Shear Wave Splitting and Subcontinental Mantle Deformation

PAUL G. SILVER

Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington, D. C.

W. WINSTON CHAN

Teledyne Geotech, Alexandria Laboratories, Alexandria, Virginia

We have made measurements of shear wave splitting in the phases SKS and SKKS at 21 broadband stations in North America, South America, Europe, Asia, and Africa. Measurements are made using a retrieval scheme that yields the azimuth of the fast polarization direction $\phi$ and delay time $\tau$ of the split shear wave plus uncertainties. Detectable anisotropy was found at most stations, suggesting that it is a general feature of the subcontinental mantle. Delay times range from 0.65 s to 1.70 s and average about 1 s. Somewhat surprisingly, the largest delay time is found in the 2.7 b.y.-old Western Superior Province of the Canadian Shield. The splitting observations are interpreted in terms of the strain-induced lattice preferred orientation of mantle minerals, especially olivine. We consider three hypotheses concerning the origin of the continental anisotropy: (1) strain associated with absolute plate motion, as in the oceanic upper mantle, (2) crustal stress, and (3) the past and present internal deformation of the subcontinental upper mantle by tectonic episodes. It is found that the last hypothesis is the most successful, namely that the most recent significant episode of internal deformation appears to be the best predictor of $\phi$. For stable continental regions, this is interpreted as "fossil" anisotropy, whereas for presently active regions, such as Alaska, the anisotropy reflects present-day tectonic activity. In the stable portion of North America there is a good correlation between delay time and lithospheric thickness; this is consistent with the anisotropy being localized in the subcontinental lithosphere and suggests that intrinsic anisotropy is approximately constant. The acceptance of this hypothesis has several implications for subcontinental mantle deformation. First, it argues for coherent deformation of the continental lithosphere (crust and mantle) during orogenies. This implies that the anisotropic portion of the lithosphere was present since the deformational episode and rules out the addition of undeformed material to this layer by subsequent "underplating" or conductive growth of the thermal boundary layer. One of the most important issues in the study of orogenies is the need to reconcile the formation of thickened lithosphere with the paradoxically high mantle temperatures often associated with orogenic episodes. Most efforts to date have focused on modes of deformation whereby the cold lithospheric mantle is removed (by convective instability or delamination) and replaced by warm asthenosphere. These models, however, are incompatible with the evidence for preserved coherent lithospheric deformation; rather, the deformed mantle appears to have been heated in place. We suggest that the elevated mantle temperatures may be due to the strain heating accompanying the deformation.

INTRODUCTION

The steadily increasing evidence that mantle anisotropy is due to the strain induced lattice preferred orientation (LPO) of upper mantle minerals [see Nicolas and Christensen, 1987], an idea originally proposed by Hess [1964], has provided a means of inferring modes of mantle deformation from observations of seismic anisotropy. There have been numerous studies of upper mantle anisotropy below the ocean basins: measurements of $P_o$ velocity from refraction surveys [e.g., Raitt et al., 1969; Shearer and Orcutt, 1986], and both azimuthal and polarization anisotropy in long-period surface waves [e.g., Forsyth, 1975; Nataj et al., 1984; Tanimoto and Anderson, 1985; Montagner and Tanimoto, 1990]. Nearly all of these studies can be successfully interpreted as due to the strain induced by plate-motion processes: seafloor spreading at the mid-ocean ridges, the differential motion of the lithosphere and asthenosphere, or more generally, the mantle flow induced by the motions of the plates.

An equally successful characterization of subcontinental mantle deformation from anisotropic observations has yet to fully emerge. Not only is plate tectonics less successful in describing the features of the continents and in predicting deformation in the subcontinental mantle, but the much shorter horizontal length scales of coherent continental deformation make the study of anisotropy more difficult using the techniques typically applied to the oceans. Notable exceptions are the few $P_o$ observations on the continents [Barnford, 1977; Barnford et al., 1979; Vetter and Minster, 1981; Beghoul and Barazangi, 1990]. The problem of smaller length scales can be addressed by the use of a particular manifestation of anisotropy, shear wave splitting [Christensen, 1966; Keith and Crampin, 1977]. Splitting in teleseismic shear waves such as $SKS$, $ScS$, and $S$, with steep arrival angles beneath the receiver, provides excellent lateral resolution in the upper mantle, and thus allows for the direct comparison of anisotropy with surface tectonic and geologic features possessing typical continental dimensions. Measurements of shear wave splitting in waves sampling the upper mantle began less than a decade ago with the pioneering work of Ando and coworkers using $S$-waves from nearby events in the descending slab [Ando and Ishikawa, 1982; Ando et al., 1983; Bowman and Ando, 1987] and $ScS$ [Ando, 1984; see also Fukao, 1984]. Because $ScS$ is most easily used in the distance range of less than $30^\circ$, it provides information primarily about the anisotropy in tectonic areas along active continental margins. $ScS$, however,
is less useful in studying stable continental interiors that are relatively devoid of seismic activity. One important advance in retrieving splitting information has been the utilization of SKS. First employed by Vinnik et al. [1984], this phase has begun to be exploited extensively in teleseismic splitting [Kind et al., 1985; Silver and Chan, 1988; Vinnik et al., 1989a; Ansel and Natal, 1989; Savage et al., 1990, hereafter called SC], because it has many advantages for the study of the continents. First, the observed anisotropy can be localized to the receiver side of the path due to the P to S conversion at the core-mantle boundary (CMB). Second, the detection of anisotropy is much simpler; since SKS is radially polarized in an isotropic, spherically symmetric Earth, detectable energy on the transverse component, SKS T, signifies deviations from these idealized properties. As the effects of anisotropy and lateral heterogeneity can be easily distinguished, SKS T becomes an excellent diagnostic for the presence of anisotropy. Third, SKS is most easily observed beyond 85°, allowing the study of stable continental interiors. Finally, SKS represents a nearly vertical ray path through the mantle, so that the propagation direction is essentially constant; this allows for a somewhat simpler analysis.

Silver and Chan [1988, hereafter called SC] began a systematic study of continental anisotropy through SKS splitting, with the ultimate goal of obtaining a general characterization of subcontinental mantle deformation. They not only found anisotropy at all stations studied, but also proposed that the splitting is predominantly caused by "fossil" strain in the subcontinental mantle due to the last major episode of tectonic activity. The present report represents a continuation of that study. In particular, it seeks to test the fossil strain hypothesis with an expanded data set, against two alternatives: plate-motion related strain, analogous to that found in the oceanic upper mantle, and present-day crustal stress. Each of these hypotheses carries implications. For example, if the plate-motion hypothesis is correct, it would argue that the strains associated with plate motion dominate subcontinental mantle deformation, so that oceanic and continental upper mantle may be treated similarly in this respect. But if the fossil strain hypothesis is true, it argues for something quite different: that the orogenic and rifting episodes responsible for producing the fabric of continental geology have also imparted this same fabric to the mantle. Anisotropy would then constitute the extension of structural geology to the mantle and help provide a geologic history of the entire continental lithosphere.

The evaluation of these hypotheses requires knowledge of three relationships: (1) between splitting and anisotropy; (2) between anisotropy and strain; and (3) between strain and geologic/tectonic processes. For a homogeneous medium, the relation between the splitting parameters, the fast polarization direction φ and delay time δt, and the elasticity tensor Cijkl is straightforward even for the most general anisotropy. While it can become quite complicated for multiple regions of anisotropy, however, the assumption of a localized homogeneous anisotropic layer appears to be justified in most cases and will be utilized for the applications in the present report.

Assuming that anisotropy is caused by the lattice preferred orientation (LPO) of mantle minerals, the relation between anisotropy and strain requires knowledge of composition and mineralogy, along with the single-crystal elastic constants and dominant slip systems for these minerals. Since these properties are reasonably well known for the upper mantle above 400 km depth, the problem is reduced to finding an orientation distribution function (ODF) for the primary anisotropic minerals (such as olivine and orthopyroxene) as a function of finite strain. In fact, seismological applications only require a relationship between a suitably averaged macroscopic elasticity tensor <Cijkl> derived from the ODF, and finite strain. Obtaining an appropriate relationship has involved the interplay of theory [McKenzie, 1979; Rice, 1989a; Rice and Yu, 1991] laboratory measurements, and observations of mantle xenoliths [see Nicolas and Poitevin, 1976; Gueguen and Nicolas, 1980; Christensen, 1984].

The third relationship requires a kinematic or dynamic description of the relevant tectonic process, and a constitutive relation. This issue has been addressed for the strain associated with oceanic plate motion (spreading, subduction, absolute plate motion) and mantle convection [McKenzie, 1979; Rice, 1989a, b] with emphasis on the resulting anisotropy. Continental strain has been examined by several studies [e.g., England and McKenzie, 1982, 1983; England, 1983; Bird, 1984, 1988; Molnar, 1988a; England and Houseman, 1989], although the implications for seismic anisotropy have not been discussed.

In this report we present shear wave splitting measurements from all presently available three-component broadband stations from the GDSN network, the NARS array, and the Chinese Digital Seismic Network (CDSN). This includes measurements from the continents of North America, South America, Eurasia, and Africa. We will also discuss two other related studies, one which examines splitting in the Basin and Range (SSM), and the other which examines IRIS/IDA stations [Givens and Silver, 1990], which have used the same measurement scheme as the present study.

