9th Workshop on Three-Dimensional Modelling of Seismic Waves Generation, Propagation and their Inversion

22 September - 4 October, 2008

Upper mantle Anisotropy from Surface Wave studies

Jean-Paul Montagner
Seismological Laboratory
Institut de Physique du Globe
Paris
France

jpm@ipgp.jussieu.fr
Upper mantle Anisotropy from Surface Wave studies

Jean-Paul Montagner\(^{(1)}\)

\(^{(1)}\) Seismological Laboratory, CNRS UMR 7154, Institut de Physique du Globe, Paris, France

Modified version of the Chapter entitled "Upper mantle Structure: Global isotropic and anisotropic elastic tomography" in Treatise on Geophysics- Volume 1- Seismology and Structure of the Earth

Paris, 19 September 2008

Abstract

Major advances in Structural Seismology during the last twenty years, are related to the emergence and development of more and more sophisticated 3D imaging techniques, usually named seismic tomography, at different scales from local to global. Progress has been made possible by the rapid developments in seismic instrumentation and by the extensive use of massive computation facilities. The scope of this chapter is limited to the tomographic elastic structure of the upper mantle. In order to obtain a good spatial coverage of this part of the Earth, it is necessary to make use of dispersive properties of surface waves. Most global tomographic models are still suffering severe limitations in lateral resolution, due to the imperfect data coverage, and to crude theoretical approximations. It is usually assumed that the propagating elastic medium is isotropic, which is a poor approximation. It is shown in this chapter how to take account of anisotropy of Earth's materials and a complete tomographic technique including the resolution of the forward problem and of the inverse problem is presented. Consequently, by including other geological constraints, it is possible to map not only the 3D temperature heterogeneities but also the flow field within the convecting mantle. In order to improve the lateral resolution of global models, the installation of ocean bottom observatories is necessary and constitutes a new challenge for this new century. The next step will also consist in systematically applying recent developments in numerical modeling and theory to seismic data, in order to use the complete information provided by seismic waveforms and to incorporate physical and chemical constraints provided by other fields of earth sciences.

Keywords: Tomography, isotropic model, seismic anisotropy, S-wave splitting, surface waves, Rayleigh wave, Love wave, broadband seismology, normal modes, inverse problem, upper mantle, lithosphere, asthenosphere, transition zone.
Introduction

Twenty years ago, the first global isotropic tomographic models of the mantle were published (Woodhouse and Dziewonski, 1984; Dziewonski, 1984). Since that time, many new tomographic models were published, and a large family of techniques was made available. This important progress was made possible by the extensive use of computers which can handle very large datasets and by the availability of good quality digital seismograms recorded by broadband seismic networks such as GEOSCOPE (Romanowicz et al., 1984), IRIS (Smith, 1986) and all networks coordinated by the FDSN (Federation of Digital Seismograph Networks; Romanowicz and Dziewonski, 1986). Thanks to the installation of modern digital networks, it is now possible to map the whole earth from the surface down to the center of the Earth by seismic tomography. However, most tomographic techniques only make use of travel times or phase information in seismograms and very few use the amplitude, even when seismic waveforms are used (Woodhouse and Dziewonski, 1984; Li and Romanowicz, 1996). Global tomographic models have been improved over years by an increase in the number of data and more importantly by using more general parameterizations, now including anisotropy (radial anisotropy in Nataf et al. (1986); general slight anisotropy in Montagner and Tanimoto (1990, 1991)) and to a lesser extent anelasticity (Tanimoto, 1989; Romanowicz, 1990). This chapter is focused on the imaging of large scale (>1000km) lateral heterogeneities of velocity and anisotropy in the upper mantle (0-660km depth) where the lateral resolution is the best thanks to surface waves providing an almost uniform lateral and azimuthal coverage, particularly below oceanic areas. We will discuss how tomographic imaging completely renewed our vision of upper mantle dynamics. It makes it possible to relate surface geology and plate tectonics to underlying mantle convection, and to map at depth the origin of geological objects such as continents, mountain ranges, slabs, ridges and plumes. The goal of this chapter is not to review all contributions to this topic, but to underline the main scientific issues, to present different approaches and to illustrate the different progress (partly subjectively) by some of my results or by other more recent models. This chapter aims to show why a major step, which takes a complete account of amplitude anomalies in the most general case and which will enable to map shorter scale heterogeneities, is now possible and presently ongoing.

1 Effects of seismic velocity and anisotropy on seismograms

For theoretical and practical reasons, the Earth was considered for a long time, as composed of isotropic and laterally homogeneous layers. While an isotropic elastic medium can be described by 2 independent
elastic parameters ($\lambda$ and $\mu$ Lamé parameters), the cubic symmetry requires 3 parameters, but the most
commonly used anisotropic medium (transverse isotropy with vertical symmetry axis) necessitates 5
independent parameters (Love, 1927; Anderson, 1961) and the most general elastic medium requires
21 independent parameters. However, since the sixties, it was recognized that most parts of the Earth
are not only laterally heterogeneous but also anisotropic. Though the lateral heterogeneities of seismic
velocities were used for a long time for geodynamical applications, the importance of anisotropy for
understanding geodynamic processes has only been recognized recently.

Seismology is an observational field based on the exploitation of seismic recordings of the displacement
(velocity or acceleration) of the Earth induced by earthquakes. Broadband 3-component high dynamic
seismometers have been installed in more than 500 stations around the world during the last 20 years
(see overview of Romanowicz and Dziewonski, this issue). Thanks to progress in instrumentation and
theoretical developments, it is now possible to observe and to take a simultaneous account of the
effects of lateral heterogeneities of velocity and anisotropy on seismograms.

1.1 First order perturbation theory

The basic equation which governs the displacement $u(r, t)$ is the elasto-dynamics equation:

$$\rho_0 \frac{d^2 u_i}{dt^2} = \sum_j \sigma_{ij,j} + F_{li} + F_{Ei}$$

(1)

$F_{li}$ et $F_{Ei}$ represent respectively the whole ensemble of applied inertial and external forces (see
Takeuchi and Saito (1972) or Woodhouse and Dahlen (1978) for a complete description of all terms).

Generally, by neglecting the advection term, this equation is written in a simple way:

$$(\rho_0 \partial_{tt} - H_0) u(r, t) = F(r_S, t)$$

(2)

where $H_0$ is an integro-differential operator and $F$ expresses all forces applied to the source volume in
$r_S$ at time $t$ (considered as external forces). $F$ is assumed to be equal to 0 for $t < 0$. In the elastic
case, there is a linear relationship between $\sigma_{ij}$ and the strain tensor $\epsilon_{kl}$. $\sigma_{ij} = \sum_k \Gamma_{ijkl} \epsilon_{kl}$ (+ terms
related to the initial stress). $\Gamma_{ijkl}$ is a 4th-order tensor, often written in its condensed form $C_{ij}$ as a
6x6 matrix. By using the different symmetry conditions $\Gamma_{ijkl} = \Gamma_{jikl} = \Gamma_{ijk} = \Gamma_{klij}$, the tensor $\Gamma$
is shown to have 21 independent elastic moduli in the most general anisotropic medium. In an isotropic
medium, this number reduces to 2, the Lamé coefficients $\lambda$ and $\mu$. 

3
When solving for the free oscillations of the Earth $F = 0$. The solution $u(r, t)$ of equation (2) can be calculated for a spherically symmetric non rotating reference Earth model associated with the operator $H_0$, according to the equation:

$$\rho_0 \partial^2_{tt} u(r, t) = H_0 u(r, t) \quad (3)$$

The solution of equation (3) is beyond the scope of this chapter and is described elsewhere in the treatise (Woodhouse). The eigenvalues of the operator $H_0$ are equal to $-\rho_0 n \omega^2_\ell$ where $n \omega_\ell$ is the eigenfrequency characterized by 2 quantum numbers $n$ and $\ell$, respectively termed radial and angular orders. The corresponding eigenfunctions $n^m_u(r, t)$ depend on 3 quantum numbers $n, \ell, m$, where $m$ is the azimuthal order, with the following property $-\ell \leq m \leq \ell$. Therefore, for a given eigenfrequency $n \omega_\ell$ calculated in a spherically symmetric Earth model, $2\ell + 1$ eigenfunctions can be defined. The eigenfrequency $n \omega_\ell$ is said to be degenerate, with a degree of degeneracy $2\ell + 1$. There is a complete formal similarity with the calculation of the energy levels of the atom of hydrogen in quantum mechanics. The eigenfunctions $n^m_u(r, t)$ of the operator $H_0$ are orthogonal and normalized.

The important point is that the basis of functions $n^m_u(r, t)$ is complete. This implies that any displacement at the surface of the Earth can be expressed as a linear combination of these eigenfunctions:

$$u(r, t) = \sum_{n, \ell, m} n^m_u(r, t)$$

Therefore, these eigenfunctions can be used to calculate the synthetic displacement at any point $r$, at time $t$, due to a force system $F$ in the source volume. For a point force $F$ at point $r_S$, a step time function and its associated moment tensor $M$, which is a good starting model for earthquakes, the solution of the equation (2) is given by (Gilbert, 1971):

$$u(r, t) = \sum_{n, \ell, m} n^m_u(r, t) \frac{1 - \cos n \omega_\ell t}{n^2 \omega^2_\ell} e^{-\frac{n \omega_\ell t}{\alpha}} M : n^m_u(r, t)$$

where \(\epsilon\) is the deformation tensor. Since equation (4) is linear in $M$, it can be easily generalized to more complex spatial and temporal source functions, and can be rewritten:

$$u(r, t) = G(r, r_S, t, t_S)M(r_S, t_S)$$

where $G(r, r_S, t, t_S)$ is the Green operator of the medium. Normal mode theory is routinely used to calculate synthetic seismograms at long periods ($T \geq 40s.$) and Centroid Moment Tensor solutions (Dziewonski et al., 1981).
An example of real and synthetic seismograms is presented in figure 1. However, there are still some discrepancies (usually frequency dependent) between the observed and synthetic seismograms. The simplest way to explain the observed phase shifts (time delays) is to remove the assumption that the Earth is spherically symmetric, i.e. there are lateral heterogeneities between the source and the receiver. The next step is to characterize these lateral heterogeneities. Since the agreement between synthetic and observed seismograms is good at long periods \((T \geq 40\, \text{s})\), we can reasonably infer that the amplitude of heterogeneities is small \(< 10\%\). Behind the surface wave train, a long coda is usually observed, interpreted as scattered waves. However, when filtering out periods shorter than 40s, this coda vanishes, which means that the scattering effect is only large in the shallowest regions of the Earth (primarily the crust, and the upper lithosphere) but that it is probably negligible at larger depths. However, some groups are starting to use the information contained in these coda waves \((\text{Aki and Richards}, 1980; \text{Snieder et al.}, 2002)\), and even from seismic noise \((\text{Shapiro et al.}, 2005)\) for imaging the crust. For the sake of simplicity, our study is limited to long period surface waves and it is hypothesized that the scale of lateral heterogeneities is large compared with the seismic wavelength. This point will be discussed in section 2.1. A second hypothesis which must be discussed, is the isotropic nature of the Earth materials. Actually, it is a poor assumption, because seismic anisotropy can be unequivocally observed at different scales. Finally, the influence of lateral variations in attenuation must also be taken into account and will be discussed in another chapter \((\text{Seismic Attenuation in the Earth}, \text{Romanowicz and Mitchell}, \text{this volume})\).

1.2 Effect of anisotropic heterogeneities on normal modes and surface waves

Different geophysical fields are involved in the investigation of the manifestations of anisotropy of Earth materials: mineral physics and geology for the study of the microscopic scale, and seismology for scales larger than, typically, one kilometer. The different observations related to anisotropy, at different scales are reviewed in \textit{Montagner} (1998) and in \textit{Park and Maupin} (this volume).

Different kinds of observations have been used for investigating anisotropy in the upper mantle: the Rayleigh-Love wave discrepancy \((\text{Anderson, 1961})\), the azimuthal variation of phase velocities of surface waves \((\text{Forsyth, 1975})\) and the shear-wave splitting particularly for SKS waves \((\text{Vinnik et al.}, 1992)\). The lack of stations in oceanic areas explains why it is necessary to use surface waves to investigate upper mantle structure (isotropic or anisotropic) at the global or regional scales.

