1. Background

The purpose of these notes is to show how body waves, and in particular, the travel times of body waves, can be used to determine 3-dimensional structure in the Earth. These notes mainly consider "ray theory" which is a high frequency approximation to wave propagation and is invalid for non-geometrical optics effects such as diffraction. In the past couple of years, so-called finite-frequency kernels have become more common and it is becoming increasingly common to model complete waveforms. We may discuss these aspects later if we have time. Even if ray theory were perfectly adequate, there are many difficulties with recovering a good high-resolution image of Earth structure. We shall try and give the reader a feel for the problems involved.

Fig. 1. A 3 component seismogram illustrating the arrival of the P, PP and S etc. body waves (we shall define what these labels mean later). Usually we measure motion in the vertical, N-S, and E-W directions. In this example, the horizontal components have been rotated to a coordinate system in which one component is in the direction from source to receiver (the "radial" direction) and the other is perpendicular to it (the transverse direction).

To make sure everyone is on the same page. At high frequencies, there are two kinds of body waves: compressional or P waves and shear or S waves. The particle motion of these waves are shown in figure 2. In a slightly anisotropic model, there are two S waves which can travel at slightly different speeds leading to the phenomenon of S-wave splitting. Mostly we shall be concerned with isotropic media.
where there is no splitting.

\[ V_p = \sqrt{\frac{K_s + \frac{4}{3} \mu}{\rho}} \]  

where \( K_s \) is the adiabatic bulk modulus, \( \rho \) is the density and \( \mu \) is the "rigidity" of the material. The velocity of shear waves (denoted \( V_s \)) is given by

\[ V_s = \sqrt{\frac{\mu}{\rho}} \]
\( \mu \) is zero in a fluid so shear waves are not transmitted through a fluid. We define \( \phi \), the seismic parameter, by

\[
\phi = \frac{K_s}{\rho} = V_p^2 - \frac{4}{3}V_s^2
\]  

(3)

The square root of \( \phi \) is called the bulk sound speed and is often used in tomographic inversions because it is purely due to compression effects (unlike compressional velocity which is a mixture of shear and compression).

For a physically realistic material \( V_p \) is always greater than \( V_s \). We can show this by noting that the bulk modulus is always positive so

\[
V_p = \sqrt{\frac{K_s + \frac{4}{3}\mu}{\rho}} > \sqrt{\frac{\frac{4}{3}\mu}{\rho}} > \sqrt{\frac{\mu}{\rho}} = V_s
\]  

(4)

Therefore \( V_p > V_s \), i.e., compressional waves travel faster than shear waves.

After an earthquake, an approximately spherical wavefront (either P or S) propagates out and we can think of rays as being lines perpendicular to the wavefront. As the wavefront meets material of higher (or lower) seismic velocity it will speed up (or slow down) and so will get distorted. An equivalent statement is that the rays are bent. In the geometrical optics limit, this ray bending follows Snells Law which allows us to get simple equations for interpretation of travel times (see below). When a wavefront meets a "sharp" interface (sharp here refers to a change in properties which is fast compared with the wavelength of the seismic wave), we can get reflection, refraction and conversion of seismic waves. "Conversion" means that a P wave incident on an interface can be converted to an S wave and vice-versa. This means that seismic energy can get from a source to a receiver by a wide variety of paths – each path is labelled according to the number and type of ray legs in the mantle and core. Here is an example:

![Fig. 3.](image)

Suppose that the ray consists of compressional wave motion in the mantle legs of the ray and, because the outer core is fluid and does not sustain shear waves, the leg of the ray in the outer core must also be compressional wave motion.

We use the following notation

<table>
<thead>
<tr>
<th>Refracted rays</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P )</td>
<td>compressional in mantle</td>
</tr>
<tr>
<td>( S )</td>
<td>shear in mantle</td>
</tr>
<tr>
<td>( K )</td>
<td>compressional in the outer core</td>
</tr>
</tbody>
</table>
There is compressional in the inner core
shear in the inner core

Reflected rays

c reflection from the core-mantle boundary

i reflection from the inner core boundary

The ray in the diagram is called PKP. Here are two other examples

An underside reflection does not have a symbol so a P-wave which bounces off the surface once between the source and receiver is called PP if it does not enter the core, e.g.,

(You may be wondering what distinguishes PP from pP. A ray leg that emerges upwards from the source is given a lower case letter while a ray leg that emerges downwards from the source is given an upper case letter.)

