet correction (Figure 4.7b) and the reconstructed S-wave velocity structure is approximately one-dimensional. The waveform residuals that are caused by the different rheology are now partly compensated by the source wavelet correction which acts as a lowpass filter and removes the high frequency parts in the forward modeled data in the FWI. As the source wavelet correction filters can only act as frequency dependent and not distance dependent filters their application leads to a worse fit for the near-offset traces because the final seismograms of the FWI are too band limited compared to the pseudo-observed data (Figure 4.9b). However, the fit of the middle and far-offset traces improves and therefore we observe a decrease of the entire final data misfit (Figure 4.8) in comparison to the corresponding test without source wavelet correction. For the reconstruction tests using elastic forward modeling this effect can be observed in Figure 4.15 where the contribution of each receiver (x plus y-component) to the final datamisfit is plotted color coded (first row of Figure 4.15). The lowpass filtering effect can be also observed in the corrected source wavelet in Figure 4.14.

The reconstructed S-wave velocity model also improves by the application of a source wavelet correction when we use viscoelastic forward modeling with Q_s=Q_p=40 in the FWI (Figure 4.13c). Again the source wavelet correction reduces the residuals between the pseudo-observed data and the synthetic data of the inversion partially. However, the lowpass effect of the source wavelet correction is not as strong as in the test where we use purely elastic wave propagation (Figure 4.14 and seismogram at 2 m offset in Figure 4.9c).

For too low quality factors of Q_s=Q_p=10 the reconstruction of the S-wave velocity improves slightly by applying the source wavelet correction (Figure 4.11a and Figure 4.13d). In this case the synthetic data in the inversion are more band limited than the pseudo-observed data. Therefore, the source wavelet correction has to increase the bandwidth of the synthetic data to compensate for the differences between the wavefields. However, this increase of the bandwidth is limited by the damping that is used in the optimization problem for determining the source wavelet correction (parameter ε^2 in Appendix F). I use in all reconstruction tests 0.1% of the energy of the synthetic data as stabilization in the deconvolution to find the source wavelet correction filters (ε^2=0.001 in Equation F.3). This value is chosen heuristically. It is not estimated by the L-curve criterion since it would require to apply the L-curve criterion to each determination of the source wavelet correction individually which is rather difficult to automate. However, as the

inversion for the correct source wavelet is not the primary goal of these tests I choose a comparably small value for the stabilization. This ensures that the waveforms are fitted well rather than that the norm of the model vector is small or that the source wavelet correction filter is as simple as possible, respectively. Because there is no noise in synthetic data I could perhaps decrease the stabilization in these tests and would still obtain a stable result for the source wavelet correction followed by a better reconstructed S-wave velocity model. However, in an FWI of field data a decrease of the stabilization is critical because of the lower signal-to-noise ratio. In the reconstruction test with Q_s=Q_p=10 the source wavelet correction acts as a highpass filter which can be observed in the final wavelet by the post-pulse oscillation (Figure 4.14).

Multiparameter tests

In comparison to the monoparameter tests I observe two major differences in the multiparameter tests. When I apply a source wavelet correction, which is required in field data inversion, the reconstructed S-wave velocity models degrades for all tests with viscoelastic modeling (Figure 4.11b). An improvement in the reconstruction of the S-wave velocity model is achieved only in the case of elastic forward modeling. Figure 4.16 shows relative model errors for the final S-wave velocity models for multiparameter reconstruction tests with a priori known source wavelet and with source wavelet correction. The displayed model errors δv are calculated at each grid point via

$$\delta v = \frac{v_s - v_{s_{true}}}{v_{s_{true}}} \tag{4.1}$$

where v_s is the obtained S-wave velocity and $v_{s_{true}}$ is the true S-wave velocity. The reconstruction of the S-wave velocity model mainly degrades in the first meter if a source wavelet correction is applied. This can be observed by larger relative model errors in the shallow part for the results obtained with source wavelet correction in comparison to the results obtained with a priori known source wavelet. It is not clear yet why the errors mainly increase in the very shallow part of the S-wave velocity model. Bretaudeau et al. (2013) observe similar artefacts that appear close to the source positions in an FWI of ultrasonic data. They use the same method to correct for the source wavelet and relate the artefacts to an imperfect reconstruction of the actual source wavelet. However, this result might also indicate undesired effects due to trade-off between S-wave velocity and P-wave velocity structure. Nevertheless, for most tested quality factors the results are still better than the result obtained with purely elastic forward modeling (Figure 4.11b). Exceptions are a quality factor of Q_s=Q_p=15 where the result is as good as in the elastic case and a quality factor of Q_s=Q_p=10 where the result is significantly worse than in the elastic case. This is presumably due to the limited capability of the source wavelet correction to increase the bandwidth of the synthetic seismograms.

The final data misfit increases in the multiparameter tests when a source wavelet correction is applied (Figure 4.8b). The fit of the near-offset traces gets worse whereas the fit of the far-offset traces does not change significantly (Figure 4.15). This is a contrast to the observation in the monoparameter tests where the data misfit for quality factors of 10, 40, and 50 as well as for elastic forward modeling decreased by the application of a source wavelet correction (Figure 4.8a).

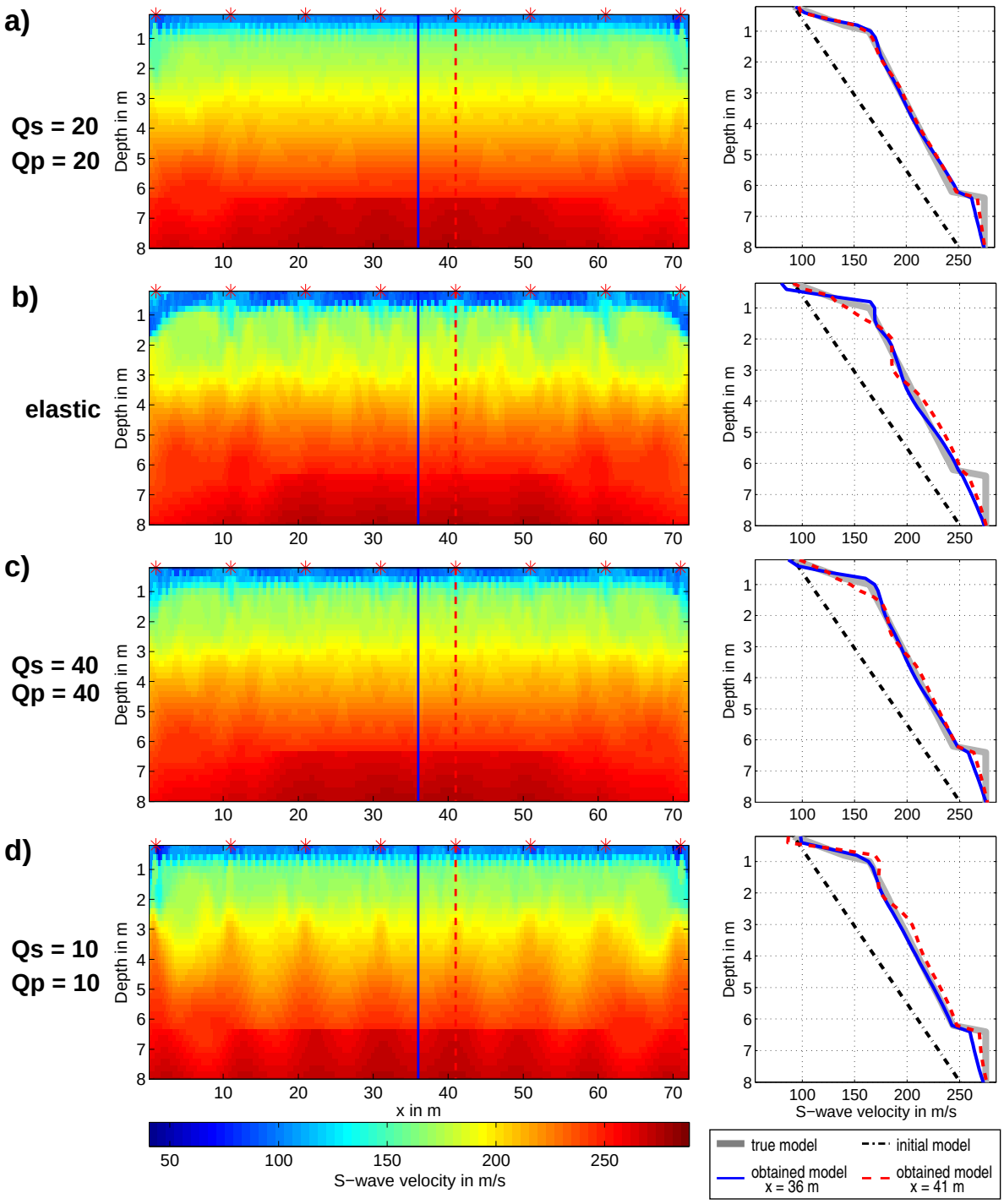


Figure 4.13: Results of monoparameter reconstruction tests using a source wavelet correction. On the left, the obtained S-wave velocity models are displayed. The source positions are marked by the red stars. On the right, vertical velocity profiles at x=36 m (solid blue line) and x=41 m (dashed red line) are shown together with the true model (thick grey line) and the initial model (dash-dotted black line). The locations of the profiles are marked by the blue and red lines in the models. CPML boundaries as well as the lower part of the models are cut off.

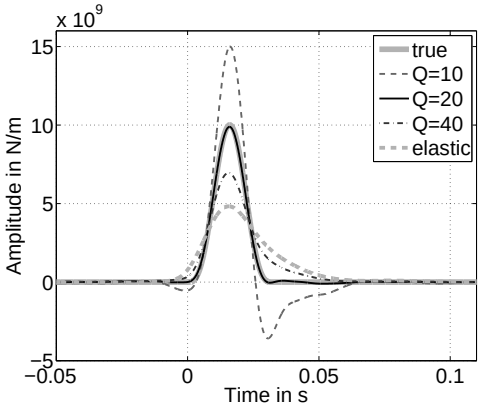


Figure 4.14: Final source wavelets obtained in the monoparameter tests with source wavelet correction for the shot located at x=11 m. The thick grey line displays the true source wavelet used to calculate the pseudo-observed data. This figure only shows a small section of the whole time series. Outside this time frame the wavelets effectively vanish.

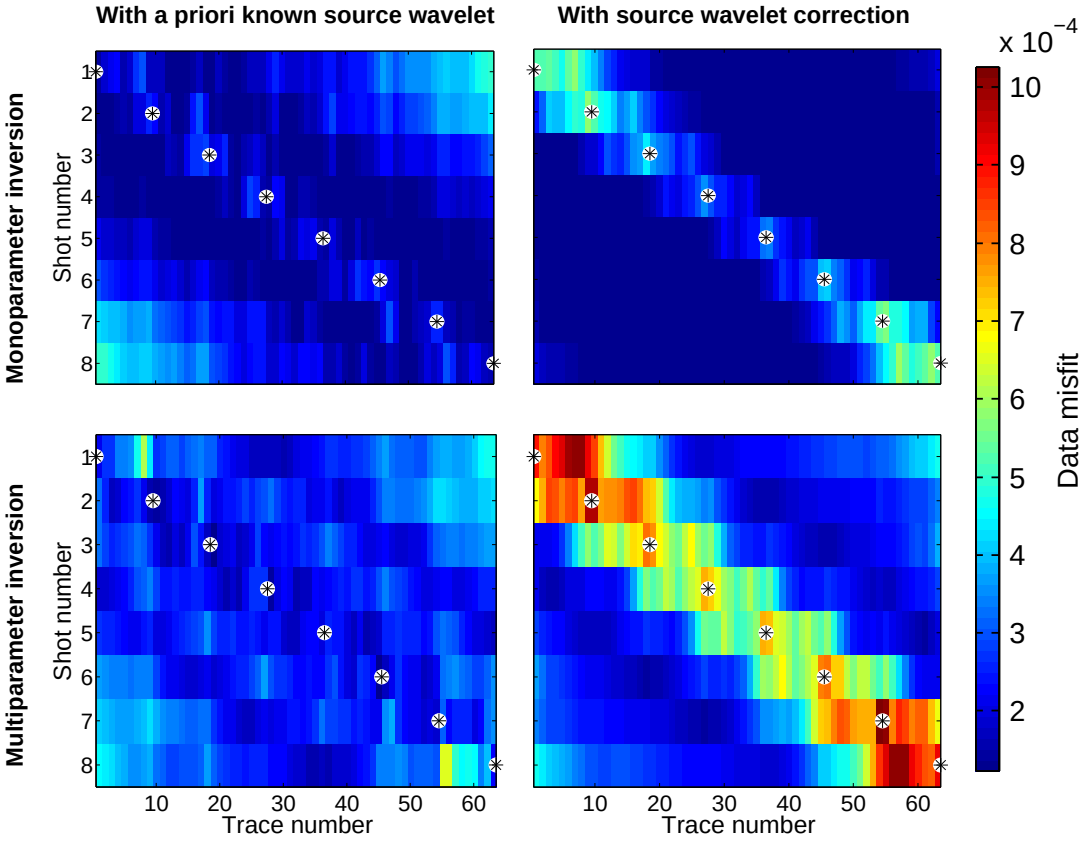


Figure 4.15: Contribution to the final data misfit of single receivers (x plus y component) for the reconstruction test where purely elastic forward modeling is used in the inversion. The first row displays the results for the monoparameter inversion test and the second row for the multiparameter inversion test. The first column corresponds to the results for the tests with a priori known source wavelet and the second to the results with source wavelet correction. The sum of the misfit contributions shown in each figure gives the final data misfit for each of the four displayed tests. The black stars mark the source positions.

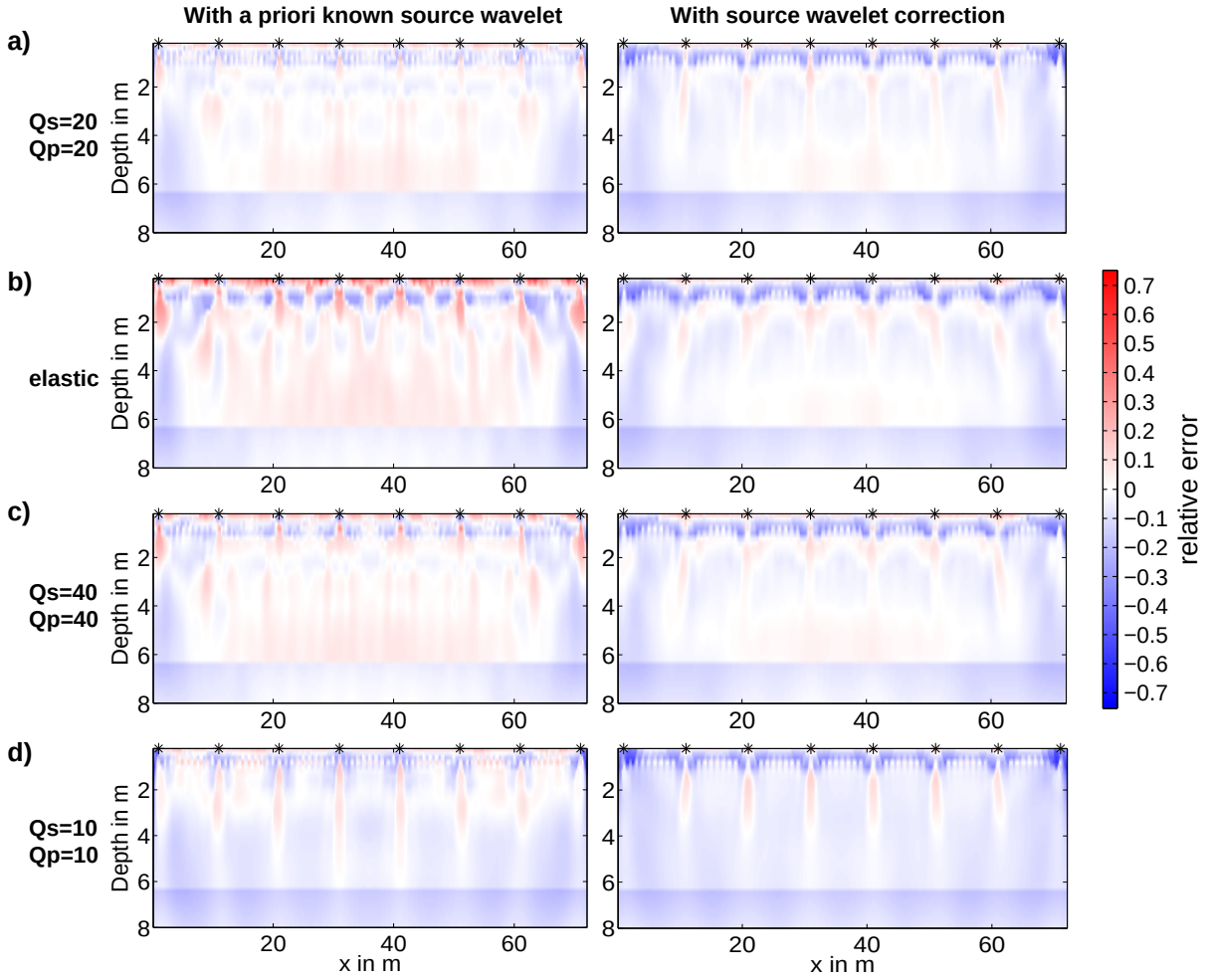


Figure 4.16: Relative errors of final S-wave velocity models for multiparameter reconstruction tests calculated by Equation (4.1). The relative errors of the results obtained with known source wavelet (left) are compared with the results obtained with source wavelet correction (right). CPML boundaries as well as the lower part of the model are cut off.

Conclusions

The tests with a priori known source wavelets show that the reconstruction of the S-wave velocity model from viscoelastic data is not possible if purely elastic modeling is used in the inversion. The result improves if additionally a source wavelet correction is applied. Nevertheless, a source wavelet correction cannot completely compensate the effects caused by anelastic damping. Therefore, the reconstruction of the S-wave velocity model is superior when a priori known quality factors are used in the inversion.

4.3 Influence of P-wave velocity model

I perform further reconstruction tests to investigate the influence of the initial P-wave velocity model on the reconstruction of the S-wave velocity model. The pseudo-observed data are inverted by FWI and the handling of the P-wave velocity model in the inversion is varied. In general, we focus on the inversion for S-wave velocity because of the high sensitivity of surface waves to this parameter. However, for the forward modeling during FWI a P-wave velocity model must be assumed as well as a strategy how to handle it. One can either invert for the P-wave velocity or the P-wave velocity model can be used as fixed model in the inversion.

Details about the true subsurface model, the acquisition geometry and the inversion setup used for the reconstruction tests are described in Section 4.1. Again viscoelastic data with Q_s=Q_p=20 are used as pseudo-observed data.

4.3.1 Results of reconstruction tests

I describe four reconstruction tests. I use the initial P-wave velocity model as fixed model in the inversion in the first two tests and invert only for S-wave velocity and density. In the second two reconstruction tests I invert for all three elastic parameters, P-wave velocity, S-wave velocity and density. There is no inversion for viscoelastic parameters applied. I use viscoelastic forward modeling with the correct Q values of 20 in all reconstruction tests.

A linear gradient for S-wave velocity as well as density is used as initial model in all reconstruction tests. The gradients start with the correct value of the corresponding model parameter at the surface and increase linearly up to the corresponding half-space values in a depth of 9 m. Below a homogeneous half-space with the correct half-space values is assumed. The model used for P-wave velocity varies between the reconstruction tests.

Test 1: P-wave velocity model fixed, true P-wave velocity model

In the first reconstruction test I only invert for S-wave velocity and density and use the true P-wave velocity model as fixed model in the inversion. The result of the reconstruction test is shown in Figure 4.17. The reconstruction of the S-wave velocity model is almost perfect (Figure 4.17a and d). The resulting structure is nearly 1D and all characteristics of the S-wave velocity model (both gradients as well as the small discontinuity) are reconstructed successfully. The reconstruction

of the density model is also possible in this test (Figure 4.17c and f) although density is not constrained as well by the data as S-wave velocity. The deviation of the density values in the layer and the half-space from the density values in the true model are small and only few 2D structures are observed in the reconstructed density model. The good reconstruction of the density model is most likely due to the almost perfect conditions in this reconstruction test. The pseudo-observed data do not contain noise and the true P-wave velocity model is used in the inversion. Figure 4.17g displays displacement seismograms of the shot located at x=1 m for the vertical component. The fit of the pseudo-observed data is perfect. There are no significant differences observable. This is confirmed by the final data misfit of E_{final}=1.3% (calculated with Equation 2.63). The initial data misfit in the inversion is $E_{initial}$= 106%.

Test 2: P-wave velocity model fixed, P-wave velocity model derived from first arrival travel times

If FWI is applied to field data the true P-wave velocity model is unknown and cannot be used in the inversion. Test 1 is unrealistic in this respect. A P-wave velocity model can be derived from the observed data by a conventional P-wave refraction analysis. I apply such an analysis to the pseudo-observed data of the two outermost shots in the model located at x=1 m and x=71 m. The P-wave velocity model shown in Figure 4.18b is obtained by averaging the results of the refraction analysis of both shots. This model is used as fixed P-wave velocity model in the second reconstruction test. Figure 4.18 displays the results of the reconstruction test. The reconstructed S-wave velocity model and the reconstructed density model contain strong artefacts. A periodic pattern which coincides with the source positions can be observed in the models. Especially the strong gradient in the topmost meter in the S-wave velocity model can not be reconstructed. However, the final data misfit of 2.5% is still low compared to the artefacts in the S-wave velocity model (initial data misfit $E_{initial}$=104%). This shows the high ambiguity of the inversion problem. In comparison to Test 1 the higher modes of the Rayleigh waves are fitted worse by the obtained model but this is hardly observable in Figure 4.18g since the differences are rather small.

Test 3: Inversion for P-wave velocity model, initial P-wave velocity model derived from first arrival travel times

In this test again the P-wave velocity model which is derived from the pseudo-observed data by an analysis of the direct and refracted P-wave onsets is used as initial model. In contrast to Test 2 I additionally invert for P-wave velocity. The results are displayed in Figure 4.19. The reconstructed S-wave velocity model is in very good agreement with the true S-wave velocity model down to approximately 4 m depth (Figure 4.19a and d). Below a small discontinuity in approximately 5.4 m depth is observed which is similar to the discontinuity in 6.3 m depth in the true S-wave velocity model. The small discontinuity in the reconstructed S-wave velocity model coincides in depth with the strong discontinuity in the initial P-wave velocity model. The depth of the discontinuity in the P-wave velocity model as well as the P-wave velocities in the half-space are not changed significantly by FWI. Thus, a clear bias of the P-wave velocity model on the reconstruction of the S-wave velocity model is observed in this test. The P-wave velocity model is only changed in the layer where a periodic pattern can be observed in the final model which coincides with the source positions. Please note that the colorbar for the P-wave velocity model

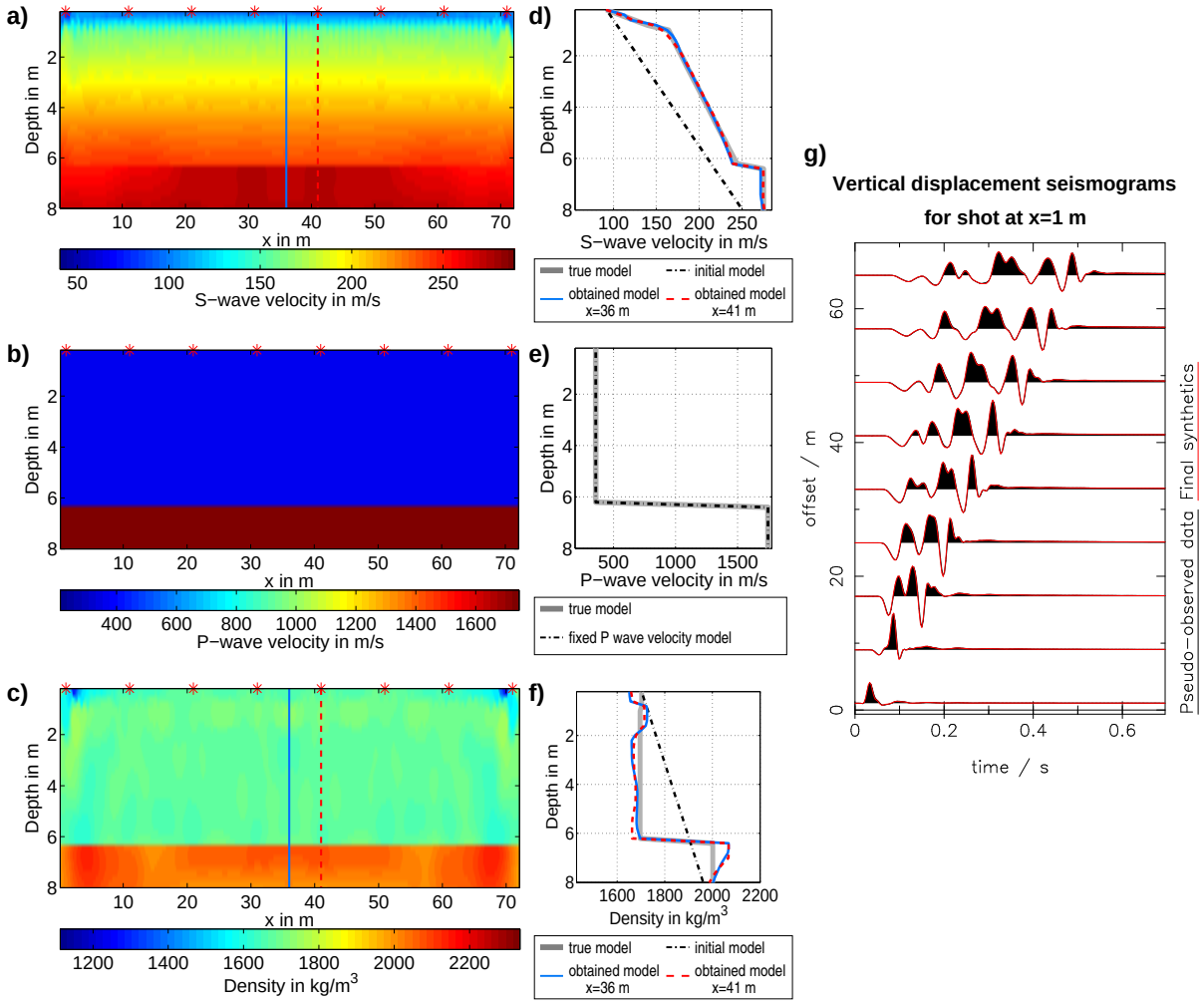


Figure 4.17: Results of reconstruction test 1. a) and c) display the obtained S-wave velocity model and density model, respectively. The red stars mark the source positions. The blue and the red dashed lines mark the position of vertical profiles which are shown in d) and f) in comparison to the true models (thick grey line) and the initial models (dash-dotted black line). b) displays the initial P-wave velocity model which is fixed in this test and which is equal to the true P-wave velocity model. g) shows vertical displacement seismograms of the left shot located at x=1 m for the pseudo-observed data (black) and the final synthetics (red). Not all seismograms of this shot are shown. The seismograms are multiplied by an offset dependent factor of $\left(\frac{r}{\mathrm{m}}\right)^{0.7}$ where r is the offset.