**MEASUREMENT OF SHEAR WAVE SPLITTING PARAMETERS**

In an isotropic, homogeneous medium a shear wave arriving at the surface at time T0 after traversing a path of length L can be written as a vector function

\[ u(\omega) = A u(\omega) \exp(-i \omega T_0) \]

where A is the complex vector amplitude, and \( u(\omega) \) is the wavelet function which is the product of the Fourier-transformed source time function, attenuation operator and instrument response. We assume that A can be approximated by A\( \hat{\Phi} \), where A is a complex scalar and \( \hat{\Phi} \) is a real unit vector pointing in the displacement direction and which is contained in the plane orthogonal to the propagation direction \( \hat{\Phi} \). Incorporating A into the definition of \( u(\omega) \), (1) is rewritten as

\[ u(\omega) = u(\omega) \exp(-i \omega T_0) \hat{\Phi} \]

In essence, we are ignoring phase shifts in the radial component due to the free surface. For the steep arrival angles used in the present report, this should be an excellent approximation.

We model the splitting due to anisotropy as the geometrical operations of projecting \( \hat{\Phi} \) onto the fast and slow polarization directions \( \hat{f} \) and \( \hat{s} \) and then time-shifting these
two components by $\delta t/2$ (earlier) and $-\delta t/2$ (later), respectively. For an elasticity tensor $C_{ijkl}$, $\mathbf{\tilde{f}}$ and $\mathbf{\tilde{b}}$ are the two quasi-shear eigenvectors of the polarization matrix $\mathbf{V}$ defined by

$$\mathbf{\nu}_V \equiv C_{ijkl} \mathbf{\tilde{b}}_j \mathbf{\tilde{b}}_k$$

(3)

[Backus, 1965] with eigenvalues $\beta_j^2$ and $\beta_k^2$ corresponding to the two squared shear velocities. For small anisotropy, $\delta t$ can be conveniently expressed in terms of a relative perturbation in shear velocity $\delta \beta = \beta_0^{-1}(\delta \beta_1 - \delta \beta_2)$ as

$$\delta t = \beta_0^{-1} \delta \beta$$

(4)

where $\delta \beta_{1,2} = \beta_1 - \beta_2$ and $\beta_0$ is the isotropic shear velocity defined such that $\beta_1 + \beta_2 = 0$. The process of producing a split shear wave $u_s(\omega)$ can be represented as the application of the splitting operator

$$\Gamma = \exp[i\omega \delta t/2  \mathbf{\tilde{f}} \hat{\mathbf{f}}] + \exp[-i\omega \delta t/2  \mathbf{\tilde{b}} \hat{\mathbf{b}}]$$

(5)

to a signal of the form (2) with the resulting split waveform

$$u_s(\omega) = w(\omega) \exp[-i\omega T_0] \Gamma(\phi, \delta t) \cdot \hat{\mathbf{p}}$$

(6)

where $\phi$ is the angle between $\mathbf{\tilde{f}}$ and $\hat{\mathbf{p}}$. Defining the tensor

$$\delta T = \delta t/2 (\mathbf{\tilde{f}} \hat{\mathbf{f}} - \mathbf{\tilde{b}} \hat{\mathbf{b}})$$

(7)

$\Gamma$ may be more compactly written as

$$\Gamma = \exp[i\omega \delta T(\phi, \delta t)]$$

(8)

in analogy with the isotropic case (2).

The two parameters that may be estimated are $\delta t$ and $\phi$. A straightforward way of doing so is to try to correct for the anisotropy and choose the pair that most nearly returns (6) into the form (2) by reversing the geometrical operations: that is, search for the inverse operator $\Gamma^{-1}$. From (5), we note that this operator is unitary so that the inverse is simply the complex conjugate $\Gamma^*$. Assuming the polarization plane of the shear wave is known, $\Gamma^{-1}$ may be found by computing the two-dimensional time-domain covariance matrix of particle motion in that plane, as its eigenvalues may then be used as a measure of linearity [Vidale, 1986]. We define the covariance $c_{ij}$ between any two orthogonal components of ground motion, making angles of $\phi$ and $\phi + \pi/2$ with $\hat{\mathbf{p}}$, 

$$c_{ij}(\phi, \delta t) = \int_{-\infty}^{\infty} u_i(t)u_j(t - \delta t)dt \quad i, j = 1, 2$$

(9)

for lag $\delta t$.

In the absence of anisotropy, $c$ will possess one nonzero eigenvalue $\lambda_1 = E_u = \int_{-\infty}^{\infty} w(t)^2dt$ and corresponding eigenvector $\hat{\mathbf{p}}$ (see (2)). In the presence of anisotropy, $c$ will have two nonzero eigenvalues, $\lambda_1$, $\lambda_2$ unless $\phi = n\pi/2$ ($n = 0, 1, 2, ...$) or $\delta t$ is zero. Thus, one may search for $\Gamma^{-1}$ such that the corrected seismogram $u_1(\omega) = \Gamma^{-1} \cdot u_s(\omega)$ possesses a singular covariance matrix. For any candidate pair of values ($\phi$, $\delta t$), the covariance matrix $\mathbf{C}(\phi, \delta t)$ for the rotated and shifted seismograms may be expressed in terms of the covariance $c(\delta t)$ of a reference coordinate system (say north-south, east-west), and will have the components

$$\tilde{c}_{11}(\phi, \delta t) = \int_{-\infty}^{\infty} u_1^2(t + \delta t/2)dt = \tilde{c}_{11}(\phi, 0)$$

$$= R_{11}(\phi)c_{11}(0)R_{11}(\phi)$$

(10)

$$\tilde{c}_{22}(\phi, \delta t) = \int_{-\infty}^{\infty} u_2^2(t - \delta t/2)dt = \tilde{c}_{22}(\phi, 0)$$

$$= R_{22}(\phi)c_{22}(0)R_{22}(\phi)$$

(11)

$$\tilde{c}_{12}(\phi, \delta t) = \int_{-\infty}^{\infty} u_1(t + \delta t/2)u_2(t - \delta t/2)dt$$

$$= R_{12}(\phi)c_{12}(0)R_{21}(\phi)$$

(12)

$$\tilde{c}_{21}(\phi, \delta t) = \tilde{c}_{12}(\phi, \delta t)$$

(13)

where the rotation tensor $\mathbf{R}$ defines the change of coordinate system from the reference frame to that defined by the trial fast and slow polarization directions. Equations (10)-(13) demonstrate that $\tilde{c}$ may be formed by taking linear combinations of the covariance and zero-lag autocorrelation in the original reference frame, which allows for efficient evaluation of candidate operators. In the presence of noise $\eta(t)$, $\tilde{c}$ will not be singular and one seeks instead the matrix that is most nearly singular. While there are apparently different eigenvalue-based measures of linearity such as maximizing $\lambda_1$, $\lambda_1/\lambda_2$, minimizing $\lambda_2$ or $\lambda_1\lambda_2$, they are in fact equivalent. This is because the two-dimensional trace $\Theta = \lambda_1 + \lambda_2$ of $\mathbf{C}$ is invariant with respect to changes in either $\phi$ or $\delta t$. Various methods of estimating splitting parameters may thus be put in this context. Maximizing the cross-correlation between components, a method used by Bowman and Ando [1987], is similar to minimizing the determinant, which can be seen from the relationship $\lambda_1\lambda_2 = c_{11}c_{22} - c_{12}^2$. Maximizing the ratio $\lambda_1/\lambda_2$ (aspect ratio) has been applied by Shih et al. [1989]. Finally, finding the minimum of $\lambda_1\lambda_2$, is equivalent to the method of SC. The importance of $\lambda_1\lambda_2$ is that it constitutes a measure of variance of the noise process. As such, it can provide the basis for calculating a confidence region for the two splitting parameters. Such a procedure has been applied by SC, SSM, and Given and Silver [1990] and is discussed below.

An important special case in splitting estimation is where the isotropic polarization vector $\hat{\mathbf{p}}$ is known, as with SKS and SKKS. If viewed on the radial and transverse components, one obtains for a noise free seismogram, the time-domain expressions

$$u_r(t) = w(t + \delta t/2)\cos^2 \phi + w(t - \delta t/2)\sin^2 \phi$$

(14)

$$u_t(t) = -\frac{1}{2}[w(t + \delta t/2) - w(t - \delta t/2)]\sin(2\phi)$$

(15)

In this case the energy $E_r = \int_{-\infty}^{\infty} u_r^2(t)dt$ on the corrected transverse component can be minimized instead of $\lambda_2$, as in SC. Equations (14) and (15) demonstrate an additional important diagnostic for these phases. For $\delta t$ small compared to the characteristic period of study, $u_r$ is only slightly broadened and distorted, while $u_t$, which is identically zero in an isotropic medium, is approximately proportional to the derivative of the radial component. Note that this relationship will hold for any pair of components in the directions parallel and orthogonal to $\hat{\mathbf{p}}$. 
Error Estimation

Error estimation is an important part of the splitting measurement. In many cases, only a few records will be available and evaluating the uncertainty of each measurement is essential. This is especially true for portable experiments, where the observing time is necessarily short. For an n-point discrete time series, $\chi^2$ is the sum-of-squares of a noise process which, it is assumed, is approximately $\chi^2$-distributed. For $\nu$ degrees of freedom and $k$ parameters, we take the confidence region at the $\alpha$ confidence level to be defined by values of $\lambda_2$ satisfying

$$\frac{\lambda_2}{\lambda_2^{\min}} \leq 1 + \frac{k}{\nu - k} f(k, \nu - k)(1 - \alpha)$$