In the simplest case (fundamental modes, no coupling between branches of Rayleigh and Love waves), the frequency shift \(\frac{\delta \omega}{\omega}\) (and the corresponding phase velocity perturbation \(\frac{\delta V}{V}\)), for a constant wavenum-
ber $k$ can be written by applying Rayleigh’s principle:
\[
\frac{\delta \omega}{\omega} \bigg|_k = \frac{\delta V}{V} \bigg|_k = \frac{1}{2\omega} \int_{\Omega} \epsilon_{ij}^s \delta\Gamma_{ijkl} \epsilon_{kl} \, d\Omega
\]
where $\epsilon_{ij}$ and $\delta\Gamma_{ijkl}$ are respectively the deformation and the deviations of elastic tensor components from a SNREI model (spherically symmetric, non rotating, elastic, isotropic), and $\epsilon_{ij}^s$ the eigenfunctions as defined in the previous section.

We only consider the propagation of surface waves in a plane-layered medium for a general slight elastic anisotropy, but it can be easily extended to the spherical Earth (Mochizuki, 1986; Tanimoto, 1986; Romanowicz and Snieder, 1988; Larson et al., 1998; Trampert and Woodhouse, 2003). Smith and Dahlen (1973, 1975), found that, to first order in anisotropy and at frequency $\omega$, the azimuthal variation of local phase velocity (Rayleigh or Love wave) can be expanded as a Fourier series of the azimuth $\Psi$ along the path and is of the form:

\[
V(\omega, \theta, \phi, \Psi) - V_0(\omega, \Psi) = \alpha_0(\omega, \theta, \phi) + \alpha_1(\omega, \theta, \phi)\cos 2\Psi + \alpha_2(\omega, \theta, \phi)\sin 2\Psi
\]
\[+ \alpha_3(\omega, \theta, \phi)\cos 4\Psi + \alpha_4(\omega, \theta, \phi)\sin 4\Psi
\]
(6)

where $V_0(\omega, \Psi)$ is the reference velocity of the unperturbed medium, and $\Psi$ is the azimuth along the path with respect to the North direction. Montagner and Nataf (1986) present the expressions for the different azimuthal coefficients $\alpha_i(\omega, \theta, \phi)$ as depth integral functions dependent on 13 simple linear combinations of standard cartesian elastic coefficients $C_{ij}$. The Appendix shows how to relate $\Gamma_{ijkl}$ to $C_{ij}$ and presents detailed calculation of azimuthal terms for Love waves in the geographical coordinate system.

Constant term ( 0 $\Psi$-azimuthal term: $\alpha_0$)
\begin{align*}
A &= \rho V_{PH}^2 = \frac{3}{8}(C_{11} + C_{22}) + \frac{1}{4}C_{12} + \frac{1}{2}C_{66} \\
C &= \rho V_{PV}^2 = C_{33} \\
F &= \frac{1}{2}(C_{13} + C_{23}) \\
L &= \rho V_{SV}^2 = \frac{1}{2}(C_{44} + C_{55}) \\
N &= \rho V_{SH}^2 = \frac{1}{8}(C_{11} + C_{22}) - \frac{1}{4}C_{12} + \frac{1}{2}C_{66}
\end{align*}

2 $\Psi$-azimuthal term:
\begin{align*}
\alpha_1 \cos 2\Psi & \quad \alpha_2 \sin 2\Psi \\
B_c &= \frac{1}{2}(C_{11} - C_{22}) & B_s &= C_{16} + C_{26} \\
G_c &= \frac{1}{2}(C_{55} - C_{44}) & G_s &= C_{54} \\
H_c &= \frac{1}{2}(C_{13} - C_{23}) & H_s &= C_{36}
\end{align*}
4 \Psi\text{-azimuthal term:}

\[ E_c = \frac{1}{8}(C_{11} + C_{22}) - \frac{1}{4}C_{12} - \frac{1}{2}C_{66} \]

\[ E_s = \frac{1}{2}(C_{16} - C_{26}) \]

where indices 1 and 2 refer to horizontal coordinates (1: North; 2: East) and index 3 refers to vertical coordinate. \( \rho \) is the density, \( V_{PH}, V_{PV} \) are respectively horizontally and vertically propagating P-wave velocities, \( V_{SH}, V_{SV} \) horizontal and vertical polarized S-wave velocities. So, the different parameters present in the different azimuthal terms are simply related to elastic moduli \( C_{ij} \).

From a practical point of view, once source phase is removed and assuming that the scale of heterogeneities is larger than the wavelength, the total phase \( \phi_t \) (and the travel time) between the epicenter \( E \) and the receiver \( R \), is easily related to the measurement of phase velocity \( V_d(\omega) \), and therefore to the local phase velocity \( V(\omega, \theta, \phi, \Psi) \):

\[ \phi_t = \omega t_{E \rightarrow R} = \frac{\omega \Delta}{V_d(\omega)} = \omega \int_{E}^{R} \frac{ds}{V(\omega, \theta, \phi, \Psi)} \]  

Therefore, equations (6) and (7) define the forward problem in the framework of first order perturbation theory. We will see in the next section how to solve the inverse problem. That means that, ideally, surface waves in the plane case have the ability to provide information on 13 elastic parameters, which emphasizes the enormous potential of surface waves in terms of geodynamical and petrological implications. There are only 13 elastic moduli among 21, since propagation of surface waves is invariant against rotation by \( \pi \) which corresponds to a monoclinic symmetry.

The 0-\Psi term corresponds to the average over all azimuths and involves 5 independent parameters, A, C, F, L, N, which represent the equivalent transversely isotropic medium with a vertical symmetry axis (more simply named VTI or radial anisotropy). It must be noted that it is possible to retrieve the equivalent isotropic shear modulus from these 5 parameters. By using a Voigt average, the shear modulus \( \mu_{iso} \) is equal to:

\[ \mu_{iso} = \rho V_{S_{iso}}^2 = \frac{1}{15}(C_{11} + C_{22} + C_{33} - C_{12} - C_{13} - C_{23} + 3C_{44} + 3C_{55} + 3C_{66}) \]

According to the expressions of A, C, F, L, N in terms of elastic moduli, \( \mu_{iso} = \frac{1}{15}(C + A - 2F + 6L + 5N) \) So we can see that the equivalent isotropic velocity depends not only on \( V_{SV} \) and \( V_{SH} \), but also on P-wave velocity and anisotropy (\( \phi = \frac{F}{A} \)) and on \( \eta = \frac{F}{2(A - 2L)} \). By rewriting this expression \( \mu_{iso} = \frac{1}{15}(C + (1 - 2\eta)A + (6 + 4\eta)L + 5N) \), neglecting anisotropy in P-wave (\( \phi = 1 \)) and assuming \( \eta = 1 \), it is found that \( \mu_{iso} = \rho V_{S_{iso}}^2 \approx \frac{2}{3}L + \frac{1}{3}N = \frac{2}{3}\rho V_{SV}^2 + \frac{1}{3}\rho V_{SH}^2 \). Naturally, this choice is partly arbitrary, since usually, there is no S-wave anisotropy without P-wave anisotropy. Another way might consist
in using correlations between anisotropic parameters for petrological models as derived by Montagner and Anderson (1989a).

The other azimuthal terms (2-Ψ and 4-Ψ) depend on 4 groups of 2 parameters, B, G, H, E respectively describing the azimuthal variation of A, L, F, N. These simple parameters make it possible to describe in a simple way the two seismically observable effects of anisotropy on surface waves, the ”polarization” anisotropy (Schlue and Knopoff, 1977) and the azimuthal anisotropy (Forsyth, 1975).

Another important point in these expressions is that they provide the partial derivatives for the radial and azimuthal anisotropy of surface waves. The corresponding kernels and their depth dependence are plotted in Montagner & Nataf (1986) (Figures A2-A3). These partial derivatives of the different azimuthal terms with respect to the elastic parameters can be easily calculated by using a radial anisotropic reference Earth model, such as PREM (Dziewonski and Anderson, 1981). The partial derivatives of the eigenperiod \( 0 T \ell \) with respect to parameter \( p \), \( \frac{p}{T} \frac{\partial T}{\partial p} \) can easily be converted into phase velocity partial derivatives by using:

\[
\frac{p}{V} \left( \frac{\partial V}{\partial p} \right)_T = - \frac{V}{U} \frac{p}{T} \left( \frac{\partial T}{\partial p} \right)_k
\]

For example, the parameters \( G_c \) and \( G_c \) have the same kernel as parameter \( L \) (related to \( V_{SV} \)) as shown by comparing the expressions of \( R_1 \), \( R_2 \) and \( R_3 \) in equation (29) of Appendix. For fundamental modes, the calculation of kernels shows that Love waves are almost insensitive to \( V_{SV} \) (Figure A2) and Rayleigh waves to \( V_{SH} \). Rayleigh waves are the most sensitive to \( SV \)-waves. However, as pointed out by Anderson and Dziewonski (1982), the influence of P-waves (through parameters \( A \) and \( C \)) can be very large in an anisotropic medium. The influence of density is also very large for Love and Rayleigh waves, but as shown by Takeuchi and Saito (1972), it is largely decreased when seismic velocities are inverted for, instead of elastic moduli and density.

### 1.3 Comparison between surface wave anisotropy and SKS splitting data

It can be noted that some of the linear combinations of elastic moduli \( C_{ij} \), derived from surface waves in the previous section, also come up when considering the propagation of body waves in symmetry planes for a weakly anisotropic medium (see for example Crampin et al. (1984) and their azimuthal dependence

\[
\rho V_{P}^2 = A + B_c \cos 2\Psi + B_s \sin 2\Psi + E_c \cos 4\Psi + E_s \sin 4\Psi
\]

\[
\rho V_{qSH}^2 = N - E_c \cos 4\Psi - E_s \sin 4\Psi
\]

\[
\rho V_{qSV}^2 = L + G_c \cos 2\Psi + G_s \sin 2\Psi
\]
where $V_{qSH}$ and $V_{qSV}$ corresponds respectively to quasi-SH and quasi-SV waves.

A global investigation of anisotropy inferred from SKS body wave splitting measurements (delay times and directions of maximum velocities) has been undertaken by different authors (Vinnik et al., 1992; Silver, 1996; Savage, 1999). Unfortunately, most SKS measurements have been done in continental parts of the Earth, and very few in oceans. It turns out that a direct comparison of body wave and surface wave datasets is now possible (Montagner et al., 2000). If the anisotropic medium is assumed to be characterized by a horizontal symmetry axis with any orientation (that is a very strong assumption which can be alleviated as shown by Chevrot et al., 2004), and for a vertically propagating SKS wave, a synthetic dataset of SKS delay times and azimuths can be calculated from the global distribution of anisotropy derived from surface waves, by using the following equations:

$$
\delta t_{SKS} = \int_0^h \frac{dz}{L(z)} \sqrt{\rho L(z)} [G_c(z) \cos(2\Psi(z)) + G_s(z) \sin(2\Psi(z))]$$

(8)

where $\delta t_{SKS}$ is the integrated travel time for the depth range $0-h$ for a propagation azimuth $\Psi$, where the anisotropic parameters $G_c(z), G_s(z)$ and $L(z)$ are the anisotropic parameters retrieved from surface waves at different depths. It is remarkable to realize that only the $G$-parameter (expressing the SV-wave azimuthal variation) is present in this equation. From equation (8), we can infer the maximum value of delay time $\delta t_{SKS}^{max}$ and the corresponding azimuth $\Psi_{SKS}$:

$$
\delta t_{SKS}^{max} = \sqrt{\left\{ \int_0^h \frac{dz}{L(z)} \sqrt{\rho L(z)} G_c(z) \right\}^2 + \left\{ \int_0^h \frac{dz}{L(z)} \sqrt{\rho L(z)} G_s(z) \right\}^2}
$$

(9)

$$
\tan(2\Psi_{SKS}) = \frac{\int_0^h \frac{dz}{L(z)} G_s(z)}{\int_0^h \frac{dz}{L(z)} G_c(z)}
$$

(10)

However, equation (8) is approximate and only valid when the wavelength is much larger than the thickness of layers. It is possible to make more precise calculations by using the technique derived for 2 layers by Silver & Savage (1994) or by using the general expressions given in Rumpker and Silver (1998), Montagner et al. (2000), Chevrot et al., 2004).