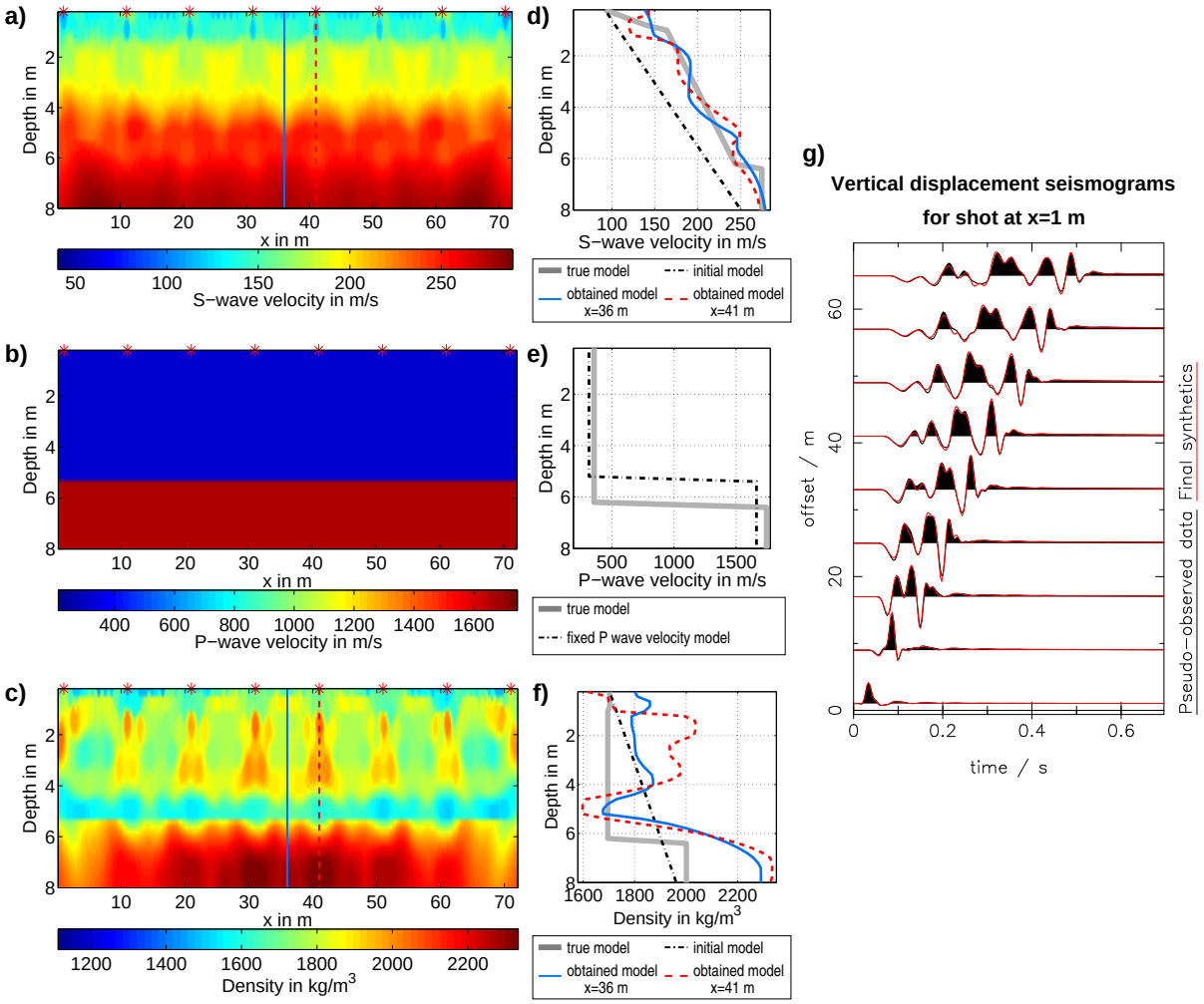


Figure 4.18: Results of reconstruction test 2. a) and c) display the obtained S-wave velocity model and density model, respectively. The red stars mark the source positions. The blue and the red dashed lines mark the position of vertical profiles which are shown in d) and f) in comparison to the true models (thick grey line) and the initial models (dash-dotted black line). b) displays the initial P-wave velocity model which is fixed in this test and e) displays a comparison of the true P-wave velocity model (thick grey line) and the initial P-wave velocity model (dash-dotted black line). g) shows vertical displacement seismograms of the left shot located at x=1 m for the pseudo-observed data (black) and the final synthetics (red). Not all seismograms of this shot are shown. The seismograms are multiplied by an offset dependent factor of $\left(\frac{r}{\mathrm{m}}\right)^{0.7}$ where r is the offset.

in Figure 4.19b is clipped. The depth of the discontinuity in the P-wave velocity model is possibly not adjusted during the FWI due to the narrow bandwidth of the pseudo-observed data (up to 70 Hz). The low frequency P-waves presumably do not constrain the depth of the discontinuity accurate enough. The reconstruction of the density model improves in comparison to Test 2. The artefacts in the layer decrease but the reconstructed density model oscillates around the true density model down to 3 m depth. A discontinuity in 5.4 m depth is observed in the density model similar to the discontinuity in the S-wave velocity model. The P-wave velocity model also bias the reconstruction of the density model in this test. In comparison to Test 2 the final data misfit decreases again and amounts to 1.6% (initial data misfit is equal to $E_{initial}$ in Test 2).

Test 4: Inversion for P-wave velocity model, linear gradient as initial P-wave velocity model

In the third test FWI is not able to shift the strong discontinuity in the P-wave velocity which is present in the initial model to the correct depth. Therefore, a last test is performed where a linear gradient is used as initial P-wave velocity model (see dash-dotted black line in Figure 4.20e). I again invert for all three elastic parameters (P-wave velocity, S-wave velocity, and density). The inversion result is shown in Figure 4.20. The initial P-wave velocity model is only changed in the first two meters of the subsurface. FWI is not able at all to reconstruct the true P-wave velocity model, in particular there is no sign of a half-space and a discontinuity on top of the half-space in the final model. This is presumably due to the poor initial P-wave velocity model used in this test. The inversion converges to a local minimum. Again a periodic pattern is observed in the S-wave velocity model which coincides with the source positions. However, in comparison to Test 2 the strong gradient in the topmost meter is reconstructed. Furthermore, the obtained S-wave velocity model deviates less from the true model compared to the reconstructed S-wave velocity model in Test 2. The small discontinuity at 6.3 m depth is not resolved. The reconstruction of the density model fails in this test. There are strong artefacts especially in the vicinity of the sources. The final data misfit of 10% in this inversion is much higher than the final data misfit in the other tests (initial data misfit in this test is $E_{initial}$=110%). The fit of the fundamental mode of the Rayleigh waves is again satisfactory (Figure 4.20g) but the fit of the P-waves and the higher modes of the Rayleigh waves is poor because the inversion was not able to reconstruct the P-wave velocity model.

4.3.2 Summary and conclusions

A significant bias of the P-wave velocity model on the reconstruction of the S-wave velocity model is observed. The wrong discontinuity in the P-wave velocity model leads to an artificial discontinuity in the reconstructed S-wave velocity model. Furthermore, the wrong P-wave velocity model that is used as fixed model in the inversion causes significant artefacts in the reconstructed S-wave velocity model.

FWI is not able to reconstruct the P-wave velocity model in these tests. The linear gradient used as initial model in Test 4 seems to be too different to the true P-wave velocity model and the inversion does not converge into the global minimum. In the initial P-wave velocity model which was estimated from first arrival P-wave travel times (Test 3) the depth of the discontinuity is not adjusted by FWI. In principle, FWI has the potential to reconstruct also the P-wave velocity

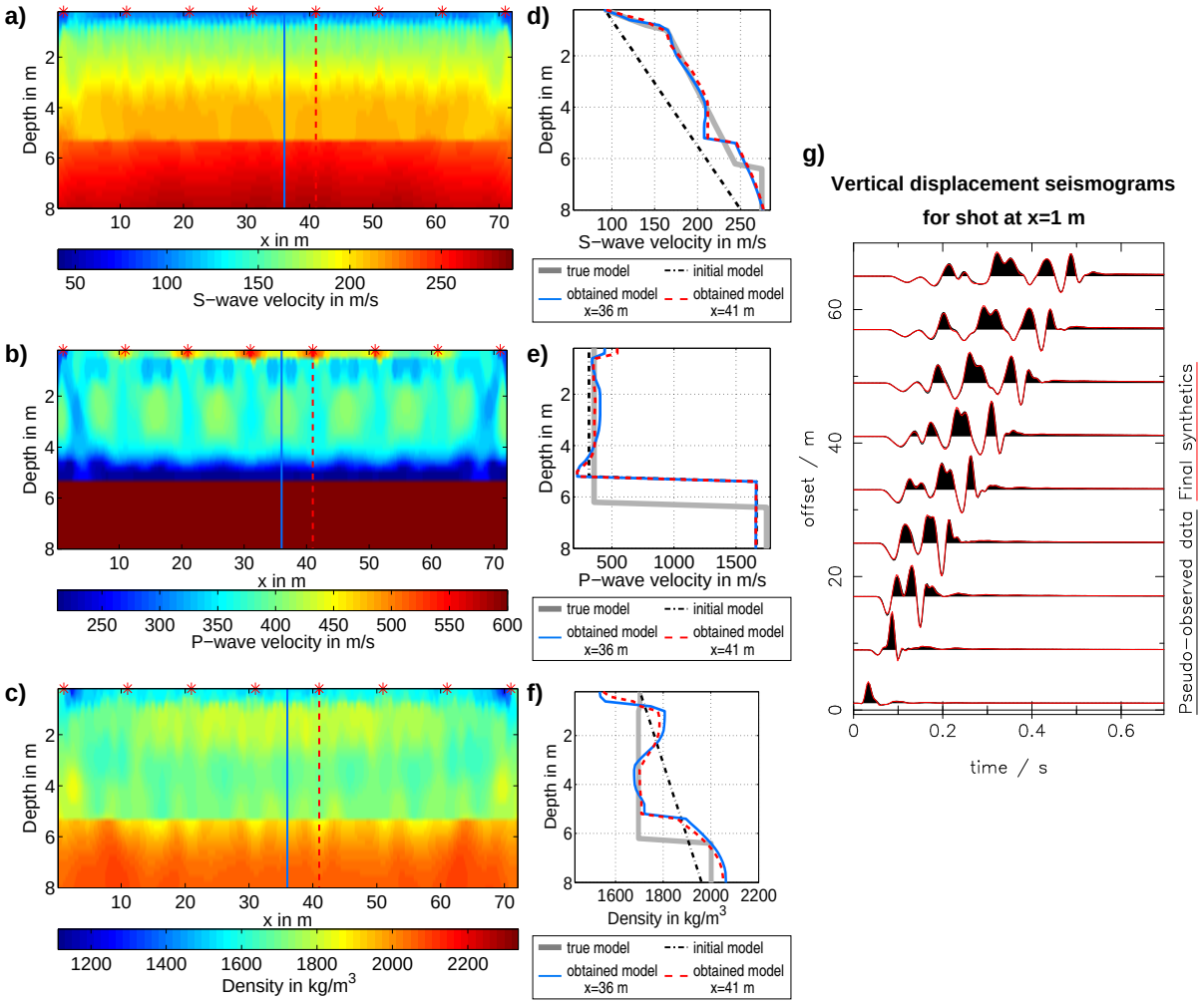


Figure 4.19: Results of reconstruction test 3. a) to c) display the obtained S-wave velocity model, P-wave velocity model, and density model, respectively. The red stars mark the source positions. The blue and the red dashed lines mark the position of vertical profiles which are shown in d) to f) in comparison to the true models (thick grey line) and the initial models (dash-dotted black line). g) shows vertical displacement seismograms of the left shot located at x=1 m for the pseudo-observed data (black) and the final synthetics (red). Not all seismograms of this shot are shown. The seismograms are multiplied by an offset dependent factor of $\left(\frac{r}{\mathrm{m}}\right)^{0.7}$ where r is the offset. The colorbar in b) is clipped. There are no significant variations of the P-wave velocity in the half-space. The P-wave velocity varies between 1656 m/s to 1664 m/s in the half-space.

model. However, as the P-waves which mainly constrain the P-wave velocity model have small amplitudes in comparison to the Rayleigh waves they do not contribute in the same extent to the data misfit as the Rayleigh waves. Applying a multi-stage inversion where in a first stage only the P-waves are inverted can possibly help to obtain a more accurate P-wave velocity model. This model can be afterwards used in an inversion of the whole data set. However, I have not tested this strategy yet.

Concluding it is desirable to use a P-wave velocity model which is as similar as possible to the true P-wave velocity model to obtain an accurate reconstruction of the S-wave velocity model as well as a fit of the higher modes of the Rayleigh waves. However, discontinuities in the reconstructed S-wave velocity model that can be related to discontinuities in the initial P-wave velocity model should be handled with care. Based on the results of these reconstruction tests I recommend to invert for P-wave velocity to reduce artefacts in the reconstructed S-wave velocity model.

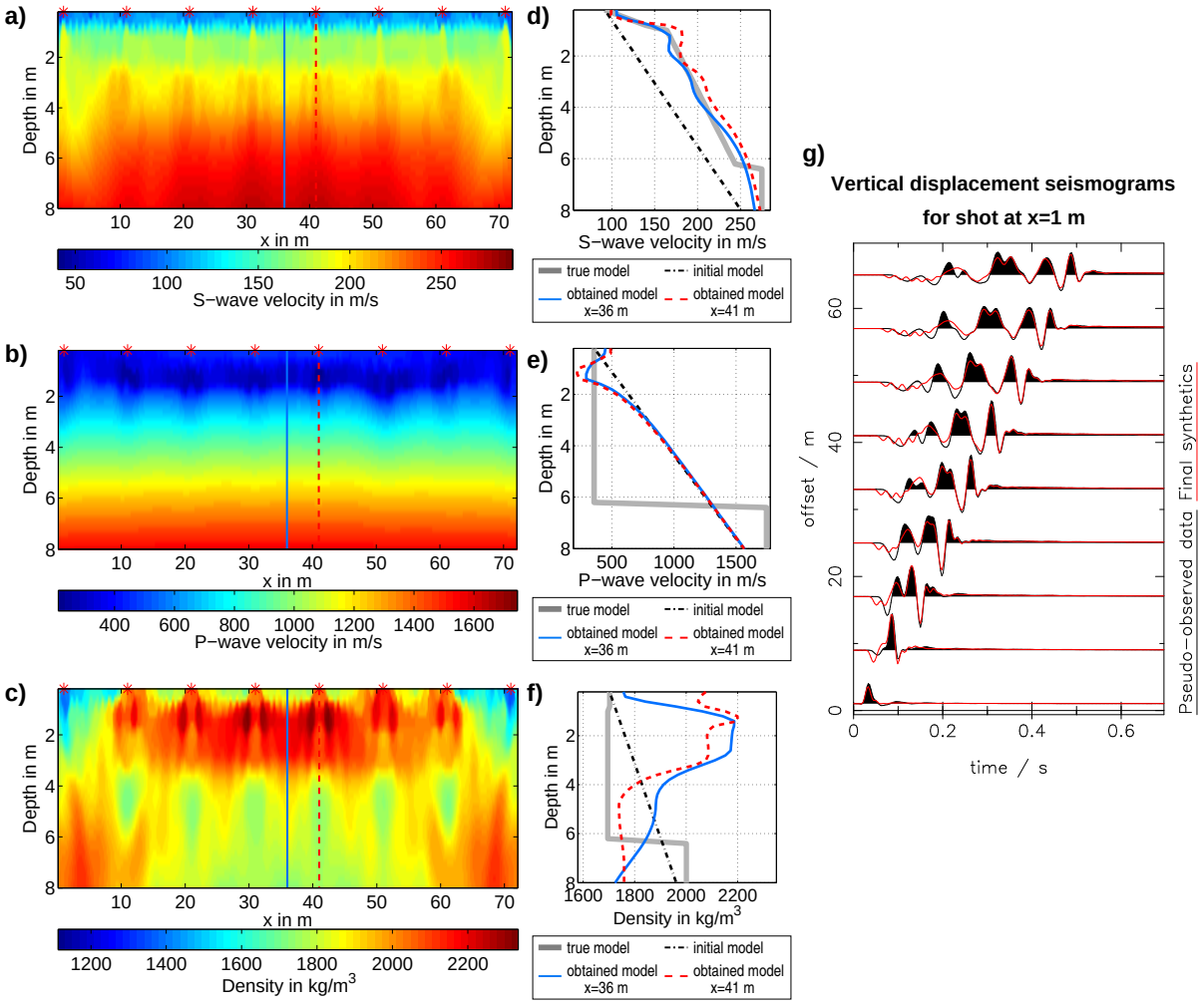


Figure 4.20: Results of reconstruction test 4. a) to c) display the obtained S-wave velocity model, P-wave velocity model, and density model, respectively. The red stars mark the source positions. The blue and the red dashed lines mark the position of vertical profiles which are shown in d) to f) in comparison to the true models (thick grey line) and the initial models (dash-dotted black line). g) shows vertical displacement seismograms of the left shot located at x=1 m for the pseudo-observed data (black) and the final synthetics (red). Not all seismograms of this shot are shown. The seismograms are multiplied by an offset dependent factor of $\left(\frac{r}{\mathrm{m}}\right)^{0.7}$ where r is the offset.

Chapter 5

Inversion of field data

Finally, I apply the discussed FWI approach to a shallow seismic field data set. After an introduction of the field data set (Section 5.1) I discuss the result of a conventional inversion of Fourier-Bessel expansion coefficients and first arrival P-wave travel times (Section 5.2). This inversion provides a 1D subsurface model and is called 1D inversion in the following. The 1D inversion is applied to the recorded data for the purpose of comparison with the 2D FWI results. Section 5.3 focuses on the results obtained by application of 2D FWI to the same data set. In Section 5.4 the 1D and 2D S-wave velocity models that are obtained by the two inversion methods are compared.

5.1 Field data set

The field data set was acquired at the gliding airfield at Rheinstetten near Karlsruhe. The seismic profile was located at the northern end of the runway and was aligned in northwest to southeast direction. The acquisition geometry consisted of a linear profile with 72 vertical geophones (eigen frequency of 4.5 Hz) with an equidistant receiver spacing of 1 m. 25 shots were excited along the profile where a vertical hammer blow was used as source. The subsurface structure consists of layered fluviatile sediments of the Late Pleistocene (Hüttner et al., 1986). Figure 5.1 displays data of a shot in the middle of the profile (shot number 13; profile meter 36.5 m). The fundamental mode as well as most phases of the higher modes coincide. Thus, it can be assumed that the subsurface structure is predominantly depth dependent (no significant lateral variations). Such a subsurface structure is called 1D structure in the following.

Although 2D FWI has the potential to reconstruct 2D subsurface models (variation of model parameters in vertical and in one horizontal direction) I applied it to a field data set recorded on a 1D subsurface structure to firstly evaluate the chosen FWI approach for field data. Furthermore, the application to a data set of a 1D subsurface structure provides the possibility to compare the FWI result to the result of a conventional 1D inversion.

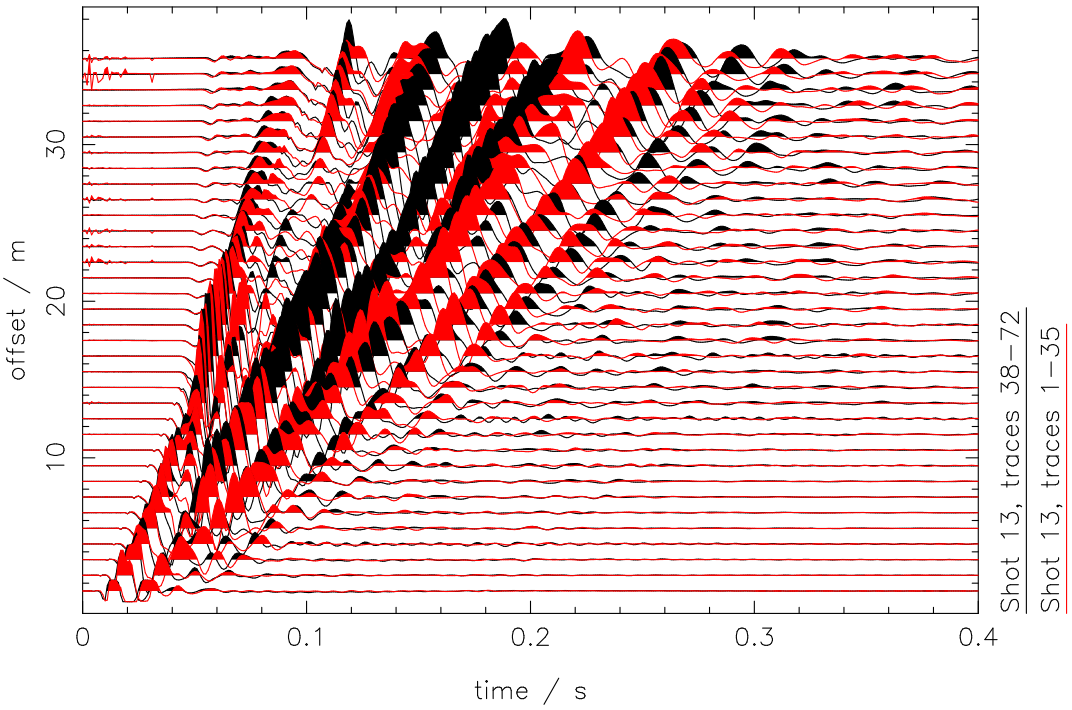


Figure 5.1: Shot gather for field data of shot 13. Displayed are the raw data (vertical velocity seismograms). Traces 38 to 72 are plotted in black and traces 1 to 35 are plotted in red. Traces 36 to 37 are clipped and therefore not shown. The seismograms are multiplied with an offset dependent factor of $(\frac{r}{\mathrm{m}})^{1.7}$ where r is the offset.

5.2 1D inversion

I apply a joint inversion of first arrival P-wave travel times and Fourier-Bessel expansion coefficients to the recorded data of shot 25 at the southeastern end of the profile. This inversion method was developed by Forbriger (2003a,b). After a short review of the inversion method (Section 5.2.1) I discuss the obtained subsurface model (Section 5.2.2) and results of a resolution analysis (Section 5.2.3).

5.2.1 Theory

If we assume a subsurface model where the material properties vary only with depth the wavefield excited by a cylindrically symmetrical point source (e. g. a vertical force source) can be expanded by Fourier-Bessel functions (see Equations 2.65 and 2.66). The expansion coefficients G can be calculated by the inverse transform from the Fourier coefficients of the wavefield $\tilde{u}$ (Equations 2.69 and 2.70). Since the wavefield is recorded only at discrete offsets r_l Forbriger (2003a) approximates the integral expressions in Equations (2.69) and (2.70) by the trapezoid rule. Furthermore, Forbriger (2003a) suggests to approximate the Bessel function $J_\eta = (H_\eta^{(1)} + H_\eta^{(2)})/2$ by the Hankel function $H_\eta^{(2)}/2$ alone, which corresponds to an expansion with waves travelling away from the source. η is the order of the Bessel and Hankel functions, respectively. η is chosen zero for the vertical component and one for the radial component. The

Fourier-Bessel expansion coefficients can be calculated by (Forbriger, 2003a, eq. 19)

$$G_\eta(\omega,p) = \frac{\omega^2}{2}\sum_{l=1}^{N} \tilde{u}_\eta(\omega,r_l)\, H_\eta^{(2)}(\omega p r_l)\, r_l\, \Delta r_l \tag{5.1}$$

with

$$\Delta r_l = \frac{1}{2}\begin{cases} r_2 - r_1 & \text{for} \quad l = 1, \\ r_N - r_{N-1} & \text{for} \quad l = N, \text{and} \\ r_{l+1} - r_{l-1} & \text{otherwise} \end{cases} \tag{5.2}$$

where $r_{l+1} \geq r_l$. According to the implementation of Forbriger (2001, 2003a,b) phase slowness $p = k/\omega$ is used here rather than wavenumber k.

The Fourier-Bessel coefficients G are inverted together with the first arrival P-wave travel times in a damped least-squares inversion (Forbriger, 2003b). Since it is not necessary to identify single modes in the Fourier-Bessel expansion coefficients the problem of misinterpretation of modes is avoided.

5.2.2 Inversion results

The subsurface model obtained by the 1D inversion of the data of shot 25 is displayed in Figure 5.2. I refer to this model in the following as 1D model. The model is parametrized by two layers overlaying a homogeneous half-space. The seismic velocities in the two layers are described by a mean velocity and a gradient. In the half-space constant seismic velocities are assumed. Furthermore, constant density values are assumed in the layers and the half-space.

The P-wave velocity model v_p is mainly constrained by the P-wave travel times and the S-wave velocity model v_s is mainly constrained by the Fourier-Bessel expansion coefficients. The depth of the discontinuity is equal for all three parameters (seismic velocities and density) due to the parametrization of the model. This depth is predominantly constrained by the first arrival P-wave travel times. Density is not well constrained by the data. The density values in the layer and the half-space are estimated based on common density values for unconsolidated sediments (Schön, 1996). At the transition to the half-space the inversion requires a strong discontinuity in the v_p/v_s ratio which is interpreted as ground water table. The v_p/v_s ratio increases from 1.56 to 5.17 at this discontinuity. The small discontinuity at 6.8 m depth in v_s might be artificially enforced by the parametrization of the model.

Figure 5.3 displays the final fit of the synthetic and observed P-wave travel times. All picked travel times are explained by the obtained P-wave velocity model within the assumed pick error of ± 2 ms. Additionally, the dispersion of the fundamental Rayleigh mode as well as the first higher mode can be explained with the derived subsurface model (Figure 5.4). The seismograms predicted by the 1D model fit the recorded seismograms well (see Section 5.4).

The derived 1D model is very similar to the model used for the reconstruction tests (Chapter 4) since both models are inferred from data sets that were acquired at the gliding airfield in Rheinstetten. The difference in the depth of the discontinuity between the second layer and the half-space is most likely due to a variation in the groundwater table.

