EARTH STRUCTURE, MAJOR DIVISIONS

Although the Earth is a complex body with pervasive three-dimensional structure in its solid portions, the dominant variation in properties is with depth. The figure of the Earth is close to an oblate spheroid with a flattening of 0.003356. The radius to the pole is 6357 km and the equatorial radius is 6378 km; for most purposes, a spherical model of the Earth with a mean radius of 6371 km is adequate. Thus, reference models for internal structure can be used in which the physical properties depend on radius. Three-dimensional variations can then be described by deviations from a suitable reference model.

The main divisions of Earth structure are illustrated in Figure E1. Beneath the thin crustal shell lies the silicate mantle which extends to a depth of 2890 km. The mantle is separated from the metallic core by a major change of seismic properties that has a profound effect on global seismic wave propagation (see Seismic phases). The outer core behaves as a fluid at seismic frequencies and does not allow the passage of shear waves, while the inner core appears to be solid. Thus only seismic compressional waves (P) can traverse the outer core.

Much of the evidence for the nature of internal structure comes from the analysis of seismograms and the patterns of propagation of seismic waves.

The existence of a discontinuity at the base of the crust was found by Mohorovičić in the analysis of the Kupatal earthquake of 1909, based on only a limited number of records from permanent seismic stations. Knowledge of crustal structure from seismic methods has developed substantially through the use of controlled sources, e.g., explosions. Indeed, most of the information on the oceanic crust comes from such work. The continental crust varies in thickness from around 20 km in rift zones to 70 km under the Tibetan Plateau. Typical values are close to 35 km. The oceanic crust is thinner with basalt pile about 7 km thick whose structure changes somewhat with the age of the oceanic crust.

The lithosphere continues from the crust into the mantle and also shows significant differences between the oceanic and continental regimes. The entire oceanic lithosphere is generated by the spreading process at mid-ocean ridges and the increase in thickness with age to at least 85 Ma is consistent with thermal cooling. Precambrian shield components of continents have a very thick but lower density lithosphere extending to 200 km (or possibly more in some places). The lithosphere beneath Phanerozoic regions tends to be thinner, about 120 km, with considerable complications in the neighborhood of active tectonic belts.

Beneath the lithosphere lies the asthenosphere that is typically characterized by a decrease in shear wavespeed, increased attenuation and electrical conductivity, and inferred lower viscosity. The transition from lithosphere to asthenosphere appears to be generally sharp in oceanic regions, but probably is more gradational under the continents.

The mantle shows considerable variation in seismic properties with depth, with strong gradients in seismic wavespeed in the top 800 km. The presence of distinct structure in the upper mantle was recognized by Jeffreys in the 1930s who noted a distinct shift in the rate of change of the travel-times of P-waves with distance near 20° from the source. He ascribed this 20° discontinuity to a rapid change in seismic wavespeed with depth. In the 1960s–1970s detailed analysis of the seismic wave field in the distance range from 15° to 25° from earthquake and nuclear explosion sources, using seismic arrays, revealed two clear discontinuities in seismic wavespeed at depths near 410 and 660 km. From mineral physics results, these discontinuities can be associated with major phase changes in the silicate minerals of the mantle. A number of other discontinuities have been proposed. Only a feature near 520 km depth has a widespread occurrence in long-period observations, and appears to represent a small change over a 30–50 km interval in depth (Shearer, 1993). Many different styles of study have shown both the global presence of discontinuities near 410 and 660 km depth, and also significant variations in seismic structure within the upper mantle (for a review see Nolet et al., 1994).

The reduction of seismic wavespeeds in the lowermost mantle led Bullen to identify a subregion D" of the lower mantle (region D). This zone situated just above the core-mantle boundary has proved to have heterogeneity comparable to that in the uppermost mantle. There are localized discontinuities in shear wavespeed, as first noted by Lay and Helmberger (1983), as well as narrow zones of "ultralow velocity" (Ganero et al., 1998).

