12

The Shape of the Earth, Heat Flow and Convection

When Galileo let his balls run down an inclined plane with a gravity which he had chosen himself. . then a light dawned upon all natural philosophers.

—I. KANT

We now come to the more global properties of the Earth. The shape of the Earth and its heat flow are both manifestations of convection in the interior and conductive cooling of the outer layers. The style of convection, however, is unknown. There are various hypotheses in this field that parallel those in petrology and geochemistry. The end-members are whole-mantle convection in a chemically uniform mantle and layered convection with little or no interchange of material between layers. Layered schemes have several variants involving a primitive lower mantle or a depleted lower mantle. In a convecting Earth we lose all of our reference systems, including the axis of rotation.

TOPOGRAPHY

The topography of the Earth's surface is now generally well known, though some areas in Tibet, central Africa and the southern oceans remain poorly surveyed. The distribution of elevations is distinctly bimodal, with a peak near 0.1 km representing the mean elevation of continents and a peak near -4.7 km corresponding to the mean depth of the oceans. This bimodal character contrasts with that of the other terrestrial planets. The spherical harmonic spectrum of the Earth's topography shows a strong peak for $n = 1$, corresponding to the distribution of most continents in one hemisphere, and a regular decrease with increasing n. The topography spectrum is similar to that of the other terrestrial planets. There are small peaks in the spectrum at $n = 3$ and $n = 9$–10, the latter apparently corresponding to the distribution of large oceanic islands and hotspots.

In general, the most recent orogenic belts such as the Alpine and Himalaya are associated with high relief, up to 5 km, while older orogenic belts such as the Appalachian and Caledonian, because of erosion, are associated with low relief, less than 1 km. Regional changes in the topography of the continents are generally accompanied by changes in mean crustal thickness. Continents stand high because of thick, low density crust, compared to oceans.

The long-wavelength topography of the ocean floor exhibits a simple relationship to crustal age. The systematic increase in the depth of the ocean floor away from the mid-ocean ridges can be explained by simple thermal models for the evolution of the oceanic lithosphere. Parsons and Sclater (1977) using data from the western North Atlantic and central Pacific Oceans showed that, for seafloor ages from 0 to 70 Ma, topography is described by

$$d(t) = 2500 + 350t^{1/2}$$

where t is age in Ma and $d(t)$ is the depth in meters. Older seafloor does not follow this simple relationship, being shallower than predicted. There are large portions of the ocean floor whose depth cannot be explained by simple thermal models; these include oceanic islands, hotspot swells, aseismic ridges and oceanic plateaus as well as other areas where the effects of surface tectonics and crustal structure are not readily apparent. Simple cooling models assume that the underlying mantle is uniform and that all of the variation in bathymetry is due to cooling of a thermal boundary layer. The North Atlantic is generally too shallow for its age, and the Indian Ocean between Australia and Antarctica is too deep. The residual depth anomaly is the departure of the depth of the ocean from the value expected for its age and is given by

$$Ad = d(t) + \frac{S(\rho_s - \rho_w)}{(\rho_m - \rho_s)} - d$$

where $d(t)$ is the expected depth based on age, S is the sediment thickness, ρ_s, ρ_w and ρ_m are the densities of sediments, water and mantle, respectively, and d is the observed depth. Residual depth anomalies observed in the ocean have dimensions of order 2000 km and amplitudes greater than 1 km. Part of the residual anomalies are due to regional changes in crustal thickness. This cannot explain all of the anomalies. Positive depth anomalies are generally associated with volcanic regions such as Bermuda, Hawaii, the Azores and the Cape Verde Islands. These might be due to thinning of the lithosphere or the presence of abnormally hot upper mantle. Hotspot swells generally have a higher heat flow than appropriate for the age of the surrounding oceanic crust.

Menard and Dorman (1977) noted that ridge-crest depths, a measure of the depth of the ocean floor at zero age, were greatest in the equatorial regions. They expanded the depths in the Pacific Ocean as a function of age and latitude:

$$d(t) = Kt^{1/2} + \sum_{n=0} C_n P_n (\sin \phi)$$

and obtained

$$K = 0.313 \text{ km (Ma)}^{-1/2}$$

$$C_0 = 2.64 \text{ km}$$

$$C_1 = 0.09 \text{ km}$$

$$C_2 = -0.26 \text{ km}$$

$$C_3 = -0.253 \text{ km}$$

Thus, the mean depth of the ocean at zero age is 2.64 km, and it deepens at an average rate of 0.313 km/(million years)$^{1/2}$. In addition to these age- and latitude-dependent effects, they found other depth anomalies having wavelengths of 4000 km and amplitudes of +0.8 to −0.5 km and a crest-to-trough relief of 0.7 to 1.0 km. The anomalies under the Pacific plate trend northwest-southeast, subparallel to the Hawaiian hotspot track, and those under the Nazca plate trend east-west. A large fraction of recent hotspots are associated with shallow anomalies. The Nazca plate is abnormally shallow.

Although cooling of the oceanic upper mantle is the first-order control of oceanic bathymetry, many large bathymetric features are not related to standard cooling and subsidence. Crough (1978, 1979) summarized these "depth anomalies" and, in particular, oceanic hotspot swells.

A few places are markedly deep, notably the seafloor between Australia and Antarctica and the Argentine Basin of the South Atlantic. Other deep regions occur in the central Atlantic and the eastern Pacific and others, most notably south of India, are not so obvious because of deep sedimentary fill. Most of the negative areas are less than 400 m below the expected depth, and they comprise a relatively small fraction of the seafloor area.

Shallow areas often exceed 1200 m in height and occupy almost the entire North Atlantic and most of the western Pacific that has been mapped. Almost every volcanic island, seamount or seamount chain surmounts a broad topographic swell. The swells generally occur directly beneath the volcanic centers and extend away from them in the downstream direction of plate-hotspot motion. Some extend a short distance in the upstream direction. Small regions of anomalously shallow depth occur in the northwestern Indian Ocean south of Pakistan, in the western North Atlantic near the Caribbean, in the Labrador Sea and in the southernmost South Pacific. They are not associated with volcanism but are slow regions of the upper mantle as determined from seismic tomography. Shallow regions probably associated with plate flexure border the Kurile Trench, the Aleutian Trench and the Chile Trench. Major volcanic lineaments without swells include the northern end of the Emperor Seamount chain, the Cobb Seamounts off the west coast of North America and the Easter Island trace on the East Pacific Rise. Bermuda and Vema, in the southeast Atlantic, are isolated swells with no associated volcanic trace. For most of the swells explanations based on sediment or crustal thickness and plate flexure can be ruled out. They seem instead to be due to variations in lithospheric composition or thickness, or abnormal upper mantle. Underplating the lithosphere by basalt or depleted peridotite, serpentinization of the lithosphere, delamination, or reheating and thinning the lithosphere are mechanisms that can decrease the density or thickness of the lithosphere and cause uplift of the seafloor. A higher temperature asthenosphere, greater amounts of partial melt, chemical inhomogeneity of the asthenosphere and upwelling of the asthenosphere are possible sublithospheric mechanisms. It would seem that the presence of a hotspot requires anomalous upper mantle, and it is likely that both lithospheric and asthenospheric properties contribute to the swell. Surface-wave tomography, although presently having low resolving capability, suggests that many swells do not have abnormally thin lithosphere but are associated with slower than average upper-mantle shear-wave velocities.

In general, smaller swells are located on younger seafloor, and the larger ones (Cape Verde, Hawaii, Great Meteor, Bermuda, Reunion) are on older crust. Within the scatter of the data, the midplate swells reach an approximately uniform depth below sea level, approximately 4250 m, or the swell height is proportional to the square root of crustal age (Crough, 1978, 1979; Menard and McNutt, 1982). Swells on ridge crests have heights that are inversely proportional to the local spreading rate; Iceland, the largest swell, is on a slow-opening ridge and Easter Island, one of the smallest, is on a fast-spreading ridge. Vogt (1975) showed that this is consistent with relatively shallow flow along the ridge axis. The widths of the subridge asthenospheric channel is proportional to the local spreading rate, so that flow along a fast-spreading ridge encounters less viscous resistance than flow along a slower spreading ridge. The pressure gradient necessary to drive the lateral flow manifests itself as a topographic gradient. Slow-spreading ridges require a large pressure gradient, and they therefore support high swells. The persistence of swells after they leave the ridge crest implies an anomalous lithosphere or additional shallow flow in the spreading direction. An alternative explanation for the relation between swell height and spreading rate is constant flux of material from the hotspot.

The great highlands of Africa including the Ethiopian and East African plateaus and the Hoggar and Tibesti massifs, the northern Rocky Mountains around Yellowstone and the Brazilian highlands are examples of possible hotspot-related continental swells. In shape and size these swells are similar to their oceanic counterparts. Western North America and northeast Africa are also associated with slow upper-mantle seismic anomalies.

The departure of the bathymetry-age relationship from a simple cooling law for Cretaceous lithosphere may reflect the extensive igneous activity that was occurring during this period. Many of the seamounts and plateaus in the western Pacific were formed in the Cretaceous.

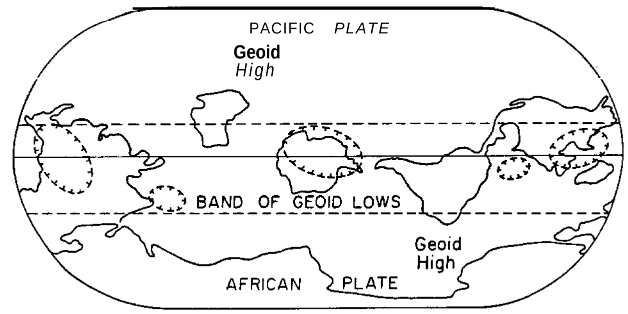

FIGURE 12-1
Geoid lows are concentrated in a narrow polar band passing through Antarctica, the Canadian Shield and Siberia. Most of the continents and smaller tectonic plates are in this band. Long-wavelength geoid highs and the larger plates (Africa, Pacific) are antipodal and are centered on the equator. The geoid highs control the location of the axis of rotation.

THE GEOID

Although the Earth is not flat or egg-shaped, as previously believed at various times, neither is it precisely a sphere or even an ellipsoid of revolution. Although mountains, ocean basins and variations in crustal thickness contribute to the observed irregular shape and gravity field of the Earth, they cannot explain the long-wavelength departures from a hydrostatic figure.

The centrifugal effect of the Earth's rotation causes an equatorial bulge, the principal departure of the Earth's surface from a spherical shape. If the Earth were covered by oceans then, apart from winds and internal currents, the surface would reflect the forces due to rotation and the gravitational attraction of external bodies, such as the Sun and the Moon, and effects arising from the interior. When tidal effects are removed, the shape of the surface is due to density variations in the interior. Mean sea level is an equipotential surface called the *geoid* or *figure* of the Earth. Crustal features, continents, mountain ranges and midoceanic ridges represent departures of the actual surface from the geoid, but mass compensation at depth, *isostasy,* minimizes the influence of surface features on the geoid. To first order, near-surface mass anomalies that are compensated at shallow depth have no effect on the geoid.

The shape of the geoid is now known fairly well, particularly in oceanic regions, because of the contributions from satellite geodesy. Apart from the geoid highs associated with subduction zones, there is little correlation of the long-wavelength geoid with such features as continents and midocean ridges. The geoid reflects temperature and density variations in the interior, but these are not simply related to the surface expressions of plate tectonics.