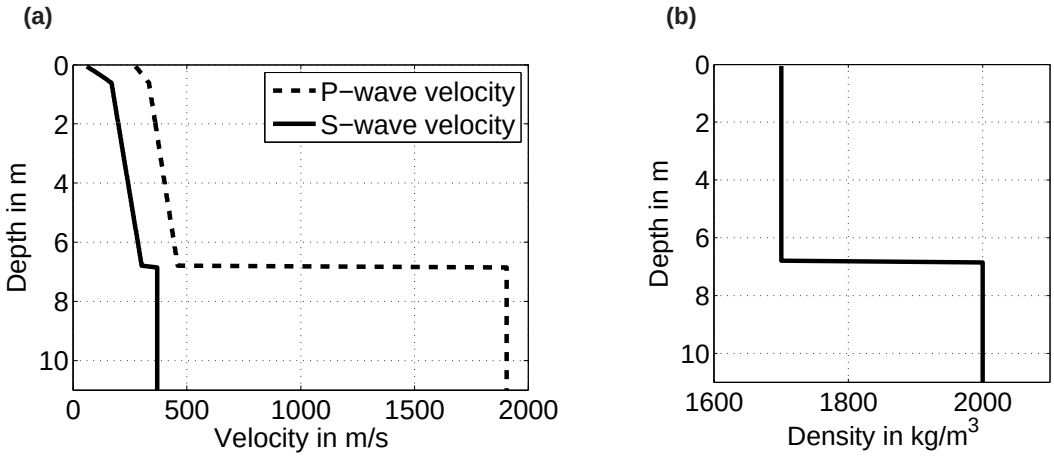


Figure 5.2: 1D subsurface model obtained from joint inversion of first arrival P-wave travel times and Fourier-Bessel expansion coefficients. a) displays the P- and S-wave velocity model and b) the density model.

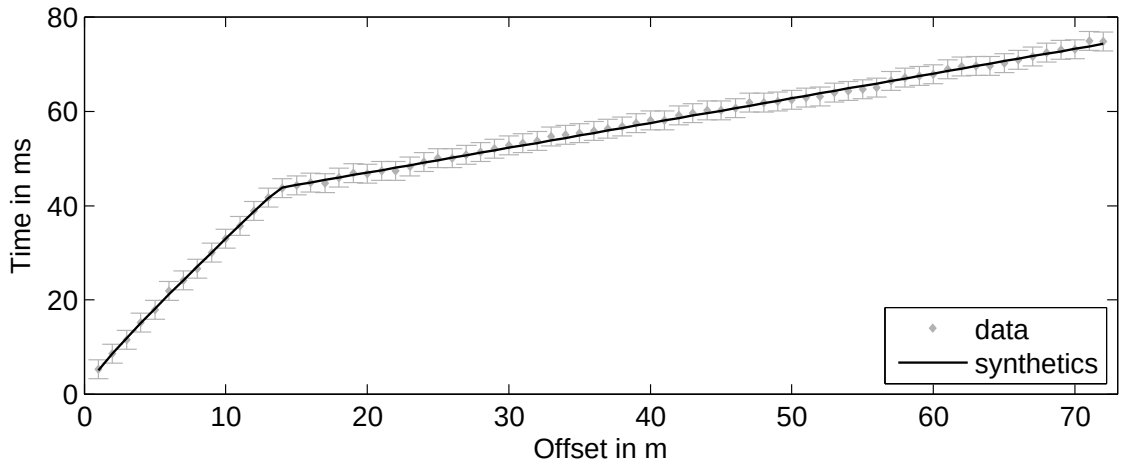


Figure 5.3: Fit of travel time data. The travel times calculated with the 1D P-wave velocity model shown in Figure 5.2a) are displayed in black and the manually picked travel times of the recorded data are displayed in grey where a constant pick error of ± 2 ms is assumed.

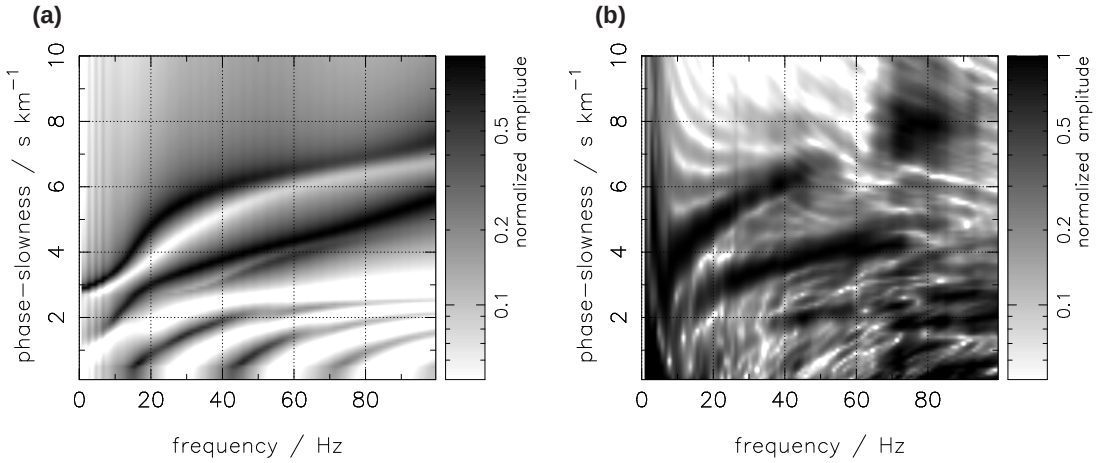


Figure 5.4: Comparison of amplitudes of complex Fourier-Bessel expansion coefficients. a) displays the amplitudes of the expansion coefficients calculated with the final subsurface model (Figure 5.2) of the 1D inversion. b) displays the amplitudes of the expansion coefficients derived from the field data of shot 25 by Equation (5.1). The expansion coefficients in both plots are scaled for each frequency individually using the maximum amplitude value.

5.2.3 Resolution analysis

Rubberband test

The 1D inversion code includes a tool for local resolution analysis which is called rubberband test (Forbriger, 2003b). In the rubberband test one starts at an optimized model that means at a minimum of the misfit function. One model parameter is now varied which causes an increase of the misfit between the synthetic and observed data. However, by changing also the other model parameters the misfit can be minimized again. This accounts for trade-off between the model parameters. In the rubberband test a maximum increase in the misfit is defined. Afterwards, the test provides the maximum model changes which are possible within the defined increase of misfit. However, the misfit increase is not calculated by actually changing the model parameters. It is found by a linear extrapolation of the predicted data using the matrix of partial derivatives. Therefore, the obtained models are no valid alternative models. For a detailed description of the rubberband test I refer to Forbriger (2003b).

Result

To apply the rubberband test the obtained subsurface model displayed in Figure 5.2 is divided into five layers overlying a homogeneous half-space (layer interfaces at 0.55 m, 2.00 m, 3.50 m, 5.00 m, and 6.84 m). Tested model parameters are the thickness of the first layer, the depth of the half-space as well as the mean P-wave velocity and S-wave velocity in each layer. The applied damping in the analysis is negligible ($\lambda < 5 \cdot 10^{-4}$ in eq. 13 in Forbriger, 2003b). The result of the rubberband test for a relative increase of the misfit of 2% is displayed in Figure 5.5. The smaller the filled areas are the better constrained is the parameter where trade-off between the parameters is taken into account (Forbriger, 2003b). The displayed models are not a crowd of alternative models since they are not obtained by an actual optimization.

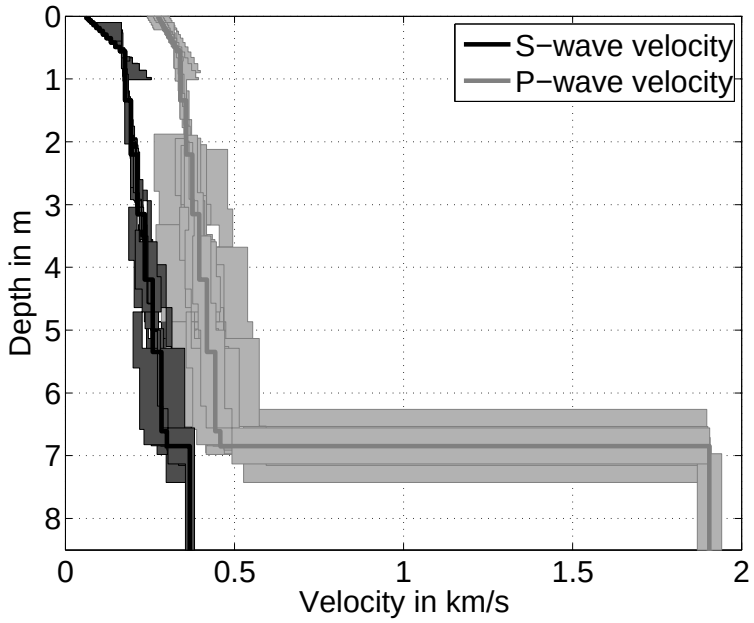


Figure 5.5: Resolution analysis with a rubberband test for the 1D inversion result. Tested parameters are the thickness of the first layer, the depth of the half-space as well as the mean P-wave and S-wave velocities in the layers and the half-space. The result is shown for a relative increase of the misfit of 2%. Small variations of a parameter can be linked to a good constrainedness of this model parameter. The thick lines display the optimized model which is obtained by the 1D inversion.

The S-wave velocity is well constrained by the Fourier-Bessel expansion coefficients. The resolution of the S-wave velocity decreases with depth. The high constrainedness of the S-wave velocity in the half-space is presumably not reliable since it is determined from only a few Fourier-Bessel expansion coefficients below 12 Hz. A trade-off between the depth of the first layer and the seismic velocities at shallow depth is observable. The P-wave velocity is constrained less than the S-wave velocity. A trade-off between the P-wave velocities in the layers exists. A faster P-wave velocity in one layer can be compensated by a slower P-wave velocity in the next layer, for example. Furthermore, a trade-off between the depth of the half-space and the P-wave velocities in the layers is observable.

5.3 2D full waveform inversion

In this section the preprocessing of the field data set is discussed (Section 5.3.1). Afterwards the inversion setup and the problem of cycle-skipping are described (Section 5.3.2 and 5.3.3). Section 5.3.4 focuses on the inversion result which is obtained by FWI.

5.3.1 Preprocessing

3D/2D transformation

I apply the hybrid transformation (Forbriger et al., 2013) to the recorded data (for details see Section 2.4.3) to account for the different geometrical spreading of the field data (3D) and the 2D forward modeled data in the FWI. Seismograms up to an offset of 10 m are transformed with the single-velocity transformation where a phase velocity of v_{ph}=300 m/s is used. This phase velocity is estimated from the seismograms with small offsets. For seismograms with offsets larger than 20 m the direct-wave transformation is applied. Between 10 m and 20 m both transformations are mixed with a continuous transition from one transformation to the other.

The convolution with $1/\sqrt{t}$ that is applied to the data within the hybrid transformation corresponds to a half integration of the seismograms (Forbriger et al., 2013). In the used field data set this causes strong artefacts in some traces due to low-frequency noise. To reduce these artefacts I apply a fourth order Butterworth highpass filter with a corner frequency of 5 Hz to the field data prior to the 3D/2D transformation. Furthermore, the seismograms are tapered. The applied taper ensures that the seismograms are zero before the P-wave onset and after the fundamental mode of the Rayleigh waves.

Estimation of Q value

The reconstruction tests presented in Section 4.2.2 show that viscoelastic wave propagation must be considered in an FWI of shallow seismic Rayleigh waves. To estimate a quality factor which is used as a priori known parameter in the FWI I compare the field data with viscoelastically forward modeled data as described in Section 4.2.1. I again assume no spatial variation of the quality factors and used $Q_p = Q_s$ where Q_p is the quality factor for P-waves and Q_s is the quality factor for S-waves.

To generate synthetic data with different quality factors the 1D model introduced in Section 5.2 is used. Prior to the comparison of field data and synthetic data a source wavelet correction filter is determined (Section 4.2.1 and Appendix F) and applied to the synthetic data. For each quality factor and each shot a separate source wavelet correction filter is estimated. In the optimization of the source wavelet correction filters a weighting of the data misfit of $(r/\mathrm{m})^{1.2}$ is used (Equations (F.4) and (F.1) in Appendix F). The final data misfit of all 25 shots after the source wavelet correction is displayed in Figure 5.6 as a function of the tested quality factors. The weighting used in the estimation of the source wavelet correction filters is included in this analysis. A minimum of the L2 misfit is observed between quality factors of 8 and 15. This result is also confirmed by a qualitative comparison of the field data and the synthetic data. For the FWI I choose a quality factor of 15 although the L2 data misfit for a quality factor of 10 is less. However, the source wavelet correction can account better for differences between field data and synthetic data that are caused by too high quality factors used to generate the synthetic data. In this case the source wavelet correction filters act additionally as lowpass filters (Section 4.2.1). If a too low quality factor is used to generate the synthetics the source wavelet correction filters must increase the bandwidth of the synthetic data. However, this is limited by the stabilization that must be used in the optimization of the source wavelet correction filter (Section 4.2.2).

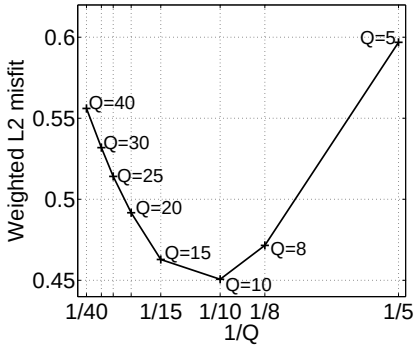


Figure 5.6: Estimation of a quality factor for field data set. All 25 shots are considered in the analysis. The weighted L2 misfit is the sum of the data misfits which are minimized in the estimation of the source wavelet correction filters.

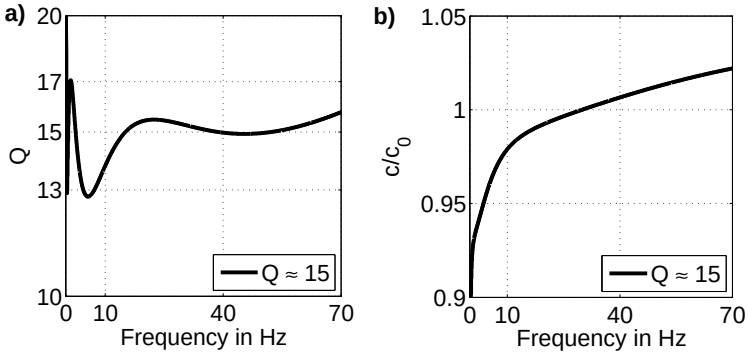


Figure 5.7: Rheological model used for FWI of field data set. The model is built by a generalized standard linear solid with three relaxation mechanisms and approximates a constant quality factor of 15. a) displays the frequency dependence of the quality factor Q and b) the corresponding phase velocity dispersion caused by damping. c is the phase velocity and c_0 is the phase velocity at the used reference frequency of 30 Hz.

Figure 5.7 displays the rheological model that is used in the FWI. The quality factor of 15 is approximated by three relaxation mechanisms of a generalized standard linear solid with the relaxation frequencies $f_{\sigma,1}$= 0.29 Hz, $f_{\sigma,2} = 5.73$ Hz, and $f_{\sigma,3} = 64.51$ Hz and $\tau = 0.15$ (see Appendix A for details about the rheological model of a generalized standard linear solid).

5.3.2 Inversion setup

In the FWI of field data I use a very similar inversion setup as in the reconstruction tests. Again a multiscale inversion approach with frequency filtering is applied. I start with a passband from 5 Hz to 10 Hz and increase the upper corner frequency sequentially to 15 Hz, 20 Hz, 30 Hz, 40 Hz, 55 Hz, and 70 Hz. The bandlimited data are obtained by the application of a fourth-order Butterworth highpass filter with a corner frequency of 5 Hz and a fourth-order Butterworth lowpass filter with different corner frequencies (10 Hz, 15 Hz, etc.). The highpass filter is already applied to the field data prior to the 3D/2D transformation. The bandwidth of the data is increased in the inversion as soon as the relative misfit change of the current iteration in comparison to the last but one iteration drops below 1%. In contrast to the reconstruction tests no minimum number of iterations per frequency interval is performed. The relevant parameters used for forward modeling are given in Table H.2 in Appendix H.

At the beginning of each frequency interval a source wavelet correction filter for each shot is determined by using synthetic data generated with the current subsurface model in a stabilized deconvolution with the field data as described in Appendix F. The source wavelet correction filters are estimated once at the beginning of each frequency band and are then used unaltered for the corresponding passband. For the determination of the source wavelet correction filters only 10 to 20 seismograms with small offsets are used per shot. The used traces are chosen manually and the number of used seismograms depends on data quality. It is undesirable that

the source wavelet correction filters account for differences in the wavefields that are caused by differences in the subsurface models. Therefore, only seismograms with small offsets are used in the estimation of the correction filters.

The misfit function is defined by the L2 norm of the normalized seismograms (Equation 2.63). For gradient preconditioning I use the method described in Section 2.3.1 where the semi-circular tapers have a radius of 3.0 m. Additionally, the gradients and the models are smoothed with a 2D median filter. The filter lengths vary for the different passbands considered in the multi-scale inversion and for the different parameters. The used filter lengths are given in Table H.3 in Appendix H. The filter lengths are smaller than half of the minimum wavelength of the current bandwidth for the S-wave velocity model in all frequency bands. For the P-wave velocity stronger smoothing must be applied in the frequency bands between 5 Hz-40 Hz, 5 Hz-55 Hz, and 5 Hz-70 Hz in comparison to the S-wave velocity to avoid instabilities during forward modeling due to small-scale heterogeneities in the model. However, the used filter lengths are still smaller than the minimum wavelength of the P-waves.

As initial P-wave velocity model for the FWI the 1D model shown in Figure 5.2a) is used as it is mainly constrained by the first arrival P-wave travel times and might be therefore also obtained by a P-wave refraction tomography of the data set. As initial density model the model shown in Figure 5.2b) is used. As initial S-wave velocity model I test two different models where one of them produces cycle-skipping and the inversion gets stuck in a local minimum. This is discussed in the following section.

5.3.3 The problem of cycle-skipping

I tested two different initial S-wave velocity models. Both consist of a linear gradient. Such initial S-wave velocity models have proofed to be appropriate in synthetic tests. The first model, named initial model 1 in the following, starts with a velocity of 60 m/s at the surface and increases linearly up to a velocity of 369 m/s in a depth of 9 m where it passes into a homogeneous half-space. In contrast, the second model (initial model 2) starts at a velocity of 100 m/s at the surface and has therefore a weaker gradient. The two initial models are displayed in Figure 5.8 in comparison to the 1D S-wave velocity model which is obtained by the 1D inversion.

Initial model 1 produces cycle-skipping (Section 2.3.4) even in the frequency interval between 5 Hz and 10 Hz which is used at the beginning of the inversion (Figure 5.9a). The initial synthetics generated with initial model 1 do not fit the observed data well enough. The misfit can be significantly decreased by fitting cycles of the waveforms that do not correspond to each other. This can be observed in the traces with an offset between 30 m and 60 m in Figure 5.9a). Figure 5.10a) displays the gradient for the S-wave velocity for the first iteration step. The gradient has predominantly positive amplitudes which causes a decrease of the S-wave velocities due to the update in the opposite direction of the gradient of the misfit function. This decrease in the S-wave velocities causes a decrease of the misfit in the offset interval between 30 m and 60 m. However, the recorded data at large offsets cannot be predicted by the obtained subsurface model. After seven iterations the inversion gets stuck in a local minimum.

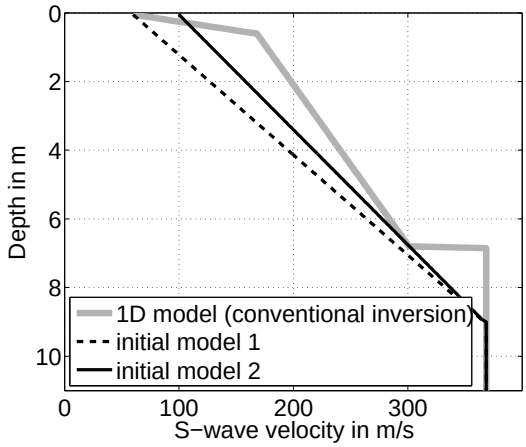


Figure 5.8: Initial S-wave velocity models used for FWI of field data (black line and black dashed line) in comparison to 1D S-wave velocity model derived by a conventional 1D inversion (inversion of Fourier-Bessel expansion coefficients and first arrival P-wave travel times; Section 5.2.2).

In contrast, the synthetics generated with initial model 2 fit the field data significantly better (Figure 5.9b) in the bandwidth between 5 Hz and 10 Hz, especially the seismograms with large offsets. For this initial model the gradient for the S-wave velocity differs clearly from the gradient for initial model 1 (Figure 5.10). For the second tested initial model the S-wave velocities between 0 m and 5 m depth are increased and the data misfit is successfully reduced for middle and far offset traces. After 15 iteration steps the inversion stops because the misfit cannot be reduced significantly any more. The fit of the recorded data is satisfactory (Figure 5.9b) and one can switch to the next broader frequency band between 5 Hz and 15 Hz.

5.3.4 2D subsurface model

Using model 2 as initial model for the S-wave velocity FWI can be successfully applied to the field data set. Figure 5.11 displays a comparison of seismograms for three representative shots. The data misfit is successfully decreased by FWI as the seismograms calculated with the final model fit the recorded data significantly better than the seismograms calculated with the initial model. Note that the seismograms displayed in Figure 5.11 have a bandwidth of 5 Hz to 70 Hz. In contrast, the seismograms displayed in Figure 5.9 have a bandwidth of only 5 Hz to 10 Hz. The decrease of the data misfit is displayed in Figure 5.12. In total, 59 iterations are performed. I run the computations with 16 CPUs on the supercomputer JUROPA at Jülich Supercomputing Centre and the total computation time for the 59 iterations is approximately 7 hours. However, the whole FWI is not carried out in this time since the source wavelet correction filters are estimated manually between the inversion of the data with different bandwidths.

Figure 5.13 displays the final 2D model obtained by FWI. There are significant 2D structures in the layer of the obtained P-wave velocity model (Figure 5.13a). In contrast, the P-wave velocity in the half-space is almost not changed during the inversion. In the half-space the P-wave velocity is 1905 m/s with a maximum variation of ± 1 m/s. The 2D structures in the P-wave velocity model influence the final data fit significantly. Synthetic seismograms that are generated with the obtained 2D S-wave velocity and density model but with the initial 1D P-wave velocity model

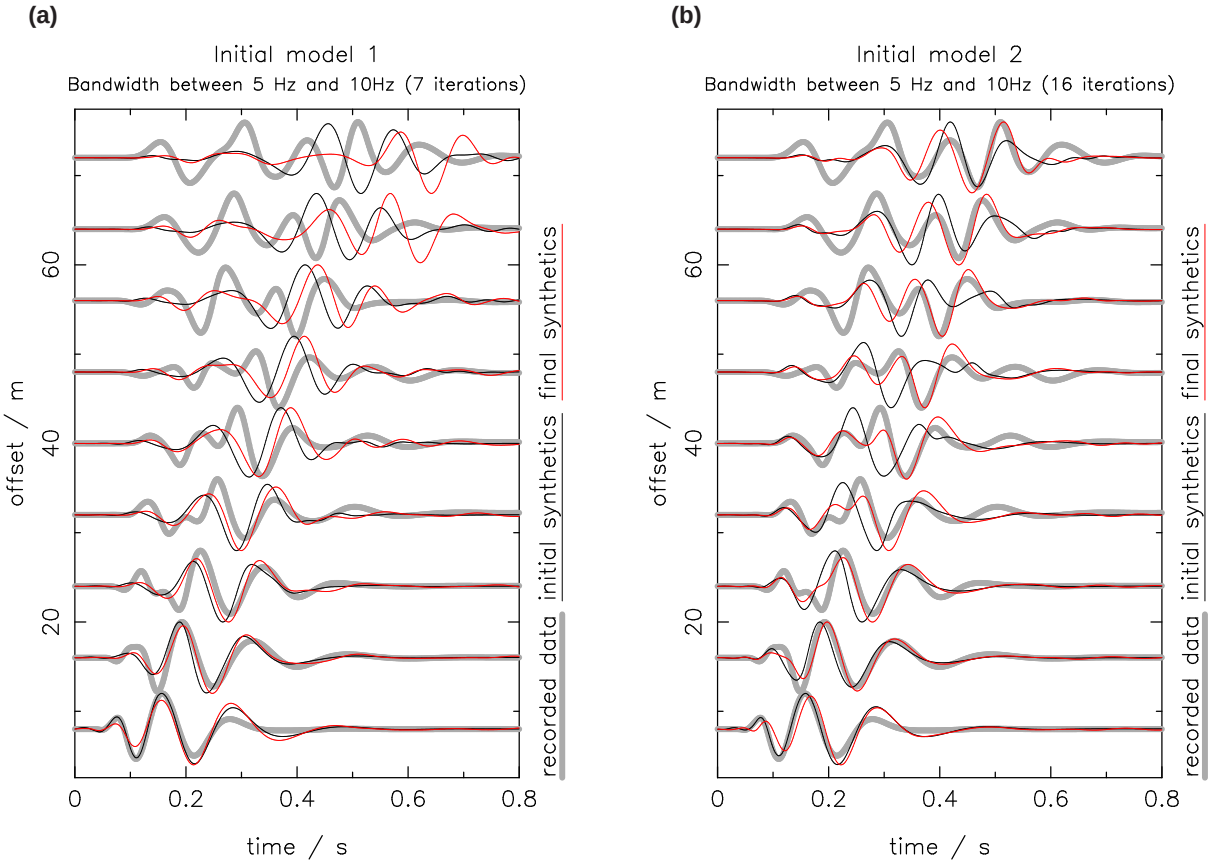


Figure 5.9: Comparison of field data (thick grey) and initial synthetics (black) for two different initial models as well as the resulting synthetics (red) after the inversion of the field data in the frequency band between 5 Hz and 10 Hz. Displayed are vertical displacement seismograms of shot 25 (shot at x=76 m). Each seismogram is normalized to its maximum amplitude. Every eighth trace is displayed. a) displays the results for initial model 1 (Figure 5.8) where cycle-skipping occurs. b) displays the results for initial model 2.

