zones just south of the southernmost extent of the glaciers, including much of the United States, are subsiding by millimeters per year. This ongoing change in shape of the entire northern hemisphere, and indeed at some level the entire globe, has direct effects on the orientation the Earth, and has the secondary effect of changing the geometry of the network that defines the terrestrial reference frame.

At an even longer time scale is a variation in LOD caused by the transfer of energy and angular momentum between the spin of the Earth and the orbital motion of the Moon. This is a highly nonlinear phenomenon, but at the current configuration of the Earth-Moon system the interaction results in the Moon receding from the Earth, with the mean distance increasing by about 3.7 centimeters per year, while the Earth’s spin rate is decreasing by approximately $6.3 \times 10^{-22}$ radians per second$^{-2}$. The ability to directly measure the change in the distance to the Moon is an important contribution of LLR to separating long period changes in Earth rotation.

Concluding Remarks

The introduction of advanced technology has completely redefined the monitoring of Earth orientation during the past decade, and the immediate future promises still further progress. By the close of this century a network of 20 to 30 globally distributed VLBI observatories will define a terrestrial reference frame that will continually be maintained at the 1 cm level, properly accounting for the effects of plate tectonics, glacial rebound, earth tides, and ocean loading. For the first time a complete suite of EOP time series accurate to a fraction of a millisecond of arc, at intervals of a few hours, will be available for geophysical applications. This breakthrough in observational capabilities has already begun a revitalization of theoretical and analytical efforts, offering new hope of unraveling some of the mysteries about the internal structure and dynamics of Earth.

WILLIAM E. CARTER

References


Cross-references: Core-Mantle Coupling; Earth Structure: Global; Equilibrium Figure of the Earth; Figure of the Earth; Mantle Convection and Plumes; Mantle Dynamics; Mantle Viscosity; Planetary Geodesy; Very-Long-Baseline Interferometry (VLBI).

EARTHQUAKE MECHANISMS

Introduction

The most common model for the origin of crustal earthquakes is based on the “elastic rebound” theory as put forth in 1911 by Reid in his study of the 1906 San Francisco earthquake: Earthquakes occur in regions subject to deformation due to external causes, generally regional tectonic stresses. When the accumulated strain at some point exceeds the strength of the rock there is brittle fracture of competent rock and/or slip on preexisting zones of weakness. The region over which the slip or displacement occurred is called the earthquake fault. Before and after the earthquake the region is in static equilibrium. In accordance with this model, the estimate of the plane of faulting and the direction of slip within that plane inferred from seismic data is referred to as the focal mechanism solution. Focal mechanism solutions determined for several earthquakes in a geological region allow both the possible identification of large-scale faults and also some insight into the spatial geometry of the in situ stresses that cause the deformation.

This model is also applied to deep-focus earthquakes in subducting lithospheric slabs where “faults” in the conventional sense may not exist. However, a single point-force model for the focal mechanism, the double-couple model, appears applicable to all earthquakes [with the possible exception of some earthquakes related to volcanoes (see Earthquakes: Volcanogenic)].

In this chapter we develop the link between the earthquake source moment tensor introduced by Madariaga (see Seismic Source: Theory) and the applications of earthquake mechanisms to constrain tectonic plate motions as discussed by Stein and Woods (see Earthquake Mechanisms and Plate Tectonics). The emphasis here is on the methods used to obtain estimates of focal mechanisms from body wave polarity data, from body wave amplitude ratios, and from waveform fitting.

A Brief History

As early as the 1920s, some seismologists noticed that the first motions for P waves from earthquakes—up or down on vertical seismograms—showed systematic patterns with regard to the source-
station geometry. In particular, the first motions could be grouped into quadrants about the focus alternating between up motion (compressions) and down motion (dilatations) as shown in Fig. 1a. Many point-force equivalent models for the earthquake source process were compared with observation, and the single-couple model (often called a type I source) seemed consistent with the data and physically reasonable, based on the elastic rebound theory described earlier.

