European Society of Computational Methods in Sciences and Engineering (ESCMSE)

Journal of Numerical Analysis, Industrial and Applied Mathematics (JNAIAM)
vol. 1, no. 1, 2006, pp. 1-3
ISSN 1790–8140

Simulation and inversion of seismic wave propagation on continental scales based on a spectral-element method

Andreas Fichtner*[1], Heiner Igel*, Hans-Peter Bunge* and Brian L. N. Kennett†

*Dept. of Earth and Environmental Sciences, Ludwig-Maximilian University, Munich, Germany
†Research School of Earth Sciences, The Australian National University, Canberra, Australia

Received 7 March, 2006; accepted in revised form 10 March, 2006

Abstract: We propose a novel technique for seismic waveform tomography on continental scales. This is based on the fully numerical simulation of wave propagation in complex Earth models, the inversion of complete waveforms and the quantification of the waveform discrepancies through a specially designed phase misfit. The numerical solution of the equations of motion allows us to overcome the limitations of ray theory and of finite normal mode summations. Thus, we can expect the tomographic models to be more realistic and physically consistent. Moreover, inverting entire waveforms reduces the non-uniqueness of the tomographic problem. Following the theoretical descriptions of the forward and inverse problem solutions, we present preliminary results for the upper mantle structure in the Australasian region.

Keywords: seismic tomography, spectral-element method, adjoint-method, Australia

Mathematics Subject Classification: 86-08, 86A15, 86A22, 35R30, 74J15, 74S05, 74S25

PACS: 91.30.Ab, 91.30.Jk, 91.35.Cb, 91.35.Pn

1 Introduction

Seismic tomography is a powerful tool that allows us to infer the Earth's structure from observations of seismic waves at the surface. It provides information about the current dynamics [1] and the long-term evolution of our planet [2]. The physical basis of seismic tomography on continental and global scales are elastic waves that are excited by natural earthquakes. The waves propagate through the Earth and are recorded in the form of seismograms that may be used for tomography. Seismic tomography differs from technical tomographies on a laboratory scale in several decisive points: (1) Sufficiently large natural earthquakes as sources of seismic waves are poorly distributed. They only occur along the few tectonically active zones of the Earth. (2) The earthquake mechanism, i.e. the details of the source, are often poorly constrained. Even the magnitude of an earthquake can often not be determined with sufficient accuracy. (3) Receivers can only be installed and maintained in the relatively small accessible and secure regions of the Earth. This

[1]Corresponding author. E-mail: fichtner@geophysik.uni-muenchen.de

mostly excludes the deep ocean floor, high mountains, deserts, polar regions and politically unstable countries. Thus, large regions of the Earth, especially in the southern hemisphere, are rather poorly illuminated by seismic waves. (4) The seismic wavefield is highly complex, consisting of compressional waves, shear waves and different types of interface and surface waves that may convert to each other. This complexity increases due to the presence of strong lateral heterogeneities, topography and oceans of varying depth. (5) Seismic waveforms, and their amplitudes in particular, may depend highly non-linearly on the structure of the Earth, therefore rendering a linearised inversion difficult. As a result of these complications, seismic tomography is a non-linear and ill-posed inverse problem that bears little resemblance with laboratory ray tomography based on variants of the central-slice theorem [3][4].
To date seismic tomography has mostly been based on simplified forward problems: either ray theory [5][6] or truncated normal mode summations [7]. The corresponding inversion techniques are computationally efficient and reliable. Nevertheless, there is a high potential for further improvements: (1) To reduce the non-uniqueness of the seismic tomographic problem, more waveform information needs to be exploited. This requires that suitable measures of waveform misfit be designed. While extracting as much information as possible from the seismograms, the misfits must respect the particularities and problems of the seismic tomographic problem in order to be successful. (2) The maximum amount of usable waveform information can be exploited only when the solution of the forward problem is sufficiently accurate. Thus, fully numerical methods need to be used to simulate the propagation of elastic waves through a highly heterogeneous Earth model.

In this paper we describe a new waveform inversion method that is applicable on continental scales. We propose to solve the forward problem, i.e. the elastic wave equation, using a spectral-element discretisation that maintains the natural spherical coordinate system. To quantify the discrepancies between the observed seismograms and the seismograms computed for a reference Earth model we introduce a phase misfit that extracts a large amount of waveform information while being both reliably measurable and quasi-linearly related to Earth structure. We apply our new approach to the imaging of upper mantle structure under the Australian continent.

The principal advancements of our method are the fully numerical computation of the seismic wavefield for continental-scale tomography, the use of full waveforms and the actual applicability to real data.

This paper is outlined as follows: We start with a review of the equations that govern the propagation of seismic waves in the Earth. Special attention is given to the implementations of anisotropy and visco-elastic dissipation. Section 3 deals with the numerical solution of the equations of motion using a spectral-element discretisation in a spherical section. The setup of a seismic waveform tomography as a discrete optimisation problem is the topic of section 4. We elaborate on the design of a suitable waveform misfit and the computation of its gradient with respect to the model parameters via the adjoint method. Finally, we illustrate the application of our new method with preliminary tomographic images of the Australian upper mantle.