The need for a core at depth with greatly reduced seismic wavespeeds was recognized at the end of the last century by Oldham in his analysis of the great Assam earthquake of 1890, because of a zone without distinct P arrivals (a "shadow zone" in PKP, see Seismic phases). By 1914, Gutenberg had obtained an estimate for the radius of the core which is quite close to the current value. The presence of the inner core was inferred by Inge Lehmann in 1932 from careful analysis of arrivals within the shadow zone (PKIKP), which had to be reflected from some substructure within the core.

The nature of the transition form the outer core to the inner core has been a matter of some debate. Early analysis by Jeffreys, based on a limited number of reports of the travel-times of PKP phases, produced a model with a decrease in P-wavespeed just above the inner core (the
Figure E1 The major divisions of the radial structure of the Earth: the gradations in tone in the upper part of the mantle indicate the presence of discontinuities at 410 and 660 km depth. The diagram is drawn to true scale, so that the Earth's crust appears only as the thin bounding surface.

Figure E2 Radial reference earth model AK135, seismic waves speeds $\alpha$ (P), $\beta$ (S): Kennett et al. (1995), attenuation parameters $Q$, density ($\rho$): Montagner and Kennett (1996). Zones marked with gray tones show significant lateral heterogeneity and some level of anisotropy.

F-region in Bullen's notation). In contrast, Gutenberg working with seismograms recorded at Pasadena, California, favored a simple discontinuity with an increase in wavespeed in the inner core. The difficulties arise because of the very complex patterns of arrival expected for the PKP phase (see Kennett, 2002, Chapter 26). Modern analysis using the differential times between arrivals following different paths through the core, or matching observed and computed seismograms, favor a decrease in wavespeed gradient at the base of the outer core, but not a low velocity zone.

The times of arrival of seismic phases on their different paths through the globe constrain the variations in P- and S-wavespeed, and can be used to produce models of the variation with radius. A very large volume of arrival time data from stations around the world have been accumulated by the International Seismological Centre and are available in digital form. This data set has been used to develop high-quality travel-time tables that can in turn be used to improve the locations of events. With reprocessing of the arrival times to improve locations and the identification of the picks for later seismic phases, a set of observations of the relation between travel-time and epicentral distance have been produced for a wide range of phases. The reference model AK135 of Kennett et al. (1995) for both P- and S-wavespeeds, illustrated in Figure E2, gives a good fit to the travel-times of mantle and core phases. The reprocessed data set and the AK135 reference model have formed the basis of much recent work.
on high-resolution travel-time tomography to determine three-dimensional variations in seismic wavespeed.

One of the difficulties in just using the travel-time relations comes from the sampling of the core by seismic phases. For arrivals traveling solely as P-waves the change in material properties at the core-mantle boundary means that the paths become steeper on entry into the core from Snell's law. Eventually, the effects of sphericity and the gradients in seismic wavespeed are sufficient to turn energy back to the surface but the outermost part of the core is not sampled by the PKP wave. In contrast, the P-wavespeed in the core is a little larger than the S-wavespeed at the base of the mantle and so the converted arrival SKS can sample the whole core. The differing patterns for PKP and SKS are shown in Figure E3 (see also Seismic phases). From travel-times alone, the P-wavespeeds at the top of the outer core depend on knowledge of the S-wavespeed throughout the mantle.

The use of the times of arrival of seismic phases enables the construction of models for P- and S-wavespeed, but more information is needed to provide a full model for Earth structure. The density distribution in the Earth has to be inferred from indirect observations and the main constraints come from the mass and moment of inertia. The mean density of the Earth can be reconciled with the moment of inertia if there is a concentration of mass toward the center of the Earth; which can be associated with a major density jump going from the mantle into the outer core and a smaller density contrast at the boundary between the inner and outer cores (Bullen, 1975). With successful observations of the free-oscillations of the Earth following the great Chilean earthquake of 1960, additional information could be extracted from the frequencies of oscillation on both the seismic wavespeeds and the density. Fortunately the inversion of the frequencies of the free-oscillations for a spherically symmetric reference model provides independent constraints on the P-wavespeed structure in the outer core. Even with the additional information from the normal modes the controls on the density distribution are not strong (Kennet, 1998) and additional assumptions such as adiabatic state in the core and lower mantle have often been employed to produce a full model. The reference model PREM of Dziewonski and Anderson (1981) combined the free-oscillation and travel-time information available at the time. A parametric representation of structure was employed in terms of simple mathematical functions to aid the inversion; thus a single cubic was used for seismic wavespeed in the outer core and again for most of the lower mantle. The PREM forms the basis of much current global seismology using quantitative exploitation of seismic waveforms at longer periods (e.g., Dahlen and Tromp, 1998).