[16]

[Jenkins and Watts, 1968; Bates and Watts, 1988], where $f$ is the inverse of the $F$-distribution. In our case $k = 2$, $\alpha = 0.05$, and $\nu$, a quantity that depends both on the instrument response and noise spectrum, is usually much smaller than $n$. A discussion of expression (16) and the number of degrees of freedom is given in the appendix. Briefly, from each minimum obtained, an estimate of $\hat{\nu}_i = \nu_i / n_i$ for the $i$th record is calculated; then all estimates for one station are averaged to obtain $\bar{\nu} > n$. For the networks we have used, the product of $\bar{\nu}$ and the sampling rate is roughly constant and equal to one degree of freedom for each second of record. The confidence region is then defined for the $i$th record using $< \bar{\nu} > n_i$ degrees of freedom. If the confidence region is approximately Gaussian, then its bounding values along the parameter axes correspond to $2\sigma$ marginal uncertainties for $\phi$ and $\delta t$. In cases where this region is not symmetric but approximately elliptical, we determine the largest symmetric region that includes the asymmetric region. In either case, we take one half of the bounding values of the confidence region (along the parameter axes) as the $1\sigma$ uncertainties. In cases of no observable splitting, there is a characteristic nonellipsoidal contour pattern that may be interpreted as allowing the three options $\delta t = 0$, $\phi = \phi_b$ or $\phi = \phi_b + 90^\circ$. The second two possibilities can be evaluated by utilising different values of $\bar{\nu}$.

In order to test the statistical assumptions that have gone into the error estimation, we have performed two simple numerical experiments, taking a synthetic shear wave, adding Gaussian random noise with known variance, and finally obtaining splitting estimates. In both cases we have used the geometry of event 85253 (Table 1) to station RSON (Table 2) with a back azimuth $\phi_b$ of $-46^\circ$ and have introduced the splitting parameters of $(\phi, \delta t) = (75.0^\circ, 1.75^\circ)$ in case one and $(-39.0^\circ, 1.75^\circ)$ in case two. (In discussing the observations, $\phi$ will be referenced to north instead of the polarization direction $\phi_b$.) We have convolved the noise with a box-car filter of length $1s$ (four samples). The filter reduces the number of degrees of freedom by a factor of about 2.5. The weighted (by standard deviation) mean for five realizations is $(\phi, \delta t) = (734.3^\circ, 1.804.15^\circ)$. Figure 1a shows one of the realizations. For case 2, $\phi_b$ and $\phi$ differ by only $7^\circ$. Note in Figure 1b the characteristic contour pattern for this case, signifying the absence of detectable splitting. A similar pattern is obtained for $\delta t \approx 0$ or $\phi \approx \phi_b + 90^\circ$.

Resolution

In making the measurements, we have concentrated on broadband digital stations because of the greatly improved resolution. The average value of $\delta t$ for teleseismic splitting is approximately $1s$. For long-period instruments with characteristic periods of $20s$, this is a subtle effect. The result is reduced detectability and increased uncertainty in the splitting parameters. For $SKS$, this is seen by writing $E_1$ (using (15)) as
where we have utilized Parseval’s theorem. The integrand is a product of $\psi(\omega)^2$ modulated by an oscillating function with maxima at $\omega = 0$ and zeros at $\omega = \pm n \pi / \delta t$. For $\delta t = 1$ s, the maxima occur at $0.3$ Hz, $1.5$ Hz, $2.5$ Hz, etc., with zeros at $0$ Hz, $1$ Hz, $2$ Hz, etc. For a broadband instrument that is flat to velocity, a wavelet function with characteristic duration $\tau_c$ [Silver and Jordan, 1983] will have a maximum at about $\tau_c^{-1}$. Thus, energetic intermediate and deep focus events with short durations close to $2\delta t$ are the most useful. The difficulty of using long-period data is that at the peak sensitivity of the instrument (about $0.05$ Hz), the integrand is down by a factor of nearly 50 compared to the maximum. Broadband data are also preferred to narrow band short-period sensors, as cycle-skipping is minimized and several maxima, rather than one frequency maximum, can be utilized.

The data for this study consist of SKS and SKKS phases in the distance range 85-140° from three-component broadband data compiled by the GDSN (including DWWSN, RSTN), GRR, NARS array, and the CDSN. We have considered all events deeper than $80$ km and $m_b$ greater than 5.7 (Table 1). For each station listed in Table 2, we have made individual measurements on all available records with sufficiently excited and isolated SKS and SKKS phases. Measurements are made in two ways: minimizing $E_i$ and minimizing $\lambda_i$. We normally take the former procedure as the measurement and use the latter as a check for the influence of lateral heterogeneity. In this way, we can allow for the possibility that the presence of SKS$_P$ can be explained by the rotation of $\delta$ away from the predicted direction $\phi_0$ (e.g., by a dipping layer) rather than splitting. In performing the minimization, we have chosen increments of $1°$ and $0.05$ s for $(\phi, \delta t)$.

**TABLE 1. List of Earthquakes Used in Study**

<table>
<thead>
<tr>
<th>Event</th>
<th>Year</th>
<th>Month</th>
<th>Day</th>
<th>Time, UT</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Depth, km</th>
<th>$m_b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>79138</td>
<td>1978</td>
<td>05</td>
<td>18</td>
<td>2018:09:50</td>
<td>24.15</td>
<td>142.41</td>
<td>598.00</td>
<td>5.8</td>
</tr>
<tr>
<td>81147</td>
<td>1981</td>
<td>09</td>
<td>04</td>
<td>2115:13:90</td>
<td>9.95</td>
<td>124.00</td>
<td>651.00</td>
<td>5.6</td>
</tr>
<tr>
<td>81150</td>
<td>1983</td>
<td>06</td>
<td>02</td>
<td>2012:50.70</td>
<td>-9.51</td>
<td>-71.25</td>
<td>598.60</td>
<td>5.9</td>
</tr>
</tbody>
</table>

**RESULTS**
The resulting estimates of \((\phi, \delta t)\), given in Table 2, were formed from weighted averages of the individual measurements. Rather than relying solely on the uncertainties derived from the noise model, we have modified the error estimates to allow for other unknown sources of error. First, for individual measurements, we place lower allowable limits on the standard deviations of \(\phi\) and \(\delta t\); these are \(\sigma_\phi > 0.2\) s. Second, \(2\sigma_\phi > 45^\circ\) is interpreted as a failure to detect splitting. Third, we estimate a reduced uncertainty of the success of the two-parameter model in removing the energy on the transverse component. (2) We superimpose the fast and slow components of the isolated SKS (or SKKS) phases, both uncorrected and corrected, and plot the particle motion for these two cases. This illustrates both the phase coherences of the two components and the degree to which the correction procedure produces linear particle motion. (3) In order to assess the uncertainty of each measurement and to check for multiple minima, a contour plot of \(E_2(\phi, \delta t)\) is produced with the 95% contour level drawn plus multiples of this level. We present some samples that illustrate the range of splitting parameters that may be obtained: some that have well constrained anisotropy, and others that show marginally detectable splitting.

Finally, for each of the stations, we display the results of all of the measurements for which there is a good radial component phase. All measurements are plotted as a function of \(\phi\) as in Figure 5. The closed symbols represent constrained measurements; the others signify null measurements (splitting not detected) and are plotted with \(\phi\) corresponding to \(\phi_b \pm 90^\circ\), whichever is closer to the positive observation. This is illustrated in Figures 5 – 7 for stations RSON, RSSD, and COL, as we have the most number of records for each. Note that in the case of RSON, the results are very consistent as a function of \(\phi\) over a range of about 120°. RSSD is at the limit of detectability for the RSTN inter-
mediate period channel ($\delta t = 0.65$ s), which is why there are so many more null measurements. Nevertheless, there are sufficient measurements over a range of $\phi_0$ to suggest that the results are well constrained. In only one case is there a suggestion of azimuthal dependence of the splitting. For COL (Figure 7a) several measurements of $\phi$ over the range $\phi_0 = 100-230^\circ$ yield a consistent value. However, one record, for which $\phi_0 = 261^\circ$ gives an answer that is clearly inconsistent with the rest; a greater sampling of this backazimuth range will be required to confirm this. In the average quoted in Table 2, we have not included this record although adding it does not change the average of $\phi$ by more than $10^\circ$.

The measurements for all stations are plotted in Figure 8. There are three categories of measurements, each denoted by a different symbol. First, those with well constrained observable anisotropy are given by a solid symbol proportional in size to $\delta t$ and a line whose orientation gives $\phi$. Second, there are open symbols where no anisotropy was detected. Given the resolution of our data, this constrains $\delta t$ to be less than about 0.5 s. Third, there are stations that exhibit no splitting but which possess a limited range of $\phi_0$, so that anisotropy is still possible but with only two allowable values of $\phi$.

In all, we have analyzed 21 stations in this study. We have also plotted the results from SSM (LAC, MNV; MNV has been assigned the average value obtained in that study for the northern Basin and Range) and Given and Silver [1990] (NNA, OBN, ARU). While the stations are concentrated in North America and Eurasia, we do have results for all continents, with the exception of Australia and Antarctica. The largest value of $\delta t$ is found for RSON, $\delta t = 1.70$ s, followed by COL with $\delta t = 1.55$ s, while the smallest detectible is $\delta t = 0.65$ s for station RSSD; the average value is 1.0 s.