2 Equations governing the propagation of seismic waves

The propagation of seismic waves through a model Earth $G \subset \mathbb{R}^3$ can be described by the linearised momentum balance equation that relates the displacement field $\mathbf{u}(\mathbf{x},t)$ to the stress tensor $\boldsymbol{\sigma}(\mathbf{x},t)$, the external force density $\mathbf{f}(\mathbf{x},t)$ and the mass density distribution in the Earth, $\rho(\mathbf{x})$ [8][9]:

$$\rho(\mathbf{x})\,\ddot{\mathbf{u}}(\mathbf{x},t) - \nabla\cdot\boldsymbol{\sigma}(\mathbf{x},t) = \mathbf{f}(\mathbf{x},t)\,, \quad \mathbf{x}\in G\,, t\in[t_0,t_1]\,. \tag{1}$$

At the surface ∂G of the Earth model the normal components of the stress tensor $\boldsymbol{\sigma}$ vanish, i.e., the Neumann condition

$$\boldsymbol{\sigma} \cdot \mathbf{n}|_{\mathbf{x}\in\partial G} = \mathbf{0} \tag{2}$$

is satisfied. The symbol $\mathbf{n}$ in equation (2) denotes the outward-pointing unit normal. Both the displacement field $\mathbf{u}$ and the velocity field $\mathbf{v} = \dot{\mathbf{u}}$ are required to be zero prior to the time $t = t_0$ when the external force $\mathbf{f}$ starts to act:

$$\mathbf{u}|_{t<t_0} = \mathbf{v}|_{t<t_0} = \mathbf{0}\,. \tag{3}$$

To obtain a complete set of equations, the stress tensor needs to be related to the displacement field. We assume a visco-elastic rheology where the current stress tensor $\boldsymbol{\sigma}$ depends linearly on the time-history of the strain tensor $\boldsymbol{\epsilon} = \frac{1}{2}(\nabla\mathbf{u} + \nabla\mathbf{u}^T)$. In symbols:

$$\boldsymbol{\sigma}(\mathbf{x}, t) = \int_{t=t_0}^{t} \dot{\mathbf{C}}(\mathbf{x}, t - t') : \boldsymbol{\epsilon}(\mathbf{x}, t')\, dt'\,. \tag{4}$$

The fourth-order tensor $\mathbf{C}$ is the elastic tensor. In the case of a perfectly elastic medium it takes the form $\mathbf{C}(\mathbf{x}, t) = \mathbf{C}(\mathbf{x})\, H(t)$, where H is the Heaviside function. We then have $\boldsymbol{\sigma} = \mathbf{C}(\mathbf{x}) : \boldsymbol{\epsilon}(\mathbf{x}, t)$. The elastic tensor components satisfy the symmetry relations $C_{ijkl} = C_{klij} = C_{jikl}$. Moreover, $\mathbf{C}$ is causal, i.e. $\mathbf{C}(t)|_{t<t_0} = \mathbf{0}$. In the following paragraphs on anisotropy and visco-elastic dissipation we will specify the components of $\mathbf{C}$ and their time dependence.

Anisotropy is the dependence of the elastic tensor on the orientation of the coordinate system. It plays a major role in the Earth where it translates to the dependence of elastic wave speeds on the propagation direction and the polarisation vector [10]. We decided to implement anisotropy with radial symmetry axis that is particulary pronounced in the crust and upper mantle of the Earth [11]. For such a medium, there are 5 independent elastic tensor components that are different from zero. Due to the symmetry of the elastic tensor they can be summarised in two 3×3 schemes [10]:

$$\begin{pmatrix} C^{(rrrr)} & C^{(rr\phi\phi)} & C^{(rr\theta\theta)} \\ C^{(\phi\phi rr)} & C^{(\phi\phi\phi\phi)} & C^{(\phi\phi\theta\theta)} \\ C^{(\theta\theta rr)} & C^{(\theta\theta\phi\phi)} & C^{(\theta\theta\theta\theta)} \end{pmatrix} = \begin{pmatrix} \lambda+2\mu & \lambda+C & \lambda+C \\ \lambda+C & \lambda+2\mu+A & \lambda+A \\ \lambda+C & \lambda+A & \lambda+2\mu+A \end{pmatrix} \tag{5a}$$

$$\begin{pmatrix} C^{(\phi\theta\phi\theta)} & C^{(\phi\theta r\theta)} & C^{(\phi\theta r\phi)} \\ C^{(r\theta\phi\theta)} & C^{(r\theta r\theta)} & C^{(r\theta r\phi)} \\ C^{(r\phi\phi\theta)} & C^{(r\phi r\theta)} & C^{(r\phi r\phi)} \end{pmatrix} = \begin{pmatrix} \mu & 0 & 0 \\ 0 & \mu+B & 0 \\ 0 & 0 & \mu+B \end{pmatrix} \tag{5b}$$

All components of $\mathbf{C}$ that do not appear in equation (5) are equal to zero. In the purely isotropic case the elastic parameters A, B and C vanish and the elastic tensor components take the simple form $C_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu\delta_{ik}\delta_{jl} + \mu\delta_{il}\delta_{jk}$, where λ and μ are the Lamé parameters.

Visco-elastic dissipation is responsible for the loss of elastic energy during the propagation of seismic waves and it is expressed mathematically through the time dependence of the elastic tensor $\mathbf{C}$. To specify this time dependence, we roughly follow [12]: Let σ, C and ϵ be representatives of some particular components of $\boldsymbol{\sigma}, \mathbf{C}$ and $\boldsymbol{\epsilon}$, respectively. Then a scalar version of the constitutive relation (4) is given by

$$\dot{\sigma}(t) = \int_{t'=t_0}^{t} \dot{C}(t - t')\dot{\epsilon}(t')\, dt'\,. \tag{6}$$

The spatial dependence has been omitted in the interest of clarity. We choose C to be of the form

$$C(t) = C_r \left[1 + \frac{\tau}{N}\sum_{p=1}^{N} e^{-t/\tau_{\sigma p}}\right] H(t)\,, \tag{7}$$

where H denotes the Heaviside function. The dimensionless parameter τ, the relaxed modulus C_r and the stress relaxation times $\tau_{\sigma p}$ with $p = 1, ..., N$, are material-specific properties. Differentiating (7) with respect to t and introducing the result into equation (6) yields

$$\dot{\sigma}(t) = C_r(1+\tau)\,\dot{\epsilon}(t) + C_r \sum_{p=1}^{N} M_p\,, \tag{8}$$

where the memory variables M_p are defined by

$$M_p := -\frac{\tau}{N\tau_{\sigma p}} \int_{t_0}^{t} e^{-(t-t')/\tau_{\sigma p}}\, H(t-t')\,\dot{\epsilon}(t')\,dt'\,. \tag{9}$$