With equations (9) and (10), a synthetic map of the maximum value of delay time $\delta t_{SKS}^{max}$ can be obtained by using an 3D anisotropic surface wave model. A detailed comparison between synthetic SKS derived from AUM (Montagner & Tanimoto, 1991) and observed SKS (Silver, 1996) was presented in Montagner et al. (2000). Figure 2 shows such a map for the Earth centered in the Pacific, by using the anisotropic surface wave model of Montagner (2002) derived from the data of Montagner and Tanimoto (1991) and Ekström et al. (1997). First of all, the comparison shows that both datasets are
compatible in magnitude but not necessarily in directions. Some contradictions between measurements derived from surface waves and from body waves have been noted. The agreement of directions is correct in tectonically active areas but not in old cratonic zones. The discrepancy in these areas results from the rapid lateral change of directions of anisotropy at a small scale. These changes stem from the complex history of these areas, which have been built by successive collages of continental pieces. It might also result from the hypothesis of horizontal symmetry axis, which was shown to be invalid in many areas (Plomerova et al., 1996). The positive consequence of this discrepancy is that a small scale mapping of fossil anisotropy in such areas might provide clues for understanding the processes of growth of continents and mountain building opening a new field, the paleoseismology.

Unlike surface waves, SKS-waves have a good lateral resolution, and are sensitive to the short wavelength anisotropy just below the stations. But their drawback is that they have a poor vertical resolution. On the other hand, global anisotropy tomography derived from surface waves only provides long wavelength anisotropy (poor lateral resolution) but enables the location at depth of anisotropy. The long wavelength anisotropy derived from surface waves will display the same direction as the short wavelength anisotropy inferred from body waves only when large scale vertical coherent processes are predominant. As demonstrated by Montagner et al. (2000), the best agreement between observed and synthetic SKS can be found when only layers in the uppermost 200km of the mantle are taken into account. Moreover, tomographic models derived from surface waves looses resolution at depths greater than 200km. In some continental areas, short scale anisotropy, the result of a complex history, might be important and even might mask the large-scale anisotropy more related to present convective processes (see for example Marone and Romanowicz, 2006 for North America). From a statistical point of view, good agreement is found between orientations of anisotropy and plate velocity motion for fast-moving plates. The differences between anisotropy and tectonic plate directions are related to more complex processes, as will be seen in section 3.

2 Upper mantle tomography of seismic velocity and anisotropy

We now show how to implement theory of section 1 from a practical but general point of view, and how to design a tomographic technique in order to invert for the 13 different elastic parameters and density. A tomographic technique necessitates to solve simultaneously a forward problem and an inverse problem. By using the results of the previous section, it successively considers how to set the forward problem, and how it is used to retrieve a set of parameters by inversion.
2.1 Forward problem

First, it is necessary to define the data space \( d \) and the parameter space \( p \). It is assumed that a functional \( g \) relating \( d \) and \( p \) can be found such that:

\[
d = g(p),
\]

where \( d \) is the set of data (which samples the data space), and \( p \) the set of parameters.

**Data Space: \( d \)**

The basic dataset is made of seismograms \( u(t) \). We can try to directly match the waveform in the time domain, or we can work in the Fourier domain, by separating phase and amplitude on each component \( u_i(t) \):

\[
u_i(t) = \int_{-\infty}^{\infty} A_i(\omega)e^{i(\omega t - \phi_i)}d\omega
\]

The approach consisting in fitting seismic waveforms is quite general but, from a practical point of view, it does not necessarily correspond to the simplest choice. In a heterogeneous medium, the calculation of amplitude and phase effects makes it necessary to calculate the coupling between different multiplets (Li and Tanimoto, 1993; Li and Romanowicz, 1995; Marquering et al., 1996), which is very time consuming. When working in Fourier domain, different time windows can be considered and the phase of different seismic trains, body waves and surface waves can be separately matched (Nolet, 1990; Léveque et al., 1991) under drastic simplifying assumptions. Figure 1 showed an example of observed and synthetic seismograms, the latter obtained by normal mode summation with the different higher modes. The fundamental wavetrain is well separated from other modes at large epicentral distances. The part of the seismogram corresponding to higher modes is more complex and shows overlap of these modes in the time domain. Therefore, from a practical point of view, the fitting of the fundamental mode wavetrain will not cause any problem and has been widely used in global mantle tomography.

The use of higher mode wavetrains and the separation of overtones is much more difficult. The first attempts were performed by Nolet (1975), Cara (1978), Okal and Jo (1985) and Dost (1990) by applying a spatial filtering method. Unfortunately, all these techniques can only be applied in areas where dense arrays of seismic stations are present, i.e. in North America and Europe. By using a set of seismograms recorded at one station but corresponding to several earthquakes located in a small source area, Stutzmann and Montagner (1993) showed how to separate the different higher modes. A similar approach was also followed by Van Heijst and Woodhouse (1997). We only detail in this paper the technique which was designed for fitting the fundamental mode wavetrain and the reader is referred to Stutzmann and Montagner (1993, 1994), Van Heijst and Woodhouse (1997) and Beucler et al. (2003) for the description of the recovery of higher mode dispersion properties and to Romanowicz
(2002) for a general overview. Figure 3 presents an example of phase velocity dispersion for different surface wave modes (fundamental and first higher modes (Beucler et al., 2003)) and how they compare with previous investigations (Cara, 1978; Van Heijst and Woodhouse, 1997). We take advantage of the fact that, according to the Fermat’s principle, the phase velocity perturbation is only dependent to second order on path perturbations, whereas amplitude perturbation are dependent, to first order, on these perturbations, which implies that the eigenfunctions must be recalculated at each iteration. Therefore, the phase is a more robust observable than the amplitude. The amplitude $A(\omega)$ depends in a complex manner on seismic moment tensor, attenuation, scattering, focusing effects, station calibration and near-receiver structure whereas the contribution of lateral heterogeneities of seismic velocity and anisotropic parameters to the phase $\phi(\omega)$ can be easily extracted. The dataset under investigation, is composed of propagation times (or phase velocity measurements for surface waves) along paths: $d = \{V(\omega)\}$.

On the other hand, the phase of a seismogram at time $t$ is decomposed, as follows: $\phi = k \cdot r + \phi'_0$, where $k$ is the wave vector, $\phi'_0$ is the initial phase including several terms: $\phi'_0 = \phi_0 + \phi_S + \phi_I$, $\phi_S$ is the initial source phase, $\phi_0$ is related to the number of polar phase shifts, $\phi_I$ is the instrumental phase. $\phi$ can be measured on seismograms by Fourier transform. We usually assume that $\phi_S$ is correctly given by the centroid moment tensor solution. For a path between epicenter $E$ and receiver $R$ with an epicentral distance $\Delta$, the phase $\phi$ is given by:

$$\phi = \frac{\omega \Delta}{V_{\text{obs}}} + \phi_0 + \phi_S + \phi_I$$  \hspace{1cm} (11)

In the general case, we want to relate the observed phase velocity $V_{\text{obs}}(\omega)$ to the parameters of the Earth model $p(r, \theta, \phi)$. Data and 3D parameters can be related through integrals over the whole volume of the Earth. But for computing reasons, it is usual to use a multistep approach, where we first retrieve the local phase velocity $V(\omega, \theta, \phi)$ including its azimuthal terms, and then perform the inversion at depth. These 2 steps can be reversed since the order of the integrations can be reversed.

It is necessary to consider the nature of the perturbed medium. Following the approach of Snieder (1988), if the perturbed medium is at the same time smooth (long wavelength heterogeneities) and weak (small amplitude of heterogeneities), the geometrical optics approximation (and ray theory) applies. This hypothesis is not necessarily met within the Earth where some geological objects (slabs, mantle plumes, ...) have a length scale which can be close to the seismic wavelength. In the approximation of ray theory, the volume integral reduces to the curvilinear integral along the geometrical ray path.
When ray theory is applicable:

$$\phi - \phi'_0 = \frac{\omega \Delta V_{\text{obs}}(\omega)}{V_0(\omega)} = \int_{E}^{R} \frac{\omega ds}{V(\omega, \theta, \phi)}$$ (12)

Where the integral is evaluated along the ray path between the epicenter E and the receiver R.

Following the results of the previous section, different approximations are implicitly made when using this expression of the phase:

- large angular order $\ell \gg 1$, but not too large (scattering problems). From a practical point of view, this means that measurements are performed in the period range $40s < T < 200s$ with seismic wavelengths between 200 and 1000km.
- geometrical optics approximation: If $\lambda$ is the wavelength of the surface wave at period $T$, and $\Lambda_S$ the spatial wavelength of heterogeneity: $\Lambda_S \gg \lambda = VT \Rightarrow \Lambda_S \approx 2000$ km. Epicentral distance $\Delta$ must be larger than seismic wavelength.
- slight anisotropy and heterogeneity: $\frac{\delta V}{V} \ll 1$. According to Smith and Dahlen (1973) for the plane case, the local phase velocity can be decomposed as a Fourier series of the azimuth $\Psi$ (equation 6): Each azimuthal term $\alpha_i(T, \theta, \phi)$ of equation (6) can be related to the set of parameters $p_i(r, \theta, \phi)$ (density + 13 elastic parameters), according to the expressions derived in Appendix.

$$\frac{\Delta V_{\text{obs}}(T)}{V_0(T)} - \frac{\Delta V_{0}(T)}{V_0(T)} = -\sum_{j=0}^{2} \sum_{i=1}^{14} \int_{E}^{R} \frac{ds}{V_0} \int_{0}^{a} \left[ \left( \frac{p_i \partial V}{V \partial p_i} \right)_j \frac{\delta p_i(r, \theta, \phi)}{p_i} \cos(2j\Psi) \right] \frac{dz}{\Delta h}$$ (13)

Equation (13) defines the forward problem in the framework of first order perturbation theory, relating the data and the parameter spaces. This approach is usually named PAVA (Path average approximation). Many terms in equation (13) are equal to zero since all parameters are not present in each azimuthal term. A last important ingredient in the inverse problem formulation is the ”structure” of the data space. It is expressed through its covariance function (continuous case) or covariance matrix (discrete case) of data $C_d$. When data $d_i$ are independent, $C_d$ is diagonal and its elements are the square of the errors on data $\sigma_{d_i}$.

Finite-frequency effects

As mentioned previously, a strong hypothesis is that in the framework of geometrical optics approximation, only large scale heterogeneities can be retrieved. But interesting geological objects such as slabs, plumes are smaller scale. To go beyond the ray theory, it is necessary to take account of the

13
finite-frequency effect when scale length has the same order of magnitude as the seismic wavelength. It is possible to use the scattering theory based on the Born or Rytov approximations (see for example Woodhouse and Girnius (1982) for normal mode approach, Snieder (1988) for surfaces waves, Yomogida (1992), Dahlen et al. (2000) for body waves). Equation (13) shows that the sensitivity kernels are 1D, meaning that only heterogeneities in the vertical plane containing the source and the receiver are taken into account, whereas, by using the scattering theory, it is possible to calculate 3D kernels and consequently to take account of off-path heterogeneities. The equations (12) and (13) must be replaced by an integral over the volume Ω:

$$\frac{\Delta V_{\text{obs}}(T)}{V_0(T)} = \sum_{i=1}^{14} \int \int \int_{\Omega} K_i(T, \theta, \phi) \frac{\delta p_i(r, \theta, \phi)}{p_i} d\Omega$$

(14)

where $K(T, \theta, \phi)$ is the scattering Fréchet sensitivity kernel, which depends on wave type (Rayleigh or Love) and on the relative location of $E$ and $R$ (see for example Romanowicz, 2002 for a review).