In addition to fitting P-wave first motion data, a successful source model must also be consistent with the observed S-wave particle motion. Because S waves are not first arrivals, S-wave first-motion data are generally inferior in quality to P-wave data. In addition, three-component seismographs are required to determine estimates of the S-wave particle motion. Hence it was not until the 1960s that there was an adequate data set of S-wave particle motion of sufficient quality to test source models. Once available, these data were consistent with the quadrupole pattern of the double couple (or type II source) as shown in Fig. 1b rather than the dipole pattern predicted by the single-couple model, ruling out the single couple as an appropriate source model. When the mathematical theory now used for source studies was developed, it was realized that a single couple was not consistent with static equilibrium before and after the earthquake—the same logic that leads to the stress tensor being symmetric. Benioff (1964) showed that the double-couple model, rather than the single-couple model, was implied by a correct interpretation of the classic elastic rebound theory.

Good reviews of the earliest days of earthquake mechanisms studies are given by Stauder (1962) and Honda (1962).

The use of P-arrival polarities remains today the most common way to constrain the focal mechanism for a given event. S-wave particle motion, in the form of polarization estimates, has been supplanted by more robust methods. Among the methods currently used are polarity comparisons for teleseismic observations of P with pP and/or sP, amplitude ratios from among pP/P, sP/P, SV/P, or SH/P, and inversion schemes based on waveform fitting. After a review of the underlying theory, these methods are described.

**Theory**

The ground displacement vector from an earthquake can be written as the convolution of a moment tensor and a gradient of a Green’s function (see *Seismic Source: Theory*). The Green’s function includes propagation effects, and the moment tensor contains information about the source process. The moment tensor can be interpreted as the volume integral of the stress drop associated with the earthquake. The trace of the moment tensor represents an isotropic source, such as an explosion. Since no volume change is observed with normal earthquakes, one generally restricts consideration to traceless moment tensors. (Removing the trace is a linear constraint and is therefore easily incorporated in an inversion procedure.) The resulting point-source moment tensor includes more general force systems than a pure double couple, but for now we consider only a pure double-couple source model. A double-couple model corresponds to a point-source moment tensor of the form

\[
M(t)_{jk} = M_0(t) \left[ n_j a_k + n_k a_j \right]
\]

where \( M_0(t) \) is the total seismic moment (see *Seismic Source: Theory*), and \( a_j \) and \( n_k \) are respectively the \( j \)th and \( k \)th components of \( \mathbf{a} \), a unit vector in the direction of the fault slip, and \( \mathbf{n} \), a unit vector normal to the fault plane (Fig. 2).

To see explicitly the connection between the fault geometry and the equivalent force picture for an earthquake, we evaluate Eq. 1 in two different coordinate frames. For the coordinate frame in which \( \mathbf{a} \) lies along \( x \) and \( \mathbf{n} \) along \( y \), the right-hand side of Eq. 1 becomes \( M_0(t) \left[ \delta_{ij} \delta_{k2} + \delta_{ij} \delta_{k3} \right] \), where \( \delta_{jk} \) is
unity for \( j = k \) and 0 otherwise. As shown in Fig. 3a, in this coordinate frame the stresses implied by the moment tensor take the form of a double couple with one couple representing the slip along \( a \). If now we represent the moment tensor in a coordinate system rotated clockwise by \( 45^\circ \) about \( z \), the right-hand side of Eq. 1 is of the form \( M_d(t) \left[ \delta_{ij} \delta_{kl} - \delta_{ik} \delta_{jl} \right] \), a purely diagonal traceless tensor with an equivalent force representation as shown in Fig. 3b. P waves are longitudinally polarized, so the direction of first motion for P arrivals will be away from the source (compressions) in the NE and SW quadrants and toward the source (dilatations) in the NW and SE quadrants. The outward-pointing arrows in the compressional quadrants are along \( T \), the tension axis, and the inward-pointing arrows in the dilational quadrants are along \( P \), the pressure axis. The direction orthogonal to both \( P \) and \( T \) is \( B \), the null axis, which is along the direction within the fault plane that is perpendicular to the slip direction. \( B \) is also orthogonal to \( a \) and \( n \).

