Chapter 13

Applications of Inverse Theory to Solid Earth Geophysics

13.1 EARTHQUAKE LOCATION AND DETERMINATION OF THE VELOCITY STRUCTURE OF THE EARTH FROM TRAVEL TIME DATA

The problem of determining the *hypocentral* parameters of an earthquake (that is, its location and its origin time) and the problem of determining the velocity structure of the earth from travel time data are very closely coupled in geophysics, because earthquakes, which occur naturally at initially unknown locations, are a primary source of structural information.

When the wavelength of the seismic waves is smaller than the scale of the velocity heterogeneities, the propagation of seismic waves can be described with *ray theory*, an approximation in which energy is assumed to propagate from source to observer along a curved path called a *ray*. In this approximation, the wave field is completely determined by the pattern of rays and the travel time along them.

In areas where the velocity structure is already known, earthquakes can be located using Geiger's (1912) method (see Section 12.9). The travel time $T(\mathbf{x})$ of an earthquake with location $\mathbf{x}$ and its ray tangent $\mathbf{t}(\mathbf{x})$ (pointing from earthquake to receiver) is calculated using ray theory. The arrival time is $t = \tau + T(\mathbf{x})$, where τ is the origin time of the earthquake. The perturbation in arrival time due to a change in the earthquake location from $\mathbf{x}_0$ to $\mathbf{x}_0 + \delta\mathbf{x}$ and a change in the origin time from τ_0 to $\tau_0 + \delta\tau$ is

$$\delta t = t - [\tau_0 + T(\mathbf{x}_0)] \approx \delta\tau - s(\mathbf{x}_0)\mathbf{t}(\mathbf{x}_0) \cdot \delta\mathbf{x} \tag{13.1}$$

where s is the slowness (reciprocal velocity). Note that this linearized equation has four model parameters, τ, and the three components of $\mathbf{x}$.

The forward problem, that of finding the ray path connecting earthquake and station and its traveltime, is computationally intensive. Two alternative strategies have been put forward, one based on first finding the ray path using a shooting or bending strategy and then calculating the travel time by an integral of slowness

Geophysical Data Analysis: Discrete Inverse Theory. DOI: 10.1016/B978-0-12-397160-9.00013-8
© 2012 Elsevier Inc. All rights reserved.

along the ray (Cerveny, 2001), or conversely, by first finding the traveltime by solving its partial differential equation (the *Eikonal* equation) and then calculating the ray path by taking the gradient of the travel time (Vidale, 1990).

The velocity structure can be represented with varying degrees of complexity. The simplest is a vertically stratified structure consisting of a stack of homogeneous layers or, alternatively, a continuously varying function of depth. Analytic formula for the travel time and its derivative with source parameters can be derived in some cases (Aki and Richards, 2002, their Section 9.3). The most complicated cases are fully three-dimensional velocity models represented with voxels or splines. Compromises between these extremes are also popular and include two-dimensional models (in cases where the structure is assumed to be uniform along the strike of a linear tectonic feature) and regionalized models (which assign a different vertically stratified structure to each tectonically distinct region but average them when computing travel times of rays that cross from one region to another).

In the simplest formulation, each earthquake is located separately. The data kernel of the linearized problem is an $N \times 4$ matrix, where N is the number of travel time observations, and is typically solved by singular-value decomposition. A very commonly used computer program that uses a layered velocity model is HYPOINVERSE (Klein, 1985). However, locations can often be improved by simultaneously solving for at least some aspect of the velocity model, which requires the earthquakes to be located simultaneously. The simplest modification is to solve for a *station correction* for each station; that is, a time increment that can be added to predicted travel times to account for near-station velocity heterogeneity not captured by the velocity model. More complicated formulations solve for a vertically stratified structure (Crosson, 1976) and for fully three-dimensional models (Menke, 2005; Zelt, 1998; Zelt and Barton, 1998). In the three-dimensional case, the inverse problem is properly one of tomographic inversion, as the perturbation in arrival time δt includes a line integral

$$\delta t \approx \delta\tau - s(\mathbf{x}_0)\mathbf{t}(\mathbf{x}_0) \cdot \delta\mathbf{x} + \int_{\text{unperturbed ray}} \delta s \, d\boldsymbol{\ell} \tag{13.2}$$

Here, δs is the perturbation in slowness with respect to the reference model and $d\boldsymbol{\ell}$ is the arc-length along the unperturbed ray. This problem is $N \times (4K + L)$, where N is the number of arrival time observations, K is the number of earthquakes, and L is the number of unknown parameters in the velocity model. One of the limitations of earthquake data is that earthquakes tend to be spatially clustered. Thus, the ray paths often poorly sample the model volume, and *a priori* information, usually in the form of smoothness information, is required to achieve useful solutions. Even so, earthquake location (and especially earthquake depth) will often trade off strongly with velocity structure.

A very useful algorithm due to Pavlis and Booker (1980) simplifies the inversion of the very large matrices that result from simultaneously inverting for the source parameter of many thousands of earthquakes and an even larger

number of velocity model parameters. Suppose that the linearized problem has been converted into a standard discrete linear problem $\mathbf{d}=\mathbf{G}\mathbf{m}$, where the model parameters can be arranged into two groups $\mathbf{m}=[\mathbf{m}_1,\mathbf{m}_2]^{\mathrm{T}}$, where $\mathbf{m}_1$ is a vector of the earthquake source parameters and $\mathbf{m}_2$ is a vector of the velocity parameters. Then the inverse problem can be written

$$\mathbf{d}=\mathbf{G}\mathbf{m}=[\,\mathbf{G}_1\;\mathbf{G}_2\,]\begin{bmatrix}\mathbf{m}_1\\ \mathbf{m}_2\end{bmatrix}=\mathbf{G}_1\mathbf{m}_1+\mathbf{G}_2\mathbf{m}_2 \tag{13.3}$$