Characterization of the Anisotropy

While the phases $SKS$ and $SKKS$, with nearly vertical ray paths through the mantle, provide excellent lateral resolution, the depth resolution is poor. One can regard the splitting parameters as representing the integrated effect along the entire mantle path on the receiver side. Information about source side anisotropy is lost through the $S$ to $P$ conversion at the core-mantle boundary (CMB). In certain cases, it is possible to obtain tighter depth constraints. For example, the stations RSON and RSSD are about 1000 km apart at the surface. While each station shows little dependence of ($\phi, \delta t$) on $\phi_0$ as shown in Figures 5 and 6, their delay times represent the largest difference in the data set. The insensitivity to $\phi_0$ argues for homogeneous anisotropy within a cone whose vertex is at the station and whose base...
has a radius of about 1500 km at the CMB. The disparity between these stations means that the anisotropic regions lie outside the volume defined by the intersection of the two cones. This constrains the anisotropy to lie in the upper half of the mantle, above about 1000 km depth. While this is not a very tight constraint, it does rule out significant contributions from such regions as D" as giving rise to the anisotropy. Support for an upper mantle source comes from a recent portable experiment which deployed 22 stations between RSON and RSSD [Silver et al., 1989; P.G. Silver and S. Kaneshima, The APT89 seismic experiment: Shear-wave splitting results, submitted to Geophysical Research Letters, 1991]. Preliminary results for one event, well recorded by most of the stations, shows variations in splitting along the traverse with length scales of a few hundred kilometers. This would suggest that the anisotropy lies well within the upper mantle.

Contribution of the Crust

In order to determine the size of the mantle anisotropy we need to estimate the extent of crustal contribution to the splitting. The presently available evidence suggests that the crustal component is in fact small and can be effectively ignored in most cases. We first note that shear wave splitting measurements that exist for the upper crust [Crampin and Booth, 1985; Savage et al., 1989] and entire crust [Kaneshima et al., 1988; Kaneshima and Ando, 1989; McNamara et al., 1989; McNamara, 1990] yield delay times in the range 0.1-0.3 s. On average this is about a factor of 5 smaller than what we observe. Second, confining the anisotropy entirely to the crust would require unreasonably high effective crustal anisotropy of about 5-16%. In order to produce the typically simple split waveforms we observe, any crustal contribution would have to be coherent with the mantle (i.e., same fast polarization direction) or else it could complicate or even diminish the mantle component.

As discussed by SC, a direct assessment of the total crustal component beneath some stations such as RSON, can be obtained by the use of the radially polarized phase, $P\hat{n}S$ [see Clarke and Silver, 1991] that converts from $P$ to $S$ at the Moho. This is because it is only sensitive to shear wave splitting in the crust. At RSON in particular, this phase is clearly observable and has been used previously to model the crustal structure under this station [Owens et al., 1987]. We have searched for very impulsive events recorded at RSON, for which $P\hat{n}S$ is clearly observed, and the backazimuth is within 5° of where we see the strongest SKS$_T$ arrivals. Shown in Figure 9 are two such events that occurred beneath the Sea of Okhotsk. While $P\hat{n}S$ is observed on the radial components of both events, the transverse com-

Fig. 2b. (Top two traces) Original transverse and radial components. (Bottom two traces) Corrected transverse and radial components. Note that in the corrected seismogram, transverse component SKS has been removed.
Fig. 2c. (Top two traces) Superposition of fast (φ direction) and slow components (φ +90°) uncorrected (left) and corrected (right). (Bottom) Particle motion for fast and slow components uncorrected (left) and corrected (right). Corrections made using estimated values of (φ, Δt).

Components shows no detectable signal during the same time interval. This strongly suggests that the coherent crustal component is no more than about 0.5 s (corresponding to 4% anisotropy throughout the entire crust) and is probably closer to the 0.1-0.3 s range quoted above. Preliminary attempts to measure crustal anisotropy using PnS in the Basin and Range [McNamara et al., 1989] appear to be successful and may lead to a general procedure for separating out the two components.

RELATION OF SPLITTING TO TECTONIC/GEOLeGIC PROCESSES

As mentioned in the introduction, the interpretation of splitting observations in terms of tectonic/geologic processes involves a chain of relations: between splitting and anisotropy, between anisotropy and strain (through LPO), and between strain and tectonic processes. We briefly review the data relevant to these relationships and state the assumptions that will be used to interpret the data.

Splitting and Anisotropy

One of the essential parameters that is required to interpret anisotropic measurements is the parameter δβ in equation (4), as it provides an estimate of the effective layer thickness L. The single-crystal elastic constants for olivine provide an absolute upper bound on the maximum allowable splitting δβmax of about 0.10 [Verma, 1960; Kumazawa and

Fig. 2d. Contour plot of E(φ, Δt). Minimum value shown along with 95% confidence region (double contour) and multiples of that contour level. For symmetric ellipsoidal regions, we take the 1σ confidence level to be half the bounding values of this region along the parameter axes. For asymmetric regions, we use the smallest symmetric region that includes the original confidence region.
Anderson, 1969] for propagation along the crystallographic axes. For certain intermediate propagation directions in the [010] plane, $\delta \beta$ can get as high as 0.17, although the average value in this plane is close to 0.10. The upper mantle, however, is only 50-70% olivine with the rest primarily orthopyroxene (see SC for a discussion), a mineral that is much less anisotropic than olivine. In addition, crystals are far from being perfectly aligned so that the strain-induced LPO of mantle minerals produces an ODF and macroscopic $\delta \beta_{\text{max}}$ that is somewhat below the single-crystal value. The most direct methods for estimating this parameter are by utilizing mantle samples derived from ophiolites, Alpine peridotites, or xenoliths from kimberlite pipes. The two most common methods are petrofabric analysis combined with estimates of the single-crystal elastic constants, and ultrasonic measurements. The most extensive data set is for ophiolites [Christensen, 1984], where $\delta \beta_{\text{max}} \approx 0.04$. This same value is also found for kimberlite xenoliths (D. Mainprice, personal communication, 1990) and will be adopted in the estimation of effective layer thickness. In this case, 1 s of delay time corresponds to 115 km. Thus the range for our data set is 75 km (RSSD) to nearly 200 km (RSON).

The single-crystal elastic constants for olivine show that propagation along [010] and [001] produces a large value of $\delta \beta$ with $\phi$ parallel to [100]. Propagation along [100] would probably not produce observable splitting. That $\phi$ can usually be taken as the direction of [100] for observable splitting (a relation that also appears to hold for aggregates that possess strong [100] preferred orientation [Christensen, 1984]), is of central importance in the interpretation of splitting observations.

**LPO-Induced Anisotropy and Finite Strain**

It is generally accepted that the primary cause of LPO is finite strain, $\epsilon$ [Nicolas and Poirier, 1976; McKenzie, 1979; Nicolas and Christensen, 1987; Ribe and Yu, 1991], and as a result it is possible to regard splitting observations as functions of finite strain as well. There are two relations of interest: one of magnitude, $\delta t = \delta t(\epsilon)$, and one of orientation, $\phi = \phi(\zeta_i)$, where the $\zeta_i$, $(i=1,2,3)$ are the three principal strain directions of $\epsilon$ corresponding to the shortening, intermediate, and extension directions, respectively. The most relevant relation involves the behavior of the [100] axis of olivine. Based on observations of naturally occurring peridotites, laboratory work and theoretical modeling [Nicolas and Poirier, 1976; McKenzie, 1979; Christensen, 1984; Nicolas and Christensen, 1987; Ribe and Yu, 1991], [100] is predicted to be parallel to $\zeta_3$ (extension direction) for nearly all kinds of finite strain: uniaxial extension (pure shear), uniaxial shortening (pure shear, [010] is parallel to
Fig. 3b, while both [100] and [001] are in the extension plane containing \( \hat{e}_2, \hat{e}_3 \), and progressive simple shear [Hobbs et al., 1976; McKenzie, 1979]. This last mode of deformation is appropriate for the differential motion between a plate and underlying mantle, or relative motion along plate boundaries (RPM); in this case [100] is contained within the flow plane, and is parallel to the flow direction [Nicolas and Poirier, 1976]. Since \( \phi \) is generally parallel to [100], we expect \( \phi \) to follow these same rules:

**Mantle Strain and Tectonic Processes**

How does the subcontinental mantle deform in response to various tectonic processes, such as orogenic episodes, extension, large-scale strike-slip faulting or plate motion? As with the case of olivine deformation, there are certain generalizations that may be made regarding continental deformation which we use in the subsequent analysis. (1) We treat collisional episodes as producing pure shear within the subcontinental mantle, with \( \hat{e}_1 \) parallel to the collision direction and both \( \hat{e}_2 \) and \( \hat{e}_3 \) in the plane perpendicular to it. (2) We regard extensional regimes such as the Basin and Range as exhibiting pure shear with \( \hat{e}_3 \) parallel to the extension direction. (3) We assume that both relative (RPM) and absolute plate motion (APM) produce progressive simple shear as a result of differential motion between the plates in the former case and between the plate and a presumed stationary mantle in the latter. Thus, for strike-slip plate boundaries such as the San Andreas Fault, \( \hat{e}_3 \) should be parallel to the fault strike (for large finite strain); for convergent margins, \( \hat{e}_3 \) should be parallel to the relative plate motion vector between the downgoing and overlying plate. Finally, for APM, \( \hat{e}_3 \) should be parallel to the APM direction.