The largest departures of the geoid from a radially symmetric rotating spheroid are the equatorial and antipodal geoid highs centered on the central Pacific and Africa (Figure 12-1). The complementary pattern of geoid lows lie in a polar band that contains most of the large shield regions of the world. The largest geoid highs of intermediate scale are associated with subduction zones. The most notable geoid high is centered on the subduction zones of the southwest Pacific near New Guinea, again near the equator. The equatorial location of geoid highs is not accidental; mass anomalies in the mantle control the moments of inertia of the Earth and, therefore, the location of the spin axis and the equator. The largest intermediate-wavelength geoid lows are found south of India, near Antarctica (south of New Zealand) and south of Australia. The locations of the mass anomalies responsible for these lows are probably in the lower mantle. Many shield areas are in or near geoid lows, some of which are the result of deglaciation and incomplete rebound. The thick continental crust would, by itself, raise the center of gravity of continents relative to oceans and cause slight geoid highs. The thick lithosphere (—150 km) under continental shields is cold, but the seismic velocities and xenoliths from kimberlite pipes suggest that it is olivine-rich and garnet-poor; the temperature and petrology have compensating effects on density. The long-term stability of shields indicates that, on average, the crust plus its underlying lithosphere is buoyant. Midocean ridges show mild intermediate-wavelength geoid highs, but they occur on the edges of long-wavelength highs. Hotspots, too, are associated with geoid highs. The long-wavelength features of the geoid are probably due to density variations in the lower mantle and the resulting deformations of the core-

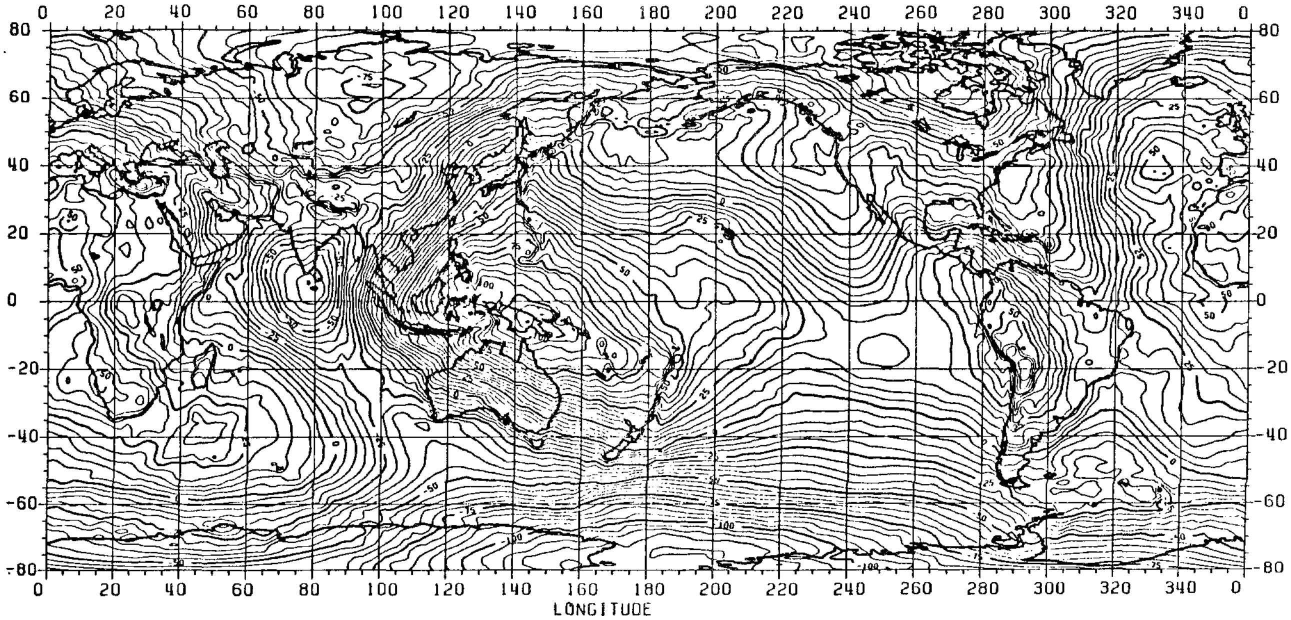

FIGURE 12-2
Geoid undulations (to degree 180) referred to a hydrostatic flattening of 1/299.638. Contour interval is 5m (after Rapp, 1981).

mantle boundary and other boundaries in the mantle (Richards and Hager, 1984).

Geoid anomalies are expressed as the difference in elevation between the measured geoid and some reference shape. The reference shape is usually either a spheroid with the observed flattening or the theoretical hydrostatic flattening associated with the Earth's rotation. The latter, used in Figure 12-2, is the appropriate geoid for geophysical purposes and is known as the nonhydrostatic geoid. By referencing the shape of the Earth's equipotential to the observed flattening by subtracting the C_{20} terms from the observed shape, one is throwing away some potentially important geophysical data. The observed C_{20} is not in fact the same as one obtains for an Earth in hydrostatic equilibrium, and this fact is telling us about the mass distribution in the mantle. By convention the origin is taken as the center of mass, so there is no C,, term. The geometric flattening of the Earth is 1/298.26. The hydrostatic flattening is 1/299.64.

The C_{ij} terms are spherical harmonic coefficients. C_{m0} are zonal harmonics, with boundaries between positive and negative features being small circles parallel to the equatorial plane. For a rotating homogeneous sphere the C_{20} term is the major term in a spherical harmonic expansion of the shape. This is sometimes referred to as the equatorial bulge. The C_{m1} are sectoral harmonics, and the general case C_{ml}, $0 < m < l$, are known as tesseral harmonics. If the origin of the coordinate system is not at the center of the Earth, then there will be a C,, term. North-south-running bands of high or low geoid anomalies will show up as strong sectoral components in a spherical harmonic expansion.

The maximum geoid anomalies are of the order of 100 m. This can be compared with the 21-km difference between the equatorial and polar radii. To a good approximation the net mass of all columns of the crust and mantle are equal when averaged over dimensions of a few hundred kilometers. This is one definition of isostasy. Smaller-scale anomalies can be supported by the strength of the crust and lithosphere. The geoid anomaly is nonzero in such cases and depends on the distribution of mass. It depends on the dipole moment of the density distribution

$$\int_0^h \Delta\rho(z) z \, dz$$

where z is the radial direction, h is the depth of the anomalous density distribution and $\Delta\rho(z)$ is the density anomaly. Clearly, if the depth extent of the anomalous mass is shallow, the geoid anomaly is small. This is why continents do not show up well in the geoid.

The geoid anomaly due to a long-wavelength isostatic density distribution is

$$N = \frac{-2\pi G}{g} \int_0^h z \Delta\rho(z) \, dz$$

The dipole moment $\Delta\rho(z)$ is nonzero, and the first moment of the density anomaly, that is the net mass anomaly, is zero for isostatic density distributions.

The elevation anomaly, *Ah*, associated with the density anomaly is

$$\int_0^h \Delta\rho(z)\,\mathrm{d}z = 0$$

or

$$\frac{\Delta h}{h} = -\frac{\Delta\rho}{\rho}$$

A negative *Ap*, caused for example by thermal expansion, will cause the elevation of the surface to increase (Ah = positive) and gives a positive geoid anomaly because the center of mass is closer to the Earth's surface. The mass deficiency of the anomalous material is more than canceled out by the excess elevation.

All major subduction zones are characterized either by geoid highs (Tonga and Java through Japan, Central and South America) or by local maxima (Kuriles through Aleutians). The long-wavelength part of the geoid is about that expected for the excess mass of the cold slab. The shorter-wavelength geoid anomalies, however, are less, indicating that the excess mass is not simply rigidly supported. Hager (1984) showed that there is an excellent correlation between the l = 4–9 geoid and the theoretical slab geoid and showed that this could be explained if the viscosity of the mantle increased with depth by about a factor of 30. The high viscosity of the mantle at the lower end of the slab partially supports the excess load. A chemical boundary near 650 km depth could also support the slab. The thick crust of island arcs and the high temperatures and partial melting of the "mantle wedge" above the slab may also contribute to the geoid highs at subduction zones. The deep trenches represent a mass deficiency, and this effect alone would give a geoid low. Other geoid highs, unrelated to slabs, appear to be associated with hotter than normal mantle, such as hotspot volcanism and low seismic velocities in the upper mantle.

The ocean floor in back-arc basins is often lower than equivalent age normal ocean, suggesting that the mass excess associated with the slab is pulling down the surface. This is not the only possibility. A thinner-than-average crust or a colder or denser shallow mantle could also depress the seafloor.

Cooling and thermal contraction of the oceanic lithosphere causes a depression of the seafloor with age and a decrease in the geoid height. Cooling of the lithosphere causes the geoid height to decrease uniformly with increasing age, symmetrically away from the ridge crest. The change is typically 5–10 m over distances of 1000 to 2000 km. The elevation and geoid offset across fracture zones is due to the age differences of the crust and lithosphere. Assuming a plate model with a fixed thickness and lower boundary temperature, a plate thickness of 66 km has been estimated for plates less than 30 Ma in age and 92 km for older plates (Cazenave, 1984; Cazenave and others, 1986). This implies a broad upwarping of geotherms in the younger parts of the Pacific Ocean.

The variation of geoid height with age is

$$\Delta h/\Delta t \approx 2\pi G/g\rho_m \alpha kT$$

for a cooling half-space model and at young ages in the plate model (where G is gravitational constant, g is surface gravity, ρ_m is upper-mantle density, *Ah* is change in geoid height, *At* is difference in age and T is boundary temperature). Observations are more consistent with a plate model in which lithospheric evolution slows down with time and plate buoyancy approaches an asymptotic value. However, it is also likely that lateral heterogeneity in the mantle beneath the plates, and extensive volcanism at certain periods of time, notably the Cretaceous, affect the geoid-age relation.

The long-wavelength topographic highs in the oceans generally correlate with positive geoid anomalies, giving 6–8 meters of geoid per kilometer of relief (Cazenave, 1984; Cazenave and others, 1986).

Tanimoto and Anderson (1985) showed that there was good correlation between intermediate-wavelength geoid anomalies and seismic velocities in the upper mantle; slow regions were geoid highs and vice versa. Subduction zones are slow in the shallow mantle, presumably due to the hot, partially molten mantle wedge under back-arc basins.

In the subduction regions the total geoid anomaly is the sum of the positive effect of the dense sinker and the negative effects caused by boundary deformations (Hager, 1984). For a layer of uniform viscosity, the net dynamic geoid anomaly caused by a dense sinker is negative; the effects from the deformed boundaries overwhelm the effect from the sinker itself. For an increase is viscosity with depth, the deformation of the upper boundary is less and the net geoid anomaly is positive.

For a given density contrast, the magnitude and sign of the resulting geoid anomaly in a dynamic Earth depends on the viscosity structure and the chemical stratification. Observation of the gravitational field of the Earth thus provides a null experiment, where the net result is a small number determined by the difference of large effects (Richards and Hager, 1984). The sign of the result depends on which of the effects is dominant. The anomaly also depends on the depth of the convecting system, with deep systems leading to larger geoid anomalies for a given density anomaly. Observations of the geoid in conjunction with observations of seismic velocity heterogeneities place constraints upon the variation of mantle viscosity and the depth of mantle convection.