There are three equivalent conventions for defining the three parameters that specify a unique fault plane solution. Two conventions involve specifying the directions of vectors on the focal sphere, a conceptual sphere of unit radius centered at the point source. These two parameterizations are to specify the trend and plunge either of \( \hat{a} \) and \( \hat{n} \) or of \( P \) and \( T \). (Orthogonality provides one constraint condition, so there are indeed only three independent parameters in each case.) The \( P \), \( T \) parameterization is the one most used in the analysis of deep-focus earthquakes for which the local stresses are modeled as tensional or compressional (see Earthquake Mechanisms and Plate Tectonics). The third way is to specify the dip (\( \delta \)) and strike (\( \phi \)) of the fault plane and the rake (\( \lambda \)) associated with the direction of motion in that place (Fig. 2). There are several conventions for these angles, but the current standard is the one used by Herrmann (1975) and Aki and Richards (1980, p. 106). A fault has two surfaces: the lower surface, shown in Fig. 2, is the foot wall, and the upper surface, not shown, is the hanging wall. The strike is the azimuth of the fault with the convention that if one faces down dip, the strike direction is to the left. The dip is measured down from the horizontal and is bounded by \( 0^\circ \) and \( 90^\circ \). The rake angle is measured within the fault plane and is bounded by \(-180^\circ \) and \(+180^\circ \) measured from the strike direction. If \( 0^\circ < \lambda < +180^\circ \), the fault is called a reverse fault or thrust fault; if \(-180^\circ < \lambda < 0^\circ \), the fault is referred to as a normal fault. If \( \lambda = +90^\circ \) or \( \lambda = -90^\circ \), the fault is called dip-slip. A right-lateral fault is one for which an observer standing on one side of the fault sees the block on the other side move to the right. If \( \lambda = 0^\circ \), the fault is left-lateral strike-slip (as in Fig. 3); if \( \lambda = +180^\circ \), the fault is designated right-lateral strike-slip. Herrmann (1975) includes a complete discussion about relationships among these different parameterizations.

As a consequence of the symmetry of \( \hat{a} \) and \( \hat{n} \) in Eq. 1, there is no way to distinguish on the basis of far-field data between the above interpretation and one in which the roles of \( \hat{a} \) and \( \hat{n} \) are reversed. The plane normal to \( \hat{a} \) is referred to as the auxiliary fault plane. The decision as to which is the "real" fault plane requires other kinds of data, such as correlation with other events in the region, location on a known fault, and/or assumptions about the regional stress.

If the far-field displacement is evaluated for the
positive up and backwards, and SH is positive to the right. Expressions for $u_P$, $u_{SV}$, and $u_{SH}$ are then of the following form: $u_P = (v_P/v_S)F_{RP}$, $u_{SV} = F_{RSV}$, and $u_{SH} = F_{RSH}$, where $v_P$ and $v_S$ are, respectively, the P and S velocities at the source, $F$ is a common factor independent of angle, and the $R$'s are normalized radiation factors of the form (e.g., Aki and Richards, 1980, pp. 113–114)

$$R_P = r/M_p N_k = 2(\hat{a} \cdot \hat{r})(\hat{n} \cdot \hat{r})$$
$$R_{SV} = r/M_p N_k = (\hat{a} \cdot \hat{r})(\hat{n} \cdot \hat{r}) + (\hat{a} \cdot \hat{r})(\hat{n} \cdot \hat{r})$$
$$R_{SH} = r/M_p N_k = (\hat{a} \cdot \hat{r})(\hat{n} \cdot \hat{r}) + (\hat{a} \cdot \hat{r})(\hat{n} \cdot \hat{r})$$

The predicted P-wave and SH-wave radiation patterns for the fault plane orientation shown in Fig. 3 are given in Fig. 1. In this case the projection of the SV radiation pattern is identical in shape to the P radiation pattern, but it has a different dependence on $\theta$. The surface upon which the radiation for a given wave type is zero is called a nodal surface. For a vertical strike-slip fault the nodal surfaces are planes; for P and SV the nodal planes are the fault and auxiliary planes; for SH the nodal planes are those containing the B-axis (z in this case) and the P- and T-axes. For P radiation, the nodal surfaces are always planes; for SV and SH the nodal surfaces are generally not planes. As can be seen from the example discussed later and shown in Fig. 4, the stereographic representation of the SV or SH nodal surfaces is much more complex than the projection of the P nodal planes.