A plot of the travel times of a variety of phases is shown in figure 6. (An individual arrival is usually called a seismic ”phase”). Some of these phases are more easily observed than others and have been
used extensively in seismic tomography. The way to read this plot is to imagine a seismic station at a certain epicentral distance from the source (say 60 degrees) then the arrival times of pulses arriving at the station can be read off by looking at where the curves intersect a vertical line at 60 degrees. Often, we measure the differential arrival times between two phases arriving on the same seismogram. These differential times have some advantages in that they tend to be less sensitive to uncertainties in source location (obviously, they are insensitive to the origin time of the event) – we shall discuss some of these later.

Fig. 6. A global set of travel times
2. Measuring travel times

Historically, seismograms were recorded either at "long" periods or "short" periods. The reason for this is that a major source of motion of the ground is the "microseisms" which are due to nonlinear interactions of ocean waves causing pressure variations on the ocean floor. Microseisms have a main peak at 14 second period and a secondary peak at 7 seconds. It is actually the secondary peak that is mainly seen on seismometers. In the past, seismic recording systems did not have the dynamic range to record both the microseisms and the small seismic signals which ride on them. Thus instruments were designed to see periods shorter than 7 seconds (usually peak response at about 1 second) or periods longer than about 15 seconds. Modern seismic recording systems have enough dynamic range to be able to record the microseisms (so-called broad-band recording) but, for most earthquakes, we must still filter out the microseisms so we can see the small seismic signals.

On the short period side, body waves of dominant period 1 second are seen and the first arriving P wave can be accurately picked. Scattering by short-wavelength heterogeneity causes large "codas" which can obscure secondary arrivals. Many observations have been made and are collected by the International Seismological Commission (ISC) which has used them to make a more comprehensive tabulation of earthquake locations. Such data are also used in tomography. P wave tomography using the ISC data has been quite successful but the S waves are more problematic. This is because S waves typically have a lower frequency content due to attenuation and are more poorly recorded by short period instruments. Furthermore, ISC picks are usually made from vertical component instruments so interference of S by the SKS phase at distances beyond 80 degrees is a problem. This makes it very difficult to image S velocity in the lowermost mantle from ISC S picks.

Long-period data offer some advantages over the ISC data – in particular, codas from scattering are nearly nonexistent so later phases can be accurately picked. One picking algorithm is discussed in detail in the paper by Bolton and Masters (2001) and we shall give some demonstrations during the lecture. Differential times can also be picked but require corrections for relative attenuation and, sometimes, corrections for waveform distortion due to propagation effects. Again, we shall give some examples in the lecture.

3. Ray theory

For simplicity, we start by considering a flat Earth. Cast your mind back to those physics labs where you shot beams of light into blocks of glass. Consider what happens to a ray of light passes from a medium of refractive index $n_1$ to a medium of refractive index $n_2$.

![Fig. 7](image)

Snell’s Law gives
\[ n_1 \sin \theta_1 = n_2 \sin \theta_2 \]

For seismic waves Snell’s Law is

\[ \frac{\sin \theta_1}{V_1} = \frac{\sin \theta_2}{V_2} \]

where \( V_1 \) is the velocity in medium 1 and \( V_2 \) is the velocity in medium 2. If we define the slowness, \( u \), to be the reciprocal of the velocity, \( V \) we have

\[ u_1 \sin \theta_1 = u_2 \sin \theta_2 \]

(In fact the refractive index of a material is the speed of light in vacuum divided by the speed of light in the medium so \( n \) is a normalized slowness.) If we now have a stack of homogeneous layers with velocity increasing in each layer we find:

\[ u_1 \sin \theta_1 = u_2 \sin \theta_2 = u_3 \sin \theta_3 = \cdots \]

