The Earth's Interior from both a Seismological and Petrological Perspective

by

Rebecca Lee Saltzer

BA Wellesley College, 1984

MS University of Oregon, 1996

Submitted to the Department of Earth, Atmospheric and Planetary Sciences

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2002

Signature of Author: _____

Department of Earth, Atmospheric and Planetary Sciences

September 10, 2001

Certified by: _____

Robert D. van der Hilst

Thesis Supervisor

Accepted by: _____

Ronald G. Prinn

Department Head

The Earth's Interior from both a Seismological and Petrological Perspective

by

Rebecca Lee Saltzer

Submitted to the Department of Earth, Atmospheric and Planetary Sciences on September 10, 2001 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Geophysics

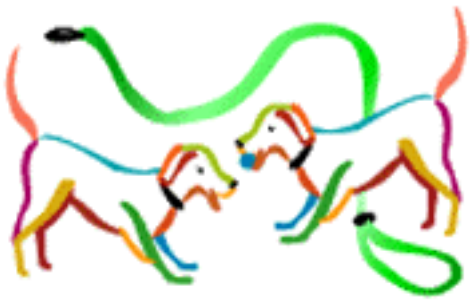
Abstract

Shear-wave splitting measurements of teleseismic shear waves, such as SKS, have been used to estimate the amount and direction of upper-mantle anisotropy worldwide. One of the basic assumptions in making these measurements is that the anisotropy is confined to a single, homogeneous layer. In this thesis, I use both numerical and analytical modeling to examine the validity of this assumption. I find that variability in the orientation of anisotropy with depth causes observable effects, such as frequency dependence in the apparent splitting parameters, and that the measured fast-axis direction is consistently different than the average of the medium.

A separate focus of this thesis is how spatial associations between minerals in a thin-section can be used to infer the evolutionary pressure-temperature history of a rock. I present a new method for textural analysis that uses digital images obtained with the electron microprobe. This method is used to characterize nine mantle xenoliths erupted from kimberlite pipes in South Africa and to test whether the pyroxenes are spatially correlated with the garnets. The observed associations can be explained by a model in which harzburgitic residues are produced by large extents of partial melting at shallow depths (~60-90 km) and high temperatures (~1300-1400° C) and are then subsequently dragged down to greater depths where garnet and clinopyroxene exsolve, perhaps in an Archean subduction zone.

The third focus of this thesis is on the seismological evidence for compositional heterogeneity in the lower mantle. Using reprocessed ISC data, I compare P and S wave tomographic models and find systematic differences between regions that have undergone subduction in the last 120 million years and those that have not below ~1500 km. This global study is followed up with a regional study using higher-quality P and S wave differential traveltimes. Beginning at depths of ~1000 km down and continuing down to the core-mantle boundary I find variability in Poisson's ratio that is greater than what would be expected by temperature variations alone. A simple explanation is that the variability includes a contribution from compositional effects, such as 2% variability in iron from one region to another.

Thesis Supervisor: Robert D. van der Hilst, Professor of Geophysics



To my wonderful friend and devoted companion, Tory (a.k.a. T-dog).

Table of Contents

Introduction

This thesis addresses three different and unrelated problems in earth science. The first is a theoretical look at wave propagation in vertically varying anisotropic media and how to interpret shear-wave splitting measurements in such a medium. Shear wave splitting measurements have been used to estimate the magnitude and direction of upper-mantle anisotropy worldwide, however one of the basic assumptions in making these measurements is that anisotropy is homogeneous throughout the layer. But what if the fast-axis direction does not point uniformly in one direction and instead varies as a function of depth? In chapter 1, I explore this question with both numerical experiments as well as analytic calculations. I investigate how variability in the fast-axis direction is "averaged" in shear-wave splitting measurements. I also examine whether there are any observables, such as frequency dependence of the shear-wave splitting measurements, that can be used to determine when the anisotropy is variable so that scientists making shear-wave splitting measurements in the real world can have more insight into how their measurements should really be interpreted.

The second problem I address in this thesis is how to use the spatial relationship between minerals to infer the pressure-temperature history of a rock. The solution to this problem required developing a new technique for 2-D textural analysis that used the raw pixel data of scanning back-scattered electron and x-ray images to determine modal amounts and average crystal sizes. This information was then used to determine whether any of the minerals were more closely associated spatially than would be expected if they were randomly oriented. I applied this technique to nine, low-temperature garnet-peridotites erupted from kimberlite pipes in South Africa that we suspected had formed at one pressure and temperature and then subsequently re-equilibrated at a different pressure and temperature. Eight of the nine rocks I analyzed showed a spatial relationship between opx, cpx, and garnet suggesting that two of the minerals had exsolved from the third. By recombining the three minerals together again, it is possible to see what this original rock might have been. In the South African samples, I found that the original rock was a harzburgite that

originated at ~100 km depth as a results of large extents of mantle melting whereas the final rock that was actually erupted had ended up (and equilibrated) at much greater depths in the lithosphere.

The third issue addressed in this thesis is whether there is seismological evidence for (or against) compositional heterogeneity in the mantle. Both geochemical and heatflow considerations suggest that compositionally distinct reservoirs have existed in the deep mantle for 1 billion years or more; however, the seismic evidence for the existence of such domains is more equivocal. I have looked at this issue (seismologically) from both a global perspective as well as a regional perspective. In the global study, I separate the world into those regions where there has been subduction in the last 120 million years and where there has not and examine $\partial \ln Vs / \partial \ln Vp$ as well as the correlation between the models to see whether the variability can simply be explained in terms of thermal variations in the mantle. One of the drawbacks of the global analysis is that it relies on routinely processed traveltime residuals with a lot of scatter so that some conclusions necessarily remain tentative. Therefore, I have also selected a regional corridor to study with higher-quality, waveform derived traveltimes. In the regional study, I use estimates from the mineral physics community to explore the relative effects of temperature and composition (Fe and Si) on the resulting tomographic images.

Chapter 2

How are vertical shear-wave splitting measurements affected by variations in the orientation of azimuthal anisotropy with depth?

Published by Blackwell Publishers in *Geophysical Journal International* by Rebecca Saltzer, James Gaherty and Tom Jordan, *141*, 374-390, 2000.

Summary

Splitting measurements of teleseismic shear waves, such as SKS, have been used to estimate the amount and direction of upper-mantle anisotropy worldwide. These measurements are usually made by approximating the anisotropic regions as a single, homogeneous layer and searching for an apparent fast direction ($\tilde{\phi}$) and an apparent split time ($\Delta\tilde{t}$) by minimizing the energy on the transverse component of the back-projected seismogram. In this paper, we examine the validity of this assumption. In particular, we use synthetic seismograms to explore how a vertically varying anisotropic medium affects shear-wave splitting measurements. We find that weak heterogeneity causes observable effects, such as frequency dependence of the apparent splitting parameters. These variations can be used, in principle, to map out the vertical variations in anisotropy with depth through the use of Fréchet kernels that we derive using perturbation theory. In addition, we find that measurements made in typical frequency bands produce an apparent orientation direction that is consistently different than the average of the medium and weighted toward the orientation of the anisotropy in the upper portions of the model. This tendency of the measurements to mimic the anisotropy at the top part of the medium may explain why shear-wave splitting measurements tend to be correlated with surface geology. When the heterogeneity becomes stronger, multiple scattering reduces the amplitude of the tangential-component seismogram and the associated split time, so that a null result may be obtained despite the fact that the waves have traveled through a strongly anisotropic medium. Regardless of the amount of vertical heterogeneity, we find that there is very little dependence on backazimuth for the measured fast-axis direction or split time if the top and bottom halves of the medium average to similar fast-axis directions. If, however, the average fast-axis direction in the top half of the model differs from that in the bottom half then split time measurements will show a significant dependence on backazimuth, but fast-axis direction measurements will remain relatively constant.

Introduction

Measurements of seismic anisotropy are used to infer mantle deformation and flow patterns. While several different methods for constraining upper-mantle anisotropy have been developed, such as *Pn* refraction surveys (e.g., Raitt et al., 1969; Shearer & Orcutt, 1986) and surface-wave polarization analyses

(e.g., Forsyth, 1975, Nataf et al., 1984; Tanimoto & Anderson, 1985; Montagner & Tanimoto, 1990), the last decade has seen an explosion in shear-wave splitting studies using vertically propagating shear waves (see reviews by Silver, 1996, and Savage, 1999). Typically, these analyses are performed on waves such as *SKS* or *SKKS* because they have a known polarization direction (S_V) as a result of passing through the liquid outer core. The standard procedure is to find the inverse splitting operator Γ^{-1} which, when applied to the observed waveform, minimizes the energy on the tangential component (Silver & Chan, 1991). When other phases such as *S* and *ScS* are used, the splitting parameters are found either by assuming a rectilinear source mechanism (Ando & Ishikawa, 1982, Ando, 1984) or by explicitly diagonalizing the covariance matrix of surface-corrected horizontal particle motions (Vidale, 1986; Fouch & Fischer, 1996).

A basic assumption in interpreting measurements using these techniques is that the splitting operator Γ corresponds to a single homogeneous layer in which the anisotropy has a horizontal symmetry axis and a constant magnitude. The parameters used to describe this model (the splitting parameters) are the polarization azimuth of the fast eigenwave, ϕ, and the travel-time difference between the fast and slow eigenwaves, Δt. It is straightforward to construct the splitting operator for an arbitrary stack of layers with depth-dependent properties and more general forms of anisotropy using propagator matrices (e.g., Keith & Crampin, 1977; Mallick & Frazer, 1990). But it is less clear how one might use such constructions to make inferences about anisotropic structure. A potentially fruitful direction is to fit the waveform data by optimizing the homogenous-layer operator and then interpret the two recovered quantities, denoted here by $\tilde{\phi}$ and $\Delta\tilde{t}$, as apparent splitting parameters that are functionals of the vertical structure. This approach was adopted by Silver & Savage (1994), who showed how an approximation to the variation of $\tilde{\phi}$ and $\Delta\tilde{t}$ with the incident polarization angle could be inverted for a two-layer anisotropic model. They also discussed the generalization of their approximate functional relations, which are valid for forward scattering at low frequencies (wave periods >> Δt), to an arbitrary layer stack. Rümpker and Silver (1998) have recently expanded this theoretical discussion of vertical heterogeneity to include expressions for the

apparent splitting parameters valid at high frequencies, as well as some statistical properties of the parameters for random layer stacks, and they have tested various aspects of their theory with numerical calculations.

In this paper, we consider several additional aspects of this interpretation problem. Using a propagator-matrix method that includes both forward- (upgoing) and back-scattered (downgoing) waves, we compute synthetic seismograms for different types of depth dependence, including smooth models as well as those with discontinuous variations in the anisotropy axis. We investigate the behavior of the apparent splitting parameters with increasing amounts of vertical heterogeneity in the azimuthal anisotropy and use the results to define three wave-propagation regimes corresponding to weak, intermediate, and strong scattering. For weakly heterogeneous media, we employ perturbation theory to calculate the sensitivity (Fréchet) kernels for band-limited, apparent-splitting measurements and show how these measurements sample the depth dependence as a function of frequency and incident polarization angle. In realistic situations, the center frequencies of the observations are sufficiently small that the kernels are one-sided, and we can define an apparent depth of sampling that we demonstrate is biased towards the upper part of the structure. In principle, the Fréchet kernels can be used to set up the problem of inverting frequency-dependent splitting measurements for depth-dependent anisotropy. We show that in practice, however, strong scattering by vertical heterogeneity can invalidate the assumptions that underlie this linearized approach, especially at higher frequencies.

Problem Formulation

Because our purpose is to investigate some elementary aspects of vertical shear-wave propagation, we adopt a very simple model for the mantle comprising a heterogeneous, anisotropic layer of thickness d overlying a homogeneous, isotropic half-space (Fig. 1).

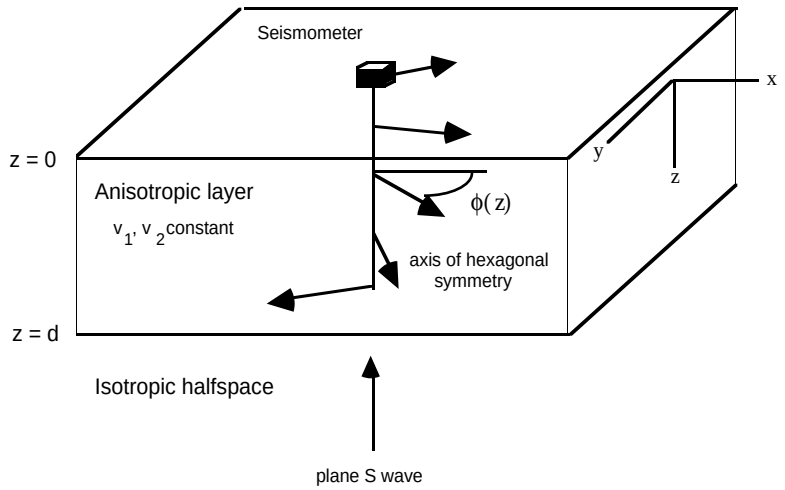


Figure 1. Model used in the calculations. A vertically traveling, rectilinearly polarized shear wave impinges at depth d on the base of a heterogeneous, anisotropic layer in which the fast axis direction ϕ varies as a function of depth z. The two eigenvelocities, v_1 and v_2, are constant throughout the layer, and the velocity of the isotropic half-space is taken to be equal to their mean.

The anisotropy is assumed to be hexagonally symmetric with a horizontal axis of symmetry, and the lateral heterogeneity is assumed to be sufficiently smooth that horizontal gradients in the wave velocities can be ignored. Vertically propagating shear waves can thus be represented as linear combinations of orthogonal eigenwaves with shear velocities v_1 and v_2 that depend on the depth coordinate z. To simplify the problem further, we assume that the mean velocity $\bar{v} = (v_1 + v_2)/2$ and the velocity difference $\Delta v = v_1 - v_2$ are constants, and we label the eigenwaves such that $\Delta v > 0$. The heterogeneity in the medium is specified by a single function of depth that we take to be the azimuth of the fast (v_1) axis, $\phi(z)$, measured clockwise from the x axis.

For the calculations in this paper, we adopt a layer thickness of $d = 200$ km and a mean velocity of $\bar{v} = 4.54$ km/s, and we take the velocity of the isotropic half-space to equal this mean velocity. The maximum splitting time for shear waves propagating from the base of the anisotropic layer to the surface—we ignore the crust—is $\Delta t = d(v_1 - v_2)/v_1 v_2 \approx d\Delta v/\bar{v}^2$, which represents the "splitting strength" of the model. We refer to Δt in some of our numerical experiments as the "true" split time.

2.1 *Forward Problem: Synthetic Seismograms*

We use a stack of thin, homogeneous layers to represent the medium and a propagator-matrix method to propagate shear waves vertically through the layers (e.g., Kennett, 1983). For all our calculations, these layers are less than one km thick to ensure that seismic wavelengths do not approach layer thickness. Boundary conditions restrict displacements and tractions to be continuous at the interfaces between the layers and for tractions to be zero at the surface. The Fourier-transformed, density-normalized stress vector $\tau(z,\omega) = \rho^{-1}\,[T_{xz}\; T_{yz}]^{\mathrm{T}}$ is related to the depth derivative of the displacement vector $\mathbf{u}(z,\omega) = [u_x\; u_y]^{\mathrm{T}}$ by the Christoffel matrix

$$\mathbf{C}(z) \;=\; \begin{bmatrix} v_1^2\cos^2\phi(z) + v_2^2\sin^2\phi(z) & (v_1^2 - v_2^2)\cos\phi(z)\sin\phi(z) \\ (v_1^2 - v_2^2)\cos\phi(z)\sin\phi(z) & v_1^2\sin^2\phi(z) + v_2^2\cos^2\phi(z) \end{bmatrix}. \tag{1}$$

The equations of motion are $\partial_z f = \mathcal{A} f$, where the displacement-stress vector and system matrix are given by

$$f(z,\omega) \;=\; \begin{bmatrix} \mathbf{u}(z,\omega) \\ \mathbf{t}(z,\omega) \end{bmatrix}, \tag{2}$$

$$\mathcal{A}(z,\omega) \;=\; \begin{bmatrix} \mathbf{0} & \mathbf{C}^{-1}(z) \\ -\omega^2\mathbf{I} & \mathbf{0} \end{bmatrix}. \tag{3}$$

The rotation operator

$$\mathbf{U}(\phi(z)) \;=\; \begin{bmatrix} \cos\phi(z) & -\sin\phi(z) \\ \sin\phi(z) & \cos\phi(z) \end{bmatrix} \tag{4}$$

diagonalizes the Christoffel matrix: $\hat{\mathbf{C}} \equiv \mathrm{diag}[v_1^2, v_2^2] = \mathbf{U}\,\mathbf{C}\,\mathbf{U}^{\mathrm{T}}$.

The propagator matrix for this problem and some of its approximations are discussed in Appendix A. For an upgoing wave $\mathbf{u}_I(\omega)$ incident at the base of the anisotropic layer, the free-surface displacement vector can be written

$$\mathbf{u}(0,\omega) \;=\; [\mathbf{P}_{uu} + i\omega\bar{\nu}\mathbf{P}_{u\tau} + (\mathbf{P}_{uu} - i\omega\bar{\nu}\mathbf{P}_{u\tau})\mathbf{R}]\mathbf{u}_I(\omega) \qquad \text{(A16)}$$

where $\mathbf{P}_{uu}$ and $\mathbf{P}_{u\tau}$ are 2×2 submatrices of the propagator matrix (A7), $\mathbf{R}$ is the 2×2 matrix of reflection coefficients (A15), and $\mathbf{U}_z = \mathbf{U}(\phi(z))$.

The pulse shape at the base of the anisotropic layer in all of our calculations is taken to be of the form $u_I(t) = \exp[-a/(t-t_0) - (t-t_0)/b]\; H(t-t_0)$ with a duration a = 2 s and a decay constant b = 4 s. The convolution of this initial pulse shape with the broadband instrument response and prefilter is given in Fig. 2a. There is no energy on the tangential component because the initial pulse is radially polarized; however, propagation of the pulse through an anisotropic layer produces energy on both the radial and tangential components of the surface seismogram. An example of a synthetic seismogram calculated for a homogeneous anisotropic layer is shown in Fig. 2b for velocity contrast of $\Delta v/\bar{v}$ = 4.54%. This corresponds to a splitting strength of Δt = 2 s, which lies toward the high end of the observations summarized by Silver (1996) and Savage (1999). The azimuth of the fast axis, measured clockwise from the radial ($\hat{\mathbf{x}}$) direction, is 45°. The seismograms in this example are "broadband" with a corner at 50 mHz and at 300 mHz and a center frequency of ~140 mHz.

A simple model of a heterogeneous anisotropic layer, used extensively in our numerical illustrations, is one in which $\phi(z)$ varies linearly with depth:

$$\phi(z) \;=\; \phi_0 + \kappa z \;=\; \phi_d + \kappa(z-d) \qquad (5)$$

The constant $\kappa = d\phi/dz$ is the vertical rotation rate. The orientation of the fast axis varies from ϕ_0 at the surface to ϕ_d at the base of the layer, and the total change in the orientation $\Delta\phi \equiv \phi_d - \phi_0 = \kappa d$ measures the strength of the heterogeneity. Figures 2c-e show synthetic seismograms for the linear- rotation model with a splitting strength of Δt = 2 s and increasing amounts of heterogeneity: $\Delta\phi$ = 30°, 100°, and 1000°, respectively. In all three examples, the incident polarization was chosen such that the mean orientation of the fast axis was 45°; i.e.,

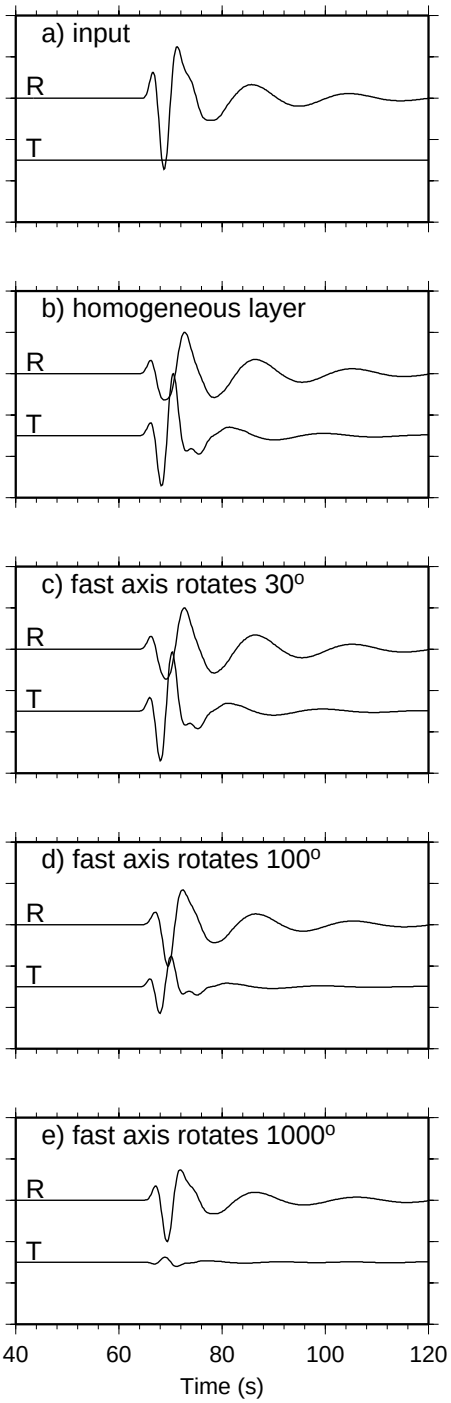


Figure 2. Radial and Tangential component seismograms at a) the base of the anisotropic layer b) the surface after passing through homogeneous anisotropic layer c) surface after passing through weakly heterogeneous anisotropic layer in which the fast axis direction ($\Delta\phi$) linearly rotates 30°) intermediate heterogeneous anisotropic layer in which the fast axis direction ($\Delta\phi$) linearly rotates 100°) strongly heterogeneous anisotropic layer in which the fast axis direction ($\Delta\phi$) linearly rotates 1000°. _ Seismograms are bandpass filtered with a butterworth filter to frequencies between 50 mHz and 300 mHz.

$$\bar{\phi} \equiv \frac{1}{d}\int_0^d \phi(z)\,dz = \phi_0 + \Delta\phi/2 = \pi/4. \quad (6)$$

In the case where $\Delta\phi$ = 30°, the surface seismograms (Fig. 2c) are very similar to those produced in the homogeneous anisotropic case (Fig. 2b). Both the radial and tangential component seismograms have a comparable amount of energy. As $\Delta\phi$ increases to 100°, the energy on the radial component becomes greater than that on the tangential component (Fig. 2d), and when $\Delta\phi$ is as large as 1000°, the net effect of the anisotropy is to put very little energy onto the tangential component (Fig. 2e).

Inverse Problem: Apparent Splitting Parameters

If the incident pulse is known to be radially polarized, then the apparent splitting parameters $\tilde{\phi}$ and $\Delta\tilde{t}$ can be defined as the values that minimize the energy on the transverse ($\hat{\mathbf{y}}$) component of the displacement field back-projected to $z = d$ using a homogeneous-layer splitting operator (Silver & Chan, 1991). Parseval's theorem allows the transverse-component energy to be written as the frequency-domain integral:

$$\varepsilon^2(\phi', \Delta t') = \int_{-\infty}^{\infty} |\hat{\mathbf{y}}^{\mathsf{T}} \Gamma_{\mathrm{h}}^{-1}(\phi', \Delta t')\,\mathbf{u}(0,\omega)|^2\, d\omega \,. \quad (7)$$

Here Γ_{h}^{-1} is the inverse of the splitting operator given by eqn. (A24). In practice, determination of these so-called splitting parameters requires a search over a grid of fast-axis directions and delay times. For typical teleseismic observations, these parameters can be determined to within +/-10° and 0.15 s (Fouch & Fischer, 1996, Silver & Chan, 1991). Figure 3 shows the "energy map" contoured as a function of ϕ' (0° to 180°) and $\Delta t'$ (0 to 4 s) for the three linear-gradient examples shown in the previous section and Figure 4 shows the inferred seismograms at the base of the layer, calculated by back-projecting the splitting parameters and assuming propagation of waves through a homogeneous, anisotropic layer.

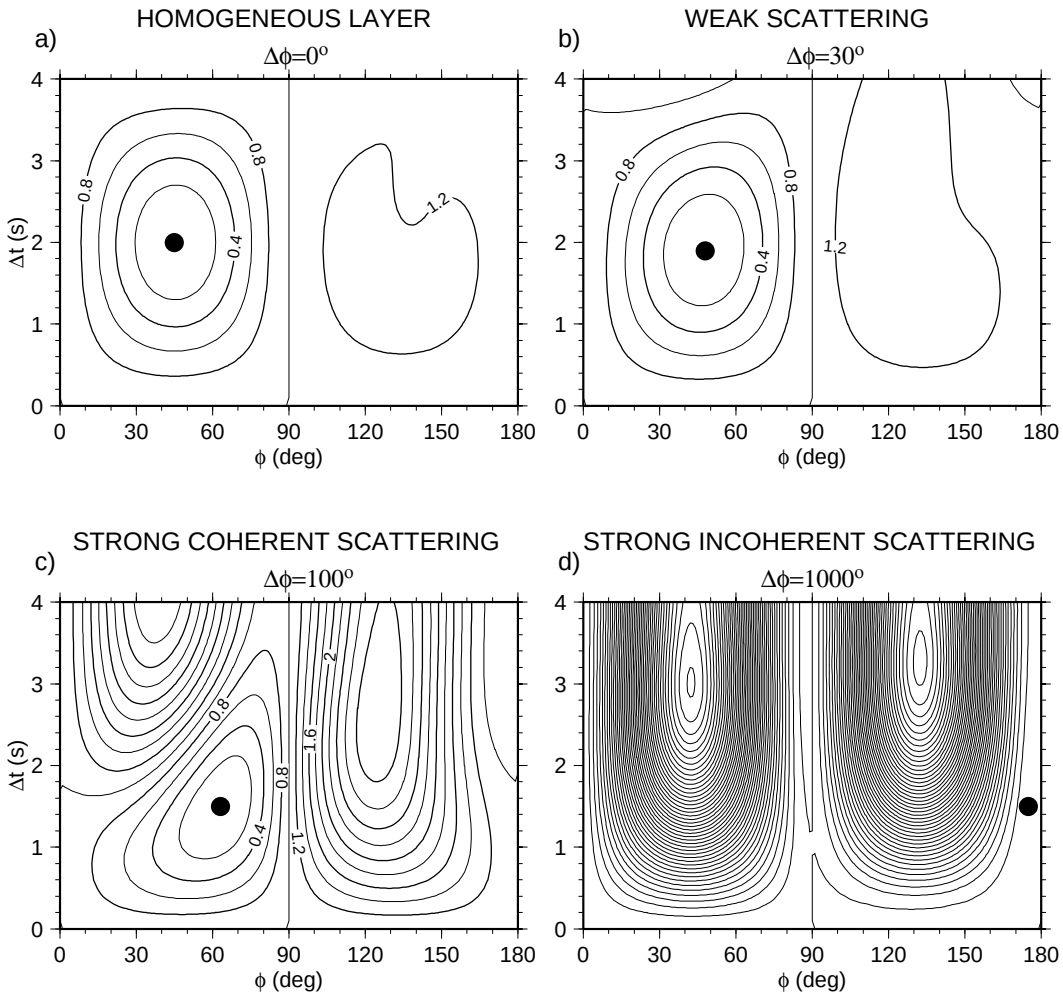


Figure 3. Energy diagrams for a) homogeneous anisotropic layer b) weakly heterogeneous anisotropic layer ($\Delta\phi$=30°) intermediate heterogeneous anisotropic layer ($\Delta\phi$=100°) and d) strongly heterogeneous anisotropic layer ($\Delta\phi$=1000°).

The energy map for the seismograms computed for a homogeneous layer (Fig. 3a) shows a well-defined minimum at the correct values of splitting parameters ($\tilde{\phi}$ = 45°, $\Delta\tilde{t}$ = 2 s). For weak heterogeneity ($\Delta\phi$ = 30°), the energy minimum remains close to the layer mean ($\tilde{\phi}$ = 48°, $\Delta\tilde{t}$ = 1.9 s), and the bulk of the tangential component energy has been removed (Fig. 4c). At intermediate values of the heterogeneity ($\Delta\phi$ = 100°), the energy minimum is still well defined (Fig. 3c), but it is displaced away from the layer mean by 18° ($\tilde{\phi}$ = 63°). Moreover, the scattering from the vertical gradients in the anisotropy is sufficient to reduce the apparent splitting time significantly below its true

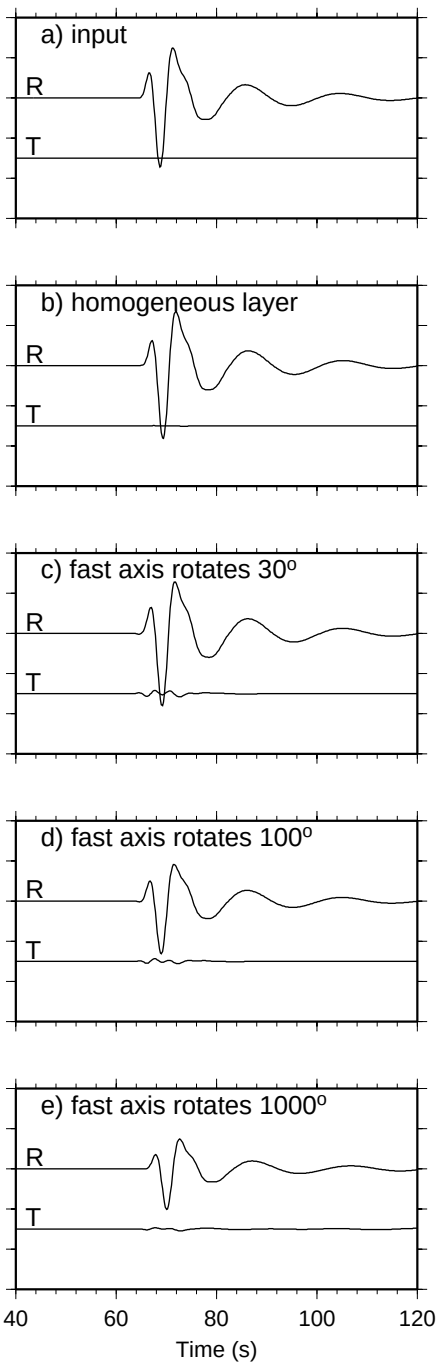


Figure 4. Radial and Tangential component seismograms. After determining the apparent splitting parameters, the seismograms can be back-projected to the base of the layer again using those same parameters and assuming a homogeneous layer. a) input pulse at the base of the anisotropic layer to be compared with back-projected seismograms from b) homogeneous anisotropic layer c) weakly heterogeneous anisotropic layer ($\Delta\phi$ linearly rotates 30°) d) intermediate heterogeneous anisotropic layer ($\Delta\phi$) linearly rotates 100°) e) strongly heterogeneous anisotropic layer ($\Delta\phi$) linearly rotates 1000°). Seismograms are bandpass filtered with a butterworth filter to frequencies between 50 mHz and 300 mHz.

value ($\Delta\tilde{t}$ = 1.5 s). When the heterogeneity gets to be very large ($\Delta\phi$ = 1000°), the scattering is sufficiently strong as to cause destructive interference that nearly wipes out the arrivals on the transverse component (Fig. 2e). The resulting energy map (Fig. 3d) is characteristic of a "null measurement", with the lowest values occurring near the horizontal axis where $\Delta t'$ = 0 and along vertical ridges corresponding to the degenerate azimuths of ϕ' = 0° and 90°.

The behaviors illustrated in Fig. 3 are typical of three scattering regimes that can be qualitatively described as "weak," "strong coherent", and "strong incoherent".

Weak-Scattering Regime

When the scattering is weak, the effects of the heterogeneity on the apparent splitting parameters can be approximated with a linearized perturbation theory. In this section, we derive analytical expressions for perturbations from a homogeneous starting model, test their applicability with numerical calculations, and use them to gain insight into how the sensitivity of the apparent splitting parameters varies with depth.

Fréchet Kernels

The linearized equations that relate a structural perturbation $\delta\phi(z)$ to perturbations in the apparent splitting parameters, $\delta\tilde{\phi}$ and $\delta\tilde{t}$, can be written as integrals over the layer:

$$\delta\tilde{\phi} \approx \int_0^d G_\phi(z)\,\delta\phi(z)\,dz, \tag{8}$$

$$\delta\tilde{t} \approx \int_0^d G_t(z)\,\delta\phi(z)\,dz \tag{9}$$

This approximation ignores terms of order $\delta\phi^2$. $G_\phi(z)$ and $G_t(z)$ are the sensitivity functions, or Fréchet kernels. They generally depend on the structural model that is being perturbed, as well as the spectral properties of the waves

being measured. For the structures considered here, a model is specified by the fast-axis orientation function $\{\phi(z)\colon 0 \le z \le d\}$ and the two velocity constants v_1 and v_2.