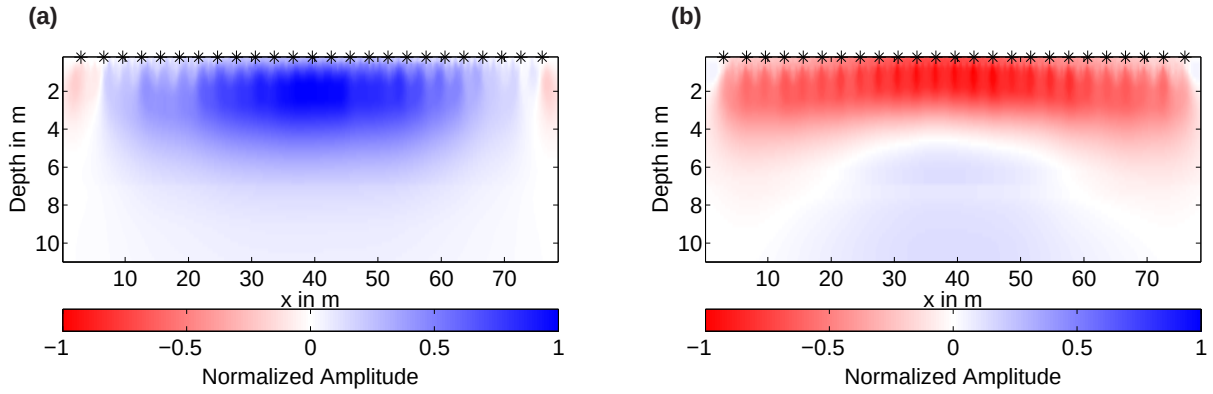


Figure 5.10: Gradients of the misfit function for the S-wave velocity in the first iteration step of the FWI using a) initial model 1 and b) initial model 2 (for models see Figure 5.8). The gradients are normalized to their maximum amplitude.

(shown in Figure 5.2a) fit the recorded data worse. This is presumably due to a trade-off between S-wave velocity and P-wave velocity. However, the changes in the S-wave velocity model influence the seismograms much stronger than the changes in the P-wave velocity model. A trade-off between P-wave velocity and S-wave velocity is possibly also influenced by the individual scaling of the gradient of each model parameter implemented in DENISE (Section 2.3.2). It is expected that the P-wave velocity is not as well constrained by the data as the S-wave velocity due to the small amplitudes of the P-waves in comparison to the amplitudes of the Rayleigh waves and due to the high sensitivity of the Rayleigh waves to the S-wave velocity structure. However, in each iteration the update of each model parameter is a percentage of the maximum value of the corresponding model parameter in the current model, independent of the absolute amplitude of the gradients for the different model parameters. This could cause a too large update of the P-wave velocity which then influences also the update of the S-wave velocity in the following iteration steps. This hypothesis should be investigated in future studies.

The S-wave velocity model (Figure 5.13b) still corresponds to a predominantly depth dependent structure although this is not enforced in the inversion by regularization. High S-wave velocities in the first two meters are observed in the left part of the model in comparison to the middle and the right part of the model (see vertical velocity profile at x=14 m in Figure 5.13b). It seems that structures that are deeper than 6.0 m cannot be resolved by the inversion. This could be possibly improved by a better preconditioning of the gradients e. g. by using the inverse main diagonal elements of the approximate Hessian matrix for preconditioning (similar to the method proposed by Sheen et al., 2006). Another possibility could be the amplification of low frequencies. The deep parts of the model are constrained by Rayleigh waves with low frequencies due to their larger penetration depth. However, the investigated data set contains at least in a few shot gathers significant low-frequency noise which is critical when low frequencies are amplified. It should be checked first if the low frequency content of the field data can provide information of the deeper parts of the S-wave velocity model or if these signals are too much disturbed by noise. Although the changes in the S-wave velocity model are small these changes have a strong influence on the wavefields (see Figure 5.11) which confirms the high sensitivity of the Rayleigh waves to the S-wave velocity model.

The obtained density model (Figure 5.13c) is even more heterogeneous than the P-wave velocity model. The density is not further interpreted because density is not well constrained by the recorded data. Synthetic seismograms that are modeled with the obtained 2D S-wave and P-wave velocity models but with the initial density model (Figure 5.13c) fit the recorded data almost equally well as the synthetics generated with final P-wave velocity model, S-wave velocity model, and density model obtained by FWI. This indicates that the influence of the density on the waveforms is rather small.

The inversion scheme does not ensure that the obtained subsurface models are physically consistent. The P-wave velocities v_p range from 220 m/s to 1905 m/s which is within the expected interval for unconsolidated sediments (Schön, 1996). The S-wave velocities v_s vary between 98 m/s and 372 m/s which is also within the expected interval for unconsolidated sediments. However, a critical parameter for this inversion result is the v_p/v_s ratio or the Poisson number ν, respectively. Reasonable values for the v_p/v_s ratio lie between 1.56 and values larger than

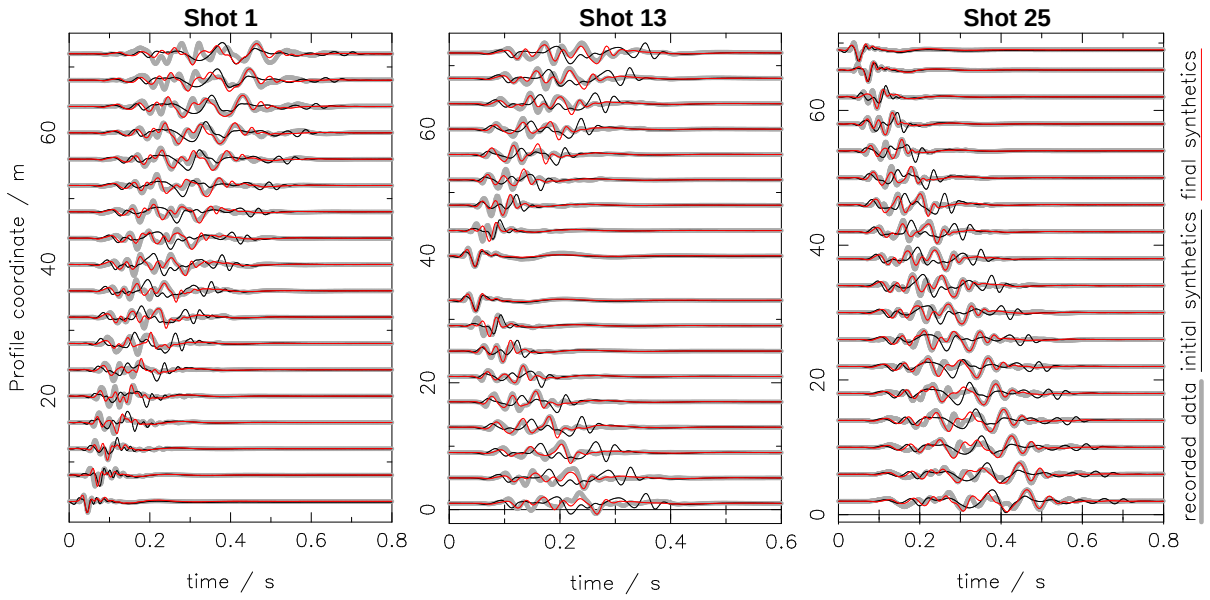


Figure 5.11: Vertical displacement seismograms for shot 1 (x=3.0 m), shot 13 (x=39.6 m), and shot 25 (x=76.0 m) in the frequency band between 5 Hz and 70 Hz. Recorded data are displayed by the thick grey line, seismograms calculated with the initial model are displayed in black and seismograms calculated with the final 2D model are displayed in red. Each trace is normalized to its maximum amplitude. Every fourth trace is displayed.

10 (Bachrach et al., 2000; Forbriger, 2003b). This corresponds to Poisson numbers between ν=0.15 and $\nu >$0.495. Stümpel et al. (1984) observe even smaller values of v_p/v_s ratios down to 1.41 (ν=0) in dry and partially saturated sands. Smaller values for the v_p/v_s ratios correspond to negative Poisson numbers. If a tension is applied to a material with negative Poisson number it expands also in the other directions. This disagrees with the physical intuition for materials that are met in the subsurface. Figure 5.14 displays the v_p/v_s ratio calculated for the 2D inversion result. The colorbar is clipped at a ratio of 1.56 and areas with higher ratios are displayed in white. The areas where the v_p/v_s ratio is smaller than 1.56 are mainly at the left and the right boundary of the model as well as in the vicinity of the discontinuity in 6.8 m depth which is caused by the groundwater table. There is especially one anomaly between x=4 m and x=10 m in a depth range between 2.5 m and 4 m where the v_p/v_s ratio drops below 1.41 which corresponds to a negative Poisson number. Therefore, this part of the model is physically not feasible. However, in the middle part of the model the obtained v_p/v_s ratios match the expected parameter range although this is not enforced by the inversion. Nevertheless, for future inversions one should implement a mechanism that avoids at least negative Poisson numbers.

5.4 Comparison of 1D and 2D S-wave velocity model

The 1D S-wave velocity model which is obtained by the 1D inversion clearly differs from the 2D S-wave velocity model obtained by FWI (vertical velocity profiles in Figure 5.13b). These differences are partly caused by the different parametrization of the models in the two inversion algorithms. In the FWI the model is parametrized by decoupled velocities at each grid point of

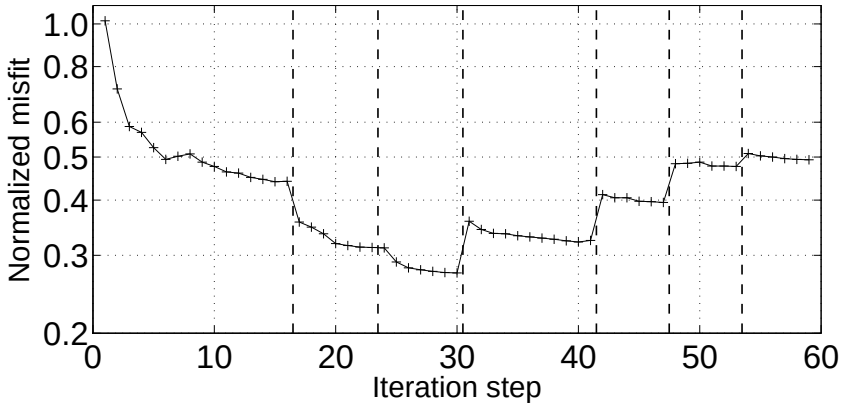


Figure 5.12: Evolution of data misfit calculated with Equation (2.63) during FWI. The dashed lines mark the increase of the bandwidth of the recorded data which are used in the inversion. At these points new source wavelet correction filters are estimated.

the FD grid (grid spacing 0.2 m) whereas in the 1D inversion code the model is divided into three layers and described by the mean velocity and a gradient in each layer. This presumably causes the differences in the first two meters of the S-wave velocity models. There is a sharp transition in the 1D model from the very steep gradient in the first layer to the weaker gradient in the second layer. In contrast, such a clear transition is not observed in the 2D model. However, in the middle part of the 2D model the increase of the S-wave velocities is stronger in the first two meters than in the deeper parts. Between 2 m and approximately 6 m depth the S-wave velocities in the middle and the right part of the 2D model are very similar to the velocities in the 1D model.

In spite of the differences in the S-wave velocity structures both models predict the recorded seismograms very well (Figure 5.15). This shows the ambiguity of the inverse problem. However, this ambiguity does not result in strong 2D artefacts in the FWI. The 2D inversion is stable in this respect. The 2D model predicts the field data better compared to the 1D model especially for the fundamental mode of the Rayleigh waves in the seismograms with large offsets. However, some parts of the higher modes of the Rayleigh waves are fitted better by the 1D model. This can be also observed in Figure 5.16 where the contribution to the total data misfit of each trace is displayed color coded. The better fit of the field data by the 2D model is not unexpected since the number of degrees of freedom in the 2D inversion is much higher than in the 1D inversion.

The differences in the S-wave velocity models are largest in the left part of the 2D model (see vertical velocity profile at x=14 m in Figure 5.13b). To investigate if these lateral variations in the 2D model are caused by the observed data, common-offset gathers are analyzed. Figure 5.17 displays common-offset gathers for an offset of 25.5 m. Figure 5.17b) displays the field data. No significant variations in the phase velocities along the profile are observable in the field data which indicates that there is no significant 2D structure. (This also holds for common-offset gathers for smaller offsets.) The common-offset gather for the field data is also compared to the corresponding common-offset gathers generated with the 1D and the 2D model. If a constant source time function is assumed all seismograms would be equal for a perfect 1D subsurface structure. The differences of the seismograms calculated with the 1D model (Figure 5.17c) are only caused by differences in the optimized source wavelet correction filters for each shot. In the

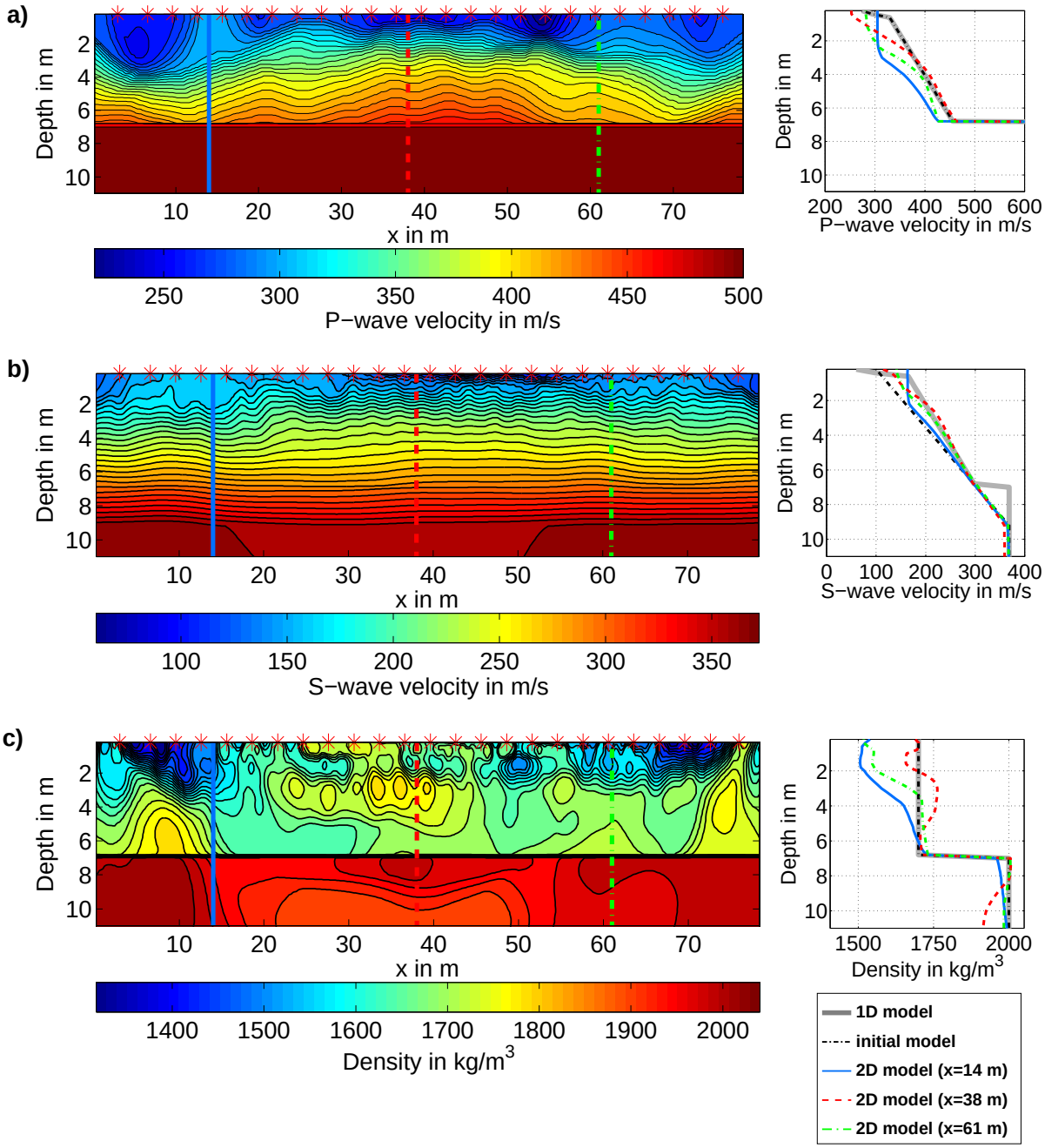


Figure 5.13: 2D subsurface model obtained by FWI. a) displays the P-wave velocity model, b) the S-wave velocity model, and c) the density model. The red stars in the 2D models on the left mark the source positions. On the right vertical profiles through the 2D model are shown in comparison to the 1D model (thick grey line) and the initial model (dash-dotted black line). The positions of these profiles are marked by the vertical lines in the 2D models on the left. Due to the high contrast in P-wave velocities between the layer and the half-space the colorbar as well as the x-coordinate are clipped in a). However, in the half-space the P-wave velocity does not vary significantly (P-wave velocity in the half-space is 1905 m/s with a maximal variation of $\pm$ 1 m/s).

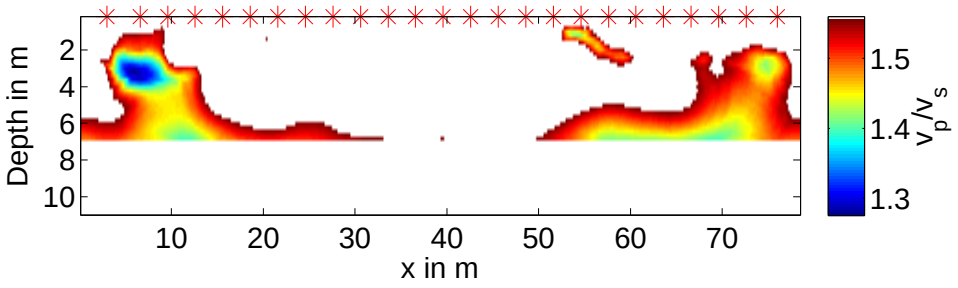


Figure 5.14: v_p/v_s ratio for the 2D subsurface model obtained by FWI. The colorbar is clipped at a ratio of 1.56. Areas with higher v_p/v_s ratios are displayed in white. The red stars mark the source positions.

field data clear differences in the seismograms are observed because of noise, 3D scattering and again differences in the source wavelets. However, the main phases of the field data are fitted well by the seismograms of both models. Only few features of the field data are fitted better by the seismograms of the 2D model like e. g. the amplitude variation of the phase at 0.14 s or the small phase at 0.22 s which is present in some traces only.

5.5 Conclusions

A successful application of 2D FWI to a field data set which was acquired on a predominantly depth-dependent structure is presented. The data misfit reduces significantly in the inversion and the obtained 2D subsurface model predicts the field data well. The resulting S-wave velocity model contains no significant 2D artefacts although a 1D structure is not enforced by regularization. For the purpose of comparison I additionally derive a 1D subsurface model by an inversion of Fourier-Bessel expansion coefficients and first arrival P-wave travel times. Although the 1D and 2D subsurface models derived by the two inversion methods differ both models are able to predict the main phases of the recorded data. However, the synthetic seismograms generated with the 2D model fit the field data better especially in some details in the recorded seismograms.

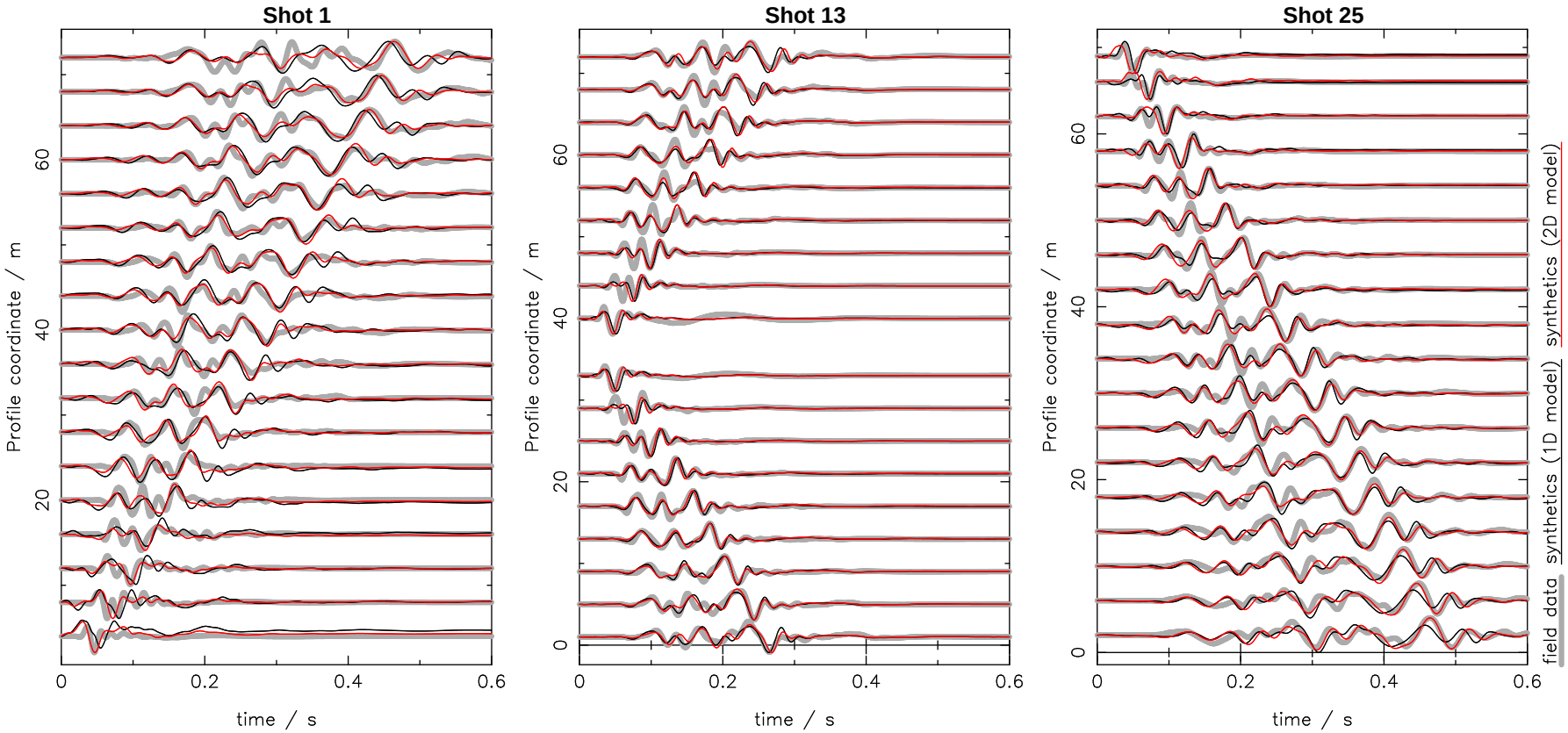


Figure 5.15: Comparison of field data (thick grey) with synthetic seismograms generated with the 1D model (black) and the 2D model (red) for shot 1 (x=3.0 m), shot 13 (x=39.6 m), and shot 25 (x=76.0 m). Displayed are vertical particle displacement seismograms. Each seismogram is normalized to its maximum amplitude. Every fourth trace is displayed. A source wavelet correction is already applied to the synthetic seismograms. For the seismograms of each shot and each subsurface model a separate source wavelet correction filter was estimated. In the optimization of the source wavelet correction filter all seismograms of a shot are used unless the clipped seismograms. A weighting is applied with $\frac{r}{\mathrm{m}}^{1.2}$.

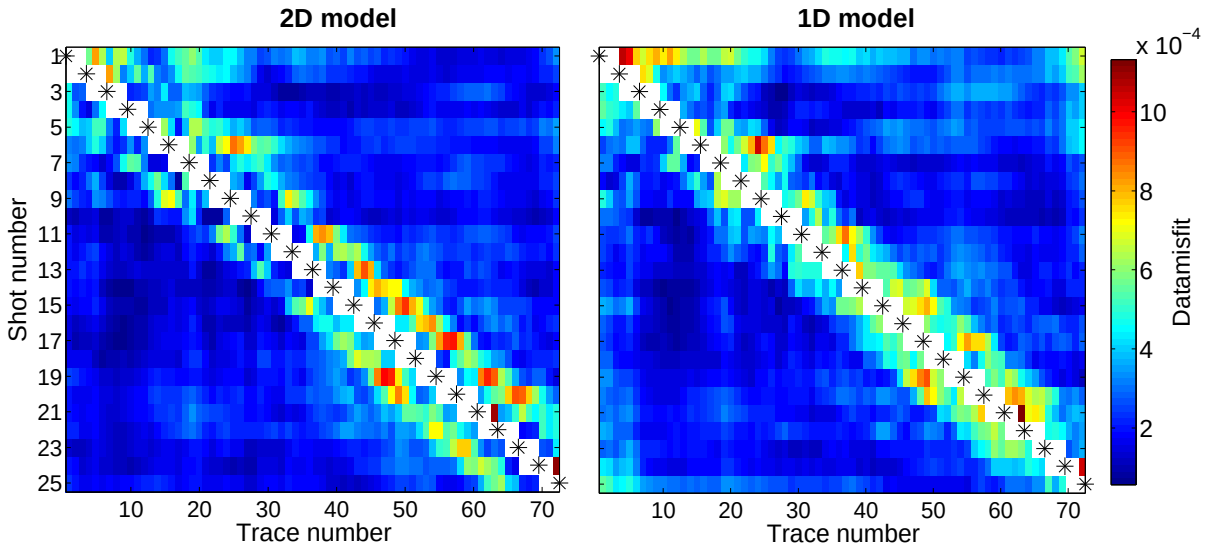


Figure 5.16: Contribution to the final data misfit of the single shots and the single receivers for the 2D model (left) and the 1D model (right). The sum of the misfit contributions shown in each figure results in the total data misfit. The black stars mark the source positions. Clipped traces which are not used in the inversion are plotted in white.