In a continuous medium we get

\[ u_1 \sin \theta_1 = u_2 \sin \theta_2 = u_3 \sin \theta_3 = \cdots \]

If \( \theta_0 \) is the take-off angle of the ray and \( u_0 = 1/V_0 \) is the surface slowness then \( u_0 \sin \theta_0 = u \sin \theta \) and is constant along the ray. \( u \sin \theta \) is called the ray parameter, \( p \), and a ray is uniquely specified by its ray parameter. At the “turning point” of the ray, when \( \theta = \frac{\pi}{2} \) we have

\[ p = u_0 \sin \theta_0 = u_{TP} \sin \frac{\pi}{2} = u_{TP} \]
Therefore $p$ is also the slowness at the turning point of the ray. Clearly if we can measure $p$, we have the velocity at the turning point. How might we go about measuring $p$? Consider a wave front propagating towards the surface in a uniform velocity medium:

From the diagram $\Delta = \delta x \sin \theta$ but we also have $\Delta = \delta t / u$ where $u$ is the slowness of the medium. Therefore

$$\delta t = \delta x u \sin \theta$$

or

$$\frac{\delta t}{\delta x} = u \sin \theta = p$$

By measuring the difference in time of arrival of a wavefront at two stations of known distance apart, we can measure $p$. This argument is only valid for a uniform velocity medium. If the velocity varies with depth then the argument generalizes to $p = dT/dX$, the slope of the travel-time curve.

The travel time curve is a plot of the observed time it takes for a particular ray to travel from the source to the receiver as a function of distance of that receiver from the source. The slope of this curve is the ray parameter. In general $p$ varies along the travel time curve which means that a different ray is responsible for the arrival at each distance.

Now, all of the previous development is actually valid for a flat Earth where $p$ has units of slowness (seconds/kilometer). The argument has to be modified slightly for a spherical Earth because the vertical
(or radial direction on a spherical Earth) varies as we go along the ray (Fig. 12)

Fig. 12 Ray geometry for spherical shells of constant velocity. Note that \( \theta_2 \) isn’t equal to \( \theta'_2 \) but \( r_1 \sin \theta_2 = r_2 \sin \theta'_2 \)

The equivalent formula for ray parameter (still constant along the ray) is now

\[
p = \frac{r}{v} \sin \theta
\]

and \( p = dT/d\Delta \) where \( \Delta \) is epicentral distance. \( p \) is now measured in seconds/radian.

Consider a ray through the Earth as shown in Figure 13.

Fig 13

Now focus on the small segment of the ray which subtends the angle \( d\xi \) at the center of the Earth (Fig.
\( i \) is the angle the ray makes with the vertical and we know that the ray parameter is related to \( i \) by

\[
p = \frac{r}{v} \sin i
\]

where \( v \) is the velocity for the ray segment at radius \( r \). The travel time of the ray is

\[
T = \int_{\Gamma} \frac{1}{v} \, d\Gamma
\]

where the integral is taken along the ray path. Fermat’s principle states that a ray path between two points is a path of stationary time. Thus the travel time will not change to first order if the ray path is slightly perturbed. If we make a small perturbation in velocity structure there will be a change in travel time due to the change in velocity structure and due to the change in the ray path but the latter term is of second order and so can be neglected. We can therefore differentiate equation 5 giving:

\[
\delta T = - \int_{\Gamma} \frac{\delta v}{v^2} \, d\Gamma
\]

From figure 14, we have

\[
\sin i = \frac{rd\xi}{d\Gamma}\quad \text{so} \quad d\Gamma = \frac{r^2}{pv} d\xi
\]

and we can rewrite equation 6 as an integral over distance:

\[
\delta T = \int_{0}^{\Delta} G(\xi) \delta v(\xi) \, d\xi
\]

where

\[
G(\xi) = - \frac{r^2}{pv^3}
\]

This kernel is evaluated by keeping track of the depth of the ray for every position of arc length \( \xi \). To do this we need an equation relating \( \xi \) to \( r \). Reconsider Figure 14 and note that (using the equation for \( p \))
\[ \frac{r}{dr} \frac{d\xi}{dr} = \tan i = \frac{\frac{vp}{r}}{\left(1 - \left(\frac{vp}{r}\right)^2\right)^{\frac{1}{2}}} \]  