### Solutions from Polarities

The method to obtain estimates of the focal mechanism from observed polarities is as follows:

1. For each observed polarity one determines the azimuth $\phi_d$ and the takeoff angle with the downward vertical $\theta$. Before about 1970, $\theta$ estimates were gotten from tables; now $\theta$ is usually calculated on the computer by a ray-tracing program. (Many earthquake location programs include $\theta$ and $\phi_d$ in the output.)

2. Symbols representing compression or dilatation are plotted for each station on a planar projection of the focal sphere. The first studies of focal mechanisms used teleseismic data, so the rays leading to the stations started downward. Hence the convention is to use a lower hemisphere projection. Geologists generally use a Wulf projection because it preserves angles, but for the teleseismic data such a projection complicates interpretation because the plot becomes too cluttered near the center. Accordingly, the equal-area Lambert-Schmidt projection is used. For this projection, if the takeoff angle $\theta$ is less than $90^\circ$, the $(r, \phi)$ position for a polarity is $\phi = \phi_d$ and $r = \sqrt{2r_0 \sin(\theta/2)}$, where $r_0$ is the radius of the circle. If $\theta > 90^\circ$, $\phi = 180^\circ + \phi_d$ and $r = \sqrt{2r_0 \sin(\theta/2)}$.
FIGURE 4. Data and focal mechanism solutions for an \( m = 2.2 \) earthquake recorded on vertical seismometers: (a) triangles represent dilatations, circles represent compressions, and the \( X \) sizes are proportional to the \( \log(\text{SV}/\text{P}) \) ratios. Solid lines are \( P \) nodal plane projections, and dashed lines are \( SV \) nodal surface projections for solutions for which no polarity errors are allowed and for which the ratio misfit is less than 70%; (b) \( P \) nodal plane projections for solutions for which no polarity errors are allowed, but one ratio misfit is allowed to be greater than 70%. The station polarity ratio circled in (a) is in error for the solutions in (b) that have one nodal plane poorly constrained.

\[ \sqrt{2} r_0 \sin(90 - \theta/2). \]  
(For a point source, rays reflected through the origin have identical properties.) Figure 6 in Earthquake Mechanisms and Plate Tectonics shows examples of focal sphere projections for strike-slip and thrust fault mechanisms. (The compressive quadrants are shaded.)

3. Determine the best-fit focal mechanisms. Previously this was done by picking visually the perpendicular nodal planes that separate the dilatational and compressional quadrants. Now possible solutions are generally calculated by computer programs that systematically search the focal sphere for all possible solutions consistent with the data. The input typically includes the polarity data and the search criteria: the number of "acceptable" errors, the region of the focal sphere to be searched, and the fineness of the grid. The output includes all solutions that fit the criteria.

An extension of the above procedure is used when polarity data are available from several events in a given region. In this procedure the data are combined and a composite solution is obtained. Such solutions are particularly useful for attempts to infer regional stresses from focal mechanisms.

Solutions from Ratios and from Waveform Fitting

If the station coverage were dense and uniform in both distance and azimuth for a given event, observed direct-arrival first motions would tightly constrain the event's focal mechanism. Generally this is not the case, but there is considerably more information derivable from seismograms which can potentially be used. For example, it follows from the above discussion that comparisons between predicted and observed amplitude ratios from among \( P, SV, \) and \( SH \) arrivals provide possible constraints. For many events data are only from vertical seismometers, which still allows observations of \( (\text{SV}/\text{P})_3 \). A disadvantage of this ratio is that \( SV \) is often not well recorded on the vertical and it is the most affected by earth structure due to \( S \)-to-\( P \) converted energy. However, for small earthquakes with observations at local or regional distances, these ratios have proved to be a useful supplement to polarity data in many studies. For discussions of \( (\text{SV}/\text{P})_3 \), see Kisslinger (1980, 1982). Other researchers have used synthetics to aid in the identification of depth phases such as \( P \) and \( S \) at teleseismic distances and then used amplitude ratios and relative polarities between them or the direct \( SH \) and the direct \( P \) to constrain the mechanisms (e.g., Langston et al., 1982).