The differentiation of (9) with respect to time yields a set of first-order differential equations for the memory variables:

$$\dot{M}_p = -\frac{\tau}{N\tau_{\sigma p}}\,\dot{\epsilon} - \frac{1}{\tau_{\sigma p}}\,M_p\,. \tag{10}$$

Hence, anelasticity can be modelled by solving N ordinary differential equations for the memory variables M_p simultaneously with the momentum equation and a slightly modified stress-strain relation. Generalising equations (8) and (10) to the three-dimensional elastic case results in a set of ODEs that can be written in matrix form:

$$\dot{\boldsymbol{\sigma}} = \mathbf{C}_1 : \dot{\boldsymbol{\epsilon}} + \mathbf{C}_2 : \sum_{p=1}^{N} \mathbf{M}_p\,, \quad \dot{\mathbf{M}}_p = \mathbf{T}_{p1} : \boldsymbol{\epsilon} + \mathbf{T}_{p2} : \mathbf{M}_p\,. \tag{11}$$

The components of the matrices $\mathbf{C}_1, \mathbf{C}_2, \mathbf{T}_{p1}$ and $\mathbf{T}_{p2}$ are

$$C_1^{(ijkl)} = C_r^{(ijkl)}(1+\tau^{(kl)})\,,\; C_2^{(ijkl)} = C_r^{(ijkl)}\,,\; T_{p1}^{(ijkl)} = -\frac{\tau^{(ij)}}{N\tau_{\sigma p}^{(ij)}}\delta_{ik}\delta_{jl}\,,\; T_{p2}^{(ijkl)} = -\frac{\delta_{ij}\delta_{kl}}{\tau_{\sigma p}^{(ij)}}\,. \tag{12}$$

In most practical applications the time dependence of the elastic parameters A, B and C is neglected. The coupled system of equations (1) and (11) constitutes the elastic wave equation in a medium with polarisation anisotropy and visco-elastic dissipation. It describes the propagation of seismic waves in an Earth model that is characterised by the distributions of mass density, elastic moduli and a set of relaxation times.
Note that the representations of the constitutive relation in terms of a convolution (equation 4) or in terms of ODEs (equation 11) are fully equivalent. However, the ODE version is easier to solve numerically.

3 Spectral-element discretisation of the governing equations

One of the principal advantages of our proposed waveform inversion technique is that the complete wavefield in a heterogeneous Earth model is simulated with high accuracy. This yields tomographic images that are physically more consistent than images that are based on approximations such as ray theory, for example. The equations of motion, summarised in the previous section, can not be solved analytically or semi-analytically in Earth models with realistic lateral heterogeneities. We therefore rely on numerical solutions that we describe and assess in the following paragraphs.

3.1 Spatial discretisation

In the course of the past two decades several methods have been developed that allow us to solve the elastic wave equation in three-dimensional, heterogeneous media. They range from classical finite-difference schemes [13] to discontinuous Galerkin methods [14]. For wave propagation on global and continental scales the spectral-element method (SEM) has proven to be the optimal compromise between efficiency and accuracy [15][16].
The centrepiece of the SEM is the decomposition of the computational domain G − in our case a spherical section of the Earth − into N_e non-overlapping elements G_e, with $e = 1, ..., N_e$. Each element is mapped onto the reference cube $\Lambda = [-1, 1]^3$ via an invertible transformation $\boldsymbol{\chi}_e : G_e \rightarrow \Lambda$. Inside the reference cube each displacement field component $u^{(p)}$ is approximated by a polynomial expansion, using Lagrange polynomials of degree N, denoted by ℓ_i, with $i = 1, ..., N+1$:

$$u^{(p)}(\boldsymbol{\chi}, t) \approx \sum_{ijk=1}^{N+1} \bar{u}_{ijk}^{(p)}(t)\, \ell_i(\xi)\ell_j(\zeta)\ell_k(\eta)\,, \quad \boldsymbol{\chi} = (\xi, \zeta, \eta)^T \in \Lambda\,. \tag{13}$$

The collocation points of the Lagrange polynomials are the Gauss-Lobatto-Legendre (GLL) points, and the polynomial degree, N, is mostly chosen between 4 and 6. With the approximation (13), the elastic wave equation can be transformed into an ordinary algebro-differential equation for the polynomial coefficients $\bar{u}_{ijk}^{(p)}(t)$ that can be solved with standard numerical techniques. Starting with the term $\rho(\mathbf{x})\ddot{\mathbf{u}}(\mathbf{x}, t)$ that appears on the left-hand side of equation (1), we compute the Galerkin projection $\mathfrak{G}_{qrs}^e[\rho\ddot{u}^{(p)}]$ for the element with index e ($e = 1, ..., N_e$):

$$\begin{aligned}
\mathfrak{G}_{qrs}^e[\rho\ddot{u}^{(p)}] &= \int_\Lambda \rho(\boldsymbol{\chi})\ddot{u}^{(p)}(\boldsymbol{\chi}, t)\, \ell_q(\xi)\ell_r(\zeta)\ell_s(\eta)\, J_e(\boldsymbol{\chi})\, d^3\boldsymbol{\chi}\,, \\
&= \sum_{ijk=1}^{N+1} \int_\Lambda \rho(\boldsymbol{\chi})\ddot{\bar{u}}_{ijk}^{(p)}(t)\ell_i(\xi)\ell_j(\zeta)\ell_k(\eta)\ell_q(\xi)\ell_r(\zeta)\ell_s(\eta)\, J_e(\boldsymbol{\chi})\, d^3\boldsymbol{\chi}\,,
\end{aligned} \tag{14}$$

where J_e denotes the Jacobian of the transformation $\boldsymbol{\chi}_e : G_e \rightarrow \Lambda$. As in the following equations we replaced '$\approx$' by '=' in the interest of clarity. The integral in equation (14) can be approximated with the GLL quadrature rule. Invoking the cardinal interpolation property of the Lagrange polynomials, $\ell_i(\xi_j) = \delta_{ij}$, then results in the following expression:

$$\mathfrak{G}_{qrs}^e[\rho\ddot{u}^{(p)}] = \rho(\boldsymbol{\chi}_{qrs})\, w_q w_r w_s J_e(\boldsymbol{\chi}_{qrs})\, \ddot{\bar{u}}_{qrs}^{(p)}\,, \quad \boldsymbol{\chi}_{qrs} = (\xi_q, \zeta_r, \eta_s)^T\,. \tag{15}$$