The Fréchet kernels can be numerically approximated for an arbitrary starting model by computing the small change in the splitting parameters due to a small perturbation in the fast-axis orientation distributed over a thin layer. However, in the special case of a homogeneous starting model ($\phi(z) = \phi_0$), the kernels for narrow-band pulses can be derived analytically. The details are relegated to Appendix B. The approximate results for a pulse with center frequency ω_0 and half-bandwidth σ are,

$$G_\phi \;=\; \frac{g_4 + [g_5 \cos 4\phi_0 + g_6](\sigma^2/\omega_0^2)}{g_1 + (g_2 \cos 4\phi_0 + g_3)(\sigma^2/\omega_0^2)} \Delta k \;+\; O\!\left(\frac{\sigma^4}{\omega_0^4}\right), \tag{10}$$

$$G_t \;=\; \frac{g_7 + [g_8 \cos 4\phi_0 + g_9](\sigma^2/\omega_0^2)}{g_1 + (g_2 \cos 4\phi_0 + g_3)(\sigma^2/\omega_0^2)} \Delta k \cot 2\phi_0 \;+\; O\!\left(\frac{\sigma^4}{\omega_0^4}\right), \tag{11}$$

where $\Delta k = \omega_0 (v_2^{-1} - v_1^{-1})$ is the differential wavenumber. The three parameters appearing in the denominator of these expressions are independent of depth z:

$$g_1 \;=\; (1 - \cos \Delta k d)^2 \tag{12a}$$

$$g_2 \;=\; \Delta k d [\sin^2 \Delta k d + (\Delta k d - 2) \sin \Delta k d \cos \Delta k d] \tag{12b}$$

$$g_3 \;=\; g_2 + 2(\Delta k d)^2 [\sin^2 \Delta k d + (1 - \cos \Delta k d) \cos \Delta k d] \tag{12c}$$

The other six can be written in terms of trigonometric functions of the height variable $r = d - z$:

$$g_4(r) \;=\; (1 - \cos \Delta k d) \sin \Delta k r \tag{13a}$$

$$g_5(r) \;=\; (\Delta k d - 1) \sin \Delta k d \cos \Delta k r + (\Delta k d - \Delta k d \Delta k r) \sin \Delta k d \sin \Delta k r - \Delta k d \cos \Delta k d \cos \Delta k r \tag{13b}$$

$$g_6(r) = g_5(r) + (\Delta kd)^2 \cos\Delta kd \sin\Delta kr + 2(\Delta kd)(\Delta kr)\sin\Delta kd \cos\Delta kr - (\Delta kr)^2(1-\cos\Delta kd)\sin\Delta kr \quad (13c)$$

$$g_7(r) = (1-\cos\Delta kd)^2 \cos\Delta kr - (1-\cos\Delta kd)\sin\Delta kd \sin\Delta kr \quad (14a)$$

$$g_8(r) = \sin\Delta kd(\sin\Delta kd - 2\cos\Delta kd)[\Delta kd\cos\Delta kr - (\Delta kd)(\Delta kr)\sin\Delta kr] + (\Delta kd)^2 \sin\Delta kd \cos\Delta kd \cos\Delta kr \quad (14b)$$

$$g_9(r) = g_8(r) + 2(\Delta kd)^2[\sin^2\Delta kd + (1-\cos\Delta kd)\cos\Delta kd]\cos\Delta kr - (1-\cos\Delta kd)^2[\Delta kr\sin\Delta kr + (\Delta kr)^2\cos\Delta kr] \quad (14c)$$

The relative bandwidth σ/ω_0, which is less than 0.5 in most seismological applications Hz (e.g. Silver, 1996; Fouch and Fischer, 1996; Wolfe and Solomon, 1998), is sufficiently small that it is safe to ignore the fourth-order terms in (10) and (11). Fig. 5 displays kernels computed under this approximation for a range of initial polarizations and center frequencies. The sinuosity of the kernels increases with frequency, reflecting the first-order trigonometric dependence on $\Delta k\, z$. The kernel for the apparent splitting azimuth satisfies the lower boundary condition, $G_\phi(d) = 0$, which can be verified from the analytical expressions.

Integration of these expressions show that the kernel for $\delta\tilde{\phi}$ is unimodular and the kernel for $\delta\tilde{t}$ averages to zero:

$$\int_0^d G_\phi(z)\,dz = 1, \quad (15)$$

$$\int_0^d G_t(z)\,dz = 0. \quad (16)$$

These properties must apply to the exact forms of the Fréchet kernels, not just to their narrow-band approximations given by (10) and (11), because a constant perturbation maintains the structural homogeneity of the starting model. In other words, setting the perturbation in (8) and (9) to a constant value of $\delta\phi$ must always yield $\delta\tilde{\phi} = \delta\phi$ and $\delta\tilde{t} = 0$.

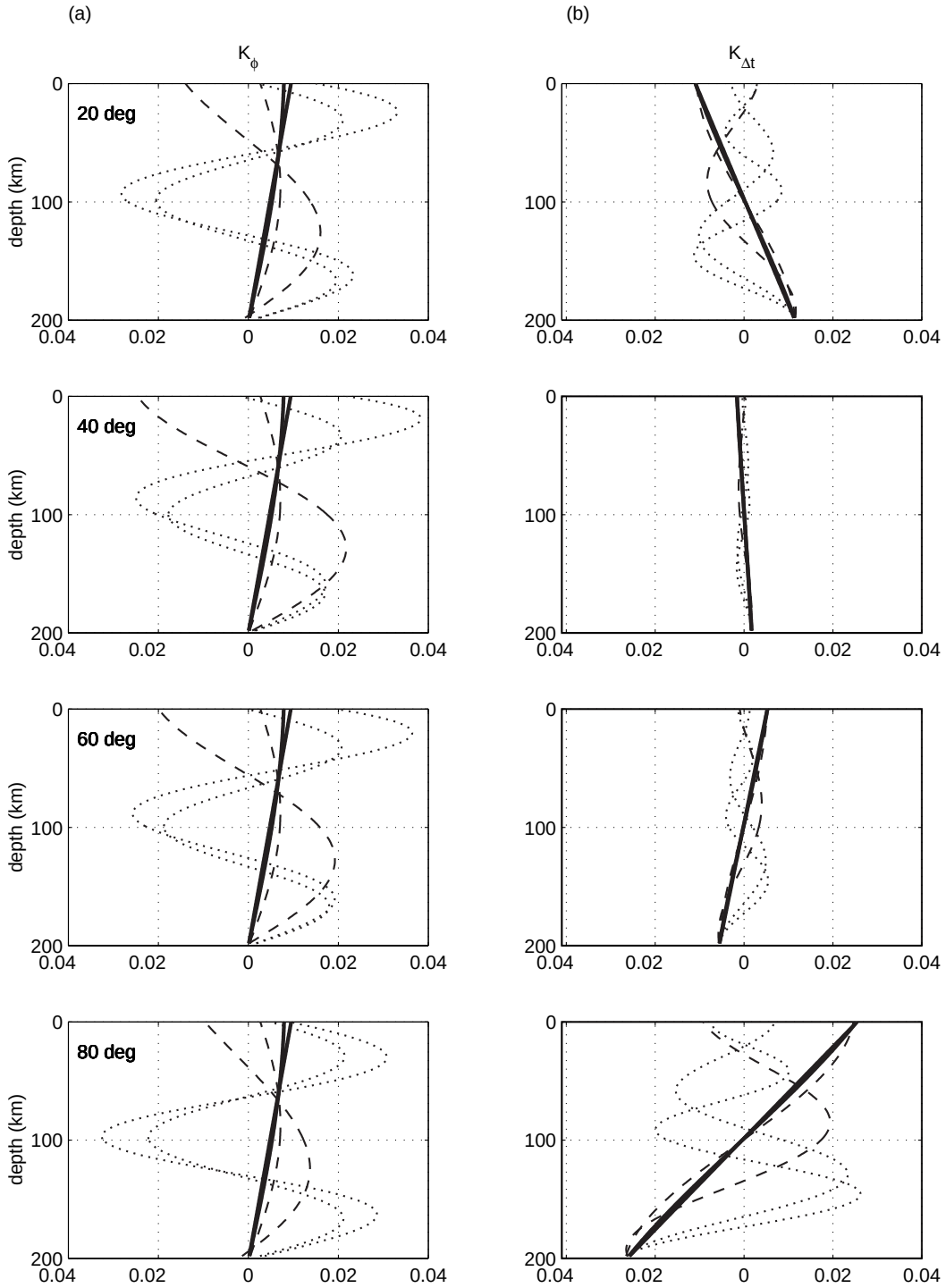


Figure 5. Fréchet kernels for the apparent splitting parameters computed from eqns. (10) and (11) to second order in the relative bandwidth σ/ω_0. The reference model is a 200-km thick homogeneous layer with Δt = 2 s. Left panels show the apparent azimuth kernels $G_\phi(z)$, and right panels show the apparent splitting time kernels $G_t(z)$. The initial polarization ϕ_0 increases downward from 20° (top panel) to 80° (bottom panel). In each panel, the center frequencies ω_0 range from 0.1 Hz to 0.8 Hz (lowest frequencies are solid lines, intermediate are dashed and highest frequencies are dotted), while the relative bandwidth is held constant at σ/ω_0 = 0.125.

Limiting Forms of the Kernels

By considering the zero-bandwidth limit, we gain additional insight into the nature of the sensitivity kernels:

$$G_\phi^0(z) \equiv \lim_{\sigma\to 0} G_\phi(z) = \frac{\sin\Delta k(d-z)}{1-\cos\Delta k d}\Delta k \tag{17}$$

$$G_t^0(z) \equiv \lim_{\sigma\to 0} G_t(z) = \left[\cos\Delta k(d-z) - \frac{\sin\Delta k d \sin\Delta k(d-z)}{1-\cos\Delta k d}\right]\Delta k \cot\phi_0 \tag{18}$$

The splitting-time kernel (18) varies like the cotangent of $2\phi_0$ and is thus singular at $\phi_0 = n\pi/2$, which corresponds to incident polarizations aligned with an eigenwave orientation in the unperturbed model. In contrast to this singular behavior, the splitting-azimuth kernel (17) is independent of the initial polarization, and $\delta\tilde{\phi}$ remains well defined even at its degenerate values.

This important theoretical point deserves special emphasis. A shear wave with a polarization aligned with one of the eigenwave directions is not split by propagation through the reference model, and the tangential-component energy of the back-projected displacement field is thus identically zero at $\phi' = \phi_0 = n\pi/2$ for arbitrary values of $\Delta t' > 0$. Consequently, the energy map displays vertical nodal lines at these azimuths, as well as a horizontal nodal line at $\Delta t' = 0$ and the inversion of the seismograms for the apparent splitting parameters via the minimization of (7) becomes unstable (e.g., Silver & Chan, 1991). Nevertheless, the apparent splitting azimuth remains formally defined in this limit by the orientation of the appropriate vertical node; i.e., $\tilde{\phi}$ equals either ϕ_0 or $\phi_0 + \pi/2$. Moreover, $\tilde{\phi}$ is Fréchet differentiable, because a small perturbation to the model will result in a small, well-defined perturbation of the node in the ϕ' direction. The splitting-time functional $\Delta\tilde{t}$, on the other hand, is not Fréchet differentiable at the nodes, which is why its kernel is singular. This behavior generalizes to pulse shapes with finite bandwidths, as is evident from eqn. (11).

The cotangent dependence of the splitting-time kernel also implies that $G_t(z) = 0$ for $\phi_0 = n\pi/4$, which means that at polarization angles near 45° the apparent

splitting time is only weakly dependent on perturbations to the local splitting orientation.

Equations (17) and (18) show that, in the zero-bandwidth limit, both kernels become unbounded at $\Delta kd = 2n\pi$, where the total splitting time Δt is an integral multiple of the wave period $2\pi/\omega_0$. A finite bandwidth introduces a weak ($\sim\sigma^2/\omega_0^2$) dependence of G_ϕ and G_t on the initial azimuth through the $\cos 4\phi_0$ term, which suppresses this resonance singularity.

The expressions for the kernels presented thus far apply to splitting strengths that are arbitrarily large. At low frequencies, when the splitting strength is much less than the wave period, the maximum phase shift between the two eigenwaves is small, $\Delta kd = \omega_0 \Delta t << 1$. If we expand the trigonometric functions and retain only the leading terms, the single-frequency kernels become linear functions of the depth:

$$G_\phi^0(z) \approx \frac{2(d-z)}{d^2}, \tag{19}$$

$$G_t^0(z) \approx 2\Delta t\left(\frac{2z-d}{d^2}\right)\cot\phi_0. \tag{20}$$

Thus, in the low-frequency limit, the apparent splitting azimuth is insensitive to heterogeneity at the base of the layer and most sensitive to heterogeneity at the top of the layer. The sensitivity of the apparent splitting time to azimuthal heterogeneity is zero in the middle of the layer, and it is of equal magnitude and opposite sign at the top and bottom of the layer.

Apparent Depth of Sampling

The previous discussion shows that the Fréchet kernel for the apparent splitting azimuth will be nonnegative when the center period of the wavegroup, $2\pi/\omega_0$, is greater than or equal to twice the splitting time Δt (i.e., $\Delta kd \le \pi$). Under this condition, which applies to most observations of teleseismic shear-wave splitting, we can define an apparent depth of sampling by the centroid of the kernel:

$$z_{\text{app}} = \int_0^d G_\phi(z)\, z\, dz. \tag{20}$$

Using (15), we obtain for the zero-bandwidth limit,

$$z_{\text{app}}^0 = \lim_{\sigma \to 0} z_{\text{app}} = \frac{\sin \Delta k\, d - \Delta k\, d \cos \Delta k\, d}{\Delta k\, d (1 - \cos \Delta k\, d)}. \tag{21}$$

In the low-frequency limit, where the $\tilde{\phi}$ kernel becomes a linear function of depth, the apparent depth of sampling goes to one-third the layer thickness. This value increases to half the layer thickness as the center period approaches the splitting time. Figure 6 shows three low-frequency $\tilde{\phi}$ kernels and their apparent depths of sampling for a 200-km thick, anisotropic layer. In the low-frequency limit, the apparent depth of sampling is just 66 km whereas for data low-pass filtered at 100 mHz, the depth of sampling increases to 71 km and for data low-pass filtered at 200 mHz, it increases to 88 km.

This bias in the sensitivity to near-surface structure explains why the recovered value for $\tilde{\phi}$ in the weak scattering case (Fig. 3b) is greater than the layer mean by about 3°. In this model, the fast axis rotates linearly from 40° at

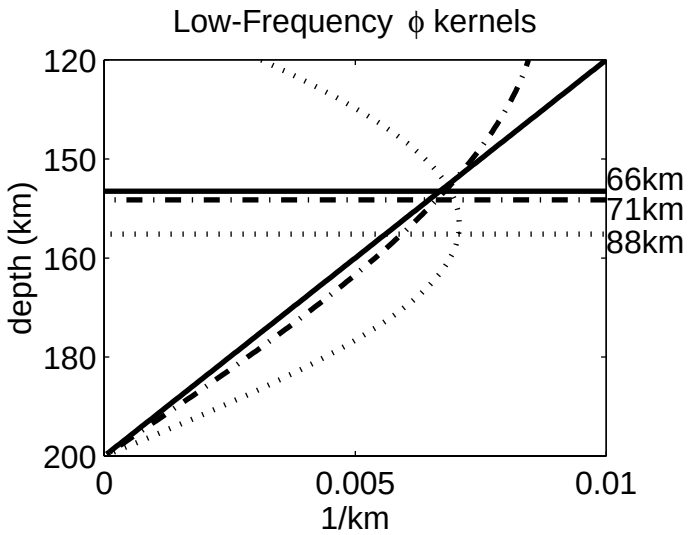


Figure 6. Low frequency Fréchet kernels and the apparent depth of sampling computed with eqn. 20. Solid line is the kernel in the low-frequency limit. Dashed line is the low-passed kernel for 100 mHz and the dotted line is the low-passed kernel at 200 mHz.

the base of the layer to 50° at the top. The kernels predict that the upper part of the model, where the fast axis ranges between 45° and 50°, will dominate the shear-wave splitting measurement and indeed, the value found numerically ($\tilde{\phi}$ = 48°) agrees with this prediction, corresponding to the fast axis direction approximately one-third of the way down the layer. This increase in sensitivity to the fast-axis direction near the surface may explain why shear-wave splitting measurements tend to correlate with tectonic deformation observed at the surface (Silver, 1996). Most shear-wave splitting measurements are made on seismograms with relatively low center frequencies (< 200 mHz), so that the apparent depth of sampling is less than the mean thickness of the layer. There are significant variations in the center frequencies used by different researchers, however, so this apparent depth varies from study to study.

Numerical Tests

We conducted a series of numerical experiments to test the perturbation theory derived for a homogeneous model. To investigate the sensitivity of the kernels to the reference structure, we have computed them by numerical perturbation to heterogeneous starting models. The results for starting models with a linear gradient and a step-wise discontinuity in $\phi(z)$ are compared with the homogeneous-layer case in Fig. 7. The average orientation was chosen to be the same for all three models, $\bar{\phi} \equiv d^{-1}\int_0^d \phi(z)\,dz$ = 45°, while the total variation in $\phi(z)$ was taken to be 20° for the two heterogeneous models. The kernels are very similar, indicating only a weak dependence on the starting model when the heterogeneity is of this magnitude. In particular, the apparent depths of sampling for the layered and linear model are 91 km and 90 km respectively, essentially the same as the value of 88 km calculated for the homogeneous model. The properties found for the homogeneous-layer kernels, such as their dependence on frequency, bandwidth, and incidence azimuth, should therefore pertain more generally in the weak scattering regime.

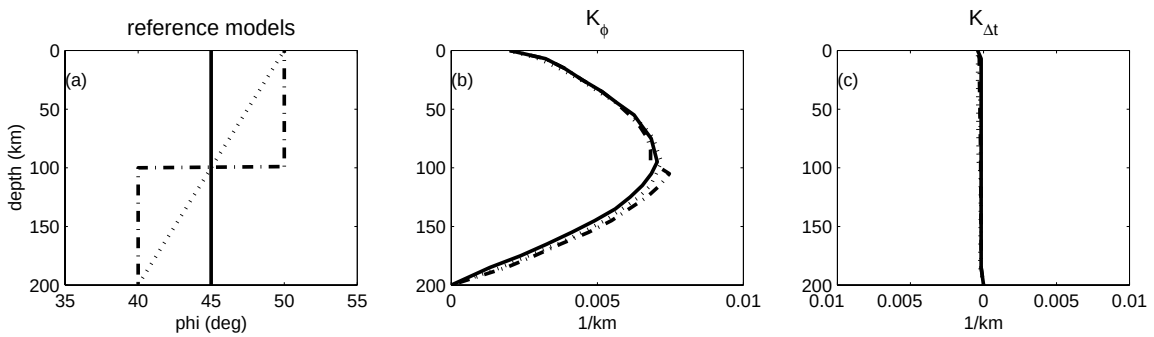


Figure 7. A comparison of Fréchet kernels for three starting models. Panel (a) is a plot of $\phi(z)$ for the homogeneous-layer (solid line), constant-gradient (dotted line), and two-layer (dashed line) models. Panels (b) and (c) show the corresponding kernels for these models, $G_\phi(z)$ and $G_t(z)$, respectively. The kernels were calculated by a numerical perturbation scheme for ω_0 = 0.14 Hz, $\sigma = \omega_0 / 6$ Hz, which are similar to the values used in the processing of teleseismic shear waves. The models have the same average azimuth, $\bar{\phi}$ = 45°. Eqn. (10) shows that $G_t(z) = 0$ for a homogeneous layer with this initial azimuth. The agreement illustrates the weak dependence of the kernels on the starting model.

We also compared the perturbations calculated from the Fréchet kernels using eqns. (8) and (9) with the results of a direct numerical calculation that minimized the tangential-component energy on back-projected synthetic seismograms. Figure 8 shows several examples of these comparisons as a function of the average polarization direction $\bar{\phi}$ for an initial pulse with a center frequency at 60 mHz and corners at 45 mHz and 75 mHz. Figure 9 plot the results for $\bar{\phi}$ = 45° for increasing values of the center frequency ω_0. We note that care must be taken in the numerical calculations when evaluating the apparent splitting parameters near the azimuthal nodes at $\phi_0 = n\pi/2$, because the energy surfaces can be very flat in the $\Delta t'$ direction, and the location of the minimum is susceptible to numerical inaccuracies that can cause a $\pi/2$ ambiguity.

When the heterogeneity is small ($\Delta\phi$ = 10°) in the linear gradient models, the kernels do a good job of predicting $\tilde{\phi}$ and $\Delta\tilde{t}$ for all backazimuths (Fig. 8a). For heterogeneity with a total rotation angle as large as 60° (Fig. 8b), the kernels typically overestimate the apparent splitting time by about 0.3 s, and, near the nodes, $\tilde{\phi}$ can be off by as much as 20°. This failure of the kernels with greater heterogeneity reflects a breakdown in the small-angle approximations (e.g., eqn. B5). Inclusion of the back-scattering terms in the calculation of synthetic

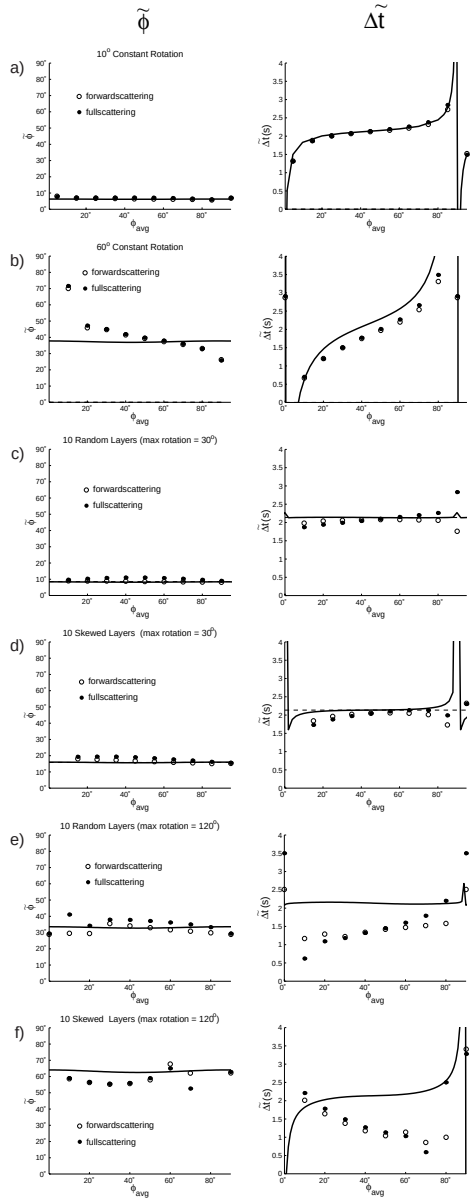


Figure 8. Comparison of numerical and analytical results. Left panels show $\tilde{\phi}$ and right panels show $\Delta\tilde{t}$, plotted as a function of backazimuth. Solid lines are analytical predictions from the kernels. Solid circles are the numerical results including back scattering and open circles include just forward scattering. a) weakly heterogeneous, ($\Delta\phi$ = 10°) linear rotation model. b) strongly heterogeneous ($\Delta\phi$ = 60°) linear rotation model. c) weakly heterogeneous 10 random layers model (maximum $\Delta\phi$ = 30°) in which the average fast-axis direction in the top and bottom halves are similar. d) weakly heterogeneous 10 random layers model (maximum $\Delta\phi$ = 30°) in which the layer orientations in the top half are different than that in the bottom half. e) strongly heterogeneous 10 random layers model (maximum $\Delta\phi$ = 120°) in which the average fast-axis direction in the top and bottom halves are similar. f) strongly heterogeneous 10 random layers model (maximum $\Delta\phi$ = 120°) in which the layer orientations in the top half are different than that in the bottom half.

seismograms (solid dots in Figure 8) does not produce significantly different results from those obtained using just the forward-scattering terms.

In a second set of comparisons, we use what is perhaps a more geologically relevant model comprising 10 layers, each 20 km thick, of differing orientations constrained such that the fast-axis directions varies between 0° and 30°. When the layer orientations ϕ_i are distributed randomly, such that the average orientation in the top half of the model was similar to the bottom half, the agreement between the numerical results and the predictions of the analytical kernels is usually very good (Fig. 8c). In addition, both the exact values of the apparent splitting parameters and perturbation-theory predictions show very little dependence on the polarization angle, with the variations in the apparent splitting time associated with nodal singularities compressed into a narrow range of azimuths.

When the layer orientations are skewed, however, such that the average orientation in the top half of the model differed significantly from the bottom half (Fig. 8d), the $\pi/2$ periodicity in $\Delta\tilde{t}$ associated with the nodal singularities becomes more pronounced. This differences in the variation of $\Delta\tilde{t}$ with initial azimuth results from the fact that $G_t(z)$ is approximately a linear function of depth that averages to zero, as seen from its low-frequency form (18). That is, the perturbation to the apparent splitting time will be small and the $\pi/2$ periodicity will be suppressed when the first moment $\int_0^d \phi(z) z\, dz$ is small. For a specified level of heterogeneity, the constant-gradient case has the largest first moment of any model, which is why the initial-azimuth dependence in Figs. 8a and 8b is so pronounced.

These results can be used to qualify Silver and Savage's [1994] argument that a $\pi/2$ periodicity in initial azimuth should be diagnostic of vertical heterogeneity. This periodicity will be relatively weak for heterogeneous structures where the azimuth of the anisotropy does not vary systematically with depth.

Increasing the heterogeneity in the 10 random layers so that the fast-axis direction ranges over 120°, we find azimuthal discrepancies of up to ±5° and splitting-time discrepancies exceeding 1 s (Figs. 8e and 8f). At this level of

heterogeneity, back-scattering effects, given by the differences between the open and solid circles, begin to become important.

A comparison of the analytic and numerical results for different frequencies at a single backazimuth, shows that the kernels do a nice job of predicting $\tilde{\phi}$ and $\Delta\tilde{t}$ up to 0.5 Hz when the heterogeneity is weak (Figs. 9a and 9c) but that when the heterogeneity gets stronger, the kernels break down more quickly at higher frequencies (Figs. 9b and 9d) than at lower frequencies. In these examples, Δt = 2.14 s. At frequencies approaching $1/\Delta t$ (~0.46 Hz) the kernels predict highly oscillatory behavior in $\tilde{\phi}$, which also corresponds to the frequency at which the kernels become unbounded in the single-frequency limit.

Strong-Scattering Regime

The numerical experiments demonstrate that the first-order perturbation theory expressed in eqn. (8) and (9) provides an accurate description of the apparent splitting parameters in situations where the magnitude of the vertical heterogeneity is small. As this magnitude increases, the perturbation theory fails because the small-angle approximations employed in obtaining the scattering matrix (B5) and the linearized minimization condition (B11) become inaccurate, owing to the accumulating effects of multiple forward-scattering. The combination of these strong-scattering effects causes the behavior of the apparent splitting parameters to deviate from the weak-scattering results.

Aspects of this behavior were noted in the previous discussion of Fig. 3, which display the numerical results for a constant-gradient model. In these calculations, Δt = 2 s, $\bar{\phi}$ = 45°, and all forward- and back-scattering terms were retained. For a 45° average polarization, the general form of the kernel (11) shows that the splitting-time perturbation should be zero to first order; i.e., the apparent splitting time $\Delta\tilde{t}$ should equal the total splitting strength Δt. Strong scattering acts to reduce $\Delta\tilde{t}$ below this theoretical limit, so that the ratio $(\Delta t - \Delta\tilde{t})/\Delta t$, to the extent it can be accurately estimated, measures the higher-order effects. For the 30° rotation in Fig. 2c, this reduction is only about 5%, consistent with the weak-scattering approximations. The 120° rotation in Fig. 2d

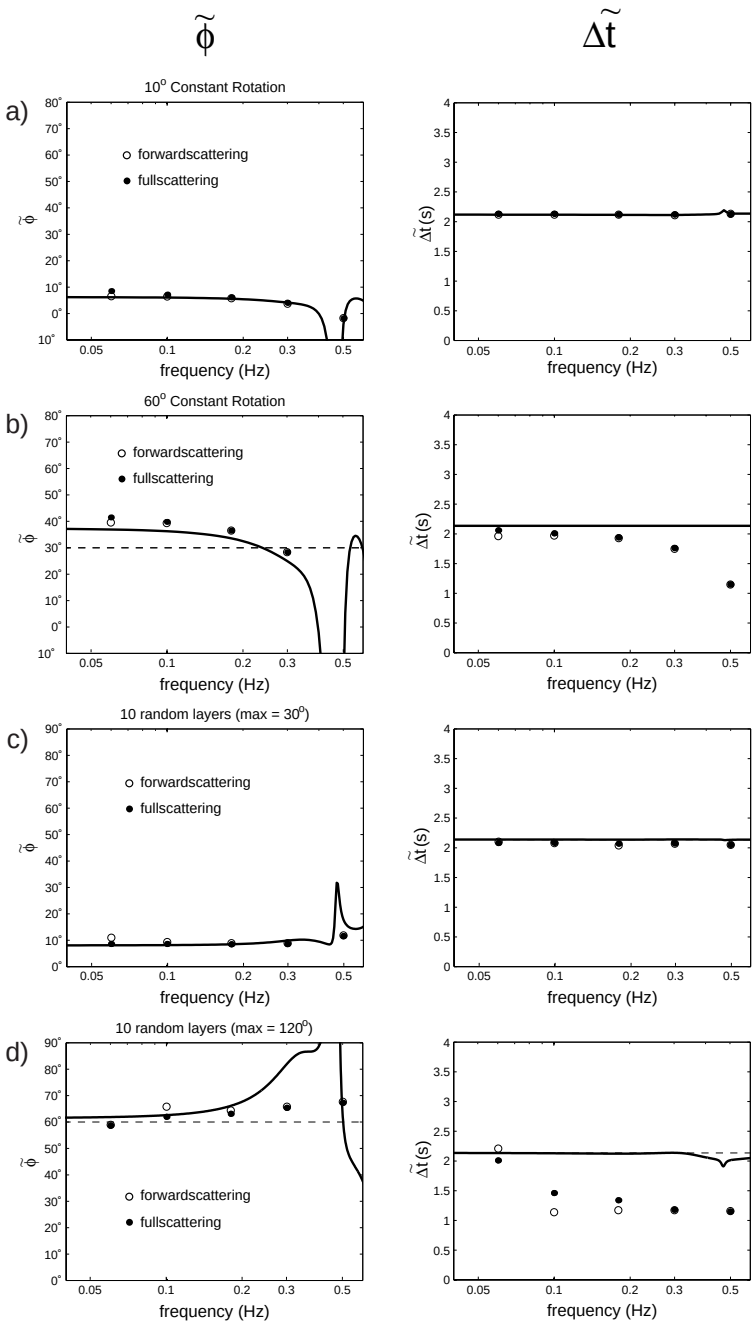


Figure 9. Comparison of numerical and analytical results. Left panels show $\tilde{\phi}$ and right panels show $\Delta\tilde{t}$, plotted as a function of frequency. Solid lines are analytical predictions from the kernels. Solid circles are the numerical results with both forward and back scattering and open circles include just forward scattering. a) weakly heterogeneous, ($\Delta\phi$ = 10°) linear rotation model. b) strongly heterogeneous ($\Delta\phi$ = 60°) linear rotation model. c) weakly heterogeneous 10 random layers model (maximum $\Delta\phi$ = 30°). d) strongly heterogeneous 10 random layers model (maximum $\Delta\phi$ = 120°).

gives a much more substantial effect (~35%), indicating that these approximations are not accurate for heterogeneity of this magnitude. For the 1000° rotation, the scattering is sufficiently large that the tangential-component arrivals are incoherent, so that the long-period amplitude is nearly zero, and the energy diagram looks nodal. In this case, there is no well-defined energy minimum, and it is difficult to measure the apparent splitting parameters.

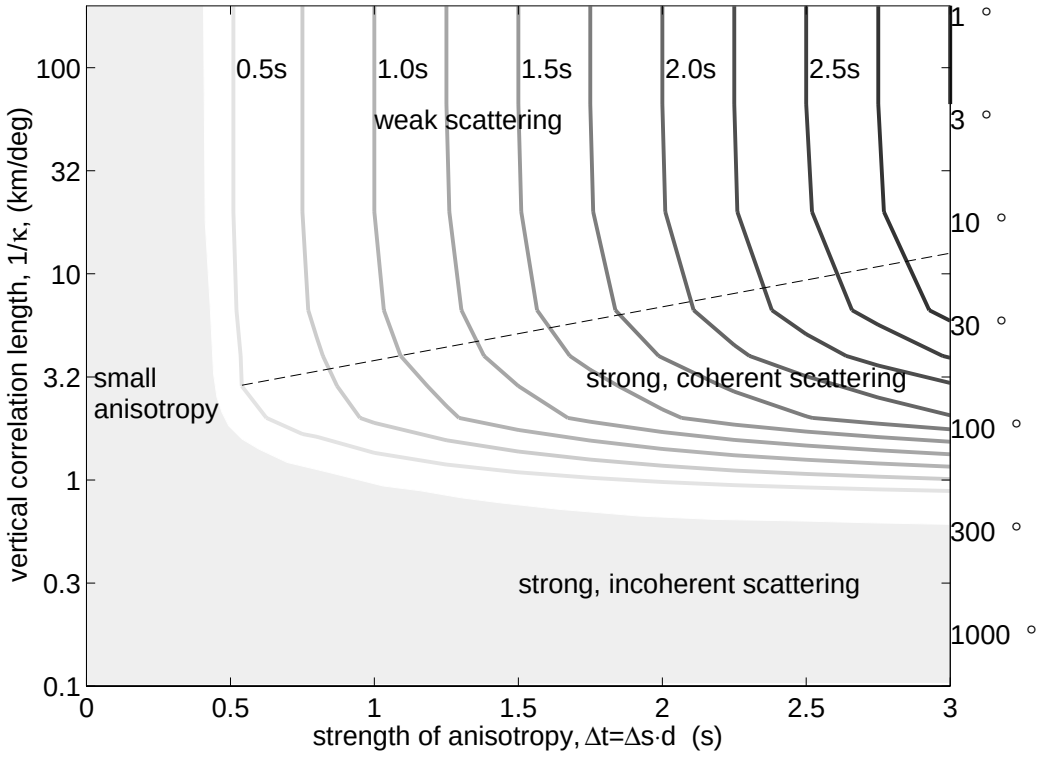


Figure 10. Scattering Diagram Contour of split time measurements $\Delta\tilde{t}$ for smoothly varying models. Horizontal axis shows the range of anisotropy in the models ($\Delta t'$) and vertical axis shows the range of vertical heterogeneity in the models (κ is the rotation rate). Splitting measurements cannot be made when $\Delta t' < 0.5$ s or $1/\kappa < 1$ km/degree.

In Fig. 10, we extend these calculations to constant gradient models with $\bar{\phi}$ = 45° and a range of splitting and heterogeneity strengths. The ordinate is taken to be $1/\kappa$, a quantity proportional to the inverse of the heterogeneity gradient,

which defines a vertical correlation length. The contours of apparent splitting time on this plot can be used to delineate the three scattering regimes. The region with nearly vertical contours at large values of the correlation length corresponds to weak scattering, where $\Delta\tilde{t} \approx \Delta t$ and the perturbation theory of Sect. 3 is valid. The deviation of these contours towards the horizontal defines a region where the scattering is too strong for perturbation theory to apply but not so strong as to prohibit the estimation of the apparent splitting parameters. The boundary between these two scattering regimes, indicated by the dashed line in Fig. 10, is given by a correlation length that increases exponentially with the anisotropy strength. As the correlation length of the vertical heterogeneity decreases at constant Δt, the apparent splitting time decreases, at first slowly then rapidly. Below some critical value of the correlation length (~50 km in this example, corresponding to κ = 1 deg/km), the scattering becomes so strong that $\Delta\tilde{t}$ cannot be defined.