Very different strategies can be followed for calculating this triple integral, by separating the surface integral and the radial integral. Marqueing

**Parameter space:** $p(r)$

It is quite important to thoroughly think of the structure of the parameter space. First of all, it is necessary to define which parameters are required to explain our dataset, how many physical parameters can be effectively inverted for, in the framework of the theory that is considered. For example, if the Earth is assumed to be elastic, laterally heterogeneous but isotropic, only 3 independent physical parameters, $V_P$, $V_S$ and density $\rho$ (or the elastic moduli $\lambda$, $\mu$ and $\rho$) can be inverted for, from surface waves. In a transversely isotropic medium with a vertical symmetry axis (Anderson, 1961; Takeuchi and Saito, 1972), the number of independent physical parameters is now 6 (5 elastic parameters + density). In the most general case of a weak anisotropy, 14 physical parameters (13 combinations of elastic moduli + density) can actually be inverted for, using surface waves. Therefore, the number of ”physical” parameters $p_i$ depends on the underlying theory which is used for explaining the dataset.

Once the number of ”physical” independent parameters is defined, we must define how many ”spatial” (or geographical) parameters are required to describe the 3D distributions $p_i(r, \theta, \phi)$. That is a difficult problem because the number of ”spatial” parameters which can be reliably retrieved from the dataset, is not necessarily sufficient to provide a correct description of $p_i(r, \theta, \phi)$, i.e. of the real Earth. The correct description of $p_i(r, \theta, \phi)$ depends on its spectral content: for example, if $p_i(r, \theta, \phi)$
is characterized by very large wavelengths, only a small number of spatial parameters is necessary, but if $p_i(r, \theta, \phi)$ presents very small-scale features, the number of spatial parameters will be very large. In any case, it is necessary to assess the range of possible variations for $p_i(r, \theta, \phi)$ in order to provide some bounds on the parameter space. This is done through a covariance function of parameters in the continuous case (or a covariance matrix for the discrete case) $C_{p_i,p_j}(r,r')$ at 2 different points $r, r'$. These a priori constraints can be provided by other fields in geosciences, geology, mineralogy, numerical modeling...

Consequently, a tomographic technique must not be restricted to the inversion of parameters $\mathbf{p} = \{p_i(r, \theta, \phi)\}$ that are searched for, but must include the calculation of the final covariance function (or matrix) of parameters $C_p$. That means that the retrieval of parameters is contingent to the resolution and the errors of the final parameters and is largely dependent on the resolving power of data (Backus and Gilbert, 1967, 1968, 1970). Finally, the functional $\mathbf{g}$ which expresses the theory relating the data space to the parameter space is also subject to uncertainty. In order to be completely consistent, it is necessary to define the domain of validity of the theory and to assess the error $\sigma_T$ associated with the theory. Tarantola and Valette (1982) showed that the error $\sigma_T$ is simply added to the error on data $\sigma_d$.

### 2.2 Inverse problem

So far, we did not make assumption on the functional $\mathbf{g}$ relating data and parameters. But in the framework of first order perturbation theory, the forward problem is usually linearized and the equation (13) can be simply written in the linear case:

$$\mathbf{d} = \mathbf{Gp}$$

where $\mathbf{G}$ is now a matrix (or a linear operator) composed of Fréchet derivatives of $\mathbf{d}$ with respect to $\mathbf{p}$, which has the dimensions $n_d \times n_p$ (number of data $\times$ number of parameters). This matrix usually is not square and many different techniques in the past have been used for inverting $\mathbf{G}$. In any case, the inverse problem will consist in finding an inverse for the functional $\mathbf{g}$, that we will write $\mathbf{g}^{-1}$, notwithstanding the way it is obtained, such that:

$$\mathbf{p} = \mathbf{g}^{-1}(\mathbf{d})$$

To solve the inverse problem, different algorithms can be used. The least-squares solution is usually solved by minimizing a cost function $J$. Making symmetric the data space and the parameter space,
Tarantola and Valette (1982) are defining the cost function $J$ as:

$$J = (d - Gp)^t C_d^{-1} (d - Gp) + (p - p_0)^t C_p^{-1} (p - p_0)$$

The first term corresponds to classical least-squares with no damping, whereas the second term corresponds to norm damping, which imposes smoothness upon the parameter space. Different choices were proposed for this second term. For example Montagner (1986b) is using a Gaussian covariance function characterized by a correlation length and an a priori error $\sigma_p$ on parameters whereas Su et al. (1984) prefer to minimize the roughness of the model. Other choices consist in taking a constant value such that $\Delta p^t C_p^{-1} \Delta p = \lambda^2 \Delta p^t \Delta p$ (Yoshizawa and Kennett, 2004). Or the covariance operator can be replaced by a Laplacian operator (see for example Zhou et al., 2006):

$$\Delta p^t C_p^{-1} \Delta p = \epsilon \left( \int \int \int |\nabla^2 (\Delta p)|^2 d\Omega \right)^{1/2}$$

A discussion about damping can be found in Trampert and Snieder (1996) who prefer Laplacian over model damping to reduce the spectral leakage.

As an example, by using the expression of $J$, a quite general and widely used algorithm has been derived by Tarantola and Valette (1982):

$$p - p_0 = (G^t C_d^{-1} G + C_p^{-1})^{-1} G^t C_d^{-1} (d - g(p) + G(p - p_0))$$

$$= C_p G^t (C_d + G C_p) G^{-1} (d - g(p) + G(p - p_0))$$

(15)

where $C_d$ is the covariance matrix of data, $C_p$ the covariance function of parameters $p$, and $G$ is the Frechet derivative of the operator $g$ at point $p(r)$. This algorithm can be made more explicit by writing it in its integral form:

$$p(r) = p_0(r) + \sum_i \sum_j \int_V dr' C_{p_0}(r, r') G_i(r') (S^{-1})_{ij} F_j$$

(16)

with $S_{ij} = C_{ij} + \int_V dr_1 dr_2 G_i(r_1) C_{p_0}(r_1, r_2) G_j(r_2)$, $F_j = d_j - g_j(p) + \int_V dr'' G_j(r'') (p(r'') - p_0(r''))$

This algorithm can be iterated and is suited for solving slightly non-linear problems. Different strategies can be followed to invert for the 3D-models $p(r)$, because the size of the inverse problem is usually enormous in practical applications and a compromise must be found between resolution and accuracy (and also computing time). For the example of mantle tomography, a minimum parameter space will be composed of 13 (+density) physical parameters multiplied by 30 layers (if the mantle is divided into 30 independent layers. If geographical distributions of parameters are searched for up to
degree 40 (lateral resolution around 1000 km), that implies a number of about 700,000 independent parameters. Such a problem is still very hard to handle from a computational point of view. A simple approach for solving this problem consists in dividing the inversion procedure into 2 steps. The first step consists in regionalizing phase (or group) velocity data in order to retrieve the different azimuthal terms, and the second step is the inversion at depth. It was implemented by Montagner (1986a,b) and a very similar technique is presented by Barmin et al. (2001). In case of a large dataset, Montagner and Tanimoto (1990) showed how to handle the inverse problem by making a series expansion of the inverse of matrix \( S \). It was recently optimized from a computational point of view by Debayle and Sambridge (2004) and Beucler and Montagner (2005). One advantage of this technique is that it can be applied indifferently to regional studies or global studies. In case of imperfect spatial coverage of the area under investigation, it does not display ringing phenomena commonly observed when a spherical harmonics expansion is used (Tanimoto, 1986).

From a practical point of view, the choice of the model parameterization is also very important and different possibilities can be considered:

- **Discrete basis of functions:**

  A simple choice consists in dividing the earth into 3D blocks with a surface block size different from the radial one. The size of block depends on the lateral resolution expected from the path coverage. A variant of this parameterization is the use of a set of spherical triangular grid points (see for example Zhou et al., 2006). The block decomposition is valid as well for global investigations as for regional studies. Usually, the Earth surface parameterization is different for the radial one. For global study, the natural basis is composed of the spherical harmonics for the horizontal variations

  \[
  p_i(r, \theta, \phi) = \sum_{\ell=0}^{\ell_{\text{max}}} \sum_{m=-\ell}^{\ell} a_{\ell m}(r) Y_{\ell m}(\theta, \phi).
  \]

  Other choices are possible, such as spherical splines (Wang and Dahlen, 1995). When data coverage is very uneven, other strategies are proposed using irregular cells or adaptative meshes (see for example Zhang and Thurber (2005)).

- **Continuous function \( p(r) \).** In that case, the function is directly inverted for. Since the number of parameters is then infinite, it is necessary to regularize the solution by defining a covariance function of parameters \( C_{p_0}(r, r') \). For the horizontal variations, a Von Mises distribution (Fisher, 1953; Montagner, 1986b) can be used for initial parameters \( p_0(r) \):

  \[
  C_{p_0}(r, r') = \sigma_p(r) \sigma_p(r') \exp \frac{\cos \Delta r^2 - 1}{L_{\text{cor}}^2} \approx \sigma_p(r) \sigma_p(r') \exp \frac{-\Delta r^2}{2L_{\text{cor}}^2}
  \]

  where \( L_{\text{cor}} \) is the correlation length which will define the smoothness of the final model. This kind of distribution is well suited for studies on a sphere and is asymptotically equivalent to a Gaussian distribution when \( L_{\text{cor}} \ll a \) (a radius of the Earth). When different azimuthal terms distributions are
It is possible to define cross-correlated covariance functions of parameters \( C_{p_i p_j}(r, r') \), but since the different terms of the Fourier expansion in azimuth correspond to orthogonal functions, the cross-correlated terms off the diagonal can be taken equal to zero.

It is interesting to note that, in equation (16), the Frechet derivatives \( G \) along the path are multiplied by the Gaussian covariance operator \( C_{p_0} \). It means that the technique, which can be named ”Gaussian” tomography, is equivalent to use ”fat” rays: when the correlation length is wider than the Fresnel zone, ray theory applies, and consequently, the finite-frequency effects can be neglected. As discussed by Ritzwoller et al. (2002) and Sieminski et al. (2004), there might be some slight differences in amplitude between Gaussian tomography and Diffraction tomography (taking account of finite-frequency effects), but not in the location of heterogeneities provided that the spatial path coverage is sufficiently dense.

The radial parameterization must be related to the resolving capability of the data at depth, according to the frequency range under consideration. For the radial variations, polynomial expansions can be used (see for example Dziewonski and Woodhouse, 1987) for Tchebyshev polynomials, or Boschi and Ekström (2002) for radial cubic splines. Since the number of physical parameters is very large for the inversion at depth, physical parameters are usually correlated. The different terms of the covariance function \( C_p \) between parameters \( p_1 \) and \( p_2 \) at radii \( r_i \) and \( r_j \) can be defined as follows:

\[
C_{p_1 p_2}(r_i, r_j) = \sigma_{p_1}(r_i)\sigma_{p_2}(r_j)\zeta_{p_1 p_2} \exp \left(\frac{(r_i - r_j)^2}{2L_{r_i} L_{r_j}}\right)
\]

Where \( \zeta_{p_1 p_2} \) is the correlation between physical parameters \( p_1 \) and \( p_2 \) inferred for instance from different petrological models (Montagner and Anderson, 1989a) such as pyrolite (Ringwood, 1975) and piclogite (Anderson and Bass, 1984; and Bass and Anderson, 1984). \( L_{r_i}, L_{r_j} \) are the radial correlation lengths which are used to smooth the inverse model.