In order to reconcile the information derived from the free-oscillations of the Earth and the travel-time of seismic phases, it is necessary to take account of the influence of anelastic attenuation within the Earth. A consequence of attenuation is a small variation in the seismic wavespeed with frequency, so that waves with frequencies of 0.01 Hz (at the upper limit of free-oscillation observations) travel slightly slower than the 1-Hz waves typical of the short period observations used in travel-time studies. The density and attenuation models shown in Figure E2 were derived by Montagner and Kennett (1996) to link the wavespeed distributions from travel-time analysis to the free-oscillation results.

Figure E4 displays the travel-times for the ax135 model superimposed on the reported phases for a set of 104 test events (83 earthquakes, 21 explosions) assembled by Kennett and Engdahl (1991). These events have well-controlled locations and origin times with a rich set of later phase readings (57655 phases in all). As we can see the reference model provides a good representation of the features
of the full set of phase picks (for explanation of the phase codes see Seismic phases).

The reference model AK135 (Figure E2) is isotropic and depends solely on radius. However, the real Earth both varies in three-dimensional and displays anisotropy in seismic parameters. The zones of largest variability are indicated in Figure E2 by bands of gray tone. These are also regions in which there is the strongest evidence for anisotropy in seismic properties, from differences in seismic wavespeeds depending on the direction of propagation or the polarization of the seismic waves.

The strongest levels of heterogeneity are found in the outermost 300 km with a strong contrast between the high wavespeeds beneath the ancient cores of continents extending to at least 250 km and the oceanic regime where the fast lithosphere is quite thin. The continental lithosphere also appears to be anisotropic, with evidence from the analysis of seismic surface waves and the shear wave splitting for SKS waves.

The process of subduction brings the cold oceanic lithosphere into the upper mantle and locally there are large contrasts in seismic wavespeeds, well imaged by detailed tomography, that extend down to at least 660 km and in some zones even deeper. Remnant subducted material can have a significant presence in some regions, e.g., above the 660 km discontinuity in the north-west Pacific and in the zone from 660 down to 1000 km beneath Indonesia.

Mantle
The nature of mantle structure varies with depth and it is convenient to divide the mantle up into four major zones (e.g., Jackson and Ridglen, 1998).

Upper mantle (depth $z < 350$ km), with a high degree of variability in seismic wavespeed (exceeding $\pm 4\%$) and relatively strong attenuation in many locations.

Transition zone ($350 < z < 800$ km), including significant discontinuities in P- and S-wavespeeds and generally high velocity gradients with depth.

Lower mantle ($800 < z < 2600$ km) with a smooth variation of seismic wavespeeds with depth that is consistent with adiabatic compression of a chemically homogeneous material.

D* layer ($2600 < z < 2900$ km) with a significant change in velocity gradient and evidence for strong lateral variability and attenuation.

The two major discontinuities in seismic wavespeeds near depths of 410 and 660 km are associated with phase transformations in silicate minerals induced by the effects of increasing pressure, and represent changes in the organization of the oxygen coordination with the silicon atoms. The change in seismic wavespeed across these discontinuities occurs quite rapidly, and they are seen in both short- and long-period observations. Other minor discontinuities have been proposed, but only one near 520 km appears to have some global presence in long-period stacks but not short-period data. This 520-km transition may occur over an extended zone, e.g., 30–50 km, so that it still appears sharp for long-period waves with wavelengths of 100 km or more. Jackson and Ridglen (1998) provide a broad ranging review of the interpretation of seismological models for the transition zone and their reconciliation with information from mineral physics.

Frequently, the lower mantle is taken to begin below the 660 km discontinuity, but strong gradients in seismic wavespeeds persist to depths of the order of 800 km and it seems appropriate to retain this region in the transition zone. There is increasing evidence for localized sharp transitions in seismic properties at depth around 900 km that
appear to be related to the penetration of subducted material into the lower mantle.