Now suppose that $\mathbf{G}_1$ has singular-value decomposition $\mathbf{G}_1=\mathbf{U}_{1p}\,\Lambda_{1p}\,\mathbf{V}_{1p}^{\mathrm{T}}$, with p nonzero singular values, so that $\mathbf{U}_1=[\mathbf{U}_{1p},\mathbf{U}_{10}]$ where $\mathbf{U}_{10}^{\mathrm{T}}\,\mathbf{U}_{1p}=0$. Premultiplying the inverse problem (Equation 13.3) by $\mathbf{U}_{10}^{\mathrm{T}}$ yields

$$\begin{aligned}&\mathbf{U}_{10}^{\mathrm{T}}\mathbf{d}=\mathbf{U}_{10}^{\mathrm{T}}\mathbf{G}_1\mathbf{m}_1+\mathbf{U}_{10}^{\mathrm{T}}\mathbf{G}_2\mathbf{m}_2=\mathbf{U}_{10}^{\mathrm{T}}\mathbf{G}_2\mathbf{m}_2\\ &\text{or}\quad \mathbf{d}'=\mathbf{G}'\mathbf{m}_2\quad\text{with}\quad \mathbf{d}'=\mathbf{U}_{10}^{\mathrm{T}}\mathbf{d}\quad\text{and}\quad \mathbf{G}'=\mathbf{U}_{10}^{\mathrm{T}}\mathbf{G}_2\end{aligned} \tag{13.4}$$

Here, we have used the fact that $\mathbf{U}_{10}^{\mathrm{T}}\mathbf{G}_1\mathbf{m}_1=\mathbf{U}_{10}^{\mathrm{T}}\mathbf{U}_{1p}\,\Lambda_{1p}\mathbf{V}_{1p}^{\mathrm{T}}\mathbf{m}_1=0$. Premultiplying the inverse problem (Equation 13.3) by $\mathbf{U}_{1p}^{\mathrm{T}}$ yields

$$\begin{aligned}&\mathbf{U}_{1p}^{\mathrm{T}}\mathbf{d}=\mathbf{U}_{1p}^{\mathrm{T}}\mathbf{G}_1\mathbf{m}_1+\mathbf{U}_{1p}^{\mathrm{T}}\mathbf{G}_2\mathbf{m}_2=\Lambda_{1p}\mathbf{V}_{1p}^{\mathrm{T}}\mathbf{m}_1+\mathbf{U}_{1p}^{\mathrm{T}}\mathbf{G}_2\mathbf{m}_2\\ &\text{or}\quad \mathbf{d}''=\mathbf{G}''\mathbf{m}_1\quad\text{with}\quad \mathbf{d}''=\mathbf{U}_{1\mathrm{p}}^{\mathrm{T}}\mathbf{d}-\mathbf{U}_{1\mathrm{p}}^{\mathrm{T}}\mathbf{G}_2\mathbf{m}_2\quad\text{and}\quad \mathbf{G}''=\Lambda_{1p}\mathbf{V}_{1p}^{\mathrm{T}}\end{aligned} \tag{13.5}$$

The inverse problem has been partitioned into two equations, an equation for $\mathbf{m}_2$ (the velocity parameters) that can be solved first, and an equation for $\mathbf{m}_1$ (the source parameters) that can be solved once $\mathbf{m}_2$ is known (see *MatLab* script gda13_01). When the data consist of only absolute travel times (as contrasted to differential travel times, described below), the source parameter data kernel $\mathbf{G}_1$ is block diagonal, since the source parameters of a given earthquake affect only the travel times associated with that earthquake. The problem of computing the singular-value decomposition of $\mathbf{G}_1$ can be broken down into the problem of computing the singular-value decomposition of the sub-matrices.

Signal processing techniques based on waveform cross-correlation are able to determine the difference $\Delta t^{\mathrm{A,B}}=t^{\mathrm{A}}-t^{\mathrm{B}}$ in arrival times of two earthquakes (say, labeled A and B), observed at a common station, to several orders of magnitude better precision than the absolute arrival times t^{A} and t^{B} can be determined. Thus, earthquake locations can be vastly improved, either by supplementing absolute arrival time data t with differential arrival time data $\Delta t^{\mathrm{A,B}}$ or by relying on differential data alone. The perturbation in differential arrival time for two earthquakes, A and B, is formed by differencing two versions of Equation (13.1)

$$\delta\Delta t^{\mathrm{A,B}}\approx\delta\tau^{\mathrm{A}}-s(\mathbf{x}_0^{\mathbf{A}})\mathbf{t}(\mathbf{x}_0^{\mathbf{A}})\cdot\delta\mathbf{x}^{\mathbf{A}}-\delta\tau^{\mathrm{B}}+s(\mathbf{x}_0^{\mathbf{B}})\mathbf{t}(\mathbf{x}_0^{\mathbf{B}})\cdot\delta\mathbf{x}^{\mathbf{B}} \tag{13.6}$$

Each equation involves eight unknowns. Early versions of this method focused on using differential data to improve the *relative* location between earthquakes (e.g., Spence and Alexander, 1968). The modern formulation, known as the

double-difference method, acknowledges that the relative locations are better determined than the absolute locations but solves for both (e.g., Menke and Schaff, 2004; Slunga et al., 1995; Waldhauser and Ellsworth, 2000). Double-difference methods have also been extended to include the tomographic estimation of velocity structure (Zhang and Thurber, 2003).

At the other extreme is the case where the origin times and locations of the seismic sources are known and the travel time T of elastic waves through the earth is used to determine the slowness $s(\mathbf{x})$ in the earth

$$\delta T = \int_{\text{perturbed ray}} \delta s \, d\boldsymbol{\ell} \approx \int_{\text{unperturbed ray}} \delta s \, d\boldsymbol{\ell} \tag{13.7}$$

The approximation relies on *Fermat's Principle*, which states that perturbations in ray path cause only second-order changes in traveltime. Tomographic inversion based on Equation (13.7) is now a standard tool in seismology and has been used to image the earth at many scales, from kilometer-scale geologic structures to the whole earth. An alternative to model estimation is presented by Vasco (1986) who uses extremal inversion to determine quantities such as the maximum difference between the velocities in two different parts of the model, assuming that the velocity perturbation is everywhere within a given set of bounds (see Chapter 6).