The a posteriori covariance function is given by:

\[
C_p = C_{p_0} - C_{p_0} G^T(C_d + GC_{p_0}G^T)^{-1}GC_{p_0} = (G^T C_{d^{-1}} G + C_{p_0^{-1}})^{-1}
\]  \( (17) \)

The resolution \( R \) of parameters can be calculated as well. It corresponds to the impulsive response of the system: \( p = \tilde{g}^{-1} d = \tilde{g}^{-1} g p' = Rp' \). If the inverse problem is perfectly solved, \( R \) is the identity function or matrix. However, the following expression of resolution is only valid in the linear case (Montagner and Jobert, 1981):

\[
R = C_{p_0} G^T(C_d + GC_{p_0}G^T)^{-1}G = (G^T C_{d} G + C_{p_0})^{-1}G^T C_{d^{-1}} G
\]  \( (18) \)

It is interesting to note, that the local resolution of parameters is imposed by both the correlation length and the path coverage, unlike the Backus-Gilbert (1967, 1968) approach, which primarily de-
pends on the path coverage. The effect of a damping factor in the algorithm to smooth the solution, is equivalent to the introduction of a simple covariance function on parameters weighted by the errors on data (Ho-Liu et al., 1989). When the correlation length is chosen very small, the algorithms of Backus-Gilbert (1968, 1970) and Tarantola and Valette (1982) are equivalent.

By considering the a posteriori covariance function and the resolution, it is possible to assess the reliability of the hypotheses made about the independence of parameters. For example Tanimoto and Anderson (1985) and Montagner and Jobert (1988) showed that there is a trade-off between azimuthal terms and constant term in case of a poor azimuthal coverage. For the inversion at depth Nataf et al. (1986) also display the trade-off between physical parameters $V_{PH}$, $V_{SV}$, $\xi$, $\phi$ and $\eta$ when only Rayleigh and Love wave $0 - \Psi$-terms are used in the inversion process.

Though 13 elastic parameters (+ density) are necessary to explain surface wave data (Rayleigh and Love waves), only 4 parameters are well resolved for small anisotropy (Montagner and Jobert, 1988): the azimuthally averaged S-wave velocity $V_S$, the radial anisotropy expressed through the $\xi$ parameter ($\xi = (V_{SH}/V_{SV})^2$) where $V_{SH}$ (resp. $V_{SV}$) is the velocity of S-wave propagating horizontally with horizontal transverse polarization (resp. with vertical polarization), and the $G$ ($G_c, G_s$) parameters expressing the horizontal azimuthal variation of $V_{SV}$. $\xi$ was introduced in the reference Earth model PREM (Dziewonski and Anderson, 1981) down to 220km in order to explain a large dataset of free oscillation eigenfrequencies and body wave travel times. The other elastic parameters can be derived, by using constraints from petrology in order to reduce the parameter space (Montagner & Anderson, 1989a). This approach was followed by Montagner & Anderson (1989b) to derive an average reference earth model, and by Montagner & Tanimoto (1991) for the first global 3-D anisotropic model of the upper mantle.

2.3 Isotropic and anisotropic images of the upper mantle

The complete anisotropic tomographic procedure has been implemented for making different regional and global studies. Many global isotropic tomographic models of the upper mantle were published since Woodhouse and Dziewonski (1984) and the recent results have been reviewed by Romanowicz (2003). Many models inverting only for radial anisotropy but neglecting azimuthal anisotropy, have also been published (Nataf et al., 1984, 1986; Ekström and Dziewonski, 1998; Shapiro and Ritzwoller, 2002, Gung et al., 2003; Panning and Romanowicz, 2004; Zhou et al., 2006) The complete anisotropic tomographic technique (including azimuthal anisotropy) has been applied for investigating the upper mantle structure either at a regional scale of the Indian Ocean (Montagner, 1986a; Montagner and
Jobert, 1988; Debayle and Lévêque, 1997), of the Atlantic Ocean (Mocquet and Romanowicz, 1989; Silveira et al., 1998; Silveira and Stutzmann, 2002), of Africa (Hadiouche et al., 1989; Debayle et al., 2001; Sebai et al., 2005; Sicilia et al., 2005), of Pacific Ocean (Nishimura & Forsyth, 1989; Bussy et al., 1993; Montagner, 2002; Ritzwoller et al., 2004), of Antarctica (Roullet et al., 1994), Australia (Debayle and Kennett, 2000; Simons et al., 2002) and Central Asia (Griot et al., 1998a,b; Villaseñor et al., 2001) or at a global scale (Montagner & Tanimoto, 1990, 1991; Montagner, 2002; Debayle et al., 2004). The reader is also referred to a quantitative comparison of tomographic and geodynamic models by Becker and Boschi (2002).

An important issue, when constructing tomographic models is the correction for crustal structure, where sedimentary thickness, Moho depth variations are so strong that they affect dispersion of surface waves at least up to 100s: it was shown (Montagner and Jobert, 1988) that standard perturbation theory is inadequate to correct for crustal correction and more rigorous approaches were proposed (Li and Romanowicz, 1996; Boschi and Ekström, 2002; Zhou et al., 2005) using the updated crustal models 3SMAC (Nataf and Ricard, 1996; Ricard et al., 1996) or CRUST2.0 (Mooney et al., 1998; Laske et al., 2001).

As an example of the results obtained after the first step of the tomographic procedure, Figure 4 shows different maps of $2 - \Psi$ azimuthal anisotropy for Rayleigh waves at 100s period for the first 3 modes, $n=0, 1, 2$ superimposed on the isotropic part ($0 - \Psi$ term) of phase velocity (Beucler and Montagner, 2006). From petrological and mineralogical considerations, Montagner and Nataf (1988) and Montagner and Anderson (1989a,b) showed that the predominant terms of phase velocity azimuthal expansion are the $0-\Psi$ and $2-\Psi$ for Rayleigh waves, and $0-\Psi$ and $4-\Psi$ for Love waves. However, Trampert and Woodhouse (2003) carefully addressed the requirement of azimuthal anisotropy, and demonstrated that Rayleigh wave data need both $2\Psi$ and $4\Psi$ terms, which is also confirmed by Beucler and Montagner (2006). It was shown that for the same variance reduction, a global parameterization of anisotropy including azimuthal anisotropy requires fewer parameters than an isotropic parameterization. This apparent paradox can be explained by the fact that the increase of physical parameters is largely compensated by the smaller number of geographical parameters, i.e. larger scale heterogeneities. Other tests have questioned whether phase data are sensitive enough to detect azimuthal anisotropies (Larson et al., 1998; Laske et al., 1998) and the use of additional polarization data has been proposed.

Most tomographic models agree that down to about 250-300km, the deep structure is closely related to plate tectonics and continental distribution. Figure 5 presents two horizontal cross-sections from the most recent model of Debayle et al. (2005), which illustrates and confirms the robust features of
the upper mantle models published so far since Montagner and Tanimoto (1991). In the upper mantle depth range around 100km, all plate boundaries are slow: ridges and back-arc areas are slow, shields are fast and seismic velocity in oceanic areas is increasing with the age of the seafloor. Except at few places, it is found that radial anisotropy expressed through the \( \xi \) parameter \( (\xi = \frac{V_{3\parallel}^2 - V_{3\perp}^2}{V_3^2}) \) is positive, as large as 10% in some oceanic areas and decreases with depth.

The amplitude of SV-wave azimuthal anisotropy \((G\) parameter) presents an average value of about 2% below oceanic areas (Figure 5b). Montagner (1994, 2002) noted a good correlation between seismic azimuthal anisotropy and plate velocity directions (primarily for fast moving plates) given by Minster and Jordan (1978) or DeMets et al. (1990). However, the azimuth of G-parameter can vary significantly as a function of depth. For instance, at shallow depths (down to 60km), the maximum velocity can be parallel to mountain belts or plate boundaries (Vinnik et al., 1991; Silver, 1996; Babuska et al., 1998), but orthogonal to them at large depth. This means that, at a given place, the orientation of fast axis is a function of depth, which explains why the interpretation of SKS splitting with a simple model is often difficult.

As depth increases, the amplitude of heterogeneities rapidly decreases, some trends tend to vanish, and some distinctive features come up: most fast ridges are still slow but slow ridges are hardly visible and back-arc regions are no longer systematically slow below 200km. Large portions of fast ridges are offset with respect to their surface signatures. Below 300km of depth (not shown here), high velocity body below the western and the eastern Pacific rim is the most striking feature, which can be related to subducting slabs.

A visual and quantitative comparison of existing models can be found in the REM (Reference Earth Model) web site at the following address: http://mahi.ucsd.edu/Gabi/rem.html

3 Geodynamic applications

The most popular application of large-scale tomographic models is the understanding of mantle convection. Seismic velocity anomalies can be converted, under some assumptions, into temperature anomalies, density anomalies but also into chemical or mineralogical heterogeneities. The application of seismic anisotropy to geodynamics in the upper mantle is straightforward, if we assume that, due to the lattice preferred orientation (LPO) of anisotropic crystals such as olivine (Christensen and Lundquist, 1982; Nicolas et al., 1973), the fast-polarization axis of mineralogical assemblages is in the flow plane parallel to the direction of flow. Figure 6 shows what is expected for the observable pa-
rameters $V_S$, $\xi$, $G$, $\psi_G$ in the case of a simple convective cell with LPO. Radial anisotropy $\xi$ expresses the vertical ($\xi < 1$) or horizontal character ($\xi > 1$) of convective flow, and the azimuthal anisotropy $G$, can be related to the horizontal flow direction. Conversely, the three maps of $V_S$, $\xi$, $G$, can be interpreted in terms of convective flow. These 3 pieces of information are necessary to correctly interpret the data. For example, upwellings or downwellings are both characterized by a weak or negative $\xi$ parameter, but a correlative positive or negative $\delta V_S$ discriminates between these possibilities. By simultaneously inverting at depth for the different azimuthal terms of Rayleigh and Love waves, it is therefore possible to separate the lateral variations in temperature from those induced by the orientation of minerals. Such an interpretation might however be erroneous in water-rich mantle regions where LPO of minerals such as olivine is not simply related to the strain field (e.g. Jung and Karato, 2001). We will only present some examples of interesting applications of anisotropy in large scale geodynamics and tectonics. Seismic anisotropy in the mantle primarily reflects the strain field prevailing in the past (frozen-in anisotropy) for shallow layers or present convective processes in deeper layers. Therefore, it makes it possible to map convection in the mantle. It must be noted that, when only the radial anisotropy is retrieved, its interpretation is non unique. A fine layering of the mantle can also generate such a kind of anisotropy, and neglecting the azimuthal anisotropy can bias the amplitude of radial anisotropy and its interpretation.

The uppermost mantle down to 410km is the depth range where the existence of seismic anisotropy is now widely recognized and well documented. Azimuthal variations have been found for body waves and surface waves in different areas of the world. During the last years, the shear wave splitting, primarily for SKS waves was extensively used to study continental deformation, but very few studies using body waves are devoted to oceanic areas. Conversely, global anisotropic upper mantle models have been primarily derived during the last 10 years from surface waves, which are sensitive to structure below oceanic areas in the absence of ocean bottom stations and consequently of dense body wave data. The intercomparison of anisotropic body wave and surface wave data is still in its infancy. However, as shown by Montagner et al. (2000), Vinnik et al. (2002), Simons et al. (2003) such a comparison is providing encouraging results.

3.1 Oceanic plates

Oceans are the areas where plate tectonics applies almost perfectly and this is particularly the case in the largest one, the Pacific plate. Figure 7 presents 3 vertical cross-sections at two different latitudes, displaying $V_{SV}$ velocity anomalies (figure 7a) and the 2 kinds of anisotropy, which can be
retrieved by simultaneous inversion of Rayleigh and Love waves constant $0 - \Psi$ and azimuthal terms of equation (1) from the model of Montagner (2002). In figure 7b, the equivalent radial anisotropy of the medium, for S-wave expressed through the $\xi$ parameter, is displayed. The maps of figure 7c are the distributions of the $G$-parameter related to the azimuthal variation of SV-wave velocity. The maximum amplitude of $G$ is around 5% and rapidly decreases as depth increases. The distribution of velocity and anisotropy are completely different for these different cross-sections. The thickening of lithosphere with the age of the seafloor is well observed on $V_{SV}$ velocity maps, but lithosphere is much thicker in the northern cross-section. When compared with the cooling half-space model, bathymetry, heat flux and lithospheric thickness flatten with age (see Ritzwoller et al., 2004 for recent results). This flattening is explained by basal reheating, especially in the Central Pacific and the birth of small-scale convection below the lithosphere (T and Jaupart, 1994; Solomatov and Moresi, 2000).