Between 800 and 2600 km, the lower mantle has, on average, relatively simple properties which would be consistent with the adiabatic compression of a mineral assemblage of constant chemical composition and phase. Although tomographic studies image some level of three-dimensional structure in this region the variability is much less than in the upper part of the mantle or near the base of the mantle.

The D' layer from 2600 km to the core-mantle boundary has a distinctive character. The nature of seismic wavespeed distribution changes significantly with a sharp drop in the average velocity gradient. There is a sharp increase in the level of wavespeed heterogeneity near the core-mantle boundary compared with the rest of the lower mantle. The base of the Earth's mantle is a complex zone with widespread indications of heterogeneity on many scales, discontinuities of variable character, and shear wave anisotropy (e.g., Gurnis et al., 1998; Kennett, 2002). The results of seismic tomography give a consistent picture of the long wavelength structure of the D' region; there are zones of markedly lowered S-wavespeed in the central Pacific and southern Africa, whereas the Pacific is ringed by relatively fast wavespeeds that may represent a "slab graveyard" arising from past subduction. Discordance between P- and S-wave results suggests the presence of chemical heterogeneity rather than just the effect of temperature (e.g., Masters et al., 2000).

There are a number of lines of evidence which suggest the presence of topography on the core-mantle boundary, but the situation is complicated by the influence of the strong heterogeneity in D'. The information from travel-times of seismic phases such as (PcP and PKP, PKKP) is compatible with long wavelength topography on the order of ±3 km, but such variations are not required by this data. As noted by Buffet (1998), topography of several kilometers on wavelengths of several thousand kilometers is not sufficient to produce the torques between mantle and core inferred from variations in the length of the day. However, such topographic coupling could become important if the core-mantle boundary is rough at smaller wavelengths as suggested from some studies of seismic scattering.

Evidence for a discontinuity at the top of the D' region comes from the presence of an additional arrival between the direct mantle phase and the core reflection in the distance range 65°-90°. This is not a universal phenomenon, but a D discontinuity is found in many parts of the world. The typical contrast in S-wavespeed is in the range 1.5%-3% at 250 ±100 km above the core-mantle boundary, with greater variability in P-wavespeed. The nature of the transition varies between different regions and in places may represent a transition zone up to 50 km thick. In some areas the fluctuations in travel-times and amplitude of the reflected phases suggest the presence of strong lateral gradients in discontinuity structure (e.g., Lay et al., 1998).

Beneath the central Pacific Ocean and in a number of other locations, there is evidence for thin zones, around 20 km thick, at the core-mantle boundary with a marked reduction in P-wavespeed of at least 10% (see Garnero et al., 1998). The inferred reduction in S-wavespeed in these ultralow velocity zones somewhat larger.

Since all waves sampling the core have to pass through the D' region, the complexity of propagation in this region influences inferences made about deeper structure.

**Core**

The core-mantle boundary at about 2890 km depth marks a substantial change in physical properties associated with a transition from the silicate mantle to the metallic core (see Figure E3). There is a significant jump in density, and a dramatic drop in P-wavespeed from 13.7 to 8.0 km s⁻¹. The major change in wavespeed arises from the absence of

---

**Figure E5** Comparison of P-wavespeed models for the core: IASP91, Kennett and Engdahl, 1991; PREM2, Song and Helmberger, 1992; SP6, Morelli and Dziewonski, 1993; AK135, Kennett et al. 1995.
shear strength in the fluid outer core, so that the P-wave speed depends just on the bulk modulus and density. No shear waves can be transmitted through the outer core so that only the K propagation legs can be present (see Seismic phases).

The process of core formation requires the segregation of heavy iron-rich components in the early stages of the accretion of the earth (e.g., O’Neill and Palme, 1998). The core is believed to be largely composed of an iron-nickel alloy, but its density requires the presence of some lighter elemental components. A wide variety of candidates have been proposed for the light components, but it is difficult to satisfy the geochronological constraints on the nature of the bulk composition of the Earth.