Interpreting the Geoid

The geoid bears little relation to present tectonic features of the Earth other than trenches. The Mesozoic supercontinent of Pangaea, however, apparently occupied a central position in the Atlantic-African geoid high (Anderson, 1982). This and the equatorial Pacific geoid high contain most of the world's hotspots. The plateaus and rises in the western Pacific formed in the Pacific geoid high, and this may have been the early Mesozoic position of Pacifica, the fragments of which are now the Pacific rim portions of the continents. Geoid highs that are unrelated to present subduction zones may be the former sites of continental aggregations and mantle insulation and, therefore, hotter than normal mantle. The pent-up heat causes rifts and hotspots and results in extensive uplift, magmatism, fragmentation and dispersal of the continents, and the subsequent formation of plateaus, aseismic ridges and seamount chains. Convection in the upper mantle would then be due to lateral temperature gradients as well as heating from below and would be intrinsically episodic.

The Earth's largest positive geoid anomalies are associated with subduction zones and hotspots and have no simple relationship to other elevated regions such as continents and ridges. When the subduction-related geoid highs are removed from the observed field, the residual geoid shows broad highs over the central Pacific and the eastern Atlantic-African regions (see Figures 8-5 and 8-11). Like the total geoid, the residual geoid does not reflect the distribution of continents and oceans and shows little trace of the ocean ridge system. Residual geoid highs, however, correlate with regions of anomalously shallow ocean floor and sites of extensive Cretaceous volcanism.

The Atlantic-African geoid high extends from Iceland through the north Atlantic and Africa to the Kerguelen plateau and from the middle of the Atlantic to the Arabian Peninsula and western Europe (Figure 8-5). Most of the Atlantic, Indian Ocean, African and European hotspots are inside this anomaly. The hotspots Iceland, Trindade, Tristan, Kerguelen, Reunion, Afar, Eiffel and Jan Mayen form the 20-m boundary of the anomaly and appear to control its shape. The Azores, Canaries, New England seamounts, St. Helena, Crozet and the African hotspots are interior to the anomaly.

Although the geoid high cuts across present-day ridges and continents, there is a remarkable correspondence of the predrift assemblage of continents with both the geoid anomaly and hotspots. Reconstruction of the mid-Mesozoic configuration of the continents reveals, in addition, that virtually all of the large shield areas of the world are contained inside the geoid high (Figure 8-5). These include the shield areas of Canada, Greenland, Fennoscandia, India, Africa, Antarctica and Brazil. Most of the Phanerozoic platforms are also in this area. In contrast, today's shields and platforms are concentrated near geoid lows. They may have drifted into, and come to rest over, these geoid lows.

The area inside the geoid high is also characterized by higher-than-normal elevations, for example in Africa, the North Atlantic and the Indian Ocean southeast of Africa. This holds true also for the axial depth of oceanic ridges. Most of the continental areas were above sea level from the Carboniferous and Permian through the Triassic, at which time there was subsidence in eastern North and South America, central and southern Africa, Europe and Arabia. The widespread uplift, magmatism, breakup and initial dispersal of the Pangaean landmass apparently occurred while the continents were centrally located with respect to the present geoid anomaly. The subsequent motions of the plates, by and large, were and are directed away from the anomaly. This suggests that the residual geoid high, hotspots, the distribution of continents during the late Paleozoic and early Mesozoic, and their uplift and subsequent dispersal and subsidence are all related. The shields are regions of abnormally thick lithosphere. The thickest lithosphere is in eastern and central North America, northeastern South America, northwestern and central Africa and northern Siberia. These regions were all within the Atlantic-African geoid high at 200 Ma and, possibly, at 350 Ma as well, assuming the geoid high was present at those times and fixed relative to hotspots.

This area experienced exceptional magmatism during the Mesozoic. The great flood basalts of Siberia and South Africa were formed during the Triassic and Jurassic, possibly at the sites of the Jan Mayen and Crozet hotspots. The plateau basalt provinces of southeast Greenland and Brazil were formed during the Cretaceous, possibly at the sites of the Iceland and Tristan or Trindade hotspots. The Deccan Traps in India were also formed in the Cretaceous, presumably at the Reunion hotspot. The Walvis Ridge and the Rio Grande Rise are mainly on Mesozoic crust. A large part of the Pacific also experienced extensive on- and off-ridge volcanism in the Cretaceous. This extensive ridge and hotspot volcanism apparently reflects itself in a rapid rise in sea level during the Jurassic and Cretaceous. If sea level can be used as a guide to the volume of the ocean basins, then the end of the Cretaceous to the end of the Oligocene was a period of less intense oceanic volcanism and subsidence of the oceanic and continental crust. Sea-level variations may therefore indicate that the thermal and geoid anomalies formed in the Paleozoic have attenuated since the early Mesozoic.

Much of the Pacific rim appears to be accreted terrane that originated in the Pacific. The possibility of a continent centrally located in the Pacific—Pacifica—has been discussed for some time, but its location has been an enigma and its size uncertain. The central Pacific geoid anomaly may mark the early Mesozoic location of the anomalous circumPacific terrane and the site of extensive Cretaceous ridge-crest and midplate volcanism. It is underlain by abnormally low seismic velocities throughout most of the upper mantle (Nataf et al., 1984).

Paleomagnetic and other data indicate that various blocks of Asia such as Kolyma, Sikhote Alin, Sino-Korea, Yangtze, Southeast Asia and Japan have moved northward by up to 32° since the Permian. It is unlikely that they were in the vicinity of Australia or associated with Gondwana,

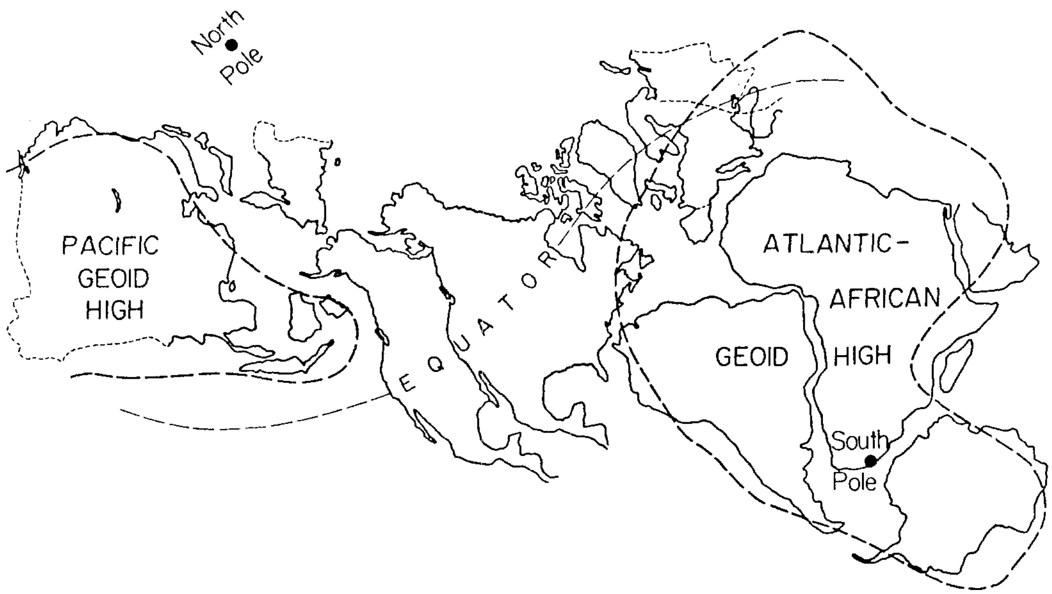

FIGURE **12-3**
A possible configuration of the continents in the Triassic relative to present geoid highs. Both continental drift and polar wander are required to bring the continents to their present positions relative to the geoid and the rotation axis.

and a central Pacific location is likely. Part or all of Alaska and northwestern North America were also far south of their present positions in the Permian. The same may be true of California, Mexico, Central America and other accreted terranes in the Pacific rim continents.

The central Pacific residual (corrected for slabs) geoid high (> 20 m) extends from Australia to the East Pacific Rise and from Hawaii to New Zealand. Its western part overlaps the subduction zones of the southwest Pacific and may be eliminated by a different slab model. The actual residual geoid anomaly may therefore be more equidimensional and is perhaps centered over the Polynesian plume province in the central Pacific. In any case it encompasses most of the Pacific hotspots and is approximately antipodal to the Atlantic-African anomaly. Figure 12-3 shows present-day geoid anomalies superposed on a hypothetical Triassic assemblage of the continents.

The western Pacific contains numerous plateaus, ridges, rises and seamounts that have been carried far to the northwest from their point of origin (see Figure 8-11). The Ontong-Java plateau, for example, presently on the equator, was formed about 4000 km further south in the mid-Cretaceous. The Hess rise, Line Islands ridge and Necker ridge were formed in the Cretaceous on the Pacific, Farallon and Phoenix ridge crests that were, at the time, in the eastern Pacific in the vicinity of the Polynesian seamount province. The ridge-crest volcanism was accompanied by extensive deep-water volcanism and, possibly, rapid seafloor spreading. The Caribbean and Bering Seas may have been formed at the same location and subsequently carried northward on the Farallon and Kula plates. It is significant that the Caribbean and the anomalous regions in the Pacific have similar geophysical and geochemical characteristics.

The Pacific geoid high also has anomalously shallow bathymetry, at least in the eastern part. This shallow bathymetry extends from the East Pacific Rise to the northwest Pacific in the direction of plate motion, and includes the central and southwest Pacific hotspots. The extensive volcanism in the central western Pacific between about 120 and 70 Ma occurred about 60° to 100° to the southeast, in the hotspot frame, which would place the event in the vicinity of the southeastern Pacific hotspots and the eastern part of the geoid and bathymetry anomaly. This suggests that the anomaly dates back to at least the Early Cretaceous. A similar thermal event in the Caribbean may mean that it was also in this region in the Cretaceous, particularly since the basalts in the Caribbean are similar to those in the western Pacific. The flattening of oceanic bathymetry after about 80 Ma may be related to extensive volcanism ending at about that time rather than due simply to the passage of time. It should be noted that there is an extensive upper-mantle low-velocity anomaly in the central Pacific.

Going back even further are the Triassic basalts in Wrangellia of northwestern North America and the Permian greenstones in Japan, both of which formed in equatorial latitudes and subsequently drifted to the north and northwest respectively. They may record the initial rift stages of continental crust. They also may have formed in the Pacific geoid high.

Episodicity in continental drift, polar wander, sea-level variations and magmatism may be due to the effect of thick continental lithosphere on convection in the mantle. The supercontinent of Pangaea insulated a large part of the mantle for more than 150 Ma. The upper mantle under Pangaea could cool, by subduction, around the edges, but the center of the continent was shielded from subduction and could only lose heat by conduction through a thick lithosphere. The excess heat caused uplift by thermal expansion and partial melting and, eventually, breakup and dispersal of Pangaea. A large geoid high is generated by this expansion, and, if this is the dominant feature of the geoid, the spin axis of the Earth would change so as to center the high on the equator, much as it is today. Surface plates experience membrane stresses during this shift, which may also contribute to breakup. Active subduction zones are also geoid highs. The continents stand high while they are within the geoid high and subside as they drift off.