Tomographic inversions are often performed separately from earthquake location but then are degraded by whatever error is present in the earthquake locations and origin times. In the case where the earthquakes are all very distant from the volume being imaged, the largest error arises from the uncertainty in origin time, since errors in location do not change the part of the ray path in the imaged volume by very much. Suppose that the slowness in the imaged volume is represented as $s(x,y,z) = s_1(z) + s_2(x,y,z)$, with (x,y) horizontal and z vertical position, and with the mean of $s_2(x,y,z)$ zero at every depth, so that it represents horizontal variations in velocity, only. Poor knowledge of origin times affects only the estimates of $s_1(z)$ and not $s_2(x,y,z)$ (Aki et al., 1976; Menke et al., 2006). Estimates of the horizontal variations in velocity contain important information about the geologic structures that are being imaged, such as their (x,y) location and dip direction. Consequently, the method, called *teleseismic tomography*, has been applied in many areas of the world.

13.2 MOMENT TENSORS OF EARTHQUAKES

Seismometers measure the motion or *displacement* of the ground, a three-component vector $\mathbf{u}(\mathbf{x},t)$ that is a function of position $\mathbf{x}$ and time t. At wavelengths longer than the spatial scale of a seismic source, such as a geologic fault, an earthquake can be approximated by a system of nine *force-couples* acting at a point $\mathbf{x}_0$, the *centroid* of the source. The amplitude of these force-couples is described by a 3×3 symmetric matrix $\mathbf{M}(t)$, called the *moment tensor*, which is a function of time, t (Aki and Richards, 2002, their Section 3.3). Ground displacement depends on the time derivative $d\mathbf{M}/dt$, called *the moment-rate tensor*.

Its six independent elements constitute six continuous model parameters $m_i(t)$. The ground displacement at $(\mathbf{x},t)$ due to a set of force-couples at $\mathbf{x}_0$ is linearly proportional to the elements of the moment-rate tensor via

$$u_i^{\text{pre}}(\mathbf{x},t) = \sum_{j=1}^{6} \int G_{ij}(\mathbf{x},\mathbf{x}_0,t-t_0) m_j(t_0) \mathrm{d}t_0 \tag{13.8}$$

Here, $G_{ij}(\mathbf{x},\mathbf{x}_0,t-t_0)$ are data kernels that describe the ith component displacement at $(\mathbf{x},t)$ due to the jth force couple at $(\mathbf{x}_0,t_0)$. The data kernels depend upon earth structure and, though challenging to calculate, are usually assumed to be known. If the location $\mathbf{x}_0$ of the source is also known, then the problem of estimating the moment-rate tensor from observations of ground displacement is completely linear (Dziewonski and Woodhouse, 1983; Dziewonski et al., 1981). The time variability of the moment-rate tensor is usually parameterized, with a triangular function of fixed width (or several overlapping such functions) being popular, so that the total number of unknowns is $M=6L$, where L is the number of triangles. The error $e_i^{(j)}(t_k)$ associated with component i, observation location j, and observation time k is then

$$e_i^{(j)}(t_k) = u_i^{\text{obs}}(\mathbf{x}^{(\mathrm{j})},t_k) - u_i^{\text{pre}}(\mathbf{x}^{(\mathrm{j})},t_k) \tag{13.9}$$

When the total error E is defined as the usual L_2 prediction error

$$E = \sum_i \sum_j \sum_k [e_i^{(j)}(t_k)]^2 \tag{13.10}$$

the problem can be solved with simple least squares. Most natural sources, such as earthquakes, are thought to occur without causing volume changes, so some authors add the linear constraint $\dot{M}_{11}+\dot{M}_{22}+\dot{M}_{33}=0$ (at all times).

Sometimes, the location $\mathbf{x}_0=[x_0,y_0,z_0]^{\mathrm{T}}$ of the source (or just its depth z_0) is also assumed to be unknown. The inverse problem is then nonlinear. This problem can be solved using a grid search over source location, that is, by solving the linear inverse problem for a grid of fixed source locations and then choosing the one with smallest total error E. The Fréchet derivatives of E with respect to hypocentral parameters are known; so alternately, Newton's method can be used. Adjoint methods are also applicable to this problem (Kim et al., 2011).

Earthquake locations are now routinely calculated by the U.S. Geological Survey (http://earthquake.usgs.gov/earthquakes/recenteqsww) and seismic moment tensors by the Global Centroid Moment Tensor (CMT) Project (http://www.globalcmt.org). They are standard data sets used in earthquake and tectonic research.

13.3 WAVEFORM "TOMOGRAPHY"

The ray approximation used in Section 13.1 is valid only when the wavelength of the elastic waves is much smaller than the spatial scale of the velocity heterogeneities. This condition is violated for "long-period" seismic waves, which

have wavelengths of hundreds of kilometers—comparable with the spatial scales of many of the earth's most significant heterogeneities, including the lithosphere and asthenosphere. In this case, the displacement field must be computed by solving its partial differential equation, a much more challenging computation than those of ray theory. The inverse problem is to find the set of the earth's material constants (Lamé parameter λ, shear modulus μ, and density ρ) that minimizes some measure of the misfit between the observed and predicted displacement, as observed at set of N locations $\mathbf{x}^{(j)}$. The simplest misfit is simply the L_2 difference between the observed and predicted wave field

$$E = \sum_{i=1}^{3} \sum_{j=1}^{N} \int_{-\infty}^{+\infty} [u_i^{\text{obs}}(\mathbf{x}^{(j)}, t) - u_i^{\text{pre}}(\mathbf{x}^{(j)}, t)]^2 \, \mathrm{d}t \tag{13.11}$$

The Fréchet derivatives $\delta u/\delta\lambda$, $\delta u/\delta\mu$, and $\delta u/\delta\rho$ (and also $\delta E/\delta\lambda$, $\delta E/\delta\mu$, and $\delta E/\delta\rho$) can be computed via adjoint methods, as described in Section 11.10 (Dahlen et al., 2000; Tromp et al., 2005). These derivatives are fully three-dimensional, in the sense that a perturbation in material constants anywhere in the model will perturb the wave field at every observer. This is in contrast to the ray-theoretical approximation, where the effect is limited to perturbations located along the ray path connecting source and receiver. Thus, waveform "tomography" is not tomography in the strict sense; nevertheless, the term is commonly used.