The symbols w_i with $i = 1, ..., N+1$ denote the GLL quadrature weights. Upon using Gauss' theorem in conjunction with the Neumann condition (2) we derive the Galerkin projection for the p-component of the second term in equation (1):

$$\begin{aligned}
\mathfrak{G}_{qrs}^e[(\nabla\cdot\sigma)^{(p)}] = &- \sum_{n=1}^{3}\sum_{i=1}^{N+1} w_i w_r w_s \frac{\partial\xi}{\partial x^{(n)}}\dot{\ell}_q(\xi_i)\, \sigma_{np}(\boldsymbol{\chi}_{irs})\, J_e(\boldsymbol{\chi}_{irs}) \\
&- \sum_{n=1}^{3}\sum_{i=1}^{N+1} w_q w_i w_s \frac{\partial\zeta}{\partial x^{(n)}}\dot{\ell}_r(\zeta_i)\, \sigma_{np}(\boldsymbol{\chi}_{qis})\, J_e(\boldsymbol{\chi}_{qis}) \\
&- \sum_{n=1}^{3}\sum_{i=1}^{N+1} w_q w_r w_i \frac{\partial\eta}{\partial x^{(n)}}\dot{\ell}_s(\eta_i)\, \sigma_{np}(\boldsymbol{\chi}_{qri})\, J_e(\boldsymbol{\chi}_{qri})\,.
\end{aligned} \tag{16}$$

Finally, the Galerkin projection of the right-hand side of equation (1) is given by

$$\mathfrak{G}_{qrs}^e[f^{(p)}] = w_q w_r w_s f^{(p)}(\boldsymbol{\chi}_{qrs})\,. \tag{17}$$

The discretisation of the constitutive relation (11) follows exactly the same steps that we outlined above, and will therefore be omitted.
Most of the computational costs are due to the evaluation of equation (16). They can, however, be significantly reduced when the natural spherical coordinate system is used inside the elements. This is possible because the spherical section excludes the centre of the Earth and because it can always be rotated to a position where it is sufficiently far from the poles. Thus, the singularity at zero radius and infinitesimally small elements at the poles can be avoided. Moreover, the Earth is approximately radially symmetric. Not conforming the shape of the elements to lateral variations of the medium properties does therefore not introduce significant errors. Based on these considerations we can choose the transformation, $\boldsymbol{\chi}_e : G_e \to \Lambda$, that is defined by the following set of equations:

$$\theta = \theta_{e,\min} + \frac{1}{2}\Delta\theta_e(1+\xi)\,, \quad \phi = \phi_{e,\min} + \frac{1}{2}\Delta\phi_e(1+\zeta)\,, \quad r = r_{e,\min} + \frac{1}{2}\Delta r_e(1+\eta)\,, \tag{18}$$

where r, θ and ϕ denote radius, co-latitude and longitude, respectively. The geometry of this transformation is visualised in figure 1. Using (18) instead of a general transformation, allows us to significantly reduce the number of operations needed to evaluate the expression $\mathfrak{G}_{qrs}[(\nabla\cdot\boldsymbol{\sigma})^{(p)}]$, which is now given by

$$\begin{aligned}\mathfrak{G}_{qrs}[(\nabla\cdot\boldsymbol{\sigma})^{(p)}] = &-\frac{2J_e}{\Delta r_e} w_s w_r \sum_{i=1}^{N+1} w_i \dot{\ell}_q(\eta_i)\,[\sigma_{rp} r^2 \sin\theta]|_{\boldsymbol{\chi}_{qri}} - \frac{2J_e}{\Delta\phi} w_s w_q \sum_{i=1}^{N+1} w_i \dot{\ell}_r(\zeta_i)\,[\sigma_{\phi p} r]|_{\boldsymbol{\chi}_{qis}} \\ &- \frac{2J_e}{\Delta\theta} w_r w_q \sum_{i=1}^{N+1} w_i \dot{\ell}_s(\xi_i)\,[\sigma_{\theta p} r\,\sin\theta]|_{\boldsymbol{\chi}_{irs}}\,,\end{aligned} \tag{19}$$

with $J_e := \frac{1}{8}\Delta\theta_e\,\Delta\phi_e\,\Delta r_e$. Combining equations (15), (17), (19) and the discretisation of the constitutive relation (4) yields the space-discretised version of the momentum equation (1). It may symbolically be written in the form of a matrix equation:

$$\ddot{\bar{\mathbf{u}}}(t) = \int_{t=t_0}^{t_1} \mathbf{M}(t-t')\bar{\mathbf{u}}(t')\,dt' + \bar{\mathbf{f}}(t)\,, \quad \bar{\mathbf{u}}, \bar{\mathbf{f}} \in \mathbb{R}^{3N_e(N+1)^3}\,, \quad \mathbf{M} \in \mathbb{R}^{3N_e(N+1)^3 \times 3N_e(N+1)^3}\,, \tag{20}$$

where the vector $\bar{\mathbf{u}}$ comprises the $3N_e(N+1)^3$ polynomial coefficients defined in equation (13). The matrix $\mathbf{M}$ is a function of the model parameters, and the vector $\bar{\mathbf{f}}$ is the discrete right-hand side of the momentum equation. In the interest of clarity we decided to express the visco-elastic dissipation as a convolution and not in terms of ODEs involving memory variables, as we did in equation (11). Both formulations are, as shown above, equivalent.