At 100 Ma Europe, North America and Africa were relatively high-standing continents. This was after breakup commenced in the North Atlantic but before significant dispersal from the pre-breakup position. North America suffered widespread submergence during the Late Cretaceous while Africa remained high. Europe started to subside at about 100 Ma. This is consistent with North America and Europe drifting away from the center of the geoid high while Africa remained near its center, as it does today.

The association of the Atlantic-Africa geoid high with the former position of the continents and the plateau basalt provinces with currently active hotspots suggests that Mesozoic convection patterns in the mantle still exist. The Mid-Atlantic and Atlantic-Indian Ridges and the Atlantic, African and Indian Ocean hotspots and the East African rift are in the Atlantic-Africa geoid high. If geoid anomalies and hotspots are due to a long period of continental insulation and if these regions had higher-than-normal heat flow and magmatism for the past 200 Ma, then we might expect that these features will wane with time.

Horizontal temperature gradients can drive continental drift. The velocities decrease as the distance increases away from the heat source and as the thermal anomaly decays. The geoid high should also decay for the same reasons. Thick continental lithosphere then insulates a new part of the mantle, and the cycle repeats. Periods of rapid polar motion and continental rift follow periods of continental stability and mantle insulation.

The relatively slow motion of the continents or the pole during the Permian and the relative stability of sea level during this period suggests that the geoid anomalies had essentially developed at this time. Continental drift velocities and sea level changed rapidly during the Triassic and Jurassic. This I interpret as the start of extensive rift and hotspot magmatism and the start of the decay of the anomaly.

The rotation axis has shifted with respect to the hotspot reference frame by about 20° in the last 200 Ma (Gordon and Cape, 1981; Morgan, 1981; Gordon, 1982; Harrison and Lindh, 1982; Andrews, 1987). This shift could represent a growth of the equatorial regions of the geoid highs, a decay of the South Pacific or North Atlantic portions of the high due to extensive ridge-axis and hotspot volcanism, or a reconfiguration of the world's subduction zones.

Chase (1979) concluded that the lack of correlation of the large geoid anomalies to plate boundaries requires that they reflect a deep convective flow regime in the mantle that is unrelated to plate tectonics. However, the correspondence of the Atlantic-African anomaly with the Mesozoic continental assemblage and of the central Pacific anomaly with extensive Cretaceous volcanism in the Pacific is remarkable. Surface-wave tomography shows a good correlation of intermediate-wavelength geoid highs ($l \approx 6$) and hot regions of the upper mantle. However, the long-wavelength components of the geoid ($l = 2,3$) correlate best with the lower mantle. Most of the present continents, except Africa, and most of the present subduction zones (except Tonga-Fiji) are in long-wavelength ($l = 2-4$) geoid lows and therefore overlie colder than average lower mantle.

With the scenario sketched above, the relationship between surface tectonic features and the geoid changes with time. Supercontinents periodically form and insulate the underlying mantle and also control the locations of mantle cooling (subduction zones). When the continent breaks up, the individual fragments move away from the hot part of the mantle, and the geoid high, and came to rest over cold mantle, in geoid lows.

The locations of hot and cold regions in the upper mantle may be controlled by convection in the lower mantle. The lower mantle contribution to the geoid is probably long-lived. Empirically, subduction zones and continents are primarily in long-wavelength geoid lows and over long-wavelength fast seismic regions of the lower mantle. This can be understood if continents come to rest in geoid lows and if subduction zones, on average, are controlled by the advancing edges of continents. Midocean ridges tend to fall between the long-wavelength highs and lows. By long wavelength, we mean features having dimensions of thousands of kilometers.

Figure 12-4 shows the approximate locations of the continents just after breakup of Pangaea commenced. The hatched regions show oceanic lithosphere that has been overridden by the advancing continents. These are labeled "fast" because these are seismically fast regions of the transition region, where cold lithosphere may have cooled off the mantle. The arrows represent the motions of the continents over the past 110 Ma. Most of the hatched regions are also geoid lows.

Involvement of the Lower Mantle

Body-wave tomographic techniques can be applied to the problem of lateral heterogeneity of the lower mantle. Tomographic results (Dziewonski, 1984; Clayton and Comer,

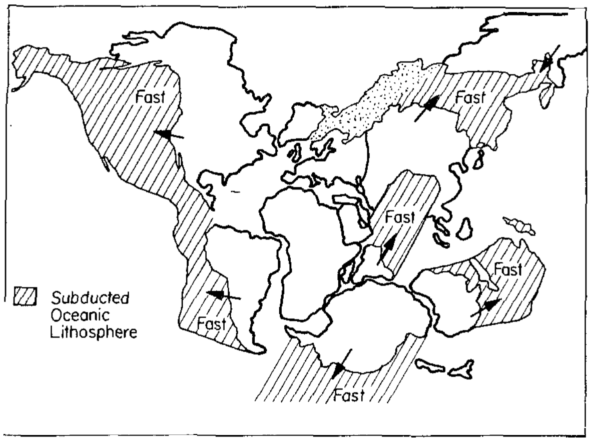

FIGURE **12-4**
Reconstruction of the continents at about 110 Ma ago. The hatched areas represent former oceanic lithosphere. These regions, in general, have high seismic velocities in the transition region, consistent with the presence of cold subducted lithosphere. They are also, in general, geoid lows. Dots represent possible convergence areas.

1988) give long-wavelength velocity anomalies that correlate well with the $l = 2,3$ geoid (Hager and others, 1985). Since one expects density to correlate with velocity, there is a strong indication that the long-wavelength geoid originates in the lower mantle. Phenomena such as tides, Chandler wobble, polar wander and the orientation of the Earth's spin axis depend on the $l = 2$ component of the geoid, so the lower mantle appears to be important in these areas. By contrast the upper mantle has only a weak correlation with the $l = 2,3$ geoid and is the reverse of the correlation with shorter wavelengths $l = 4-6$.

Regions of high velocity, and presumably high density, in the lower mantle correlate with geoid lows—the same sense of correlation found for shorter-wavelength features in the upper mantle. This is the relation expected for an isostatic mantle, with equal mass in equal columns. The deformation of the density interfaces counteracts the effect of high density. Low-density regions ride high. The $l = 2,3$ correlation of the geoid with surface-wave or upper-mantle velocities is weak and in the opposite sense.

Since hotspots occur preferentially in the long-wavelength geoid highs, as defined by the $l = 2,3$ geoid, they also occur over regions of the lower mantle that have lower than average velocities, averaged over tens of thousands of kilometers. The correlation breaks down at shorter wavelengths. This suggests that, in general terms, the locations of hot regions of the upper mantle are controlled by the locations of hot regions in the lower mantle, but upper-mantle and plate processes impose their own shorter length scales. Subduction of cold slabs and continental insulation may also influence the style of convection in the lower mantle, even if no transfer of material takes place. Subduction confined to the upper mantle can influence lower-mantle convection by deforming the boundary, cooling the lower mantle from above or by shear coupling. A thermal boundary layer and a large contrast in viscosity would favor thermal coupling over shear coupling.

HEAT FLOW'

The average heat flow through the surface of a planet is one of its fundamental properties. It is controlled by the radioactivity in the interior and the secular cooling of the planet, and it is a constraint on temperatures and styles of heat transport in the interior. Most of the heat transport in the deep interior is by convection, but heat convected toward the surface must ultimately be conducted through a cold conductive layer at the surface.

The terrestrial heat flow is defined as the quantity of heat escaping per unit time across each unit area of the Earth's solid surface. This quantity varies from place to place over the Earth's surface and with time throughout Earth history. The total heat being lost from the Earth at a given time is the integral of the heat flow over the Earth's surface. The heat arriving from the interior is a small fraction of the heat arriving from the Sun but the latter, of course, is not significant in controlling the internal temperature. The solar input, the surface albedo and atmospheric properties control the mean surface temperature of the Earth, which is about 0°C.

The heat flow through the crust and lithosphere is mainly by conduction and is governed by Fourier's Law,

$$\boldsymbol{q} = -\mathcal{K} \text{ grad } \boldsymbol{T}$$

where q is the heat flux vector, $\mathcal{K}$ is the thermal conductivity of the material through which the heat is being conducted and grad T is the local temperature gradient. Heat flow is usually measured in deep boreholes to get below the effects of seasonal and longer variations. One must measure both the thermal gradient and the rock conductivity. Temperatures in shallow holes are perturbed to varying degrees by ground-water circulation, climatic variations, irregularities of local topography, vegetation patterns, slope orientation and Sun angle. Even deep holes can be affected by local geological structures, erosion, tectonic uplift and conductivity variations that distort the isothermal surfaces.

The thermal diffusivity, κ, of a material is given by $\kappa = \mathcal{K}/\rho C_P$ where $\mathcal{K}$ is the thermal conductivity, ρ the density and C_P the specific heat. Rocks have relatively low conductivities so that they respond slowly to any change in ambient temperature. The conductive time constant is approximately

$$\tau = l^2/\kappa$$

where l is the characteristic length over which heat must be transferred. This relation shows that over the age of the Earth, conductive cooling can only extend to a depth of about 400 km. The long time constant means that as crustal rocks are buried or eroded, they tend to carry their thermal structure with them as they move. This is a form of convective transport of heat.

The present heat loss of the Earth is 10^{13} cal/cm²/s (4.2×10^{13} W). The heat loss through the surface is greatest near the midocean ridges and in tectonic regions and is least through continental Precambrian shields. Table 12-1 provides a summary of various components of the terrestrial heat-flow budget.

The mean heat flow through the ocean basins, based on a theoretical fit to oceanic heat-flow observations, decreases from about 8 μcal/cm²/s at the ridge crest, with a high variability, to a relatively constant value of about 1.2 μcal/cm²/s (50 mW/m²) for old ocean. The unit μcal/cm²/s or 10^{-6} cal/cm²/s is called a Heat Flow Unit or HFU.

On continents there is a rough correlation between heat flow and age, or age of last tectonic or magmatic event (Sclater and others, 1980, 1981). In young regions, heat flow has a high mean and large scatter. The heat flow in active or recently active areas ranges from about 0.5 to over 3 HFU. The mean heat flow decreases from 1.8 HFU (77 mW/m²) in the youngest provinces to a constant value of 1.1 HFU after about 800 Ma. A few continental heat-flow values are below 0.6 HFU, and strangely, most of these are in young areas. The scatter in older areas seems to be due to variations in crustal radioactivity.

It has been discovered that there is a simple linear relationship between the heat flow and surface radioactivity in specific areas, termed "heat-flow provinces":

$$q_o = q_r + DA,$$

where q_o and A, are the heat flow and the radioactivity at the surface, q_r is the "reduced" heat flow and D is a depth scale for the vertical distribution of radioactivity (Birch and others, 1968). This depth scale is probably different for potassium, thorium and uranium. The reduced heat flow, due to heat flow primarily from the lower crust and mantle (Sclater and others, 1980, 1981), ranges from about 0.5 to 0.67 HFU. (See Table 12-2.) This can be compared with the estimated equilibrium flux through the base of the oceanic crust, 0.9 HFU.

The steady-state temperature at moderate depth z can be estimated from

$$T(z) = T_o + q \int_0^z \frac{dz}{K}$$

The temperature gradient decreases with depth. If heat production due to radioactivity is significant over the temperature interval of interest, then this must be allowed for. Table 12-3 gives the heat productivity of various rock types and of the continental crust.