The inner core appears to be solid and formed by crystallization of material from the outer core, but it is possible that it could include some entrained fluid in the top 100 km or so. The shear wave speed for the inner core inferred from free-oscillation studies is very low and the ratio of P- to S-wave speeds is comparable to a slurry-like material. The structure of the core is both anisotropic and shows three-dimensional variation (e.g., Creager, 1999). The variations are complex with some apparent variation between hemispheres, but may be influenced by the phase of phases such as PKIKP through the strong heterogeneity at the base of the mantle in D".

The fluid outer core is conducting and motions within the core create a self-sustaining dynamo which generates the main component of the magnetic field at the surface of the Earth. The dominant component of the geomagnetic field is dipolar but with significant secondary components. Careful analysis of the historic record of the variation of the magnetic field has led to a picture of the evolution of the flow in the outer part of the core (e.g., Bloxham and Gubbins, 1989). The presence of the inner core may well be important for the action of the dynamo, and electromagnetic coupling between the inner and outer cores could give rise to differential rotation between the two parts of the core (Glazmaier and Roberts, 1996). Efforts have been made to detect this differential rotation using the time history of different classes of seismic observations but the results are currently inconclusive.

Even though the main features of the variation in seismic wave speeds through the core are well established, there are noticeable differences in the details of models proposed by different authors for the regions near the core-mantle boundary and the boundary between the inner core and outer core (Figure E5). The detailed structure just below the core-mantle boundary is primarily controlled by the properties of the multiple reflections SKKS, SKKKS, ... from the underside of the core-mantle boundary. The most detailed models of the structure of the boundary region between the inner and outer cores come from matching of observed and calculated waveforms as in the PREM2 model of Song and Helmberger (1992). This work indicates the need for a reduction in seismic wave speeds just above the inner core boundary, and is supported by studies of the differential times between the branches of the PKP phase that take different paths through the core. However, such studies are susceptible to the influence of heterogeneity and anisotropy in the inner core.

Differential time information was used in the construction of the AK135 model, which is close to PREM2, and differs from S66 (Morelli and Dziewonski, 1993) that uses a single cubic representation for P-wave speeds throughout the whole core, in a similar way to PREM (Dziewonski and Anderson, 1981).

Bibliography


Cross-references
Core, Boundary Layers
Core Properties, Physical
Core-Mantle Boundary Topography, Seismology
D°, Anisotropy
D°, as a Boundary Layer
D°, Composition
D°, Seismic Properties
Inner Core Anisotropy
Inner Core Seismic Velocities
Interiors of Planets and Satellites
Lehmann, Inge (1888-1993)
Oldham, Richard Dixon (1858-1936)
Seismic Phases

ELSASSER, WALTER M. (1904-1991)

About 2000 years ago in China it was discovered that a lump of magnetic iron oxide (magnetite), or lodestone, takes up a preferred orientation when freely suspended. Ultimately this discovery led to the magnetic compass for navigation and surveying. Four centuries ago Gilbert (1600) (q.v.), in England, carried out precise laboratory measurements of the magnetic field around a lodestone. The magnetic records accumulated by the seafaring captains of the time made it clear to Gilbert that the magnetic field of Earth has the same form as that of a lodestone, and he concluded that Earth was itself a huge lodestone.

In 1835, Gauss demonstrated mathematically that the east-west, north-south, and vertical components of the field are observed to vary relative to each other in such a way as to place the origin of the field inside, rather than outside, the body of Earth.

Gilbert's plausible interpretation of the origin of the geomagnetic field stood for about three centuries, until Kelvin in 1862 established the incandescence of the interior of Earth, and Pierre Curie in 1895 studied the abrupt disappearance of magnetic effects in metals and minerals upon heating to hundreds of degrees. The origin of the geomagnetic field became a mystery once again.

Thermoelectric effects in the hot inhomogeneous interior of Earth were a favorite notion, creating electric currents and the associated magnetic fields. However, it was not obvious how to account for a magnetic field as strong as the observed polar 0.6 G, or for the observed secular variations of the field, with characteristic times as short as a few centuries.