3.2 Verification

To assess the accuracy of the spectral-element solutions, we compare them to semi-analytical solutions [17] that exist for radially symmetric Earth models. We choose a source located at 90 km depth near the Solomon Islands and a station in eastern Australia. This source-receiver configuration and the radially symmetric Earth model ak135 [18] are shown in figure 2.
For the numerical solution we used elements that are $0.5^\circ \times 0.5^\circ$ wide and 38 km deep. The Lagrange polynomial degree is 6. This allows us to accurately model waveforms with periods as low as 7 s. A comparison of the semi-analytic and spectral-element solutions is shown in figure 3. In general, the agreement between the two solutions is excellent. The difference seismogram, plotted as 25 times amplified dashed line, is smallest for the early-arriving compressional waves. Surface waves, appearing later in the seismograms, are more difficult to model due to their comparatively

short wavelengths. A remarkable detail are the well-modelled small-amplitude waveforms in the interval between 340 s and 500 s. They partly originate from reflections at the surface and the discontinuities located at 410 km and 660 km depth.

4 Seismic waveform tomography as a discrete non-linear optimisation problem

The objective of seismic tomography is to infer the Earth's structure from a set of seismograms $\mathbf{u}_0(\mathbf{x}^k, t)$ recorded at the receiver positions $\mathbf{x}^k$, with $k = 1, ..., N_r$. This inverse problem is ill-posed due to the scarcity and noise contamination of the data. Since analytical solutions are again not available, we set the seismic tomography in the form of a discrete optimisation problem. For this we assume that all structural parameters (density, elastic parameters and relaxation times) are summarised in an N_m-dimensional model vector $\mathbf{m} \in \mathbb{R}^{N_m}$. Our goal is to find $\mathbf{m}$ such that a suitably chosen objective functional − written as a time integral from $t = t_0$ to $t = t_1$ − is minimal:

$$\mathfrak{E}(\bar{\mathbf{u}}, \mathbf{u}_0) = \int_{t=t_0}^{t_1} \mathfrak{e}(\bar{\mathbf{u}}, \mathbf{u}_0)\, dt = \min\, ! \tag{21}$$

The objective functional quantifies the difference between the observed seismograms $\mathbf{u}_0$ and the corresponding synthetic seismograms $\bar{\mathbf{u}}$ that we computed by solving the equations of motion. To solve the non-linear minimisation problem iteratively we need (1) a physically meaningful definition of the misfit, $\mathfrak{E}$, between synthetic and observed seismograms and (2) an efficient algorithm for the computation of the gradient of $\mathfrak{E}$ with respect to the model parameters $\mathbf{m}$.

4.1 Definition of suitable waveform misfits

A physically reasonable definition of the misfit functional introduced in equation (21) is the centrepiece of any successful waveform tomography. The misfit functional should extract as much waveform information as possible while conforming to the following particularities and problems of continental-scale seismic waveform tomography: (1) Amplitudes of seismic waves are strongly and non-linearly affected by structure in the immediate vicinity of the receiver. Loose sediments, for example, typically increase the amplitudes of seismic waves and therefore mask amplitude variations that may be caused by the deep structure of the Earth. (2) The magnitudes of natural earthquakes are often not sufficiently well known. Amplitude discrepancies between synthetic and observed waveforms may therefore result from inaccurate knowledge of the source rather than being due to yet undiscovered Earth structure. (3) Relevant and robust information about the structure of the Earth is often contained in the phase of seismic waveforms with small amplitudes. The best known example are compressional waves − P waves in seismological terms − that have small amplitudes compared to shear waves or surface waves.
In conclusion, phase information that is reliably measurable and quasi-linearly related to Earth structure needs to be emphasised with respect to the less usable amplitude information. To achieve this goal, we propose to first compute time-frequency representations of both the observed seismograms and the synthetic seismograms via the transformation [19]

$$\tilde{g}(t, \omega) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(\tau) h^*(\tau - t)\, e^{-\mathbf{i}\omega\tau}\, d\tau\,, \tag{22}$$

where h^* is the complex conjugate of a window function. We choose h to be a Gaussian with a half width approximately equal to the dominant period of the observed seismograms. The complex

time-frequency representations of the N_r observed seismograms, $\tilde{u}^0(\mathbf{x}^k,t,\omega)$, and the synthetic seismograms, $\tilde{u}(\mathbf{x}^k,t,\omega)$, allow us to define the phase misfit

$$\mathfrak{E}_p(\bar{\mathbf{u}},\mathbf{u}_0)=\sum_{k=1}^{N_r}\int_{\mathbb{R}^2}W(t,\omega)[\phi(\mathbf{x}^k,t,\omega)-\phi_0(\mathbf{x}^k,t,\omega)]^2\,dt\,d\omega\,. \tag{23}$$

The phases $\phi(\mathbf{x}^k,t,\omega)$ and $\phi_0(\mathbf{x}^k,t,\omega)$ correspond to $\tilde{u}(\mathbf{x}^k,t,\omega)$ and $\tilde{u}^0(\mathbf{x}^k,t,\omega)$, respectively. The weighting function $W(t,\omega)$ allows us to emphasise those parts of the time-frequency seismograms that are known from experience to be particularly informative about Earth structure in the upper mantle.

4.2 The computation of the misfit gradient via the adjoint method

To solve the minimisation problem iteratively, using for example a conjugate-gradient algorithm, we need to compute the derivatives

$$\frac{\partial\mathfrak{E}(\bar{\mathbf{u}},\mathbf{u}_0)}{\partial m_i}=\int_{t=t_0}^{t_1}\frac{\partial\mathfrak{e}(\bar{\mathbf{u}},\mathbf{u}_0)}{\partial\bar{u}_k}\,\frac{\partial\bar{u}_k}{\partial m_i}\,dt\,,\quad i=1,...,N_m\,, \tag{24}$$

where the summation over repeated indeces is implicitly assumed. Equation (24) involves the derivatives of the displacement $\bar{u}_k$ with respect to the model parameters m_i, $i=1,...,N_m$. The finite-difference approximation of all partial derivatives $\partial\bar{u}_k/\partial m_i$ requires the solution of at least N_m forward problems. Since this is computationally too expensive, we compute $\partial\mathfrak{E}(\bar{\mathbf{u}},\mathbf{u}_0)/\partial m_i$ using the adjoint method [20][21][22]: Integrating over a test function $\bar{\mathbf{u}}^\dagger$ dotted with the derivative of equation (20) with respect to m_i provides the relation