Velocity of Love waves shows a high degree of correlation with global heat-flow maps (Nakanishi and Anderson, 1984; Tanimoto and Anderson, 1984, Nataf and others, 1986). Love waves are sensitive to shear velocity in the shallow mantle. This correlation makes it possible to identify regions of possibly anomalous heat flow, which is useful since global heat-flow coverage is far from complete. Low Love-wave velocities and, therefore, regions of high upper-mantle temperature occur along the East Pacific Rise, western North America and Central America, Red Sea–East African Rift–Gulf of Aden, eastern Asia, southeast Asia, the Indian Ocean triple junction, the Tasman Sea, the south-central Pacific, the far North Atlantic, southern Europe and the Tristan da Cunha–Rio Grande Rise. Fast regions, and regions of presumed low heat flow are most con-

TABLE 12-1
Heat Loss of the Earth

Component	Value
Heat loss through continents	2.8×10^{12} calls
Heat loss through oceans	7.3×10^{12} cal/s
Total heat loss through surface	10^{13} cal/s
Heat loss by hydrothermal circulation	2.4×10^{12} calls
Heat loss by plate creation	6.3×10^{12} calls
Mean heat flow	2.0×10^{-6} callcm²/s
Continents	1.4×10^{-6} cal/cm²/s
Oceans	2.4×10^{-6} callcm²/s
Radioactive decay in crust	−17 percent of heat loss
Convective heat transport by surface plates	–65 percent of heat loss

Sclater and others (1981).

TABLE 12-2
Measured and Reduced Heat Flow

Province	Heat Flow (mW/m²) Measured	Reduced
Shields		
Superior	34 ± 1.0	22 ± 2.0
Brazil	56 ± 1.5	29 ± 3.5
Baltic	33 ± 1.4	22 ± 2.0
Ukraine	37 ± 0.4	24 ± 0.7
Western Australia	39 ± 1.2	30 r 1.8
Mean	27 ± 4.0	
Other		
Basin and Range	77 ± 1.8	63 ± 5.3
Southeast Appalachians	49 ± 0.0	28 ± 0.1
Sierra Nevada	36 ± 0.7	15 ± 1.3
Eastern United States	64 ± 0.7	33 ± 1.3
England	69 ± 3.1	24 ± 5.4
India	71 ± 1.2	38 ± 4.4
Upper Crust	12 to 56	

Pujol and others (1985).

TABLE 12-3
Heat Production of Common Rocks

Rock type	Uranium (ppm)	Thorium (ppm)	Potassium (pct.)	Heat Production (10^{-6} W/m^3)
Granite	3.9	16.0	3.6	2.5
Granodiorite	2.3	9.0	2.6	1.5
Andesite	1.7	7.0	1.1	1.1
Basalt	0.5	1.6	0.4	0.3
Peridotite	0.02	0.06	0.006	0.01
Dunite	0.003	0.01	0.0009	0.002
Continental crust	1.25	4.8	1.25	0.8

Pollack (1980)

tinental shield areas, the North Pacific, south and northwest Africa, the eastern Indian Ocean and much of the Atlantic.

Figure 12-5 shows the low-order geoid and heat-flow maps.

ROTATION OF THE EARTH AND POLAR WANDER

The axis of rotation of the Earth is not fixed, either in space or relative to the body of the Earth, nor is the speed of rotation constant. The movement of the axis of rotation with respect to the crust is called polar wander or polar motion. Variation in the speed of rotation about the instantaneous axis causes changes in the length of day. ***Precession*** is the motion of the rotation axis with respect to inertial space— an oscillation of the rotation axis about the pole of the ecliptic with a period of 26,000 years. In addition, there are shorter period nutations. Various geophysical phenomena cause the pole to wander and the Earth to slow down.

Astronomical data, including records of ancient eclipses, show that the Earth has been slowing down at a rate of about 5×10^{22} rad/s, meaning that the length of the day is increasing at a rate of 1.5 ms per century. To balance the change in angular momentum, the Moon is receding from the Earth at a rate of about 5 cm per year. Paleontological evidence suggests that these accelerations have persisted for at least the past 500 Ma. The rotation rate of the Earth is also controlled by the internal rearrangement of mass.

The tidal deceleration of the Earth's rotation is due to the attraction of the Moon on the Earth's tidal bulge, which itself is primarily due to the Moon's gravitational pull. Since the Earth is not perfectly elastic, the tidal bulge is not directly beneath the Moon but is carried beyond the Earth–Moon line by the Earth's rotation. The attraction of the Moon on this displaced bulge then serves to brake the Earth's rotation and speed up the Moon in its orbit.

The eccentricity of the lunar orbit increases with time. The inclination of the lunar orbit also evolves and was greater in the past. The history of the lunar orbit is such that there was a minimum Earth–Moon distance at some time in the past, the inclination of the lunar orbit was substantially greater in the past, and the eccentricity of the lunar orbit increases as the Moon spirals away from the Earth. The tidal dissipation in the Earth occurs mainly in the oceans, especially shallow seas and along north-south coastlines. This effect is not constant with time, and it is therefore difficult to extrapolate backwards. In some calculations the Moon approaches the Roche limit, 2.8 Earth radii, early in Earth history. Within this limit the self-gravitation of the Moon is exceeded by the gravitational attraction of the Earth, and the Moon cannot cohere. The high tides raised on the Earth would cause extensive melting of the crust. The geological effect of close approach on both the Earth and the Moon would rival that of the early bombardments of these bodies.

The spin axis of an irregular, rotating body such as the Earth is controlled by the integrated mass distribution in the interior. For an evolving, dynamic planet, or for a planet being bombarded from outside, the spin axis is not fixed to any particular part of the planet but changes with time. It is no accident that the largest mass anomalies, those associated with subduction zones, are symmetric about the equator. The largest positive mass anomalies are centered over New Guinea–Borneo and the northern Andes, and these seem to control the present axis of rotation. However, if subduction in these regions were to cease and old oceanic lithosphere continued to pile up under the Aleutians and Tonga, these regions would eventually define the equator. The processes that change the mass distribution are slow, limited by the thermal time constant of the mantle and by the reconfiguration of continents and subduction zones. The reorientation of the spin axis, however, can be relatively fast, limited by the viscosity of the mantle, if the magnitudes of the principal moments of inertia interchange their order (Figure 12-6).

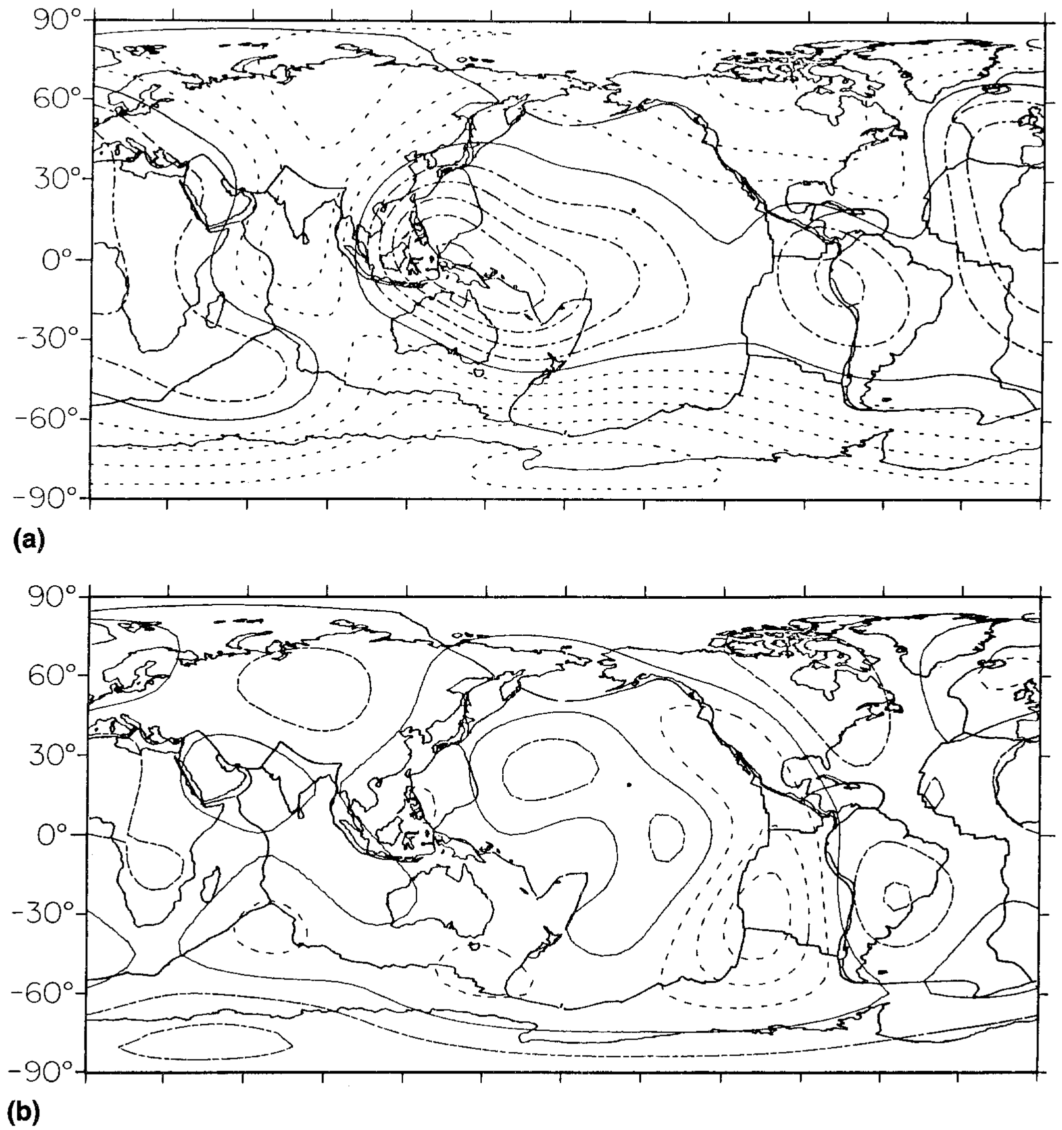

FIGURE 12-5
(a) The low-order nonhydrostatic geoid ($l = 2$–6). Contour interval 20 m. Long-wavelength geoid highs are centered on the equator near New Guinea and central Africa. **(b)** Heat flow map of Chapman and Pollack (see Pollack, 1980), $l = 2$–6. Contour interval 10 mW/m^2. Dashed lines are higher than average heat flow (after Nakanishi and Anderson, 1984).

The axis of rotation of the Earth is currently moving at a rate of 9 ± 2 cm/yr toward eastern Canada (75°–85° W). This is primarily due to isostatic rebound associated with ice and water redistribution following the last ice age and therefore contains information about the viscosity of the mantle. True polar wander must be distinguished from the apparent polar wander that is the result of plate motions; both phenomena are occurring, and it is sometimes difficult to separate the two effects. The concepts of mobile plates and a convecting mantle are now well engrained, but the concept of true polar wander, the physics of which is better understood (e.g. Goldreich and Toomre, 1969), has received less attention.