At intermediate periods, seismic wave fields often appear to contain pulse-like arrivals (or *phases*), which are reminiscent of ray-like behavior. Ray theory correctly predicts that energy has traveled along a volumetrically restricted path but does not correctly predict the details. The ray-theoretical notion that a phase has a travel time is still very powerful, because it captures the underlying physics that the main effect of changing the slowness along the path is to advance or delay the arrival of the wave. A differential travel time ΔT can be defined by comparing time windows of the observed and predicted wave fields, centered about a particular phase, and determining whether one is advanced or delayed with respect to the other. These differential travel times constitute a new type of data. While they contain less information than the wave field itself, they are more robust.

Cross-correlation is typically used to define the differential travel time between an observed and predicted phase

$$\Delta T = \arg\max_{\tau} \; c(\tau) \quad \text{where} \quad c(\tau) = \int_{t_1}^{t_2} u^{\text{pre}}(t-\tau) u^{\text{obs}}(t) \mathrm{d}t \tag{13.12}$$

Here, the phase is assumed to arrive in the time window $t_1 < t < t_2$. The travel time difference ΔT (positive when the observed signal is delayed with respect to the predicted signal) is the time lag τ at which the cross-correlation $c(\tau)$ attains its maximum, that is, the time lag at which the two signals best align. Suppose now that the predicted wave field is perturbed by a small amount $\delta u(t)$, so that

$u^{\text{pre}}(t) = u_0(t) + \delta u(t)$. Marquering et al. (1999) show that the resulting change in differential travel time $\Delta T = \Delta T_0 + \delta T$ is

$$\delta T = \frac{1}{N}\int_{t_1}^{t_2} \dot{u}_0(t)\delta u(t)\mathrm{d}t \quad \text{with} \quad N = \int_{t_1}^{t_2} \ddot{u}_0(t)u_0(t)\mathrm{d}t \tag{13.13}$$

Here, the dot denotes differentiation in time. If we define a *window function* $W(t)$ which is unity within a time interval $t_1 < t < t_2$ that encloses a particular seismic arrival, such as the P wave, and zero outside it, then

$$\begin{gathered}\delta T = (h(t), \delta u(t)) \\ \text{with} \quad h(t) = \frac{1}{N}W(t)\dot{u}_0(t) \quad \text{and} \quad N = \int_{-\infty}^{+\infty} W(t)\ddot{u}_0(t)u_0(t)\mathrm{d}t\end{gathered} \tag{13.14}$$

Hence, the travel time data are related to the perturbed displacement field by an equation of the form 11.58, for which we have already derived a method for calculating the Fréchet derivative.

As the period of the seismic waves is decreased, the high-amplitude part of the travel time derivative becomes more and more concentrated along the ray-theoretical path, taking on a curved shape reminiscent of a banana. Counter-intuitively, although the derivative's amplitude is large *near* the ray path, it is exactly zero *on* the ray path. This behavior is due to the Fréchet derivative lacking an intrinsic length scale. A heterogeneity with an infinitesimal diameter can perturb the travel time only to the extent that the travel time from source to heterogeneity to observer is different than the travel time from source to receiver. For heterogeneities located along the ray path, the two travel times are exactly the same and the derivative is zero. In cross-section, the banana has a zero at its center; that is, it looks like a doughnut—hence the term *banana-doughnut* kernel. Another counter-intuitive property of these data kernels is that they contain sign reversals. A positive velocity perturbation near the ray path causes (as expected) a travel time advance but can cause a delay at other locations. This is a wave interference effect (Figure 13.1).

13.4 VELOCITY STRUCTURE FROM FREE OSCILLATIONS AND SEISMIC SURFACE WAVES

The earth, being a finite body, can oscillate only with certain characteristic patterns of motion (eigenfunctions) at certain corresponding discrete frequencies ω_{knm} (eigenfrequencies), where the indices (k,n,m) increase with the complexity of the motion. Most commonly, the frequencies of vibration are measured from the spectra of seismograms of very large earthquakes, and the corresponding indices are estimated by comparing the observed frequencies of vibration to those predicted using a reference model of the earth's structure. The inverse problem is to use small differences between the measured and predicted frequencies to improve the reference model.

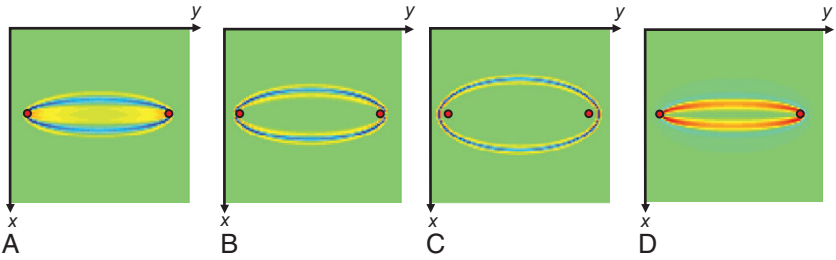

FIGURE 13.1 (A–C) Fréchet derivative $\delta p/\delta m$ of pressure p with respect to $\delta m = 2v_0\delta v$, for an acoustic wave propagation problem with constant background velocity v_0 and velocity perturbation δv, as a function of spatial position (x,y), and at three different times. Source and receiver, separated by a distance R, are shown as red circles. The time shown in (A) is shortly after the ray-theoretical time $t = R/v_0$, with (B) and (C) corresponding to subsequent times. Note that the region of space containing heterogeneities that can affect the pressure at a given time t expands with time. (D) Fréchet derivative $\delta T/\delta m$ of travel time, called a banana-doughnut kernel. Note that the derivative is zero on the line connecting source and receiver, which corresponds to the geometrical ray path. Note also the sign reversal (blue). *MatLab* script gda13_01.