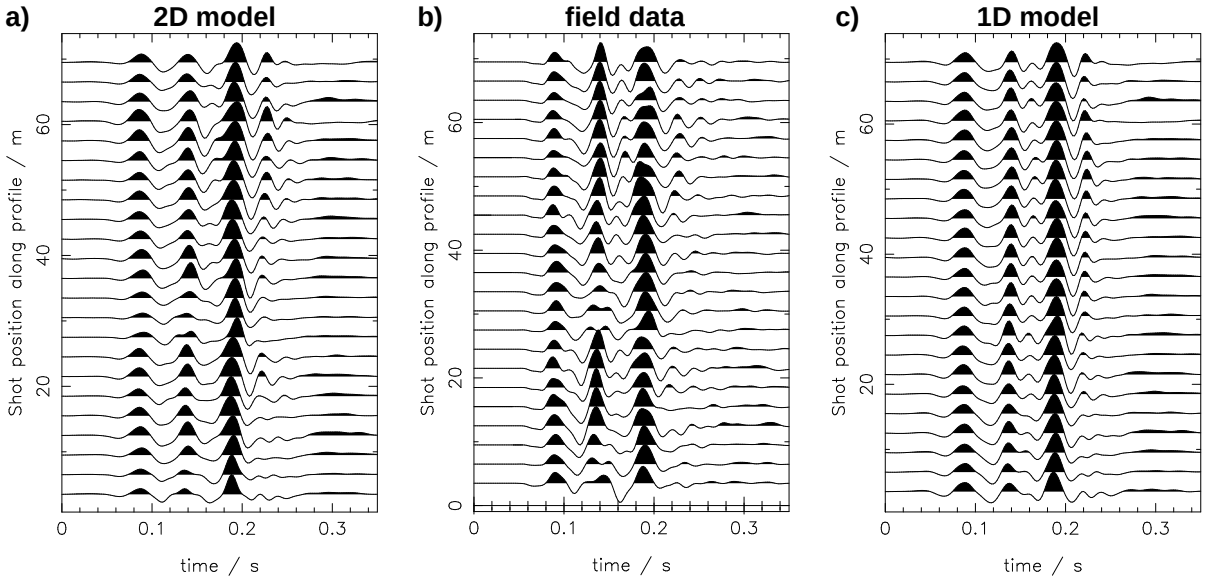


Figure 5.17: Common-offset gathers for an offset of 25.5 m. a) displays the seismograms calculated with the 2D model, b) displays the recorded seismograms, and c) displays the seismograms calculated with the 1D model. All traces are vertical displacement seismograms with a bandwidth of 5 Hz to 70 Hz. Each trace is normalized to its maximum amplitude.

Chapter 6

Summary and conclusions

Main objective of this work is the application of 2D FWI to recorded shallow seismic Rayleigh waves. Studies which are applied in preparation of such an inversion of field data are presented to investigate the significance of different effects (e. g. accuracy of the forward modeling code, effects of anelastic damping, influence of P-wave velocity model). Finally, I show a successful application of 2D FWI to field data.

Chapter 2 introduces the basics of FWI using the adjoint approach. The calculation of the gradient of the misfit function considering viscoelastic wave propagation is discussed where the rheology of the medium is described by a generalized standard linear solid. The gradients of the misfit function with respect to the unrelaxed moduli can be divided into two parts. The first part is equal to the corresponding gradient in the elastic case. The second contribution contains convolutions between the relaxation functions and the forward propagated wavefield. Therefore, the calculation of the gradients in case of viscoelastic wave propagation is computationally more expensive. An inversion for dissipative properties has most likely a high ambiguity, especially in field data applications where noise is present in the observed data and the amplitudes are also influenced by receiver coupling and differences in the transfer functions of the geophones. Thus, I have decided to invert only for elastic properties or unrelaxed moduli, respectively and not for dissipative properties. Furthermore, the gradients of the misfit for the unrelaxed moduli are approximated by their elastic counterparts in the presented studies. The additional term which appears due to the assumption of anelastic wave propagation is neglected. No difficulties due to this approximation are observed in the applied FWI tests. However, this possibly does not hold for other models or other data sets. The contribution of the neglected term should be investigated in more detail in future studies.

Furthermore, a 3D/2D transformation is discussed in Chapter 2. It is based on the Fourier-Bessel expansion of the recorded point-source wavefields. The expansion coefficients are afterwards used in an expansion with plane waves (Amundsen & Reitan, 1994). The transformation accounts for both, body and surface waves. It is shown that the described transformation is analytically exact for wavefields that are excited by a cylindrical symmetric source in a depth dependent (1D) structure. This is also confirmed by a synthetic test. The transformation assumes a 1D subsurface structure. Due to the expansion of the Fourier-Bessel expansion coefficients with plane waves the resulting transformed seismograms equal seismograms recorded on a 1D subsurface structure which means that vertical component seismograms on both sides of the

line source are identical for equal offsets and radial component seismograms just differ in sign in a Cartesian coordinate system. Effects in the point-source wavefield caused by 2D and 3D subsurface structures are shifted into a resulting pseudo 1D subsurface structure by the transformation (Schäfer et al., 2013). Although only field data sets from predominantly depth dependent subsurface structures are considered in this thesis it is undesired to remove effects in the wavefield possibly caused by 2D structures already by the 3D/2D transformation prior to a 2D FWI. We discuss this in Schäfer et al. (2013) and present in Forbriger et al. (2013) a transformation that is applicable to single traces and that is superior to the transformation using the Fourier-Bessel expansion coefficients in case of 2D structures (Schäfer et al., 2013). Therefore, this single-trace transformation is used in the presented field data applications.

The benchmark tests described in Chapter 3 confirm that the time-domain finite-difference algorithm, which is implemented in the used 2D FWI code, can accurately simulate Rayleigh waves and is therefore suited for FWI. The modeled seismograms are independent of the spatial and temporal grid spacing if a correct scaling is applied to the exciting sources (provided that significant effects due to grid dispersion are avoided). The implemented stress-velocity finite-difference scheme is based on the standard staggered grid. Therefore, the wavefield variables (components of the particle velocity and the stress tensor) are not all located at full grid points but are distributed over the grid cell. However, the benchmark tests for a pressure source in a homogeneous full-space confirm that the influence of the staggered grid can be observed only very close to the source (one-fifth of the minimum wavelength in the performed test) and is therefore negligible in an FWI of Rayleigh waves. The free surface can be implemented by either adding an air or vacuum layer, respectively, in the model or by using the image technique suggested by Levander (1988). Both implementations work accurately for vertical force sources in my tests. However, depending on the model the different location of the free surface in the staggered grid can be observed in the waveforms. The free surface is located at the first full grid row if the image technique is used. In contrast, the free surface is located between the two grid rows where the upper one has the material properties of air and the lower one has the material properties of the subsurface for an implementation of the free surface by an air layer. I observe a scaling problem in wavefields that are excited by a radial force source directly at the free surface when the image technique is applied. Further investigations for single force sources show that the observed problem is not specific for radial force sources but for single force sources in general which are excited directly at the free surface in connection with an implementation of the free surface by the image technique. By shifting the staggered grid half a grid point in vertical direction the described problem can be reproduced also for vertical force sources. Based on my studies I suggest to apply an additional scaling factor of 2 for such a configuration. A clear physical explanation why this factor must be applied is missing yet and should be further investigated. I have implemented viscoelastic forward modeling in the 2D FWI code according to the implementation suggested by Bohlen (1998). A generalized standard linear solid is used as rheological model. The results of the finite-difference scheme are successfully compared to results obtained with the reflectivity method. In the subsequent FWI tests three relaxation mechanisms are used to approximate a constant quality factor in the bandwidth of the analyzed signals.

In shallow seismic field data the effects of anelastic damping are significant. This is shown in a comparison of field data with elastically and viscoelastically simulated data (Chapter 4). Fur-

thermore, a trade-off between the source wavelet correction and anelastic effects is observed. A source wavelet correction filter can compensate a significant fraction of the residuals between viscoelastic and elastic wavefields by narrowing the bandwidth of the signals. Nevertheless, this effect is not sufficient. Viscoelastic modeling is essential to simulate the amplitude decay with offset of the field data. In a further study reconstruction tests are performed to investigate the effects of anelastic damping on the reconstruction of the S-wave velocity model by FWI. Viscoelastically forward modeled data with Q_s=Q_p=20 are used as pseudo-observed data. Dissipative properties are used as a priori known parameters in the inversion. No inversion for them is applied. Firstly, reconstruction tests are performed where the correct source wavelet of the pseudo-observed data is used as a priori known source wavelet in the inversion. These tests show that the reconstruction of the S-wave velocity model is not possible if purely elastic forward modeling is applied in the inversion. The result with elastic forward modeling improves if additionally a source wavelet correction is applied in the FWI. Nevertheless, the compensation effect of the source wavelet correction is not convincing. The reconstruction of the S-wave velocity model is superior when approximately correct quality factors are considered in the forward modeling of the inversion.

Chapter 4 also discusses reconstruction tests where the handling of the P-wave velocity model is varied and the reconstruction of the S-wave velocity model is studied. In these tests a bias of the P-wave velocity model on the reconstruction of the S-wave velocity model is observed. Strong discontinuities contained in the initial P-wave velocity model can cause artificial discontinuities in the reconstructed S-wave velocity model. Furthermore, these tests show that it is advisable to additionally invert for P-wave velocity in an FWI of shallow seismic Rayleigh waves. Strong artefacts can occur in the reconstructed S-wave velocity model if a wrong P-wave velocity model is used as fixed model in the inversion.

Finally, 2D FWI is successfully applied to a field data set which was acquired on a predominantly depth dependent structure in unconsolidated sediments (Chapter 5). The data misfit is reduced significantly during the inversion. The resulting 2D S-wave velocity model is predominantly depth dependent although this is not enforced by regularization. The initial model is only changed down to 6 m depths. The resolution in larger depth could be possibly improved by a better preconditioning of the gradients or by an amplification of low frequencies. Although the changes in the obtained S-wave velocity model are small these changes have a strong influence on the wavefields. This confirms the high potential of FWI of shallow seismic Rayleigh waves to image lateral heterogeneities in the very shallow subsurface. Similar to the observation made in many other studies about FWI the choice of the initial model is very important to avoid the problem of cycle-skipping. A 1D model is derived from the field data set by a joint inversion of first arrival P-wave travel times and Fourier-Bessel expansion coefficients for the purpose of comparison. Moderate differences between the 1D and the 2D S-wave velocity model are observed. However, the seismograms generated with both models fit the field data almost equally well which shows the high ambiguity of the inverse problem. The seismograms generated with the 2D model fit the field data better especially in small details of the recorded wavefield.

Concluding, the results presented in this thesis confirm that 2D FWI is applicable to shallow seismic Rayleigh waves. The used approach has a high potential to image lateral heterogeneities

in the very shallow S-wave velocity structure. However, further studies should investigate methods to increase the robustness of the inversion. The application of other misfit definitions which especially increase the influence of low frequencies can possibly reduce the dependence of the inversion result from the initial model and mitigate the problem of cycle-skipping. The implementation of appropriate regularization constraints is likely to ensure a better convergence to the global minimum.

Bibliography

Aki, K. & Richards, P. G., 2002. *Quantitative Seismology*. University Science Books, Sausalito, 2nd edn.

Amundsen, L. & Reitan, A., 1994. Transformation from 2-D to 3-D wave propagation for horizontally layered media, *Geophysics*, **59**(12), 1920–1926.

Aster, R. C., Borchers, B. & Thurber, C. H., 2013. *Parameter estimation and inverse problems*. Elsevier, Amsterdam, 2nd edn.

Auras, M., Bilgili, F., Bohlen, T., Butzer, S., Christen, S., Cristiano, L., Forbriger, T., Friederich, W., Giese, R., Groos, L., Igel, H., Köllner, F., Krompholz, R., Lüth, S., Mauerberger, S., Meier, T., Mosca, I., Niehoff, D., Richter, H., Schäfer, M., Schuck, A., Schumacher, F., Sigloch, K., Vormbaum, M. & Wuttke, F., 2013. Toolbox for applied seismic tomography (TOAST). In *Advanced Technologies in Earth Sciences*. Springer, in preparation.

Bachrach, R., Dvorkin, J. & Nur, A. M., 2000. Seismic velocities and Poisson's ratio of shallow unconsolidated sands, *Geophysics*, **65**(2), 559–564.

Ben-Menaḥem, A. & Singh, S. J., 1981. *Seismic waves and sources*. Springer, New York.

Blanch, J. O., Robertsson, J. O. A. & Symes, W. W., 1995. Modeling of a constant Q: Methodology and algorithm for an efficient and optimally inexpensive viscoelastic technique, *Geophysics*, **60**(1), 176–184.

Bohlen, T., 1998. *Viskoelastische FD-Modellierung seismischer Wellen zur Interpretation gemessener Seismogramme*. Dissertation, Christian-Albrechts-Universität zu Kiel.

Bohlen, T., 2002. Parallel 3-D viscoelastic finite difference seismic modeling, *Computers and Geosciences*, **28**(8), 887–899.

Bohlen, T., Kugler, S., Klein, G. & Theilen, F., 2004. 1.5D inversion of lateral variation of Scholte-wave dispersion, *Geophysics*, **69**(2), 330–344.

Bohlen, T. & Saenger, E. H., 2006. Accuracy of heterogeneous staggered-grid finite-difference modeling of Rayleigh waves, *Geophysics*, **71**(4), T109–T115.

Bozdağ, E., Trampert, J. & Tromp, J., 2011. Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements, *Geophys. J. Int.*, **185**(2), 845–870.

Bretaudeau, F., Brossier, R., Leparoux, D., Abraham, O. & Virieux, J., 2013. 2D elastic full-waveform imaging of the near-surface: application to synthetic and physical modelling data sets, *Near Surface Geophysics*, **11**(3), 307–316.

Bunks, C., Saleck, F. M., Zaleski, S. & Chavent, G., 1995. Multiscale seismic wave-form inversion, *Geophysics*, **60**(5), 1457–1473.

Butzer, S., Kurzmann, A. & Bohlen, T., 2013. 3D elastic full waveform inversion of small-scale heterogeneities in transmission geometry, *Geophysical Prospecting*, **61**(6), 1238–1251.

Carcione, J. M., 2001. *Wave fields in real media: wave propagation in anisotropic, anelastic, and porous media*. Pergamon, Amsterdam.

Carcione, J. M., Kosloff, D. & Kosloff, R., 1988. Wave propagation simulation in a linear viscoelastic medium, *Geophysical Journal*, **95**(3), 597–611.

Cerjan, C., Kosloff, D., Kosloff, R. & Reshef, M., 1985. A nonreflecting boundary condition for discrete acoustic and elastic wave equations, *Geophysics*, **50**(4), 705–708.

Charara, M., Barnes, C. & Tarantola, A., 2000. Full waveform inversion of seismic data for a viscoelastic medium. In *Methods and Applications of Inversion*, eds. P. Hansen, B. Jacobsen & K. Mosegaard, vol. 92 of *Lecture Notes in Earth Sciences*, pp. 68–81. Springer Berlin / Heidelberg. doi: 10.1007/BFb0010284.

Chen, P., Jordan, T. H. & Zhao, L., 2007. Full three-dimensional tomography: a comparison between the scattering-integral and adjoint-wavefield methods, *Geophys. J. Int.*, **170**(1), 175–181.

Choi, Y. & Alkhalifah, T., 2012. Application of multi-source waveform inversion to marine streamer data using the global correlation norm, *Geophysical Prospecting*, **60**, 748–758.

Christensen, R. M., 1971. *Theory of viscoelasticity: an introduction*. Academic Press, New York.

Emmerich, H. & Korn, M., 1987. Incorporation of attenuation into time-domain computations of seismic-wave fields, *Geophysics*, **52**(9), 1252–1264.

Ewing, W., Jardetzky, W. & Press, F., 1957. *Elastic waves in layered media*. McGraw-Hill Book Company, Inc., New York-Toronto-London.

Fichtner, A., 2011. *Full seismic waveform modelling and inversion*. Advances in geophysical and environmental mechanics and mathematics. Springer, Heidelberg, 1st edn.

Forbriger, T., 1996. *Interpretation von Oberflächenwellen in der Flachseismik*. Diplomarbeit, Institut für Geophysik, Universität Stuttgart. http://digbib.ubka.uni-karlsruhe.de/volltexte/1000007825 (last access: September 2013).

Forbriger, T., 2001. *Inversion flachseismischer Wellenfeldspektren*. Dissertation, Institut für Geophysik, Universität Stuttgart. http://digbib.ubka.uni-karlsruhe.de/volltexte/1000007822 (last access: September 2013).

Forbriger, T., 2003a. Inversion of shallow-seismic wavefields. Part 1: Wavefield transformation, *Geophys. J. Int.*, **153**, 719–734.

Forbriger, T., 2003b. Inversion of shallow-seismic wavefields. Part 2: Inferring subsurface properties from wavefield transforms, *Geophys. J. Int.*, **153**, 735–752.

Forbriger, T., 2003c. Notizen zur Berechnung synthetischer Seismogramme mit der Reflektivitätsmethode. Skriptum zum SZGRF-Kurs “Berechnung von synthetischen Seismogrammen und seismischen Strahlen mit praktischen Übungen”, Geophysikalisches Institut der Universität Karlsruhe am Geowissenschaftlichen Gemeinschaftsobservatorium der Universitäten Karlsruhe und Stuttgart. http://digbib.ubka.uni-karlsruhe.de/volltexte/1000007851 (last access: September 2013).

Forbriger, T., 2004. Dynamics of the hammer blow. In *Symposium in Memoriam of Prof. Gerhard Müller*, ed. J. Schweitzer, vol. I/2004, pp. 93–97. Deutsche Geophysikalische Gesellschaft. http://digbib.ubka.uni-karlsruhe.de/volltexte/1000007818 (last access: September 2013).

Forbriger, T., 2009. Die Berechnung partieller Ableitungen über einen störungstheoretischen Ansatz. Unpublished notes. Black Forest Observatory (BFO), Schiltach.

Forbriger, T., Groos, L. & Schäfer, M., 2013. Line-source simulation for shallow-seismic data. Part 1: Theoretical background, *Geophys. J. Int.,* submitted.

Friederich, W., 1995. Theorie seismischer Wellen. Skriptum zur Vorlesung, Universität Stuttgart & ETH Zürich. http://www.geophys.uni-stuttgart.de/oldwww/skripten/tsw/index.html (last access: September 2013).

Fuchs, K. & Müller, G., 1971. Computation of synthetic seismograms with the reflectivity method and comparison with observations, *Geophys. J. R. astr. Soc.*, **23**(4), 417–433.

Gradshteyn, I. S. & Ryzhik, I. M., 1980. *Table of integrals, series and products*. Academic Press, Orlando, Florida.

Graves, R. W., 1996. Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences, *Bull. Seism. Soc. Am.*, **86**(4), 1091–1106.

Groos, L., Schäfer, M., Forbriger, T. & Bohlen, T., 2013. 2D elastic full-waveform inversion of shallow seismic Rayleigh waves in the presence of attenuation, in preparation.

Hatherly, P., 1986. Attenuation measurements on shallow seismic refraction data, *Geophysics*, **51**(2), 250–254.

Hicks, G. J. & Pratt, R. G., 2001. Reflection waveform inversion using local descent methods: Estimating attenuation and velocity over a gas-sand deposit, *Geophysics*, **66**(2), 598–612.

Hüttner, R., Konrad, H.-J. & Zitzmann, A., 1986. *Geologische Übersichtskarte 1:200000, Blatt CC7110 Mannheim*. Bundesanstalt für Geowissenschaften und Rohstoffe in Zusammenarbeit mit den Geologischen Landesämtern der Bundesrepublik Deutschland und benachbarter Staaten.

Köhn, D., 2011. *Time domain 2D elastic full waveform tomography*. Ph.D. thesis, Christian-Albrechts-Universität zu Kiel.

Köhn, D., De Nil, D., Kurzmann, A., Groos, L., Schäfer, M. & Heider, S., 2013. *DENISE - User manual*. Christian-Albrechts-Universität zu Kiel und Karlsruher Institut für Technologie, http://www.gpi.kit.edu/downloads/manual_DENISE.pdf (last access: August 2013).

Köhn, D., De Nil, D., Kurzmann, A., Przebindowska, A. & Bohlen, T., 2012. On the influence of model parametrization in elastic full waveform tomography, *Geophys. J. Int.*, **191**(1), 325–345.

Komatitsch, D. & Martin, R., 2007. An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation, *Geophysics*, **72**(5, S), SM155–SM167.

Levander, A. R., 1988. Fourth order finite-difference P-SV seismograms, *Geophysics*, **53**(11), 1425–1436.

Liu, H.-P., Anderson, D. L. & Kanamori, H., 1976. Velocity dispersion due to anelasticity; implications for seismology and mantle composition, *Geophys. J. R. astr. Soc.*, **47**(1), 41–58.

Malinowski, M., Operto, S. & Ribodetti, A., 2011. High-resolution seismic attenuation imaging from wide-aperture onshore data by visco-acoustic frequency-domain full-waveform inversion, *Geophys. J. Int.*, **186**(3), 1179–1204.

Moczo, P., 1989. Finite-difference technique for SH-waves in 2-D media using irregular grids application to the seismic response problem, *Geophys. J. Int.*, **99**(2), 321–329.

Moczo, P., Kristek, J., Vavryčuk, V., Archuleta, R. J. & Halada, L., 2002. 3D heterogeneous staggered-grid finite-difference modeling of seismic motion with volume harmonic and arithmetic averaging of elastic moduli and densities, *Bull. Seism. Soc. Am.*, **92**(8), 3042–3066.

Mora, P., 1987. Nonlinear two-dimensional elastic inversion of multioffset seismic data, *Geophysics*, **52**(9), 1211–1228.

Mosca, I., Bilgili, F., Meier, T., Forbriger, T. & Sigloch, K., 2013. A Monte Carlo approach for the inference of structure on ultra-sonic and shallow-seismic scales. In *Tagungsband der 73. Jahrestagung der Deutschen Geophysikalischen Gesellschaft*, p. 119. http://www.dgg-online.de/tagungen/DGG-2013-komprimiert.pdf (last access: October 2013).

Müller, G., 1985. The reflectivity method: A tutorial, *J. Geophys.*, **58**, 153–174.

Oconnell, R. & Budiansky, B., 1978. Measures of dissipation in viscoelastic media, *Geophys. Res. Lett.*, **5**(1), 5–8.

O'Neill, A., Campbell, T. & Matsuoka, T., 2008. Lateral resolution and lithological interpretation of surface-wave profiling, *The Leading Edge*, **27**(11), 1550–1563.

Pratt, R. G., 1990. Inverse theory applied to multi-source cross-hole tomography, Part 2: Elastic wave-equation method, *Geophysical Prospecting*, **38**(3), 311–329.

Pratt, R. G., 1999. Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model, *Geophysics*, **64**(3), 888–901.

Pratt, R. G. & Worthington, M. H., 1990. Inverse theory applied to multi-source cross-hole tomography, Part 1: Acoustic wave-equation method, *Geophysical Prospecting*, **38**(3), 287–310.

Robertsson, J., 1996. A numerical free-surface condition for elastic/viscoelastic finite-difference modeling in the presence of topography, *Geophysics*, **61**(6), 1921–1934.

Robertsson, J. O. A., Blanch, J. O. & Symes, W. W., 1994. Viscoelastic finite-difference modeling, *Geophysics*, **59**(9), 1444–1456.

Robertsson, J. O. A., Levander, A., Symes, W. W. & Holliger, K., 1995. A comparative study of free-surface boundary conditions for finite-difference simulation of elastic/viscoelastic wave propagation, *65th Annual International Meeting, SEG, Expanded Abstracts*, pp. 1277–1280.

Romdhane, G., Grandjean, G., Brossier, R., Rejiba, F., Operto, S. & Virieux, J., 2011. Shallow-structure characterization by 2D elastic full-waveform inversion, *Geophysics*, **76**(3), R81–R93.

Schäfer, M., Groos, L., Forbriger, T. & Bohlen, T., 2013. Line-source simulation for shallow-seismic data. Part 2: Full-waveform inversion - a 2D case study, *Geophys. J. Int.,* submitted.

Schön, J. H., 1996. *Physical properties of rocks: fundamentals and principles of petrophysics*. Handbook of geophysical exploration: Section I, Seismic exploration: vol. 18. Pergamon, Oxford, 1st edn.

Schumacher, F. & Friederich, W., 2013. Full waveform inversion using waveform sensitivity kernels. In *Tagungsband der 73. Jahrestagung der Deutschen Geophysikalischen Gesellschaft*, p. 90. http://www.dgg-online.de/tagungen/DGG-2013-komprimiert.pdf (last access: October 2013).

Sheen, D.-H., Tuncay, K., Baag, C.-E. & Ortoleva, P. J., 2006. Time domain Gauss-Newton seismic waveform inversion in elastic media, *Geophys. J. Int.*, **167**, 1373–1384.

Sirgue, L. & Pratt, R. G., 2004. Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies, *Geophysics*, **69**(1), 231–248.

Socco, L. V., Foti, S. & Boiero, D., 2010. Surface-wave analysis for building near-surface velocity models - Established approaches and new perspectives, *Geophysics*, **75**(5), A83–A102.

Stümpel, H., Kähler, S., Meissner, R. & Milkereit, B., 1984. The use of seismic shear waves and compressional waves for lithological problems of shallow sediments, *Geophysical Prospecting*, **32**(4), 662–675.

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, **49**(8), 1259–1266.

Tarantola, A., 1988. Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation, *Pure Appl. Geophys.*, **128**(1-2), 365–399.

Tran, K. T. & McVay, M., 2012. Site characterization using Gauss-Newton inversion of 2-D full seismic waveform in the time domain, *Soil Dyn. Earthq. Eng.*, **43**, 16–24.

Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.*, **160**(1), 195–216.

Ungerer, J., 1990. *Berechnung von Nahfeldseismogrammen mit der Reflektivitätsmethode*. Diplomarbeit, Institut für Geophysik, Universität Stuttgart, Germany.

Virieux, J., 1986. P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method, *Geophysics*, **51**(4), 889–901.

Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, *Geophysics*, **74**(6, S), WCC1–WCC26.

Wapenaar, C. P. A., Verschnuur, D. J. & Herrmann, P., 1992. Amplitude processing of single and multicomponent seismic data, *Geophysics*, **57**(9), 1178–1188.

Wathelet, M., Jongmans, D. & Ohrnberger, M., 2004. Surface-wave inversion using a direct search algorithm and its application to ambient vibration measurements, *Near Surface Geophysics*, **2**(4), 211–221.

Zahradník, J. & Priolo, E., 1995. Heterogeneous formulations of elastodynamic equations and finite-difference schemes, *Geophys. J. Int.*, **120**(3), 663–676.

List of Figures

List of Tables

Appendix A

Generalized standard linear solid and τ method

A.1 Stress strain relationship

The stress strain relationship for a viscoelastic material is given by (e. g. Christensen, 1971)

$$\sigma_{jk}(\vec{x},t) = \Psi_{jkml}(\vec{x},t) * \varepsilon_{ml}(\vec{x},t) \tag{A.1}$$

with the components of the stress tensor σ_{jk} at a point in space $\vec{x}$ and time t, the components of the strain tensor ε_{ml} and the tensor Ψ_{jkml} of rank four which contains the rate of relaxation functions (the time derivatives of the relaxation functions).

In the frequency domain the convolution simplifies to a multiplication and the stress strain relationship can be written as

$$\bar{\sigma}_{jk}(\vec{x},\omega) = \bar{M}_{jkml}(\vec{x},\omega)\ \bar{\varepsilon}_{ml}(\vec{x},\omega), \tag{A.2}$$

with the Fourier-transformed components of the stress tensor $\bar{\sigma}_{jk}$ and the strain tensor $\bar{\varepsilon}_{ml}$ and the complex moduli $\bar{M}_{jkml}$, which are the Fourier transforms of the single elements of the tensor Ψ_{jkml}.