Radial cross-sections (Figure 7b) show that the $\delta \xi = \xi - \xi_{PREM}$ parameter is usually negative and small, where flow is primarily radial (mid-ocean ridges and subduction zones). For the east Pacific Rise, Gu et al. (2005) found that a negative radial anisotropy is observed at least down to 300km. Between plate boundaries, oceans display very large areas with a large positive radial anisotropy such as in the Pacific Ocean (Ekström & Dziewonski, 1998), characteristic of an overall horizontal flow field. This very large anisotropy in the asthenosphere might be the indication of a strong deformation field at the base of the lithosphere (Gung et al., 2003), corresponding to the upper boundary layer of the convecting mantle (Anderson and Regan, 1983; Montagner, 1998).

Since convective flow below oceans is dominated by large scale plate motions, the long wavelength anisotropy found in oceanic lithospheric plates and in the underlying asthenosphere, should be similar to the high-resolution anisotropy measured from body waves. Incidentally, one of the first evidences of azimuthal anisotropy was found in the Pacific Ocean by Hess (1964) for Pn-waves. So far, there are very few measurements of anisotropy by SKS splitting in the oceans. Due to the lack of seismic stations on the sea floor (with the exception of H2O halfway between Hawaii and California), the only measurements available for SKS were performed in stations located on ocean islands (Ansel & Nataf, 1989; Kuo & Forsyth, 1992; Russo & Okal, 1999; Wolfe & Silver, 1998), which are by nature anomalous objects, such as volcanic hotspots, where the strain field is perturbed by the upwelling material and not necessarily representative of the main mantle flow field. SKS splitting was measured during the temporary MELT experiment on the East-Pacific Rise (Wolfe & Solomon, 1998) but the orientation of the splitting is in disagreement with the petrological predictions of Blackman et al. (1996). Walker et al. (2001) presented a first measurement of SKS splitting at H2O, but it is in disagreement with
independent SKS splitting measurements at the same station by Vinnik et al. (2002) and with surface wave anisotropy (Montagner, 2002).

The large scale azimuthal anisotropy within and below lithosphere in the depth range 100-300km is closely related to plate motions (Montagner, 1994; Ekström, 2000) and modeled in this framework (Tommasi et al., 1996). Fast moving oceanic plates are zones where the comparison between directions of plate velocities (Minster & Jordan, 1978) or NUVEL-1 (DeMets et al., 1990) and directions of $G$-parameter is the most successful (Figure 8). Conversely, such a comparison is more difficult and controversial below plates bearing a large proportion of continents, such as the European-Asian plate, characterized by a very small absolute motion in the hotspot reference frame and probably a large influence of inherited anisotropy.

The map with the G-parameter at 100km (Figure 5) as well as the cross-sections of figure 7c show that the azimuthal anisotropy is very large along spreading ridges with a large asymmetry for the East Pacific rise. The direction of anisotropy is in very good agreement with plate motion, which is also found in all other available models (Ekström, 2000; Smith et al., 2004; Debayle et al., 2005).

The anisotropy is also large in the middle of the Pacific plate, but a line of very small azimuthal anisotropy almost parallel to the EPR is observed there (see also Figure 2 for synthetic SKS). This linear area of small anisotropy was named Low Anisotropy Channel (LAC) by Montagner (2002).

When calculating the variation of the amplitude of azimuthal anisotropy as a function of depth, a minimum comes up between 40 and 60Ma age of the seafloor (Figure 9a). The LAC is presumably related either to cracking within the Pacific plate and/or to secondary convection within and below the rigid lithosphere, predicted by numerical and analog experiments and also translated in the $V_S$ velocity structure (Ritzwoller et al., 2004; Figure 9b). These new features provide strong constraints on the decoupling between the plate and asthenosphere. The existence and location of these LACs might be related to the current active volcanoes and hotspots (possibly plumes) in Central Pacific. LACs, which are dividing the Pacific Plate into smaller units, might indicate a future reorganization of plates with ridge migrations in the Pacific Ocean. They call for more thorough numerical modelling.

### 3.2 Continents

Differences in the thickness of high-velocity layer underlying continents as imaged by seismic tomography have fuelled a long debate on the origin of continental roots (Jordan, 1975, 1978). Some global tomographic models provide a continental thickness of about 200-250km in agreement with heat-flow analysis or electrical conductivity, but others suggest thicker zones up to 400km.
Seismic anisotropy can provide fundamental information on the structure of continents, their root and the geodynamic processes involved in mountain building and collision between continents (Vinnik et al., 1992; Silver, 1996) such as in Central Asia (Griot et al., 1998a,b). Radial anisotropy $\xi$ is usually very heterogeneous below continents in the first 150-200km of depth with positive or negative areas according to geology. But it seems to display a systematic tendency of being positive at larger depth (down to 300km), whereas it is very large in the oceanic lithosphere in the depth range 50-200km and decreases rapidly at larger depths (Montagner, 1994). Conversely, radial anisotropy displays a maximum (though smaller than in oceanic lithosphere) below very old continents (such as Siberian and Canadian Shield) in the depth range 200-400km (Montagner and Tanimoto, 1991). Seismic anisotropy below continents, sometimes confined to the upper 220km (Gaherty & Jordan, 1995) can still be significant below. A more quantitative comparison of radial anisotropy between different continental provinces is presented in Babuska et al. (1998), and demonstrates systematic differences according to the tectonic context. The existence of positive large scale radial anisotropy below continents at depth might be a good indicator of the continental root which was largely debated since the presentation of the model of tectosphere by Jordan (1975, 1978, 1981). If this maximum of anisotropy is assumed to be related to an intense strain field in this depth range, it might be characteristic of the boundary between continental lithosphere and "normal" upper mantle material. Gung et al. (2003) showed that it is possible to reconcile different isotropic tomographic models by taking into account seismic anisotropy. They find that significant radial anisotropy (with $V_{SH} > V_{SV}$) under most cratons in the depth range 250-400km, similar to that found at shallower depths (80-250km) below oceanic basins. Such a result is also in agreement for the Australian continent (Debayle and Kennett, 2000; Simons et al., 2002). So, all results seem to show that the root of continents as defined by radial anisotropy is located between 200 and 300km. However, this result is not correlated with a maximum in azimuthal anisotropy in this depth range (Debayle and Kennett, 2005): the fast-moving Australian plate seems to be the only continental region with a sufficiently large deformation at its base to be transformed into azimuthal anisotropy. They propose that, for continents other than Australia, weak influence of basal drag on the lithosphere may explain why azimuthal anisotropy is observed only in a layer located in the uppermost 100 km of the mantle. This layer shows a complex organisation of azimuthal anisotropy suggesting a frozen-in origin of deformation, compatible with SKS splitting.

The difference in radial and azimuthal anisotropies between oceans and continents might reflect a difference of coupling between lithosphere and asthenosphere, through the basal drag. The coupling might be weak below most continental roots, in contrast with the Pacific plate, where the coupling
(reflected by plate direction) is the first order effect in the uppermost 200km for young ages, before thermal instabilities take place at the base of the lithosphere, as evidenced by the existence of low anisotropy channels. These results on the difference between oceanic and continental anisotropies are illustrated in Figure 10.

3.3 Velocity and Anisotropy in the transition zone

The transition zone plays a key role in mantle dynamics, particularly the 660km-discontinuity which might inhibit the passage of matter between the upper and the lower mantle. Its seismic investigation is made difficult on the global scale by the poor sensitivity of fundamental surface waves in this depth range and by the fact that teleseismic body waves recorded at continental stations from earthquakes primarily occurring along plate boundaries have their turning point below the transition zone. For body waves, many different techniques using SS-precursors (Shearer, 1991) or P-to-S converted waves (Chevrot et al., 1999) were used at global scale to investigate the thickness of the transition zone. In spite of some initial controversies, a recent model by Lawrence and Shearer (2006) provides a coherent large-scale image of the transition zone thickness.

Whatever the type of data (normal mode, higher modes of surface waves, body waves), an important feature of the transition zone is that, contrarily to the rest of the upper mantle, the upper transition zone is characterized by a large degree 2 pattern (Masters et al., 1982), and to a less extent, a strong degree 6. The degree 2 pattern (as well as degree 6) can be explained by the predominance of a simple large-scale flow pattern characterized by two upwellings in central Pacific Ocean and Eastern Africa and two downwellings in the Western and Eastern Pacific Ocean (Montagner & Romanowicz, 1993), proposed initially in the lower mantle (Busse, 1983). This scheme was corroborated by the existence, in the upper transition zone, of a slight but significant degree 4 radial anisotropy displayed by Montagner & Tanimoto (1991) and Roult et al. (1990) in agreement with the prediction of this model. Therefore, the observations of the geographical distributions of degrees 2, 4, 6 in the transition zone are coherent and spatially dependent. Montagner (1994) compared these different degrees to the corresponding degrees of the hotspot and slab distribution. In this simple framework, the distribution of plumes (degree 2+6) are merely a consequence of the large scale simple flow in the transition zone. The degree 6 of velocity in the transition zone is well correlated with the distribution of hotspots and might indicate that many mantle plumes might originate in the transition zone. Ritsema et al. (2004) observe lower-than-average shear velocity at eight hotspots in this depth range (Figure 11). These results suggest that there are different families of plumes, some of them originating in the transition
As for anisotropy in the transition zone, **Montagner & Kennett (1996)**, by using eigenfrequency data, display some evidence of radial anisotropy in the upper (410-660km) and lower (660-900km) transition zones. **Gung and Romanowicz (2004)** also display a slight maximum of the degree 0 $\xi$ in the transition zone. The existence of anisotropy close to the 660km-discontinuity was also found by **Vinnik & Montagner (1996)** below Germany and **Vinnik et al. (1998)** in central Africa. By studying P-to-S converted waves at the GRF network and at GEOSCOPE station BNG in central Africa, they observed that part of the initial P-wave is converted into SH-wave. This signal can be observed on the transverse component of seismograms. The amplitude of this SH-wave cannot be explained by a dipping 660km-discontinuity and it constitutes a good evidence for the existence of anisotropy just above this discontinuity. However, there is some evidence of lateral variation of anisotropy in the transition zone as found by the investigation of several subduction zones (**Fischer & Yang, 1994; Fischer & Wiens, 1996**). **Fouch & Fischer (1996)** present a synthesis of these different studies and show that some subduction zones such as Sakhalin Islands require deep anisotropy in the transition zone, whereas others such as Tonga do not need any anisotropy. They conclude that their data might be reconciled by considering the upper transition zone (410-520km) intermittently anisotropic, and the rest of the transition zone might be isotropic.

Anisotropy in the transition zone was also advocated by 2 independent studies, using different datasets. The observations of **Wookey et al. (2002)**, though controversial, present evidence of very large S-wave splitting (up to 7.s) in the vicinity of the 660km discontinuity between Tonga-Kermadec subduction zone and Australia. On a global scale, **Trampert and van Heijst (2002)** show a long-wavelength azimuthal anisotropic structure in the transition zone. The rms amplitude of lateral variations of G is about 1%. **Beghein and Trampert (2003)** using probability density functions and separating $\xi$, $\phi$ and $\eta$ anisotropies suggest a chemical component to explain these different parameters. The interpretation of these new tentative results is not obvious and new data are necessary to close the debate on the nature of velocity and anisotropy heterogeneities in the transition zone. The transition zone might be a mid-mantle boundary layer, and a detailed and reliable tomographic model of S-wave velocity and anisotropy in the transition zone will provide fundamental insights into the dynamic of the whole mantle.
4 Numerical Modeling and Perspectives

In the previous sections, we have highlighted the presence of lateral heterogeneities in seismic velocity and anisotropy in different parts of the earth’s upper mantle. However, anisotropy is not present in all depth ranges nor at all scales. There is some consensus for the presence of radial anisotropy in many parts of the upper mantle in order to simultaneously explain Love-wave and Rayleigh-wave dispersion and even in the lower mantle (Panning et al., 2004). The existence of azimuthal anisotropy is more controversial, though, from petrological reasons, it turns out that radial anisotropy and azimuthal anisotropy are intimately related and should be searched for simultaneously. Additional data such as polarization data might help to provide additional constraints on both kinds of anisotropy (Yu and Park, 1993; Pettersen and Maupin, 2002). But it requires the development of improved theoretical and numerical methods in order to work on the amplitude of seismograms.