Imagine a rotating homogeneous sphere. Now place a dense mass anywhere on the surface or in the interior. This mass will eventually end up in the equatorial plane since this minimizes the kinetic energy of rotation. Although the mass is fixed to the sphere, it ends up on the equator by tilting the whole sphere. Now paint a dot on the north pole and consider a mass that can move around the interior or the surface. If it moves slowly, with respect to the time scale of viscous deformation of the sphere, it will never leave the equatorial plane. The sphere will tilt, with the point defining the north pole moving toward the mass while the rotation axis remains fixed in space. The mass will always be

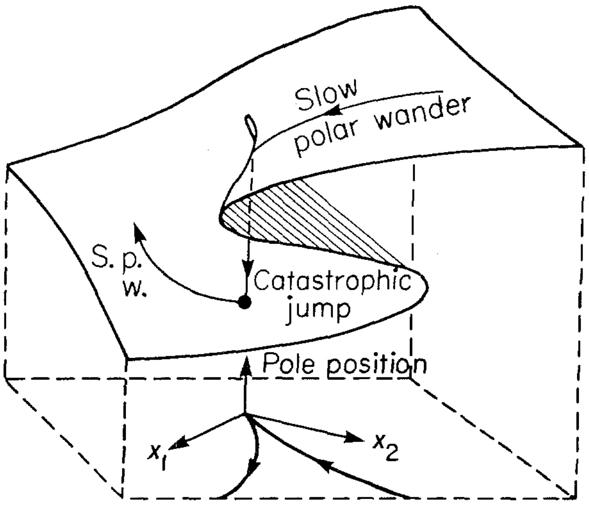

FIGURE **12-6**
The principal moments of inertia shown on a cusp catastrophe diagram. As the moments of inertia vary, due to convective processes in the interior, the pole will slowly wander unless the ratios of the moments *x*, and x_2 pass through unity, at which point a catastrophe will occur, leading to a rapid change in the rotation axis.

at the farthest point from the spin axis. If there are many such masses, they will, on average, define the equator. The masses do not have to physically move. They can change their density by, for example, cooling or melting. The rotation axis of the Earth is determined by the integrated effect. The dominant effect on the present Earth appears to be the large amount of subduction of old oceanic lithosphere that has occurred in the western Pacific, and the broad upwellings in the lower mantle under the central Pacific and Africa.

Because the Earth is a dynamic body with a constantly shifting surface, it is impossible to define a permanent internal reference frame. There are three in common use: the rotation axis, the geomagnetic reference frame and the hotspot reference frame. The rotational frame is a function of the size of the mass anomalies and their distance from the axis of rotation. Upper-mantle effects are important because lateral heterogeneity is greater than lower-mantle or core heterogeneity and because they are farthest from the center of the Earth. The lower mantle is also important because of its large volume, but a given mass anomaly has a greater effect in the upper mantle. The location of the magnetic pole is controlled by convection in the core, which in turn is influenced by the rotation of the Earth and the temperatures at the base of the mantle. On average the rotational pole and the magnetic pole are close together, but the instantaneous poles can be quite far apart. The hotspot frame is based on the observation that hotspots seem to move relative to one another less rapidly than plates move relative to one another. The 40 or so long-lived hotspots have been used with some success as reference markers to track the motions of surface plates. This has led to the view that hotspots are anchored deep in the mantle and may reflect a different kind of convection than that which is responsible for large-scale convection in the mantle. Actually, the return flow associated with the motion of thin plates is spread out over a larger volume than the plates themselves, and the average velocities required to maintain mass balance are correspondingly lower. For example, if plates are 100 km thick and the return flow is spread out over the rest of the upper mantle, the average flow velocity beneath the plates is 4.5 times less than the plate velocity. Such a small relative motion beneath hotspots cannot be ruled out. A variety of evidence suggests that the lower mantle has a higher viscosity than the upper mantle. Convection in the lower mantle is therefore sluggish, and any embedded hot areas will tend to heat the upper mantle, giving rise to hotspots even if no material exits the lower mantle. Thus, there are at least two possible explanations for the small motion between hotspots that do not require a source of magma in the lower mantle, rapid radial ascent of plumes from the core-mantle boundary, or whole-mantle convection. On the other hand, material from a deep plume source must traverse a convecting mantle, and some relative motion is expected between plumes.

A relative displacement of 22" ± 10° has been deduced between the hotspot frame and the geomagnetic frame over the past 180 Ma (Andrews, 1987). The motion has not been smooth. Episodes of rapid motion (– 1 cm/yr) include the present, 65–50 Ma, 115–85 Ma and 180–160 Ma.

The rotation axis has apparently wandered about 8" in the past 60 Ma and 20" in the past 200 Ma (Gordon and Cape, 1981; Morgan, 1981; Harrison and Lindh, 1982; Gordon, 1982). During these periods there have been major changes in the configurations of continents and subduction zones. These apparently had minor effects on the principal moments of inertia, suggesting a relatively stable and slowly evolving geoid. Even a slowly changing geoid, however, can cause a rapid shift in the whole mantle relative to the spin axis if the moment of inertia along this axis becomes less than some other axis of inertia. This may have happened sometime between 450 and 200 Ma ago, when Gondwana experienced a large latitude shift.

On the familiar Mercator projection the continents appear to be more or less randomly disposed over the surface of the globe. On other projections, however, there is a high degree of symmetry (for example, Figure 12-1). The centers of the largest (African and Pacific) plates are antipodal, and the other smaller plates, containing most of the large shield areas, occupy a relatively narrow polar band. The geoid lows in Siberia, the Indian Ocean, Antarctica, the Caribbean, and the Canadian shield also fall in this band. Long-wavelength geoid highs are centered on the large antipodal plates with hot spots, and it appears that shield areas have emigrated from these thermal and geoid highs. The polar band of discontinuous geoid lows correlates with areas of Cretaceous subduction and subsequent tertiary subsidence and areas in the upper mantle where seismic velocities are fast. Isostatically compensated cold mantle generates geoid lows. The geoid lows found in the wake of the Americas, India, and Australia exhibit fast surface-wave velocities, indicating that the anomalies are in the upper mantle, at least in part.

Density inhomogeneities in the mantle grow and subside, depending on the locations of continents and subduction zones. The resulting geoid highs reorient the mantle relative to the spin axis. Whenever there was a major continental assemblage in the polar region surrounded by subduction, as was the case during the Devonian through the Carboniferous, the stage was set for a major episode of true polar wandering.

The outer layers of the mantle, including the brittle lithosphere, do not fit properly on a reoriented Earth. Membrane stresses generated as the rotational bugle shifts may be responsible for the breakup and dispersal of Pangaea as it moves toward the equator. In this scenario, true polar wandering and continental drift are intimately related. A long period of continental stability allows thermal and geoid anomalies to develop. A shift of the axis of rotation causes plates to split, and the horizontal temperature gradient causes continents to drift away from the thermal anomalies that they caused. The continents drift toward cold parts of the mantle and, in fact, make the mantle cold as they override oceanic lithosphere.

Polar wandering can occur on two distinct time scales. In a slowly evolving mantle the rotation axis continuously adjusts to changes in the moments of inertia. This will continue to be the case as long as the major axis of inertia remains close to the rotation axis. If one of the other axes becomes larger, the rotation vector swings quickly to the new major axis (Figure 12-5). The generation and decay of thermal perturbations in the mantle are relatively gradual, and continuous small-scale polar wandering can be expected. The interchange of moments of inertia, however, occurs more quickly, and a large-scale 90° shift can occur on a time scale limited only by the relaxation time of the rotational bulge. The rate of polar wandering at present is much greater than the average rate of relative plate motion, and it would have been faster still during an interchange event. The relative stability of the rotation axis for the past 200 Ma suggests that the geoid highs related to hotspots have existed for at least this long. On the other hand, the rapid polar wandering that started 500 Ma ago may indicate that the Atlantic-African geoid high was forming under Gondwana at the time and had become the principal axis of inertia. With this mechanism a polar continental assemblage can be physically rotated to the equator as the Earth tumbles.

The southern continents all underwent a large northward displacement beginning sometime in the Permian or Carboniferous (280 Ma ago) and continuing to the Triassic (190 Ma). During this time the southern periphery of Gondwana was a convergence zone, and a spreading center is inferred along the northern boundary. One would expect that this configuration would be consistent with a stationary or a southern migration of Gondwana, unless a geoid high centered on or near Africa was rotating the whole assemblage toward the equator. The areas of very low surface-wave velocities in northeast Africa and the western Indian Ocean may be the former site of Gondwana.

Thus, expanding the paradigm of continental drift and plate tectonics to include continental insulation and true polar wandering may explain the paradoxes of synchronous global tectonic and magmatic activity, rapid breakup and dispersal of continents following long periods of continental stability, periods of static pole positions separated by periods of rapid polar wandering, sudden changes in the paths of the wandering poles, the migration of rifting and subduction, initiation of melting, the symmetry of ridges and fracture zones with respect to the rotation axis, and correlation of tectonic activity and polar wandering with magnetic reversals. Tumbling of the mantle presumably affects convection in the core and orientation of the inner core and offers a link between tectonic and magnetic field variations.

The largest known positive gravity anomaly on any planet is associated with the Tharsis volcanic province on Mars. Both geologic and gravity data suggest that the positive mass anomaly associated with the Tharsis volcanoes reoriented the planet with respect to the spin axis, placing the Tharsis region on the equator (Melosh, 1975a,b, 1980). There is also evidence that magmatism associated with large impacts reoriented the Moon. The largest mass anomaly on Earth is centered over New Guinea, and it is also almost precisely on the equator. The long-wavelength part of the geoid correlates well with subduction zones and hotspot provinces, and these in turn appear to control the orientation of the mantle relative to the spin axis. Thus, we have the possibility of a feedback relation between geologic processes and the rotational dynamics of a planet. Volcanism and convergence cause mass excesses to be placed near the surface. These reorient the planet, causing large stresses that initiate rifting and faulting, which in turn affect volcanism and subduction. Curiously, Earth scientists have been more reluctant to accept the inevitability of true polar wandering than to accept continental drift, even though the physics of the former is better understood.

The lower mantle plays an important role in the dynamics of the Earth. The seismically fast regions of the lower mantle are presumably cold, downwelling regions. Subduction zones (except Tonga-Fiji) and continents (except Africa) are located over these features. The colder parts of the lowermost mantle will have the highest thermal gradients in D" and heat will be removed from the core most efficiently in these places. Downwellings in the core will therefore occur under cold parts of the lower mantle. Since, in general, hot regions of the lower mantle will be associated with geoid highs and mass excesses, when boundary deformation is taken into account, they will tend to lie along the equator, putting the colder regions of the lower mantle into high latitudes. Thus, convection in the core will have a preferential alignment relative to the spin axis. This accentuates the symmetry caused directly by rotation of the Earth on convection. Details of convection in the core may be controlled by the thermal structure in the lower mantle.

CONVECTION

The Earth's mantle transmits short-period seismic shear waves and responds as an elastic solid to tidal stresses. However, to long sustained loads it behaves as a viscous fluid. In a viscous fluid, stresses relax with time, and heat is transported by convection as well as by conduction. Both plate tectonics and the thermal history of the Earth are controlled by the fluid-like behavior of the mantle over long periods of time. A fluid will convect if heated from below or within, or cooled from above, or driven by loading at the boundaries or in the interior. Fluid-like behavior is not only a rheology-dependent process but is also a size-dependent characteristic: Large objects are intrinsically weak.