As this diagram makes clear, it is not possible to distinguish on the basis of low-frequency splitting observations the difference between highly heterogeneous anisotropy (Δt and κ large) and weak anisotropy (Δt small). Analysis of horizontally propagating surfaces waves (e.g., Jordan and Gaherty, 1996) is one way to distinguish between these two cases.

Discussion

For the case of weak scattering, differences in the sensitivity kernels as a function of frequency can, in principle, be exploited to invert frequency-dependent shear-wave splitting measurements for a picture of anisotropic variation with depth. To do so, one applies standard splitting analysis to broadband recordings in order to obtain an apparent fast direction to be used as a starting estimate for $\bar{\phi}$. Frequency-dependent apparent splitting parameters are then extracted by applying the back-projection procedure to narrow-band filtered seismograms, and their kernels constructed from (10) and (11). Adherence to the weak-scattering regime can be checked by confirming that minimal signal remains on the tangential-component seismogram after back-projection via the apparent splitting parameters (Fig. 4). The frequency-

dependent splitting parameters can then be inverted for $\phi(z)$. (In all of our calculations, the difference between the speeds of the two eigenwaves remained constant throughout the model. In the real world, this parameter, like the anisotropy orientation, probably varies with depth. It is a simple matter, however, to extend the theory to depth-dependent wave speeds.)

The applicability of this procedure is likely to be limited by difficulties in extracting apparent splitting parameters at higher frequencies. Above 0.1 Hz, observed shear waveforms become increasingly complex due to microseisms, crustal scattering, and other sources of "noise". In addition, split shear waves have distinct "holes" in their amplitude spectra at frequencies with integer multiples of $1/\Delta t$ (Silver and Chan, 1991), which complicate analysis at higher frequency. As a result, most shear-wave splitting analyses in the literature utilize center frequencies that fall within a relatively narrow frequency band of approximately 0.05-0.2 Hz (e.g. Silver, 1996; Fouch and Fischer, 1996; Wolfe and Solomon, 1998). Our numerical experiments in the weak-scattering regime indicate that across this bandwidth, variations in $\Delta\tilde{t}$ and $\tilde{\phi}$ are generally less than 0.1 s and 5°, respectively (Figs. 8a, 9a). These variations are smaller than the typical error estimates in observational studies, and thus they cannot resolve changes in anisotropy with depth.

In the case of strong scattering, the approximations made in deriving the kernels are no longer valid, so that the kernels cannot be used to solve the inverse problem. We can, however, utilize the numerical results in the interpretation of splitting observations. Marson-Pidgeon and Savage (1997) report frequency-dependent shear wave splitting results from New Zealand (between 50 and 200 mHz) that are consistent with our numerical results for the strong coherent scattering regime (Figs. 9b, 9d), implying significant vertical heterogeneity with depth. In addition, splitting results from cratons in South Africa (Gao et al, 1998), Australia (Clitheroe and van der Hilst, 1998; Özalaybey and Chen, 1999), India (Chen and Özalaybey, 1998), and Tanzania (Owens et al., 1999) all find splitting times that are smaller (generally < 0.6 s) than many other continental environments (e.g. Silver, 1996). Such observations are typically interpreted as evidence for little or no anisotropy. Our calculations provide an

alternative explanation for these null results in terms of strong incoherent scattering in an upper mantle that is anisotropic, but has a high degree of vertical heterogeneity. Other data, such as horizontally propagating surface waves, are necessary to distinguish between these two possibilities.

In at least two cratonic regions where null or near-null results are reported (Australia and Southern Africa), analyses using surface waves find that the upper mantle is anisotropic between the Moho and 200-250 km depth, with $\Delta v_s/v_s$ of approximately 3-4% (Gaherty and Jordan, 1995; Saltzer et al., 1998). If perfectly aligned, such anisotropy would produce over 2 s of splitting. We interpret the apparent discrepancy between the splitting and surface-wave results as evidence for strongly heterogeneous anisotropy in these regions. In particular, both sets of observations can be explained by a model in which the local anisotropy axis remains, on average, close to horizontal, but varies in azimuth as a function of position, with a vertical correlation length on the order of 50 km (Jordan et al., 1999).

Conclusions

In both weakly and strongly heterogeneous media we find that shear-wave splitting measurements made in the low-frequency bands typically used to make observations are more sensitive to the upper portions of the model than to the lower portions. Consequently, if the orientation of the anisotropy is heterogeneous, the measured splitting direction will be more reflective of the fast-axis direction near the top of the upper mantle and the measured split time will vary as a function of backazimuth. This effect may explain why global shear-wave splitting measurements tend to correspond to the local tectonic fabric in crust beneath the observing station (Silver, 1996).

We have derived analytic expressions for sensitivity kernels that relate a perturbation in the measured splitting parameters to a perturbation in the anisotropy of the model. For weakly heterogeneous media, frequency-dependent shear-wave splitting measurements can be inverted using these kernels to determine how the anisotropy varies as a function of depth.

Variations in the strength of the anisotropy as a function of depth can be accounted for with a weighting function. Practically speaking, however, a sufficient number of measurements may not be available at high enough frequencies or with enough precision (errors are typically ±10° or more) for such an inversion to be feasible.

Strongly heterogeneous media (i.e., where the fast axis direction varies anywhere between 0° and 180°) have the additional property that they cause strong scattering. This scattering will cause the tangential component seismograms to have very little energy and a null-like energy measurement will be made despite the fact that the medium may be highly anisotropic. Our numerical experiments show that the greatest diagnostic of strong vertical heterogeneity is the drop-off in $\Delta\tilde{t}$ relative to Δt.

Acknowledgements

Many thanks to Martha Savage, Steve Ward and an anonymous reviewer for useful comments that improved the paper. Many of the figures were generated using GMT software freely distributed by Wessel & Smith (1991). This research was funded by NSF Grant EAR-9526702.

Chapter 3

The Spatial Distribution of Garnets and Pyroxenes in Mantle Peridotites: Pressure - Temperature History of Peridotites from the Kaapvaal Craton

To be published by Oxford University Press in the *Journal of Petrology* by Rebecca Saltzer, Neel Chatterjee and Tim Grove, December 2001.

Abstract

We present a new method for textural analysis of mineral associations that uses digital backscattered electron and x-ray images obtained with the electron microprobe to determine the spatial properties of minerals on a two dimensional surface of the rock at different scale lengths. We determine modal amounts and average grain sizes of each mineral in the thin section without resorting to ellipsoidal approximations of grain boundaries, and investigate the spatial relationship of mineral pairs. The method is used to characterize nine mantle xenoliths erupted from kimberlite pipes in South Africa and to test whether the pyroxenes are spatially correlated with the garnets. The spatial association of these minerals is used to develop a model for the evolutionary history of the Kaapvaal peridotites. The observed distributions can be explained by a two-stage model. In stage 1, harzburgitic residues are produced by large extents of partial melting at shallow depths (~60-90 km) and high temperatures (~1300-1400° C). The melting process leading to this depletion occurs in the garnet stability field where garnet, clinopyroxene and olivine are consumed and orthopyroxene and liquid are produced. The Kaapvaal sample suite shows modal and compositional variations consistent with a progressive melt depletion event. In stage 2, the residuum is dragged down to greater depths by mantle corner flow adjacent to a subducted slab. The most depleted harzburgites descend to 140 to 160 km depth and are cooled. The least depleted harzburgites end up at shallower depths. The resulting stratigraphy is the opposite of what would be expected for a preserved mantle melt column and is consistent with inversion of the melt column as it was dragged around the wedge corner and cooled by the subducted slab. The cooling process causes clinopyroxene and garnet to exsolve from the orthopyroxene. Therefore, the depleted cratonic peridotites of the Kaapvaal preserve a temperature-pressure path consistent with an origin in an Archaean subduction zone.

Introduction

Mantle xenoliths erupted from kimberlite pipes in South Africa may provide clues to the formation and evolution of the depleted mantle that makes up the Kaapvaal craton. One hypothesis that has been proposed is that the peridotite originally resides deep in the mantle (at depths of perhaps 300 or 400 km) and is transported upwards into the cratonic lithosphere (~180 km depth), where re-equilibration occurs and clinopyroxene (cpx) unmixes from the garnet phase

[*Haggerty & Sautter*, 1990]. An alternative idea is that the garnet lherzolites were originally high-temperature harzburgites that originated at depths between 100 and 250 km and that both the cpx and garnet subsequently exsolved from the Al-rich orthopyroxene (opx) when the rocks cooled and re-equilibrated [*Cox et al.*, 1987]. If exsolution of minerals has indeed occurred, then a spatial relationship between grains of minerals that were formerly dissolved in each other should be apparent in the present-day xenoliths. In particular, the first model implies a spatial relationship between cpx and garnet for rocks of pyroxenite or eclogitic compositions, whereas the second implies cpx, opx and garnet should be correlated.

The question of whether two minerals are spatially related arises frequently in petrologic studies. Despite the availability of digital images of thin sections, textural studies described in the literature are typically crystal-oriented rather than pixel-oriented [*Cashman & Ferry*, 1988; *Kretz*, 1993; *Jerram et al.*, 1996]. These analyses generally take a "nearest neighbor approach", measuring the distance between the center of a grain and the center of its nearest neighbor, to determine whether the grains are randomly distributed, clustered, or ordered [*Kretz*, 1969, 1993; *Carlson et al.*, 1995; *Jerram et al.*, 1996; *Miyake*, 1998]. The most sophisticated of these involve fitting ellipsoids to a digital image of the grains, while others simply project the thin section on a screen and trace the crystals by hand. Recently, studies of digital images have used Markov-chain analysis to look for the same sort of clustering or anti-clustering of minerals [*Kruse & Stünitz*, 1999]. A prior spatial analysis of Kaapvaal peridotites [*Cox et al.*, 1987] specifically tested the hypothesis that two different minerals were spatially associated and found a strong correlation between the presence of garnet and cpx with opx, but their method required labor-intensive manual outlining, identification and counting of crystals.

We describe a new technique for textural analysis in which the digital image of a thin section is statistically analyzed over a range of spatial scales. By analyzing the pixel data, we determine the modal amounts and average crystal size for each mineral type and examine whether any of the minerals are more closely associated spatially than would be expected if they were randomly

distributed. This method may be applicable to different textural studies in which average crystal sizes need to be determined or the relationship between mineral types, and in particular their spatial relation due to metamorphism or metasomatism, needs to be assessed. We apply this statistical method to nine mantle xenoliths erupted from kimberlite pipes in South Africa to determine which, if any, of the constituent minerals are correlated. In this way, we test the two hypotheses regarding the history of South African mantle xenoliths and, by implication, the structure and evolution of the Kaapvaal craton.

Sample Selection

We selected three garnet-peridotite xenoliths from the Bulfontein kimberlite pipes (Kimberley, South Africa), four from the Jagersfontein pipes (130 km SSE of Kimberley), one from the Premier Pipes (25 Km NE of Pretoria, South Africa) and one from the Letseng pipes in northern Lesotho (Fig. 1).

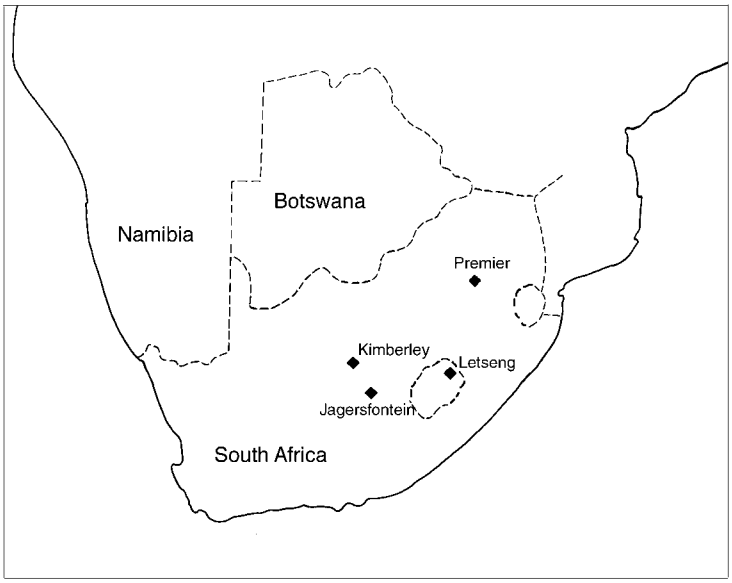


Figure 1. Map of southern Africa. Nine, low-temperature, garnet-peridotite xenoliths were collected from the four kimberlite pipes shown.

Several criteria were applied in selecting the sample suite. All samples contained dominantly olivine and opx with the coarse texture that characterizes low-temperature peridotites [*Boyd*, 1987]. Peridotite nodules (~25 × 25 × 10 cm) were

cut into multiple slabs (3 to 5 slabs each ~1cm thick) and samples were chosen from nodules that had a uniform distribution of garnet and cpx in all slabs. Samples that contained clots of garnet and cpx were avoided, as were samples that contained veins of phlogopite or samples that were serpentinized. We also avoided samples that appeared to have undergone metasomatic modification. A suite of representative low-temperature peridotites that spanned the range of modal olivine percentages observed by *Boyd* [1989] were supplied by F.R. Boyd.

Mineral compositions and modal analysis

From each slab we prepared two by three inch polished thin sections for chemical and modal analysis. We obtained a back-scattered electron image (brightness corresponding to atomic number) as well as x-ray concentration maps of Si, Al, Ca, and Mg) of each thin section using the four wavelength dispersive spectrometer-JEOL JXA-733 Superprobe electron microprobe at MIT. We used an accelerating voltage of 15 kV, a beam current of 100 nA and a dwell time of 5 ms per spot. By comparing the various images, we were able to identify each pixel in the image as either olivine, orthopyroxene, clinopyroxene or garnet. Eight of the samples we mapped and analyzed are shown at the same scale in Fig. 2. The image resolution ranged between 62 microns/pixel to 200 microns/pixel and the areas imaged were approximately 3 cm × 5 cm. Depending on the image resolution and the area mapped, the image collection time varied between 12 to 16 hours per sample.

Chemical Compositions

Chemical compositions of the minerals were obtained for five of the samples with the same electron microprobe at MIT using wavelength dispersive spectrometry. Phases were analyzed at an accelerating voltage of 15kV, a beam current of 10nA, a beam diameter of ~1 micron and typical counting times between 20 and 40 seconds per element. Data were reduced with the CITZAF program [*Armstrong,* 1995] using the atomic number correction of Duncumb and Reed, Heinrich's tabulation of mass absorption coefficients and the fluorescence correction of Reed. A minimum of 10 measurements were made for each

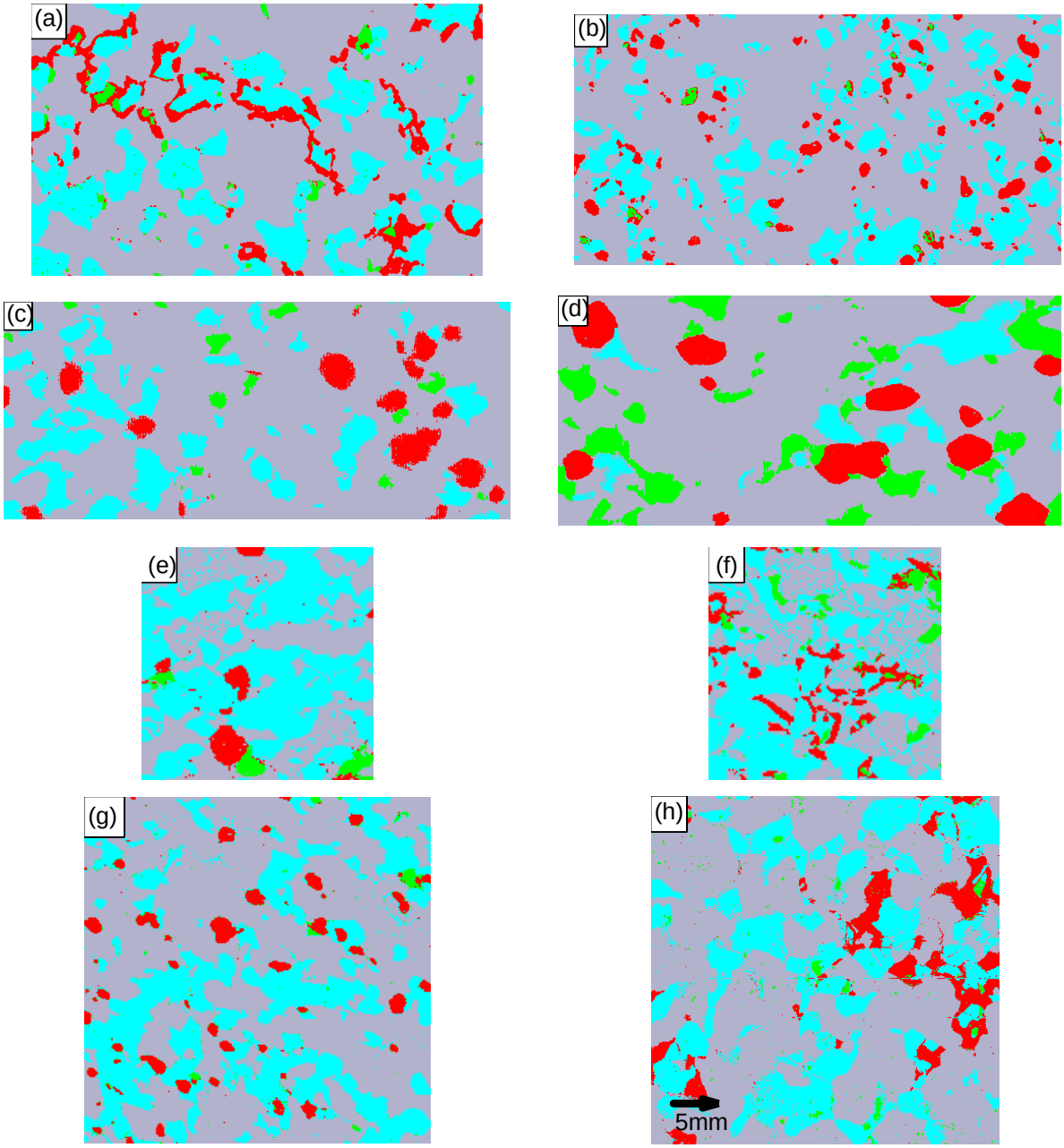


Figure 2. Eight of the nine samples we mapped and analyzed. All are shown at the same scale. Grey=olivine, blue=opx, red=garnet, and green=cpx. (a) UX497 (b) K1 (c) JAG93_8 (d) JAG90_72 (e) FRB4265 (f) FRB1350 (g) FRB347 (h) FRB1009.

mineral on every thin section. The average and standard deviation of those measurements are summarized in Table 2. Mineral compositions for the other

four samples were provided by F.R. Boyd of the Geophysical laboratory at the Carnegie Institution of Washington.

Classification as Low-Temperature Peridotites

Garnet-peridotites are broadly classified as either low- or high-temperature in origin, based on estimated equilibration temperatures [*Boyd,* 1987, 1989]. Xenoliths classified as low-temperature have been interpreted as representing lithospheric samples from the upper mantle, originating from depths of less than 150 km whereas the high-temperature xenoliths have been interpreted as representing asthenospheric samples from depths of 175 km and greater [*Nixon & Boyd,* 1973; *Boyd,* 1987, 1989; *Harte,* 1983; *Finnerty & Boyd,* 1987; *Nixon,* 1987; *Herzberg,* 1993; *Brey & Köhler,* 1990]. The high-temperature xenoliths tend to have a sheared or deformed texture [*Nixon & Boyd,* 1973; *Bouillier & Nicolas,* 1975; *Harte,* 1977, 1983; *Herzberg,* 1993; *Boyd,* 1987] making them poor candidates for a spatial study such as ours, while the low-temperature xenoliths have presumably retained their original spatial configuration making them ideal candidates for a study of unusual mineral clustering.

Table 1 lists the modal proportions and olivine Mg numbers (atomic MgO/FeO+MgO) for the samples we analyzed with low- and high-temperature averages for comparison and Table 2 gives the mineral compositions. All nine xenoliths have olivine Mg numbers greater than 91.5, a coarse, granular texture with no preferred dimensional orientation or other obvious signs of shearing, and estimated equilibration temperatures of between 800° and 1100° C and pressures of 30 to 55 kilobars (Table 1, using the thermometer of *Brey & Köhler* [1990]) consistent with a classification of low-temperature peridotites [*Nixon & Boyd,* 1973; *Harte,* 1975, 1983; *Finnerty & Boyd,* 1987; *Boyd,* 1987, 1989]. In addition, the bulk rock, major element geochemistry of the samples is typical of low-temperature peridotites from South Africa, which tend to be depleted in FeO, TiO_2, CaO and Al_2O_3 relative to higher-temperature xenoliths [*Nixon & Boyd*, 1973; *Harte,* 1977, 1983; *Nixon,* 1987; *Boyd,* 1989; *Herzberg,* 1993].

Table 1. Modal analyses and Olivine Mg numbers of low-T garnet lherzolite xenoliths from South Africa

Pipe	Rock #	olivine	opx	garnet	cpx	Mg #	T °c	P kb
Jagersfontein	JAG90_72	71%	7%	9%	12%	93.2	1,055°	51
Jagersfontein	JAG93_8	76%	16%	6%	2%	92.6	1,131°	53
Jagersfontein	UX497	66%	25%	8%	2%	92.1	800°	36
Kimberley	K1	73%	21%	5%	1%	93.0	977°	45
Kimberley	K12	79%	11%	9%	2%	92.5		
Kimberley	FRB347	64%	27%	7%	2%	93.1	1,092°	49
Jagersfontein	FRB1009	61%	29%	7%	3%	91.8	865°	39
Premiere	FRB 1350	56%	33%	5%	5%	91.7	770°	27
Lesotho	FRB4265	51%	42%	5%	1%	92.5	1,110°	52
low T avg [1]		61%	31%	6%	2%	92.6		
high T avg [2]		77%	16%	6%	3%	91.1		

Modal results are expressed as volume percentages. Olivine Mg number is atomic MgO/(MgO+FeO). Kelyphite minerals have been counted as garnet and serpentinized minerals have been counted as olivine.

(1) [*Boyd,* 1989]

(2) [*Boyd & Mertzman,* 1987]

(3) Final equilibration P and T estimates using the thermobarometer of *Brey and Köhler* [1990]

The oxide compositions of our samples are also consistent with low-temperature xenoliths. Fig. 3 shows the weight percent oxides for the individual minerals of the xenoliths normalized to the average for the corresponding minerals of thirteen low-temperature garnet lherzolites from southern Africa analyzed in a previous study by Cox *et al.,* [1987]. One of the rocks (JAG90_72) has 2 to 3 times as much TiO_2 as the average 13 low-temperature rocks do, however it is still within the range of what is found for low-temperature xenoliths [*Boyd,* 1987].

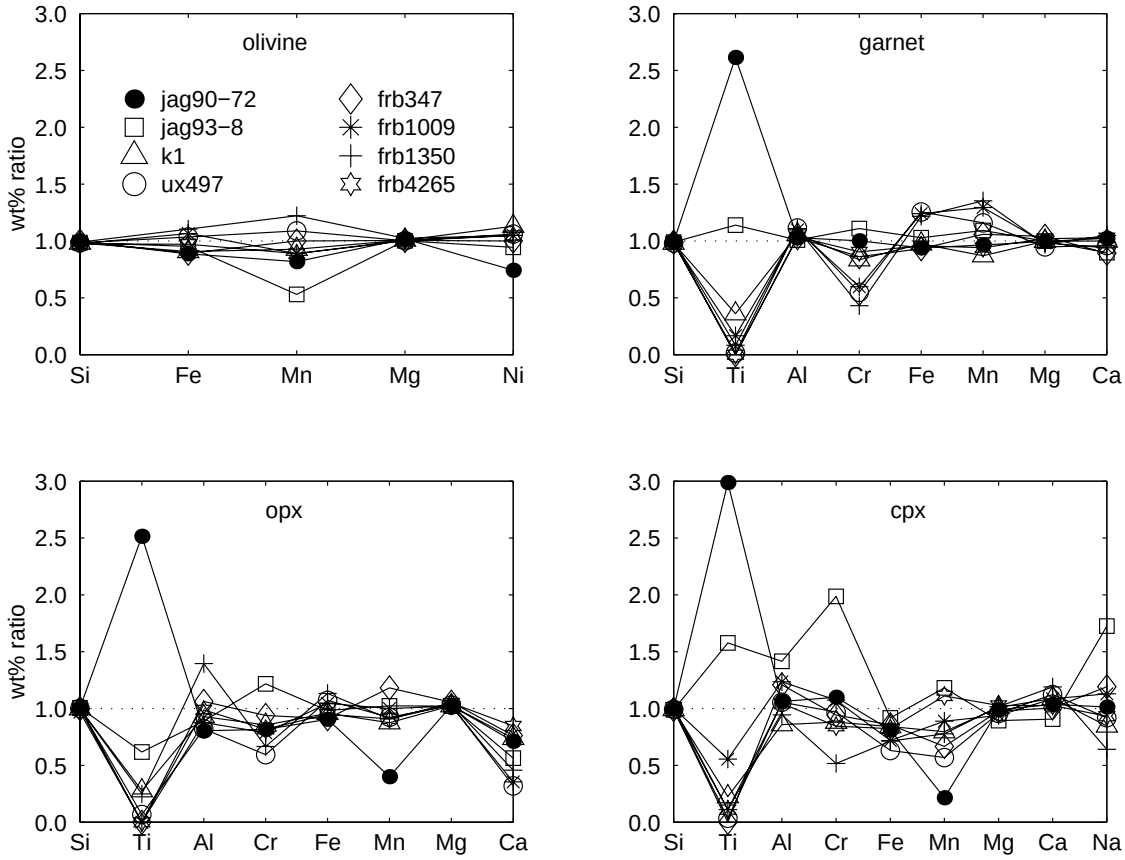


Figure 3. Spider plot of weight percent of oxides in individual minerals normalized by the average of 13 low-temperature garnet lherzolites from southern Africa analyzed by Cox *et al.* (1987).

The only way in which some of the rocks we analyzed differ from the low-temperature classification is in their modal concentrations of opx and cpx. Three of the rocks (JAG90_72, JAG93_8, and K12) contain unusually low amounts of opx and one (JAG90_72) has very high amounts of cpx, much more akin to high-temperature xenoliths than low-temperature ones. Except for this variability in modal amounts, these rocks are, otherwise, very typical of low-temperature, garnet peridotites from kimberlite pipes in South Africa. Moreover, they span almost the entire range of modal olivine and olivine Mg numbers that is found in the Kaapvaal craton xenoliths (Fig. 4) so they should provide a representative and potentially contrasting set of low-temperature samples.

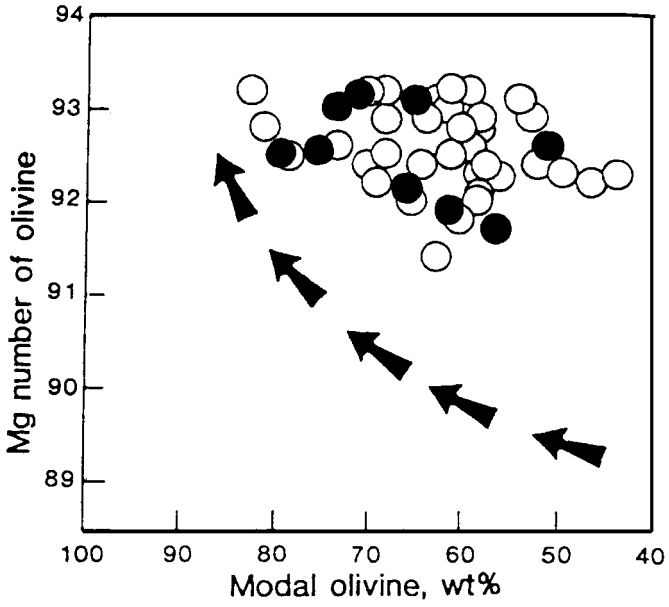


Figure 4. Plot of modal olivine versus Mg number of olivine from *Boyd* [1989]. The figure has been modified to show the eight samples we analyzed (solid black circles) in addition to those previously analyzed (open circles) by *Boyd* [1987]. Arrows show the oceanic trend inferred by *Boyd* [1989].

Textural Analysis Method

In order to test for a preferential spatial correlation between two minerals, we employ a box-counting method (described in the spatial analysis section below) on a series of different scales, to determine some of the statistical properties of the rock. We compare these results with those obtained from analysis of computer-simulated rocks that contain randomly-distributed crystals with the same modal proportions and average grain sizes as determined in our analysis of the real rock. In this way, we can test whether there is a statistically unusual spatial association between any of the mineral pairs in the real rocks.

Spatial Analysis

We divide the digital image of the thin section into squares. We choose the square size to have a total number of pixels $N_n = 2^{2n}$ where n is initially the largest possible integer allowed by the image size. Fig. 5 shows the back-scattered electron image and x-ray maps for an example thin section (FRB347) as

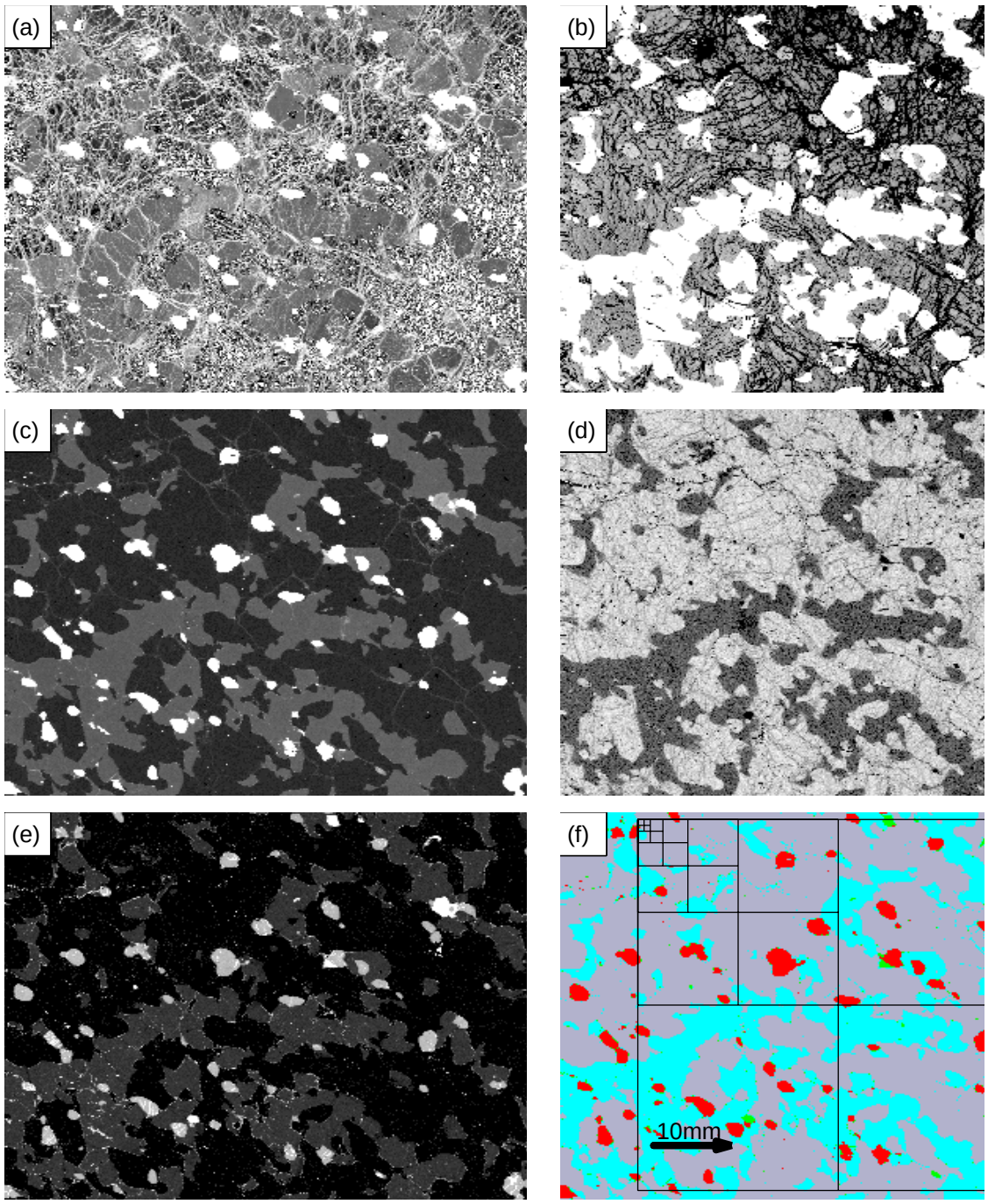


Figure 5. Example thin section. a) BSE image b) Si map c) Al map d) Fe map e) Ca x-ray maps for sample FRB347 and f) final results of mineral identification. Grey=olivine, blue=opx, red=garnet, and green=cpx. The sample has been subdivided into 4 squares that are 128 pixels by 128 pixels in size. The upper lefthand corner shows successively smaller boxes for $n = 6,\ldots 1$. When $n = 0$, each box contains a single pixel.

well as the final digital image (the final identification of each pixel as opx, cpx, garnet or olivine). In this case, we begin with $n = 7$ (corresponding to a square 128 pixels by 128 pixels in size with $N_7 = 16{,}384$) and 4 such squares. We then calculate the number of squares that contain a large fraction of each of the minerals. To make this determination, we define a correlation condition

$$\frac{n_x}{N_n} \geq \alpha \tag{1}$$

where n_x is the total number of pixels of mineral x in the box and α is the fraction of the box the crystal must occupy. In our analysis, we chose $\alpha = 0.75, 0.80$ and 0.875.

Fig. 6 shows a plot of these results on a ternary diagram for the case of a three mineral system; however, the garnet-opx-cpx-olivine system has four components, so in our analyses we actually plot the results on a tetrahedron (not shown) which is comprised of 4 ternary diagrams joined on each side. The distance from each corner is equivalent to $(1 - \frac{n_x}{N_n})$ and α plots as a line near each of the three corners defining a region we call a. The boxes that satisfy the condition defined in equation (1) are represented as points that lie within the corner regions a.