(9)

or

\[ \frac{d\xi}{dr} = \frac{p}{r} \left(\frac{r^2}{v^2} - p^2\right)^{-\frac{1}{2}} \]  

(10)

On an aspherical Earth where we have used a local block parameterization, we step finely along in distance starting from a specific source position to a specific receiver position. At each point, we compute the radius we are at using equation 10 and then evaluate the kernel using equations 7 and 8. We also keep track of which block we are in at each step along the ray then integrate the contributions to each block at the end.

4. The 1D Earth

When velocity is a function of radius only, integrals along the ray path for travel time, \( T \), and epicentral distance, \( \Delta \), can easily be converted into integrals over radius. For a surface source and receiver, equation 10 becomes

\[ \Delta = 2p \int_{r_p}^{R} \frac{1}{r} \left(\frac{r^2}{v^2} - p^2\right)^{-\frac{1}{2}} dr \]  

(11)

where \( r \) is radius, \( v \) is velocity and \( r_p \) and \( R \) are the turning point radius and surface radius respectively. The factor of 2 arises from the contribution of the downgoing and upgoing legs. From fig 14., the time, \( dT \), it takes to travel the distance \( d\Gamma \) is just \( d\Gamma/v \) and we have the relation \( d\Gamma = (r^2/pv)d\xi \) so we can get an expression for \( dT/dr \) by multiplying equation 10 by \( r^2/(pv^2) \) and then integrating to give

\[ T = 2 \int_{r_p}^{R} \frac{r}{v^2} \left(\frac{r^2}{v^2} - p^2\right)^{-\frac{1}{2}} \]  

(12)

Given a velocity distribution as a function of radius, you can now compute \( T \) and \( X \) for any particular choice of ray parameter. This is how we compute the curves shown in fig. 6.

You will often here the statement that a ray is most sensitive to the structure near its turning point (where \( \sin(i) = 1 \) so \( p = r/v \)). Consideration of the above equations shows that the integrands would actually become infinite at the turning point (certainly a strong sensitivity!). Don’t panic! This is an example of a square root singularity which is integrable. For example, consider the integral

\[ F = \int_{0}^{2} \frac{1}{\sqrt{y}} dy \]

Even though the integrand is infinite when \( y = 0 \), the area under the curve is finite (what is it?). The sensitivity to structure at the turning point is why we sometimes bin travel time residuals at ray turning points to see if we can see patterns in the data.

5. The inverse problem for the spherical harmonic basis (optional)
If we have parameterized the model laterally in spherical harmonics, the following algorithm gives an efficient way of computing the matrix elements. Suppose we represent the velocity perturbation by:

$$\delta v(r, \theta, \phi) = \sum_{s,t} \delta v^t_s(r) Y^t_s(\theta, \phi) \tag{13}$$

[s = 0 gives the perturbation to the spherically averaged Earth]. We now have:

$$\delta T = \sum_{s,t} x_s \int x_r G(\xi) \delta v^t_s(r) Y^t_s(\theta, \phi) d\xi \tag{14}$$

where \(r, \theta, \phi\) are points along the ray path and \(x_s\) is the source position and \(x_r\) is the receiver position.

Equation 14 is easy to evaluate if we transform the coordinate system (i.e., move the North pole) so that the ray is in the equatorial plane. Spherical harmonics in one coordinate system transform to spherical harmonics in another according to a formula of the form:

$$Y^t_s(\theta', \phi') = \sum_m B_{mt} Y^m_s(\theta, \phi) \tag{15}$$

Note that the harmonic degree, \(s\), is not included in the sum so that, for example, a degree 2 harmonic in the rotated coordinate system is a linear combination of the 5 degree 2 harmonics in the original coordinate system. The computation of the transformation matrix, \(B_{mt}\), is tedious but straightforward (Edmonds, 1960, ch.4) and an efficient algorithm for its computation is given in Masters and Richards-Dinger (1999).