Walter M. Elsasser became interested in the problem in the late 30s and investigated the possibility that the thermoelectric effects in the temperature variations in the convective liquid iron and nickel core might provide the observed secular fluctuations of the field observed at the surface of Earth (Elsasser, 1939). Two years later Elsasser (1941) studied the secular variations in some detail and showed that they could be modeled by several randomly oriented dipoles if those dipoles were confined to the liquid core (r < 0.5R_{\oplus}).

The essential role of convection in providing the secular variation turned Elsasser's attention to the possibility that the magnetic field is induced directly by the convective motions in the electrically conducting core. Convinced by then that the thermoelectric effect was too weak to provide the observed magnetic field, Elsasser set about developing the formal mathematical theory of magnetic induction by fluid motions in a sphere (Elsasser, 1946a, b, 1947). Elsasser recognized the fact that the large-scale (10^4 km) of the principal convective motions in the liquid iron core, combined with the substantial electrical conductivity, τ = 10^{26} s^{-1} (resistive diffusion coefficient η = c^2/4πα = 10^{-8} cm s^{-1}), coupled the magnetic field and the liquid iron over periods of 10^3-10^4 years. This matched the comparable characteristic turnover and nonuniform rotation times of the liquid iron core. So he decomposed the geomagnetic field B into what he called the toroidal (azimuthal) and poloidal (meridional) components and expressed the poloidal field in terms of a toroidal vector potential, expanding the whole in spherical harmonics. The fluid velocity v in the core was expanded in a similar form so that the magnetohydrodynamic induction equation

\[ \frac{\partial B}{\partial t} - \eta \nabla^2 B = \nabla \times (v \times B) \]

providing the growth and decay of each magnetic mode in terms of the field and fluid interaction matrix on the right-hand side.

Elsasser showed that the principal magnetic field in the core is the toroidal magnetic field, created by the interaction of the nonuniform rotation of the convecting core with the main poloidal (dipole) component of the field. Up to this point the poloidal and toroidal fields had rotational symmetry, and Elsasser was aware of Cowling's (1933) theorem (see Cowling's theorem) that a magnetic field with rotational symmetry cannot be sustained by fluid motions. Elsasser added quadrupole terms to the fluid motions, thereby breaking the rotational symmetry.

Unfortunately the system of equations for the mode amplitudes showed no signs of converging. The mathematical effort was carried forward, using computers to implement the mathematical analysis, by Bullard (see Bullard, Edward Crisp and Dynamo, Bullard-Gellman) and others (see Rikitake, Tsuneki) but without success. Pekeris, Accad, and Shkoller (1973) got around the difficulties by using a fluid velocity field with strong helicity, so that it required only a small velocity to regenerate the field, and their formal result evidently converges satisfactorily. Gubbin (1973) was able to apply computers to the formal analysis to show that the general method could be made to converge, providing a reliable physical result.

Parker, working as a research associate with Elsasser, noted that the net magnetic circulation in meridional planes (representing the dipole field) is created directly by cyclonic convective cells that push up local Ω-shaped loops in the azimuthal field. The cyclonic cells rotate the many Ω-loops into the meridional planes, where they spread out and merge through resistive diffusion, producing large-scale magnetic circulation, i.e., the dipole component of the poloidal field. The effect is described by the sum of the azimuthal vector potentials A_φ contributed by each small rotated -loop. The net effect is described by the dynamo equation

\[ \frac{\partial A_φ}{\partial t} - \eta \nabla^2 A_φ = \Gamma(r) B_φ \]

where the scalar function \( \Gamma(r) \) (nowadays written \( a(r) \)) is essentially the rotational velocity of the convective cells multiplied by the fraction of space occupied by the cells. Combining this with the Elsasser' equation for the azimuthal field \( B_φ \) yields a complete set of dynamo equations (see Dynamos, mean field) easily solved in rectangular geometry in terms of dynamo waves (see Dynamos, kinematic; Dynamo waves), and showing in simple terms how the physics of Elsasser's induction works out in Earth and in the Sun (Parker, 1955).

Elsasser's fundamental dynamo concept, that the magnetic field of Earth is generated by the convection in the liquid iron core, was greeted with outspoken doubt by many eminent physicists. This negative attitude began to weaken only with the publication of his review.