The slow uplift of the surface of the Earth in response to the removal of a load such as an ice cap or drainage of a large lake is not only proof of the fluid-like behavior of the mantle but also provides the data for estimating its viscosity. In contrast to everyday experience, however, mantle convection has some unusual characteristics. The container has spherical geometry. The "fluid" has stress-, pressure- and temperature-dependent properties. It is heated from within and from below. The boundary conditions and heat sources change with time. Melting and phase changes contribute to the buoyancy and provide additional heat sources and sinks. Convection may be driven partly by plate motions and partly by sinking and rising chemical differences. The boundaries are deformable rather than rigid. None of these characteristics are fully treated in numerical calculations, and we are therefore woefully ignorant of the style of convection to be expected in the mantle.

The theory of convection in the mantle has, by and large, been decoupled from the theory of solids and petrological constraints. The non-Newtonian rheology, the pressure and temperature sensitivity of viscosity, thermal expansion, and thermal conductivity, and the effects of phase changes and compressibility make it dangerous to rely too much on the intuition provided by oversimplified calculations. There are, however, some general characteristics of convection that transcend these details.

The energy for convection is provided by the decay of the radioactive isotopes of uranium, thorium and potassium and the cooling and crystallization of the Earth. This heat is removed from the interior by the upwelling of hot, buoyant material to the top of the system where it is lost by conduction and radiation. The buoyancy is provided by thermal expansion and phase changes including partial melting. In a chemically layered Earth the heat is transferred by convection to internal thermal boundary layers across which the heat is transferred by the slow process of conduction. Thermal boundary layers at the top of a layer cool the interior by becoming unstable and sinking, displacing new material toward the surface. Thermal boundary layers at the base of convecting systems warm up and *can* also become unstable, generating hot upwellings or plumes. Adiabatically ascending hot plumes are likely to cross the solidus in the upper mantle, at least, thereby buffering the temperature rise and magnifying the buoyancy. Because of the divergence, with pressure, of melting curves and the adiabat, deep plumes are probably subsolidus, at least initially.

In the simplest model of convection in a homogeneous fluid heated from below, hot material rises in relatively thin sheets and spreads out at the surface where it cools by conduction, forming a cold surface thermal boundary layer that thickens with time. Eventually the material achieves enough negative buoyancy to sink back to the base of the system where it travels along the bottom, heating up with time until it achieves enough buoyancy to rise. This gives the classical cellular convection with most of the motion and temperature change occurring in thin boundary layers, which surround the nearly isothermal or adiabatic cores. It has not escaped the attention of geophysicists that *midocean* ridges and subducting slabs resemble the edges of a convection cell and that the oceanic lithosphere thickens as a surface boundary layer cooling by conduction.

The planform of the convection depends on the Rayleigh number. At high Rayleigh number three-dimensional patterns result that resemble hexagons or spokes in plan view. Upwellings and downwellings can be different in shape.

In an internally heated fluid the heat cannot be completely removed by narrow upwellings. The whole fluid is heating up and becoming buoyant, so very broad upwellings result. On the other hand, if the fluid has a stress-dependent rheology, or a component of buoyancy due to phase changes such as partial melting, then the boundary layers can become thinner. A temperature-dependent rheology can force the length scales of the surface boundary to be larger than the bottom boundary layer. The lower boundary layer, having a higher temperature and, possibly, undergoing phase changes to lighter phases, can go unstable and provide upwellings with a smaller spacing than the downwellings. The effect of pressure and phase changes on the rheology may reinforce or reverse this tendency. Finally, the presence of accumulations of light material at the surface—continents—can affect the underlying motion. Subduction, for example, depends on more than the age of the oceanic lithosphere. Convection in the mantle is therefore unlikely to be a steady-state phenomenon. Collection of dense material at the base of the mantle, or light material at the top of the core, may help explain the unusual characteristics of D".

Although convection in the mantle can be described in general terms as thermal convection, it differs considerably from convection in a homogeneous Newtonian fluid with constant viscosity and thermal expansion. The temperature dependence of viscosity gives a "strong" cold surface layer. This layer must break or fold in order to return to the interior. When it does, it can drag the attached "plate" with it, a sort of surface tension that is generally not important in normal convection. In addition, light crust and depleted lithosphere serve to decrease the average density of the cold thermal boundary layer, helping to keep it at the surface. Buoyant continents and their attached, probably also buoyant, lithospheric roots move about and affect the underlying convection. The stress dependence of the strain rate in solids gives a stress-dependent viscosity. This concentrates the flow in highly stressed regions; regions of low stress flow slowly. Mantle minerals are anisotropic, tending to recrystallize with a preferred orientation dictated by the local stresses. This in turn gives an anisotropic viscosity, probably with the easy flow direction lined up with the actual flow direction. The viscosity controlling convection may therefore be different from the viscosity controlling postglacial rebound.

One of the major controversies regarding mantle convection is the radial scale length. There are proponents of whole-mantle convection and stratified convection. A relatively small chemical contrast can prevent whole-mantle convection. For example, if the lower mantle has a higher SiO_2/MgO content than subducting material, the density at a given depth will be different. This density contrast must be overcome by thermal expansion in order for material to puncture the boundary. For a temperature contrast of 10^3 K and a coefficient of thermal expansion of 3×10^{-6}/K, an intrinsic density contrast at a discontinuity of 3 percent will prevent circulation of material through the boundary. Differences in FeO, CaO and Al_2O_3 content also affect intrinsic density. The problem is somewhat more complicated than this since the boundary will be warped by the convection and the cold upper material can be pushed down to greater depths and pressures. If substantial enough phase changes occur as a result of this increased density, without a large increase in temperature associated with latent heats, then subducted material may be able to overcome the intrinsic density contrast. Convecting material in one layer may also entrain material from an adjacent layer (Figure 9-6).

The upper mantle itself may be chemically stratified. In this case, temperatures are more likely to be close to the solidus, and partial melting in the deeper layer provides additional buoyancy. Buried layers that become hot can become unstable.

Although the viscosity of the mantle increases with depth, because of pressure, and although the lower mantle may also have a higher viscosity, because the mineralogy is different and stresses may be lower, this is not sufficient to prevent whole-mantle convection. The actual viscosity structure of the mantle is uncertain because of the low resolving power of the data and the ambiguities regarding the contribution of non-Newtonian and transient rheology. If the viscosity of the upper mantle is less than that of the lower mantle, circulation will be faster and there is more opportunity to recirculate crustal and upper-mantle material through the shallow melting zone. Differentiation processes would therefore change the composition of the upper mantle, even if the mantle were chemically homogeneous initially. The separation of light and dense material, and low melting point and high melting point material, however, probably occurred during accretion and the early high-temperature history of the Earth. The differentiation will be irreversible if the recycled products of this differentiation (basalt, eclogite, depleted peridotite, continental crust) are unable to achieve the densities of the lower mantle or the parts of the upper mantle through which they must pass. A small radial difference in the intrinsic density of the mantle can prevent whole-mantle convection. The mean atomic weight, a measure of FeO content, is not an adequate measure of composition, and this parameter is poorly resolved from seismology. It is not clear if there is a difference in mean atomic weight between the upper and lower mantles, but even if this parameter were constant, there could be significant chemical, and intrinsic density, differences. MgO, SiO_2 and Al_2O_3 and minerals composed of these oxides have nearly the same mean atomic weight, but small variations in the proportions of these components change the stable mineral assemblages and phase transition boundaries and, therefore, the intrinsic densities. For example, if the lower mantle is orthopyroxene-rich, then perovskite will be the major mineral. An upper mantle richer in MgO and Al_2O_3 will have more olivine and garnet, and this assemblage will be unable to subduct into the lower mantle even when it has transformed to high-pressure assemblages and even if it has a higher FeO content than the lower mantle. The plausibility of whole-mantle convection then depends on the plausibility of avoiding separation of mantle components during accretion and the early high-temperature history of the Earth.

Considering these uncertainties in even what calculations to perform and in the parameters, much less the uncertainties in the pressure, temperature and stress dependence of the parameters, observational constraints will continue to be the most important diagnostics of the style of mantle convection. These constraints include surface topography, heat flow, gravity field, lateral variations in seismic velocities and depths of mantle discontinuities, and the distribution of earthquakes. Lateral variations in density and seismic velocity are due to temperature and temperature-induced phase changes. In general, phase changes cause larger variations than temperature alone.

There are several variants of the stratified mantle convection idea. In the hot accretion–magma ocean scenario, the mantle is not homogeneous in either the trace elements or the major elements. The incompatible elements are concentrated in the crust and shallow mantle, and this includes the radioactive elements. The lower mantle is composed of the denser, refractory silicates and is depeleted in radioactive~The upper mantle itself may be chemically zoned. In this model there are several thermal boundary layers, and most of the mantle cools very slowly.

Some geochemical models assume that the crust formed only from the upper mantle and that the lower mantle is still primitive, somehow surviving the accretional melting and differentiation and retaining primitive abundances of the heat-producing elements.

Other models combine whole-mantle circulation with embedded inhomogeneities to provide isotropic heterogeneity, or assume that the mantle is homogeneous except for the D" region, which is chemically distinct.

The whole-mantle convection models have received the most attention from the convection community. The various layered models, particularly with deformable boundaries and pressure- and temperature-dependent properties and phase changes, have received little attention. It is not yet even clear if layers are primarily coupled by thermal or mechanical phenomena.

I have discussed in previous chapters a three-layer mantle: shallow mantle, transition region and lower mantle with the transition region being potentially unstable (buoyant) at high temperature. The oceanic and continental lithospheres and D" are additional layers that may be chemically distinct. The surface layer is partially chemically buoyant (shields) and partially potentially unstable (oceans). It moves around relative to underlying features and both affects and is affected by the temperature in the upper mantle. This type of model needs to be tested by numerical calculation.

Dimensionless Numbers

The theory of convection is littered with dimensionless numbers named after prominent dead physicists. The relative importance of conduction and convection is given by the Péclet number

$$Pe = \frac{vl}{\kappa}$$

where v is a characteristic velocity, l a characteristic length, and κ the thermal diffusivity,

$$\kappa = \frac{\mathcal{K}}{\rho C_P}$$

expressed in terms of conductivity $\mathcal{K}$, density and specific heat at constant pressure. The Péclet number gives the ratio of convected to conducted heat transport. For the Earth Pe is about 10^3 and convection dominates conduction. For a much smaller body, conduction would dominate; this is an example of the scale as well as the physical properties being important in the physics. There are regions of the Earth, however, where conduction dominates, such as at the surface where the vertical velocity vanishes.

The Rayleigh number

$$Ra = \frac{g\alpha\rho d^3\Delta T}{\kappa\eta}$$

is a measure of the vigor of convection due to thermally induced density variations, $\alpha\Delta T$, in a fluid of viscosity η operating in a gravity field g. This is for a uniform fluid layer of thickness d with a superadiabatic temperature difference of AT maintained between the top and the bottom. If the fluid is heated internally, the AT term is replaced by $d^2A/\mathcal{K}$ where A is the volumetric heat production. Convection will occur if Ra exceeds a critical value of the order of 10^3. For large Ra the convection and heat transport are rapid. The Rayleigh number of the mantle is thought to be of the order of 10^5 to 10^7. This depends on the scale of convection as well as the physical properties.