Using equations 13 and 15, we can easily rotate coordinate systems:

$$\delta v(r, \theta', \phi') = \sum_{s,t} \delta v^t_s(r) Y^t_s(\theta', \phi') = \sum_{s,t} \delta v^t_s(r) \sum_{m=-s}^s B_{mt} Y^m_s(\theta, \phi)$$

If we let

$$\delta v^m_s = \sum_t \delta v^t_s B_{mt}$$

we get

$$\delta v(r, \theta', \phi') = \sum_{s,m} \delta v^m_s(r) Y^m_s(\theta, \phi)$$

thus changing the coordinate system gives us the same form as equation 13 for \(\delta v\) but with new expansion coefficients \(\delta v^t_s\). If we rotate the ray path to the equatorial plane we have

$$Y^t_s(\theta, \phi) \rightarrow Y^t_s(\frac{\pi}{2}, \xi) = X^t_s(\frac{\pi}{2}) e^{it\xi}$$

where \(X^t_s\) is now a constant along the ray path. In the rotated system, equation 14 is

$$\delta T = \sum_{s,t} X^t_s(\frac{\pi}{2}) \int_0^\Delta G(\xi) \delta v^t_s(r) e^{it\xi} d\xi \tag{16}$$

where \(\delta v^t_s = B^{-1}\delta v_s\)
The integral can be evaluated numerically providing we keep track of \( r(\xi) \) (by integrating \( dr/d\xi \)). In fact it is most efficient to calculate tables of the form

\[
I_s^t(\Delta) = \int_0^\Delta G(\xi) \delta v_s^t(r) e^{it\xi} d\xi
\]

If we define \( B' = B^{-1} \) then our equation becomes

\[
\delta T = \sum_{s,t} X_s^t \left( \frac{\pi}{2} \right) \sum_m B_{t,m}' I_s^m(\Delta)
\]

which can be rapidly evaluated.

When we come to doing the inverse problem, we would expand the coefficients for velocity, \( \delta v_s^t(r) \) in a set of basis functions (e.g. B-splines or shells) so we can write

\[
\delta v_s^t(r) = \sum_k a_{ks}^t g_k(r)
\]

where the \( g_k \)'s are some known functions and the \( a \)'s are some constants. Then

\[
I_s^t(\Delta) = \sum_k a_{ks}^t \int_0^\Delta G(\xi) g_k(r) e^{it\xi} d\xi = \sum_k a_{ks}^t J_{ks}^t(\Delta) \quad \text{say}
\]

and our equation for a travel time residual becomes:

\[
\delta T = \sum_k a_{ks}^t \sum_{s,t} X_s^t \left( \frac{\pi}{2} \right) \sum_m B_{t,m}' J_{ks}^t(\Delta)
\]

which is in a form suitable for inversion.

### 6. Moving discontinuities

So far, we have been considering the effect of a “volume perturbation” in velocity. There may also be perturbations in the levels of discontinuities which, if the velocity is different on either side, produce travel time anomalies. The formula for \( \delta T \) for a transmitted ray if a boundary is moved by \( \delta r \) is

\[
\delta T = -\frac{\delta r}{r} \left[ \left( \frac{r^2}{v^2} - \frac{p^2}{r^2} \right) \right] ^+ \quad (17)
\]

where \([f]^+\) indicates the value of \( f \) below subtracted from the value of \( f \) above the discontinuity. For a reflected ray, we get

\[
\delta T = -\frac{2\delta r}{r} \left( \frac{r^2}{v^2} - \frac{p^2}{r^2} \right) ^{\frac{1}{2}} \quad (18)
\]

for a topside reflection (so \( v \) is the velocity just above the discontinuity) and

\[
\delta T = \frac{2\delta r}{r} \left( \frac{r^2}{v^2} - \frac{p^2}{r^2} \right) ^{\frac{1}{2}} \quad (19)
\]
for a bottomside reflection (so $v$ is the velocity just below the discontinuity). Note that the perturbation in the radius of a discontinuity can also be expanded in spherical harmonics:

$$\delta r = \sum_{s,t} \delta r^s_{t} Y^s_{t}(\theta, \phi)$$

and, in general, there are different expansion coefficients for each discontinuity in the model.