It has been proposed that there is often a large strike-slip component associated with major continental orogenies such as the Hercynian, and that much of the deformation occurs by simple, rather than pure shear [Vauchez and Nicolas, 1991]. This mode of deformation has been referred to as transpression and has also been proposed for the Archean deformation in the Canadian Shield [see Cord, 1990]. In this case, there will be strike-slip faulting that is parallel to orogenic structures. Thus, the strike-slip component will tend to align \( \hat{e}_3 \) perpendicular to the collision direction, as in the pure shear case, although \( \hat{e}_3 \) is restricted to be in the horizontal plane.

**Candidate Processes**

Putting together the above relationships, we are in a position to make predictions of \( \phi \) (for vertical propagation of \( SKS \)) for various sources of mantle strain. This provides us with a means of distinguishing between modes of
deformation, localizing the anisotropy and ultimately identifying the dominant processes. For this purpose, we consider three candidate hypotheses for producing the observed anisotropy. The first is inspired by what is believed to occur in the oceanic upper mantle: namely, that the anisotropy is the result of absolute plate motion, concentrating strain in the asthenosphere but possibly producing strain in the lithosphere as well. This will be referred to as the APM hypothesis. This model predicts that the fast polarization direction $\phi_{\text{APM}}$ is parallel to the APM direction for the plate. It furthermore predicts that $\phi$ should be smoothly varying over the plate; assuming $\delta t$ is solely a function of plate speed, $\delta t$ should be smoothly varying as well.

The second hypothesis is that the present-day stress measured in the continental crust reflects lithospheric-wide stress that ultimately produces strain-induced anisotropy. The global survey of crustal stress measurements by Zoback et al. [1989] provides a means of testing this possibility. The crustal stress (CS) hypothesis is conceptually different from APM in that it does not invoke a particular physical process. Rather, it states that the processes responsible for crustal stress (basal drag from APM, ridge push, mountain-building, rifting...), are also responsible for mantle anisotropy. This hypothesis is more difficult to evaluate than APM, however, because LPO-induced anisotropy is a function of finite strain [Nicolas and Christensen, 1987], not instantaneous stress. Thus one must assume a constitutive relation between the two and unfortunately there is no unique relationship. The most straightforward would be to equate the compression, intermediate, and tension stress axes with the shortening, intermediate, and extension finite strain axes respectively. This would be most appropriate for pure shear in tensional and collisional regimes as the result of membrane stresses (i.e., those derived from forces parallel to the plane of the plate). Using the maximum horizontal stress (maximum compression) direction $\phi_{\text{max}}$ provided by Zoback et al. [1989], the predicted fast polarization direction $\phi_\text{c}$ should then be orthogonal to $\phi_{\text{max}}$. Predictions are much more difficult to make in regions that are deforming by progressive simple shear (strike-slip fault-
ing) and characterized by $\sigma_{\text{max}}$ and minimum horizontal shear stress $\sigma_{\text{min}}$ magnitudes that are comparable. The orientation of the finite strain principal axes would depend on which of the two planes of maximum shear stress has actually failed. An additional ambiguity is introduced if the strains are produced by a basal shear stress as would be the case with APM-related resistive drag forces on the plate. Assuming that this would cause $\sigma_{\text{max}}$ to be in the direction of absolute plate motion, then $\phi_c$ should be parallel to the $\sigma_{\text{max}}$ direction.

The third hypothesis is that mantle anisotropy is dominated by the last significant episode of internal coherent deformation (ICD) of the continental lithosphere by tectonic processes such as orogenies, rifting episodes, and strike-slip deformation. These strains are assumed to be the result of membrane stresses, in contrast to basal shear stresses. It is assumed that the anisotropy is preserved in the lithosphere unless there is subsequent deformation that overprints it. The coherence predicts that the strain field revealed in the structural geology of the Earth’s surface constitutes a reliable measure of mantle strain and thus provides a geologically derived prediction, $\phi$, based on the expected relation between $\phi$ and strain, as discussed above. In particular, we take $\phi_p$ to be perpendicular to the inferred collisional direction (or parallel to the structural trends) for orogenic episodes, and parallel to the extension direction for regions undergoing extension. Finally, for deformation near major strike-slip plate boundaries, $\phi_c$ is predicted to be parallel to the strike of the fault.

The three candidate processes can be grouped in the following way. Both the APM and CS hypotheses imply that anisotropy is due to present day processes, as does the ICD hypothesis in tectonically active areas. However, ICD predicts that “fossil” anisotropy is dominant in regions that are tectonically stable, with the important implication that the anisotropy would then provide a history of deformation in the mantle. APM would predict that anisotropy is localized in the asthenosphere with perhaps some anisotropy in the lithosphere as well. CS and ICD would both argue for lithospheric deformation. With regard to the lithosphere, APM implies that mantle strains are produced by basal shear stresses, whereas ICD predicts that they are caused by membrane stresses. The CS hypothesis would allow for either type of deformation.

**Absolute Plate Motion**

In order to test this possibility, we have computed the APM velocity vectors for the model AM1-2 [Minster and Jordan, 1978] at each of the stations for which anisotropy has been clearly observed. The predictions are shown in Figure 10 and may be compared to Figure 8. For many of the North American stations, the general ENE trend of the ob-
servations appears to correspond to $\phi_{APM}$. The notable exceptions are COL and LAC (SSM). For the other continents, there are several stations that are at a high angle to $\phi_{APM}$: NNA in South America, SLR in Africa, and TOL and GRA1 in Europe. In the case of active continental margins, such as NNA and COL, what is plotted is the $\phi_{APM}$ for the continental plate. One could rightfully argue that it would be more appropriate to consider either $\phi_{APM}$ for the descending plate or $\phi_{RPM}$, corresponding to the relative motion vector between the continental and oceanic plates. If these directions are chosen, however, the discrepancy remains because for both stations, the three directions are roughly parallel. For COL, $\phi_{APM}$ for the Pacific plate is $-26^\circ$ and $\phi_{RPM}$ is approximately NS. Similarly, in South America $\phi_{APM}$ for both the Nazca and South American plates are roughly EW as is $\phi_{RPM}$. A histogram of the angle $\phi - \phi_{APM}$ is shown in Figure 11. As one can see, the agreement between predicted and observed directions is not good. The APM prediction of plate-wide coherence in $\phi$ meets with limited success as well. Of the two plates for which we have a significant amount of data, North America and Eurasia, North America appears to show plate-wide coherence; however, Eurasia does not. This is especially apparent for the stations in China (HIA, WMQ) and the Soviet Union (ARU, OBN).

The APM hypothesis would have to be substantially modified to reconcile these difficulties. First, it would have to be abandoned for tectonic areas, leaving stable areas as the possible domain of applicability. This would work for stable North America, but apparently not for Eurasia. It could be argued, however, that in the case of Eurasia, the APM velocity is so low that (1) the direction may be in error, and (2) APM-related strains may be too small to produce anisotropy. With regard to the first point, a recently proposed absolute plate motion model HS2-NUVEL1 [Gripp and Gordon, 1990], based on the relative plate motion model NUVEL-1 [DeMets et al., 1990], predicts plate motions for Europe that differ in direction by $90^\circ$ from AM1-2. The errors are so large, however, that the differences between models for the Eurasian plate are not statistically significant [Gripp and Gordon, 1990]. Thus, it appears that this issue may not be resolved for some time. Regardless of what the true APM velocities are for Eurasia, lack of coherence of $\phi$ values within this Eurasian plate suggest that APM strains are not dominant. This may be due to the second point, that APM velocities are low. This implies, however, that the only region left for the APM hypothesis is stable North America or other undeformed, fast-moving plates.

Even within stable North America, there are difficulties with this hypothesis. The large variation in $\delta t$ between RSSD and RSON (which are only 1000 km apart) argues that plate speed could not be the only variable controlling $\delta t$. Lithospheric thickness might be proposed as another important parameter. If the anisotropy is localized in the asthenosphere, one would expect anisotropy to be strongest
where the lithosphere is thinnest, as is apparently true in the ocean basins (the observations of anisotropy in the oceanic upper mantle from long period surface waves find the largest anisotropy in the youngest lithosphere near the mid-ocean ridges [Montagner and Tanimoto, 1990]). In contrast, if the anisotropy is the result of the uniform straining of the continental lithosphere, one would expect larger delay times to occur in regions with thicker lithosphere. Since RSON appears to have thicker lithosphere than RSSD [see Grand, 1987], the latter possibility would seem more reasonable.

Present-Day Crustal Stress

Since it is not possible to obtain crustal stress estimates at the locations of the stations analyzed, the comparison with the splitting observations must be done in a more qualitative manner. Such an evaluation can be made by concentrating on particularly coherent regions of crustal stress where interpolation is warranted. From Zoback et al. [1989], the two regions with the most data and most uniform horizontal stress orientations are North America and Western Europe. Assuming that $\phi_c \perp \delta_{h_{max}}$, the expected relationship for a pure shear response to membrane stresses, we find that $\phi_c$ is at a high angle to $\delta$ for all of the stations in stable North America, since $\delta_{h_{max}}$ ranges from NE to ENE; the CS hypothesis is more successful for tectonic areas to the west, especially COL and LON, where $\delta_{h_{max}}$ is nearly NS. For Western Europe, assuming a range of values of $\delta_{h_{max}}$ of NW to NNW, $\phi_c$ should range from NE to ENE. Both of our measurements are E to ENE, as are most of the $\phi$ measurements from $SKS$ splitting obtained at other Western European stations [Vinnik et al., 1989b], and would thus be marginally consistent with this prediction.