$$\int_{t=t_0}^{t_1}\bar{u}_k^\dagger(t)\frac{\partial\ddot{\bar{u}}_k(t)}{\partial m_i}\,dt-\int_{t=t_0}^{t_1}\int_{t'=t_0}^{t_1}\left[\bar{u}_k^\dagger(t)\frac{\partial M_{kl}(t-t')}{\partial m_i}\bar{u}_l(t')+\bar{u}_k^\dagger(t)M_{kl}(t-t')\frac{\partial\bar{u}_l(t')}{\partial m_i}\right]dt\,dt'=0\,, \tag{25}$$

that may be added to equation (24):

$$\begin{aligned}\frac{\partial\mathfrak{E}(\bar{\mathbf{u}},\mathbf{u}_0)}{\partial m_i}=&\int_{t=t_0}^{t_1}\bar{u}_k^\dagger(t)\frac{\partial\ddot{\bar{u}}_k(t)}{\partial m_i}\,dt+\int_{t=t_0}^{t_1}\frac{\partial\bar{u}_k(t)}{\partial m_i}\left[\frac{\partial\mathfrak{e}(\bar{\mathbf{u}},\mathbf{u}_\circ)}{\partial\bar{u}_k}-\int_{t'=t_0}^{t_1}\bar{u}_k^\dagger(t')M_{kl}(t'-t)\,dt'\right]dt\\&-\int_{t=t_0}^{t_1}\int_{t'=t_0}^{t_1}\bar{u}_k^\dagger(t)\frac{\partial M_{kl}(t-t')}{\partial m_i}\bar{u}_l(t')\,dt\,dt'\,.\end{aligned} \tag{26}$$

Repeatedly integrating the first term on the right-hand side by parts and imposing the terminal conditions $\bar{\mathbf{u}}^\dagger|_{t=t_1}=\dot{\bar{\mathbf{u}}}^\dagger|_{t=t_1}=\mathbf{0}$, yields a modified version of equation (26):

$$\begin{aligned}\frac{\partial\mathfrak{E}(\bar{\mathbf{u}},\mathbf{u}_0)}{\partial m_i}=&\int_{t=t_0}^{t_1}\frac{\partial\bar{u}_k(t)}{\partial m_i}\left[\frac{\partial\mathfrak{e}(\bar{\mathbf{u}},\mathbf{u}_\circ)}{\partial\bar{u}_k}+\ddot{\bar{u}}_k^\dagger(t)-\int_{t'=t_0}^{t_1}\bar{u}_k^\dagger(t')M_{kl}(t'-t)\,dt'\right]dt\\&-\int_{t=t_0}^{t_1}\int_{t'=t_0}^{t_1}\bar{u}_k^\dagger(t)\frac{\partial M_{kl}(t-t')}{\partial m_i}\bar{u}_l(t')\,dt\,dt'\,.\end{aligned} \tag{27}$$

Furthermore imposing that $\bar{\mathbf{u}}^\dagger$ be a solution of the adjoint equation

$$\ddot{\bar{u}}^\dagger(t) = \int_{t'=t_0}^{t_1} \mathbf{M}^T(t'-t)\bar{\mathbf{u}}^\dagger(t')\,dt' - \nabla_{\bar{\mathbf{u}}}\mathfrak{e}(\bar{\mathbf{u}},\mathbf{u}_0)\,, \tag{28}$$

allows us to give an expression for $\partial\mathfrak{E}/\partial m_i$ that does not explicitly contain the unknown derivative $\partial\bar{\mathbf{u}}/\partial m_i$:

$$\frac{\partial\mathfrak{E}(\bar{\mathbf{u}},\mathbf{u}_0)}{\partial m_i} = -\int_{t=t_0}^{t_1}\int_{t'=t_0}^{t_1} \bar{u}_k^\dagger(t)\frac{\partial M_{kl}(t-t')}{\partial m_i}\bar{u}_l(t')\,dt\,dt'\,. \tag{29}$$

The computation of the complete gradient of $\mathfrak{E}$ therefore reduces to the solution of the adjoint equation (28) and the subsequent evaluation of (29). Both steps can be performed efficiently on modern parallel computers even for large three-dimensional problems [23].
With the derivative computed via equation (29) we can iteratively minimise the misfit $\mathfrak{E}(\bar{\mathbf{u}},\mathbf{u}_0)$ by using the iterative gradient method

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \mathbf{P}_n\cdot\nabla_{\mathbf{m}}\mathfrak{E}\,,\quad (\nabla_{\mathbf{m}}\mathfrak{E})_i = \frac{\partial\mathfrak{E}(\bar{\mathbf{u}},\mathbf{u}_0)}{\partial m_i}\,,\quad n=0,1,... \tag{30}$$

where $\mathbf{P}_n$ is a suitably chosen pre-conditioner. We stop the iteration when the remaining waveform misfits approach the estimated level of seismic noise.

5 Application to the imaging of the Earth's structure in the Australasian region

We are currently testing our newly developed waveform inversion technique with data collected on the Australian continent. The principal advantages of this choice are the availability of high-quality seismic recordings and the relatively even distribution of sources (earthquakes) along the tectonically actives zones surrounding the Indo-Australian plate. The data coverage is visualised in the left panel of figure 4.
To keep the model simple during the first iterations, we restrict our attention to isotropic and non-dissipative media where the components of the elastic tensor are given by

$$C_{ijkl}(\mathbf{x},t) = [\lambda(\mathbf{x})\,\delta_{ij}\delta_{kl} + \mu(\mathbf{x})\,\delta_{ik}\delta_{jl} + \mu(\mathbf{x})\,\delta_{il}\delta_{jk}]\,H(t)\,. \tag{31}$$

In equation (31) the symbols μ and λ denote the Lamé parameters. A preliminary model of the shear wave velocity distribution, $c_s(\mathbf{x}) = \sqrt{\mu(\mathbf{x})\rho^{-1}(\mathbf{x})}$, at 100 km depth is displayed in the right panel of figure (4). As initial model we used a smoothed version of a classical surface wave tomography [24]. Fast velocities appear mostly in the western and central parts of the Australian continent that formed more than 550 million years ago. Younger parts of the continent, located in the east, are marked by slow to intermediate velocities. Tectonically active regions, such as the subduction zones in the north and the east, are distinctively slow. A detailed interpretation of the tomographic images will become possible after further iterations and refinements.