Thanks to the access to very powerful computers, we are at the beginning stage of a new era for seismology. The twentieth century was dominated by the use of ray theory and later on, of normal mode theory. Since it is now feasible to numerically compute synthetic seismograms in complex 3D structures in global spherical geometry (Komatitsch and Vilotte, 1998; Komatitsch and Tromp, 1999; Capdeville et al., 2003), it is possible to model the complex interaction between seismic waves and 3D heterogeneity, particularly in anisotropic, anelastic media. Some new and sophisticated tomographic methods are presently developed (Montelli et al., 2004; Capdeville et al., 2005; Tromp et al., 2005; Zhou et al., 2006), which should provide access to the complexity of the Earth mantle by the mapping of short-scale heterogeneities such as mantle plumes, in anisotropic and anelastic media.

A second important challenge is the complete understanding of the origin of anisotropy from the mineral scale up to global scale in the different layers of the earth. In the upper mantle, seismic anisotropy is due to LPO of anisotropic minerals such as olivine at large scales, requiring several strong conditions, starting with the presence of anisotropic crystals up to the existence of an efficient large scale present or past strain field. In order to fill the gap between grain scale modeling (McKenzie, 1979; Ribe, 1989; Kaminski and Ribe, 2001) and large scale anisotropy measurements in a convective system (Tommasi et al., 2000), there is now a real need to make more quantitative comparisons between seismic anisotropy and numerical modeling. Gaboret et al. (2003) and Becker et al. (2003) calculated the convective circulation in the mantle by converting perturbations of S-wave velocity into density perturbations. Figure 12 shows 2 cross-sections through the Pacific hemisphere and the associated flow lines (Gaboret et al., 2003) derived from the tomographic model of Ekström and Dziewonski (1998). This kind of modeling makes it possible to calculate the strain tensor and to test
different hypotheses for the prevailing mechanisms of alignment, by comparison with seismic data. The upper mantle is the best known of the deep layers of the earth, where there is now good agreement between many isotropic global tomographic models. But the account of seismic anisotropy is mandatory to avoid biased isotropic heterogeneities. The main application of anisotropy is the mapping of mantle convection and of its boundary layers (Karato, 1998; Montagner, 1998). The finding of anisotropy in the transition zone (if confirmed) will provide strong constraints on the flow circulation and the exchange of matter between the upper and the lower mantle. Pursuing the first pioneering efforts, the systematic modeling of the complete seismic waveform in 3D heterogeneous, anisotropic and anelastic media associated with new techniques of numerical modeling of seismograms will probably enhance our vision of the whole mantle.

In parallel to these theoretical and numerical challenges, there is a crucial need for instrumental developments since there are still many areas at the surface of Earth devoid of broadband seismic stations. These regions are primarily located in southern hemisphere and more particularly in oceanic areas where no islands are present. Therefore, an international effort is ongoing, coordinated through I.O.N. (International Ocean Network) in order to promote the installation of geophysical ocean bottom observatories in order to fill the enormous gaps in the station coverage (for a description of I.O.N., http://seismo.berkeley.edu/seismo/ion).

REFERENCES


1971.


Shapiro N.M., M. Campillo M., Stehly L., Ritzwoller M.H. 2005 High-resolution surface-wave tomography from ambient


Appendix: Effect of Anisotropy on surface waves in the plane-layered medium

Figure 1: Definition of the Cartesian coordinate system \((x, y, z)\) used in the calculations. \(\Psi\) is the azimuth of the wavevector with respect to North.

The half-space is assumed to be homogeneous and may be described by its density \(\rho(z)\) and its 4th-order elastic tensor \(\Gamma(z)\) with 21 independent elastic coefficients. All these parameters are so far supposed independent of \(x\) and \(y\) coordinates \((z\) is the vertical component\). This condition will be released in the next section. The unperturbed medium is assumed isotropic with an elastic tensor \(\Gamma_0(z)\). In that medium, the two cases of Love and Rayleigh wave dispersion can be successively considered.

The unperturbed Love wave displacement is of the form:

\[
\mathbf{u}(\mathbf{r}, t) = \begin{pmatrix} -W(z) \sin \Psi \\ W(z) \cos \Psi \\ 0 \end{pmatrix} \exp(i[k(x \cos \Psi + y \sin \Psi) - \omega t]) \tag{19}
\]

where \(W(z)\) is the scalar depth eigenfunction for Love waves, \(k\) is the horizontal wave number, and \(\Psi\) is the azimuth of the wave number \(k\) measured clockwise from the North.

The unperturbed Rayleigh wave displacement is of the form:

\[
\mathbf{u}(\mathbf{r}, t) = \begin{pmatrix} V(z) \cos \Psi \\ V(z) \sin \Psi \\ iU(z) \end{pmatrix} \exp(i[k(x \cos \Psi + y \sin \Psi) - \omega t]) \tag{20}
\]

where \(V(z)\) and \(U(z)\) are the scalar depth eigenfunctions for Rayleigh waves. The associated strain tensor \(\epsilon(\mathbf{r}, t)\) is defined by:

\[
\epsilon_{ij}(\mathbf{r}, t) = 1/2(u_{i,j} + u_{j,i}) \tag{21}
\]

where \(,\) denotes the differentiation with respect to the \(j\)-th coordinate. The medium is perturbed from \(\Gamma_0(z)\) to \(\Gamma_0(z) + \gamma(z)\), where \(\gamma(z)\) is small compared to \(\Gamma_0(z)\) but quite general in the sense that there is no assumption on the kind of anisotropy. This means that in this approximation, we can still consider quasi-Love modes and quasi-Rayleigh modes (Crampin, 1984). From Rayleigh’s principle, the first order perturbation \(\delta V(k)\) in phase velocity dispersion is (Smith and Dahlen, 1973, 1975):

\[
\delta V(k) = \frac{V}{2\omega^2} \int_0^\infty \gamma_{ijkl} \epsilon_{ij} \epsilon_{kl}^* dz \tag{22}
\]

where \(u_i\) and \(\epsilon_{ij}\) are respectively the displacement and the strain for the unperturbed half-space and the asterisk denotes complex conjugation. Now because of the symmetry of the tensors \(\gamma(z)\) and \(\epsilon\),
we use the simplified index notation \( c_{ij} \) and \( \epsilon_i \) for the elements \( \gamma_{ijkl} \) and \( \epsilon_{ij} \), but the number \( n_{ij} \) of coefficients \( \gamma_{ijkl} \) for each \( c_{ij} \) must be taken into account. The simplified index notation for the elastic tensor \( \gamma_{ijkl} \) is defined in a coordinate system \((x_1, x_2, x_3)\) by:

\[
\gamma_{ijkl} \rightarrow c_{pq} \begin{cases} 
  \text{if } i = j & \Rightarrow p = i \\
  \text{if } k = l & \Rightarrow q = k \\
  \text{if } i \neq j & \Rightarrow p = 9 - i - j \\
  \text{if } k \neq l & \Rightarrow q = 9 - k - l 
\end{cases} \quad (23)
\]

This kind of transformation enables us to relate the 4\(^{th}\) order tensor \( \gamma \) (3x3x3x3) to a matrix \( c \) (6x6). The same simplified index notation can be applied to the components of the strain tensor \( \epsilon_{ij} \), transforming the 2\(^{nd}\) order tensor \( \epsilon \) (3x3) into a vector with 6 components. However, it is necessary to be careful, because to a given \( c_{pq} \) correspond several \( \gamma_{ijkl} \), and \( \gamma_{ijkl} \) must be replaced by \( n_{pq} c_{pq} \), where \( n_{pq} \) is the number of \( \gamma_{ijkl} \) giving the same \( c_{pq} \). Therefore, the equation (22) expressing Rayleigh’s principle can be rewritten as:

\[
\delta V(k) = \frac{V}{2\omega^2} \int_0^\infty \sum_{ij} n_{ij} c_{ij} \epsilon_i^* \epsilon_j^* \, dz 
\]

(24)

We only detail the calculations for Love waves.

- Love waves.

By using previous expressions for \( u(r, t) \) (19) and \( c_{ij}(r, t) \) (21), the various expressions of strain are:

\[
\begin{align*}
\epsilon_1 &= -i \cos \Psi \sin \Psi \, k W \\
\epsilon_2 &= i \cos \Psi \sin \Psi \, k W \\
\epsilon_3 &= 0 \\
\epsilon_4 &= 1/2 \cos \Psi \, W' \\
\epsilon_5 &= -1/2 \sin \Psi \, W' \\
\epsilon_6 &= 1/2(\cos^2 \Psi - \sin^2 \Psi) \, k W \\
\end{align*}
\]

(25)

where \( W' = \frac{dW}{d\Psi} \). In table 1, the different terms \( n_{ij} c_{ij} \epsilon_i \epsilon_j^* \) are given. We note that when \( c_{ij} \epsilon_i \epsilon_j^* \) is a purely imaginary complex, its contribution to \( \delta V(k, \Psi) \) is null. When all the contributions are summed, the different terms \( \cos^k \Psi \sin^l \Psi \) are such that \( k + l \) is even, which is not surprising in the light of the reciprocity principle. Therefore, each term can be developed as a Fourier series in \( \Psi \) with only even terms. Finally it is found:

\[
\delta V_L(k, \Psi) = \frac{V}{2\omega^2 L_0} \int_0^\infty dz \left[ k^2 W^2 \frac{1}{8} (c_{11} + c_{22} - 2c_{12} + 4c_{66}) + W'^2 \frac{1}{2} (c_{44} + c_{55}) \right. \\
+ \cos 2\Psi \, W'^2 \left( \frac{1}{2} (c_{44} - c_{55}) \right) - \sin 2\Psi \, W'^2 c_{45} \\
- \cos 4\Psi \, k^2 W^2 \frac{1}{8} (c_{11} + c_{22} - 2c_{12} - 4c_{66}) \\
\left. + \sin 4\Psi \, k^2 W^2 \frac{1}{2} (c_{26} - c_{16}) \right] 
\]

(26)

In the particular case of a transversely isotropic medium with a vertical symmetry axis (also named radial anisotropic medium), we have: \( c_{11} = c_{22} = \delta A \), \( c_{33} = \delta C \), \( c_{12} = \delta (A - 2N) \), \( c_{13} = c_{23} = \delta F \), \( c_{44} = c_{55} = \delta L \), \( c_{66} = \delta N \) and \( c_{14} = c_{24} = c_{15} = c_{25} = c_{16} = c_{26} = 0 \). The local azimuthal terms vanish and the previous equation (26) reduces to:

\[
\delta V_L(k, \Psi) = \frac{1}{2V_L L_0} \int_0^\infty \left( W^2 \delta N + \frac{W'^2}{k^2} \delta L \right) dz 
\]

(27)
Therefore, the same expressions as in Takeuchi and Saito (1972, p. 268) are found in the case of radial anisotropy. The $0 - \Psi$ term of equation (26) corresponds to the averaging over azimuth $\Psi$, which provides the equivalent transversely isotropic model with vertical symmetry axis by setting:
\[
\begin{align*}
\delta N &= \frac{1}{2}(c_{11} + c_{22}) - \frac{1}{4}c_{12} + \frac{1}{2}c_{66} \\
\delta L &= \frac{1}{2}(c_{14} + c_{55})
\end{align*}
\]
If we call $C_{ij}$ the elastic coefficients of the total elastic tensor, we can set:
\[
N = \rho V_{SH}^2 = \frac{1}{8}(C_{11} + C_{22}) - \frac{1}{4}C_{12} + \frac{1}{2}C_{66}
\]
\[
L = \rho V_{SV}^2 = \frac{1}{2}(C_{44} + C_{55})
\]
According to equation (26), the first order perturbation in Love wave phase velocity $\delta V_L(k, \Psi)$ can then be expressed as:
\[
\delta V_L(k, \Psi) = \frac{1}{2V_0(k)}[L_1(k) + L_2(k) \cos 2\Psi + L_3(k) \sin 2\Psi + L_4(k) \cos 4\Psi + L_5(k) \sin 4\Psi]
\]
where
\[
\begin{align*}
L_0(k) &= \int_0^\infty \rho W^2dz \\
L_1(k) &= \frac{1}{L_0} \int_0^\infty (W^2\delta N + \frac{W^2}{k^2} \delta L)dz \\
L_2(k) &= \frac{1}{L_0} \int_0^\infty -G_e(\frac{W^2}{k^2})dz \\
L_3(k) &= \frac{1}{L_0} \int_0^\infty -G_s(\frac{W^2}{k^2})dz \\
L_4(k) &= \frac{1}{L_0} \int_0^\infty -E_c.W^2dz \\
L_5(k) &= \frac{1}{L_0} \int_0^\infty -E_s.W^2dz
\end{align*}
\]
- Rayleigh waves.