The most fundamental part of this inverse problem is the Fréchet derivatives relating small changes in the earth's material properties to the resultant changes in the frequencies of vibration. These derivatives are computed through the perturbative techniques described in Section 12.10 and are three-dimensional analogs of Equation (12.44) (Takeuchi and Saito, 1972). All involve integrals of the perturbation in earth structure with the eigenfunctions of the reference model, so the problem is one in linearized continuous inverse theory. The complexity of this problem is determined by the assumptions made about the earth model—whether it is isotropic or anisotropic, radially stratified or laterally heterogeneous—and how it is parameterized. A radially stratified inversion that incorporates both body wave travel times and eigenfrequencies and that has become a standard in the seismological community is the Preliminary Reference Earth Model of Dziewonski and Anderson (1981).

Seismic *surface waves*, such as the horizontally polarized *Love* wave and the vertically polarized *Rayleigh* wave, occur when seismic energy is trapped near the earth's surface by the rapid increase in seismic velocities with depth. The velocity of these waves is very strongly frequency dependent, since the lower-frequency waves extend deeper into the earth than the higher-frequency waves and are affected by correspondingly higher velocities. The surface wave caused by an impulsive source, such as an earthquake, is not itself impulsive, but rather is dispersed, since its component frequencies propagate at different phase velocities.

The phase velocity c of the surface wave is defined as $c = \omega(k)/k$, where ω is angular frequency and k is wavenumber (wavenumber is 2π divided by wavelength). The function $c(\omega)$ (or alternatively, $c(k)$) is called the *dispersion function* (or, sometimes, the *dispersion curve*) of the surface wave. It is easily measured from observations of earthquake wave fields and easily predicted for vertically stratified earth models. The inverse problem is to infer the earth's material properties from observations of the dispersion function.

As in the case of free oscillations, Fréchet derivatives can be calculated using a pertubative approach. For Love waves, the perturbation in phase velocity c due to a perturbation in shear modulus μ and density ρ is (Aki and Richards, 2002, their Equation 7.71):

$$\left(\frac{\delta c}{c}\right)_{\omega} = \frac{\int_0^{\infty}\left[k^2 l_1^2 + \left(\frac{\mathrm{d}l_1}{\mathrm{d}z}\right)^2\right]\delta\mu\,\mathrm{d}z - \int_0^{\infty}\left[\omega^2 l_1^2\right]\delta\rho\,\mathrm{d}z}{2k^2\int_0^{\infty}\mu_0 l_1^2\,\mathrm{d}z} \tag{13.15}$$

Here, shear modulus $\mu(z)$ and density $\rho(z)$ are assumed to be vertically stratified in depth z, and $u_y = l_1(\omega,k,z)\exp(\mathrm{i}kx - \mathrm{i}\omega t)$ is the horizontal displacement of the Love wave. The corresponding formula for the Rayleigh wave (Aki and Richards, 2002, their Equation 7.78) is similar in form but involves perturbations in three material parameters, Lamé parameter λ, shear modulus μ, and density ρ.

When the earth model contains three-dimensional heterogeneities, the inverse problem of determining earth structure from surface waves is just a type of waveform tomography (see Section 13.3). The dispersion function, now understood to encode the frequency-dependent time shift between an observed and predicted surface wave field, is still a useful quantity. Adjoint methods can be used to compute its Fréchet derivatives (Sieminski et al., 2007).

Ray-theoretical ideas can also be applied to surface waves, which at sufficiently short periods can be approximated as propagating along a horizontal ray path connecting source and receiver (Wang and Dahlen, 1995). Each point on the earth's surface is presumed to have its own dispersion function $c(\omega,x,y)$. The dispersion function $c^{(i)}(\omega)$ observed for a particular source-observer path i is then calculated using the *pure path approximation*

$$\frac{1}{c^{(i)}(\omega)} = \frac{1}{S_i}\int_{\text{path } i}\frac{\mathrm{d}\ell}{c(\omega,x,y)} \tag{13.16}$$

Here, S_i is the length of path i. The path is often approximated as the great circle connecting source and receiver. This is a two-dimensional tomography problem relating line integrals of $c(\omega,x,y)$ to observations of $c^{(i)}(\omega)$—hence the term *surface wave tomography*. The $c(\omega,x,y)$ then can be inverted for estimates of the vertical structure using the vertically stratified method described above (e.g., Boschi and Ekström, 2002). This two-step process is not as accurate as full waveform tomography but is computationally much faster.

13.5 SEISMIC ATTENUATION

Internal friction within the earth causes seismic waves to lose amplitude as they propagate. The amplitude loss can be characterized by the *attenuation factor* $\alpha(\omega,\mathbf{x})$, which can vary with angular frequency ω and position $\mathbf{x}$. In

a regime where ray theory is accurate, the fractional loss of amplitude A of the seismic wave is

$$\frac{A}{A_0} = \exp\left\{-\int_{(\text{ray})} \alpha(\omega, x)\mathrm{d}\ell\right\} \tag{13.17}$$

Here, A_0 is the initial amplitude. After linearization through taking its logarithm, this equation is a standard tomography problem (identical in form to the X-ray problem described in Section 1.3.4)—hence the term *attenuation tomography*. Jacobson et al. (1981) discuss its solution when the attenuation is assumed to vary only with depth. Dalton and Ekström (2006) perform a global two-dimensional tomographic inversion using surface waves. Many authors have performed three-dimensional inversions (see review by Romanowicz, 1998).

A common problem in attenuation tomography is that the initial amplitude A_0 (or equivalently, the CMT of the seismic source) is unknown (or poorly known). Menke et al. (2006) show that including A_0 as an unknown parameter in the inversion introduces a nonuniquess that is mathematically identical to the one encountered when an unknown origin time is introduced into the velocity tomography problem.