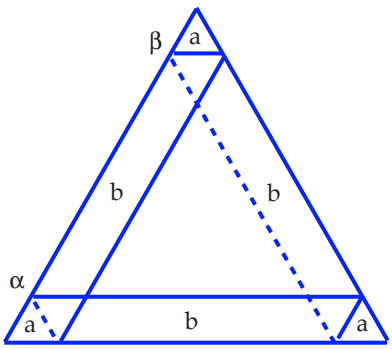


Figure 6. Triangular plot of counting results. The regions marked "a"contain the number of boxes that have a large fraction (greater than α) of the mineral identified in the corner. Each region marked "b"contains the number of boxes that have a large proportion of the two minerals identified at the two corners that bound the region. The unidentified region in the center contains the number of boxes with a small proportion of all three minerals.

We then count the number of boxes that contain a large proportion of any two minerals. These points plot in the sidebars (regions b). The correlation condition for this determination is

$$\frac{n_x}{N_n} \leq \beta \tag{2}$$

where n_x is the total number of pixels of the mineral in the corner opposite the sidebar being determined, and β is a chosen correlation threshold. In our analysis, we chose $\beta = (0.25, 0.20 \text{ and } 0.125)$ and found that all gave similar results. We repeat the counting procedure for successively smaller boxes $(n = 6,\ldots,1)$ and plot the results for each n on a separate ternary diagram. For each decrement of n, the box area is reduced by one quarter. This process yields a "mineral correlation spectrum" as a function of box size. When $n = 0$, each box contains a single pixel and all regions of the ternary diagram are empty except for the corners, from which the modal amounts of each mineral comprising the rock can be determined.

As n increases, the percentage of boxes plotted in each of the corners P_{nx} will gradually decrease. When the box size exceeds the average crystal size, P_{nx} will drop abruptly. A plot of P_{nx} as a function of box size (Fig. 7) will show a break in slope, which indicates that this critical size has been reached. For equant

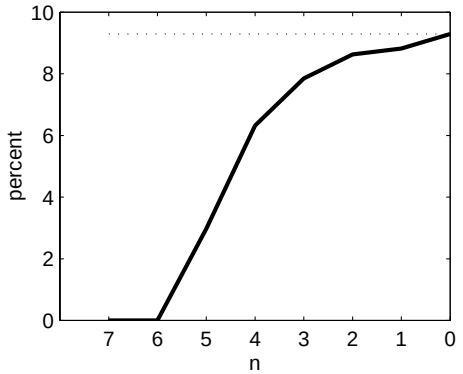


Figure 7. Percentage of boxes containing garnet as a function of n for sample JAG90_72. The break in slope at $n = 6$ indicates that the grid size has become larger than the average crystal size of $n = 5$. The percentage of boxes at n=0 indicates the modal percent of garnet in this rock (~9%).

minerals, this size will correspond to the average crystal size $\bar{n}$. In some cases, the break in slope is less evident and so we used the additional criterion $P_{nx} \geq 0.25M_x$, where M_x is the modal proportion of mineral x.

Given the above results, we can create a thin section of a synthetic rock containing the same modal proportions and average grain sizes as the real rock, but with randomly distributed crystals. In this way, the synthetic rocks match the real rocks for $n = 0$, and the average number appearing in each of the ternary diagram corners is the same as the real rock for all n. We then test whether significantly more points occur in the sidebars of the real rocks than in the randomly distributed synthetic rocks.

Synthetic Rock Calculations

In general, average crystal sizes differ between minerals in the rocks we examined. Therefore, we create the synthetic thin section in two stages. First, we lay out a grid of the same size as the largest average crystal size. Since the larger crystals will fill the boxes completely while that the smaller ones will not, we re-normalize the modal proportions to account for the various size differences. The re-normalized modal amount for each mineral is

$$R_x = \frac{M_x}{4^{(\bar{n}_x - 1)} \cdot \sum_{i=1}^{4} \frac{M_i}{4^{(\bar{n}_i - 1)}}} \tag{3}$$

where i denotes the minerals comprising the rock (e.g., olivine, opx, cpx, and garnet), x denotes the particular mineral being re-normalized, M_x is the modal proportion of mineral x and $\bar{n}_x$ is the average crystal size of mineral x. Using a random number generator, we fill the boxes according to the R_x proportions with the appropriate sizes n_x. This will produce a partially filled rock with the same modal proportions as the real rock. The second stage is to go back and fill the remaining blank spaces according to the modal proportions M_x of the rock. In this way, the final rock produced has appropriate modal proportions, representative crystal sizes, and the minerals are randomly distributed. Fig. 8

shows an example of a synthetic rock that was generated according to this scheme (both stages of its creation) and the real rock it represents.

We create 1000 synthetic thin sections for each of the real rocks imaged and then analyze each of the synthetic rocks in exactly the same manner as the real rock. Therefore, for each n, we get a distribution of synthetic results for the

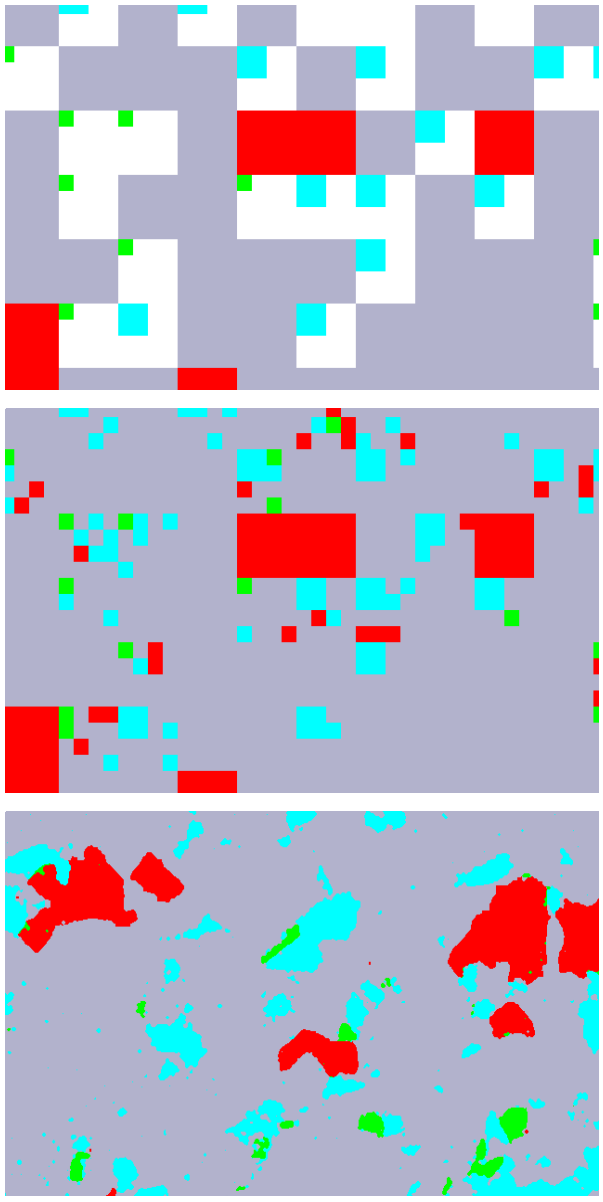


Figure 8. Partially filled synthetic (a), fully filled synthetic (b), and real rock (c). Grey=olivine, blue=opx, red=garnet, and green=cpx.

three corners of the ternary diagram and the three sidebars. Because we have created the synthetic rocks using the average crystal sizes, the results plotting in the corners of the ternary diagram for the real rock are well within the range of 95% of the results obtained for the synthetic rocks for all n. We focus on the sidebars to see if statistically significant differences arise. For example, when examining the garnet-pyroxene bar, we look at the ternary diagram for box size n corresponding to the $\bar{n}$ for garnet and pyroxene. If $\bar{n}$ differs for the two minerals, we take the average of the two and if they differ by just one order of magnitude, then we consider both $\bar{n}$. When the value in the sidebar of the real rock exceeds 95% of the expected range found in the synthetic rocks, we conclude that the two minerals are spatially associated.

Results and Interpretation

We found statistically significant relationships between cpx, opx and garnet in eight of the samples and between cpx and opx in the remaining sample, as summarized in Table 3. On the basis of these spatial relationships, it is possible that eight of the nine xenoliths we analyzed (JAG90_72, JAG93_8, UX497, K1, FRB347, FRB850, FRB1350, and FRB4265) were originally two-phase rocks consisting of primary olivine and opx (harzburgites) and that later in the sample's history the garnet and cpx unmixed from the opx. Assuming this history, we can reconstruct what the composition of these "original" orthopyroxenes would have been using the microprobe analyses of the weight percent oxides and relative densities of the garnet, cpx and opx.

There are several sources of error in making such a reconstruction including errors in the chemical analyses and errors in the estimates of modal abundances. We have calculated formal errors for the measurements of weight percent oxides and for the modal amounts imaged in the thin sections and found both to be negligible. The largest source of error is in the assumption that the modal amounts contained in the particular slice of rock we analyzed is representative of the bulk rock. In order to estimate the size of this error, we subdivide the thin sections into 10 strips and calculate the modal abundances of each mineral for every strip. We then use a Student's T distribution to determine the error

Table 3. Correlation of mineral pairs in garnet lherzolite xenoliths from South Africa

rock	cpx/opx	opx/garnet	cpx/garnet
JAG90_72	√	√	√
JAG93_8	√	√	√
UX497	√	√	√
K1	√	√	√
K12	√		
FRB347	√	√	√
FRB1009	√	√	√
FRB1350	√	√	√
FRB4265	√	√	√

bounds on the modal amounts with 95% confidence. Table 4 lists the chemical compositions of the reconstructed orthopyroxenes and the average estimated errors in these reconstructions given our modal error estimates.

The reconstructed Kaapvaal orthopyroxenes are chemically similar to the orthopyroxenes that are produced in both anhydrous and hydrous peridotite melting experiments. In Fig. 9a the compositions of pyroxenes from liquids saturated with a harzburgite (oliv + opx) residue are plotted along with the Kaapvaal reconstructed orthopyroxenes. The anhydrous melting experiments span a range of pressures from 2.2 GPa [*Parman*, 2001] to between 3 and 6 GPa [*Walter*, 1998]. Over this pressure range there is a systematic decrease in the Ca and Al proportions that is dominantly correlated with an increase in melt fraction. The effect of increasing pressure under anhydrous conditions is to increase the amount of melting required to exhaust cpx and garnet from the residual assemblage (Walter, 1998). Therefore, both higher temperatures and higher melt fractions are required in order to reach oliv + opx saturation at higher pressures. The effect of increasing H_2O is to increase the melt fraction and depress the temperature at which high melt fractions are stable with an oliv + opx residue [*Parman*, 2001]. Also shown in Fig. 9 are the orthopyroxenes in equilibrium with olivine and a Barberton komatiite liquid (24 wt. % MgO, 6.5 wt.

Table 4. Recombined opx compositions and olivine Mg numbers

Wt % Oxides	JAG90_72	JAG93_8	UX497	K1	frb347	frb1009	frb1350	frb4265	average +/-
SiO_2	54.0	53.8	53.7	54.5	56.7	53.5	55.5	56.4	1.7
TiO_2	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Al_2O_3	5.4	5.9	6.2	5.1	2.4	4.8	4.0	3.3	0.1
Cr_2O_3	1.7	1.9	0.9	1.1	0.6	0.9	0.5	0.8	0.0
FeO	4.1	5.0	5.7	4.8	4.3	5.2	5.2	4.6	0.2
MnO	0.1	0.2	0.2	0.1	0.2	0.2	0.1	0.1	0.0
MgO	27.7	30.5	31.0	32.6	35.6	31.2	32.3	34.8	1.1
CaO	6.6	2.8	2.5	1.8	1.1	2.6	3.4	1.4	0.1
Na2O	0.5	0.3	0.1	0.1	0.2	0.2	0.2	0.1	0.0
NiO	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.0
Mg #	93.2	92.6	92.1	93.0	93.1	91.8	91.7	92.5	

Olivine Mg numbers are atomic MgO/(MgO+FeO)

% H_2O) at 2.4 GPa. The low Al in the opx is a consequence of the higher temperature of oliv + opx saturation in this liquid compared to the hydrous peridotite melts.

Most of the Kaapvaal reconstructed orthopyroxenes parallel the trend defined by the experimental data. The compositions of reconstructed pyroxenes *from Cox et al.* [1987] also parallel the trend defined by our sample suite. The exception is the cpx + garnet rich sample JAG90_72 that plots at much higher Ca contents than the experimental data. Sample JAG90_72 will not be considered further in the discussion, and we assume that it formed by a different process than the one discussed below. When the remaining reconstructed orthopyroxenes are compared to the experimentally produced orthopyroxenes, the Ca and Al contents of the anhydrous opx compositions match the reconstructed compositions most closely. However, the differences between the hydrous and anhydrous Ca contents are small and our assumption that opx was the only pyroxene present initially in these samples may be artificially inflating the reconstructed Ca contents to higher values than were actually present in the

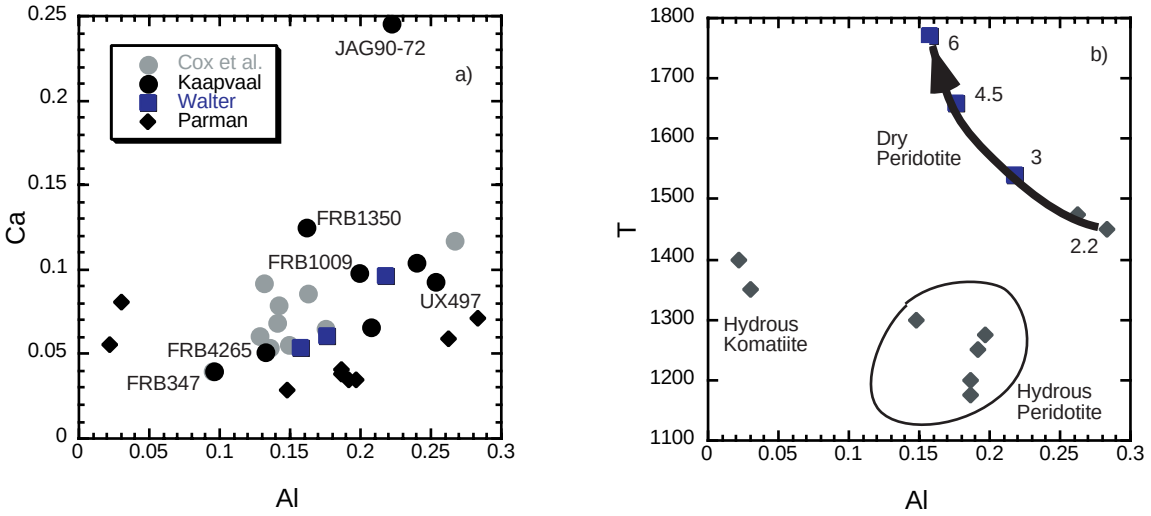


Figure 9. a) Comparison of Ca and Al in the reconstructed Kaapvaal orthopyroxenes from our study (black circles) and from *Cox et al.* [1987] (grey circles) with anyhdrous [*Walters,* 1998; *Parman,* 2001] and hydrous [*Parman,* 2001] peridotite melting experiments. Units are the number of cations based on 6 oxygens. b) Temperature in °C plotted against the number of Al cations per 6 oxygens found in the experimentally produced orthopyroxenes that are in equilibrium with olivine. Numbers indicate pressure in GPa. Hydrous experiments were performed at 1.5 and 2 GPa.

original, high temperature protolith (see below). There is an inferred 500° C difference between the anyhydous and hydrous orthopyroxenes containing the lowest Al (Fig. 9b), so that temperature variations have a small effect on the pyroxene Ca content. The important first-order observation is that neither an anhydrous or hydrous melting origin for the Kaapvaal orthopyroxenes can be excluded.

Kinzler & Grove [1999] (K&G) explored the compositional trends that would be expected in a peridotite residue as it underwent a near-fractional melting event beginning in the garnet stability field. Using experimental data from *Walter* [1998] and *Kinzler* [1997], they calculated the reaction coefficients for the melt reaction and the composition of residues expected for melting up until the exhaustion of garnet and cpx. Their model extracts melt over the pressure range of 3.7 to 1.5 GPa and is reproduced in Fig. 10. The plot of modal olivine vs olivine Mg# (Fig. 10a) shows an enlargement of the low-T peridotite data plotted

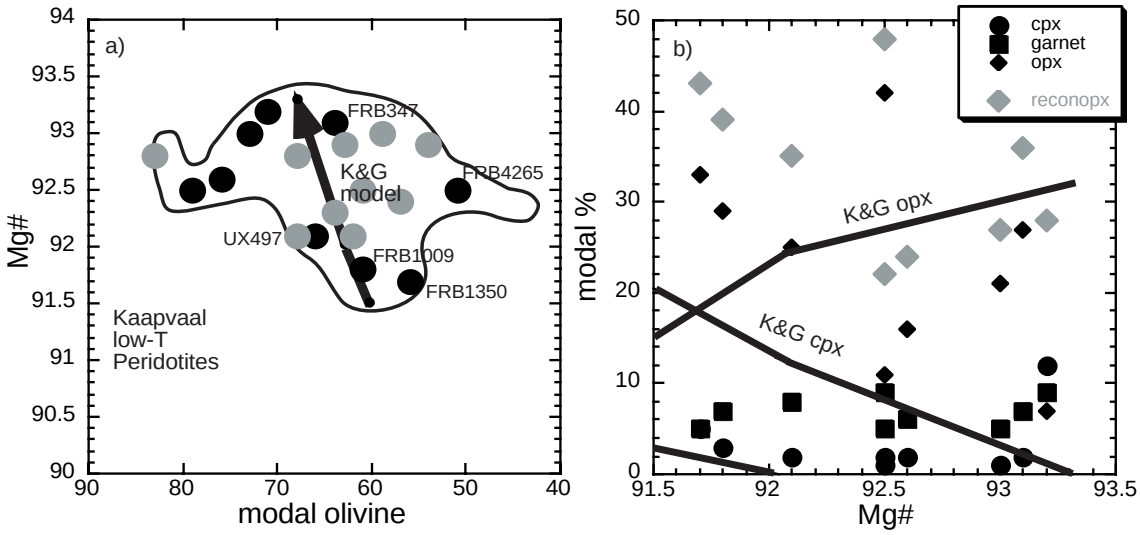


Figure 10. a) *Kinzler & Grove* [1999] model predictions for the evolution of modal olivine and garnet, cpx and opx (lower diagram) vs Mg#. Variations are produced by near-fractional decompression melting that begins in the garnet stability field (at 3.7 GPa) and ends with the exhaustion of cpx (1.5 GPa). Outlined field defines the field of Kaapvaal low-T garnet peridotites from *Boyd* [1989] (see Fig. 3). Arrow indicates the path followed by the mantle residue in the *Kinzler & Grove* [1999] model. b) The solid lines show the variation in modal garnet (lowest line that reaches zero at Mg# of 92), cpx and opx for near fractional melting predicted by the *Kinzler & Grove* [1999] model. Actual Mg# and modal amounts are plotted as dark symbols and the recombined modal opx (grey diamonds) is shown for comparison. The cpx calculated by K&G is a subcalcic augite.

in Fig. 3, and Fig. 10b shows the evolution of the opx, subcalcic cpx and garnet in the residue as melt is extracted. K&G demonstrate that fractional melting can produce part of the trend that extends through the center of the modal olivine vs Mg# variation (Fig. 10a). The extension of that trend to more olivine-rich compositions is not explained by fractional melting and requires an alternative process [*Boyd et al*, 1997; *Kelemen et al.*, 1998]. *Boyd et al.* [1997] propose that the extension of the modal variations to high opx is a consequence of a later stage metamorphic process. K&G propose that olivine enrichment samples are formed by passage of a melt that crystallizes olivine after the fractional melting event has occurred.

When the K&G model predictions are compared with the reconstructed Kaapvaal opx compositions, some interesting systematics emerge (Fig. 10 b). The trend of increasing modal opx with increasing degree of melt extraction predicted by K&G is shown as the solid line. The trend of the more depleted Kaapvaal peridotites (olivine Mg# > 92.5) follows the prediction closely. Sample FRB4265 is the exception and it plots on the olivine-poor side of the predicted fractional melting trend (Fig. 10a). The less depleted samples (with olivine Mg# < 92) show a trend of increasing modal opx with decreasing Mg#, which is the opposite of the modeled trend. However, the K&G model predicts that in this part of the Mg# - modal olivine space a sub-calcic cpx is present in the residue. Therefore, our initial assumption that opx was the only residual pyroxene may be incorrect for these low Mg# samples.

The Kaapvaal samples that plot at the low Mg#, low modal percentage olivine-end of the K&G model trend (FRB 1009 and FRB 1350) plot at the high-Ca and high-Al end of the opx trend in Fig. 9a, consistent with lower extents of melting. Note that the least depleted Kaapvaal opx (UX497) also plots in the appropriate part of Mg# - modal olivine space. The Kaapvaal reconstructed opx with the lowest Ca and Al contents plots at the end of the K&G model trend, where the most depleted residue would be expected to plot.

Another prediction of the K&G model is that the fractional melting residue over the range of olivine Mg# spanned by the Kaapvaal peridotites should have contained both opx and cpx as residual phases. Only at the end of the model

residue trend (where it coincides with FRB 347) would cpx be exhausted from the residue. Therefore, the modeled Ca contents of the reconstructed Kaapvaal orthopyroxenes could be artificially high because some of the cpx that was recombined should have remained apart. Melting experiments suggest that the high-temperature protolith would have contained cpx with ~9-10 wt. % CaO and 7-9 wt. % Al_2O_3 [*Walter,* 1998; *Parman,* 2001]. Our statistical model does not test the presence of mixtures of opx + cpx + garnet that exsolved from both high and low Ca pyroxenes. It is also possible that the reconstructed residue was more similar to the lower pressure, lower temperature trend predicted by K&G and can be accounted for by a shallow anhydrous or hydrous melting process.

Another interesting correlation is the depth from which the samples were derived (Fig. 11). The most depleted samples (FRB 347 and FRB4265) come from the greatest depths whereas the least depleted samples (FRB 1350, FRB1009 and UX497) all come from the shallowest depths. The *Cox et al.* [1987] sample suite shows a trend that parallels the trend defined by our sample suite, but it does not extend to the low temperature, undepleted end member defined by our samples. This relationship is not completely followed by all of the samples because undepleted samples JAG93_8 and JAG90_72 have final equilibration depths equivalent to FRB347, but there are no depleted samples in our dataset that come from shallow depths. If these mantle samples were derived from a melt extraction column that was left behind after near-fractional melting, one would expect that the deepest mantle residues should be the most fertile and the shallowest mantle residues would be the most depleted. The opposite correlation is observed.

Our conclusions are basically similar to those of *Cox et al.*, [1987] who found a spatial relationship between grains of opx, cpx and garnet in 10 low-temperature xenoliths from the Bulfontein and Mothae pipes and compared the reconstructed opx compositions of those rocks to the Al_2O_3 and CaO isopleths of *Yamada & Takahashi* [1984]. In their analysis, *Cox et al.* [1987] concluded that the peridotites were formed as residual products of melting at temperatures near the dry peridotite solidus of Kushiro (1973) in the pressure range of 4–8 GPa. However,

a comparison of their 10 reconstructed orthopyroxenes (Fig. 9) with more recent mantle melting experiments [*Walter,* 1998; *Parman,* 2001] shows they are also consistent with large extents of melting as discussed above.

Following *Cox et al.,* [1987], we have plotted the final equilibration temperatures and pressures of our mantle samples on a pressure-temperature grid derived from the geothermobarometer of *Brey & Köhler* [1990] (Fig. 11). All of them plot in one section of the figure between temperatures of ~770° and 1130° C and pressures of ~2.7-5.3 GPa (100-165 km depth). For comparison, we have also calculated the final equilibration temperatures of the 10 rocks analyzed by *Cox et al.* [1987] and found that they too plot in the same temperature range as our samples (Fig. 11). Distinctly different from these final rocks, the reconstructed pyroxenes have compositions that are consistent with dry melting experiments that would place them at ~1300-1500° C and pressures of 2-3 GPa (60-90 km depth) or experimentally produced orthopyroxenes in equilibrium with a hydrous mantle melts from 2 GPa [*Parman,* 2001].

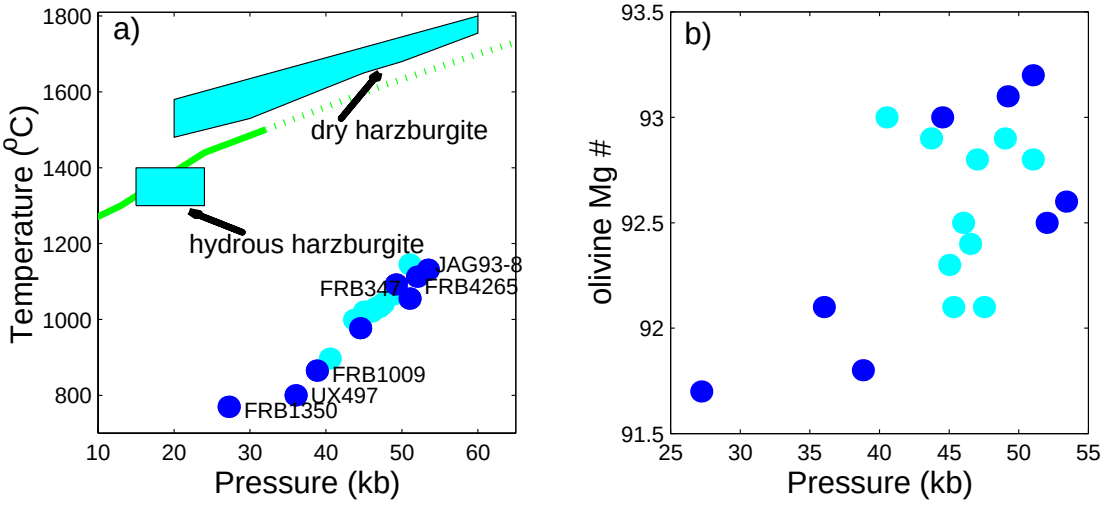


Figure 11. a) Pressure-Temperature plot of all low-temperature samples after final equilibration using the geothermobarometer of *Brey & Köhler* [1990]. The opx stability field is shown for both dry peridotite as determined by melting experiments of *Walter* [1998], and for hydrous peridotite [*Parman,* 2001]. The dry peridotite solidus is from *Kinzler* [1997]. B) Depth of final equilibration from geothermometry is plotted against Mg# of olivine in the sample for the data set from *Cox et al.* [1987] (grey circles) and this study (black circles). Depleted samples are from greater depths. This correlation is the opposite of what would be expected if the samples had preserved a mantle melt column produced by near-fractional decompression melting.

Together, these pieces of information can be reconciled with a scenario in which large degrees of shallow mantle melting (~60-90 km) produce a harzburgitic residuum that is subsequently pushed down to greater depths (~160 km) as a result of cratonic formation. At these greater depths, the high-aluminum opx is no longer stable and the garnet and cpx exsolve, forming the garnet-peridotite rocks that are then erupted as kimberlite xenoliths (Fig. 12). We see no evidence that the xenoliths in our sample suite originated at greater depths (300 to 400 km), as suggested by *Haggerty & Sautter* [1990] since none of the rocks we analyzed showed a relationship between cpx and garnet or opx and garnet alone. However, we cannot rule this scenario out since it is possible that

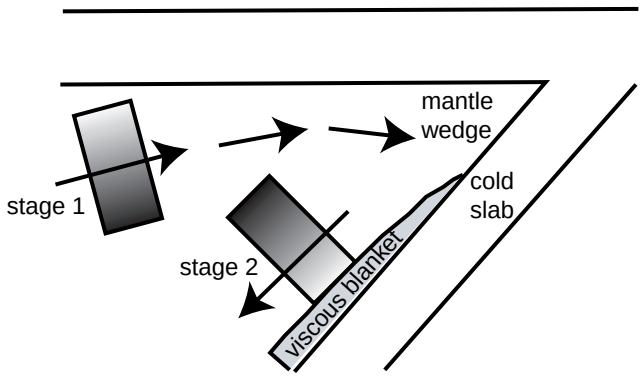


Figure 12. Pressure-Temperature Time Scenario. Large degrees of shallow mantle melting (~100 km) produce a harzburgitic residuum that is variably depleted by the melting event. This is schematically represented by a rectangular section of a melting column in which the deeper (dark grey) mantle residue is less depleted and the most shallow mantle residue (lightest grey) is most depleted. This melt column is subsequently pushed down to greater depths (~160 km) as a result of mantle flow around the corner of the mantle wedge in a subduction zone, where it cools against the cold subducting slab. After the mantle has been delivered to these greater depths, and is cooled by the slab, it re-quilibrates in the cratonic geothermal gradient. High-aluminum opx is no longer stable and the garnet and cpx exsolve forming the garnet-peridotite rocks that are then erupted as kimberlite xenoliths.

majorite would decompose to the same residual products. It should be realized that the ultradeep samples appear distinctly different in texture and mineralogy than the low-temperature harzburgites and that they constitute about 10% of the nodule population at the Jagersfontein pipe [*Sautter et al.*, 1991]. Therefore, these "deeper" samples may be representative of another process that is also preserved in the Kaapvaal peridotite suite. Our results are also consistent with previous studies that suggest the low-temperature xenoliths were formed as

residual products of extensive amounts of melting (>20%) in the Archaean at shallow depths (~100 km) [*Boyd et al.*, 1997; *Keleman et al.*, 1998; *Walter*, 1998].

The thermal history inferred for these mantle samples could have been imposed by mantle flow that occurs in a subduction zone. In modern subduction zones, anhydrous and hydrous mantle melting results in extreme depletion of the mantle lithosphere, and subduction zone melts are extracted from shallow mantle depths [*Baker et al.*, 1994]. After this mantle is depleted it could be entrained in the flow around the wedge corner, where it would undergo cooling as it descended into the mantle above the subducted slab [*Kincaid & Sacks*, 1997; *Davies & Stevenson*, 1992]. The presence of the most depleted samples at the greatest depths is consistent with this model. If the melt column that was left from adiabatic decompression melting in the backarc was dragged around the mantle wedge corner and cooled against the subducted slab, the column may have been inverted (Fig. 12) as it was cooled and dragged to greater depths. If the subduction process shut down or the trench migrated, the depleted mantle would record an initial, hot, shallow equilibration followed by a later, cooler, deeper re-equilibration event. Thus, the depleted Kaapvaal harzburgites may provide a record of Archean subduction processes, and may form a complement to komatiite magmas that have also been proposed to record Archean subduction zone processes [*Parman et al.*, 1997].

Conclusions

We have developed a new technique for 2-D textural analysis that uses the raw pixel data of scanning backscattered electron and x-ray images to determine modal amounts, average crystal sizes and whether any of the minerals are more closely associated spatially than would be expected if they were randomly oriented.

We apply this technique to thin sections of nine, low-temperature garnet-peridotites erupted from kimberlite pipes in South Africa. Eight of the nine rocks we analyzed show a spatial relationship between opx, cpx, and garnet. By recombining these three minerals together, we infer that the original rock was a harzburgite that originated at ~100 km depth from large amounts of mantle melting. The rocks were subsequently pushed down to greater depths in the

creation of the craton and the cpx and garnet then exsolved when they re-equilibrated.

Acknowledgements

We are deeply indebted to Prof. Maarten de Wit of the University of Cape Town and to DeBeers for assistance in collecting the Kimberley samples. Prof. Steven Haggerty at the University of Massachusetts and Dr. Joe Boyd at the Carnegie Institution of Washington generously provided the other xenolith samples for this study. Thoughtful comments from Ben Harte, Martin Drury and a third anonymous reviewer greatly improved this paper. We also appreciate the efforts of Prof. T. H. Jordan who assisted in the development of the textural analysis method. Support for this research was provided by NSF Grant EAR-9526702 under the Continental Dynamics Program.