A.2 Generalized standard linear solid

A.2.1 Relaxation functions in time domain

Generally, a relaxation function describing a generalized standard linear solid (GSLS) with L relaxation mechanisms can be expressed by

$$\psi(t) = M_r \left[1 + \sum_{l=1}^{L} \left(\frac{\tau_{\varepsilon l}}{\tau_{\sigma l}} - 1 \right) e^{-t/\tau_{\sigma l}} \right] H(t) \tag{A.3}$$

where M_r is the relaxed modulus. M_r can be either the P-wave modulus π_r or the shear modulus μ_r (compare with Equation 2.10).

Limit $t \to \infty$
Considering Equation (A.3) for $t \to \infty$, one obtains

$$\lim_{t\to\infty} \psi(t) = M_r. \tag{A.4}$$

Limit $t \to 0$
In contrast, in the limit $t \to 0$ one obtains

$$\lim_{t\to 0} \psi(t) = M_r \left[1 + \sum_{l=1}^{L} \left(\frac{\tau_{\varepsilon l}}{\tau_{\sigma l}} - 1\right)\right] \tag{A.5a}$$

$$= M_u \tag{A.5b}$$

where M_u is the unrelaxed modulus.

A.2.2 Complex moduli in frequency domain

To derive the complex moduli the Fourier transforms of the rate of relaxation function $\dot{\psi}(t)$ must be calculated. According to Carcione et al. (1988) this leads to

$$\bar{M}(\omega) = M_r \left[1 - L + \sum_{l=1}^{l=L} \frac{1 + i\omega\tau_{\varepsilon l}}{1 + i\omega\tau_{\sigma l}}\right]. \tag{A.6}$$

Limit $\omega \to \infty$
In the limit $\omega \to \infty$ the complex modulus

$$\lim_{\omega\to\infty} \bar{M}(\omega) = \lim_{\omega\to\infty} M_r \left[1 - L + \sum_{l=1}^{L} \frac{\frac{1}{i\omega} + \tau_{\varepsilon l}}{\frac{1}{i\omega} + \tau_{\sigma l}}\right] \tag{A.7a}$$

$$= M_r \left[1 - L + \sum_{l=1}^{L} \frac{\tau_{\varepsilon l}}{\tau_{\sigma l}}\right] \tag{A.7b}$$

$$= M_r \left[1 + \sum_{l=1}^{L} \left(\frac{\tau_{\varepsilon l}}{\tau_{\sigma l}} - 1\right)\right] \tag{A.7c}$$

$$= M_u \tag{A.7d}$$

equals the unrelaxed modulus.

Limit $\omega \to 0$
According to Equation (A.6) one obtains

$$\lim_{\omega\to 0} \bar{M}(\omega) = M_r. \tag{A.8}$$

As a reasonable approximation we assume that materials present in the subsurface structure behave elastically for both, very low and very high frequencies. This behaviour is described correctly by a GSLS. Carcione et al. (1988) choose the behaviour in the limit of low frequencies as elastic behaviour. Related to the GSLS an elastic behaviour is obtained when the dashpot

vanishes that means in the limit of $\tau_{\varepsilon l} \to 0$ and $\tau_{\sigma l} \to 0$ (Carcione et al., 1988). This limit is related to the limit $\omega \to 0$ of Equation (A.6). Therefore, the relaxed modulus and the elastic modulus coincide. However, as also mentioned by Carcione et al. (1988) in practice we are not limited to simple mechanical models to describe the rheology of different materials. Therefore, also the behaviour for very high frequencies could be chosen as elastic properties. This is e. g. discussed by Ben-Menahem & Singh (1981, chapter 10.3).

A.2.3 Quality factor Q

According to Oconnell & Budiansky (1978) the quality factor Q is defined by

$$Q = \frac{\Re\left[\bar{M}(\omega)\right]}{\Im\left[\bar{M}(\omega)\right]} \tag{A.9}$$

with the complex modulus $\bar{M}$. $\Re[\bar{M}]$ denotes the real part and $\Im[\bar{M}]$ denotes the imaginary part. For a GSLS the quality factor is therefore given by (e. g. Blanch et al., 1995)

$$Q(\omega) = \frac{1 - L + \sum_{l=1}^{L} \frac{1+\omega^2 \tau_{\sigma l} \tau_{\varepsilon l}}{1+\omega^2 \tau_{\sigma l}^2}}{\sum_{l=1}^{L} \frac{\omega(\tau_{\varepsilon l} - \tau_{\sigma l})}{1+\omega^2 \tau_{\sigma l}^2}}. \tag{A.10}$$

A.2.4 τ method and approximation of frequency independent quality factors

Blanch et al. (1995) introduced the τ method to reduce the number of variables used for describing a GSLS. Thus, memory consumption as well as computing time during viscoelastic forward simulations are reduced.

The τ method is based on the observation that anelastic damping caused by a standard linear solid (SLS) can be described by a dimensionless and frequency independent variable τ (Blanch et al., 1995). It is defined by

$$\tau = \frac{\tau_{\varepsilon l}}{\tau_{\sigma l}} - 1. \tag{A.11}$$

By inserting this definition into Equation (A.10) we obtain the quality factor

$$Q(\omega) = \frac{1 + \sum_{l=1}^{L} \frac{\omega^2 \tau_{\sigma l}^2}{1+\omega^2 \tau_{\sigma l}^2} \tau}{\sum_{l=1}^{L} \frac{\omega \tau_{\sigma l}}{1+\omega^2 \tau_{\sigma l}^2} \tau} \tag{A.12}$$

for a GSLS or

$$Q(\omega) = \frac{1 + \omega^2 \tau_\sigma^2 (1+\tau)}{\omega \tau_\sigma \tau}. \tag{A.13}$$

for a single SLS (Blanch et al., 1995). Figure A.1 displays the quality factor for a single SLS for varying values of τ and τ_σ or f_σ, respectively, with $\tau_\sigma = 1/(2\pi f_\sigma)$. The variation of τ only

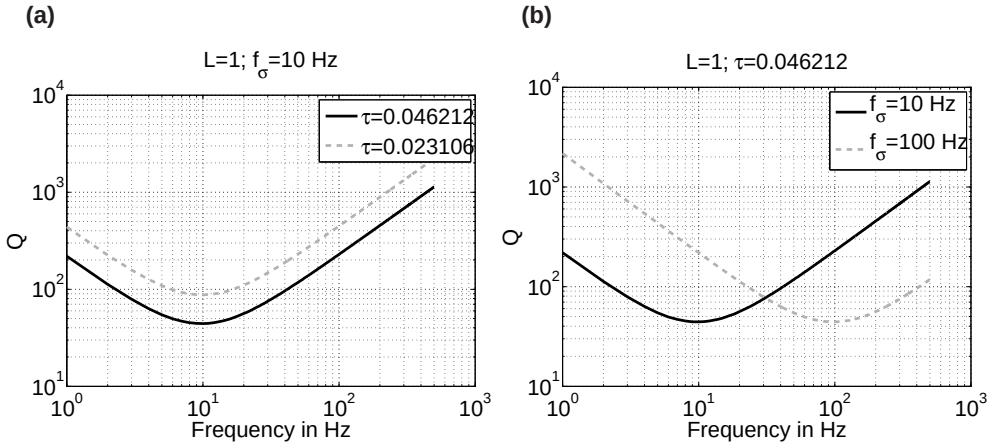


Figure A.1: Comparison of Q models for a single SLS calculated with Equation (A.13) according to Blanch et al. (1995). a) displays Q for varying τ values and b) for varying relaxation frequencies f_σ or relaxation times τ_σ, respectively.

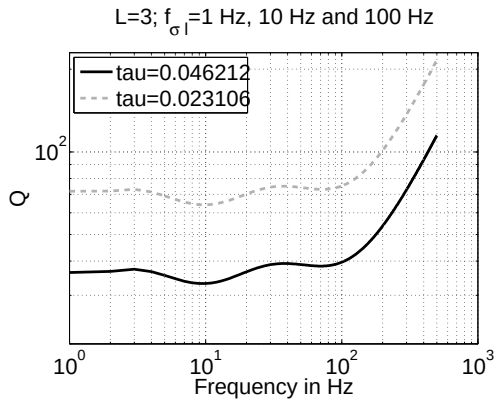


Figure A.2: Frequency dependent Q value for a GSLS with three relaxation mechanisms with frequencies at $f_{\sigma l} = 1\,\text{Hz}$, $10\,\text{Hz}$ and $100\,\text{Hz}$. Equation (A.12) was used for the calculation with two different τ values.

shifts the function along the Q axis or ordinate whereas a variation of f_σ shifts the function along the frequency axis. By adding several SLSs with a constant τ to a GSLS an almost constant Q value can be approximated in a desired frequency interval. An example is displayed in Figure A.2 where I used three relaxation mechanisms with relaxation frequencies of $f_{\sigma l} = 1\,\text{Hz}$, $10\,\text{Hz}$ and $100\,\text{Hz}$ and calculated the quality factors for two different values of τ. Again it can be observed that the relaxation frequencies determine the interval where the quality factor is approximately constant and τ determines the value of the frequency independent quality factor.

Appendix B

Auxiliary calculations for the derivation of the gradient of the misfit function

B.1 Partial derivative of synthetic seismograms with respect to the model parameters

The partial derivatives of the synthetic seismograms with respect to the model parameters are obtained by Equations (2.17) and (2.24). The partial derivatives can be expressed by

$$
\begin{aligned}
&\left.\frac{\partial}{\partial m_k}\right|_{\vec{m}=\vec{m}_0} s_j(\vec{x}_r,t,\vec{m}) = \left.\frac{\partial}{\partial \delta m_k}\right|_{\delta\vec{m}=0} s_j(\vec{x}_r,t,\vec{m}) \\
&= \int_V \int_{-\infty}^{\infty} \left[\left.\frac{\partial}{\partial \delta m_k}\right|_{\delta\vec{m}=0} G_{jn}(\vec{x}_r,t-t';\vec{x}',0) \right] f_n(\vec{x}',t')\,\mathrm{d}t'\,\mathrm{d}^3\vec{x}' \\
&= \int_V \int_{-\infty}^{\infty} \left[-\int_V \int_{-\infty}^{\infty} G^{(0)}_{jm}(\vec{x}_r,t-t'';\vec{x}'',0) \left.\frac{\partial}{\partial \delta m_k}\right|_{\delta\vec{m}=0} \hat{L}^{(1)}_{mi}(\vec{x}'',\delta\vec{m}) \right. \\
&\qquad \left. G^{(0)}_{in}(\vec{x}'',t''-t';\vec{x}',0)\,\mathrm{d}t''\,\mathrm{d}^3\vec{x}'' \right] f_n(\vec{x}',t')\,\mathrm{d}t'\,\mathrm{d}^3\vec{x}' \\
&= -\int_V \int_{-\infty}^{\infty} G^{(0)}_{jm}(\vec{x}_r,t-t'';\vec{x}'',0) \left.\frac{\partial}{\partial \delta m_k}\right|_{\delta\vec{m}=0} \hat{L}^{(1)}_{mi}(\vec{x}'',\delta\vec{m}) \\
&\qquad \underbrace{\int_V \int_{-\infty}^{\infty} G^{(0)}_{in}(\vec{x}'',t''-t';\vec{x}',0) f_n(\vec{x}',t')\,\mathrm{d}t'\,\mathrm{d}^3\vec{x}'}_{s_{0_i}(\vec{x}'',t'')}\,\mathrm{d}t''\,\mathrm{d}^3\vec{x}'' \\
&= -\int_V \int_{-\infty}^{\infty} G^{(0)}_{jm}(\vec{x}_r,t-t'';\vec{x}'',0) \left.\frac{\partial}{\partial \delta m_k}\right|_{\delta\vec{m}=0} \hat{L}^{(1)}_{mi}(\vec{x}'',\delta\vec{m}) s_{0_i}(\vec{x}'',t'')\,\mathrm{d}t''\,\mathrm{d}^3\vec{x}''.
\end{aligned}
\tag{B.1}
$$

The limits of the integral over time can be adjusted because the synthetic wavefield $s_{0_i}(\vec{x}'',t'')$ vanishes for $t''<0$ due to the initial conditions (Equation 2.4) and the causality of the Green's function. This leads to

$$
\left.\frac{\partial}{\partial m_k}\right|_{\vec{m}=\vec{m}_0} s_j(\vec{x}_r,t) = -\int_V \int_0^t G^{(0)}_{jm}(\vec{x}_r,t-t'';\vec{x}'',0) \left.\frac{\partial}{\partial \delta m_k}\right|_{\delta\vec{m}=0} \hat{L}^{(1)}_{mi}(\vec{x}'',\delta\vec{m}) s_{0_i}(\vec{x}'',t'')\,\mathrm{d}t''\,\mathrm{d}^3\vec{x}''.
\tag{B.2}
$$

B.2 Gradient of the misfit function

Inserting the perturbation of the differential operator in Equation (2.38) results in

$$\begin{aligned}
\frac{\partial E}{\partial m_k} = & \\
& -\int_V \int_{t''=0}^{T} s_m^\dagger(\vec{x}'', T-t'') \frac{\partial}{\partial \delta m_k} \left(\delta\rho(\vec{x}'') \delta_{mi} \frac{\partial^2}{\partial t''^2} - \frac{\partial}{\partial x_j''} \delta\Psi_{jmni}(\vec{x}'', t'') * \frac{\partial}{\partial x_n''} \right) s_{0_i}(\vec{x}'', t'') \, \mathrm{d}t'' \, \mathrm{d}^3\vec{x}'' \\
& = -\int_V \int_{t''=0}^{T} s_m^\dagger(\vec{x}'', T-t'') \left(\frac{\partial \delta\rho(\vec{x}'')}{\partial \delta m_k} \delta_{mi} \frac{\partial^2}{\partial t''^2} \right) s_{0_i}(\vec{x}'', t'') \\
& \qquad - s_m^\dagger(\vec{x}'', T-t'') \left(\frac{\partial}{\partial x_j''} \frac{\partial \delta\Psi_{jmni}(\vec{x}'', t'')}{\partial \delta m_k} * \frac{\partial}{\partial x_n''} \right) s_{0_i}(\vec{x}'', t'') \, \mathrm{d}t'' \, \mathrm{d}^3\vec{x}''
\end{aligned} \tag{B.3}$$

where $\delta\rho$ and $\delta\Psi$ are the perturbations of the model parameters.

Integration by parts and application of Gauss's theorem (divergence theorem) results in

$$\begin{aligned}
\frac{\partial E}{\partial m_k} = & \\
& -\int_V \int_{t''=0}^{T} \frac{\partial}{\partial t''} \left(s_m^\dagger(\vec{x}'', T-t'') \frac{\partial \delta\rho(\vec{x}'')}{\partial \delta m_k} \delta_{mi} \frac{\partial s_{0_i}(\vec{x}'', t'')}{\partial t''} \right) \\
& \qquad - \frac{\partial s_m^\dagger(\vec{x}'', T-t'')}{\partial t''} \frac{\partial \delta\rho(\vec{x}'')}{\partial \delta m_k} \delta_{mi} \frac{\partial s_{0_i}(\vec{x}'', t'')}{\partial t''} \\
& \qquad\qquad - \frac{\partial}{\partial x_j''} \left(s_m^\dagger(\vec{x}'', T-t'') \frac{\partial \delta\Psi_{jmni}(\vec{x}'', t'')}{\partial \delta m_k} * \frac{\partial s_{0_i}(\vec{x}'', t'')}{\partial x_n''} \right) \\
& \qquad\qquad\qquad + \frac{\partial s_m^\dagger(\vec{x}'', T-t'')}{\partial x_j''} \frac{\partial \delta\Psi_{jmni}(\vec{x}'', t'')}{\partial \delta m_k} * \frac{\partial s_{0_i}(\vec{x}'', t'')}{\partial x_n''} \, \mathrm{d}t'' \, \mathrm{d}^3\vec{x}'' \\
& = \int_V \int_{t''=0}^{T} \frac{\partial s_m^\dagger(\vec{x}'', T-t'')}{\partial t''} \frac{\partial \delta\rho(\vec{x}'')}{\partial \delta m_k} \delta_{mi} \frac{\partial s_{0_i}(\vec{x}'', t'')}{\partial t''} \\
& \qquad\qquad - \frac{\partial s_m^\dagger(\vec{x}'', T-t'')}{\partial x_j''} \frac{\partial \delta\Psi_{jmni}(\vec{x}'', t'')}{\partial \delta m_k} * \frac{\partial s_{0_i}(\vec{x}'', t'')}{\partial x_n''} \, \mathrm{d}t'' \, \mathrm{d}^3\vec{x}'' \\
& \quad \underbrace{- \int_V \left[s_m^\dagger(\vec{x}'', T-t'') \frac{\partial \delta\rho(\vec{x}'')}{\partial \delta m_k} \delta_{mi} \frac{\partial s_{0_i}(\vec{x}'', t'')}{\partial t''} \right]_{t''=0}^{T} \mathrm{d}^3\vec{x}''}_{(1)} \\
& \quad + \underbrace{\int_S \int_{t''=0}^{T} n_j(\vec{\xi}) s_m^\dagger(\vec{\xi}, T-t'') \frac{\partial \delta\Psi_{jmni}(\vec{\xi}, t'')}{\partial \delta m_k} * \frac{\partial s_{0_i}(\vec{\xi}, t'')}{\partial x_n''} \, \mathrm{d}t'' \, \mathrm{d}S}_{(2)}
\end{aligned} \tag{B.4}$$

with $\vec{\xi}$ being a point on the surface S of the integration volume V. $n_j(\vec{\xi})$ is the outward pointing unit normal vector on the surface S.

Integral (1) is zero due to the initial conditions (Equation 2.4) $\frac{\partial s_{0_i}}{\partial t''} = 0$ for $t'' = 0$ and due to the final conditions (Equation 2.37) $s_m^\dagger = 0$ for $t'' = T$. Furthermore, I assume that the size of the volume V is large enough so that the wavefields at the surface S are almost zero. Then, the contribution of integral (2) can be neglected.

B.3 Expression of gradients in a stress-velocity scheme

Using the stress strain relationship the spatial derivatives of the displacement field $\vec{s}_0 = (s_{0_1}, s_{0_2})^T$ and the adjoint field $\vec{s}^\dagger = (s_1^\dagger, s_2^\dagger)^T$ in the gradients for the Lamé parameters λ and μ can be expressed in terms of stresses σ_{ij} for the forward simulated wavefield and in terms of wavefield variables $\sigma_{ij}^\dagger$ for the adjoint wavefields. For the adjoint field the physical unit of the wavefield variables $\sigma_{ij}^\dagger$ depends on the physical unit of the adjoint sources or the misfit function, respectively, and is $\frac{\mathrm{kg \cdot s \cdot}[g]}{\mathrm{m}^3}$.

In general the stresses in the 2D case (2D subsurface structure and assumption of line sources or cylindrical waves) can be expressed by

$$\sigma_{11} = \lambda \left(\frac{\partial s_{0_1}}{\partial x_1} + \frac{\partial s_{0_2}}{\partial x_2} \right) + 2\mu \frac{\partial s_{0_1}}{\partial x_1} \tag{B.5a}$$

$$\sigma_{22} = \lambda \left(\frac{\partial s_{0_1}}{\partial x_1} + \frac{\partial s_{0_2}}{\partial x_2} \right) + 2\mu \frac{\partial s_{0_2}}{\partial x_2} \tag{B.5b}$$

$$\sigma_{12} = \mu \left(\frac{\partial s_{0_1}}{\partial x_2} + \frac{\partial s_{0_2}}{\partial x_1} \right). \tag{B.5c}$$

Summation and subtraction of Equations (B.5a) and (B.5b) results in

$$\frac{\partial s_{0_1}}{\partial x_1} + \frac{\partial s_{0_2}}{\partial x_2} = \frac{\sigma_{11} + \sigma_{22}}{2(\lambda + \mu)} \tag{B.6a}$$

$$\frac{\partial s_{0_1}}{\partial x_1} - \frac{\partial s_{0_2}}{\partial x_2} = \frac{\sigma_{11} - \sigma_{22}}{2\mu}. \tag{B.6b}$$

Repeated summation and subtraction of Equations (B.6a) and (B.6b) provides the spatial derivatives of the displacement field

$$\frac{\partial s_{0_1}}{\partial x_1} = \frac{1}{2} \left(\frac{\sigma_{11} + \sigma_{22}}{2(\lambda + \mu)} + \frac{\sigma_{11} - \sigma_{22}}{2\mu} \right) \tag{B.7a}$$

$$\frac{\partial s_{0_2}}{\partial x_2} = \frac{1}{2} \left(\frac{\sigma_{11} + \sigma_{22}}{2(\lambda + \mu)} - \frac{\sigma_{11} - \sigma_{22}}{2\mu} \right). \tag{B.7b}$$

Inserting in Equations (2.42) and (2.43) the gradients with respect to λ and μ are

$$\frac{\partial E}{\partial \lambda_k} = -2 \int_0^T \left(\frac{\sigma_{11}^\dagger + \sigma_{22}^\dagger}{2\lambda + 2\mu} \cdot \frac{\sigma_{11} + \sigma_{22}}{2\lambda + 2\mu} \right) \mathrm{d}t'' \Delta x''^3 \tag{B.8}$$

$$\frac{\partial E}{\partial \mu_k} = -2 \int_0^T \left[\frac{\sigma_{12}^\dagger}{\mu} \cdot \frac{\sigma_{12}}{\mu} + \left(\frac{(\sigma_{11}^\dagger + \sigma_{22}^\dagger)\,(\sigma_{11} + \sigma_{22})}{4(\lambda+\mu)^2} + \frac{(\sigma_{11}^\dagger - \sigma_{22}^\dagger)\,(\sigma_{11} - \sigma_{22})}{4\mu^2} \right) \right] \mathrm{d}t'' \Delta x''^3 \tag{B.9}$$

where $\sigma_{ij} = \sigma_{ij}(\vec{x}'', t'')$ and $\sigma_{ij}^\dagger = \sigma_{ij}^\dagger(\vec{x}'', T - t'')$.

B.4 Derivative of the misfit function for L2 norm of normalized wavefields

For convenience I assume only one receiver and only one component. Equation (2.57) yields

$$E = \frac{2}{\frac{1}{T}} - 2 \int_0^T \frac{s(\vec{x}, t, \vec{m}) \cdot d(\vec{x}, t)}{S^{1/2} D^{1/2}}\, \mathrm{d}t \tag{B.10}$$

where the abbreviations

$$S = \frac{1}{T} \int_0^T s^2(\vec{x}, t'', \vec{m})\, \mathrm{d}t'' \qquad \text{and} \qquad D = \frac{1}{T} \int_0^T d^2(\vec{x}, t'', \vec{m})\, \mathrm{d}t'' \tag{B.11}$$

are used. The derivative with respect to the model parameters is

$$\frac{\partial E}{\partial m_k} = \frac{\partial}{\partial m_k} \frac{-2 \int_0^T s(\vec{x}, t, \vec{m}) d(\vec{x}, t)\, \mathrm{d}t}{S^{1/2} D^{1/2}} \tag{B.12a}$$

$$= \frac{-2}{S\,D} \left[S^{1/2}\, D^{1/2} \int_0^T d(\vec{x}, t) \frac{\partial s(\vec{x}, t, \vec{m})}{\partial m_k}\, \mathrm{d}t - D^{1/2} \int_0^T s(\vec{x}, t, \vec{m}) d(\vec{x}, t)\, \mathrm{d}t \frac{\frac{1}{T} \int_0^T 2 s(\vec{x}, t'', \vec{m}) \frac{\partial s(\vec{x}, t'', \vec{m})}{\partial m_k}\, \mathrm{d}t''}{2 S^{1/2}} \right] \tag{B.12b}$$

$$= \frac{-2 \int_0^T d(\vec{x}, t) \frac{\partial s(\vec{x}, t, \vec{m})}{\partial m_k}\, \mathrm{d}t}{S^{1/2}\, D^{1/2}} + \frac{\frac{2}{T} \int_0^T \left(\int_0^T s(\vec{x}, t'', \vec{m}) d(\vec{x}, t'')\, \mathrm{d}t'' \right) s(\vec{x}, t, \vec{m}) \frac{\partial s(\vec{x}, t, \vec{m})}{\partial m_k}\, \mathrm{d}t}{S^{3/2}\, D^{1/2}} \tag{B.12c}$$

$$= \int_0^T -2 \left[\frac{d(\vec{x}, t)}{S^{1/2}\, D^{1/2}} - \frac{\frac{1}{T} \left(\int_0^T s(\vec{x}, t', \vec{m}) d(\vec{x}, t')\, \mathrm{d}t' \right) s(\vec{x}, t, \vec{m})}{S^{3/2}\, D^{1/2}} \right] \frac{\partial s(\vec{x}, t, \vec{m})}{\partial m_k}\, \mathrm{d}t \tag{B.12d}$$

Appendix C

3D/2D transformation using the Fourier-Bessel expansion

C.1 Vertical component

The expressions for the vertical component are derived by Thomas Forbriger (Forbriger et al., 2013). Using Equations (2.65) and (2.67) the Fourier coefficients of the vertical-component seismograms of a line source can be expressed by

$$\tilde{u}_{L_z}(y,\omega) = \int_{-\infty}^{\infty}\int_{0}^{\infty} G_z(\omega,k)\, J_0(k\sqrt{x^2+y^2})\, k\,\mathrm{d}k\, \mathcal{C}\,\mathrm{d}x \tag{C.1a}$$

$$= \int_{0}^{\infty} G_z(\omega,k) \int_{-\infty}^{\infty} J_0(k\sqrt{x^2+y^2})\, \mathcal{C}\,\mathrm{d}x\, k\,\mathrm{d}k \tag{C.1b}$$

$$= \int_{0}^{\infty} G_z(\omega,k)\, 2 \int_{0}^{\infty} J_0(k\sqrt{x^2+y^2})\, \mathcal{C}\,\mathrm{d}x\, k\,\mathrm{d}k. \tag{C.1c}$$

Applying the substitution

$$\beta(x) = \sqrt{\frac{x^2}{y^2}+1} \tag{C.2}$$

results in

$$\tilde{u}_{L_z}(y,\omega) = \int_{0}^{\infty} 2\, G_z(\omega,k) \int_{1}^{\infty} J_0(k\beta|y|) \frac{\beta|y|}{\sqrt{\beta^2-1}}\, \mathcal{C}\,\mathrm{d}\beta\, k\,\mathrm{d}k \tag{C.3a}$$

$$= \int_{0}^{\infty} 2\, G_z(\omega,k)\, |y| \int_{1}^{\infty} J_0(k|y|\beta) \frac{\beta}{\sqrt{\beta^2-1}}\, \mathcal{C}\,\mathrm{d}\beta\, k\,\mathrm{d}k. \tag{C.3b}$$