Attenuation also has an effect on the free oscillations of the earth, causing a widening in the spectral peaks associated with the eigenfrequencies ω_{knm}. The amount of widening depends on the relationship between the spatial variation of the attenuation and the spatial dependence of the eigenfunction and can be characterized by a quality factor Q_{knm} (that is, a fractional loss of amplitude per oscillation) for that mode. Pertubative methods are used to calculate the Fréchet derivatives of Q_{knm} with the attenuation factor. One example of such an inversion is that of Masters and Gilbert (1983).

13.6 SIGNAL CORRELATION

Geologic processes record signals only imperfectly. For example, while variations in oxygen isotopic ratios $r(t)$ through geologic time t are recorded in oxygen-bearing sediments as they are deposited, the sedimentation rate is itself a function of time. Measurements of isotopic ratio $r(z)$ as a function of depth z cannot be converted to the variation of $r(t)$ without knowledge of the sedimentation function $z(t)$ (or equivalently $t(z)$).

Under certain circumstances, the function $r(t)$ is known *a priori* (for instance, oxygen isotopic ratio correlates with temperature, which can be estimated independently). In these instances, it is possible to use the observed $r^{\text{obs}}(z)$ and the predicted $r^{\text{pre}}(t)$ to invert for the function $t(z)$. This is essentially a problem in signal correlation: distinctive features that can be correlated between $r^{\text{obs}}(z)$ and $r^{\text{pre}}(t)$ establish the function $t(z)$. The inverse problem

$$r^{\text{obs}}(z) = r^{\text{pre}}[t(z)] \tag{13.18}$$

is therefore a problem in nonlinear continuous inverse theory. The unknown function $t(z)$—often called the *mapping function*—must increase monotonically with z. The solution of this problem is discussed by Martinson et al. (1982) and Shure and Chave (1984).

13.7 TECTONIC PLATE MOTIONS

The motion of the earth's rigid tectonic plates can be described by an *Euler vector* $\boldsymbol{\omega}$, whose orientation gives the pole of rotation and whose magnitude gives its rate. Euler vectors can be used to represent relative rotations, that is, the rotation of one plate relative to another, or absolute motion, that is, motion relative to the earth's mantle. If we denote the relative rotation of plate A with respect to plate B as $\boldsymbol{\omega}_{AB}$, then the Euler vectors of three plates A–C satisfy the relationship $\boldsymbol{\omega}_{AB}+\boldsymbol{\omega}_{BC}+\boldsymbol{\omega}_{CA}=0$. Once the Euler vectors for plate motion are known, the relative velocity between two plates at any point on their boundary can easily be calculated from trigonometric formulas.

Several geologic features provide information on the relative rotation between plates, including the faulting directions of earthquakes at plate boundaries and the orientation of transform faults, which constrain the direction of the relative velocity vectors, and spreading rate estimates based on magnetic lineations at ridges, which constrain the magnitude of the relative velocity vectors.

These data can be used in an inverse problem to determine the Euler vectors (Chase and Stuart, 1972; DeMets et al., 2010; Minster et al., 1974). The main differences between various authors' approaches are in the manner in which the Euler vectors are parameterized and the types of data that are used. Some authors use Cartesian components of the Euler vectors; others use magnitude, azimuth, and inclination. The two inversions behave somewhat differently in the presence of noise; Cartesian parameterizations seem to be more stable. Data include seafloor spreading rates, azimuths of strike-slip plate boundaries, earthquake moment tensors, and Global Positioning System strain measurements.

13.8 GRAVITY AND GEOMAGNETISM

Inverse theory plays an important role in creating representations of the earth's gravity and magnetic fields. Field measurements made at many points about the earth need to be combined into a smooth representation of the field, a problem which is mainly one of interpolation in the presence of noise and incomplete data (see Section 3.9.3). Both spherical harmonic expansions and various kinds of spline functions (Sandwell, 1987; Shure et al., 1982, 1985; Wessel and Becker, 2008) have been used in the representations. In either case, the trade-off of resolution and variance is very important.

Studies of the earth's core and geodynamo require that measurements of the magnetic field at the earth's surface be extrapolated to the core-mantle

boundary. This is an inverse problem of considerable complexity, since the short-wave-length components of the field that are most important at the core-mantle boundary are very poorly measured at the earth's surface. Furthermore, the rate of change of the magnetic field with time (called "secular variation") is of critical interest. This quantity must be determined from fragmentary historical measurements, including measurements of compass deviation recorded in old ship logs. Consequently, these inversions introduce substantial *a priori* constraints on the behavior of the field near the core, including the assumption that the root-mean-square time rate of change of the field is minimized and the total energy dissipation in the core is minimized (Bloxam, 1987). Once the magnetic field and its time derivative at the core-mantle boundary have been determined, they can be used in inversions for the fluid velocity near the surface of the outer core (Bloxam, 1992).

The earth's gravity field $\mathbf{g}(\mathbf{x})$ is determined by its density structure $\rho(\mathbf{x})$. In parts of the earth where electric currents and magnetic induction are unimportant, the magnetic field $\mathbf{H}(\mathbf{x})$ is caused by the magnetization $\mathbf{M}(\mathbf{x})$ of the rocks. These quantities are related by

$$
\begin{aligned}
g_i(\mathbf{x}) &= \int_V G_i^g(\mathbf{x},\mathbf{x}_0)\rho(\mathbf{x}_0)\mathrm{d}V_0 \qquad \text{and} \qquad H_i(\mathbf{x}) = \int_V G_{ij}^H(\mathbf{x},\mathbf{x}_0)\, M_j(\mathbf{x}_0)\, \mathrm{d}V_0 \\
G_i^g(\mathbf{x}) &= -\frac{\gamma(x_i - x_{0i})}{|\mathbf{x}-\mathbf{x}_0|^3} \\
G_{ij}^H(\mathbf{x}) &= \frac{1}{4\pi\mu_0}\left[\frac{\delta_{ij}}{|\mathbf{x}-\mathbf{x}_0|^3} - \frac{3(x_i - x_{0i})(x_j - x_{0j})}{|\mathbf{x}-\mathbf{x}_0|^5}\right]
\end{aligned}
\tag{13.19}
$$

Here, γ is the gravitational constant and μ_0 is the magnetic permeability of the vacuum. Note that in both cases, the fields are linearly related to the sources (density and magnetization), so these are linear, continuous inverse problems. Nevertheless, inverse theory has proved to have little application to these problems, owing to their inherent underdetermined nature. In both cases, it is possible to show (e.g., Menke and Abbott, 1989, their Section 5.14) that the field outside a finite body can be generated by an infinitesimally thick spherical shell of mass or magnetization surrounding the body and below the level of the measurements. The null space of these problems is so large that it is generally impossible to formulate any useful solution (as we encountered in Figure 7.7), except when an enormous amount of *a priori* information is added to the problem.