Waveform-fitting procedures to find focal mechanisms include implicit polarity and ratio comparison techniques, and they are the only methods available for the analysis of complex events. The development and implementation of these procedures have grown considerably in the last decade because of the advent of broadband three-component digital data, refined estimates of local and regional velocity structures, and advances in synthetic seis-
EARTHQUAKE MECHANISMS

Seismograms (see Seismograms: Synthetic). Waveform-fitting procedures to invert for focal mechanisms have become almost routine for large \((m_b \geq 5.5)\) events and are increasingly used for smaller events as well.

For large earthquakes recorded at telescismic distances the best-quality data have been long-period seismograms. Such data are easier to model than are short-period data because long-period waves are less sensitive to source and receiver site effects, attenuation, and details of the rupture process. Green’s functions for synthetics have been constructed using normal mode theory (e.g., Dziewonski et al., 1981), WKBJ theory (e.g., Sipkin, 1982), and generalized ray theory (e.g., Langston et al., 1982).

As discussed in detail by Sipkin (1982), there are two different inversion procedures commonly used: the construction method and the appraisal method. Both methods start with the same set of equations: Each far-field displacement component at each station is written as linear superpositions from among five terms (assuming a purely deviatoric moment tensor), where the terms are proportional to linear combinations of elements of the moment rate tensor—the time derivative of the moment tensor (see Seismograms: Synthetic and Aki and Richards, 1980, p. 79). For the construction method, one varies the relative weighting of the moment rate tensor terms to optimize the waveform fit at all stations and for all components simultaneously. Solutions have been evaluated by a systematic grid search, by a least squares inversion procedure, or by some combination of the two. In the appraisal method, which utilizes linear inverse theory, filters are constructed that, when convolved with the time series data, yield unique time averaged estimates of the individual elements of the moment rate tensor. A disadvantage of this method is that all three components cannot be used simultaneously in the formal inversion, because SH is decoupled from P and SV in the far field. The appraisal method, as developed by Dziewonski et al. (1981), is applied routinely by the National Earthquake Information Center in preparing its preliminary determinations of epicenters.

The outputs from both the construction and appraisal methods are estimates of elements of the moment rate tensor. The necessary condition that a traceless moment rate tensor correspond to a single double-couple point source is that one of the principle values be zero. Such a condition cannot be included as a linear constraint in the inversion process, so the moment rate tensor found by inversion will in general not correspond to a single double-couple. As Dziewonski et al. (1981, p. 2837) point out, there is no unique way to project out a double-couple solution from a purely deviatoric moment tensor; the method currently used by the National Earthquake Information Center is to identify T with the maximum eigenvalue and P with the minimum eigenvalue.

As noted by Sipkin (1982), focal mechanisms found from the inversion of entire waveforms correspond to averages over the spatiotemporal dimensions of the source, so the results may not be entirely consistent with solutions implied by first motions that include information about only the initiation of rupture.

Reliability of Focal Mechanisms

The reliability of focal mechanisms inferred from polarity and amplitude ratio data depends on the quality of the data. Aspects to be considered include the reliability of the individual polarities and/or amplitudes themselves, the accuracy of the location of the event, the correctness of the assumed velocity structure, and the distribution and density of the data on the focal sphere. Arrivals may be impulsive or emergent, and for emergent arrivals the pick of a direction for the initial motion may be in error. A low signal-to-noise ratio, particularly on short-period records, may cause errors in polarity picks, as may small amplitudes caused by the proximity to a nodal surface. If the station coverage of the focal sphere is not uniform, the nodal surfaces may be poorly constrained. In addition to the problems inherent in polarity data, ratio data may be in error because of interference among arrivals and/or inaccurate anelasticity or free surface corrections. Errors in hypocenter estimates will shift the positions of the stations on the focal sphere and will accordingly distort the patterns in the projected data. This can be quite important for stations located near nodal surfaces. Changing the velocity structure at the source has a usually small but systematic effect on the takeoff angles. For crustal events observed at regional distances, using a velocity structure composed of constant-velocity layers (a practice commonly used in locating earthquakes), the takeoff angles will appear unnaturally “quantized” at the critical refraction takeoff angles for each boundary below the hypocenter. This feature can be removed by using gradient velocity models.