Acknowledgments

The authors are grateful to the seismology groups at the Research School of Earth Sciences (ANU, Canberra, Australia) and the Institut de Physique du Globe (Paris, France) for their support. This work would not have been possible without the Elite Network of Bavaria, the German Academic Exchange Service, the Leibniz Rechenzentrum in Garching and the SPICE network. Many thanks also to Yann Capdeville, Peter Moczo and Stewart Fishwick for inspiring discussions.

References

[1] R. D. van der Hilst, S. Widiyantoro and E. R. Engdahl, Evidence for deep mantle circulation from global tomography, *Nature*, **386**, 578-584 (1997).

[2] H.-P. Bunge, C. R. Hagelberg and B. J. Travis, Mantle circulation models with variational data assimilation: inferring past mantle flow and structure from plate motion history and seismic tomography, *Geophysical Journal International*, **152**, 280-301 (2003).

[3] A. M. Cormack, Representation of a function by its line integrals, with some radiological applications, *Journal of Applied Physics*, **34**(9), 2722-2727 (1963).

[4] W. H. K. Lee and V. Pereyra, Mathematical introduction to seismic tomography, *in: Seismic Tomography: Theory and Practice (edited by H. M. Iyer and K. Hirahara)*, Chapman & Hall, London, 1993.

[5] S. P. Grand, Mantle shear structure beneath the Americas and surrounding oceans, *Journal of Geophysical Research*, **99**, 11,591-11,621 (1994).

[6] B. L. N. Kennett and A. Gorbatov, Seismic heterogeneity in the mantle - strong shear wave signature of slabs from joint tomography, *Physics of the Earth and Planetary Interiors*, **146**, 87-100 (2004).

[7] M. Panning and B. Romanowicz, A three-dimensional radially anisotropic model of shear velocity in the whole mantle, *Geophysical Journal International*, **167**, 361-379 (2006).

[8] B. L. N. Kennett, *The seismic wavefield: Volume 1. Introduction and theoretical development*, Cambridge University Press (2001).

[9] K. Aki and P. G. Richards, *Quantitative Seismology, 2nd edition*, University Science Books (2002).

[10] V. Babuska and M. Cara, *Seismic anisotropy in the Earth*, Kluwer Acad. Pub. (1991).

[11] A. M. Dziewonski and D. L. Anderson, Preliminary reference Earth model, *Physics of the Earth and Planetary. Interiors*, **25**, 297-356 (1981).

[12] J. O. A. Robertsson, J. O. Blanch and W. Symes, Viscoelastic finite-difference modeling, *Geophysics*, **59**(9), 1444-1456 (1994).

[13] J. Kristek and P. Moczo, Seismic-wave propagation in viscoelastic media with material discontinuities: A 3D fourth-order staggered-grid finite-difference modeling, *Bulletin of the Seismological Society of America*, **93**(5), 2273-2280 (2003).

[14] M. Dumbser, M. Käser and J. de la Puente, Arbitrary high-order finite volume schemes for seismic wave propagation on unstructured meshes in 2D and 3D, *Geophysical Journal International*, **171**(2), 665-694 (2007).

[15] E. Faccioli, F. Maggio, R. Paolucci and A. Quarteroni, 2D and 3D elastic wave propagation by a pseudo-spectral domain decomposition method, *Journal of Seismology*, **1**(3), 237-251 (1997).

[16] D. Komatitsch and J. Tromp, Spectral-element simulations of global seismic wave propagation, I: validation, *Geophysical Journal International*, **149**(2), 390-412 (2002).

[17] W. Friederich and J. Dalkolmo, Complete synthetic seismograms for a spherically symmetric earth by a numerical computation of the Green's function in the frequency domain, *Geophysical Journal International*, **122**, 537-550 (1995).

[18] B. L. N. Kennett, E. R. Engdahl and R. Buland, Constraints on seismic velocities in the Earth from traveltimes, *Geophysical Journal International*, **122**, 108-124 (1995).

[19] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press (1996).

[20] J. L. Lions, *Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles*, Dunod Gauthier-Villars (1968).

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

[22] A. Fichtner, H.-P. Bunge, H. Igel, The adjoint method in seismology - I. Theory, *Physics of the Earth and Planetary Interiors*, **157**, 86-104 (2006).

[23] J. Oeser, H.-P. Bunge and M. Mohr, Cluster design in the Earth sciences: Tethys, *in: High Performance Computing and Communications, Proceedings*, **4208**, 31-40 (2006).

[24] S. Fishwick, B. L. N. Kennett, A. M. Reading, Contrasts in lithospheric structure within the Australian craton - insights from surface wave tomography, *Earth and Planetary Science Letters*, **231**, 163-176 (2006).

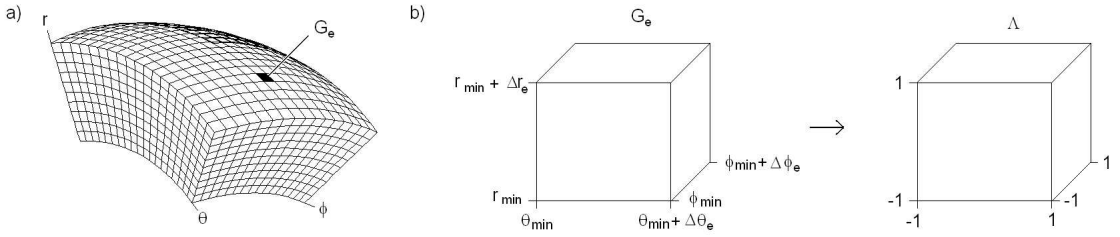

Figure 1: **Left**: Coordinate lines in a spherical section parameterised with the natural spherical coordinates r, θ, ϕ. **Right**: Illustration of the mapping from the element G_e to the reference cube Λ. See equation (18).