Table 2. Mineral Compositions

Jag90_72	olivine	+/-	garnet	+/-	opx	+/-	cpx	,,
SiO_2	41.16	0.19	41.83	0.48	58.52	0.24	55.15	0.46
TiO_2	0.00	0.01	0.32	0.02	0.10	0.01	0.27	0.04
Al_2O_3	0.00	0.00	20.47	0.26	0.65	0.01	2.29	0.06
Cr_2O_3	0.02	0.01	4.81	0.22	0.27	0.01	1.92	0.06
FeO	6.67	0.16	6.35	0.16	4.15	0.08	1.93	0.09
MnO	0.07	0.02	0.33	0.03	0.04	0.02	0.02	0.02
MgO	51.49	0.32	20.61	0.18	36.19	0.27	16.80	0.17
CaO	0.03	0.01	5.30	0.15	0.42	0.02	19.92	0.14
Na_2O	0.00	0.00	0.00	0.01	0.01	0.02	1.94	0.07
NiO	0.31	0.03	0.00	0.00	0.00	0.00	0.00	0.00
Total	99.75		100.03		100.36		100.25	

Jag93_8	olivine	+/-	garnet	+/-	opx	+/-	cpx	+/-
SiO_2	40.73	0.19	41.77	0.26	58.45	0.32	54.68	0.46
TiO_2	0.01	0.01	0.14	0.02	0.03	0.01	0.14	0.03
Al_2O_3	0.01	0.01	20.10	0.10	0.72	0.03	3.06	0.11
Cr_2O_3	0.02	0.01	5.33	0.06	0.40	0.01	3.46	0.12
FeO	7.17	0.16	6.96	0.16	4.51	0.16	2.19	0.14
MnO	0.05	0.02	0.38	0.02	0.11	0.01	0.11	0.03
MgO	50.55	0.40	20.63	0.11	36.04	0.17	15.17	0.14
CaO	0.00	0.00	4.65	0.13	0.33	0.01	17.43	0.28
Na_2O	0.00	0.00	0.00	0.00	0.11	0.03	3.31	0.10
NiO	0.39	0.03	0.00	0.00	0.00	0.00	0.00	0.00
Total	98.94		99.96		100.71		99.54	

K1	olivine	+/-	garnet	+/-	opx	+/-	cpx	+/-
SiO_2	40.59	0.18	41.45	0.33	58.07	0.42	54.80	0.11
TiO_2	0.03	0.01	0.04	0.01	0.01	0.02	0.02	0.02
Al_2O_3	0.01	0.01	21.07	0.09	0.75	0.03	1.85	0.07
Cr_2O_3	0.00	0.01	4.01	0.07	0.28	0.02	1.53	0.02
FeO	6.82	0.13	6.59	0.14	4.40	0.12	2.00	0.09
MnO	0.08	0.02	0.30	0.02	0.10	0.02	0.07	0.02
MgO	50.98	0.41	20.64	0.17	36.17	0.21	17.14	0.13
CaO	0.03	0.01	5.18	0.11	0.44	0.04	20.93	0.21
Na_2O	0.00	0.00	0.00	0.00	0.05	0.03	1.62	0.06
NiO	0.47	0.03	0.00	0.00	0.00	0.00	0.00	0.00
Total	99.00		99.29		100.26		99.98	

UX497	olivine	+/-	garnet	+/-	opx	+/-	cpx	+/-
SiO_2	40.69	0.24	41.41	0.35	58.18	0.34	54.86	0.47
TiO_2	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01
Al_2O_3	0.01	0.01	22.22	0.13	0.66	0.04	2.28	0.07
Cr_2O_3	0.01	0.01	2.63	0.07	0.19	0.03	1.69	0.06
FeO	7.79	0.37	8.49	0.15	4.92	0.11	1.50	0.06
MnO	0.10	0.02	0.40	0.04	0.10	0.02	0.05	0.01
MgO	50.63	0.52	19.59	0.15	36.33	0.22	16.25	0.16
CaO	0.01	0.01	4.97	0.13	0.19	0.02	21.58	0.18
Na_2O	0.00	0.02	0.01	0.01	0.00	0.00	1.76	0.09
NiO	0.44	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total	99.67		99.72		100.58		99.98	

Table 2. Mineral Compositions Continued

FRB347	olivine	garnet	opx	cpx
SiO_2	41.30	41.80	58.00	54.80
TiO_2	0.00	0.00	0.00	0.00
Al_2O_3	0.00	20.50	0.86	2.61
Cr_2O_3	0.00	4.14	0.31	1.64
FeO	6.71	6.30	4.18	2.04
MnO	0.09	0.36	0.13	0.06
MgO	50.80	21.50	37.20	16.60
CaO	0.00	4.62	0.45	19.30
Na_2O	0.00	0.00	0.12	2.29
NiO	0.41	0.00	0.00	0.00
Total	99.31	99.22	101.25	99.34

FRB1009	olivine	garnet	opx	cpx
SiO_2	40.39	41.71	56.29	55.57
TiO_2	0.00	0.02	0.00	0.05
Al_2O_3	0.00	21.88	0.71	2.67
Cr_2O_3	0.02	2.87	0.26	1.87
FeO	8.01	8.39	4.80	1.71
MnO	0.08	0.44	0.11	0.08
MgO	50.58	20.32	36.03	15.93
CaO	0.02	4.87	0.21	20.87
Na_2O	0.00	0.01	0.04	2.17
NiO	0.43	0.00	0.06	0.04
Total	99.53	100.51	98.51	100.96

FRB1350	olivine	garnet	opx	cpx
SiO_2	40.96	41.96	57.97	54.87
TiO_2	0.00	0.01	0.01	0.01
Al_2O_3	0.04	22.10	1.13	2.04
Cr_2O_3	0.00	2.07	0.22	0.90
FeO	8.30	8.23	5.20	1.70
MnO	0.11	0.46	0.10	0.07
MgO	51.67	20.09	36.53	17.26
CaO	0.02	5.43	0.27	22.87
Na_2O	0.00	0.02	0.03	1.23
NiO	0.43	0.00	0.10	0.04
Total	101.53	100.37	101.56	100.99

FRB4265	olivine	garnet	opx	cpx
SiO_2	40.28	41.91	58.43	55.91
TiO_2	0.00	0.00	0.00	0.01
Al_2O_3	0.00	20.66	0.80	2.26
Cr_2O_3	0.02	4.34	0.27	1.47
FeO	7.30	6.47	4.36	1.96
MnO	0.08	0.32	0.10	0.10
MgO	50.84	21.10	37.02	17.43
CaO	0.06	5.33	0.50	19.74
Na_2O	0.00	0.02	0.09	1.79
NiO	0.43	0.00	0.12	0.05
Total	99.01	100.15	101.69	100.72

Chapter 4

Comparing P and S wave Heterogeneity in the Mantle

Published by the American Geophysical Union in *Geophysical Research Letters* by Rebecca Saltzer, Rob van der Hilst and Hrafnkell Kárason, *28*, 1335-1338, 2001.

Abstract

From the reprocessed data set of Engdahl and co-workers we have carefully selected matching P and S data for tomographic imaging. We assess data and model error and conclude that our S model uncertainty is twice that of the P model. We account for this in our comparison of the perturbations in P and S-wavespeed. In accord with previous studies we find that P and S perturbations are positively correlated at all depths. However, in the deep mantle systematic differences occur between regions that have undergone subduction in the last 120 million years and those that have not. In particular, below 1500 km depth $\partial \ln Vs / \partial \ln Vp$ is significantly larger in mantle regions away from subduction than in mantle beneath convergent margins. This inference is substantiated by wavespeed analyses with random realizations of the slab/non-slab distribution. Through much of the mantle there is no significant correlation between bulk sound and S-wave perturbations, but they appear to be negatively correlated between 1700 and 2100 km depth, which is also where the largest differences in $\partial \ln Vs / \partial \ln Vp$ occur. This finding supports convection models with compositional heterogeneity in the lowermost mantle.

Introduction

Systematic differences between P and S-wave velocity models can be used to infer mantle properties because the bulk (κ) and shear (μ) moduli have different sensitivities to temperature and mineral composition. In particular, similar behavior in the moduli is consistent with a thermal origin for velocity variations while an inverse relationship suggests chemical heterogeneity or the presence of volatiles [*Stacey*, 1998].

Robertson & Woodhouse [1996] have found that P and S-wave models are correlated to ~2100 km depth but that the patterns of P and S-wave heterogeneity are different below that depth. Other studies focusing explicitly on bulk-sound and S-wave heterogeneity find that variations in κ and μ are decorrelated below ~2000 km depth [*Su et al.*, 1997; *Kennett et al.*, 1998]. Furthermore, *Su et al.* [1994] and *Li and Romanowicz* [1996] suggested that somewhere between 1500 and 2000 km depth imaged velocity structures become longer wavelength. In accordance

with these and other observations [*van der Hilst & Kárason*, 1999], and in an attempt to reconcile geophysical convection models with geochemical constraints, *Kellogg et al.* [1999] suggested the presence of a compositionally distinct region in the lower mantle that is hotter, but nonetheless more dense than the overlying mantle, due to compositional differences.

Previous studies used spherical harmonic representations of long wavelength radial and lateral variations [*Robertson & Woodhouse* , 1996; *Su et al.*, 1997; *Masters et al.*, 2000] or relatively small constant wavespeed cells [*Kennett et al.*, 1998]. Here we take a different approach and investigate whether subduction leaves a discernible imprint on the large-scale distribution of compositionally distinct domains in the deep mantle, as implied by *Kellogg et al.* [1999]. We determine as a function of depth the difference in average P, S and bulk-sound wavespeed perturbations between regions where subduction has occurred in the last 120 million years and where it has not. For the upper mantle we also distinguish between oceans, active tectonic regions, and Precambrian cratons and shields, but we emphasize the results for the lower mantle.

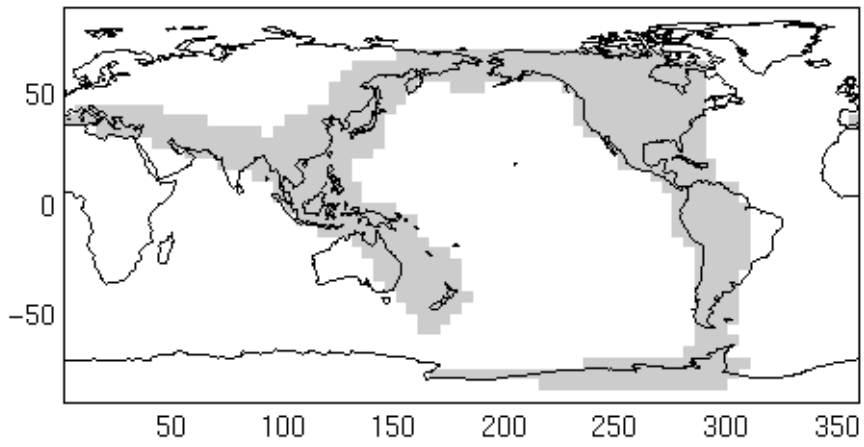


Figure 1. Map depicting regions (light grey) where there has been subduction in the last 120 million years (after *Wen & Anderson* [1995]).

Data and Tomographic Models

Following *Robertson & Woodhouse* [1996] and *Kennett et. al.* [1998] we select bodywave traveltime residuals with common source-receiver pairs. We use the most recent global dataset of reprocessed ISC traveltime residuals [*Engdahl et. al.*,

1998] (EHB), which are better than the original in both hypocenter determination (EHB includes S phases also) and phase identification. Owing to the interference of the S and SKS wave fields near 84° epicentral distance, S picks associated with rays turning in the lowermost mantle are prone to phase misidentification that can affect shear wave models at depths exceeding 2000 km [*Robertson & Woodhouse,* 1996; *Su et al.*, 1997; *Kennett et al.*, 1998; *van der Hilst & Kárason,* 1999]. *Engdahl et. al.* [1998] paid special attention to this problem and we use this dataset to infer aspherical variations in P and S-wave models throughout the entire mantle.

Restricting the P and S-wave datasets to match one another allows us to construct P and S-wave models with similar ray coverage and, hence, similar resolution properties. The individual P and S models could be improved by exploiting all EHB P wave data or by adding constraints from surface wave propagation, but the one to one correspondence between the P and S sampling would then be lost and comparisons between the resulting models more problematic. We further limit the P and S-wave residuals to those that are less than 5 s and 10 s respectively, group the data associated with nearby events and recorded within 1° by 1° regions into summary rays, and take the median of the repeated measurements. This clustering reduces the tomographic system of equations and produces robust residuals for well-traveled paths. The clustered rays contribute more heavily to the solution with a weighting that depends on the total number of rays contributing to the bundle. Bundles containing more than 10 rays are limited to the equivalent weight of just 10 rays so that they do not dominate the solution [*Kárason & van der Hilst,* 2001]. We recognize that the level of noise (picking errors) in the P and S datasets is different, and we account for this in our present study.

Rays are traced through the one-dimensional Earth model ak135 [*Kennett,* 1995], and the P and S-wave data are inverted separately to obtain tomographic images with 1° by 1° by 100 km constant velocity blocks using an iterative, conjugate gradient algorithm. From the independently derived P and S-wave velocity models it is possible to extract a bulk-wavespeed model $V\phi$ using the relation

$$V\phi^2 = Vp^2 - \tfrac{4}{3}Vs^2 = \frac{\kappa}{\rho}$$

which, unlike the P wavespeeds, depends on just one elastic modulus, κ. While this model is perhaps noisier than one that would be generated in a joint inversion of the data [*Kennett et al.*, 1998] it makes explicit the differences between the P and S-wave models.

Even though the P and S-wave models are constructed from a similar ray set and subjected to identical damping and smoothing constraints, we have more confidence in the P-wave model than we do in the S-wave model because the traveltime residuals used to construct the model are of higher quality. We quantify this by determining the scatter of the residuals in ray bundles containing at least 25 paths. On average, we found that the P-wave traveltime residuals for earthquakes originating from within a 50 km square region to any single station have ~0.6 s of scatter. In contrast, the S-wave residuals show ~2.0 s of scatter along the same path. Part of this difference is structural signal but S-waves also tend to have larger picking errors because they often arrive in the coda of P and PP and they may be more prone to effects of attenuation and anisotropy. As regards the latter, we note that ISC does not report on which sensor component an S wave pick was made. The inferred bulk-sound model is noisier still, since it is derived by differencing the P and S-wave models. Nonetheless, we find that the resulting models are qualitatively similar to one another, and when we take the ratio of velocity perturbations from one model to the next the large scale features in the tomographic images remain evident.

To quantify the confidence we have in the P-wave model relative to the S-wave model we add uniformly distributed random noise to the traveltime residuals in an amount equal to the estimated picking error associated with each data type and then calculate how much the model changes when the noisier data is inverted. In over 300 inversion runs we find that with the addition of 0.6 s maximum P-wave noise and 2.0 s maximum S-wave noise the average change in the S-wave model (the block by block difference between the noise-free and noisy models) is just twice (0.02%) that of the P-wave model (0.01%), even though the variability in the raw S-wave data is 3.3 times that of the P-wave data.

This is due to the larger signal in S and the effects of smoothing and regularization imposed upon inversion. In the following we assume that the uncertainty in the S-wave model is twice that of the P-wave model.

To quantify the correlation between the models and to calculate $\partial \ln Vs / \partial \ln Vp$ and $\partial \ln V\phi / \partial \ln Vs$ for a given region and depth of interest, we plot for each cell the magnitude of the perturbations against one another and determine the slope of the best-fitting line by iterative linear regression. After rescaling the axes according to the results of the noise tests, so that both have similar estimated uncertainties, we make an initial guess of the slope and y-intercept and then

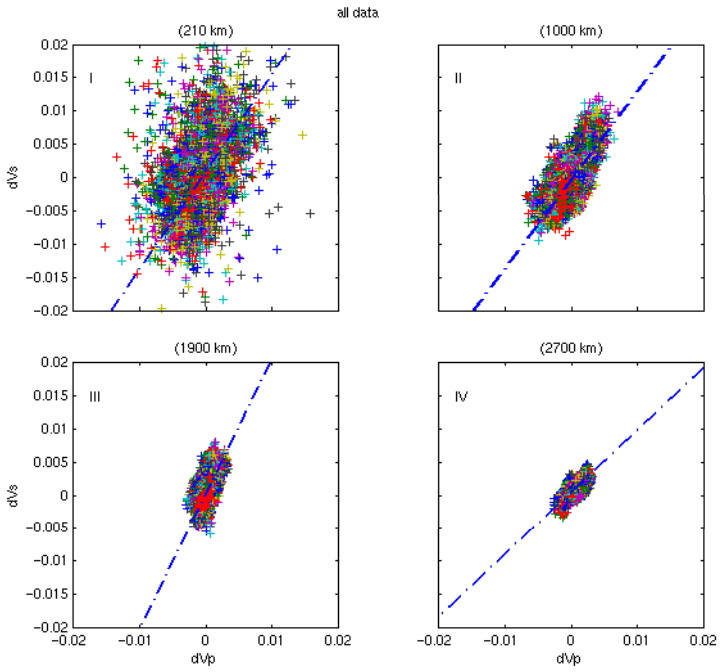


Figure 2. Plots of P-wave versus S-wave perturbations from each of the four depth ranges. Note how the slope of the fitted line changes with depth.

iteratively refit the line by weighting the perturbations that are outside of 1σ by the inverse of their distance from this line. Perturbations within 1σ are weighted equally. This weighting scheme minimizes the effects of outliers and at the same time prevents the data closest to the initial guess from dominating the solution. Unsampled regions are excluded from the analysis. The bulk-sound errors ($\sigma_{V\phi}$) are inferred from the errors in the P and S wave models according to

$$\sigma^2_{V\phi} = \sigma^2_{Vp}(\frac{Vp}{V\phi})^2 + \frac{16}{9}\sigma^2_{Vs}(\frac{Vs}{V\phi})^2$$

which yields an estimate of bulk-sound model uncertainty of ~1.2 times the S-wave model.

Results

We find that the P and S-wave models are strongly correlated at all depths (R~0.6-0.7) confirming earlier results [*Robertson & Woodhouse,* 1996; *Kennett et al.,* 1998; *Masters et al.,* 2000]. However, except for the very deep mantle and the top several hundred km near the surface, we find that through much of the mantle there is no significant correlation between the bulk-sound and S-wave perturbations, although the magnitude and slope of the correlation coefficients as a function of depth is similar to that of *Masters et al.* [2000]. It is possible that this lack of correlation is due to compositional heterogeneity throughout the lower mantle or that the contribution of the bulk modulus is small. It is also possible that our bulk-sound wavespeed model is too noisy for this comparison.

Next, we distinguish regions where there has been subduction in the last 120 million years from those where there has not and loosely divide the mantle into four depth intervals: I (0-660 km), II (660-1500 km), III (1500-2400 km), and IV (2400-CMB). In the upper mantle (depth I), which is best sampled beneath earthquakes and stations, and the mid-mantle (depth II), which is generally well sampled, we find no significant differences in $\partial \ln Vs / \partial \ln Vp$ between regions where there has and has not been subduction in the last 120 million years. However, in the lower mantle the curves have different slopes and begin to diverge at ~1000 km depth. The lack of a statistically significant difference between the slab and non-slab regions to at least 1200 km depth is consistent with an interpretation that the anomalies have a thermal origin.

Between ~1500 and ~2100 km depth, $\partial \ln Vs / \partial \ln Vp$ increases slightly in the slab regions and dramatically in the non-slab regions before decreasing (along with data coverage) toward the core mantle boundary. Previous studies [*Robertson & Woodhouse,* 1996] found a similar increase in $\partial \ln Vs / \partial \ln Vp$ and argued that it results from the increase in pressure with depth, which causes a reduction in the sensitivity of the bulk modulus to changes in temperature [*Agnon and Bukowinski,*

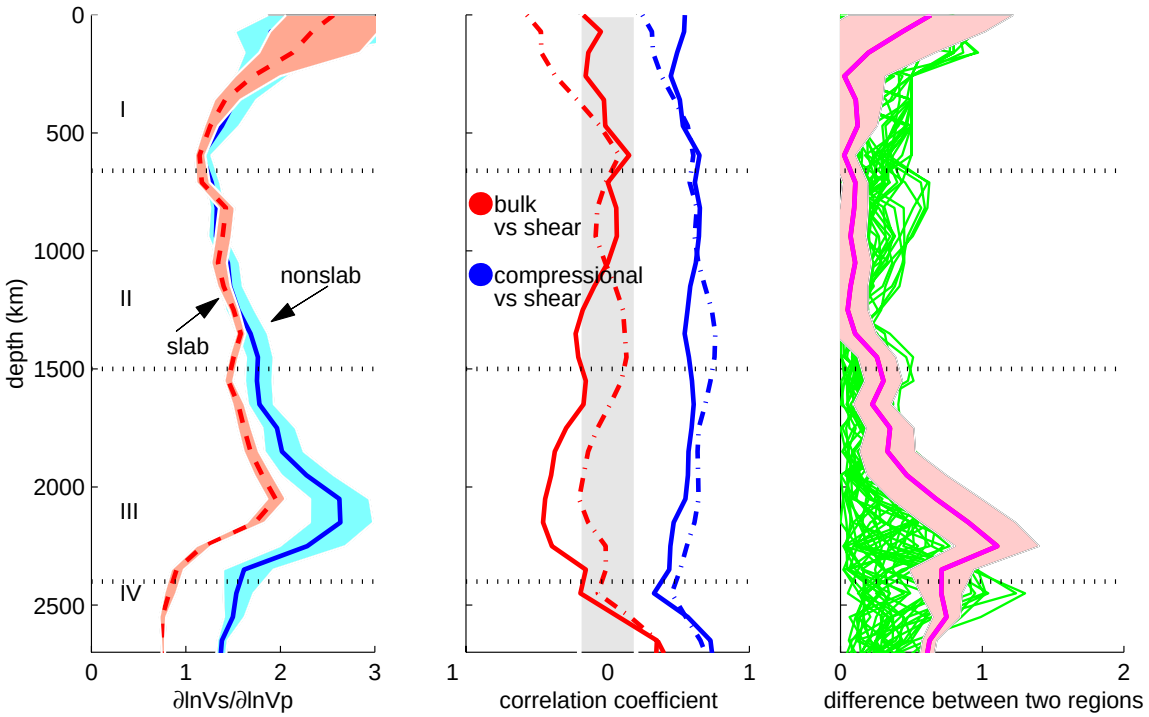


Figure 3. (a) S-wave versus P-wave model perturbations in non-slab regions (solid line) versus slab regions (dashed line) as a function of depth. The shaded areas indicate the 1σ uncertainty and encompass the models allowed by the data. (b) Correlation coefficients for bulk and shear (pair on the left) and P and S-waves (pair on the right) as a function of depth. Slab regions shown with dashed and non-slab-regions with solid lines. The gray zone around zero depicts the range within which the correlations are not thought to be significant. Throughout most of the mantle the bulk and shear models are not significantly correlated except for a negative correlation between 1700 and 2400 km depth. The P and S-wave models are positively correlated throughout the mantle. (c) Difference between the slab and non-slab regions for a series of regionalizations shifted by 10 degree intervals around the globe (green). The regionalization shown in Figure 1 is the one that produces the largest difference between the slab and non-slab regions (pink curve) in the lower mantle, demonstrating that the differences are not due to a bias in sampling or random chance.

1990; *Isaak et al.*, 1992]. While that effect may be occuring, it is unlikely the dominant factor here because at those same depths we also find the bulk and shear wave perturbations are negatively correlated in the non-slab regions (Figure 3b), which suggests, instead, a predominance of compositional heterogeneity [*Stacey*, 1998].

In addition to our uncertainty analysis, we conducted two tests to assess the significance and robustness of the result shown in Figure 3a. First, we applied the above analysis to a large number of different regionalizations. Figure 3c

shows the result of shifting the slab vs. non-slab regions randomly around the globe. The deep mantle difference in $\partial \ln Vs / \partial \ln Vp$ peaks for the regionalization that is based on the actual distribution of slabs, that is, the regionalization as shown in Figure 1. For any other regionalization the differences are smaller, except in the shallow mantle. Second, we have tested whether preferential sampling of low-velocity regions beneath the southern Pacific ocean and Africa as suggested by *Masters et al.,* [2000] could bias our results and found that excluding those regions from the analysis does not significantly change them. These results suggest that the deep mantle difference in $\partial \ln Vs / \partial \ln Vp$ between the slab and non-slab regions and the negative correlation in $\partial \ln V\phi / \partial \ln Vs$ are robust and causally related to the pattern of subduction in the lower mantle.

The boundary layers (regions I and IV) are another part of the globe where we find significant differences. In the very top of the upper mantle beneath cratons and shields (not shown) we find a negative correlation in $\partial \ln V\phi / \partial \ln Vs$ suggesting they are compositionally distinct, in support of the tectosphere hypothesis [*Jordan,* 1975]. In addition, $\partial \ln Vs / \partial \ln Vp$ is significantly higher in both tectonically active and young, oceanic regions than in more stable, continental regions, which is consistent with high temperature gradients or the presence of partial melt and volatiles. At the very bottom of the mantle (the lower boundary layer) we find a positive correlation between $\partial \ln Vs / \partial \ln Vp$, which is consistent with thermal effects at the core-mantle boundary.

Conclusions

The deep mantle (>1000 km depth) beneath regions that have undergone subduction in the last 120 million years have consistently lower $\partial \ln Vs / \partial \ln Vp$ ratios than regions that have not. In addition, largest difference occurs between 1700 and 2400 km depth where a significant negative correlation between $\partial \ln Vs / \partial \ln Vp$ suggests widespread chemical heterogeneity. These results do not dictate the nature of the compositional heterogeneity and are consistent with mantle models involving anomalous domains, as envisioned by *Kellogg et al.* [1999], or slab accumulations, as suggested by *Christensen and Hofmann* [1994], provided the latter extend far enough above the core mantle boundary.

The boundary layers (regions I and IV) are another part of the globe where we find significant differences. Beneath Pre-cambrian cratons and shields to ~250 km depth, we find a negative correlation in $\partial \ln V\phi / \partial \ln Vs$ suggesting compositional heterogeneity. In tectonically active and young oceanic regions $\partial \ln Vs / \partial \ln Vp$ is significantly higher than in more stable, continental regions, consistent with elevated temperature gradients and partial melt or volatiles. We find a positive correlation between $\partial \ln V\phi / \partial \ln Vs$ at the base of the mantle (the lower boundary layer) which is consistent with thermal effects at the core mantle boundary.

We realize that higher-quality traveltime residuals are necessary to provide a more complete and robust picture. In particular, the S-wave model is less robust than the P-wave model despite similar ray sets, damping etc., which we attribute to lower quality S-wave residuals. Higher-quality S-wave and bulk-sound wavespeed models are required to determine whether the lack of correlation in bulk and shear wavespeed models throughout much of the mantle is due to physical properties in the mantle or is just an artifact of noise in the models.

Acknowledgments

We thank Frank Stacey and Jeannot Trampert for thoughtful reviews of the manuscript. This research was supported by NSF under grant EAR-9905779 and the David and Lucille Packard Foundation.

Chapter 5

Poisson's Ratio beneath Alaska from the surface to the Core-Mantle Boundary

Submitted for publication to the American Geophysical Union in the *Journal of Geophysical Research* by Rebecca Saltzer, Eléonore Stutzmann and Robert D. van der Hilst, 2001.

Abstract

Using waveform cross-correlation, we have picked P and S wave differential travel times from earthquakes and receivers lying along a great circle path that spans Japan-Alaska and the Western United States. We jointly invert the data using 2° × 2° × 200 km constant velocity blocks and then calculate Poisson's ratio in each of those blocks. In the mid mantle, the magnitude of the lateral variations in Poisson's ratio is large (~8%) and the dominant length scales over which these variations occur are ~2000 km and ~4000 km in size. In the bottom 1000 km of the mantle where laterally coherent, slab-like structures are absent, the magnitude of the lateral variations in Poisson's ratio decrease to 3% and the dominant length scale over which variations take place is ~4000 km. The variability in Poisson's ratio is greater than what would be expected by temperature variations alone, assuming that the four different estimates of $\frac{\partial V_p}{\partial T}$ and $\frac{\partial V_s}{\partial T}$ we have used [*Karato*, 1993; *Karki & Stixrude*, 1997; *Stacey*, 1998; *Trampert et al.*, 2001] are anywhere in the ballpark. A simple explanation is that the variability in Poisson's ratio includes a contribution from compositional effects. One effect that could explain the difference is the depletion of the subducted slabs in perovskite relative to the convecting mantle as suggested by *Ringwood* [1991]. Another possibility that can explain the differences is that we are seeing variability in iron content from one region to another.

1. Introduction

An issue that has been receiving a lot of attention lately is whether there might be compositionally distinct domains in the lower mantle. Both geochemical and heatflow considerations suggest that such mantle reservoirs have existed for 1 billion years or more [*Zindler and Hart*, 1986; *Albarède and van der Hilst*, 1999], however the seismic evidence for the existence of such domains is more equivocal. In an effort to reconcile these views, *van der Hilst and Kárason* [1999], *Kellogg et al*. [1999] and *Davaille* [1999] have proposed the existence of a compositionally distinct layer in the lower mantle that is hotter but at the same time intrinsically more dense as a result of compositional differences, which renders it essentially "invisible" to traditional seismic detection. In a similar vein, *Tackley* [1998, 2000] has suggested "piles" of material at the base of the core-

mantle boundary (CMB) and *Becker et al.* [1999] have suggested blobs in the lowermost mantle.

Although none to date really proves it, an emerging body of seismic evidence is consistent with the idea of compositional heterogeneity in the lower mantle. For example, studies of global P and S-wave models find that the patterns of P and S-wave heterogeneity are strongly correlated down to ~2100 km depth but that below that depth the correlation declines [*Robertson and Woodhouse,* 1996]. In addition, below ~2000 km depth other studies find that $\partial \ln Vs / \partial \ln Vp$ is greater than 2.5 [*Romanowicz*, 2001] and that the bulk (κ) and shear (μ) moduli behave significantly differently from one another [*Kennett et al.,* 1998; *Su and Dziewonski,* 1997; *Masters et al.,* 2000]. Studies that consider lateral variability find that the high values for $\partial \ln Vs / \partial \ln Vp$ found in the 1-D profiles are characteristic only for deep mantle regions away from regions of recent subduction (< 120 million years) and that $\partial \ln Vs / \partial \ln Vp$ maintains expected values near downwellings [*Saltzer et al.*, 2001]. Moreover, they find that bulk and shear models for the non-slab regions are negatively correlated at the same depths where $\partial \ln Vs / \partial \ln Vp$ becomes abnormally high.

Although still debated [e.g., *Resovsky and Ritzwoller,* 1999; *Romanowicz,* 2001] a recent joint inversion of seismological and gravity data finds high densities below two major low-velocity regions of upwelling in the lower mantle [*Ishii and Tromp,* 1999, 2001] which is consistent with lower mantle heterogeneity. In addition, global tomographic studies have found that the P and S-wave velocity structures change significantly from predominantly shorter to longer wavelengths, somewhere between 1500 and 2000 km depth [*Su et. al.,* 1994; *Li and Romanowicz,* 1996] and that in the current snapshot of convection few slabs seem to descend below ~1700 km depth [*van der Hilst and Kárason*, 1999*; Kárason and van der Hilst*, 2001]. While these various studies are consistent with compositional heterogeneity, they do not unequivocally prove the existence of such domains. Moreover, there has been no evidence to date of reflected or converted phases that might be expected if a compositionally distinct layer existed [*Vidale et al.,* 2001; *Castle and van der Hilst,* 2001]. Consequently, the nature and existence of such domains remains debated.

In this paper, we investigate the idea of compositional heterogeneity further by looking at the variability in Poisson's ratio both laterally and as a function of depth. Our study differs from most Vp/Vs studies to date in that it focuses on a specific region rather than drawing inferences from global inversions that are plagued with uneven data coverage and differences in sampling of the P and S waves. We have chosen a great-circle transect between Japan, Alaska and the western US to study propagation differences between P and S waves. We selected this particular study area for three reasons. First, the entire transect lies along continental or island arc plate boundaries, so there is a fairly continuous distribution of earthquakes and stations (Figure 1). This geometry produces a

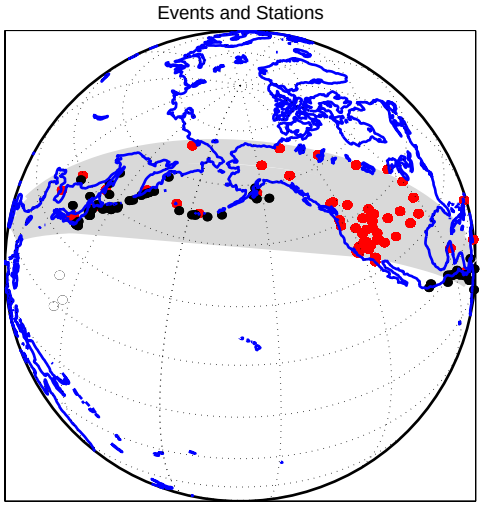


Figure 1. Location map of earthquakes (black circles) and stations (red circles) along transect. Shaded grey region shows width of corridor investigated.

large number of crossing rays within the corridor of interest that can be used to illuminate the region well and allow us to investigate the idea of deep mantle compositional heterogeneity better. Second, Wadati Benioff zone seismicity and body-wave travel time tomography [e.g., *van der Hilst et al.*, 1997; *Bijwaard et al.*, 1998; *Kárason and van der Hilst, in prep.*, 2001] delineate the trajectories of mantle flow in the upper mantle beneath Alaska and the Aleutian arc and in the middle mantle only to depths of ~1700 km beneath North America. Figure 2 depicts the predicted locations of slab material subducted between 90 and 120 Myr according to a recent numerical mantle flow calculation that accounts for lateral advection ("mantle wind") [*Steinberger*, 2000]. The predicted location of slab

fragments is in excellent agreement with the images of P and S wave speeds in the mid mantle. However, at depths beneath 1700 km or so, the tomographic models reveal more scatter and laterally coherent slab structures appear absent in P-wave models [*van der Hilst et al.,* 1997; *Kárason and van der Hilst, in prep,* 2001]. The dense data coverage produced by the distribution of sources and receivers (Figure 1) allows us to investigate in more detail than hitherto possible if, and how, slab fragmentation is reflected in the images of shear and bulk-sound speed and Poisson's ratio. This information may contain important clues as to how downwellings interact with deep mantle domains, if present.

Third, there have been a number of interesting but also puzzling observations in this region. In particular, near the CMB it has been observed that S-waves are fast while P-waves are slow [*Wysession et al.,* 1992; *van der Hilst et al.,* 1997; *Grand et al.,* 1997; *Li and Romanowicz,* 1996; *Masters et al.,* 1996; *Bijwaard et al.,* 1998; *Kárason and van der Hilst,* 2001], which is difficult to explain by thermal variations alone. In addition, there is both evidence for [*Revenaugh and Meyer,* 1997; *Garnero and Helmberger,* 1998] and against [*Castle and van der Hilst*, 2000; *Persh et al.,* 2001] an ultra-low velocity zone at the CMB directly beneath and to the south of the Aleutians (Figure 2) and evidence of increased lateral variation in anisotropy in the bottom 200 km from beneath Alaska eastwards [*Matzel et al.,* 1996; *Garnero*

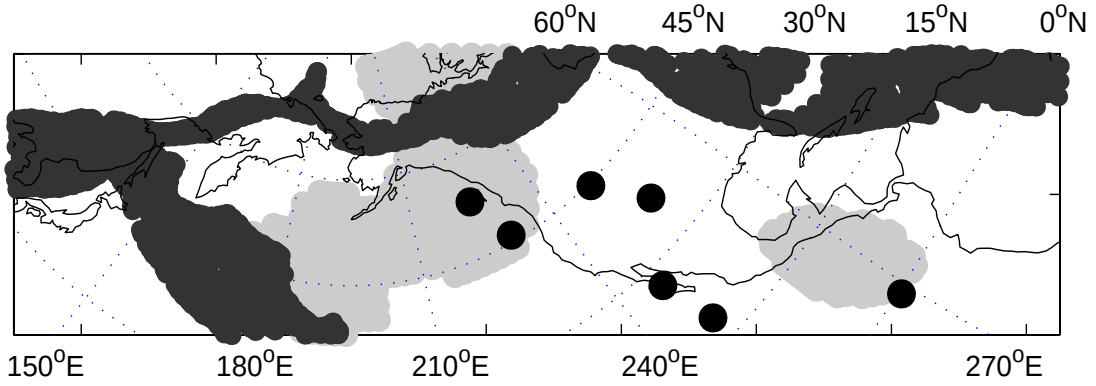


Figure 2. Locations of ultra low-velocity zones reported in previous studies (light grey) [e.g., *Revenaugh & Meyer,* 1997; *Garnero & Helmberger,* 1998], inferred locations of slabs >90 million years old from plate reconstructions [*Steinberger,* 2000], and the surface projection of hotspots at the core-mantle boundary (black circles) [*Richards et al.,* 1988]. Figure is aligned along the great circle transect investigated in our study. Dashed lines show parallels and meridians for geographical correspondence.

and Lay, 1997; *Lay et al.*, 1998].