In some currently available focal mechanism computer programs the user may specify an allowed number of polarity errors and an allowed range in amplitude ratio error for a solution to be acceptable. Such programs may also allow for weighted errors, depending upon the proximity to the predicted nodal surface, the logic being that errors far from the nodal surface should be more significant. One generally plots all acceptable solutions on a single plot: P nodal planes and/or projections of the P, T+, and B-axes. The range in acceptable solutions will then show—qualitatively at least—how well constrained the solutions are.

The effects of sparse and nonuniform coverage can be seen from the following example. Figure 4 shows polarity and ratio data along with inferred focal mechanisms for a small \((m_b = 2.2)\) Virginia earth-
Earthquake recorded on vertical seismometers in the Virginia Tech Seismic Network. There are five reliable P polarities and nine (SV/P); ratios, and the coverage of the focal sphere is far from uniform. Because four of the polarities are dilatations, the focal mechanism would be very poorly constrained on the basis of polarities alone. Shown in Fig. 4a are the P nodal planes, the SV nodal surfaces, and the projections of the P-, T-, and B-axes for a grid search of the focal sphere with no allowed polarity or ratio errors (where a ratio error is here defined as a ratio mismatch greater than 70%). The three solutions shown are all essentially the same. Allowing one polarity error but no ratio errors produces no significantly new solutions. Figure 4b shows the P nodal planes for no polarity errors but one ratio error. Although one nodal plane remains well constrained, the other is now almost unconstrained. For this station event geometry, the fact that one observed ratio (circled in Fig. 4a) is very small requires that it lie near an SV nodal surface. Allowing that ratio to be in error removes that constraint and allows for many different solutions that are consistent with the data. A (qualitative) interpretation is that one nodal plane is well constrained, but the other one is only marginally constrained.

At present there is no standard method to assess quantitatively the reliability of a solution. The discrete nature of polarity data precludes the use of least squares inversion procedures [as in earthquake locations (see Earthquakes: Location Techniques)]. Kisslinger (1980) employs a nonlinear least squares inversion procedure for the dip, strike, and rake when only amplitude ratio data (which are continuously varying) are used, which allows him to specify confidence regions for these angles. Julian (1986) shows how linear programming procedures can be applied to solve for focal mechanisms; such procedures can include both continuous and discrete data. Other researchers have formulated probability models in which a likelihood function can be constructed based on polarity and/or S-wave polarization data. Such a formulation allows the calculation of confidence regions for focal mechanisms, but the implicit assumptions in such a procedure are most likely to be valid when there are many observations.

The reliability of focal mechanisms inferred from waveform fitting is affected by all the factors described here, except that errors in focal depth and in the assumed near-station velocity structure are more important than in the other inversion schemes. Only through waveform fitting is it possible to deal with events that are too complex to be approximated adequately by a point source (see Seismic Source: Observations). For solutions found by the appraisal method, confidence regions for the solutions can be obtained directly. For focal mechanisms found from waveform fitting by the construction method, a $5 \times 5$ covariance matrix can be obtained for the elements of the deviatoric moment rate tensor, and then estimates of the variances in the focal mechanisms can be calculated.

J. A. SNOKE

References
Herrmann, R. B., 1975, A student’s guide to the use of P and S wave data for focal mechanism determination, Earthquake Notes 46, 29–39.

Cross-references: Earthquake Mechanisms and Plate Tectonics; Earthquakes: Hazards and Prediction; Earthquake Seismology; Seismic Monitoring of Nuclear Explosions; Seismic Source: Observations; Seismic Source: Theory.

Earthquake seismology has played a major role in the development of our current understanding of global plate tectonics. Because earthquakes occur primarily at the boundaries between lithospheric plates, their distribution is used to map plate boundaries and their focal mechanisms provide information about the motion at individual boundaries.

Plate boundaries are divided into three types (Fig. 1). Oceanic lithosphere is formed at spreading centers, or mid-ocean ridges, and is destroyed at subduction zones, or trenches. Thus, at spreading centers plates move away from the boundary, whereas at subduction zones the subducting plate moves toward the boundary. At the third boundary type, transform faults, plate motion is parallel to the