Available online at www.sciencedirect.com
SCIENCE DIRECT®

Physics of the Earth and Planetary Interiors 157 (2006) 105–123

PHYSICS OF THE EARTH AND PLANETARY INTERIORS

www.elsevier.com/locate/pepi

The adjoint method in seismology—II. Applications: traveltimes and sensitivity functionals

A. Fichtner*, H.-P. Bunge, H. Igel

Department of Earth and Environmental Sciences, Ludwig-Maximilians University Munich, Theresienstrasse 41, D-80333 Munich, Germany

Received 20 December 2005; received in revised form 17 March 2006; accepted 22 March 2006

Abstract

Sensitivity functionals which allow us to express the total derivative of a physical observable with respect to the model parameters, are defined on the basis of the adjoint method. The definition relies on the existence of Green's functions for both the original and the adjoint problem. Using the acoustic wave equation in a homogeneous and unbounded medium, it is shown that the first derivative of the wavefield u with respect to the model parameter $p = c^{-2}$ (c is the wave speed) does not contain traveltime information. This property also influences objective functions defined on u and in particular the least squares objective function. Therefore, a waveform inversion should either be complemented by a traveltime tomography or work with initially very long wavelengths that decrease in the course of the iteration. The definition of the sensitivity functionals naturally introduces waveform sensitivity kernels. Analytic examples are shown for the case of an isotropic, elastic and unbounded medium. In the case of a double couple source there are three classes of sensitivity kernels, one for each of the parameters λ, μ (Lamé parameters) and ρ (density). They decompose into P→P, P→S, S→P and S→S kernels and can be described by third-order tensors incorporating the radiation patterns of the original wavefield and the adjoint wavefield. An analysis of the sensitivity kernels for density suggests that a waveform inversion procedure should exclude either the direct waves or the source and receiver regions. S-wave sensitivity kernels, corresponding to conversions from or to S-waves, are larger than the kernels corresponding to P-waves only. This implies that S-wave residuals caused by parameter differences between the true Earth model and the numerical model will dominate a waveform inversion that does not account for that effect.

Keywords: Inversion; Fréchet derivative; Waveform analysis; Sensitivity

1. Introduction

Seismological observations play an important role in the imaging of the Earth's interior structure. Radial models of density and seismic velocities, based primarily on the arrival times of seismic waves at the surface, were developed by Dziewonski and Anderson (1981), Kennett and Engdahl (1991) and Kennett et al. (1995). The one-dimensional

* Corresponding author. Tel.: +49 89 2180 4215; fax: +49 89 2180 4205.
E-mail address: andreas.fichtner@geophysik.uni-muenchen.de (A. Fichtner)..

0031-9201/$ – see front matter
doi:10.1016/j.pepi.2006.03.018

models reflect the static increase of pressure and temperature with increasing depth and the associated phase transitions of mantle minerals. Moreover, they serve as reference models for maps of global three-dimensional variations of density and seismic velocities obtained on the basis of ray theory and finite mode summations (e.g. Dziewonski, 1984; Grand, 1994; Masters et al., 1996; Grand et al., 1997; Kennett and Gorbatov, 2004). Today, it is possible to simulate the propagation of seismic waves in realistic Earth models numerically (e.g. Igel et al., 1995; Komatitsch et al., 2000). This suggests that we replace ray theory and finite normal mode summations as forward models in inversion schemes by numeric solutions of the wave equation, thus, substantially increasing the amount of exploitable information.

The inversion of seismic data can be based on different concepts. In probabilistic inverse theory the solution of an inverse problem is defined as a marginal probability distribution in the model parameter space P (Tarantola, 1987). This definition is very general and elegant. However, a complete characterisation of this probability distribution usually requires us to evaluate a large number of parameter space elements $\mathbf{p} \in P$. Therefore, the probabilistic approach becomes impractical if the model space is large and if the solution of the forward problem is time-consuming. As an alternative, the solution of an inverse problem may be defined as the parameter set $\mathbf{p}_{\min}$ that minimises an objective function $\mathfrak{E}(\mathbf{u}(\mathbf{p}))$ defined on the observable $\mathbf{u}(\mathbf{p})$. In symbols: $\mathbf{p}_{\min} = \{\mathbf{p} \in P; \mathfrak{E}(\mathbf{u}(\mathbf{p})) = \min\}$. Possible observables are arrival times or waveforms. Typical parameters are seismic velocities and density. In the context of the minimisation approach the sensitivity plays a central role. It is defined as the first derivative of the observable $\mathbf{u}(\mathbf{p})$ with respect to its parameters $\mathbf{p}$ and is denoted by $D_p\mathbf{u}$. The sensitivity is the basis of resolution studies and minimisation algorithms. Moreover, it influences the choice of data that one wishes to include in the inversion procedure.

Often the computation of $D_p\mathbf{u}$ or $D_p\mathfrak{E}(\mathbf{u})$ by means of classical finite differencing techniques is impractical due to the large number of parameters in realistic models. A solution of this problem is the adjoint method (Tarantola, 1984, 1988). It allows us to obtain the first derivative by solving the original forward problem and its adjoint problem only once, therefore being very efficient. As a consequence of the continuously increasing computation power the adjoint method has received much attention in many geo-scientific fields, such as for example meteorology (Talagrand and Courtier, 1987), geodynamics (Bunge et al., 2003), seismology (Crase et al., 1990; Igel et al., 1996; Tromp et al., 2005) or groundwater modelling (Sun, 1994).

For any application that employs the derivative $D_p\mathbf{u}$ it is important to know the type of information that it contains. This is particularly true for inversion algorithms that exclusively rely on $D_p\mathfrak{E}(\mathbf{u})$ and therefore on $D_p\mathbf{u}$. Gauthier et al. (1986) used such an algorithm (the method of steepest descent) for the purpose of acoustic waveform inversion in two dimensions. They discovered that the derivative of the least squares objective function (the time integral over the squared waveform residuals) primarily contains information about where the edges of parameter perturbations are located. However, the interior of their famous ‘Camembert’ shaped bulk modulus perturbation could hardly be reconstructed during the inversion.

A particularly interesting aspect of the adjoint method is that it allows us to find analytic expressions for the different sensitivities (one for each parameter) in terms of Green’s functions. This directly leads to the formulation of sensitivity kernels, i.e., volumetric sensitivity densities. Asymptotic approximations for such sensitivity kernels for body waves have been derived on the basis normal modes (Li and Tanimoto, 1993; Li and Romanowicz, 1995).

The objective of this study is the analysis of sensitivity functionals, expressing the sensitivity $D_p\mathbf{u}$ in terms of a functional that is independently linear in the source time function and the model perturbations. Based on the example of the acoustic wave equation in an unbounded and homogeneous medium, we show that the first derivative of the least squares objective function only contains information about scatterers and reflectors. Traveltime information about a signal with a given wavelength is not contained in the first derivative. A simple numerical example in two dimensions confirms this result, which has important implications for any seismological waveform inversion procedure that is founded on the gradient method alone. We find that even in the simple case of an unbounded, homogeneous, elastic medium, the waveform sensitivity kernels corresponding to density perturbations exhibit surprising complexity. Their shape depends on the source orientation, the observation direction and the type of wave considered. In contrast to numerically computed kernels the analytic kernels automatically decompose into P→P, P→S, S→P and S→S kernels, meaning that they are sensitivity densities for different types of conversions and therefore for different types of waveform residuals. The results derived from an analysis of the sensitivity kernels in such a simplified scenario yield important information concerning the types of waves that one should consider for a waveform inversion.