If instead we assume $\phi_c \parallel \delta_{h_{max}}$, appropriate for APM-related basal shear stresses, we again meet with mixed success. The CS hypothesis does reasonably well in stable North America as $\phi_c$ and $\phi$ both point in the direction NE to ENE, although it is inconsistent with the tectonic stations
Internal Coherent Deformation

To test the ICD hypothesis, we need to determine the strain field of the last major tectonic episode, as inferred from the surface geology. We have compiled geologic information for each of the stations for which splitting has been detected and for which fairly unambiguous geologic information can be obtained; a predicted fast polarization direction, \( \phi_d \), is determined as given in Table 3 using the relationships between modes of continental deformation and mantle strain. Where precise numerical values were not available, we have converted compass directions as follows: NS: 0°, EW: 90°, NE-SW: +45° ENE-WSW: +67.5° etc. There is ultimately some subjectivity in making such an estimate. However, in most of the areas where anisotropy is observed there is a relatively unambiguous fabric. There are some exceptions to this: first, there is presently insufficient geologic information available to determine \( \phi_d \) for either the Soviet stations (OBN, ARU) or the Chinese stations (WMQ, HIA), and these stations have not been included. For RSSD there is both Archean and Proterozoic fabric of roughly equal strength whose strikes are at right angles to each other. We have reported both directions. For MNV in the Basin and Range we have reported the direction for the present day extension as well as for the somewhat stronger pre-Miocene extension (see SSM). In Figures 12a and 12b are displayed a histogram of \( \delta = \phi - \phi_d \) for our data set. Figure 12a corresponds to the younger deformation, 12b to the older. We note that in most cases there is a striking agreement in the directions. The places where the APM and crustal stress hypotheses appeared to be good predictors of \( \phi_d \), namely eastern North America, are precisely the places in which \( \phi_{apm} \), \( \phi \), and \( \phi_d \) are nearly parallel. In addition, the
obvious discrepancies with $\phi_{opt}$: COL, LAC, NNA, SLR, TOL, and GRA1 are reasonably well accounted for by this model.

The ICD hypothesis also explains some of the variations in $\delta t$ in stable North America. Given equal values of $\delta \beta$ and assuming the anisotropy is in the lithosphere, $\delta t$ should correlate with lithospheric thickness. One plausible measure of thickness, at least in stable continental areas, is the lateral variation in shear wave vertical delay times. Such information is available in North America from the model of Grand [1987]. Plotted in Figure 13 is a contour map of the vertical delay times $\delta T$ (integrated from 400 km depth to the surface) along with the locations of the the five stations in stable North America, as well as portable stations that were deployed along a traverse between RSSD and RSON [Silver et al., 1989; P.G. Silver and S. Kaneshima, submitted manuscript, 1991]. In Figure 14 is a plot of $\delta T$ versus $\delta t$ for these stations. Note that there is indeed a good correlation both for the permanent and portable stations. The existence of such a correlation is consistent with the proposition that the anisotropy has a lithospheric rather than deeper mantle source and that $\delta \beta$ is roughly constant.

Of the three hypotheses, the ICD appears to be the most successful in accounting for the entire data set. Ultimately, the only region where the APM hypothesis has any success is stable North America, a region where the ICD hypothesis is equally successful. The CS hypothesis was the most difficult to evaluate, due to the need to assume a constitutive relationship. No one relation, however, appears to satisfy the data, even for stable areas.

**Specific Stations of Interest**

Some of the stations are of particular interest because of their location and because of their unusual splitting parameters. We therefore discuss RSON and RSSD, which possess the largest and smallest delay times in the data set, COL, which possesses the second largest delay time and is located in a region of large-scale tectonic activity, SLR, which is situated within the Archean Kaapvaal Craton, and GRA1, which shows evidence for two sources of anisotropy. Some of the other stations have been discussed previously by SC.

**RSON: Red Lake, Ontario**

This station, first discussed by SC, yields the largest delay time in our data set and thus warrants special attention. Paradoxically, it is situated in one of the most geologically stable regions in the world: an 800 km by 1000 km block of E-W trending gneiss and greenstone terrains in the Western Superior province of the Canadian Shield; it is also distinguished by unusually fast shear velocities [Grand, 1987]. It was noted in SC that there is an excellent correlation between $\phi$ and the dominant fabric present in both small-scale and large-scale features (see their Figure 4). This region experienced a large-scale episode of north-south compressional deformation at the end of the Archean, formerly referred to as the Kenoran Orogeny. The deformation, which has been extensively dated, took place from 2.73 to 2.70 b.y ago, with the older deformation occurring to the north of RSON and the younger deformation to the south [Card, 1990]. The portable data [Silver et al., 1989] suggest that the large delay times persist about 300 km to the south of RSON.

**RSSD: Black Hills, South Dakota**

RSSD is distinguished by having the smallest delay time in the data set (0.65 s). It is also the closest station to RSON. RSSD sits on the eastern edge of the Wyoming Craton along the boundary between the Craton and the Trans-Hudson orogenic zone. While it took part in late Archean deformation, it has subsequently been exposed to the Proterozoic Trans-Hudson deformation as well as the Laramide Orogeny. The structural fabric in the Black Hills is fairly complex and is dominated by the NE-SW Archean and NW-SE Proterozoic events [Gosselin et al., 1988; Dewitt et al., 1986]. As these two directions are virtually perpendicular to each other, they would tend to cancel each other out. This may partially account for the small $\delta t$. The value of $\phi$ is closer to $\phi_{opt}$ for the Archean direction and would suggest that, at least in the mantle, the Archean deformation is the stronger of the two.

**COL: Fairbanks, Alaska**

Outside of RSON, this station possesses the largest delay time in the data set, ($\phi, \delta t$)=(-82°,1.55 s); it is also situated in the most tectonically active region, the result of a convergent plate boundary. Crustal stress measurements indicate that the convergence produces roughly NS compression [Nakamura et al., 1980; Gedney, 1985; Zoback and Zoback, 1980] and associated strain throughout most of Alaska [Estabrook et al., 1988]. The combination of parallel strike-slip motion along major faults and orogenic features suggests a "transpressional" mode of deformation. The closest features to COL are the strike-slip Denali Fault and
the Alaska Range to the south of COL and the strike-slip Tintina-Kaltag fault system to the north. At the longitude of the station, all features strike approximately EW. This direction is the same as $\phi$, strongly suggesting that the subcontinental mantle is being deformed in the same way.

The thickness of subcontinental mantle is constrained by the presence of the subducting slab. The closest slab-related seismicity is about 50 km to the southwest, and approaches depths of 140 km [Gedney and Davies, 1986], so that the presumed aseismic extension of the slab under COL would have a depth of about 200 km. The delay time $\delta t$ predicts a layer thickness of about 170 km of mantle material and is thus consistent with the deformation of the entire subcontinental mantle. Similar behavior is seen for other convergent plate margins. NNA, on the west coast of Peru (Figure 8) yields a value of $\phi$ that is parallel to the coastal mountains. This same result is seen in two other splitting studies in analogous areas. Using data from portable stations in Columbia, South America recording events within the Bucharamanga nest, Savage et al. [1989] found that estimates of $\phi(-0^\circ)$ were parallel to coastal mountains. Similarly, below Honshu, Japan estimates of $\phi$ from direct $S$ waves coming up vertically from events in the downgoing slab yield values of $\phi$ parallel to the dominant mountain ranges in Honshu (S. Kaneshima, personal communication, 1990). These results are somewhat surprising. The two studies that have modeled flow-induced anisotropy associated with subduction, McKenzie [1979] and Ribe [1989b], both predict large strains and strong anisotropy above the slab, but with $\phi$ parallel to the direction of convergence. Our results suggest that the dominant effect is instead the shortening and deformation within the overlying subcontinental mantle.

**SLR: Silverton, South Africa**

SLR is situated within the Kaapvaal Craton just north of the Witwatersrand Basin. It is only about 20 km from the Premier mine, the oldest Kimberlite pipe in the world (1 b.y. old). This site then provides perhaps the best opportunity to compare the properties of mantle xenoliths with anisotropy. While the rocks from this mine have yet to be analyzed for anisotropy, some mines several hundred km to the south have been analyzed (D. Mainprice and X. Christensen, personal communications, 1990) with values of $\delta \beta_{\max}$ that are consistent with those found elsewhere. The value of 1.1s for $\delta t$ corresponds to a layer thickness of about 120 km. The last large-scale deformational episode to affect this part of the Kaapvaal Craton is related to the the 2.6-2.7 b.y.-old north-south collision between the Kaapvaal
and Zimbabwe cratons that gave rise to the Limpopo belt to the north. Volcanism and deformational features found within the Kaapvaal Craton have been interpreted by Burke et al. [1986] as an Andean-style arc, in which the Craton was shortened and thickened during the collision. The Witwatersrand basin has been interpreted as a foreland basin; its northern boundary marks the southward extent of the shortened block and is characterized by east-west trending thrust faults. The detailed zircon-based dating of sediments in the Witwatersrand basin are consistent with this picture [Robb et al., 1990]. SLR is just at the southern edge of the collisional plateau and the fast polarization direction of 81° is virtually parallel to the trend of the Witwatersrand northern boundary, strongly suggesting that this event was the primary cause of the observed anisotropy.