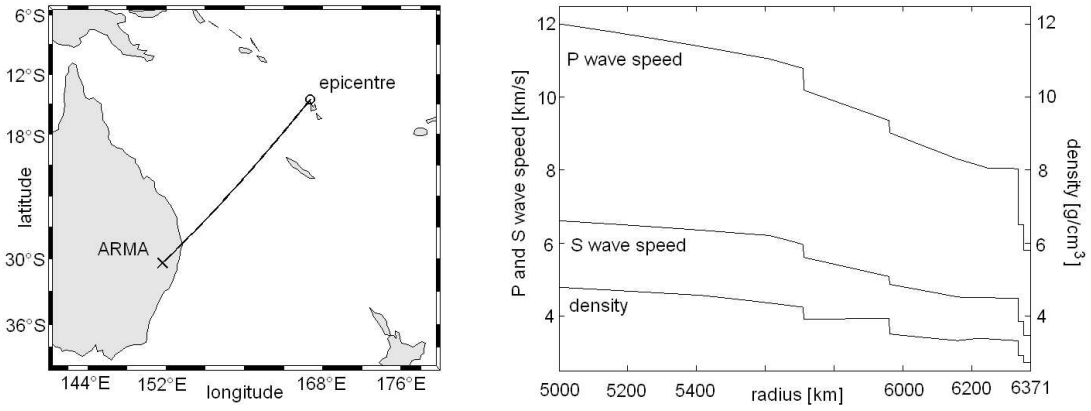

Figure 2: **Left:** Source-receiver configuration. The distance between the source ($\circ$) and station ARMA ($\times$) is $21.1°$. **Right:** Radially symmetric Earth model ak135 [18] parameterised in terms of density and the seismic P and S wave speeds.

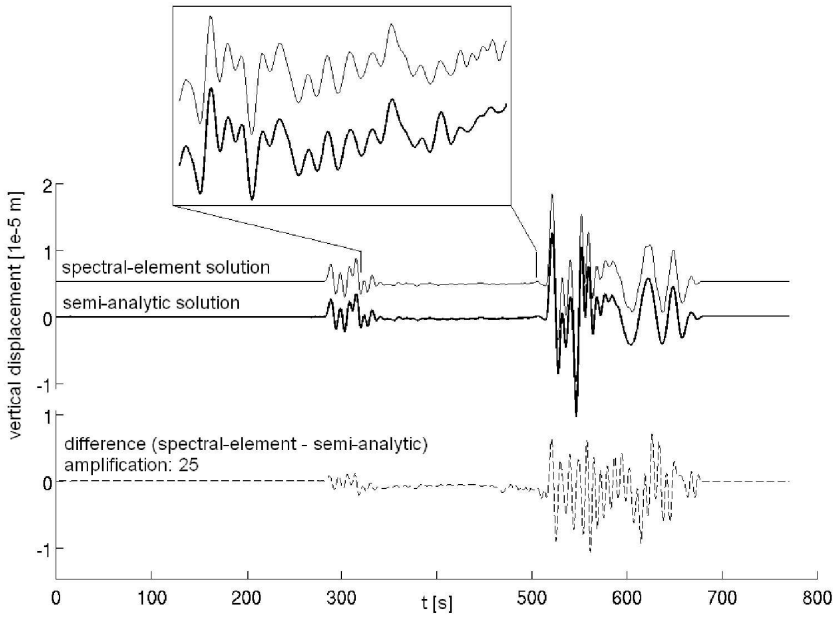

Figure 3: Comparison of the semi-analytic solution (bold line) with the spectral-element solution (thin line) for station ARMA, located at an epicentral distance of $21.1°$ from the epicentre. The difference between the two solutions, amplified by a factor of 25 to enhance visibility, is plotted as a dashed line below the seismograms. Also the magnified parts of the seismograms, between 340 s and 500 s, agree remarkably well.

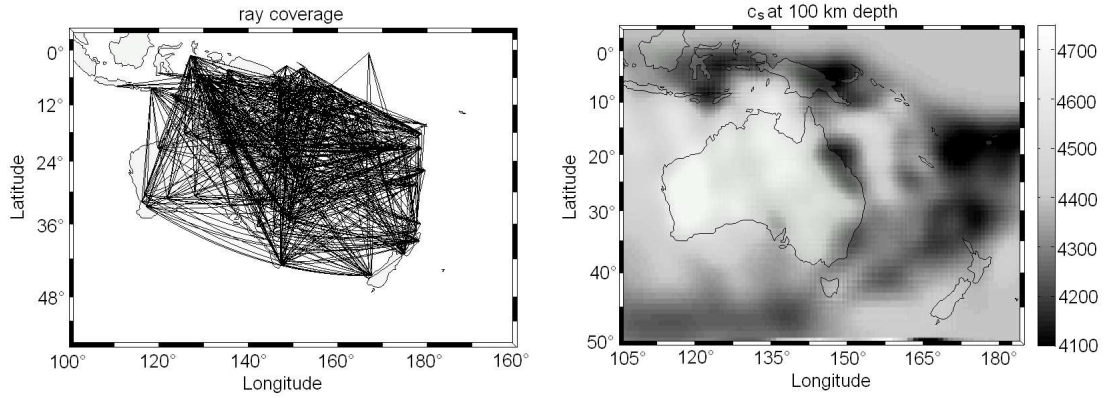

Figure 4: **Left**: Ray coverage (great circle segments between the sources and receivers used in this study). **Right**: Preliminary shear velocity model (c_s in m/s) at 100 km depth.