In this paper, all map views are aligned with the great circle arc along which the corridor we have investigated. In Figure 2, we have superimposed parallels and meridians so that the correspondence between images and geographical locations can be made more easily.

2 Data

We used waveform cross-correlation on unfiltered, broadband seismograms to measure a total of 16,926 differential travel times from 128 events and 76 stations (Figure 1). The data were obtained from the archives of the Data Management Center (DMC) of the Incorporated Research Institutions for Seismology (IRIS) and were comprised of large events (Mb>5.8) between the years 1993 and 2000 that occurred anywhere along the corridor of interest. Events outside of this corridor were excluded in order to avoid the erroneous mapping of structure from outside the plane. The dominant period of the broadband data was ~1 s for the P-waves and ~10 s for the S-waves. This difference has a negligible effect on the elastic moduli.

For a given event, we selected a reference station and cross-correlated the P or S waveform with the direct arrival of all the other stations to obtain P or S differential travel times. The PcP-P and ScS-S travel times were made for a given source-receiver pair by selecting the P or S waveform, applying a t* operator to account for differences in attenuation along the raypaths, and then cross-correlating that waveform with the PcP or ScS phase. We also tried to measure PP-P and SS-S traveltimes, but found that those phases were very weakly excited and that we obtained so few measurements (and of questionable quality) that we did not include them in the inversion. The shifted traces were inspected visually so that errors due to cycle skipping could be corrected and noisy and otherwise unpickable traces discarded. The Harvard Centroid Moment Tensor database of source mechanisms were used to determine when the polarity of a given seismogram should be reversed relative to the reference waveform and corrections for the earth's ellipticity were made after the measurements were complete.

We calculated differential travel time residuals by taking the difference between the ak135 [*Kennett et al.*, 1995] predicted travel time difference and the observed differential times. Only when both the P wave and S wave differential times for a given source-receiver pair (and phase) could be successfully measured would those traveltimes be included in the inversion so that the raypaths and geometries (and ultimately model resolution) would be as similar as possible for both resulting tomographic models. Unfortunately, many events that excited beautiful ScS did not excite any discernible PcP and vice versa. Therefore, we ended up with just 99 matched pairs of PcP-P and ScS-S traveltime measurements (198 in total) with the remainder of our measurements (16,728) comprised of relative P and relative S wave traveltime measurements.

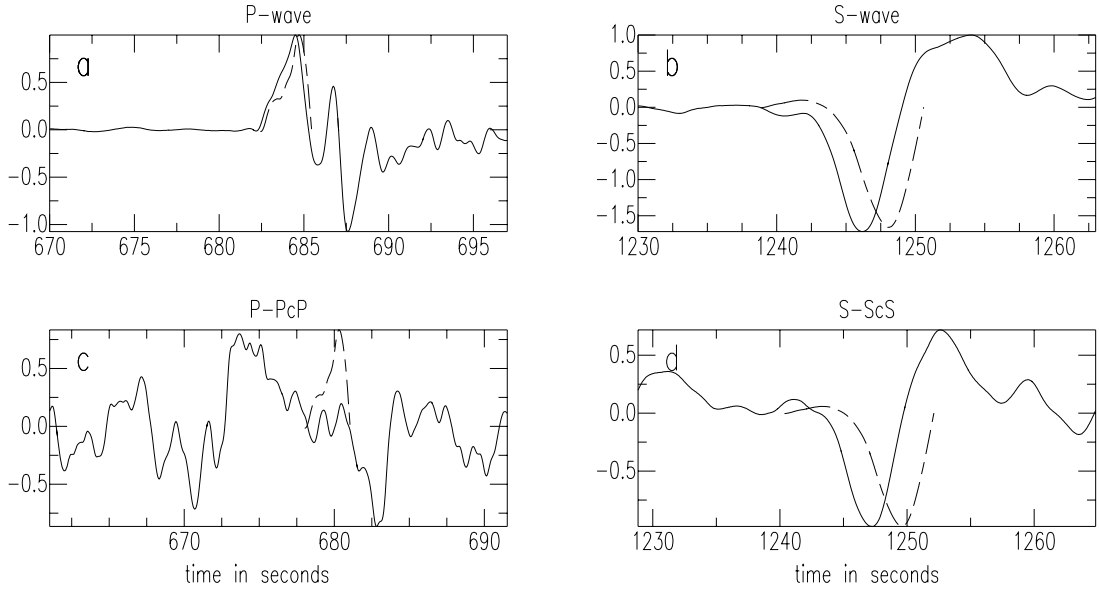


Figure 3. Example of reference trace (dashed) and data (solid) for an mb 6.2 event in the Kuril Islands on January 28, 2000. The reference trace has been shifted to the ak135 predicted arrival time for the P wave (a) and S wave (b). Note that the differential P-wave arrival time is ~0.3 s earlier than the predicted differential time from ak135 whereas the differential S-wave arrival times for the same station pair is ~2 s. Bottom panels show PcP along with the shifted P-wave reference trace (c) and ScS along with the shifted S-wave reference trace (d) for a single station. In general, the PcP data are much noisier than the ScS data for all events, limiting the number of differential travel times that could be measured.

3 Methodology

3.1 Inversion

Rays were traced through the radially stratified one-dimensional Earth model ak135 [*Kennett et al.*, 1995] using the improved earthquake hypocenters of *Engdahl et al.* [1998]. The P and S wave data were jointly inverted to obtain tomographic images of P and S-wave models using the following formulation

$$\begin{matrix} G & & m & = & d \end{matrix}$$
$$\begin{bmatrix} P-matrix & 0 \\ 0 & S-matrix \end{bmatrix} \begin{bmatrix} P \\ S \end{bmatrix} = \begin{bmatrix} \delta t_p \\ \delta t_s \end{bmatrix}. \quad (1)$$

The G-matrix contains the partial derivatives relating the functional dependence of the traveltime residuals to the slowness perturbations, m is the slowness model vector, and d is the vector of measured differential traveltime residuals. Each row of the G-matrix corresponds to a particular raytraced source-receiver pair and each column corresponds to one of the model parameters.

From the P and S wavespeeds, the bulk-sound wavespeed can be derived.

$$V_\phi^{\,2} = \frac{\kappa}{\rho} = V_p^{\,2} - \tfrac{4}{3} V_s^{\,2} \quad (2)$$

One attractive property of the bulk-sound wavespeed is that it depends on just the bulk modulus, (κ), and density (ρ) alone, so that the separate effects of the bulk and shear moduli can be evaluated. In addition, the bulk-sound wavespeed can be directly compared with experimental and theoretical mineral physics results and is relatively insensitive to anelastic effects. Because of its inherent relationship to the P and S wavespeeds, it is possible to obtain tomographic images of bulk-sound and shear wavespeed models from joint inversion of P and S wave data using a similar formulation to the one above

$$\begin{matrix} G & & m & = & d \end{matrix}$$
$$\begin{bmatrix} bulk-matrix & 0 \\ 0 & S-matrix \end{bmatrix} \begin{bmatrix} \phi \\ S \end{bmatrix} = \begin{bmatrix} \delta t_p \\ \delta t_s \end{bmatrix}. \quad (3)$$

The partial derivative of the P-wave travel times with respect to the bulk sound and shear slowness (the bulk-matrix) are obtained by rewriting equation (2) in terms of slowness (inverse of velocity) and computing the appropriate derivatives which yields

$$\delta t_p = \int_{ray} ds \delta a = \int_{ray} ds \left(\frac{a^3}{c^3} \delta c + \frac{4a^3}{3b^3} \delta b \right) \quad (4)$$

where $a = \frac{1}{Vp}$, $b = \frac{1}{Vs}$, $c = \frac{1}{V\phi}$, δt_p is the P-wave traveltime residual and δa, δb, and δc are the perturbations in P-wave, S wave and bulk-sound slownesses respectively.

We inverted the traveltime data using both formulations (equations (1) and (3)) with 2° by 2° by 200 km constant velocity blocks. The model grid contained 5083 parameters and was aligned along the great-circle corridor containing the majority of the sources and the receivers. An iterative, conjugate gradient algorithm solved the system of equations [*Paige and Sanders,* 1982] which included additional model parameters to absorb the effects of earthquake mislocations. We found that all events moved less than 10 km and generally only 1 or 2 km in any direction. In part, this is due to the use of *Engdahl et al.'s* [1998] improved hypocenter locations, which we took as our starting locations. However, the differential travel times used here have a reduced sensitivity to errors in source location and origin time, which also explains why the events moved such small distances. Optimal amounts of damping and smoothing were determined by a series of tests with synthetic data in which the model misfit error and amount of damping/smoothing were compared. The tomographic models presented here were obtained after 200 iterations and produced a 90% variance reduction.

3.2 Uncertainties

There are several sources of error in our results. We estimate our P and S wave picking errors at 0.2 s and 0.4 s respectively. On the basis of noise tests in previous work [*Saltzer et al.,* 2001] we estimate the effect of these errors on the resulting tomographic models to be less than 0.01% (when measured on % deviation tomographic images) and on calculations of Poisson's ratio in the

following results section to be less than 0.04%. Since we employed just one-dimensional ray tracing (we did not re-raytrace through imaged anomalies) we are implicitly assuming that the velocity anomalies do not distort the raypaths. The maximum velocity variations we found were just a few percent and so this assumption seems justified. We also carried out various synthetic tests to understand problems associated with our sampling geometry and to verify the resolution of our model (i.e. checkerboard, random slabs, etc).

Figure 4 shows an example of a synthetic model containing velocity

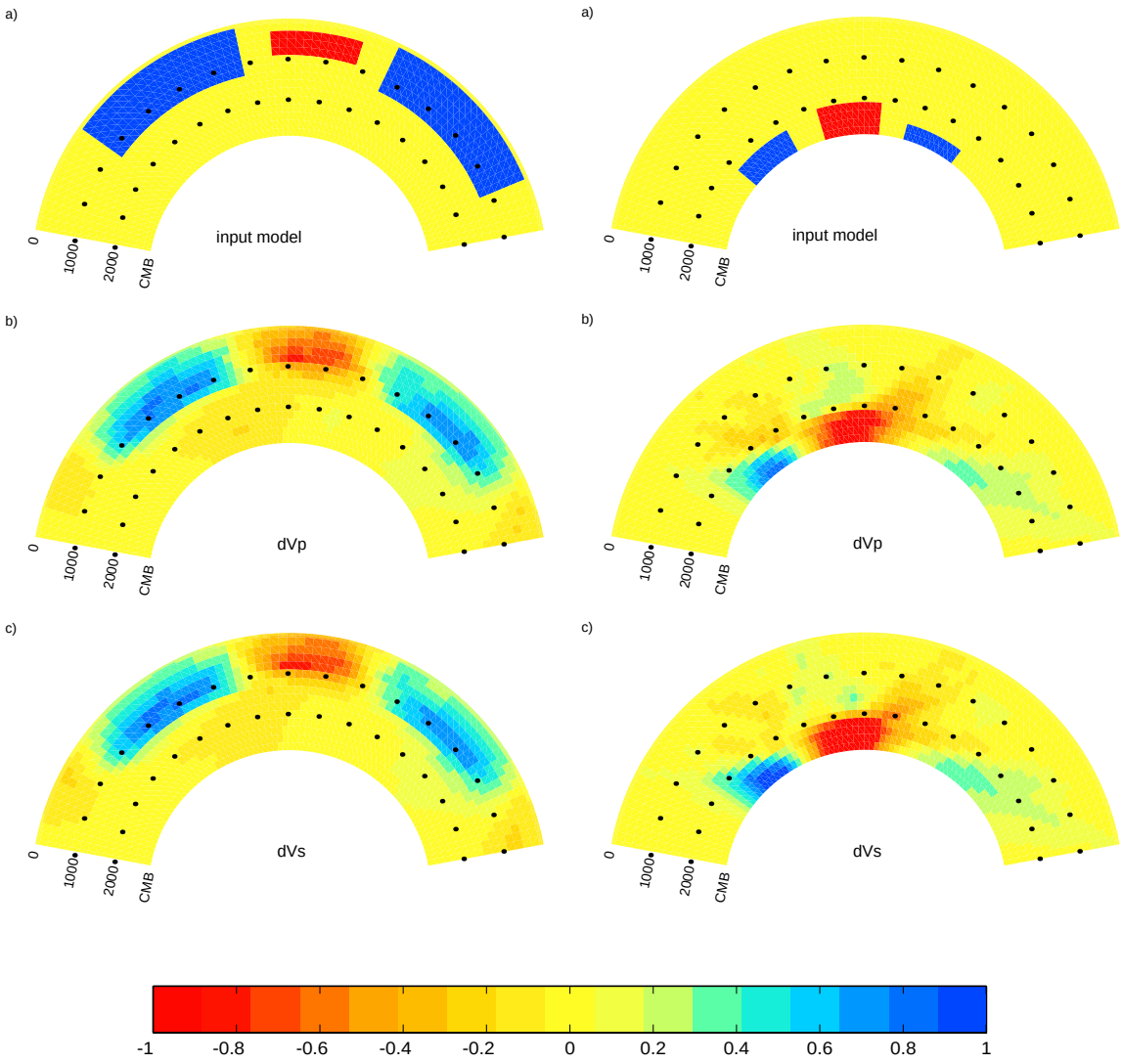


Figure 4. Results of two synthetic tests. Rays are traced through a known velocity model (a) and then the resulting traveltime anomalies are jointly inverted to see what structure is actually recovered for the (b) P wave model and (c) S wave model. First set shows anomalies in the shallow mantle and second set shows smaller anomalies in the very deep mantle.

anomalies in just the upper mantle and an example in which velocity anomalies are in just the bottom of the mantle. Figure 5 shows the number of times each block was hit. As a result of dense path coverage, streaking of anomalies in the radial direction from the surface downwards is small from which we conclude that the lower mantle anomalies we image are indeed required by the data. However, some smearing does occur upwards when anomalies are placed at the very base of the mantle which is not surprising given the arcuate nature of ray paths and decline in data coverage with depth.

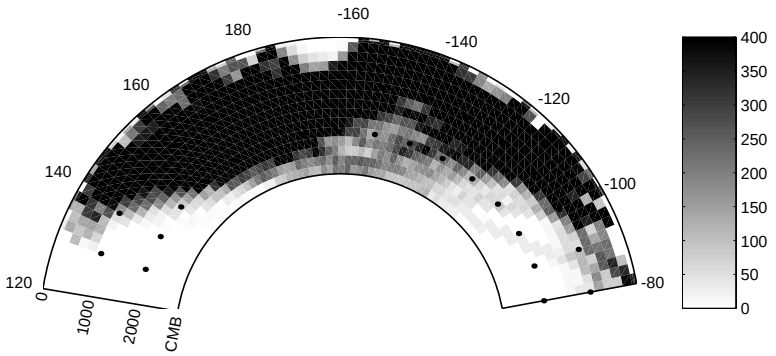


Figure 5. Hit count. Number of times a ray passes through each cell of the tomographic models. Since the source-receiver geometry for the P and S waves is the same, the hitcount pattern looks the same for both as well. Note that even at the base of the mantle, the blocks in the central region are hit at least 100 times.

4 Results of inversion for elastic parameters

We find that the bulk-sound model obtained through explicit joint inversion is virtually indistinguishable from the inferred bulk-sound wavespeed models using equation (2) confirming the results of *Kennett et al.* [1998], but that the amplitudes of the perturbations are slightly greater for the inferred bulk-sound wave speed model than the one that is directly inverted for. Neither type of bulk-sound model consistently shows a significant correlation with the shear-wave models (correlation coefficient < 0.2) whereas on average we find a high degree of correlation (correlation coefficient =0.8) between the P and S wave models except at depths below ~2300 km (Figure 6). We consider correlations above 0.25 to be significant. This high-degree of P and S wave correlation is similar to results obtained for many global inversions of P and S wave data [i.e. *Robertson and Woodhouse*, 1996; *Kennett et al.*, 1998; *Masters et al.*, 2000; *Saltzer et al.*,

2001]. The low degree of correlation between bulk-sound and shear (and the value of the correlation coefficient) is consistent with our previous global analysis [*Masters et al.*, 2000; *Saltzer et al.*, 2001].

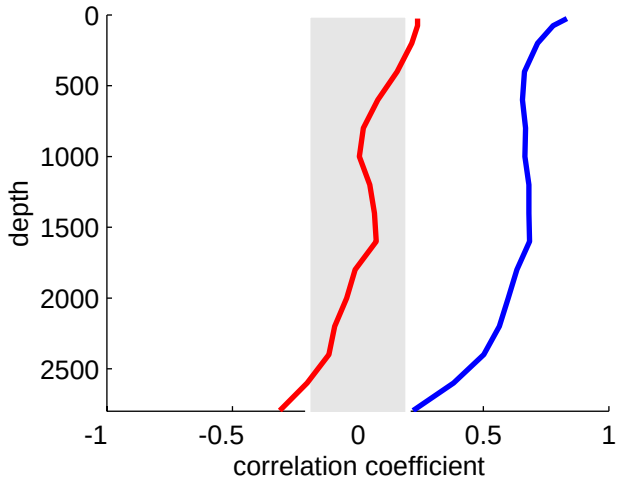


Fig 6. Correlation coefficients for P and S models (blue) and bulk sound wavespeed and shear models (red).

In order to compare the P and S wave tomographic images quantitatively, we calculate the Poisson's ratio (σ) in every block through which at least two P and two S rays passed. Poisson's ratio has the following relationship to Vp and Vs:

$$\sigma = 0.5 * \left[1 - \frac{1}{\left(\frac{V_p}{V_s}\right)^2 - 1} \right] = \frac{\lambda}{2(\lambda + \mu)}, \qquad (5)$$

where λ and μ are the Lamé parameters (μ=rigidity). Figure 7(a-f) shows a map of shear-wave (Vs) and bulk-sound (Vϕ) velocity perturbations at various example depth slices and Figure 7(g-l)) shows the corresponding values for Poisson's ratio along with the sampling density at the same depths.

In the upper mantle, ray coverage is fairly sparse since it is restricted primarily to blocks directly beneath the sources. Therefore, we choose not to interpret the calculated Poisson's ratios in the shallow mantle. Beginning at ~650 km depth the ray coverage is much more dense and coherent patterns begin to emerge. Near the Aleutian arc and near the Kuriles there are regions of lower Poisson's ratio (implying greater rigidity) in the same locations where there has been recent subduction [e.g. *Steinberger*, 2000] with ± 5% variability. It seems likely that this variability is simply a result of thermal anomalies from the cold

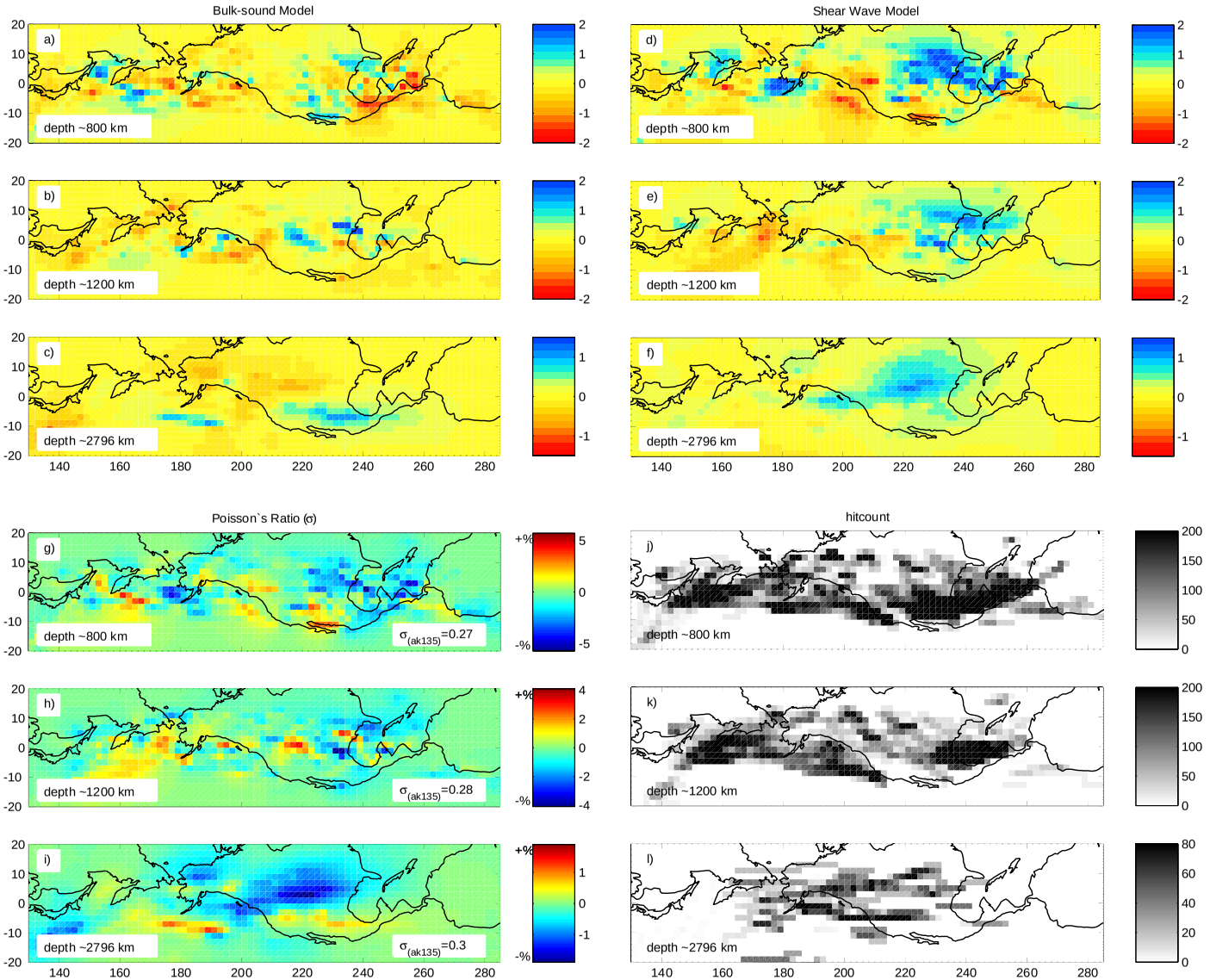


Figure 7. Depth slices at ~700 km, ~1100 km and the bottom 200 km of the mantle. through the Bulk-sound (a-c) and Shear-wave (d-f) models. Poisson's ratio is computed for every block according to equation (5) for both the ak135 reference model [Kennett et al., 1995] and for the inferred Alaska corridor compressional and shear wave models (g-i). The difference between the two are shown as percent deviations from ak135. The number of times each cell is hit (hitcount) is shown for comparison (j-l). Latitude and Longitude coordinates are in reference frame that is aligned along the great-circle corridor (Pole of rotation is at latitude = 32 ° and longitude = 15°).

slabs subducting in the mantle. Consistent with this interpretation, the shear-wave images show seismically fast regions where recent subduction has occurred (what we interpret as "slab fragments"). However, these fragments do not show up as visibly on the bulk-sound maps. This observation is similar to what *Widiyantoro et al.*, [1999] found for the northwest Pacific slab systems they studied.

At depths of ~1050 km, in the region behind the Aleutian arc we observe a shift in Poisson's ratio from low to high (implying softer or warmer material properties). The higher Poisson's ratios persist to ~2000 km depth. In the bottom 200 km, above the core-mantle boundary and within what most people would call D", we see a band of lower Poisson's ratio (more rigid) material more or less in the predicted locations of ancient subducted slabs (>90 million years old). At these same locations, the bulk-sound and shear wave models appear to be anti-correlated. The regions of lower Poisson's ratio surrounds a small area of higher Poisson's ratio that lies within a region that has been identified in two studies as an ultra-low velocity zone [*Revenaugh and Meyer*, 1997; *Garnero and Helmberger*, 1998] (Figure 2). The value of Poisson's ratio in this central zone is ~0.29 which is similar to but not quite as low as the value found by *Wysession et al,* [1999] with diffracted waves, while the value in the surrounding regions is ~0.31. It should be noted that the results from the diffracted wave studies represent values averaged over large regions owing to the broad fresnel zones of the long-period diffracted waves. By taking a regional approach in our study, we get much better spatial resolution of structure over smaller length scales.

Figures 8(a,b) show lateral averages of Poisson's ratio as a function of depth. Poisson's ratio has a much larger dynamic range at 1000 km depth (±4%) than below 1500 km depth (±2%), where it remains fairly constant down to the core-mantle boundary. At all depths, the average Poisson's ratio value is within a fraction of a percent of the value given by the ak135 reference model.

Figure 8c shows the dominant wavelengths of imaged structure (along the great-circle arc of our transect) as a function of depth. These were determined by taking the Fourier transform of the center row of each depth slice as well as the two adjoining rows on each side and computing the power density spectrum at each depth (accounting for the decrease in block size that occurs as a function of

depth when determining the wavelength). The two strongest peaks are found at wavelengths of ~4000±800 km and ~8500±2000 km (corresponding to features of ~2000 km and ~4250 km in size); however, the relative power of the two varies as

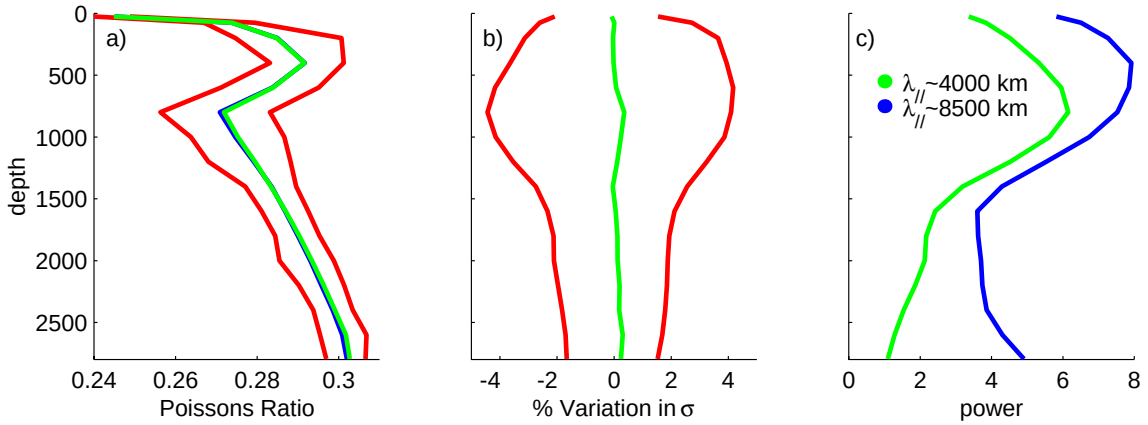


Figure 8. a) Poisson's ratio as a function of depth for reference model (blue), mean of each layer (green) and the range inferred from our data (red). b) Same as in a) except shown as percent variations from ak135 reference model. (c) Power of the two dominant wavelength structures found in the corridor. The shorter wavelength structures (green line) have similar power to the longer wavelength structure (blue line) down to about 1500 km depth. Below that depth, the shorter wavelength structures start to diminish and the longer wavelength structures become more dominant. At the base of the mantle, the structure is entirely long wavelength in nature.

a function of depth. At shallow depths, both wavelengths are similarly represented in our images, but below ~1500 km depth, the dominant wavelength of the imaged structures shifts to longer wavelength, with very small power in the short-wavelength structures. This change in dominant wavelength can also be seen directly in the depth slices themselves (Figure 6). These wavelengths are consistent with the size of domes (2000-4000 km) observed in laboratory tank experiments of thermo-chemical convection in a fluid with stratified density and viscosity [*Lebars and Davaille*, 2001].

5 Dependence of Poisson's Ratio on temperature and composition

Our seismic images of Poisson's ratio along the Alaska corridor provide information on the physical state of the mantle. From the seismic inversion we infer that the magnitude of the lateral variation in Poisson's ratio is ~8% at 1000 km depth, 4% at 1500 km depth and ~3% just above the core-mantle boundary.

Variability in Poisson's ratio can be caused by temperature or compositional differences. We discuss below how each of those factors might contribute to the variability we image. Since the composition of the lower mantle is still a subject of debate, we explore two simple possibilities. The first is that the amount of perovskite ($MgSiO_3$) in the perovskite-magnesiowüstite assemblage is variable in the lower mantle, perhaps comprising 95% in some regions and just 80% in others. It has been argued that subducted slabs are depleted in perovskite [Ringwood, 1991] relative to a convecting pyrolitic mantle, which might give rise to such differences. The other effect we explore is the addition of iron to certain regions of the lower mantle.

5.1 Effect of Temperature

We have calculated $\frac{\partial V_p}{\partial T}$ and $\frac{\partial V_s}{\partial T}$ at various depths using the quantum mechanics calculations of Karki & Stixrude [1999] for the elastic properties of the major minerals considered likely constituents of the lower mantle and assuming a lower mantle comprised of 90% perovskite, 10% magnesiowüstite and X_{Fe}=0.1. With these derivatives, we can compute the effect that temperature perturbations would have on Poisson's ratio. Figure 9 shows the results of these calculations for a 300° thermal anomaly as a function of pressure (depth). For comparison, we have also plotted the percent change in Poisson's ratio predicted by several other independent calculations of lower mantle P and S wave temperature-velocity derivatives [*Karato*, 1993; *Stacey*, 1998; *Trampert et al.*, 2001], all of which were derived by extrapolating mineral physics data from relatively shallow pressures to lower mantle conditions. The extrapolations were done using various 3rd order Birch-Murnaghan equations of state with varying assumptions (i.e. quasi-harmonic approximations vs. including higher order anharmonic terms vs. including anelastic effects) and all assumed a simplified model of the lower mantle consisting of perovskite and magnesiowüstite only. In addition, only the *Karato* [1993] derivatives include an estimate for the effects of anelasticity, the effect of which may be substantial, but is also relatively unknown. At pressures representative of the core-mantle boundary (130 GPa), a given temperature difference has less of an effect on Poisson's ratio than at

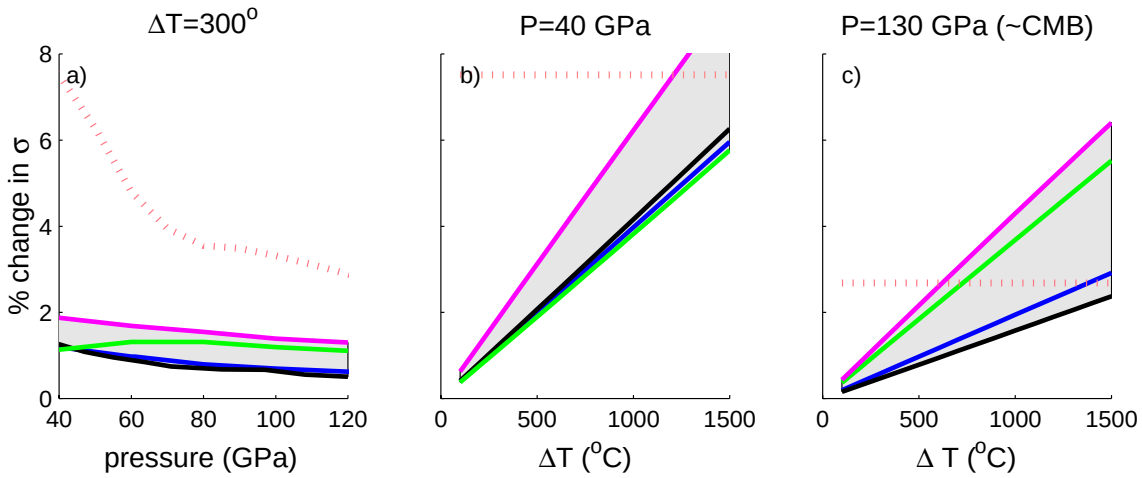


Figure 9. a) Effect of a 300° thermal anomaly on Poisson's ratio at various presures. Poisson's ratio is computed with equation (5) and the temperature-velocity derivatives of *Stacey* [1998] (pink), *Karato* [1993] (green), *Trampert et al.* [2001] (black) and *Karki and Stixrude* [1997] (blue). For cmparison, the dotted orange line shows Poisson's ratio inferred from the tomographic images. Also shown is the effect of various thermal anomalies on Poisson's ratio at a constant pressure of b) 40 GPa (~1000 km depth) and c) 130 GPa (the core-mantle boundary) using the same derivatives as in a).

shallower depths (Figures 9b, c); however, the magnitude of the effect varies by a factor of three depending upon which derivative you choose. A temperature difference of 300° can explain 1-2% Poisson's ratio variability at 1000 km depths, and 0.5-1.5% at CMB depths. If one were to explain the entire peak-to-peak variability we image, temperature differences of at least 1300° at 1000 km depth and of 700° at the CMB would be required.

5.2 Effect of partitioning between perovskite and magnesiowüstite

One of the problems in investigating what happens when the amount of perovskite in a magnesiowüstite-perovskite assemblage varies is that the relative velocities are affected by the presence of iron. Not only is the total amount of iron in the lower mantle debated but results from experimental studies of iron partitioning between the two minerals do not agree. Many studies have shown that iron tends to concentrate in the magnesiowüstite [*Yagi et al.*, 1979, *Ito et al.*, 1984, *Guyot et al.*, 1988; *Ito & Takahashi*, 1989; *Mao et al.*, 1997] but more recent studies have suggested that the partitioning may be more equal [*Wood & Rubie*, 1996; *Kesson et al.*, 1998], and with the inclusion of Aluminum in the mix the iron may actually concentrate entirely in the perovskite instead [*McCammon*, 1997].