Gradshteyn & Ryzhik (1980, eq. 6.554.3) provide

$$\int_{1}^{\infty} J_0(k|y|\beta) \frac{\beta}{\sqrt{\beta^2-1}}\, \mathcal{C}\,\mathrm{d}\beta = \frac{1}{k|y|}\cos(k|y|)\mathcal{C} \tag{C.4}$$

for $k|y| > 0$. In Equation (C.1a) $k \geq 0$. Therefore, $k|y| > 0$ can be satisfied by excluding the case of vertical incidence and zero offset (Forbriger et al., 2013). The vertical component wavefield is then given by

$$\tilde{u}_{L_z}(y,\omega) = \int_{0}^{\infty} 2\, G_z(\omega,k)\cos(ky)\, \mathcal{C}\,\mathrm{d}k. \tag{C.5}$$

C.2 Horizontal component

Using Equations (2.66) and (2.68) the Fourier coefficients of the horizontal-component seismograms of a line source can be expressed by

$$\tilde{u}_{L_y}(y,\omega) = \int_{-\infty}^{\infty}\int_{0}^{\infty} G_r(\omega,k)\, J_1(k\sqrt{x^2+y^2})\, k\,\mathrm{d}k\, \frac{y}{\sqrt{x^2+y^2}}\, \mathcal{C}\,\mathrm{d}x \tag{C.6a}$$

$$= \int_{0}^{\infty} G_r(\omega,k) \int_{-\infty}^{\infty} J_1(k\sqrt{x^2+y^2})\, \frac{y}{\sqrt{x^2+y^2}}\, \mathcal{C}\,\mathrm{d}x\, k\,\mathrm{d}k \tag{C.6b}$$

$$= \int_{0}^{\infty} G_r(\omega,k)\, 2 \int_{0}^{\infty} J_1(k\sqrt{x^2+y^2})\, \frac{y}{\sqrt{x^2+y^2}}\, \mathcal{C}\,\mathrm{d}x\, k\,\mathrm{d}k. \tag{C.6c}$$

Applying the substitution of Equation (C.2) results in

$$\tilde{u}_{L_y}(y,\omega) = \int_{0}^{\infty} G_r(\omega,k)\, 2 \int_{1}^{\infty} J_1(k|y|\beta)\, \frac{y}{|y|\beta} \frac{|y|\beta}{\sqrt{\beta^2-1}}\, \mathcal{C}\,\mathrm{d}\beta\, k\,\mathrm{d}k \tag{C.7a}$$

$$= \int_{0}^{\infty} G_r(\omega,k)\, 2\, y \int_{1}^{\infty} J_1(k|y|\beta)\, \frac{1}{\sqrt{\beta^2-1}}\, \mathcal{C}\,\mathrm{d}\beta\, k\,\mathrm{d}k. \tag{C.7b}$$

Gradshteyn & Ryzhik (1980, eq. 6.552.6) provide

$$\int_{1}^{\infty} J_1(k|y|\beta)\, \frac{1}{\sqrt{\beta^2-1}}\, \mathcal{C}\,\mathrm{d}\beta = -\frac{\pi}{2}\, J_{1/2}\left(\frac{1}{2}k|y|\right)\, N_{1/2}\left(\frac{1}{2}k|y|\right)\, \mathcal{C} \tag{C.8}$$

for $k|y| > 0$. $J_{1/2}$ and $N_{1/2}$ are Bessel functions of the first kind and of the second kind (also called Neumann function), respectively, of fractional order 1/2. Again the case of vertical incidence and zero offset must be excluded to satisfy $k|y| > 0$. Furthermore, Gradshteyn & Ryzhik (1980, eq. 8.464.1 and eq. 8.469.1) provide

$$J_{1/2}\left(\frac{1}{2}k|y|\right) = \sqrt{\frac{4}{\pi k|y|}}\sin\left(\frac{1}{2}k|y|\right) \tag{C.9a}$$

$$N_{1/2}\left(\frac{1}{2}k|y|\right) = -\sqrt{\frac{4}{\pi k|y|}}\cos\left(\frac{1}{2}k|y|\right). \tag{C.9b}$$

Using these expressions in Equation (C.8) together with the addition theorem $2\sin\left(\frac{1}{2}k|y|\right)\cos\left(\frac{1}{2}k|y|\right) = \sin(k|y|)$ yields to

$$\int_{1}^{\infty} J_1(k|y|\beta)\, \frac{1}{\sqrt{\beta^2-1}}\, \mathcal{C}\,\mathrm{d}\beta = \frac{2}{k|y|}\, \sin\left(\frac{1}{2}k|y|\right)\, \cos\left(\frac{1}{2}k|y|\right)\, \mathcal{C} \tag{C.10}$$

$$= \frac{1}{k|y|}\, \sin(k|y|)\, \mathcal{C} \tag{C.11}$$

and with Equation (C.7) to

$$\tilde{u}_{L_y}(y,\omega) = \int_0^\infty 2\, G_r(\omega,k)\, \frac{y}{k|y|}\, \sin(k|y|)\, \mathcal{C}\, k\, \mathrm{d}k \tag{C.12a}$$

$$= \int_0^\infty 2\, G_r(\omega,k)\, \frac{y}{|y|}\, \sin(k|y|)\, \mathcal{C}\, \mathrm{d}k \tag{C.12b}$$

$$= \int_0^\infty 2\, G_r(\omega,k)\, \mathrm{sgn}(y)\, \sin(k|y|)\, \mathcal{C}\, \mathrm{d}k \tag{C.12c}$$

$$= \int_0^\infty 2\, G_r(\omega,k)\, \sin(ky)\, \mathcal{C}\, \mathrm{d}k. \tag{C.12d}$$

Appendix D

Benchmark results for pressure seismograms

I run the benchmark test described in Section 3.4 also for pressure seismograms.

D.1 Analytic solution

With the divergence operator in cylindrical coordinates and the bulk modulus k the pressure can be derived from Equation (3.18) via

$$P = -k \bigtriangledown \vec{u} = -(\lambda + \frac{2}{3}\mu)\frac{1}{r}\frac{\partial(ru_r)}{\partial r} \tag{D.1a}$$

$$= -(\lambda + \frac{2}{3}\mu)\left[\frac{1}{r}u_r + \frac{\partial u_r}{\partial r}\right] \tag{D.1b}$$

$$= (\lambda + \frac{2}{3}\mu)\frac{M_0' S(\omega)\omega}{4iv_p^3\rho}\left[-\frac{\omega}{v_p}H_2^{(2)}\left(\frac{\omega}{v_p}r\right) + \frac{2}{r}H_1^{(2)}\left(\frac{\omega}{v_p}r\right)\right] \tag{D.1c}$$

where λ and μ are the Lamé parameters.
For the Lamé parameters it holds that

$$v_s = \sqrt{\frac{\mu}{\rho}} \quad \rightarrow \quad \mu = \rho v_s^2 \tag{D.2a}$$

$$v_p = \sqrt{\frac{\lambda + 2\mu}{\rho}} \quad \rightarrow \quad \lambda = \rho v_p^2 - 2\rho v_s^2 \tag{D.2b}$$

and therefore the pressure is

$$P = \left(v_p^2 - \frac{4}{3}v_s^2\right)\frac{M_0' S(\omega)\omega}{4iv_p^3}\left[-\frac{\omega}{v_p}H_2^{(2)}\left(\frac{\omega}{v_p}r\right) + \frac{2}{r}H_1^{(2)}\left(\frac{\omega}{v_p}r\right)\right]. \tag{D.3}$$

D.2 Scaling of pressure wavefield in 2D FD code

In case of a line source $\left(\frac{\partial u_z}{\partial z} = 0\right)$ the pressure P can be expressed by

$$P = -\frac{\sigma_{xx} + \sigma_{yy} + \sigma_{zz}}{3} \tag{D.4}$$

with the normal stresses

$$\sigma_{xx} = \lambda \left(\frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} \right) + 2\mu \frac{\partial u_x}{\partial x} \tag{D.5a}$$

$$\sigma_{yy} = \lambda \left(\frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} \right) + 2\mu \frac{\partial u_y}{\partial y} \tag{D.5b}$$

$$\sigma_{zz} = \lambda \left(\frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} \right). \tag{D.5c}$$

From Equation (D.5) it follows

$$\sigma_{zz} = \lambda \frac{\sigma_{xx} + \sigma_{yy}}{2\lambda + 2\mu}. \tag{D.6}$$

Inserting σ_{zz} in Equation (D.4) provides together with equation (D.2)

$$P = -\frac{3v_p^2 - 4v_s^2}{2v_p^2 - 2v_s^2} \frac{\sigma_{xx} + \sigma_{yy}}{3}. \tag{D.7}$$

The normal stresses σ_{xx} and σ_{yy} are calculated within the FD code. However, to obtain correct amplitudes the scaling factor derived in Equation (D.7) must be applied.

D.3 Results

Figure D.1 displays the comparison of analytically calculated pressure seismograms with the simulated seismograms. All seismograms coincide in amplitudes and phases. No influence of the staggered grid is observed because both the source as well as the receivers are located on full grid points in this case.

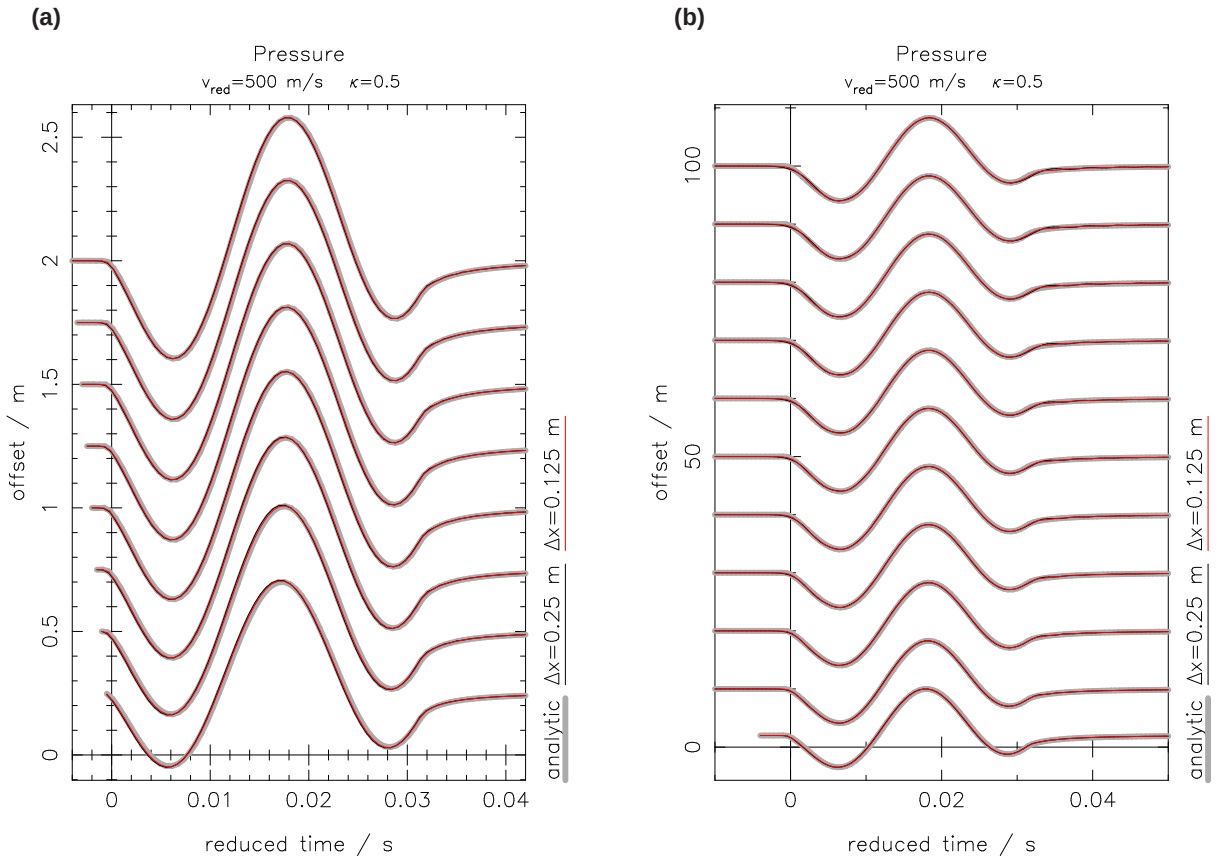


Figure D.1: Pressure seismograms for an explosive line source in a homogeneous full-space. Comparison of analytic results with seismograms simulated with different spatial grid spacing Δx (see modelings no. 1 and 2 from Table 3.1). The time axes in both plots are reduced with a velocity of 500 m/s. The seismograms are not normalized but they are multiplied by an offset dependent factor of $(r/\mathrm{m})^{\kappa}$ where κ is given in the head of each figure. a) displays seismograms with very small offsets and b) displays seismograms over the whole offset range.

Appendix E

Image technique together with force sources at the free surface

In the following I compare two FD schemes which can be used to calculate seismograms excited by a vertical force source directly at the free surface. The first FD scheme uses an implicit definition of the free surface by adding a vacuum or air layer on top of the model. In the second FD scheme the wavefield variables are shifted by half a grid point in vertical direction (y-direction) in comparison to the staggered grid used in the 2D FWI code DENISE. The free surface is implemented by an adjusted image technique. For the comparison the most simple model of a vertical point force at the surface of a homogeneous half-space is used.

E.1 Implementation of free surface by a vacuum layer

With the FD scheme used in the 2D FWI code (Section 3.1) it is possible to excite a vertical force source directly at the free surface in case of an implementation of the free surface by a vacuum or air layer. The free surface is located directly between the two full grid points where the upper one has the material parameters of the vacuum layer and the lower one has the material properties of the half-space (Section 3.6). The boundary condition $\sigma_{xy} = 0$ at the free surface is implicitly satisfied within this scheme because the Lamé parameter $< \mu >$ at the position of σ_{xy} is calculated by a harmonic average of the values of μ at the four surrounding full grid points and is therefore almost zero.[1] Since the material parameters ρ, λ, and μ are very small or even zero in the vacuum layer the waves are not able to penetrate into this layer.

E.2 Implementation of free surface by image technique

If the staggered grid displayed in Figure 3.2 is used in combination with the image technique, vertical point forces are injected half a grid point below the free surface. To obtain a staggered grid with the vertical particle velocity field at full grid points I shift the staggered grid half a grid point in vertical direction. The resulting staggered grid is shown in Figure E.1.

[1] Actually, in the DENISE code the averaged value of $< \mu >$ is set to zero as soon as one of the four surrounding μ values is zero to avoid numerical problems.

FD scheme

In this FD scheme the stress wavefield is updated by

$$\sigma_{xx}^{n+}(i,j^+) = \sigma_{xx}^{n-}(i,j^+) + \Delta t\left(\Big(\lambda(i,j^+)+2\mu(i,j^+)\Big)D_x^-\left[v_x^n(i^+,j^+)\right] + \lambda(i,j^+)D_y^+\left[v_y^n(i,j)\right]\right) \tag{E.1a}$$

$$\sigma_{yy}^{n+}(i,j^+) = \sigma_{yy}^{n-}(i,j^+) + \Delta t\left(\lambda(i,j^+)D_x^-\left[v_x^n(i^+,j^+)\right] + \Big(\lambda(i,j^+)+2\mu(i,j^+)\Big)D_y^+\left[v_y^n(i,j)\right]\right) \tag{E.1b}$$

$$\sigma_{xy}^{n+}(i^+,j) = \sigma_{xy}^{n-}(i^+,j) + \Delta t\ \mu(i^+,j)\Big(D_y^-\left[v_x^n(i^+,j^+)\right] + D_x^+\left[v_y^n(i,j)\right]\Big) \tag{E.1c}$$

(E.1d)

and the particle velocity wavefield by

$$v_x^{n+1}(i^+,j^+) = v_x^n(i^+,j^+) + \frac{\Delta t}{\rho(i^+,j^+)}\Big(D_x^+\left[\sigma_{xx}^{n+}(i,j^+)\right] + D_y^+\left[\sigma_{xy}^{n+}(i^+,j)\right] + f_x^{n+}(i^+,j^+)\Big) \tag{E.2a}$$

$$v_y^{n+1}(i,j) = v_y^n(i,j) + \frac{\Delta t}{\rho(i,j)}\Big(D_x^-\left[\sigma_{xy}^{n+}(i^+,j)\right] + D_y^-\left[\sigma_{yy}^{n+}(i,j^+)\right] + f_y^{n+}(i,j^+)\Big) \tag{E.2b}$$

where

$$\lambda(i,j^+) = <\lambda_y> = 2\left[\frac{1}{\lambda(i,j)} + \frac{1}{\lambda(i,j+1)}\right]^{-1} \tag{E.3a}$$

$$\mu(i,j^+) = <\mu_y> = 2\left[\frac{1}{\mu(i,j)} + \frac{1}{\mu(i,j+1)}\right]^{-1} \tag{E.3b}$$

$$\mu(i^+,j) = <\mu_x> = 2\left[\frac{1}{\mu(i,j)} + \frac{1}{\mu(i+1,j)}\right]^{-1} \tag{E.3c}$$

$$\rho(i^+,j^+) = \overline{\rho} = \frac{1}{4}\left[\rho(i,j)+\rho(i+1,j)+\rho(i,j+1)+\rho(i+1,j+1)\right] \tag{E.3d}$$

are the averaged material parameters. The density is averaged arithmetically and the Lamé parameters are averaged harmonically (Moczo, 1989; Graves, 1996).

Free surface conditions

In this scheme the free surface is located at the first row of grid points. To satisfy the free surface conditions the component σ_{xy} of the stress tensor is set explicitly to zero. Furthermore, in analogy to the technique proposed by Levander (1988) the stress components σ_{yy} and σ_{xy} are imaged via

$$\sigma_{yy}(i,j-(m-1/2)) = -\sigma_{yy}(i,j+(m-1/2)) \qquad m = 1,\ldots,N_{FD}/2 \tag{E.4a}$$

$$\sigma_{xy}(i,j-1) = -\sigma_{xy}(i,j+1) \qquad m = 1,\ldots,N_{FD}/2-1 \tag{E.4b}$$

where N_{FD} is the order of the spatial FD operators used for the simulation. The velocity field is not imaged as suggested by Robertsson et al. (1995).

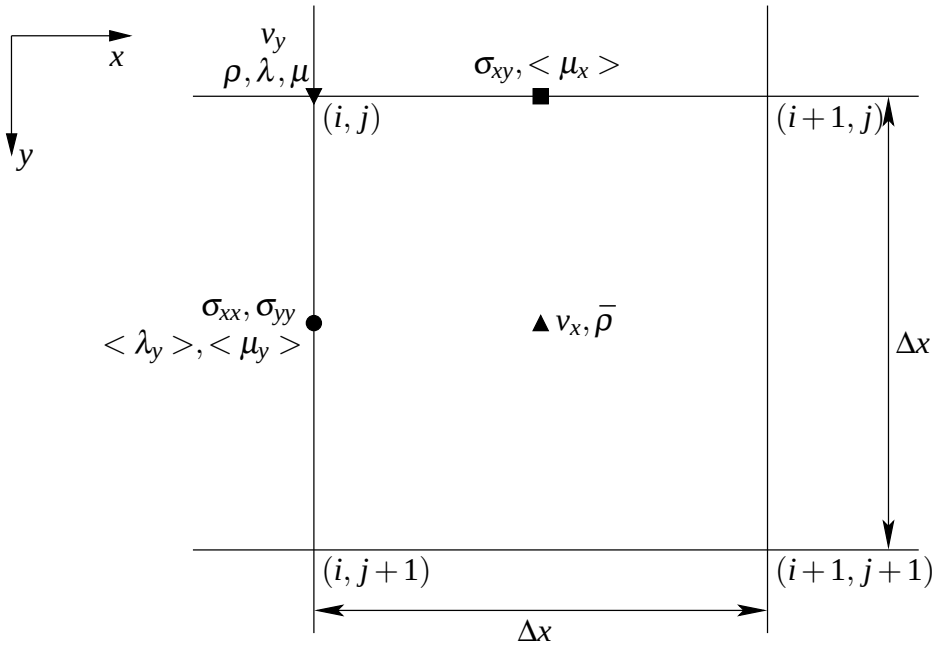


Figure E.1: Staggered grid used in the test for investigating the problem concerning the image technique in combination with forces excited directly at the free surface.

E.3 Comparison

I compare the previously described FD schemes for a homogeneous half-space. Figure E.2 sketches a small section of the model with the corresponding distribution of the material properties for the FD scheme where the free surface is implemented by a vacuum layer ($v_p = v_s = \rho = 0$ and therefore $\lambda = \mu = 0$). In the half-space I assume $\lambda = \lambda_1$, $\mu = \mu_1$ and $\rho = \rho_1$. Analogously, Figure E.3 sketches the distribution of material properties for the FD scheme where the free surface is implemented by the image technique.

E.3.1 Evaluation of FD schemes by hand for an initial value problem

First I will evaluate both FD schemes by hand to solve an initial value problem. For convenience I use spatial FD operators of second order. In this case both implementations provide identical results. As initial condition I set the vertical velocity wavefield at a grid point at the free surface for the first time step to the initial value v_{ini}. Figure E.4 to E.7 show the evolution of the wavefields for the first two time steps.

Time step 1:
After the first time step there is only one non-zero value in the vertical velocity wavefield which is given by the initial condition (Figure E.4 and E.5).

Time step 2:
During time step 2 first the stress field is updated. According to Figure E.6 and E.7 they are

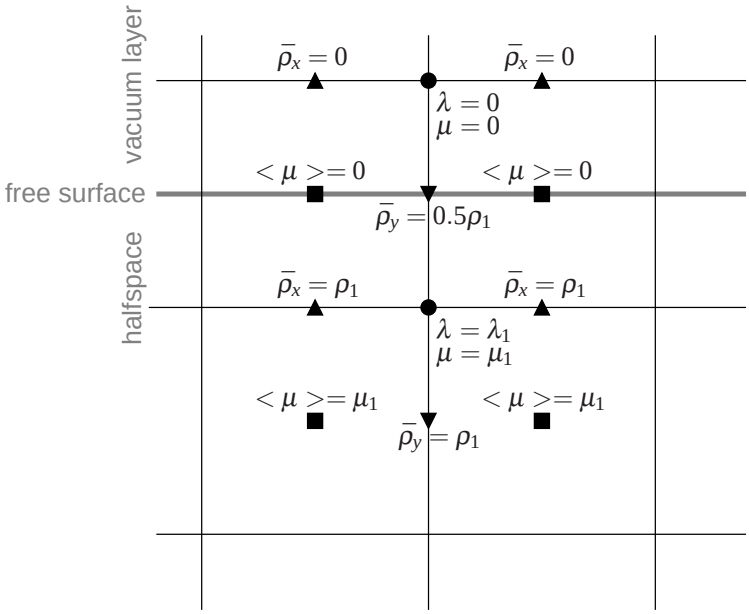


Figure E.2: Small section of the model which displays the parameter distribution in the vicinity of the free surface. The free surface is implemented implicitly by adding a vacuum layer at the top of the model. The elastic material properties in the half-space are λ_1, μ_1 and ρ_1. In the vacuum layer we assume $\lambda = \mu = \rho = 0$.

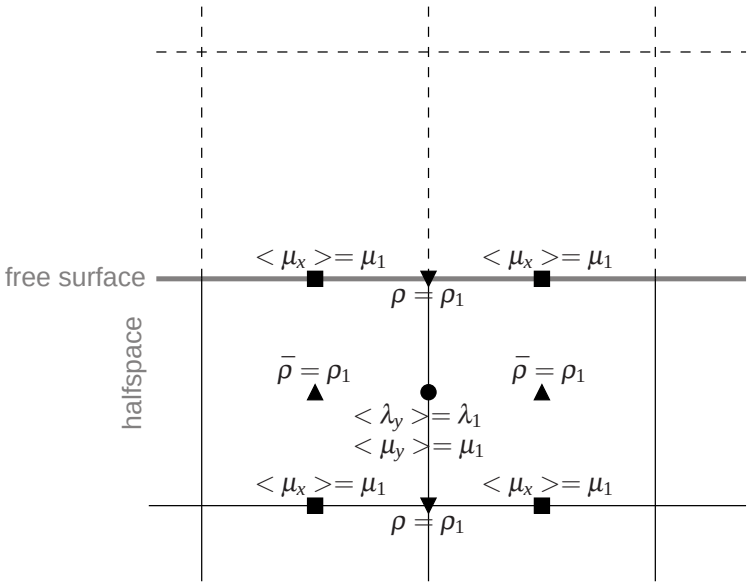


Figure E.3: Small section of the model which displays the parameter distribution in the vicinity of the free surface. The free surface is implemented by the imaging technique. The elastic material properties in the half-space are λ_1, μ_1 and ρ_1.

equal. In a second step the particle velocity wavefield is evaluated. To calculate the vertical particle velocity at the free surface only the derivative of σ_{yy} with respect to the vertical or y-component is required because σ_{xy} is zero at the free surface and therefore also its derivative with respect to the x-component is zero. If a vacuum layer is used to implement the free surface (Figure E.6) σ_{yy} is zero above the free surface. However, because of the averaging of the density values in this FD scheme the density value at the source point is only half the density of the half-space. The vertical velocity can be calculated via (see Equation 3.8e)

$$v_y(i,j^-) = v_{ini} + \frac{\Delta t}{0.5\rho_1}\frac{1}{\Delta x}\sigma_{yy}(i,j). \tag{E.5}$$

In case of an implementation of the free surface by the image technique (Figure E.7) the vertical velocity wavefield is calculated at full grid points and therefore the density values are not averaged. However in contrast to the FD scheme including the vacuum layer the normal stress σ_{yy} is not zero above the free surface but is imaged from the point below the free surface so that $\sigma_{yy}(i,j^-) = -\sigma_{yy}(i,j^+)$. Technically this compensates the fact that the full density value is used in the update of the vertical velocity at the free surface and we finally obtain (see Equation E.2b)

$$v_y(i,j) = v_{ini} + \frac{\Delta t}{\rho_1}\frac{1}{\Delta x}2\sigma_{yy}(i,j^+) \tag{E.6}$$

which is equal to Equation (E.5).

Since the initial value problem is solved equivalently by both FD schemes the problem must arise in the source implementation.