7. Importance of earthquake location in tomography

It turns out that our (in)ability to locate earthquakes accurately means that we have a source of noise in our tomographic problem which can rival the signal from 3D structure (at least for P-wave tomography). We can estimate the uncertainty due to event mislocation by considering the following equation

$$\delta t = \frac{\partial t}{\partial x} \delta x + \frac{\partial t}{\partial y} \delta y + \frac{\partial t}{\partial z} \delta z + \delta t_0,$$

where $\delta x, \delta y, \delta z$ are errors in event location, $\delta t_0$ is the error in origin time, and $\delta t$ is the resulting error in travel time. Analyses of mislocations of events located by independent means leads to estimates of the length of a typical mislocation vector, $\epsilon_X$, of $\approx 14–18$ km. To convert this number to a typical change in epicentral distance, we assume that the stations are uniformly distributed around the event so that stations in a direction perpendicular to the mislocation vector see no change in epicentral distance while stations in the direction of mislocation will see the full value. Assuming a cosinusoidal dependence as a function of azimuth suggests that, on average, the error in epicentral distance is $\approx \epsilon_X / \sqrt{2}$. Assuming that $\delta x$ and $\delta y$ do not co-vary (as suggested by an analysis of the differences of the NEIC and ISC locations), the error in the travel time due to the error in each of $x$ and $y$ is

$$\frac{p \epsilon_X}{\sqrt{2}},$$

where $p$ is the ray parameter. It is well known that errors in depth and origin time do covary with $\delta t_0 \approx \delta z/9$ (depth in kilometers). Since $\partial t / \partial z$ is negative, the errors in origin time and depth tend to cancel in their contribution to the total error and the errors in $x$ and $y$ dominate the error budget. We now assume a typical depth uncertainty of about 10 km and find that $\sigma_X$ is $0.6–1.2$ seconds for $P$ waves at epicentral distances of about $70^\circ$ for mislocation vectors of length 10–20 km. The corresponding estimate of $\sigma_X$ for $S$ waves is $1.6–2.5$ seconds. As we shall see below, these numbers rival the signals from 3D structure.

These results mean that we cannot ignore earthquake mislocation in our tomography and we must either relocate events or make our data insensitive to event location. Consider the travel time residuals for one event:

$$\delta t = A \delta h + B \delta v$$

where $\delta h$ is a 4-vector of earthquake mislocation terms and $\delta v$ is either a vector of velocity perturbations for a block inversion or a vector of coefficients in the expansion defining the velocity model in a global basis inversion. We could iteratively solve this equation first by relocating the events then solving separately for velocity structure then relocating again but now including the new velocity structure. Convergence is usually attained after a few iterations. Alternatively, we can seek linear combinations of the data for each event which, to first order, are insensitive to the event location. This reduces to finding $P$ such that
\[ P\delta t = PA\delta h + PB\delta v = PB\delta v \]
i.e., we want \( PA = 0 \). Note that if \( A \) has the SVD \( A = U\Lambda V^T \) then \( P = G(I -UU^T) \) where \( G \) is any matrix. We choose \( G \) so that the new data \( \delta t' = P\delta t \) are statistically independent. If \( \delta t \) has a covariance matrix \( I \) then \( \delta t' \) has covariance matrix \( G(I-UU^T)G^T \) (since \( I = (I-UU^T)^T \) and \( (I-UU^T)(I-UU^T) = (I-UU^T) \)). Thus, if \( I -UU^T \) has the eigenvalue decomposition \( R\Omega R^T \) then choosing \( G = \Omega^{-\frac{1}{2}}R^T \) leads to the desired covariance matrix which is \( I \). It is interesting that the eigenvalues of \( I -UU^T \) are one or zero and we lose four eigenvalues during the projection process – we have effectively used up four data to remove sensitivity to location.