Some progress has been made in special cases where *a priori* constraints can be sensibly stated. For instance, the magnetization of sea mounts can be computed from their magnetic anomaly, assuming that their magnetization vector is everywhere parallel (or nearly parallel) to some known direction and that the shape of the magnetized region closely corresponds to the observed bathymetric expression of the sea mount (Grossling, 1970; Parker, 1988; Parker et al., 1987).

13.9 ELECTROMAGNETIC INDUCTION AND THE MAGNETOTELLURIC METHOD

An electromagnetic field is *induced* within a conducting medium when it is illuminated by a plane electromagnetic wave. In a homogeneous medium, the field decays exponentially with depth, since energy is dissipated by electric currents induced in the conductive material. In a heterogeneous medium, the behavior of the field is more complicated and depends on the details of the electrical conductivity $\sigma(\mathbf{x})$.

Consider the special case of a vertically stratified earth illuminated by a vertically incident electromagnetic wave (e.g., from space). The plane wave consists of a horizontal electric field E_x and a horizontal magnetic field H_y, with their ratio on the earth's surface (at $z=0$) being determined by the conductivity profile $\sigma(z)$. The quantity

$$Z(\omega) = \frac{E_x(\omega, z=0)}{\mathrm{i}\omega H_y(\omega, z=0)} \tag{13.20}$$

that relates the two fields is called the *impedance*. It is frequency dependent, because the depth of penetration of an electromagnetic wave into a conductive medium depends upon its frequency, so different frequencies sense different parts of the conductivity structure. The *magnetotelluric* problem is to determine the conductivity $\sigma(z)$ from measurements of the impedance $Z(\omega)$ at a suite of angular frequencies ω. This one-dimensional nonlinear inverse problem is relatively well understood. For instance, Fréchet derivatives are relatively straightforward to compute (Parker, 1970, 1977). Oldenburg (1983) discusses the discretization of the problem and the application of extremal inversion methods (see Chapter 6).

In three-dimensional problems, all the components of magnetic $\mathbf{H}$ and electric $\mathbf{E}$ fields are now relevant, and the impedance matrix $\mathbf{Z}(\omega,\mathbf{x})$ that connects them via $\mathbf{E}=\mathbf{ZH}$ constitutes the data. The inverse problem of determining $\sigma(\mathbf{x})$ from observations of $\mathbf{Z}(\omega,\mathbf{x})$ at a variety of frequencies and positions is much more difficult than the one-dimensional version (Chave et al., 2012), especially in the presence of large heterogeneities in electoral resistivity caused, for instance, by topography (e.g., Baba and Chave, 2005).

REFERENCES

Aki, K., Richards, P.G., 2002. Quantitative Seismology, second ed. University Science Books, Sausalito, California 700pp.

Aki, K., Christoffersson, A., Husebye, E., 1976. Three-dimensional seismic structure under the Montana LASA. Bull. Seism. Soc. Am. 66, 501–524.

Baba, K., Chave, A., 2005. Correction of seafloor magnetotelluric data for topographic effects during inversion. J. Geophys. Res. 110, B12105, doi:10.1029/2004JB003463, 16pp.

Bloxam, J., 1987. Simultaneous inversion for geomagnetic main field and secular variation, 1. A large scale inversion problem. J. Geophys. Res. 92, 11597–11608.

Bloxam, J., 1992. The steady part of the secular variation of the earth's magnetic field. J. Geophys. Res. 97, 19565–19579.

Boschi, J., Ekström, G., 2002. New images of the Earth's upper mantle from measurements of surface wave phase velocity anomalies. J. Geophys. Res. 107, 2059.

Cerveny, V., 2001. Seismic Ray Theory. Cambridge University Press, New York 607pp.

Chase, E.P., Stuart, G.S., 1972. The N plate problem of plate tectonics. Geophys. J. R. Astronom. Soc. 29, 117–122.

Chave, A., Jones, A., Mackie, R., Rodi, W., 2012. The Magnetotelluric Method, Theory and Practice. Cambridge University Press, New York, 590pp.

Crosson, R.S., 1976. Crustal structure modeling of earthquake data: 1. Simultaneous least squares estimation of hypocenter and velocity parameters. J. Geophys. Res. 81, 3036–3046.

Dahlen, F.A., Hung, S.-H., Nolet, G., 2000. Fréchet kernels for finite frequency traveltimes—i. Theory. Geophys. J. Int. 141, 157–174.

Dalton, C.A., Ekström, G., 2006. Global models of surface wave attenuation. J. Geophys. Res. 111, B05317.

DeMets, C., Gordon, R.G., Argus, D.F., 2010. Geologically current plate motions. Geophys. J. Int. 181, 1–80.

Dziewonski, A.M., Anderson, D.L., 1981. Preliminary reference earth model. Phys. Earth Planet. Interiors 25, 297–358.

Dziewonski, A.M., Woodhouse, J.H., 1983. An experiment in systematic study of global seismicity: centroid-moment tensor solutions for 201 moderate and large earthquakes of 1981. J. Geophys. Res. 88, 3247–3271.

Dziewonski, A.M., Chou, T.-A., Woodhouse, J.H., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. J. Geophys. Res. 86, 2825–2852.

Geiger, L., 1912. Probability method for the determination of earthquake epicenters from the arrival time only (translated from Geiger's 1910 German article). Bull. St. Louis Univ. 8, 56–71.