E.3.2 Difference in source scaling between the two FD schemes

In the DENISE code a force source is defined by a line force density F_0' and a vector $\vec{s} = (s^1, s^2, s^3, \dots)^T$ which contains the discrete time samples of the source time function. This external force is internally converted to a contribution to the particle velocity wavefield at the source point as discussed in Section 3.3.2. If we again consider the example of a homogeneous half-space which was discussed in the previous section the contribution to the vertical velocity wavefield at the source point for the first time step would be

$$v_y^1(i,j^-) = \frac{\Delta t}{0.5\rho_1}\frac{F_0'}{\Delta x^2}s^{1+} \tag{E.7}$$

for the FD scheme where the free surface is implemented by the vacuum layer and

$$v_y^1(i,j) = \frac{\Delta t}{\rho_1}\frac{F_0'}{\Delta x^2}s^{1+} \tag{E.8}$$

for the FD scheme where the free surface is implemented by the image technique. The contributions differ by a factor of two. Technically this is caused by the fact that for the free surface implemented by a vacuum layer the density at the source point is half of the density of the homogeneous half-space (Figure E.2) due to the arithmetic averaging of the density values. For the image technique the vertical velocity wavefield is located at the full grid points and therefore the density values are not averaged.

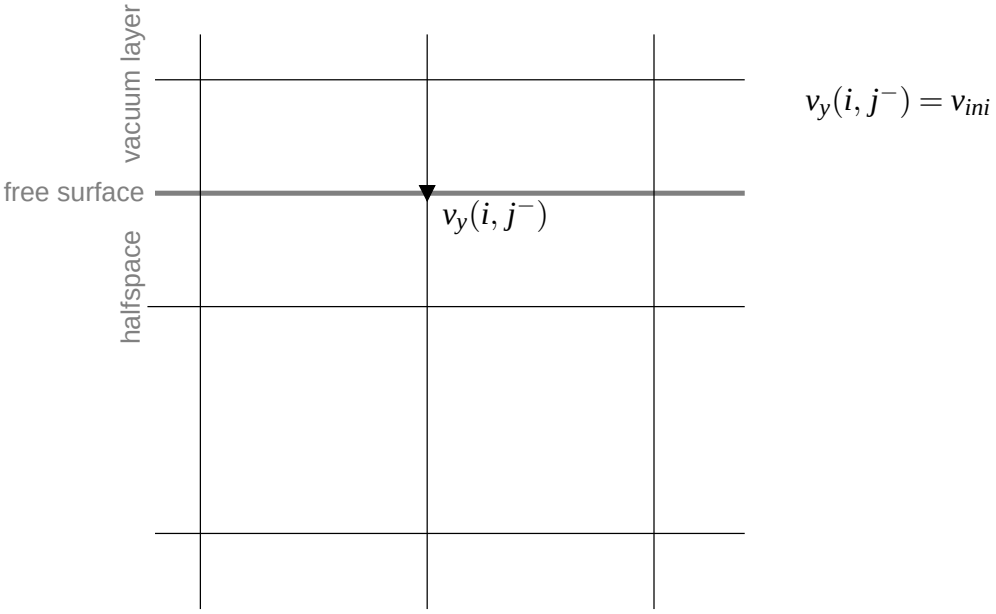


Figure E.4: Wavefields for time step 1. Free surface implemented by a vacuum layer. Vertical force source injected at grid point i, j^-. Only non zero wavefield variables are displayed.

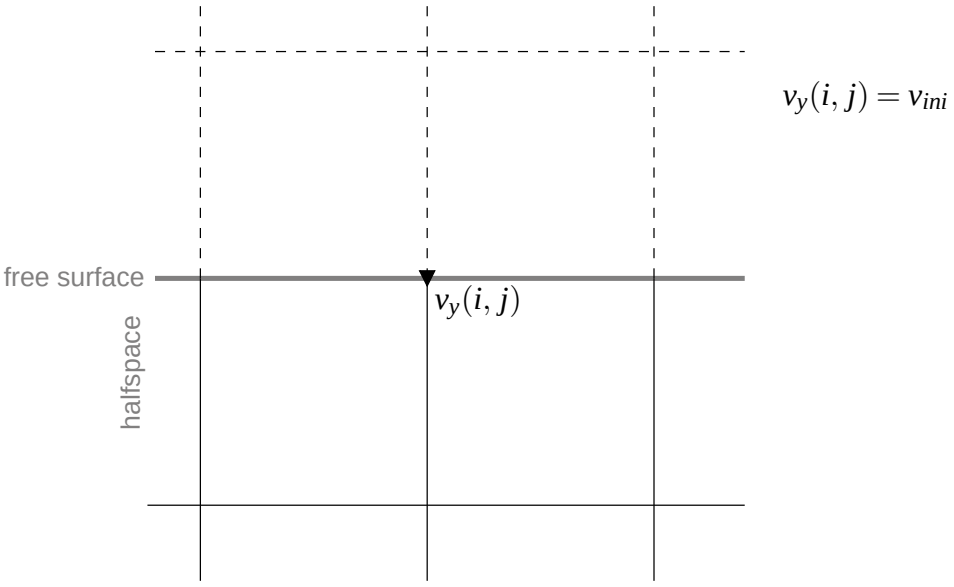
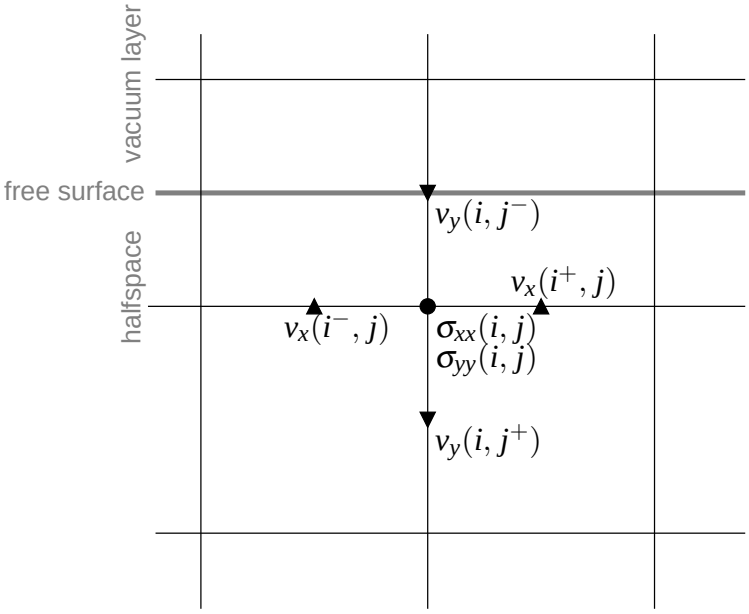


Figure E.5: Wavefields for time step 1. Free surface implemented by image technique. Vertical force source injected at grid point i, j. Only non zero wavefield variables are displayed.



$$\sigma_{xx}(i,j) = -\frac{\Delta t}{\Delta x}\lambda_1 v_{ini}$$

$$\sigma_{yy}(i,j) = -\frac{\Delta t}{\Delta x}(\lambda_1 + 2\mu_1) v_{ini}$$

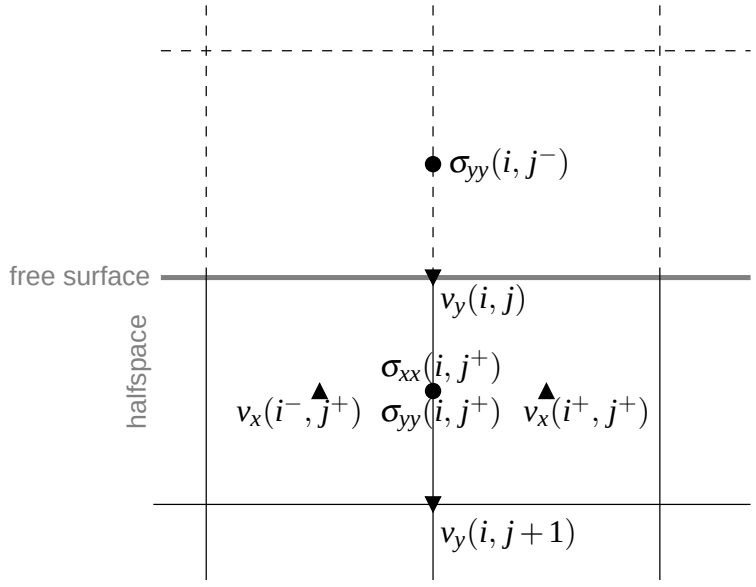
$$v_y(i,j^-) = v_{ini} - \frac{\Delta t^2}{\Delta x^2}\frac{\lambda_1+2\mu_1}{0.5\rho_1} v_{ini}$$

$$v_y(i,j^+) = \frac{\Delta t^2}{\Delta x^2}\frac{\lambda_1+2\mu_1}{\rho_1} v_{ini}$$

$$v_x(i^-,j) = -\frac{\Delta t^2}{\Delta x^2}\frac{\lambda_1}{\rho_1} v_{ini}$$

$$v_x(i^+,j) = -v_x(i^-,j)$$

Figure E.6: Wavefields for time step 2. Free surface implemented by an air layer. Vertical force source injected at gridpoint i, j^+. Only non zero wavefield variables are displayed. Note that the displayed velocity wavefields are half a time step later than the displayed stress fields.



$$\sigma_{xx}(i,j^+) = -\frac{\Delta t}{\Delta x}\lambda_1 v_{ini}$$

$$\sigma_{yy}(i,j^+) = -\frac{\Delta t}{\Delta x}(\lambda_1 + 2\mu_1) v_{ini}$$

$$\sigma_{yy}(i,j^-) = -\sigma_{yy}(i,j^+)$$

$$v_y(i,j) = v_{ini} - 2\frac{\Delta t^2}{\Delta x^2}\frac{\lambda_1+2\mu_1}{\rho_1} v_{ini}$$

$$v_y(i,j+1) = \frac{\Delta t^2}{\Delta x^2}\frac{\lambda_1+2\mu_1}{\rho_1} v_{ini}$$

$$v_x(i^-,j^+) = -\frac{\Delta t^2}{\Delta x^2}\frac{\lambda_1}{\rho_1} v_{ini}$$

$$v_x(i^+,j^+) = -v_x(i^-,j^+)$$

Figure E.7: Wavefields for time step 2. Free surface implemented by image technique. Vertical force source injected at gridpoint $i, j+1$. Only non zero wavefield variables are displayed. Note that the displayed velocity wavefields are half a time step later than the displayed stress fields.

Appendix F

Estimation of source wavelet correction filter

To estimate a source wavelet correction filter I apply a method proposed by T. Forbriger. In the following, I will only shortly review this method. For more details I refer to Groos et al. (2013).
The Fourier coefficients $\tilde{c}_l$ of the filter for angular frequencies $\omega_l = \Delta\omega_l$ are estimated by solving a damped linear least-squares problem which is described by minimizing the objective function

$$F(\tilde{c}_l;\varepsilon) = \sum_{l=0}^{N-1}\sum_{k=1}^{M} f_k^2 \left|\tilde{d}_{lk} - \tilde{c}_l\tilde{g}_{lk}\right|^2 + M\bar{E}\,\varepsilon^2 \sum_{l=0}^{N-1}|\tilde{c}_l|^2 \tag{F.1}$$

with respect to the real and imaginary parts of all $\tilde{c}_l$. Thereby, $\tilde{d}_{lk}$ are the complex Fourier expansion coefficients for a time series recorded at offset r_k to the source and $\tilde{g}_{lk}$ is the corresponding Fourier coefficient for the synthetic seismogram. The synthetic seismograms are obtained by a forward simulation for an assumed subsurface model and a source wavelet which is described by the Fourier coefficients $\tilde{s}_l$. The sum over l corresponds to the sum over the N Fourier coefficients and the sum over k corresponds to the sum over the M receiver. Furthermore,

$$\bar{E} = \frac{1}{MN}\sum_{l=0}^{N-1}\sum_{k=1}^{M} f_k^2\,|\tilde{g}_{lk}|^2 \tag{F.2}$$

is the average power of the Fourier coefficients of the synthetic seismograms scaled by f_k. The filter coefficients are calculated via

$$\tilde{c}_l = \frac{\sum_{k=1} M f_k^2\,\tilde{g}_{lk}^*\tilde{d}_l k}{M\bar{E}\,\varepsilon^2 + \sum_{k=1} M f_k^2\,|\tilde{g}_{lk}|^2} \tag{F.3}$$

which is essentially a stabilized or water-level deconvolution, respectively.
The Fourier coefficients of an optimized source wavelet are given by $\tilde{s}_l^{opt} = \tilde{c}_l\tilde{s}_l$.
The scaling coefficients f_k can be used to ensure that all receivers, independently of their offset r_k contribute to an equal average amount to the least-squares objective function in Equation F.1. The coefficients are defined by

$$f_k = \left(\frac{r_k}{1\,\mathrm{m}}\right)^{\kappa} \tag{F.4}$$

where κ is used to adjust a compensation for a power law attenuation.

This method is similar to the method proposed by Pratt (1999). However, Pratt (1999) only discusses the case for $\varepsilon^2 = 0$ and does not apply a weighting of the data in the least-squares objective function which is minimized.

Appendix G

Application of L-curve criterion

A source wavelet correction filter must be applied to the synthetic data prior to the comparison of field data with elastically and viscoelastically simulated data. Since I also compare the source wavelet correction filters in this investigation a bias in the results due to an inappropriate damping should be avoided. Therefore, I use the L curve criterion (Aster et al., 2013, section 4.1) to determine an appropriate damping ε^2 for each of the two least-squares optimization problems used to determine the source wavelet correction filter for the elastically and viscoelastically simulated data. The objective function minimized for the determination of the source wavelet correction filter is given by Equation F.1. According to Aster et al. (2013) I solve the linear least-squares problem for different damping values ε^2. Afterwards, the solution norm

$$\chi = \left(M\overline{E} \sum_{l=0}^{N-1} |\tilde{c}_l|^2 \right)^{1/2} \tag{G.1}$$

is plotted against the residual norm

$$\psi = \left(\sum_{l=0}^{N-1} \sum_{k=1}^{M} f_k^2 \left| \tilde{d}_{lk} - \tilde{c}_l \tilde{g}_{lk} \right|^2 \right)^{1/2} \tag{G.2}$$

on a log-log scale. (Note that the whole objective function of Equation F.1 is given by $F = \chi^2 + \varepsilon^2 \psi^2$.) The results are displayed in Figure G.1. In the anelastic case (Figure G.1a) the shape of the curve is very similar to an "'L"' and I chose a value of $\varepsilon^2 = 0.15$ at the break point of the curve as damping value. Furthermore, it can be observed that the residual norm or data misfit, respectively, does not change as much as in the elastic case (Figure G.1b) for different tested damping values. This is also true for the residual norm or length of the model vector. In the elastic case there is no clear "'L"'-shape observable. This is presumably caused by the fact that in the elastic case I try to explain the effect of anelastic damping by the source wavelet correction filter. This is physically not correct. As appropriate damping I choose $\varepsilon^2 = 0.05$ where the data misfit is still small in addition to a comparable small length of the model vector.

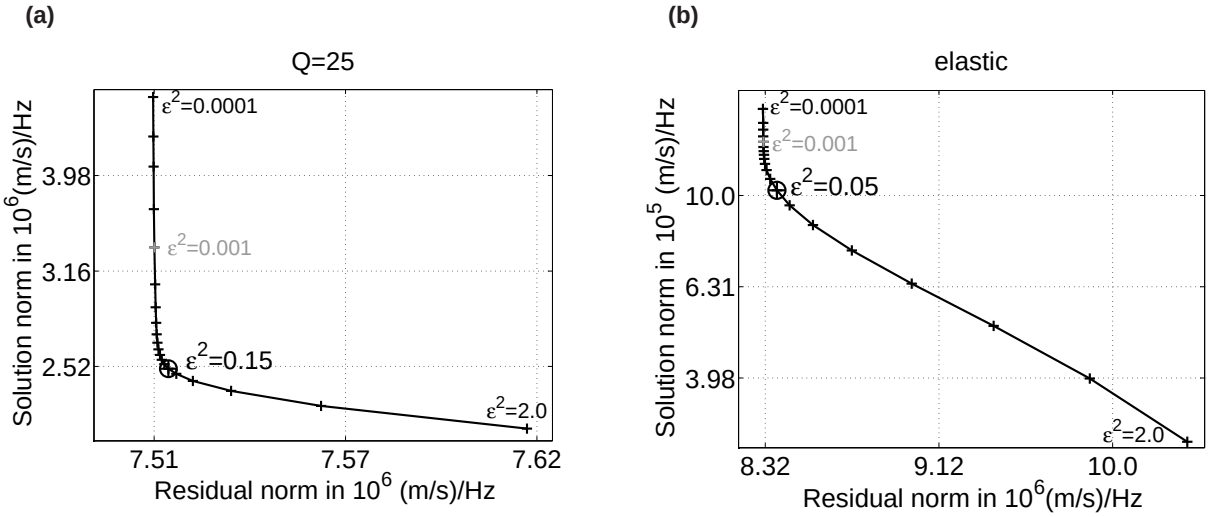


Figure G.1: Results of L-curve analysis for a) viscoelastic forward simulation with Q=25 and b) elastic forward simulation. The finally chosen damping values are marked by the larger cross and the circle whereas the corresponding values for ε^2 are given close beside them. The smallest and largest values used for ε^2 in this analysis are also given in the figures. Furthermore, the value used for the reconstruction tests (Section 4.2.2) is marked in grey.

Appendix H

Modeling parameters

Modeling parameters	
Model size in grid points ($N_x \times N_y$)	400 $\times$ 75
Spatial grid spacing Δx in m	0.2
Width of CPML layers in grid points	20
Order of spatial FD operator	4
Number of CPUs in x-direction	4
Number of CPUs in y-direction	1
Sampling interval Δt in s	5.0e^{-5}
Number of time steps per shot	14000
Implementation of free surface	image technique
Acquisition geometry	
Number of shots	8
Shot spacing in m	10
Number of receivers	63 (no receivers at source positions)
Receiver spacing in m	1
Minimum offset in m	1
Maximum offset in m	69

Table H.1: Relevant modeling parameters used for the reconstruction tests presented in Chapter 4.

Modeling parameters	
Model size in grid points ($N_x \times N_y$)	432 $\times$ 76
Spatial grid spacing Δx in m	0.2
Width of CPML layers in grid points	20
Order of spatial FD operator	4
Number of CPUs in x-direction	8
Number of CPUs in y-direction	2
Sampling interval Δt in s	5.0e^{-5}
Number of time steps per shot	18000
Implementation of free surface	image technique
Acquisition geometry	
Number of shots	25
Shot spacing in m	3
Number of receivers	72
Receiver spacing in m	1
Minimum offset in m	1
Maximum offset in m	72

Table H.2: Relevant modeling parameters used for FWI of field data set presented in Chapter 5.

Frequency band	Filter length for P-wave velocity model	Minimum wavelength of P-waves	Filter length for S-wave velocity model	Minimum wavelength of Rayleigh waves	Filter length for density
5 Hz-10 Hz	1.4 m	33 m	1.4 m	25 m	1.4 m
5 Hz-15 Hz	1.4 m	21 m	1.4 m	14 m	1.4 m
5 Hz-20 Hz	1.4 m	15 m	1.4 m	10 m	1.4 m
5 Hz-30 Hz	1.4 m	10 m	1.4 m	6 m	1.4 m
5 Hz-40 Hz	2.6 m	7 m	0.6 m	4 m	0.6 m
5 Hz-55 Hz	2.6 m	5 m	0.6 m	3 m	0.6 m
5 Hz-70 Hz	2.6 m	3 m	0.6 m	2 m	0.6 m

Table H.3: Filter lengths of the median filters applied in FWI of field data set to the gradients of the misfit function and the models of the different model parameters. The minimum wavelengths of the P-waves are estimated by the minimum P-wave velocities occurring in the central part of the obtained 2D model after the inversion of the corresponding frequency band. The minimum wavelengths of the Rayleigh waves are estimated from the Fourier-Bessel expansion coefficients of shot 25 displayed in Figure 5.4b). For the considered subsurface model the minimum wavelengths of the S-waves are most likely larger than the minimum wavelengths of the Rayleigh waves.

Appendix I

Used hard- and software

I.1 Software

The FWI results were computed with the 2D FWI code DENISE. DENISE has been mainly developed by Daniel Köhn. It has been extended by Martin Schäfer, Sven Heider, and myself. This code is available at `www.opentoast.de` or `http://www.gpi.kit.edu/Software.php` under the terms of GNU General Public License.

For the inversion of Fourier-Bessel expansion coefficients and first arrival P-wave travel times which was applied to the field data set I used the inversion code `gremlin`. This code was provided by Thomas Forbriger. It is part of `TFSoftware`. Many other tools of `TFSoftware` were used for the preprocessing and the visualization of seismic data. `TFSoftware` includes, for example, tools for filtering and tapering seismic data, the calculation of Fourier-Bessel expansion coefficients, the estimation of appropriate source wavelet correction filters, the application of 3D/2D transformation, and so forth. It is maintained and mainly developed by Thomas Forbriger. These tools are available under the terms of the GNU General Public License. For visualization `TFSoftware` uses `PGPLOT` (written by Tim Pearson).

`Matlab for Linux (The Mathworks)` was used to visualize and further analyze FWI results. Furthermore, the calculation of analytic solutions for the benchmark tests were performed with a program in `Matlab`. For the conversion of data `Seismic Un*x` was used. This thesis is written in LaTeX 2_ε. Figures are created with `Matlab`, `PGPLOT` (used in `TFSoftware`) and `Xfig`.

I.2 Hardware

Data processing and analysis as well as most FWI applications were done on personal computers with Linux operating systems openSUSE 11.1 and openSUSE 12.1.

FWI results of the field data set presented in this thesis were computed on the high-performance supercomputer JUROPA at Jülich Supercomputing Centre. Further FWI tests were computed on the high-performance computer InstitutsCluster II (ic2) at Steinbuch Centre for Computing at Karlsruhe Institute of Technology.

Danksagung

Thomas Bohlen danke ich für die Übernahme des Hauptreferats dieser Arbeit, sowie für seine hilfreichen Hinweise, vor allem im Hinblick auf Finite-Differenzen-Modellierungen. Dadurch wurde so manches verwirrende Ergebnis doch noch plausibel. Außerdem danke ich ihm für sein Vertrauen und für die vielen Freiheiten, die ich bei der Durchführung meiner Promotion hatte.

Wolfgang Friederich danke ich für die Übernahme meines Korreferats, das er ohne zu zögern übernommen hat. Weiterhin möchte ich mich für die Einladung nach Bochum und den Tag dort mit vielen spannenden Diskussionen bedanken.

Daniel Köhn hat mit seinem 2D FWI Code DENISE diese Arbeit überhaupt erst ermöglicht. Neben der freien Verfügung über seinen Code hat er vor allem zu Beginn meiner Arbeit meine Fragen schnell und vor allem ausführlich beantwortet. Dafür möchte ich mich bei ihm bedanken.

Ein großer Dank geht an alle Mitarbeiterinnen und Mitarbeiter des GPI, vor allem an die Mitglieder der Arbeitsgruppe der Angewandten Seismik. Dank Ihnen bin ich über vier Jahre lang jeden Tag gerne ans GPI gekommen. Claudia Payne hat meine Fragen immer geduldig beantwortet, egal ob zu Dienstreiseabrechnungen oder zu anderen Verwaltungsprozessen am KIT. Petra Knopf war bei IT-Problemen immer schnell zur Stelle. Mit ihrer Hilfe hat mein treuer Begleiter gpiag4 doch noch bis zum Ende durchgehalten. Sven, Anja, Martin, Simone, Ines und lange Zeit Anna und André waren immer heitere Promotionsmitstreiter. Ganz nach dem Motto „Geteiltes Leid ist halbes Leid“ habe ich so manchen Tiefpunkt durch sie schnell überwunden. Auch Jürgen und Stefan standen bei Problemen stets mit Rat und Tat zur Seite. Aber auch alle Diplom-, Master- und Bachelorstudenten, sowie alle Hiwis haben zur guten Atmosphäre in der Arbeitsgruppe beigetragen. Hilfreich waren für mich die zahlreichen Arbeitsgruppentreffen, bei denen ich viele Anregungen und hilfreiche Tipps bekommen habe. Die sehr ausführlich diskutierten Probevorträge waren stets eine sehr gute Vorbereitung auf Konferenzen.

Danken möchte ich auch allen, die bei der Messung des Datensatzes geholfen haben, der ein Kapitel dieser Arbeit füllt. Allen voran Stefan, der trotz vergessener Ohrschützer über 100 Mal den Hammer geschwungen hat. Aber natürlich auch Martin, Matthias, Johannes, André und Florian aus Leipzig, die bei den Messungen in Rheinstetten dabei waren. Weiterhin gilt mein Dank auch Zacharias. Er hat im Rahmen seiner Diplomarbeit den Segelflugplatz in Rheinstetten als Messplatz entdeckt und mit seiner Auswertung den Grundstein für weitere Messungen dort gelegt.

Ein weiterer Dank geht an alle Korrekturleserinnen und Korrekturleser dieser Arbeit. Durch sie habe ich wertvolle Hinweise erhalten, die diese Arbeit verbessert haben. Bedanken möchte ich mich aber auch bei allen, die angeboten haben, beim Korrekturlesen zu helfen. Auch wenn ich nicht auf alle Angebote zurückgegriffen habe, stellten sie für mich eine große Beruhigung dar.

Martin möchte ich danken für viele spannende Diskussionen über die Wellenforminversion von flachseismischen Oberflächenwellen. Ich habe es sehr genossen, dass ich nicht allein an diesem Thema gearbeitet habe und ich mich immer mit ihm austauschen konnte.

Besonderer Dank gilt Ines und Jörn. Sie haben an den Tiefpunkten dieser Arbeit am meisten mitgelitten. Auf sie war immer Verlass. Egal ob es sich um fachliche Fragen handelte oder ob ich einfach nur jemanden zum Reden brauchte, weil unerwartete Probleme den Tag bestimmten. Ihre Bürotüren waren für mich immer geöffnet.

Thomas Forbriger hat diese Arbeit mit sehr viel Interesse begleitet. Er war bei drängenden Problemen sofort zur Stelle und hat sich für meine Fragen immer viel Zeit genommen. So manches Fax und so manche E-Mail, die mich von ihm erreichten, lichteten meinen Blick, wenn ich den Wald vor lauter Bäumen nicht mehr sah. Sein Enthusiasmus sorgte dafür, dass ich immer wieder frisch motiviert weitergearbeitet habe. Ohne ihn würde es diese Arbeit ganz sicher nicht in dieser Art geben.

Acknowledgment

This work was funded by the Federal Ministry of Education and Research grant 03G0752 (project TOAST). It was also kindly supported by the sponsors of the Wave Inversion Technology (WIT) Consortium, Germany. The author gratefully acknowledges the computing time granted on the supercomputer JUROPA at Jülich Supercomputing Centre (JSC).