GRA1: Grafenberg Array, West Germany

As first discussed in SC, φ for GRA1 is nearly east-west. The strongest deformation seen in these rocks are associated with Hercynian deformation [see Behr et al., 1984]. GRA1 is situated on a ENE-WSW trending thrust fault separating the Saxothuringikum and Moldanubikum zones. The close correspondence between φ and φp suggests that this deformation is producing the anisotropy. This region is of special interest as there was a study in Germany of P wave anisotropy using $P_n$ [Bamford, 1977], where the fast propagation direction was found to be $φ_p = 22.5°$. Since this should also be the fast polarization direction for split shear waves, if caused by the a-axis of olivine, then there is a clear discrepancy between the two results. Fuchs [1983] noted that the direction of maximum shear stress, as inferred from the source mechanisms of crustal events, is essentially parallel to the $P_n$ fast propagation direction and concluded that this relation would hold if the crustal stress pattern penetrated into the uppermost mantle and if the a-axis of olivine aligns itself in this direction. The $SKS$ results, corroborated by Vinnik et al. [1989b], would suggest that the $P_n$ results are sampling a superficial feature, while the $SKS$ results are sampling lithospheric-wide, earlier deformation.

A discrepancy between $SKS$ splitting and $P_n$ anisotropy is also found in the Basin and Range. As discussed by SSM, measurements of shear wave splitting yield a value for φ that is about 75° and is more closely associated with the larger pre-Miocene extension ($φ_o \approx 68°$) rather than the present-day extension ($φ_o \approx -60°$). However, splitting results for the crust [McNamara et al., 1989] and $P_n$ observations [Beghoul and Barazangi, 1990] for the uppermost mantle are more consistent with the present extension being the cause. The discrepancy can again be explained by a superficial and relatively weak deformation superimposed upon a deeper stronger one.

DISCUSSION AND CONCLUSIONS

Characterizing Deformation in the Subcontinental Mantle

Shear wave splitting observations constitute an important way in which to characterize subcontinental mantle defor-
TABLE 3. Predicted Values of the Fast Polarization Direction, $\phi_f$, Based on the Surface Geology

<table>
<thead>
<tr>
<th>Station</th>
<th>$\phi_f$</th>
<th>Ref.</th>
<th>Basis</th>
</tr>
</thead>
<tbody>
<tr>
<td>COL</td>
<td>+90</td>
<td>1</td>
<td>Strike of Denali fault zone and Alaska range both of which strike almost exactly EW directly south of COL; maximum compressive stress direction</td>
</tr>
<tr>
<td>RSNT</td>
<td>+45</td>
<td>2</td>
<td>Foliation direction in Yellowknife greenstone belt and adjacent western Granodiorite</td>
</tr>
<tr>
<td>RSON</td>
<td>+83</td>
<td>3,4</td>
<td>Average foliation direction in Archean Granite-Greenstone terrains; strike of subprovince boundaries</td>
</tr>
<tr>
<td>RSSD</td>
<td>+45/-45</td>
<td>5</td>
<td>Dominant Archean/Proterozoic fabric. Both fabrics appear to be of equal strength in surface rocks</td>
</tr>
<tr>
<td>RSCP</td>
<td>+45</td>
<td></td>
<td>Strike of Grenville deformation</td>
</tr>
<tr>
<td>RSNY</td>
<td>+67.5</td>
<td>6</td>
<td>Local strike of Adirondaks</td>
</tr>
<tr>
<td>LAC</td>
<td>-45</td>
<td>7,8</td>
<td>Strike of San Andreas Fault system</td>
</tr>
<tr>
<td>MNV</td>
<td>-67.5/ +67.5</td>
<td></td>
<td>Strike of present-day/ pre-Miocene extension</td>
</tr>
<tr>
<td>TOL</td>
<td>-90</td>
<td>9</td>
<td>Strike of fabric from Hercynian deformation</td>
</tr>
<tr>
<td>GRA1</td>
<td>+67.5</td>
<td>10</td>
<td>Strike of fabric from Hercynian deformation</td>
</tr>
<tr>
<td>MNA</td>
<td>-45</td>
<td></td>
<td>Approximate strike of coastal ranges in Peru</td>
</tr>
<tr>
<td>SLR</td>
<td>+80</td>
<td>11,12</td>
<td>Strike of fabric from 2.6-b.y.-old collision with Zimbabwe Craton</td>
</tr>
<tr>
<td>LON</td>
<td>+90</td>
<td>13</td>
<td>Perpendicular to present-day most compressive stress direction</td>
</tr>
</tbody>
</table>

Where precise numerical estimates are not available, compass directions are converted to numerical values: EW = 90; NS = 0; NE-SW = 45; NWSE = -45; ENE-WSW = +67.5; WNW-ESE = -67.5. Where two values are given (RSSD, MNV) there are two predominant directions.

References are 1, Nakamura et al. [1980]; 2, Henderson [1985]; 3, Beakhouse [1985]; 4, G. Stott (personal communication, 1988); 5, Gosselin et al. [1988]; 6, Whitney et al. [1989]; 7, Zoback et al. [1981]; 8, Eaton [1982]; 9, Julivert et al. [1972]; 10, Behr et al. [1984]; 11, Burke et al. [1988]; 12, Robb et al. [1990]; 13, Zoback and Zoback [1980].

formation and ultimately to examine the dynamics and evolution of continents. As we have shown, the measurements collected thus far are most successfully interpreted as due to strain-induced mantle anisotropy that is dominantly produced by the past and present internal coherent deformation of the continental lithosphere. In particular, the splitting parameters are best predicted by the last significant episode of continental activity (as seen in the structural geology of surface rocks). This applies to appear equally well to presently active areas (such as Alaska) and areas that were last active in the Archean (such as the Canadian Shield and Kaapvaal Craton).

This result is in marked contrast to the ocean basins, where mantle anisotropy appears to be caused by the progressive simple shear accompanying past and present-day plate motions. This is not to say that APM-related strains are completely absent beneath continents. Rather, it suggests that the internal deformation of the plate, a feature of continents but not of oceans, is a very strong source of upper mantle anisotropy. As more shear wave splitting results are obtained from the rapidly expanding number of portable and permanent broadband stations, the full extent of oceanic-style anisotropy in the subcontinental mantle can be more extensively evaluated.

The existence of coherent lithospheric deformation should not be very surprising. Continental breakup and collision, the essential components of the Wilson Cycle, are, after all, interactions between plates. Coherent deformation over the thickness of the continental lithosphere would be a most likely outcome. The long-term survival of this deformation in the form of "fossil anisotropy" [SC] requires, above all, the absence of subsequent deformation. This, in turn, is enhanced by temperatures that are below the critical temperature $T_c$ where creep in mantle materials is expected to be unimportant over geologic time scales. Based on published data for creep in olivine, $T_c$ is about 900°C [Goetze and Kohlstedt, 1973; Nicolas, 1989]. Thus, the depth to...
Fig. 13. Map of North America with contours of shear-velocity delay time obtained by integrating the three-dimensional velocity model of Grand [1987] from 400 km to the surface. Contour interval is 0.5 s. More negative delay times correspond to higher velocities. Also shown are the locations of the five permanent stations in stable North America plus locations of the stations from a portable experiment [Silver et al., 1989; P. G. Silver and S. Kaneshima, unpublished manuscript, 1991] between RSON and RSSD for which splitting has been detected.

Fig. 14. Plot of $\delta t$ versus $\Delta T$ for stations shown in Figure 13. Solid circles are the permanent stations. Triangles denote the results from the portable stations. Solid and open symbols represent broadband and short-period sensors, respectively. Note the good negative correlation between $\Delta T$ and $\delta t$.

The shear wave splitting measurements provide a means for evaluating the ways in which the subcontinental mantle evolves over time. For example, if the inferred mantle deformation can be associated with a dated surface deformation, this effectively dates the mantle material down to $D_o$, as having been present since the deformation. This constraint is similar to that obtained from diamond inclusion data [see Richardson et al., 1984]. Specifically, the anisotropic portion of the subcontinental mantle had to have been present since the time of the deformational event. This rules out formation by subsequent "underplating." Also, the anisotropic layer could not have been formed merely by the conductive growth of the thermal boundary layer. The material added to the lithosphere by this process would have had to participate in the same deformational event and translate coherently with the surface since the event.

Another application of the splitting data is in assessing the relative strength and penetration depth of geologic episodes. For two stations, GRA1 in West Germany and MNV in the Basin and Range, the effect of present-day activity appears to be limited to the crust and uppermost mantle, while the previous episode has ostensibly deformed most of the lithosphere. This would imply that for both
stations, contemporary tectonic activity is not sufficiently strong to produce lithospheric-wide deformation.