Grossling, B.F., 1970. Seamount Magnetism. In: Maxwell, A.E. (Ed.), The Sea, vol. 4. Wiley, New York, pp. 129–156.

Jacobson, R.S., Shor, G.G., Dorman, L.M., 1981. Linear inversion of body wave data—part 2: attenuation vs. depth using spectral ratios. Geophysics 46, 152–162.

Kim, Y., Liu, Q., Tromp, J., 2011. Adjoint centroid-moment tensor inversions. Geophys. J. Int. 186, 264–278.

Klein, F.W., 1985. User's guide to HYPOINVERSE, a program for VAX and PC350 computers to solve for earthquake locations. U.S. Geologic Survey Open File Report 85–515, 24pp.

Marquering, H., Dahlen, F.A., Nolet, G., 1999. Three-dimensional sensitivity kernels for finite frequency traveltimes: the banana-doughnut paradox. Geophys. J. Int. 137, 805–815.

Martinson, D.G., Menke, W., Stoffa, P., 1982. An inverse approach to signal correlation. J. Geophys. Res. 87, 4807–4818.

Masters, G., Gilbert, F., 1983. Attenuation in the earth at low frequencies. Philos. Trans. R. Soc. Lond. Ser. A 388, 479–522.

Menke, W., 2005. Case studies of seismic tomography and earthquake location in a regional context. In: Levander, A., Nolet, G. (Eds.), Seismic Earth: Array Analysis of Broadband Seismograms. Geophysical Monograph Series157, American Geophysical Union, Washington, D.C., pp. 7–36.

Menke, W., Abbott, D., 1989. Geophysical Theory (Textbook). Columbia University Press, New York, 458pp.

Menke, W., Schaff, D., 2004. Absolute earthquake location with differential data. Bull. Seism. Soc. Am. 94, 2254–2264.

Menke, W., Holmes, R.C., Xie, J., 2006. On the nonuniqueness of the coupled origin time-velocity tomography problem. Bull. Seism. Soc. Am. 96, 1131–1139.

Minster, J.B., Jordan, T.H., Molnar, P., Haines, E., 1974. Numerical modeling of instantaneous plate tectonics. Geophys. J. R. Astronom. Soc. 36, 541–576.

Oldenburg, D.W., 1983. Funnel functions in linear and nonlinear appraisal. J. Geophys. Res. 88, 7387–7398.

Parker, R.L., 1970. The inverse problem of the electrical conductivity of the mantle. Geophys. J. R. Astronom. Soc. 22, 121–138.

Parker, R.L., 1977. The Fréchet derivative for the one dimensional electromagnetic induction problem. Geophys. J. R. Astronom. Soc. 39, 543–547.

Parker, R.L., 1988. A statistical theory of seamount magnetism. J. Geophys. Res. 93, 3105–3115.

Parker, R.L., Shure, L., Hildebrand, J., 1987. An application of inverse theory to seamount magnetism. Rev. Geophys. 25, 17–40.

Pavlis, G.L., Booker, J.R., 1980. The mixed discrete-continuous inverse problem: application to the simultaneous determination of earthquake hypocenters and velocity structure. J. Geophys. Res. 85, 4801–4809.

Romanowicz, B., 1998. Attenuation tomography of the earth's mantle: a review of current status. Pure Appl. Geophys. 153, 257–272.

Sandwell, D.T., 1987. Biharmonic spline interpolation of GEOS-3 and SEASAT altimeter data. Geophys. Res. Lett. 14, 139–142.

Shure, L., Chave, A.D., 1984. Comments on "An inverse approach to signal correlation." J. Geophys. Res. 89, 2497–2500.

Shure, L., Parker, R.L., Backus, G.E., 1982. Harmonic splines for geomagnetic modeling. Phys. Earth Planet. Interiors 28, 215–229.

Shure, L., Parker, R.L., Langel, R.A., 1985. A preliminary harmonic spline model for MAGSAT data. J. Geophys. Res. 90, 11505–11512.

Sieminski, A., Liu, Q., Trampert, J., Tromp, J., 2007. Finite-frequency sensitivity of surface waves to anisotropy based upon adjoint methods. Geophys. J. Int. 168, 1153–1174.

Slunga, R., Rognvaldsson, S.T., Bodvarsson, B., 1995. Absolute and relative locations of similar events with application to microearthquakes in southern Iceland. Geophys. J. Int. 123, 409–419.

Spence, W., Alexander, S., 1968. A method of determining the relative location of seismic events. Earthquake Notes 39, 13.

Takeuchi, H., Saito, M., 1972. Seismic surface waves. In: Bolt, B.A. (Ed.), Methods of Computational Physics 11, Academic Press, New York, 217–295.

Tromp, J., Tape, C., Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. Geophys. J. Int. 160, 195–216.

Vasco, D.W., 1986. Extremal inversion of travel time residuals. Bull. Seismol. Soc. Am. 76, 1323–1345.

Vidale, J., 1990. Finite difference calculation of travel times in 3-D. Geophysics 55, 521–526.

Waldhauser, F., Ellsworth, W., 2000. A double-difference earthquake location algorithm; method and application to the northern Hayward Fault, California. Bull. Seism. Soc. Am. 90, 1353–1368.

Wang, Z., Dahlen, F.A., 1995. Validity of surface-wave ray theory on a laterally heterogeneous earth. Geophys. J. Int. 123, 757–773.

Wessel, P., Becker, J., 2008. Gridding of spherical data using a Green's function for splines in tension. Geophys. J. Int. 174, 21–28.

Zelt, C.A., 1998. FAST: 3-D First Arrival Seismic Tomography programs. http://terra.rice.edu/department/faculty/zelt/fast.html.

Zelt, C.A., Barton, P.J., 1998. 3D seismic refraction tomography: a comparison of two methods applied to data from the Faeroe Basin. J. Geophys. Res. 103, 7187–7210.

Zhang, H., Thurber, C.H., 2003. Double-difference tomography; the method and its application to the Hayward Fault, California. Bull. Seismol. Soc. Am. 93, 1875–1889.