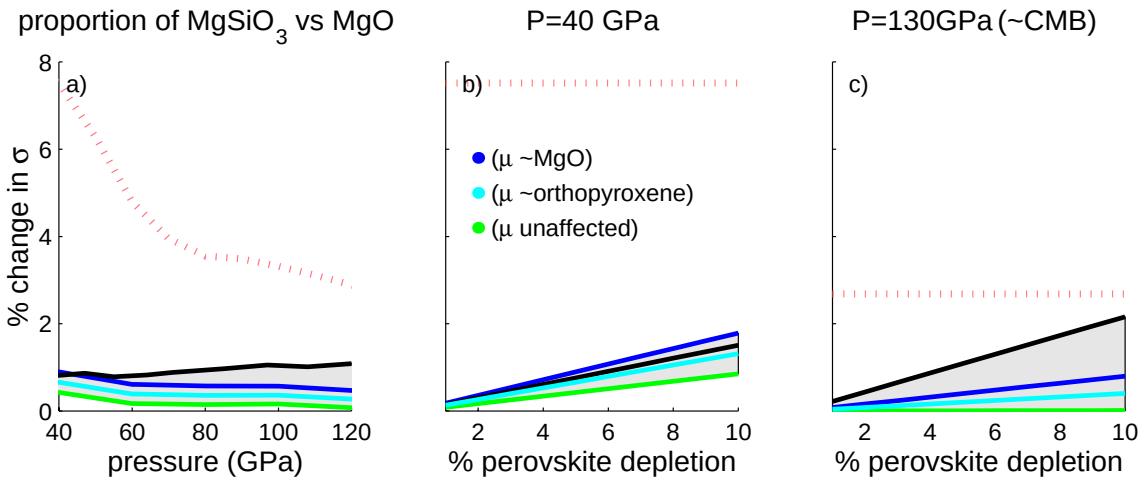


Figure 10. a) Effect of a 5% change in the proportion of $MgSiO_3$ in an MgO-$MgSiO_3$ assemblage on Poisson's ratio. Poisson's ratio is computed with equation (5) and the compositional-velocity derivatives of *Trampert et al.* [2001] (solid black line) and for several different behavioral effects of the shear modulus with the *Karki & Stixrude* [1997] derivatives since the behavior of the shear modulus in perovskite is not constrained experimentally. First, we allow μ to remain unchanged, so that the seismic velocities are affected only by the density change associated with iron (green), second we change μ in a manner similar to what has been observed experimentally for orthopyroxene (light blue) and third we change μ in a manner similar to what has been observed experimentally for MgO (dark blue). Dotted orange line shows range of inferred Poisson's ratio from our tomographic images. b) Effect of different amounts of perovskite depletion at depths of 1000 km and c) at the core-mantle boundary for the same examples as in a).

Bearing these experimental uncertainties in mind, Figure 10 shows the effect of the relative amount of perovskite using the *Karki & Stirxrude* [1999] and the *Trampert et al.* [2001] formulation. The two differ by a factor of two on the magnitude of the effect. According to *Karki & Stixrude* [1999], a change in the proportion of perovskite does not really affect Poisson's ratio in any significant manner unless (1) there is a significant amount of iron in the lower mantle (~10%) (2) all of that iron partitions exclusively into either the magnesiowüstite or the perovskite minerals, (3) the shear modulus is strongly affected by the addition of iron and (4) the bulk modulus is unaffected by the presence of iron. If those criteria are met, the *Karki and Stixrude* [1997] derivatives show that a 10% difference in the volume proportion of perovskite would give rise to ~1.75% variability in Poisson's ratio at ~1000 km depth (40 GPa) and ~0.5% variability at the base of the mantle (Figure 9). Using the *Trampert et al.* [2001] derivatives, that same variability in Poisson's ratio can be explained by less than half the

difference in volume proportion of perovskite; however, their compositional derivatives do not come with a clear statement of how iron was partitioned between the two minerals nor what assumptions were made concerning the effect of iron on the shear and bulk moduli, which makes it difficult to assess the differences between these two sets of predictions. Recognizing the inherent uncertainties in both methods, these derivatives show that 10% perovskite depletion can explain ~20% of the variability in Poisson's ratio that we see at 1000 km depth and between 40-80% of the variability at CMB depths.

5.3 Effect of iron content

Figure 11 shows the effect of iron depletion (or concentration) for both perovskite and magnesiowüstite. One problem in doing this calculation is that its effect on the elastic moduli is not well known. Experiments have shown that the effect of iron on the bulk modulus is negligible [*Mao et al*, 1991; *Fei et al.*, 1992] so we do not vary that parameter for either mineral. For the shear modulus, experimental results show that magnesiowüstite is weakened according to

$$\mu = \mu_0(1 - 0.59X_{Fe}) \tag{6}$$

where $X_{Fe} = Fe/(Fe + Mg)$ is the Fe-number [*Duffy & Anderson*, 1989] and μ_0 is the shear modulus with $X_{Fe} = 0$. There are no similar constraints regarding the effects of iron on the shear modulus for perovskite so we consider several possibilities. One possibility is that there is no effect at all, so the only change to the velocities results from the density change associated with changes in iron, which we calculate following *Jeanloz & Thompson* [1983]. Another is to assume that the effect of iron on the shear modulus is similar to that observed in orthopyroxene [*Duffy & Anderson*, 1989], and a third is that the effect of iron is the same as that observed for magnesiowüstite [*Duffy & Anderson*, 1989].

Using the quantum mechanics calculations of *Karki & Stixrude* [1999], accounting for mantle temperatures and then varying the shear modulus in the three different manners described, we find that if the shear-modulus in perovskite behaves similarly to what has been observed for magnesiowüstite, then 1% variability in Poisson's ratio can be explained by 2% variation in total iron content. In order to explain the observed peak-to-peak variability in

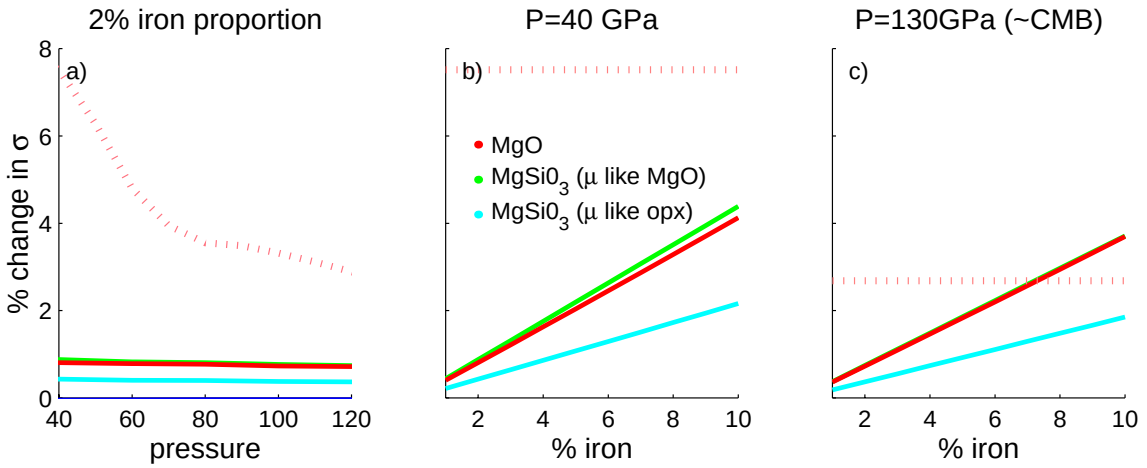


Figure 11. Effect of 2% change in the iron content on Poisson's ratio. The effect of the shear modulus is not constrained experimentally in perovskite and so we try several variations, assuming the bulk modulus is unaffected. If the shear modulus is strongly affected (such as has been observed in MgO) then the elastic equations of *Karki & Stixrude* [1997] predict that the effect of iron on perovskite (green) and on MgO (red) velocities can be significant. If the shear modulus in perovskite is similarly to orthopyroxene (light blue) then the effect is less significant. Also shown is the effect of different amounts of iron enrichment/depletion at pressures of b) 40 GPa (~1000 km depth) and c) 130 GPa (~CMB pressures) for the same set of cases. For comparison, the dotted orange line shows the range in Poisson's ratio inferred from our tomographic images.

Poisson's ratio, more than 8% iron variation is required at all depths. If the observed range in σ were to be explained entirely by variations in Fe content alone, a change of ΔX_{Fe}=16% would be required at 1000 km depth and ΔX_{Fe}=8% near the core-mantle boundary.

6 Discussion

The results for shear wave speeds obtained here are generally in good agreement with the recent P-wave model of *Kárason and van der Hilst* [in prep. 2001] and consistent with *Widiyantoro et al.*, [1999]. What we interpret as slab fragments produce strong variations in shear wavespeed but have only a very weak bulk-sound signature. In the mid mantle, the magnitude of the lateral variation in Poisson's ratio is large (~8%) and the dominant length scales over which these variations occur are ~2000 km and ~4000 km. In the bottom 1000 km of the mantle beneath Alaska and North America where the laterally coherent, slab-like structures are absent, the magnitude of the lateral variations in

Poisson's ratio decrease to 3% and the dominant length scale over which variations take place is ~4000 km. Interestingly, a recent study by *Lebars and Davaille* [in prep. 2001] of the dynamics of compositional stratification in the deep mantle suggests that oscillations of compositionally distinct domains occur on spatial wave length scales that are very similar to those inferred from our maps. Because of the limited size of our study region and the practical difficulty of accurately mimicking an earthlike parameter range in the analog experiments, a discussion of the causal relationship between these two is premature. However an intriguing scenario could be considered in which downwelling slabs interact with compositionally distinct domains whose vertical oscillations organize slab fragmentation on the wavelengths inferred from our study.

At 1000 km depth, the observed variability in Poisson's ratio is ~8%. From the previous discussion we see that temperature differences of 800° can explain 3%-5% of the variability and then differences in perovskite depletion can explain another 2% if the *Karki & Stixrude*, [1997] derivatives are used. An additional 2% of the anomaly can be explained by 2% iron depletion.

The peak-to-peak variability in Poisson's ratio at the CMB is ~3%. In order to explain this variability by temperature alone, temperature differences between 700°C and 1500°C are required, depending on which velocity-temperature derivatives are used [*Karato*, 1993; *Stacey*, 1998; *Karki & Stixrude* , 1997; *Trampert et al.*, 2001], which seems unlikely, in particular, if the colder, more rigid regions correspond to ancient subducted slabs. A finite difference calculation (solving the heatflow equation) of the thermal differences between a 120 million year-old subducted slab and the surrounding, convecting mantle shows that these differences should be of the order of 300°C (see Table 1 for values used in this calculation.)

Assuming that the velocity-temperature derivatives are anywhere in the ballpark, even a 500° temperature anomaly between the old subducted slabs and the surrounding mantle, would leave at least half (~1.5%) if not 5/6 of the variability in Poisson's ratio unexplained. Therefore, it seems likely that we are seeing some sort of compositional effect at the base of the mantle.

Table 1. Values used in slab temperature calculation

	Length of Time spent	ρ (kg/m^3)	Ambient Mantle Temp (°C)	κ (W/mK)
Upper mantle	17 Myr	3400	1350°	4.5
Top of lower mantle	40 Myr	4000	1800°	7.0
Bottom of lower mantle	40 Myr	5500	2700°	9.6
D″	20 myr	5600	3450°	10.0

In the diffusion equation we use:
Cp=1250 (J/Kg K) at all depths
Slab thickness = 100 km
Temperature profile decreases linearly from 200° to 1200°C

Ringwood [1991] argued that slabs are composed primarily of depleted harzburgite as a result of the large extents of melting they have undergone in formation. Since garnet and pyroxenes transform exclusively to perovskite, one consequence of this idea is that subducted slabs having less of these minerals may end up depleted in perovskite relative to the convecting mantle. The *Trampert et al*, [2001] compositional derivatives suggest that 5% perovskite depletion could cause a 1% change in Poisson's ratio although *Karki & Stixrude* [1997] would predict a less dramatic effect. The *Karki & Stixrude* [1999] iron-depletion derivatives show that the remaining 1% variability we see in Poisson's ratio can be explained by just 2% variation in total iron content (slabs being 2% depleted or the region in the center of our image being 2% enriched in iron), if the shear-modulus in perovskite behaves similarly to what has been observed for magnesiowüstite. Therefore, the observed variability in Poisson's ratio of 3% can be explained by a combination of a 300° temperature perturbation, 5% perovskite depletion and 2% iron variability.

Despite the many questions regarding the magnitude of the effect of iron, it seems clear that none of the published temperature-velocity derivatives are able to fully explain the variability in Poisson's ratio that we see in the lowermost mantle. Variability in the depletion of iron and/or depletion of perovskite seem to be possible explanations that are compatible with the limited set of compositional-velocity derivatives currently at our disposal. As additional high-

pressure, high-temperature laboratory results become available, it should become possible to answer this question more strongly.

7 Conclusions

From waveform cross-correlation of broadband data, we have obtained high-resolution tomographic images of the shear wave and bulk-sound velocities along a great-circle transect between Japan, Alaska, and the western US. We infer lateral variability in Poisson's ratio of 8% at depths of 1000 km, 4% at depths of 1500 km and 3% at the CMB, which we have interpreted in terms of temperature and compositional effects (both Si and Fe).

The entire transect lies along continental or island arc plate boundaries and so we consider the effect of subducted slab fragments on the observed variability in Poisson's ratio. If slab fragments are carried all the way down to the core mantle boundary, then the associated temperature anomaly should be ~300°. However, the variability in Poisson's ratio that we see at the bottom of the mantle (~3%) would imply temperature variations of 700°-1500° (if they were purely thermal in origin), assuming that the four different estimates of $\frac{\partial V_p}{\partial T}$ and $\frac{\partial V_s}{\partial T}$ we have used [*Karato,* 1993; *Karki & Stixrude,* 1997; *Stacey,* 1998; *Trampert et al.,* 2001] are realistic. A simple explanation is that the variability includes a contribution from compositional effects. One effect that could explain the difference is if the subducted slabs are 10% depleted in perovskite relative to the convecting mantle as suggested by *Ringwood* [1991]. Alternatively, variability in iron content from one region to another by 3% could equally well explain the differences. Combinations of the two in lesser amounts also work equally well (eg. 5% perovskite depletion with 2% iron depletion). At depths of 1000 km we find a variation in Poisson's ratio of 8%, which can be explained by an 800° thermal perturbation in addition to 5% perovskite depletion and 2% iron variation.

Acknowledgements

We thank Hrafnkell Kárason who generously provided both raytracing and inversion code. This research was supported by the National Science Foundation

under grant EAR-9905779, the David and Lucille Packard Foundation, and the Institut de Physique du Globe de Paris.

Appendix A: Propagator Matrices and Splitting Operators

The system matrix $\boldsymbol{\mathcal{A}}(z)$ is rotated into the eigenwave coordinate system by the 4×4 block-diagonal matrix $\boldsymbol{\mathcal{U}} = \text{diag}[\mathbf{U}, \mathbf{U}]$,

$$\hat{\boldsymbol{\mathcal{A}}} = \boldsymbol{\mathcal{U}} \boldsymbol{\mathcal{A}} \boldsymbol{\mathcal{U}}^{\mathrm{T}} = \begin{bmatrix} \mathbf{0} & \mathbf{C}^{-1} \\ -\omega^2 \mathbf{I} & \mathbf{0} \end{bmatrix}, \tag{A1}$$

and is diagonalized by a matrix whose columns represent the downgoing and upgoing eigenwaves: $\boldsymbol{\mathcal{D}}^{-1} \hat{\boldsymbol{\mathcal{A}}} \boldsymbol{\mathcal{D}} = \mathrm{i}\,\text{diag}[k_1, k_2, -k_1, -k_2]$. Here, $k_j = \omega / \mathrm{v}_j$ (i = 1, 2) are the eigenwavenumbers, and

$$\boldsymbol{\mathcal{D}} = \begin{bmatrix} \mathbf{D}_u & \mathbf{D}_u \\ \mathbf{D}_\tau & \mathbf{D}_\tau^* \end{bmatrix}, \quad \boldsymbol{\mathcal{D}}^{-1} = \begin{bmatrix} \mathbf{D}_u^{-1} & \mathbf{D}_\tau^{-1} \\ \mathbf{D}_u^{-1} & \mathbf{D}_\tau^{*-1} \end{bmatrix}, \tag{A2}$$

where the 2×2 blocks are

$$\mathbf{D}_u = \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \varepsilon_2 \end{bmatrix}, \quad \mathbf{D}_\tau = \begin{bmatrix} \mathrm{i}\omega \mathrm{v}_1 \varepsilon_1 & 0 \\ 0 & \mathrm{i}\omega \mathrm{v}_2 \varepsilon_2 \end{bmatrix}. \tag{A3}$$

Kennett's (1983, eqn. 2.63) energy normalization procedure yields $\varepsilon_j = (2\mathrm{v}_j)^{-1/2}$.

In a homogeneous layer ($\phi(z)$ = constant), the eigenwave propagator from z_0 to z is

$$\boldsymbol{\mathcal{E}}(z, z_0) = \begin{bmatrix} \mathbf{E}(z, z_0) & \mathbf{0} \\ \mathbf{0} & \mathbf{E}^*(z, z_0) \end{bmatrix}, \tag{A4}$$

where (*) indicates complex conjugation and

$$\mathbf{E}(z, z_0) = \begin{bmatrix} \mathrm{e}^{\mathrm{i} k_1 (z - z_0)} & 0 \\ 0 & \mathrm{e}^{\mathrm{i} k_2 (z - z_0)} \end{bmatrix}. \tag{A5}$$

Therefore, the displacement-stress propagator in the eigenwave coordinate system is $\hat{\boldsymbol{\mathcal{P}}}^{\mathrm{h}}(z,z_0) = \boldsymbol{\mathcal{D}}\,\boldsymbol{\mathcal{E}}(z,z_0)\,\boldsymbol{\mathcal{D}}^{-1}$, and the general homogeneous-layer propagator matrix can be written as

$$\boldsymbol{\mathcal{P}}^{\mathrm{h}}(z,z_0) \equiv \begin{bmatrix} \mathbf{P}^{\mathrm{h}}_{uu}(z,z_0) & \mathbf{P}^{\mathrm{h}}_{u\tau}(z,z_0) \\ \mathbf{P}^{\mathrm{h}}_{\tau u}(z,z_0) & \mathbf{P}^{\mathrm{h}}_{\tau\tau}(z,z_0) \end{bmatrix} = \boldsymbol{\mathcal{U}}(\phi)\,\hat{\boldsymbol{\mathcal{P}}}^{\mathrm{h}}(z,z_0)\,\boldsymbol{\mathcal{U}}^{\mathsf{T}}(\phi). \tag{A6}$$

The propagator for the anisotropic, heterogeneous interval $0 \le z \le d$ can be constructed by approximating the heterogeneity as a stack of homogeneous layers and multiplying propagators in the form of (A6). Better analytical insight into the effects of the heterogeneity can be obtained by proceeding in a different fashion, however. We make the ansatz (cf. Kennett, 1983, p. 53),

$$\boldsymbol{\mathcal{P}}(z,z') = \boldsymbol{\mathcal{U}}(\phi(z))\,\boldsymbol{\mathcal{D}}\,\boldsymbol{\mathcal{E}}(z,z_0)\,\mathbf{Q}(z,z_0)\,\boldsymbol{\mathcal{D}}^{-1}\,\boldsymbol{\mathcal{U}}^{\mathsf{T}}(\phi(z_0)), \tag{A7}$$

and find that (A7) obeys the system equation $\partial_z \boldsymbol{\mathcal{P}} = \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}}$ if and only if Q satisfies

$$\partial_z \mathbf{Q}(z,z_0) = \boldsymbol{S}(z)\mathbf{Q}(z,z_0) \tag{A8a}$$

$$\mathbf{Q}(z_0,z_0) = \mathbf{I} \tag{A8b}$$

where

$$\boldsymbol{S}(z) = \begin{bmatrix} 0 & a(z) & 0 & b^*(z) \\ -a^*(z) & 0 & b^*(z) & 0 \\ 0 & b(z) & 0 & a^*(z) \\ b(z) & 0 & -a(z) & 0 \end{bmatrix} \frac{d\phi}{dz}. \tag{A9}$$

The depth dependent coefficients a and b depend exponentially on the wavenumber difference $\Delta k = k_2 - k_1$ and the wavenumber average $\bar{k} = (k_1 + k_2)/2$, respectively:

$$a(z) = \frac{\bar{\mathrm{v}}}{\sqrt{\mathrm{v}_1\mathrm{v}_2}} \mathrm{e}^{\mathrm{i}\Delta k z}, \tag{A10a}$$

$$b(z) = \frac{\Delta v}{2\sqrt{v_1 v_2}} e^{2i\bar{k}z}. \tag{A10b}$$

The propagator $\mathbf{Q}(z, z_0)$ is a 4×4 differential scattering matrix for the eigenwaves, which we write in block form as

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{++} & \mathbf{Q}_{+-} \\ \mathbf{Q}_{-+} & \mathbf{Q}_{--} \end{bmatrix}. \tag{A11}$$

The submatrices $\mathbf{Q}_{++}$ and $\mathbf{Q}_{--}$ describe, respectively, the forward scattering of downgoing (+) and upgoing (–) eigenwaves by gradients in $\phi(z)$, and $\mathbf{Q}_{-+}$ and $\mathbf{Q}_{+-}$ describe the corresponding backward scattering. These scattering operators satisfy the reciprocal relations

$$\mathbf{Q}_{--} = \mathbf{Q}_{++}^{*}, \tag{A12a}$$

$$\mathbf{Q}_{+-} = \mathbf{Q}_{-+}^{*}. \tag{A12b}$$

From eqns. (A8)-(A10) we can see that the strength of forward-scattering in a depth increment dz is proportional to $a_0 e^{i\Delta k z} d\phi(z)$, while the strength of the back-scattering goes like $b_0 e^{2i\bar{k}z} d\phi(z)$. If the relative difference in the eigenvelocities is small, then forward-scattering will tend to dominate because $b_0 << a_0 \sim 1$. Moreover, in this situation of small anisotropy, $\Delta k << 2\bar{k}$, so that the back-scattering kernel will be more oscillatory and its integral contributions will tend to cancel if $\phi(z)$ is smooth.

At zero frequency, $\boldsymbol{S} = \boldsymbol{S}_0$ depends on z only through $\dot{\phi} \equiv d\phi/dz$. Therefore, $\boldsymbol{S}_0$ commutes with its integral, and the solution to (11) is $\mathbf{Q}(z, z_0) = \exp(\boldsymbol{S}\Delta\phi)$, where $\Delta\phi = \phi(z) - \phi(z_0)$. Using the fact that $\boldsymbol{S}_0^{2n} = (-1)^n \Delta\phi^{2n} \mathbf{I}$ and $\boldsymbol{S}_0^{2n+1} = (-1)^n \Delta\phi^{2n} \mathbf{S}_0$, we can sum the exponential series. This yields a good approximation to the propagator across a layer that is thin compared to a wavelength; i.e., for $\bar{k}\Delta z << 1$,

$$\mathbf{Q}_{++}(z+\Delta z,z) \approx \begin{bmatrix} \cos\Delta\phi & a(z)\sin\Delta\phi \\ -a^*(z)\sin\Delta\phi & \cos\Delta\phi \end{bmatrix}. \tag{A13a}$$

$$\mathbf{Q}_{-+}(z+\Delta z,z) \approx \begin{bmatrix} 0 & b(z)\sin\Delta\phi \\ b(z)\sin\Delta\phi & 0 \end{bmatrix} \tag{A13b}$$

We note that this approximation is independent of the form of $\phi(z)$ and, for example, does not require $\phi(z)$ to be a smooth function of depth. Indeed, it provides the generalization of the propagator to self-affine (fractal) media for which $\dot{\phi}$ may not be well defined. Equations (A7) and (A13) are the basis of our computational algorithm.

For an upgoing wave $\mathbf{u}_I(z) \sim \mathrm{e}^{-\mathrm{i}\bar{k}z}$ incident at the base of the anisotropic layer, the displacement-stress vector in the half-space can be expressed as

$$f'(d) = \begin{bmatrix} (\mathbf{I}+\mathbf{R}) \\ \mathrm{i}\omega\bar{\nu}(\mathbf{I}-\mathbf{R}) \end{bmatrix} u_I(d), \tag{A14}$$

where $\mathbf{R}$ is a 2×2 matrix of reflection coefficients. Satisfying the zero-traction boundary conditions at the surface yields

$$\mathbf{R} = \mathbf{U}_d\,(i\omega\bar{\nu}\hat{\mathbf{P}}_{\tau\tau} - \hat{\mathbf{P}}_{\tau u})^{-1}(i\omega\bar{\nu}\hat{\mathbf{P}}_{\tau\tau} + \hat{\mathbf{P}}_{\tau u})\mathbf{U}_d^{\mathsf{T}}. \tag{A15}$$

From here on, $\mathbf{U}_z = \mathbf{U}(\phi(z))$, and it is understood that, unless otherwise specified, the propagators are taken from the base of the anisotropic layer to the surface; e.g., $\hat{\mathbf{P}}_{uu} \equiv \mathbf{U}_0^{\mathsf{T}}\mathbf{P}_{uu}(0,d)\mathbf{U}_d$. In this notation, the free-surface displacement vector is

$$\mathbf{u}(0) = \mathbf{U}_0[\hat{\mathbf{P}}_{uu} + i\omega\bar{\nu}\hat{\mathbf{P}}_{u\tau} + (\hat{\mathbf{P}}_{uu} - i\omega\bar{\nu}\hat{\mathbf{P}}_{u\tau})\hat{\mathbf{R}}]\mathbf{U}_d^{\mathsf{T}}\,\mathbf{u}_I'(d) \tag{A16}$$

The symmetries in (A2), (A4) and (A12) can be used to express the propagator submatrices as the following (real-valued) expressions:

$$\hat{\mathbf{P}}_{uu} = \mathbf{D}_u\,\mathrm{Re}[\mathbf{E}(\mathbf{Q}_{++} + \mathbf{Q}_{+-})]\mathbf{D}_u^{-1} \tag{A17a}$$

$$\hat{\mathbf{P}}_{u\tau} = \mathrm{i}\mathbf{D}_u \,\mathrm{Im}[\mathbf{E}(\mathbf{Q}_{++} + \mathbf{Q}_{+-})]\mathbf{D}_\tau^{-1} \tag{A17b}$$

$$\hat{\mathbf{P}}_{\tau u} = \mathrm{i}\mathbf{D}_\tau \,\mathrm{Im}[\mathbf{E}(\mathbf{Q}_{++} + \mathbf{Q}_{+-})]\mathbf{D}_u^{-1} \tag{A17c}$$

$$\hat{\mathbf{P}}_{\tau\tau} = \mathbf{D}_\tau \,\mathrm{Re}[\mathbf{E}(\mathbf{Q}_{++} + \mathbf{Q}_{+-})]\mathbf{D}_\tau^{-1} \tag{A17d}$$

In the case of a homogeneous layer, these expressions reduce to $\hat{\mathbf{P}}^{\mathrm{h}}_{uu} = \hat{\mathbf{P}}^{\mathrm{h}}_{\tau\tau} = \mathrm{diag}[\cos k_j(z - z_0)]$, $\hat{\mathbf{P}}^{\mathrm{h}}_{u\tau} = \mathrm{diag}[(\omega \mathrm{v}_j)^{-1} \sin k_j d]$, and $\hat{\mathbf{P}}^{\mathrm{h}}_{\tau u} = \mathrm{diag}[-(\omega \mathrm{v}_j)\sin k_j d]$.

A useful approximation, almost always employed in vertical shear-wave splitting analysis, is to ignore back-scattering and reverberations within the anisotropic layer. This amounts to ignoring terms of order $\Delta \mathrm{v} / \bar{\mathrm{v}}$ in eqn. (A17). Under this approximation, $a_0 = 1$ and $b_0 = 0$, so that $\mathbf{Q}_{+-} = \mathbf{Q}_{-+} = 0$. A little algebra obtains

$$\left.\begin{aligned} \mathbf{R} &= \mathbf{U}_d \,\mathbf{Q}_{++}^{\mathsf{T}} \mathbf{E}^2 \,\mathbf{Q}_{++}\, \mathbf{U}_d^{\mathsf{T}} \\ \mathbf{u}(0) &= 2\,\mathbf{U}_0\, \mathbf{E}\, \mathbf{Q}_{++}\, \mathbf{U}_d^{\mathsf{T}}\, \mathbf{u}_I'(d) \end{aligned}\right\} \text{ (no back-scattering).} \tag{A18a, A18b}$$

Eqn. (A18b) shows that, when back-scattering can be ignored, the eigenwave propagator is just $\mathbf{E}\mathbf{Q}_{++}$. It will be convenient to pull out the phase factor corresponding to the mean travel time through the layer, $\bar{t} = \bar{k} d / \omega$, and rewrite these expressions in terms of the eigenwave splitting matrix,

$$\mathbf{H} = \mathrm{e}^{-\mathrm{i}\omega\bar{t}}\,\mathbf{E} = \begin{bmatrix} \mathrm{e}^{-\mathrm{i}\omega\Delta t/2} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\omega\Delta t/2} \end{bmatrix}, \tag{A19}$$

which is unimodular; i.e., $\det[\mathbf{H}] = 1$. We define the splitting operator,

$$\Gamma = \mathbf{U}_0\, \mathbf{H}\, \mathbf{Q}_{++}\, \mathbf{U}_d^{\mathsf{T}}. \tag{A20}$$

The surface displacement is thus $\mathbf{u}(0) = 2\,\mathrm{e}^{\mathrm{i}\omega\bar{t}}\Gamma\,\mathbf{u}_I'(d)$, and the reflection matrix is $\mathbf{R} = \mathrm{e}^{2\mathrm{i}\omega\bar{t}}\Gamma^{\mathsf{T}}\Gamma$. The factor of two in the former comes from the constructive interference of the upward-going wave and its surface reflection. In the case of a homogeneous layer, $\mathbf{Q}_{++} = \mathbf{I}$, and (A20) becomes

$$\Gamma_h(\phi, \Delta t) = \mathbf{U}(\phi)\, \mathbf{H}(\Delta t)\, \mathbf{U}^{\mathsf{T}}(\phi) \quad \text{(homogeneous layer).} \tag{A21}$$

All of the matrices in (A20) are both unitary and unimodular; e.g., $\Gamma^{-1} = \Gamma^{\dagger} \equiv (\Gamma^{*})^{\mathsf{T}}$, $\det[\Gamma] = 1$. (The unimodularity of $\mathbf{Q}_{++}$ follows from $\mathrm{tr}[\boldsymbol{S}_{++}(z)] = 0$; see Kennett, 1983, p. 42.) Therefore, all of the matrix operations associated with forward scattering belong to the group SU(2). This symmetry can be used to simplify the analysis. Any member of this group can be written in it terms of two complex numbers,

$$\begin{bmatrix} \alpha & \beta \\ -\beta^{*} & \alpha^{*} \end{bmatrix}, \quad \text{where } |\alpha|^2 + |\beta|^2 = 1. \tag{A22}$$

An SU(2) matrix thus depends on three real parameters and can be written in the following general forms (Varshalovich et al., 1988):

$$\Gamma = \begin{bmatrix} e^{-i(\alpha+\gamma)/2} \cos\beta/2 & e^{i(\alpha+\gamma)/2} \sin\beta/2 \\ e^{-i(\alpha+\gamma)/2} \sin\beta/2 & e^{i(\alpha+\gamma)/2} \cos\beta/2 \end{bmatrix}$$

$$= \begin{bmatrix} \cos\Omega/2 - i\cos\Theta\sin\Omega/2 & -i e^{i\Phi} \sin\Theta\sin\Omega/2 \\ -i e^{-i\Phi} \sin\Theta\sin\Omega/2 & \cos\Omega/2 + i\cos\Theta\sin\Omega/2 \end{bmatrix}. \tag{A23}$$

SU(2) is homomorphic (with a two-fold ambiguity) to $O^{+}(3)$, the group of proper orthogonal transformations in 3-space, which allows the splitting operations to be visualized as 3D rotations. In the first form in (A23), the parameters (α, β, γ) correspond to the Euler angles of the 3D rotation; in the second, the 3D rotation is through an angle _ about an axis with polar coordinates (Θ, _). Equation (A21) can be recast as

$$\Gamma_h = \begin{bmatrix} \cos\omega\Delta t/2 - i\cos 2\phi \sin\omega\Delta t/2 & -i\sin 2\phi \sin\omega\Delta t/2 \\ -i\sin 2\phi \sin\omega\Delta t/2 & \cos\omega\Delta t/2 + i\cos 2\phi \sin\omega\Delta t/2 \end{bmatrix} \tag{A24}$$

The splitting matrix for a homogeneous layer thus corresponds to a 3D rotation through an angle $\omega\Delta t$ about an axis located at co-latitude 2ϕ and zero longitude.

Appendix B: Fréchet Kernels for Homogeneous Starting Model

To calculate the Fréchet kernels defined by (8) and (9), we perturb the splitting orientation function $\phi(z)$ by a small constant amount $\delta\phi$ in a thin layer of thickness δz at a depth $0 < z < d$ and compute the corresponding perturbations $\delta\tilde{\phi}(z)$ and $\delta\tilde{t}(z)$. The kernels $G_\phi(z)$ and $G_t(z)$ are then given as the limiting values of the ratios $\delta\tilde{\phi}(z)/\delta\phi\delta z$ and $\delta\tilde{t}(z)/\delta\phi\delta z$, respectively. This calculation can be done numerically for arbitrary starting models and pulse shapes (e.g., Fig. 7). Here we present an analytical derivation for narrow-band pulses in the special case of a homogeneous starting model.

Because the perturbations are small, the forward-scattering approximation applies. Vertical propagation through a homogeneous anisotropic layer with a fast-axis orientation ϕ_0 and splitting time Δt yields the Fourier-transformed vertical displacement $\mathbf{u}(0,\omega) = 2e^{i\omega\bar{t}}\Gamma_h(\phi_0,\Delta t)\mathbf{u}_I(\omega)$, where the homogeneous-layer splitting operator Γ_h is given by eqn. (A24). We assume the incident pulse is radially polarized, $\mathbf{u}_I(\omega) = u_I(\omega)\hat{\mathbf{x}} = [u_I(\omega)\ 0]^T$, and we approximate its energy spectrum by a Gaussian:

$$|u_I(\omega)|^2 = \frac{1}{\sigma\sqrt{8\pi}}e^{-(\omega-\omega_0)^2/2\sigma^2} + \frac{1}{\sigma\sqrt{8\pi}}e^{-(\omega+\omega_0)^2/2\sigma^2}. \tag{B1}$$

This spectrum has peaks of half-bandwidth σ centered at frequencies of $\pm\omega_0$, and it is normalized to have unit total energy: $\int_{-\infty}^{\infty}|u_I(\omega)|^2\,d\omega = 1$. If pulse is narrow-band in the sense that $\sigma << \omega_0$, then the integral of its energy spectrum against any reasonably smooth function can be approximated by integrating a truncated Taylor expansion of the function about the center frequency ω_0:

$$\begin{aligned} 2\int_0^\infty f(\omega)|u_I(\omega)|^2\,d\omega &\approx 2\int_0^\infty [f(\omega_0)+\omega\dot{f}(\omega_0)+\omega^2\ddot{f}(\omega_0)]|u_I(\omega)|^2\,d\omega \\ &\approx f(\omega_0) + \sigma^2\ddot{f}(\omega_0). \end{aligned} \tag{B2}$$

The terms dropped in this approximation are of order $(\sigma/\omega_0)^4$.