**The Orogeny Paradox**

The simplest model for subcontinental mantle deformation based on the ICD hypothesis is that the continental plate is shortened and thickened by orogenies. The expected characteristics of this deformed plate would be a thickened crust and a low-temperature, high (seismic) velocity, thickened lithospheric mantle. Yet, this model raises a serious paradox. Present and recent orogenies do not fit this description. Instead, they commonly have thick crust associated with what appears to be thinned lithosphere: low mantle seismic velocities, high heat flow, and surface volcanism. The two best known examples are north Central Tibet [see Molnar, 1988b and references therein], formed by the collision between India and Eurasia, and the Altiplano [Isacks, 1988], formed by the convergence between the Nazca and South American plates. In fact orogenies appear to leave the mantle hotter, not cooler in general. Scater et al. [1980] have shown that one of the important predictors of surface heat flow is the age since the last orogenic episode, suggesting that orogenies reheat the lithosphere. Where is the thickened subcontinental lithospheric mantle?

To date, efforts to resolve this paradox have focussed on more complex modes of orogenic deformation that serve to remove the cold subcontinental mantle. One way is by arguing that a thickened thermal boundary layer indeed forms, but is removed by the development of a convective instability and replaced by warm asthenosphere [e.g., Houseman et al., 1981; England and Houseman, 1989]. The other is by asserting that the subcontinental mantle delaminates from the crust as a result of interactions with a subducted slab [Bird, 1978, 1988; Isacks, 1988].

Such models, however, are inconsistent with the preserved record of coherent lithospheric-wide deformation suggested by the shear wave splitting results given in the present report. These observations suggest that despite elevated temperatures, the original subcontinental mantle has not been removed and has instead been heated in place. How might we now resolve the paradox? One possibility is that we have misinterpreted the splitting observations. The most direct way of testing this possibility would be to obtain splitting parameters for regions that clearly show this paradoxical character, such as north-central Tibet and the Altiplano. The ICD hypothesis would predict a large value of $\delta t$ with $\phi$ orthogonal to the collision direction. The "mantle removal" models would predict very little anisotropy.

Assuming the ICD hypothesis is correct, another means of resolving this paradox is required. One possibility is the strain heating that must accompany deformation. In order to satisfy the observed properties of the mantle, a temperature increase of several hundred degrees would be required. Assuming that the finite strain is of order unity (consistent with doubling crustal thickness), the orogeny is rapid ($<<100$ m.y.), and all mechanical work is converted to heat, then 100 MPa (1 Kbar) would produce a temperature increase of about $35^\circ-40^\circ$C. Thus orogenic stresses of several hundred MPa would be required. The importance of this effect then hinges on the actual level of orogenic stresses. Although such stresses have been traditionally thought to be somewhat less than 100 MPa [e.g., Molnar and Lyon-Caen, 1988], we note that estimates of lithospheric stress in the literature range from tens of MPa to several hundred. There are several studies, in fact, that argue for NS compressional stresses of several hundred MPa in the North-east Indian Ocean, that appear to be directly related to the collision between India and Eurasia. These are based on evidence for the lithospheric buckling of oceanic lithosphere [McAdoo and Sandwell, 1985; Stein et al., 1989] and on the numerical modeling of the net forces on the Indo-Australian plate [Cloetingh and Wortel, 1986]. In addition, thin viscous sheet models of the formation of the Tibetan plateau [England and Houseman, 1986] appear to require stresses of several hundred MPa. This at least suggests that such stress levels are not implausible. If orogenic stresses are in fact low, then the orogeny paradox remains, and some other mechanism for heating the mantle in place will be required.

**APPENDIX: CALCULATION OF CONFIDENCE REGION**

The confidence region is determined from the sum-of-squares function $S(\phi, \delta t)$ for an $n$-point time series: $E_2(\phi, \delta t)$ in the case of $SKS$ and $\lambda_2(\phi, \delta t)$ in the case of other $S$ phases. If the minimum value of $S(\phi, \delta t)$, $S_{\text{min}}(\phi, \delta t)$ may be regarded as a $\chi^2$ variable with $n$ degrees of freedom, which would be true for a Gaussian white noise process, then the confidence region may be constructed for the $\alpha$ confidence level from the following expression

$$
\frac{\lambda_2}{\lambda_{\text{min}}} \leq 1 + \frac{k}{n-k} S_{b,n-1}(1-\alpha)
$$

[Jenkins and Watts, 1968; Bates and Watts, 1988], where $f$ is the inverse of the $F$ probability distribution, $k$ is the number of parameters, two in our case, and $\alpha = 0.05$ corresponds to the 95% confidence level. If $S(\phi, \delta t)$ is not distributed $\chi^2$, appropriate modifications must be made. We assume that $S(\phi, \delta t)$ is approximately $\chi^2$-distributed where $\nu$, which depends both on the instrument response and noise spectrum, is usually much smaller than $n$. Parameter $\nu$ is found in the following way. We define

$$
E = \frac{T}{2} \int_{-\infty}^{\infty} \eta^2(t) dt (A1)
$$

where $\eta(t)$ is a finite-length sample of the noise process of length $T$. It is assumed that we can write $\eta(t)$ as the convolution of an uncorrelated Gaussian noise process $g(t)$, and a filter, $f(t)$. By Parseval's theorem, (A1) can then be written in the frequency domain as

$$
E = \frac{T}{2} \int_{-\infty}^{\infty} \left| f(\omega) \right|^2 \left| g(\omega) \right|^2 d\omega (A2)
$$

For digital data, (A2) may be approximated by a finite sum of terms up to the Nyquist frequency and with frequency spacing $1/T$:

$$
E \approx \sum_{n=1}^{N-1} \left| f_n g_n \right|^2 + \frac{1}{2} \left( \left| f_n \right|^2 \left| g_0 \right|^2 + \left| f_N \right|^2 \left| g_N \right|^2 \right) (A3)
$$

If $E$ is approximately $\chi^2$ distributed then we may estimate the degrees of freedom from the expression

$$
\nu = 2 \frac{<E>^2}{\text{Var}[E]} (A4)
$$

[Jenkins and Watts, 1968]. From the definition of $E$, and assuming $g_n$ has zero mean and variance $\sigma^2$, we can write
\[< E > = \sigma^2 \left( \sum_{n=1}^{N-1} |f_n|^2 + \frac{1}{2} (|f_0|^2 + |f_1|^2) \right) \equiv \sigma^2 F_2 \quad (A5)\]

Since
\[< E^2 > = \sigma^2 (F_2^2 + F_4), \quad (A6)\]

it follows that
\[\text{Var}[E^2] = \sigma^4 F_4 \quad (A7)\]

where
\[F_k = \sum_{n=1}^{N-1} |f_n|^k + \frac{1}{2} (|f_0|^k + |f_1|^k) \quad (A8)\]

In (A6), the assumption that the real and imaginary parts of \(g_n\) are independent and normally distributed has been used in computing its fourth moment. Substituting these results into (A4),
\[\nu = 2 \frac{F_2^2}{F_4} \quad (A9)\]

Finally, it is necessary to determine the quantities \(F_2^2\) and \(F_4\) in (A9). One could accomplish this by knowledge of the instrument response and assumptions about the noise process. We have chosen, however, to determine this directly from the data. In analogy with the definition of \(E\), we define the frequency domain sum
\[E_k = \sum_{n=1}^{N-1} |f_n|^k |g_n|^k + \frac{1}{3} (|f_0|^k |g_0|^k + |f_1|^k |g_1|^k) \quad (A10)\]

whose expected value is
\[< E_k > = 2 \sigma^4 F_k \quad (A11)\]

Consider the quantity
\[\tilde{\nu} \equiv 2 \frac{F_2^2}{E_4} - 1 \quad (A12)\]

Writing \(E^2 = < E^2 > + \delta E^2\) and \(E_4 = < E_4 > + \delta E_4\) and using expressions (A6),(A9), and (A11), we find that to first order in \(\delta E^2/< E^2 >\) and \(\delta E_4/< E_4 >\), \(\tilde{\nu}\) is an unbiased estimate of \(\nu\). Since both \(E^2\) and \(E_4\) can be determined from the data, we can estimate \(\tilde{\nu}\) for each record and then take the average value of several records, assuming a stationary noise process. Expressed as the ratio \(\tilde{\nu} = \nu / n\), it is found that \(\tilde{\nu}\) is typically 0.3 for RSTN intermediate channel and about 0.12 for DWSSN intermediate channel and for NARS. In general, this comes out to one degree of freedom per second for all of the networks, and thus appears to be independent of digitizing interval for the data we are using.

Acknowledgments. We thank S. Shirey, T. Kusky, M. Kingston, A. Rizza, and J. Boyd for help with the geology of various stations, N. Hibe, P. Tapponnier, and Ines Cifuentes for useful discussions, C. Bina for help with a contour plotting routine, and S. Grand for providing the vertical delay times for his model. Helpful reviews were provided by A. Nicolas, S. Kaneshima, M. Savage, N. Valette-Silver, G. Helffrich, D. Wiens, and an anonymous reviewer. We also thank M. Acieno for providing computer support and J. Dunlap for assistance with manuscript preparation. This research was sponsored by the National Science Foundation grant EAR87210001 and by the Carnegie Institution of Washington, Department of Terrestrial Magnetism.

References


W. W. Chan, Teledyne Geotech Alexandria Laboratories, 314 Montgomery Street, Alexandria, VA 22314.

P. G. Silver, Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road NW, Washington, DC 20015.

(Received August 2, 1990; revised March 11, 1991; accepted November 19, 1990.)