The alternative process is to relocate initially, solving

\[ \delta t = A\delta h \]

which, if we have used a SVD would lead to a mislocation vector

\[ \delta\hat{h} = VA^{-1}U^T\delta t \]

and equation 21 would become

\[ \delta t - A\delta h \equiv (I-UU^T)\delta t = B\delta v \]

Note this is similar to the projection method where \( P = G(I-UU^T) \) except that we have not taken account of the fact that we have "used some data up" in doing the relocation and we have not consistently operated on the \( B \) part of equation 21 as we did with the projection method. Ignoring these niceties does still leave us with a \( B \) matrix which is sparse whereas, in the projection method, each new travel time is a linear combination of all the travel times for that event so that \( PB \) is no longer as sparse as we would like.

8. Interpretation of long-period body wave travel times

We are almost ready to invert our travel times – ideally, we would like to incorporate both short- and long-period data in our inversions but, first, we must think about the relation between the two.

In our previous tomographic studies, we have noted a significant offset between the travel times of long-period data and those of traditional 1Hz waves. Our P data are offset from the predictions of PREM (using NEIC locations) by about 4 seconds. This is reduced to less than 2 seconds if we use consistently determined locations and 1D model (e.g. the EHB locations and model AK135: Engdahl et al, 1998, Kennett et al, 1995). This remaining offset is a clear function of moment of the event. Fig 15 shows the mean travel time residual per event binned by moment of the event (red line). There is a clear linear relationship between the offset and log moment, and the offset is almost zero at a moment of \( 10^{24} \) dyne cm – this implies agreement with the ISC data from which AK135 was built. If we correct for lateral heterogeneity using a 3D model, we find that AK135 is too fast by about a second – this is because of the preponderence of fast paths in the ISC data set – and we have an offset of 1 second at a moment of \( 10^{24} \) dyne cm (black line in fig. 15).

The linear relationship is not that expected from the usual source scalings (e.g. Kanamori and Anderson, 1975) where source time is expected to go like \((\text{moment})^{1/3}\). This relationship is clearly seen in the centroid times of the CMT solutions. If we plot the centroid time (corrected for a 2.5 second shift to account for the shift induced by using NEIC locations with the PREM model) versus log moment, we find the expected \((\text{moment})^{1/3}\) variation with a 1 second shift at a moment of \( 10^{24} \) dyne cm (Ekstrom,
Figure 15. Mean residual offset (relative to AK135) per event plotted as a function of event moment. The red line is for raw P data using the EHB locations. The back line is for P data but corrected for 3D structure. The green line is the empirical (moment)\(^{1/3}\) expected for the source process time. The blue line is the same as the black line but now for S.

personal communication). This relationship is plotted on Fig. 15 as the green lines. Clearly, our measured offset agrees with the predicted shift from the centroid times at relatively low moments but diverges for the bigger events. This behavior is to be expected when looking at band limited data and is accurately predicted using measurements from synthetic seismograms including source time functions based on the CMT centroid times.

If we empirically correct for the source offset and for 3D structure, we again find that AK135 gives quite a good fit to the long-period teleseismic P data except for systematic deviations at both long and short distances (Fig 16). Note that deep events behave a little differently and a different source time/moment relationship is needed that predicts shorter offsets than for shallow events implying that deep events are systematically "faster" than shallow events.

Also shown on Fig. 15 are the average S residuals per event plotted as a function of binned moment (blue line). As perhaps should be expected, the slope of the relationship is the same as that for P but there is an offset. Clearly, differential S-P times are insensitive to source process time (to first order) and can be used to isolate other effects. For example, the difference between S and P shown in Fig. 15 may be due to physical dispersion or it may be due to imperfections in AK135.

The reader may be wondering what all these source effects have on the recovery of 3D structure. All successful tomographic inversions include some treatment for the relocation of earthquakes (as discussed above) and relocation takes care of much of the effect we have just described. Even a directivity effect can be largely modelled by a relocation of the event. Of course, the locations derived from short and
Figure 16. Average residuals relative to AK135 after correction for 3D structure and an empirical source size correction (taken from the black line in Fig. 2). The data have been grouped by event depth. Black (0–15km); red (15–35km); blue (35–55km); green (55–300km); gold (300-800km).

long period data will be different and so separate locations must be determined.

9. Ray coverage and patterns in the data

See paper by Masters et al. (2000).