Since the perturbed model is homogeneous above z and below $z + \delta z$, its splitting operator can be written in the form

$$\Gamma = \mathbf{U}(\phi_0)\mathbf{H}(\Delta t)\mathbf{Q}_{++}\,\mathbf{U}^{\mathsf{T}}(\phi_0), \tag{B3}$$

where **H** is given by (A19). For a constant perturbation $\delta\phi$ in a layer (z, $z + \delta z$), **Q** satisfies (A8) with $a_0 = 1$, $b_0 = 0$, and

$$\dot{\phi}(\zeta) = \delta\phi[\delta(\zeta - z) - \delta(\zeta - z - \delta z)]. \tag{B4}$$

Integrating up from the base of the layer across the discontinuities yields an expression for $\mathbf{Q}_{++}$ that is product of two matrices in the form of (A13a), one for an azimuthal change of $\delta\phi$ at $z + \delta z$ and one for a change of $-\delta\phi$ at z. Multiplying these out and using the small-angle approximations, we can express the forward scattering matrix in terms of a perturbation parameter $\delta X \equiv \Delta k\,\delta z\,\delta\phi\,\mathrm{e}^{\mathrm{i}\Delta k(d-z)}$:

$$\mathbf{Q}_{++} = \begin{bmatrix} 1 & -\mathrm{i}\delta X \\ -\mathrm{i}\delta X^* & 1 \end{bmatrix}. \tag{B5}$$

The apparent splitting parameters minimize the energy on the transverse component of the back-projected displacement field, given by the quadratic form (7). In the present notation, this integral becomes

$$\varepsilon^2(\phi', \Delta t') = 2\int_0^\infty |\hat{\mathbf{y}}^{\mathsf{T}}\,\Gamma_{\mathrm{h}}^{-1}(\phi', \Delta t')\,\Gamma\,\hat{\mathbf{x}}|^2\,|u_I(\omega)|^2\,d\omega \tag{B6}$$

The homogeneous-layer splitting operator in this expression corresponds to the perturbed apparent splitting parameters $\phi' = \phi_0 + \delta\phi'$ and $\Delta t' = \Delta t + \delta t'$, which can be expressed in a form similar to (B3):

$$\Gamma_{\mathrm{h}}(\phi', \Delta t') = \mathbf{U}(\phi_0)\mathbf{H}(\Delta t)\mathbf{Q}'_{++}\,\mathbf{U}^{\mathsf{T}}(\phi_0). \tag{B7}$$

Equating (B6) with $\Gamma_{\mathrm{h}}(\phi', \Delta t') = \mathbf{U}(\phi')\mathbf{H}(\Delta t')\mathbf{U}^{\mathsf{T}}(\phi')$ yields the scattering matrix,

$$\mathbf{Q}'_{++} = \begin{bmatrix} c - \mathrm{i}s\cos 2\delta\phi' & -\mathrm{i}s\sin 2\delta\phi' \\ -\mathrm{i}s^*\sin 2\delta\phi' & c^* + \mathrm{i}s^*\cos 2\delta\phi' \end{bmatrix}, \tag{B8}$$

with complex coefficients,

$$c = e^{i\omega\Delta t/2}\cos[\omega(\Delta t+\delta t')/2], \tag{B9a}$$
$$s = e^{i\omega\Delta t/2}\sin[\omega(\Delta t+\delta t')/2]. \tag{B9b}$$

Eqn. (B8) is exact and does not require the perturbations $\delta\phi'$ and $\delta t'$ to be small. When $\phi_0 = \Delta t = 0$, for example, $c = \cos\omega\delta t'/2$, $s = \sin\omega\delta t'/2$, and (B7) reduces to $\Gamma_h(\delta\phi',\delta t')$ in the form given by (A24).

From (B3) and (B6) and the fact that $\mathbf{Q}'_{++}$ is unitary, we obtain $\Gamma_h^{-1}(\delta\phi',\delta t')\Gamma = \mathbf{U}(\phi_0)\mathbf{Q}'^{\dagger}_{++}\mathbf{Q}_{++}\mathbf{U}^{\mathsf{T}}(\phi_0)$. The energy (B6) involves an integration over the (2,1) component of this matrix. The product $\mathbf{Q}'^{\dagger}_{++}\mathbf{Q}_{++}$ can be expressed in the general SU(2) form (A22), where

$$\begin{aligned}\alpha &= c^* + s^*(\delta X^*\sin 2\delta\phi' + i\cos 2\delta\phi'),\\ \beta &= -i\delta X c^* + s^*(\delta X\cos 2\delta\phi' + i\sin 2\delta\phi').\end{aligned} \tag{B10}$$

Working out the appropriate matrix element in terms of the real and imaginary parts of these coefficients, we find

$$\varepsilon^2(\delta\phi',\delta t') = 2\int_0^\infty\{\mathrm{Re}(\beta)^2 + [\mathrm{Im}(\beta)\cos 2\phi_0 + \mathrm{Re}(\alpha)\sin 2\phi_0]^2\}|u_I(\omega)|^2\,d\omega. \tag{B11}$$

To find the energy minimum, we differentiate (B11) with respect to the perturbations $\delta\phi'$ and $\delta t'$ and set the results equal to zero, which gives two equations for the apparent splitting parameters. Linearizing these equations in $\delta\tilde{\phi}$ and $\delta\tilde{t}$, we obtain a 2×2 system for the Fréchet kernels:

$$\begin{bmatrix} A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}\begin{bmatrix} G_\phi(z)\\ G_t(z)\end{bmatrix} = \Delta k\begin{bmatrix} D_1(z)\\ D_2(z)\end{bmatrix}, \tag{B11}$$

$$A_{11} = \sin^2 2\phi_0\int_0^\infty(\omega/\omega_0)|u_I(\omega)|^2\,d\omega, \tag{B12a}$$
$$A_{12} = \sin 2\phi_0\cos 2\phi_0\int_0^\infty\sin\omega\Delta t\,|u_I(\omega)|^2\,d\omega, \tag{B12b}$$
$$A_{21} = \sin 2\phi_0\cos 2\phi_0\int_0^\infty(\omega/\omega_0)\sin\omega\Delta t\,|u_I(\omega)|^2\,d\omega, \tag{B12c}$$

$$A_{22} = \int_0^\infty (1-\cos\omega\Delta t)^2 |u_I(\omega)|^2 d\omega + \cos^2 2\phi_0 \int_0^\infty \sin^2 \omega\Delta t |u_I(\omega)|^2 d\omega, \quad \text{(B12d)}$$

$$D_1(z) = \sin 2\phi_0 \cos 2\phi_0 \int_0^\infty (\omega/\omega_0)\cos\Delta k(d-z)|u_I(\omega)|^2 d\omega, \quad \text{(B13a)}$$

$$\begin{aligned} D_2(z) = & \int_0^\infty (1-\cos\omega\Delta t)\sin\Delta k(d-z)|u_I(\omega)|^2 d\omega \\ & + \cos^2 2\phi_0 \int_0^\infty (\omega/\omega_0)\sin\Delta k(d-z)\cos\omega\Delta t' |u_I(\omega)|^2 d\omega. \end{aligned} \quad \text{(B13b)}$$

Approximating these integrals using (B2) and solving for the kernels leads to the expressions (10)-(14) given in the text.

References

Agnon, A., and M. S. T. Bukowinski, δs at high-pressure and ∂lnVs\∂lnVp in the lower mantle, *Geophys. Res. Lett., 17,* 1149-1152, 1990.

Albarède, F., and R. D. van der Hilst, New mantle convection model may reconcile conflicting evidence, *EOS Trans. Am. Geophys. Un., 45,* 535-539, 1999.

Ando, M., *ScS* polarization anisotropy around the Pacific Ocean, *J. Phys. Earth*, *32*, 179-195, 1984.

Ando, M, and Y. Ishikawa, Observations of shear-wave velocity polarization anisotropy beneath Honshu, Japan: two masses with different polarizations in the upper mantle, *J. Phys. Earth*, *30*, 191-199, 1982.

Armstrong J. T., CITZAF – A package of correction programs for the quantitative electron microbeam x-ray analysis of thick polished materials, thin-films and particles, *Microbeam Anal. 4*, 177-200, 1995.

Baker, M. B., T. L. Grove, and R. Price, Primitive basalts and andesites from the Mt. Shasta region, N. California: Products of varying melt fraction and water content. *Contrib. Mineral. Petrol., 118,* 111-129, 1994.

Becker, T. W., J. B. Kellogg, and R. J. O'Connell, Thermal constraints on the survival of primitive blobs in the lower mantle, *Earth Planet. Sci., Lett., 171,* 351-365, 1999.

Bijwaard, H., W. Spakman, and E. R. Engdahl, Closing the gap between regional and global traveltime tomography, *J. Geophys. Res., 103,* 30,055-30,078, 1998.

Bouillier, A. M., and A. Nicolas, Classification of textures and fabrics of peridotite xenoliths from south African kimberlites, *Phys. Chem. Earth, 9,* 467-475, 1975.

Boyd, F. R., High- and low-temperature garnet peridotite xenoliths and their possible relation to the lithosphere-asthenosphere boundary beneath southern Africa. In: P. H. Nixon (ed.) *Mantle Xenoliths*, New York: John Wiley & Sons, pp 403-412, 1987.

Boyd, F. R., Compositional distinction between oceanic and cratonic lithosphere, *Earth Planet. Sci. Lett*., 96, 15-26., 1989.

Boyd, F. R., and S. A. Mertzman, Composition and structure of the Kaapvaal Lithosphere, southern Africa. In: B.A. Mysen (ed.) *Magmatic Processes: Physiochemical Principles,. Spec. Publ. 1*, University Park, PA: The Geochemical Society, pp. 13-24., 1987.

Boyd, F. R., N. P. Pokhilenko, D. G. Pearson, S. A. Mertztman, N. V. Sobolev, and L. W. Finger, Composition of the Siberian cratonic mantle: evidence from Udachnaya peridotite xenoliths, *Contrib. Mineral. Petrol., 128,* 228-246, 1997.

Brey, G. P., and T. Köhler, Geothermobarometry in four-phase Lherzolites II. New thermobarometers, and practical Assessment of Existing thermobarometers, *J. Petrol. 31,* 1353-1378, 1990.

Carlson, W. D., C. Denison, and R. A. Ketcham, Controls on the nucleation and growth of porphyroblasts: kinetics from natural textures and numerical models, *Geol. J., 30,* 207-225, 1995.

Cashman, K. V., and J. M. Ferry, Crystal size distribution (CSD) in rocks and the kinetics and dynamics of crystallization, *Contrib. Mineral. Petrol., 99,* 401-415., 1988.

Castle, J. C., and R. D. van der Hilst, The core-mantle boundary under the Gulf of Alaska: No ULVZ for shear waves, *Earth Planet. Sci. Lett., 176,* 311-321, 2000.

Castle, J. C., and R. D. van der Hilst, Searching for seismic scattering off mantle discontinuities between 800 and 2000 km depth, *J. Geophys. Res., under revision,* 2001.

Chen, W.-P., and S. Özalaybey, Correlation between seismic anisotropy and Bouguer gravity anomalies in Tibet and its implications for lithospheric structures, *Geophys. J. Int., 135,* 93-101, 1998.

Christensen, U. R., and A. Hofmann, Segregation of subducted oceanic-crust in the convecting mantle, *J. Geophys. Res., 95,* 19,867-19,884, 1994.

Clitheroe, G., R., and R. D. van der Hilst, Complex anisotropy in the Australian lithosphere from shear-wave splitting in broad-band *SKS* records, in *Structure and Evolution of the Australian Continent*, J. Braun, J. Dooley, B. Goleby, R. van der Hilst, and C. Klootwijk (eds), Am. Geophys. Union Geodynamic Series, **26**, 39-57, 1998.

Cox, K. G., M. R. Smith, and S. Beswetherick, Textural studies of garnet lherzolites: evidence of exsolution origin from high-temperature harzburgites. In: . P. H. Nixon (ed.) in *Mantle Xenoliths,* New York: John Wiley & Sons., pp 537-550, 1987.

Davaille, A., Simultaneous generation of hotspots and superswells by convection in a heterogeneous planetary mantle, *Nature, 402,* 756-760, 1999.

Davies, J. H., and D. J. Stevenson, Physical model of source region of subduction zone volcanics, *J. Geophys. Res. 97,* 2037–2070, 1992.

Duffy, T. S., and D. L. Anderson, Seismic velocities in mantle minerals and the mineralogy of the upper mantle, *J. Geophys. Res., 94,* 1895-1912, 1989.

Engdahl, E. R., R. D. van der Hilst, and R. Buland, Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seis. Soc. Amer., 88,* 722-743, 1998.

Fei, Y., H. K. Mao, J. Shu, and J. Hu, P-V-T equation of state of magnesiowüstite ($Mg_{0.6}Fe_{0.4}O$), *Phys. Chem. Miner., 18,* 416-422, 1992.

Finnerty, A. A., and F. R. Boyd, Thermobarometry for garnet peridotites: basis for the determination of thermal and compositional structure of the upper mantle. In: P. H. Nixon (ed.) *Mantle Xenoliths,* New York: John Wiley & Sons., pp 381-402, 1987.

Forsyth, D. W., The early structural evolution and anisotropy of the oceanic upper-mantle, *Geophys. J. R. Astron. Soc., 43* , 103-162, 1975.

Fouch, M. J., and K. M. Fischer, Mantle anisotropy beneath northwest Pacific subduction zones, *J. Geophys. Res., 101*, 15,987-16,002, 1996.

Gaherty, J. B., and T. H. Jordan, The Lehman discontinuity as the base of an anistropic layer beneath the continents, *Science, 268,* 1468-1471, 1995.

Gao, S. S., P. G. Silver, D. E. James, and the Kaapvaal Working Group, Seismic structure and tectonics of southern Africa—progress report, Eos, Transactions, AGU 79, 45, 1998.

Garnero, E. J., and D. V. Helmberger, Further constraints and uncertainties in modeling a thin varying ultralow velocity layer at the base of the mantle, *J. Geophys. Res, 103,* 12,495-12,509, 1998.

Garnero, E. J., and T. Lay, lateral variations in lowermost mantle shear wave anisotropy beneath the north Pacific and Alaska, *J. Geophys. Res., 102,* 8121-8135, 1997.

Grand, S. P., R. D. van der Hilst, and S. Widiyantoro, Global seismic tomography: a snapshot of convection in the Earth, *GSA Today, 7*, 1-7, 1997.

Grove, T. L., S. W. Parman, and J. C. Dann, Conditions of magma generation for Archean komatiites from the Barberton Mountainland, South Africa. In: Y Fei, C.M. Bertka & B.O. Mysen (ed.) *Mantle Petrology: Field Observations and High Pressure Experimentation: A tribute to Francis R. (Joe) Boyd*, The Geochemical Society, Special Publication 6, pp. 155-167., 1999.

Guyot, F., M. Madon, J. Peyronneau, and J. P. Poirier, X-ray microanalysis of high-pressure/high-temperature phases synthesized from natural olivine in a diamond-anvil cell, *Earth. Planet. Sci. Lett., 90,* 52-64, 1988.

Haggerty, S. E., and V. Sautter, Ultradeep (greater than 300 kilometers), ultramafic upper mantle xenoliths, *Science, 248,* 993-996, 1990.

Harte, B., Rock nomenclature with particular relation to deformation and recrystallisation textures in olivine-bearing xenoliths, *J. Geology, 85,* 297-288, 1977.

Harte, B., Mantle peridotites and processes–the kimberlite sample, In: Hawkesworth and Norry (ed.) *Continental Basalts and Mantle Xenoliths*, Nantwich, Cheshire: Shiva Publishing Ltd., pp 46-91, 1983.

Harte, B., K. G. Cox, and J. J. Gurney, Petrograph and geological history of upper mantle xenoliths from the Matsoku kimberlite pipe, *Phys. Chem. Earth, 9*, 447-506, 1975.

Herzberg, C. T., Lithosphere peridotites of the Kaapvaal craton, *Earth. Planet. Sci. Lett., 120,* 13-29, 1993.

Isaak, D. G., O. L. Anderson, and R. E. Cohen, The relationship between shear and compressional velocities at high pressures: reconciliation of seismic tomography and mineral physics, *Geophys. Res. Lett., 8,* 741-744, 1992.

Ishii, M., and J. Tromp, Normal-mode and free-air gravity constraints on lateral variations in velocity and density of earth's mantle, *Science, 285,* 1231-1236, 1999.

Ishii, M., and J. Tromp, Even degree lateral variations in the Earth mantle constrained by free oscillations and the free-air gravity anomaly, *Geophys. J. Int.*, *145*, 77-96, 2001.

Ito, E., E. Takahashi, and Y. Matsui, The mineralogy and chemistry of the lower mantle: an implication of he ultrahigh-pressure phase relations in the system MgO-FeO-SiO_2, *Earth Planet. Sci. Lett., 67,* 238-248, 1984.

Ito, E., and E. Takahashi, Post-spinel transformations in the system Mg_2SiO_4-Fe_2SiO_4 and some geophysical implications, *J. Geophys. Res., 94,* 10,637-10,646, 1989.

Jeanloz, R., and A. B. Thompson, Phase transitions and mantle discontinuities, *Rev. Geophys., 21,* 51-74, 1983.

Jerram, D. A., M. J. Cheadle, R. H. Hunter, and M. T. Elliott, The spatial distribution of grains and crystals in rocks, *Contrib. Mineral. Petrol., 125,* 60-74, 1996.

Jordan, T. H., The continental tectosphere, *Rev. Geophys., 13,* 1-12, 1975.

Jordan, T. H., and J. B. Gaherty, Stochastic modeling of small-scale, anisotropic structure in the continental upper mantle in *Proceedings of the 17th Annual Seismic Research Symposium*, PL-TR-95-2108, Phillips Laboratory, Massachusetts, pp433-451, 1996.

Jordan, T. H., R. L. Saltzer, and J. B. Gaherty, Small-scale anisotropic heterogeneity in the continental upper mantle, *Seism. Res. Lett., 70,* 259, 1999.

Kárason, H., and R. D. van der Hilst, Tomographic imaging of the lowermost mantle with differential times of refracted and diffracted core phases (PKP, Pdiff), *J. Geophys. Res., 106,* 6569-6588, 2001.

Karato, S., Importance of anelasticity in the interpretation of seismic tomography, *Geophys. Res. Lett., 20,* 1623-1626, 1993.

Karki, B. B., and L. Stixrude, Seismic velocities of major silicate and oxide phases of the lower mantle, *J. Geophys. Res., 104,* 13,025-13,033, 1999.

Keith, C. M., and S. Crampin, Seismic body waves in anisotropic media: synthetic seismograms, *Geophys. J. R. Astron. Soc., 49,* 225-243, 1977.

Keleman, P. B., S. R. Hart, and S. Bernstein, Silica enrichment in the continental upper mantle via melt/rock reaction, *Earth Planet. Sci. Lett., 164,* 387-406, 1998.

Kellogg, L. H., B. H. Hager, and R. D. van der Hilst, Compositional stratification in the deep mantle, *Science, 283,* 1881-1884, 1999.

Kennett, B. L. N., E. R. Engdahl, and R. Buland, Constraints on seismic velocities in the Earth from traveltimes, *Geophys, J. Int., 122,* 108-124, 1995.

Kennett, B. L. N., S. Widiyantoro, and R. D. van der Hilst, Joint seismic tomography for bulk sound and shear wave speed in the Earth's mantle, *J. Geophys. Res., 103,* 12,469-12,493, 1998.

Kennett, B. L. N., Seismic wave propagation in stratified media, Cambridge University Press, 1983.

Kesson, S. E., J. D. Fitzgerald, and J. Shelley, Mineralogy and dynamics of a pyrolite lower mantle, *Nature, 393,* 252-255, 1998.

Kincaid, C., and I. S. Sacks, Thermal and dynamical evolution of the upper mantle in subduction zones, *J. Geophys. Res., 102,* 12,295-12,315, 1997.

Kinzler, R. J., Melting of mantle peridotite at pressures approaching the spinel to garnet transition: Application to mid-ocean ridge basalt petrogenesis, *J. Geophys. Res., 102,* 853-874, 1997.

Kinzler, R. J., and T. L. Grove, Origin of depleted cratonic harzburgite by deep fractional melt extraction and shallow olivine cumulate infusion. *In Proceedings of the 7th International Kimberlite Copnference*, J. J. Gurney et al., (ed), Red Roof Design, Cape Town, pp 437-443, 1999.

Kretz, R., On the spatial distribution of crystals in rocks, *Lithos, 2,* 39-66, 1969.

Kretz, R., (1993) A garnet population in Yellowknife schist, Canada, *J. Metamorphic. Geol., 11,* 101-120, 1993.

Kruse, R., and H. Stünitz, Deformation mechanisms and phase distribution in mafic high-temperature mylonites from the Jotun Nappe, southern Norway, *Tectonophysics, 303,* 223-249., 1999.

Kushiro, I., Partial melting of garnet lherzolites from kimberlite at high pressures. In: P. H. Nixon (ed.) *Lesotho Kimberlites,* Lesotho National Development Corporation, pp. 249-299, 1973.

Lay, T., Q. Williams, E. J. Garnero, L. Kellogg, and M. E. Wysession, Seismic wave anisotropy in the D" region and its implications, *in The Core-Mantle Boundary Region, AGU,* 299-318, 1998.

Lebars, M., and A. Davaille, The oscillatory egime of thermo-chemical convection in a viscous fluid, *J. Fluid. Mech., in prep.,* 2001.

Li, X. D., and B. Romanowicz, Global mantle shear velocity model developed using nonlinear asymptotic coupling theory, *J. Geophys. Res., 101,* 22,245-22,272, 1996.

Mallick, S., and N. L. Frazer, Computation of synthetic seismograms for stratified, azimuthally anisotropic media, *J. Geophys. Res.*, *95*, 8513-8526, 1990.

Mao, H. K., R. J. Henley, Y. Fei, J. F. Shu, C. Chen, A. P. Jephcoat, Y. Wu, and W. A. Bassett, Effect of pressure, temperature, and composition on lattice parameters and density of (Fe,Mg)SiO_3-perovskites to 30 GPa, *J. Geophys. Res., 96,* 1069-1079, 1991.

Mao, H. K., G. Shen, and R. J. Hemley, Multivariable dependence of Fe-Mg partitioning in the lower mantle, *Science, 278,* 2098-2100, 1997.

Marson-Pidgeon, K., and M. K. Savage, Frequency-dependent anisotropy in Wellington, New Zealand, *Geophys. Res. Lett., 24,* 3297-3300, 1997.

Masters, G., G. Laske, H. Bolton, and A. Dziewonski, The relative behavior of shear velocity, bulk sound speed, and compressional velocity in the mantle: implications for chemical and thermal structure *in Earth's deep interior: Mineral physics and tomography from the atomic to the global scale, AGU Monograph 117*, 63-87, 2000.

Masters, G., S. Johnson, G. Laske, and H. Bolton, A shear-velocity model of the mantle, *Philos. Trans. R. Soc. Lond. A, 354,* 1385-1411, 1996.

Matzel, E., M. K. Sen, and S. P. Grand, Evidence for anisotropy in the deep mantle beneath Alaska, *Geophys. Res. Lett., 23,* 2417-2420, 1996.

McCammon, C., Perovskite as a possible sink for ferric iron in the lower mantle, *Nature, 387,* 694-696, 1997.

Miyake, A., Monte Carlo simulation of normal grain growth in 2- and 3-dimensions: the lattice-model-independent grain size distribution, *Contrib. Mineral. Petrol., 130,* 121-133., 1998.

Montagner, J. P., and T. Tanimoto, Global anisotropy in the upper mantle inferred from the regionalization of phase velocities, *J. Geophys. Res., 95,* 4797-4819, 1990.

Nataf, H. C., I. Nakanishi, and D. L. Anderson, Anisotropy and shear-velocity heterogeneities in the upper mantle, *Geophys. Res. Lett., 11,* 109-112,1984 .

Nixon, P. H., Kimberlitic xenoliths and their cratonic setting. In: P.H. Nixon (ed.) *Mantle Xenoliths,* John Wiley & Sons, pp 215-239, 1987.

Nixon, P. H., and F. R. Boyd, Petrogenesis of the granular and sheared ultrabasic nodule suite in kimberlites. In: P. H. Nixon (ed.) *Lesotho Kimberlites,* Cape Town: Cape and Transvaal Printers Ltd., pp. 48-56, 1973.

Özalaybey, S., and W.-P. Chen, Frequency-dependent analysis of SKS/SKKS waveforms observed in Australia: Evidence for null birefringence, *Phys. Earth Planet. Interior,* in press, 1999.

Paige, C. C., and M. A. Saunders, LSQR: an algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Soft., 8,* 43-71, 1982.

Parman, S. W., J. C. Dann, T. L. Grove, and M. J. de Wit, Emplacement conditions of komatiite magmas from the 3.49 Ga Komati formation, Barberton Greenstone Belt, South Africa, *Earth Planet. Sci. Lett. 150,* 303-323, 1997.

Parman, S. W., Petrology and Geochemistry of high degree mantle melts, *PhD thesis, MIT,* 2001.

Persh, S. E., J. E. Vidale, and P. S. Earle, Absence of short-period ULVZ precursors to PcP and ScS from two regions of the CMB, *Geophys. Res. Lett., 28,* 387-390, 2001.

Raitt, R. W., G. G. Shor, T. J. G. Francis and G. B. Morris, Anisotropy of the Pacific upper mantle, *J. Geophys. Res., 74,* 3095-3109, 1969.

Resovsky, J. S., and M. H. Ritzwoller, Regularization uncertainty in density models estimated from normal mode data, Geophys. Res. Lett., 26, 2319-2322, 1999.

Revenaugh, J. S., and R. Meyer, Seismic evidence of partial melt within a possibly ubiquitous low-velocity layer at the base of the mantle, *Science, 277,* 670-673, 1997.

Richards, M. A., B. H. Hager, and N. H. Sleep, Dynamically supported geoid highs over hotspots: observations and theory, *J. Geophys. Res., 89,* 5987-6002, 1988.

Ringwood, A. E., Phase transformations and their bearing on the constitution and dynamics of the mantle, *Geochim, Cosmochim, Acta 55,* 2083-2110, 1991.

Robertson, G. S., and J. H. Woodhouse, Ratio of relative S to P velocity heterogeneity in the lower mantle, *J. Geophys. Res., 101,* 20,041-20,052, 1996.

Romanowicz, B., Can we resolve 3D density heterogeneity in the lower mantle?, *Geophys. Res. Lett., 28,* 1107-1110, 2001.

Rümpker, G., and P. G. Silver, Apparent shear-wave splitting parameters in the presence of vertically varying anisotropy, *Geophys. J. Int., 135,* 790-800, 1998.

Saltzer, R. L., T. H. Jordan, J. B. Gaherty, L. Zhao, and the Kaapvaal Working Group, Anisotropic structure of the Kaapvaal craton from surface wave analysis, *EOS, Transactions, AGU 79,* 45,1998.

Saltzer, R. L., J. B. Gaherty, and T. H. Jordan, How are vertical shear-wave splitting measurements affected by variations in the orientatin of azimuthal anisotropy with depth?, *Geophys. J. Int., 141,* 374-390, 2000.

Saltzer, R. L., N. Chatterjee, and T. L. Grove, The spatial distribution of garnets and pyroxenes in mantle xenoliths: Pressure-temperature history of peridotites from the kaapvaal craton, *J. Petrology, in press,* 2001.

Saltzer, R. L., R. D. van der Hilst, and H. Kárason, Comparing P and S wave heterogeneity in the mantle, *Geophys. Res. Lett., 28,* 1335-1338, 2001.

Saltzer, R. L., E. Stutzmann, and R. D. van der Hilst, Poisson's ratio beneath Alaska from the surface to the core-mantle boundary, *J. Geophys. Res., under review,* 2001.

Sautter, V., S. E. Haggerty, and S. Field, Ultradeep (>300 Kilometers) Ultramafic xenoliths: petrological evidence from the transition zone, *Science, 252,* 827-830, 1991.

Savage, M. K., Seismic anisotropy and mantle deformation: what have we learned from shear wave splitting?, *Rev. Geophys., 37*, 65-106, 1999.

Shearer, P. M., and J. A. Orcutt, Compressional and shear wave anisotropy in the oceanic lithosphere — the Ngendei seismic refraction experiment, *Geophys. J. R. Astron. Soc., 87,* 967-1003, 1986.

Silver, P. G., Seismic anisotropy beneath the continents: probing the depths of geology, *Annu. Rev. of Earth. Planet. Sci., 24,* 385-432, 1996.

Silver, P. G., and W. W. Chan, Shear wave splitting and mantle deformation, *J. Geophys. Res., 96,* 16,429-16,454, 1991.

Silver, P. G., and M. K. Savage, The interpretation of shear-wave splitting parameters in the presence of two anisotropic layers, *Geophys. J. Int., 119*, 949-963, 1994.

Stacey, F. D., Thermoelasticity of a mineral composite and a reconsideration of lower mantle properties, *Phys. Earth. Planet. Int., 106,* 219-236, 1998.

Steinberger, B., Slabs in the lower mantle – results of dynamic modeling compared with tomographic images and the geoid, *Phys. Earth. Planet. Int., 118,* 241-257, 2000.

Su, W., and A. M. Dziewonski, Simultaneous inversion for 3-D variations in shear and bulk velocity in the mantle, *Phys. Earth. Planet. Int., 100,* 135-156, 1997.

Su, W., R. L. Woodward, and A. M. Dziewonski, Degree 12 model of shear velocity heterogeneity in the mantle, *J. Geophys. Res., 99,* 6945-6980, 1994.

Tackley, P. J., Three-dimensional simulations of mantle convection with a thermo-chemical basal boundary layer: D"? in *The Core-Mantle boundary Region,* M. Gurnis, M. E. Wysession, E. Knittle, B. A. Buffett Eds, *AGU, Washington DC,* 231-253, 1998.

Tackley, P. J., Mantle convection and plate tectonics: toward an integrated physical and chemical theory, *Science, 288,* 2002-2007, 2000.

Tanimoto, T., and D. L. Anderson, Lateral heterogeneity and azimuthal anisotropy of the upper mantle: Love and Rayleigh waves 100-250 s, *J. Geophys. Res., 90,* 1842-1858, 1985.

Thomson, W.T., Transmission of elastic waves through a stratified solid medium, *Jour. Appl. Phys., 21,* 89-93, 1950.

Trampert, J., P. Vacher, and N. Vlaar, Sensitivities of seismic velocities to temperature, pressure and composition in the lower mantle, *Phys. Earth Planet. Int., in press,* 2001.

Van der Hilst, R. D., and H. Kárason, Compositional heterogeneity in the bottom 1000 kilometers of Earth's mantle: toward a hybrid convection model, *Science, 283,* 1885-1888, 1999.

Van der Hilst, R. D., S. Widiyantoro, and E. R. Engdahl, Evidence for deep mantle circulation from global tomography, *Nature, 386,* 578-584, 1997.

Varshalovich, D. A., and A. N. Moskalev, *Quantum Theory of Angular Momentum* ed. Kheronski, V. K. World Scientific, 1988.

Vidale, J. E., Complex polarzation analysis of particle motion, *Bull. Seismol. Soc. Am., 76,* 1393-1405, 1986.

Vidale, J. E., G. Schubert, and P. S. Earle, Unsuccessful initial search for a midmantle chemical boundary with seismic arrays, *Geophys. Res. Lett., 28,* 859-862, 2001.

Vinnik, L.P., V. Farra, and B. Romanowicz, Azimuthal anisotropy in the earth from observations of *SKS* at GEOSCOPE and NARS broadband stations, *Bull. Seismol. Soc. Am., 79,* 1542-1558, 1989.

Walter, M. J., Melting of garnet peridotite and the origin of komatiite and depleted lithosphere, *J. Petrol., 39,* 29-60., 1998.

Wen, L., and D. L. Anderson, The fate of slabs inferred from seismic tomography and 130 million years of subduction, *Earth Planet. Sci. Lett., 133,* 185-198, 1995.

Wessel, P., and W. H. F. Smith, Free software helps map and display data, *Eos Trans AGU, 72,* 441-446, 1991.

Widiyantoro, S., B. L. N. Kennett, and R. D. van der Hilst, Seismic tomography with P and S data reveals lateral variations in the rigidity of slabs, *Earth Planet. Sci. Lett., 173,* 91-100, 1999.

Wolfe, C. J., and S.C. Solomon, Shear-wave splitting and implications for mantle flow beneath the MELT region of the East Pacific Rise, *Science, 280,* 1230-1232.

Wood, B. J., and D. C. Rubie, The effect of alumina on phase transformations at the 660 km discontinuity from Fe-Mg partition experiments, *Science 273,* 1522-1524, 1996.

Wysession, M. E., E. A. Okal., and C. R. Bina, The structure of the core-mantle boundary from diffracted waves, *J. Geophys. Res., 97,* 8749-8764, 1992.

Wysession, M. E., A. Langenhorst, M. J. Fouch, K. M. Fischer, and I. Ghassan, Lateral variations in compressional/shear velocities at the base of the mantle, *Science, 284,* 120-125, 1999.

Yagi, T., P. M. Bell, and H. K. Mao, Phase relations in the system MgO-FeO-SiO_2 between 150 and 700 kbars at 1000° C, *Carnegie Inst. Washington Yearb., 78,* 614-618, 1997.

Yamada, H., and E. Takahashi, Subsolidus phase relations between coexisting garnet and two pyroxenes at 50 to 100 kbar in the system CaO-MgO-Al_2O_3-SiO_2. In: J. Kornprobst (ed.) *Kimberlites, II. The mantle and crust-mantle relationships, proceedings of the third international kimberlite conference,* New York: Elsevier, pp. 247-255., 1984.

Zindler, A., and S. Hart, Chemical Geodyanmics, *Annu. Rev. Earth Planet. Sci., 14*, 493-571, 1986.