The same procedure holds for the local Rayleigh wave phase velocity perturbation $\delta V_R$, starting from the displacement given previously (Montagner and Nataf, 1986).
\[
\delta V_R(k, \Psi) = \frac{1}{2V_0(k)}[R_1(k) + R_2(k) \cos 2\Psi + R_3(k) \sin 2\Psi + R_4(k) \cos 4\Psi + R_5(k) \sin 4\Psi]
\]
where
\[
\begin{align*}
R_0(k) &= \int_0^\infty \rho(U^2 + V^2)dz \\
R_1(k) &= \frac{1}{R_0} \int_0^\infty [V^2.\delta A + \frac{U^2}{k^2}.\delta C + \frac{2U^2V}{k^2}.\delta F + (\frac{V'}{k} - U)^2.\delta L]dz \\
R_2(k) &= \frac{1}{R_0} \int_0^\infty [V^2.B_c + \frac{2U^2V}{k^2}.H_c + (\frac{V'}{k} - U)^2.G_c]dz \\
R_3(k) &= \frac{1}{R_0} \int_0^\infty [V^2.B_s + \frac{2U^2V}{k^2}.H_s + (\frac{V'}{k} - U)^2.G_s]dz \\
R_4(k) &= \frac{1}{R_0} \int_0^\infty E_c.V^2dz \\
R_5(k) &= \frac{1}{R_0} \int_0^\infty E_s.V^2dz
\end{align*}
\]
The 13 depth-dependent parameters $A, C, F, L, N, B_c, B_s, H_c, H_s, G_c, G_s, E_c, E_s$ are linear combinations of the elastic coefficients $C_{ij}$ and are explicitly given as follows:

\[
\begin{align*}
A &= \rho V_{PH}^2 = \frac{3}{8}(C_{11} + C_{22}) + \frac{1}{4}C_{12} + \frac{1}{2}C_{66} \\
C &= \rho V_{PV}^2 = C_{33} \\
F &= \frac{1}{2}(C_{13} + C_{23})
\end{align*}
\]
\[
L = \rho V_{SV}^2 = \frac{1}{2}(C_{44} + C_{55}) \\
N = \rho V_{SH}^2 = \frac{1}{8}(C_{11} + C_{22}) - \frac{1}{4}C_{12} + \frac{1}{2}C_{66}
\]

2 \( \Psi \) -azimuthal term:

\[
\begin{align*}
\cos 2\Psi & \quad \sin 2\Psi \\
B_c & = \frac{1}{2}(C_{11} - C_{22}) & B_s & = C_{16} + C_{26} \\
G_c & = \frac{1}{2}(C_{55} - C_{44}) & G_s & = C_{54} \\
H_c & = \frac{1}{2}(C_{13} - C_{23}) & H_s & = C_{36}
\end{align*}
\]

4 \( \Psi \) -azimuthal term:

\[
\begin{align*}
\cos 4\Psi & \quad \sin 4\Psi \\
E_c & = \frac{1}{8}(C_{11} + C_{22}) - \frac{1}{4}C_{12} - \frac{1}{2}C_{66} & E_s & = \frac{1}{2}(C_{16} - C_{26})
\end{align*}
\]

where indices 1 and 2 refer to horizontal coordinates (1: North; 2: East) and index 3 refers to vertical coordinate. \( \rho \) is the density, \( V_{PH}, V_{PV} \) are respectively horizontal and vertical propagating P-wave velocities, \( V_{SH}, V_{SV} \) horizontal and vertical polarized S-wave velocities. We must bear in mind that \( A, C, L, N \) anisotropic parameters can be retrieved from measurements of the P- and S-wave velocities propagating perpendicular or parallel to the axis of symmetry.
Table 1: Calculation of the various $c_{ij}e_ie_j$ for Love waves, with the simplified index notation.
\[ \alpha = \cos \Psi; \quad \beta = \sin \Psi \]

<table>
<thead>
<tr>
<th>n</th>
<th>ij</th>
<th>$c_{ij}e_ie_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
<td>$c_{11} \alpha^2 \beta^2, k^2 W^2$</td>
</tr>
<tr>
<td>1</td>
<td>22</td>
<td>$c_{22} \alpha^2 \beta^2, k^2 W^2$</td>
</tr>
<tr>
<td>1</td>
<td>33</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>$-c_{12} \alpha^2 \beta^2, k^2 W^2$</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>23</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>24</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>14</td>
<td>$c_{14}(-i\alpha^2\beta), \frac{kW W'}{2}$</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>$c_{15}(i\alpha^2\beta), \frac{kW W'}{2}$</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>$c_{16}(-\alpha\beta)(\alpha^2 - \beta^2), \frac{k^2 W^2}{2}$</td>
</tr>
<tr>
<td>4</td>
<td>24</td>
<td>$c_{24}(-i\alpha^2\beta), \frac{kW W'}{2}$</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>$c_{25}(-i\alpha\beta^2), \frac{k^2 W^2}{2}$</td>
</tr>
<tr>
<td>4</td>
<td>26</td>
<td>$c_{26}(\alpha\beta)(\alpha^2 - \beta^2), \frac{k^2 W^2}{2}$</td>
</tr>
<tr>
<td>4</td>
<td>34</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>35</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>36</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>44</td>
<td>$c_{44} \alpha^2 W^2$</td>
</tr>
<tr>
<td>8</td>
<td>45</td>
<td>$c_{45}(-\alpha\beta), \frac{W^4}{4}$</td>
</tr>
<tr>
<td>8</td>
<td>46</td>
<td>$c_{46}(-i\alpha)(\alpha^2 - \beta^2), \frac{kW W'}{2}$</td>
</tr>
<tr>
<td>4</td>
<td>55</td>
<td>$c_{55} \beta^2 W^2$</td>
</tr>
<tr>
<td>8</td>
<td>56</td>
<td>$c_{56}(i\beta)(\alpha^2 - \beta^2), \frac{kW W'}{2}$</td>
</tr>
<tr>
<td>4</td>
<td>66</td>
<td>$c_{66}(\alpha^2 - \beta^2), \frac{k^2 W^2}{2}$</td>
</tr>
</tbody>
</table>

**Acknowledgments.** I would like to thank Barbara Romanowicz and an anonymous reviewer for constructive reviews, as well as my colleagues from IPG of Paris.
Figure 1: example of real and synthetic seismograms used for retrieving Rayleigh wave dispersion curve for the fundamental mode and overtones (Beucler et al., 2003). Behind body waves, the signal is composed of surface waves. The complex phase before the high amplitude wave packet corresponding to the fundamental mode of Rayleigh wave (n=0) can be synthetized by summing the first overtones.
Figure 2: Map of synthetic SKS splitting derived from the anisotropic surface wave model of Montagner (2002). The delay time is expressed in seconds.
Figure 3: Phase velocity of the fundamental mode and the first 6 higher modes of Rayleigh compared with PREM (right plot) and with results (center) obtained in previous studies along the same path between Vanuatu and California (SCZ Geoscope station) (Beucler et al., 2003).
Figure 4: Rayleigh wave Phase velocity maps at period $T=100$S for the first 3 modes ($n=0, 1, 2$) after Beucler and Montagner (2006)
Figure 5: 2 cross-sections at 100km (top) and 200km (bottom) depths of the global tomographic model of Debayle et al. (2005). Directions of azimuthal anisotropy are superimposed on S-wave velocity heterogeneities. The length of bars is proportional to its amplitude (<2%)
Figure 6: The seismic observable parameters $V_S$, $\xi$, $G$, $\Psi_G$ associated with a simple convecting cell in the upper mantle, assuming lattice-preferred orientation of anisotropic minerals such as olivine. A vertical flow is characterized by a negative $\xi$ radial anisotropy (ratio between $V_{SH}$ and $V_{SV}$) and a small azimuthal anisotropy ($G \approx 0$). An upwelling (resp. downwelling) is characterized by a large positive (resp. negative) temperature anomaly inducing $\delta V_S < 0$ (resp. $\delta V_S > 0$).

A predominant large scale horizontal flow will be translated into a significant amplitude of the $G$ azimuthal anisotropy and its orientation will reflect the direction of flow (with a 180° ambiguity) (After Montagner, 2002).
Figure 7: Vertical cross-sections of the distribution of $V_S$, $\xi - \xi_{ref}$, $G$ in the Pacific Plate at $-20^\circ$ south and $20^\circ$ north between radii 6000km (370km depth) and 6350km (20km depth) (after Montagner, 2002). The color scales in percents are the same for $V_S$ and $\xi$.

The vertical scale is exaggerated to make the figures more readable.

a) $V_S$

b) $\delta \xi = \xi - \xi_{ref}$: deviation of $\xi$ with respect to a reference model ($\xi_{ref}$ PREM model)

c) $G$: amplitude of azimuthal anisotropy - parameters
Figure 8: Histogram of the difference between plate velocities directions and synthetic SKS anisotropy azimuths in the Pacific plate (After Montagner, 2002). It is calculated by summing the contributions of each grid point (5°x5°) weighted by the latitude and the amplitude of anisotropy.
Figure 9: variations of average S-wave velocity and azimuthal anisotropy (through the delay time of synthetic SKS splitting) plotted versus the age of lithospheric sea floor. In both cases the structure below the plate for age between 60 to 100Ma looks anomalous.
Figure 10: Scheme illustrating the difference in the location of maximum anisotropy between oceans and continents. Adapted from Gung et al. (2003).
Figure 11: Transition zone Heterogeneities
Model of shear velocity in the transition zone at 575km depth derived by inverting fundamental-mode and overtone Rayleigh wave phase velocities, and teleseismic body wave data (from Ritsema et al., 2004).
Figure 12: Mantle heterogeneities and convective flow below the Pacific Ocean. Adapted from Gaboret et al. (2003)
Figure A1: Definition of the Cartesian coordinate system (x,y,z) used in the calculations. Y is the azimuth of the wavevector with respect to North.

Figure A2: Partial derivatives for Love waves of the period of fundamental normal modes $0T_{40}$ (left) and $0T_{120}$ (right) with respect to the elastic coefficients of a transversely isotropic earth $L$, $N$ and density $\rho$, as a function of depth in the upper mantle (from Montagner and nataf, 1986). The partial derivatives with respect to $A$, $C$, $F$ are null for these modes. The plots are normalized to their maximum amplitudes, given for a $\Delta h=1000\text{km}$ thick perturbed layer. The combinations of elastic coefficients that have the same partial derivative as $L$ are $-G_c$, $-G_s$ for the azimuthal terms $2\Psi$, and as $N$ are $E_c$, $E_s$ for the azimuthal term $4\Psi$. Note that the amplitude of the $L$-partial is very small for the fundamental modes, which is not the case for higher modes.
Figure A3: kernels for Rayleigh waves. Same conventions as for figure A1 in the same depth range. The partial derivative with respect to \( N \) has not been plotted since its amplitude is very small for fundamental modes. Note that 3 partials contribute to the 2Y-azimuthal terms, A-partial for \( B_c, B_s, F \)-partial for \( H_c, H_s \) and the largest one L-partial for \( G_c, G_s \).