The Nusselt number gives the relative importance of convective heat transport:

Nu = total heat flux across the layer/conducted heat flux in the absence of convection

$$= (Ra/Ra_c)^{1/3}$$

For an internally heated layer Nu is the ratio of the temperature drops across the layer with and without convection or equivalently, the ratio of the half-depth of the layer to the thermal boundary layer thickness.

The Prandtl number Pr

$$Pr = \eta/\kappa$$

for the mantle is about 10^{24}, which means that the viscous response to a perturbation is instantaneous relative to the thermal response.

The Reynolds number, ***Re,*** is

$$Re = vl/\eta = Pe/Pr$$

For v $\approx 10^{-7}$cm/s, a typical plate tectonic rate and $l \approx 10^8$ cm, the dimension of the mantle, $Re \approx 10^{-21}$. For $Re \ll 1$ inertial effects are negligible, and this is certainly true for the mantle.

Convection can be driven by heating from below or within, or by cooling from above. The usual case treated is where the convection is initiated by a vertical temperature gradient. When the vertical increase of temperature is great enough to overcome the pressure effect on density, the deeper material becomes buoyant and rises. An adiabatic gradient simply expresses the condition that the parcel of fluid retains the same density contrast as it rises. Horizontal temperature gradient can also initiate convection. These can be caused by the presence of continents or variations in lithosphere thickness such as at fracture zones or between oceans and continents.

The stability of a fluid layer heated from below is a classic problem. In the simplest, Boussinesq, approximation, the fluid is assumed to be elastically incompressible, changing volume only by thermal expansion, and the onset of convection is controlled entirely by the Rayleigh number. In a thin layer, or on a small planet, the temperature gradient in the Rayleigh number is the actual gradient. In the presence of gravity the gradient is the excess over the adiabatic gradient, which for the upper mantle is about 0.3°C/km. An adiabatic gradient is neutrally stable, and only deviations from it can provide buoyancy. For a uniform fluid, convection is initiated for $Ra > 10^3$ and is characterized by a horizontal wavelength comparable to the thickness of the convecting layer. This is the same as saying that a parcel of fluid traveling along the bottom and becoming warmer becomes buoyant enough to rise when it has traveled a distance about equal to the depth of the layer. Likewise, a parcel moving along the top, becoming colder and denser, becomes unstable after traveling a comparable distance. If the surface layer is prevented from sinking because of its high strength or viscosity, then it may travel greater distances. The spacing of upwelling blobs may therefore be less than the spacing between downwellings. The lower part of the surface boundary may also detach and sink. In a system with a strongly temperature-dependent viscosity, there may therefore be several scales of convection. Midplate volcanism may be a manifestation of the second scale length.

For high Ra and Pr most of the temperature contrast ΔT occurs across boundary layers of thickness δ. In a boundary layer the horizontal convection of heat is balanced by vertical diffusion,

$$\frac{u\Delta T\delta}{H} \approx \kappa\frac{\Delta T}{\delta}$$

where u is the horizontal velocity and H is the cell width. The boundary layer thickness can be written

$$\delta \approx H(Ra_c/Ra)^{1/3}$$

The convective heat flux Q is

$$\dot{Q} \approx \frac{\rho C_P v \Delta T \delta}{H}$$

where $v \approx u$.

Scale of Mantle Convection

The dimensions of the major lithospheric plates are of the order of 10^3 to 10^4 km. This has sometimes been taken as evidence for whole-mantle convection. Plates, however, are strong and have high viscosity and because of the crustal and depleted lithospheric component have some nonthermal buoyancy. They, therefore, can impose their own scale length on the problem. Both the lower mantle and the upper mantle are convecting, but the important question is how they interact with each other. The lower mantle certainly differs in mineralogy and pressure from the upper mantle and therefore has different, probably higher, viscosity and thermal conductivity. The non-Newtonian (stress-dependent) rheology of solids also affects the viscosity profile. Convection is probably much faster in the shallow mantle than in the deeper mantle. The geoid and seismicity both provide evidence that subducting slabs encounter a barrier near 650 km. This could be due to an increase in viscosity or intrinsic density due to chemical differences.

The mantle contains two prominent discontinuities near 400 and 650 km. If these are equilibrium phase boundaries in a homogeneous mantle, they may inhibit but will not necessarily prevent whole-mantle convection. Cold downwelling material will transform to the denser assemblage unless it is unable to reach the appropriate pressure. If it does get deep enough to transform, it should have no trouble sinking through the lower mantle unless the latent heat associated with the phase change causes it to warm up to temperatures in excess of the surrounding mantle. Flow through a phase boundary is governed by two competing effects, density change and latent heat. The olivine-spinel transition is exothermic: The transition pressure decreases with a decrease in temperature. Cold subducted olivine-rich material transforms to the denser phase while the surrounding hotter mantle is still in the less dense phase. The slab therefore has an additional contribution to its negative density, which is only partially removed by the latent heating. The olivine-spinel transition, therefore, does not inhibit convection unless the low slab temperatures prevent the phase change from happening because of slow reaction rates. Even in this case, the slab will eventually warm up and be able to sink through the phase boundary.

In the case of an endothermic reaction, with a negative Clapeyron slope, the slab must push its way into the deeper mantle before it can reach pressures great enough to transform it to the denser phase. When it does it will cool off and may become denser than the surrounding mantle. Some of the phase changes predicted to occur near 650 km have negative Clapeyron slopes. Whether convection is precluded across the interface in this case depends not only on the volume change and latent heat but also on the total deformation of the boundary. In a chemically layered mantle the interface is depressed under cold downwellings in the upper layer by an amount that depends on the integrated density contrast across the upper mantle. If cold, chemically distinct, upper-mantle material is to penetrate the boundary, the boundary deformation must be enough to allow the phase change in the colder material, and the density increase must be enough to provide a negative buoyancy with respect to the lower mantle, otherwise upper-mantle material will be confined to the upper mantle.

The upper and lower mantles will not behave completely independently even if there is a chemical stratification. Hot upwellings in the lower mantle will elevate the boundary and heat the upper mantle. Cold downwellings in the upper layer will depress the boundary and cool the top of the lower mantle. The presence of a conductive boundary layer will decrease the viscosity near the boundary and reduce the shear coupling across it. A large viscosity contrast across the boundary will also inhibit shear coupling.

Three-dimensional tomographic studies of the upper mantle in the northwest Pacific show that the high-velocity slabs become nearly horizontal at depths just beneath the deepest earthquakes (Zhou and Clayton, 1988). This is consistent with slab confinement to the upper mantle.

COOLING OF THE EARTH

The principal contributions to the surface heat flow of the Earth are cooling of the Earth and the heating due to the decay of radioactive isotopes. In 1856 Lord Kelvin used the surface heat flow and a conduction calculation to estimate the age of the Earth as 20–400 Ma. This ignored radioactive heating and convection. Later, it was noted that the present heat flow is about that expected from the radioactivity of chondritic meteorites, and it was suggested that the Earth is in steady-state heat balance between heat generation and heat loss. However, since heat productivity decreases with time, it is likely that convective velocities and temperatures have also decreased with time. The cooling contributes a specific heat component to the observed heat flow, the term treated by Lord Kelvin, and this is not negligible. For example, a cooling rate of only 200 K in 10^9 years gives a surface heat flow of more than 1.5 μcal/cm^2/s.

For a fluid layer of thickness h heated from within a dimensionless temperature, Θ, can be written

$$\Theta = 2\mathcal{K}\overline{T}/\rho Q h^2$$

where $\overline{T}$ is the mean temperature, ρ is the density and Q is the heating rate per unit mass. For a fluid heated from below at rate q,

$$\Theta = \mathcal{K}\overline{T}/qh$$

Turcotte (1980) takes the empirical scaling

$$\Theta \sim Ra^{-n},\ n \approx 1/4$$

and an average heat productivity

$$Q = Q_o \exp(t \ln 2/\tau)$$

$$q = q_o \exp(t \ln 2/\tau)$$

and a temperature-dependent viscosity

$$\eta = \eta_o \exp(E^*/RT)$$

to deduce a cooling rate of

$$\frac{dT}{dt} = -(\ln 2)\left(\frac{1-n}{n}\right)\frac{RT_o^2}{E\tau}$$

where τ is an average radioactive half-life, t is time backwards from the present and T_o is the present mean mantle temperature $(t = 0)$. The calculated cooling rate varies from 90 K/Ga at 3 Ga ago to 36 K/Ga at present. The average temperature of the Earth has therefore decreased at least several hundred degrees since the Earth stopped accreting, and this cooling contributes about 20 percent to the current surface heat flow. (For thermal modeling of accretion itself see Turcotte and Pflugrath, 1985.) The thickness of the surface thermal boundary has increased with time since a lower average thermal gradient suffices to conduct out the internal heat. This means that the temperature of the upper mantle has decreased more than the deeper temperatures. In a chemically stratified mantle, the deeper layers cool more slowly. Upper-mantle melting was more extensive in the past, and higher temperature melts could rise closer to the surface. Today's thick lithosphere prevents high-temperature and high-density melts (such as komatiites) from rising directly to the surface. Some cooling and fractionation in the shallow mantle is probably inevitable. Komatiites and other olivine-rich magmas may, in fact, still be forming in the upper mantle by adiabatic ascent from deeper layers but would be unable to penetrate the present crust and lithosphere.

Estimates of the cooling rate of the Earth, such as the above, are highly uncertain. Cooling may contribute somewhere between 20 and 50 percent of the present heat flow, and this uncertainty makes it difficult to constrain the radioactive abundance of the mantle from the observed heat flow. A chemically layered mantle will cool more slowly than a homogeneous mantle because of the necessity of conducting heat through internal thermal boundary layers. The heat loss from the Earth may also be episodic. The present rate of heat loss may not be typical.

References

Anderson, D. L. (1982) Hotspots, polar wander, Mesozoic convection and the geoid, ***Nature,*** *297,* 391–393.

Andrews, J. (1987) True polar wander, ***J. Geophys. Res.,*** *90,* 7737.

Birch, F., R. F. Roy and E. R. Decker (1968) Heat flow and thermal history in New England and New York. In ***Studies of Appalachian Geology*** (E. Anzen, ed.), 437–451, Interscience, New York.

Cazenave, A. (1984) Thermal cooling of the oceanic lithosphere; new constraints from geoid height data, ***Earth Planet. Sci. Lett.,*** *70,* 395–406.

Cazenave, A., C. Rosemberg-Borot and M. Rabinowicz (1986) Geoid lows over deep-sea trenches, ***J. Geophys. Res., 91,*** 1989–2003.

Chase, C. G. (1979) Subduction, the geoid and lower mantle convection, ***Nature,*** *282,* 464–468.

Clayton, R. W. and R. P. Comer (1988) Reconstruction of mantle heterogeneity by iterative back-projection of travel times: 2. Results for P-waves, to be submitted to *J. Geophys. Res.,* 1988.

Crough, S. T. (1978) Thermal origin of mid-plate hotspot swells, *Geophys. J. R. Astron. Soc.*, *55*, 451.

Crough, S. T. (1979) Hotspot epeirogeny, *Tectonophysics, 61,* 321–333.

Dziewonski, A. M. (1984) Mapping the lower mantle: Determination of lateral heterogeneity in P-velocity up to degree and order 6, *J. Geophys. Res., 89,* 5929–5952.

Goldreich, P. and A. Toomre (1969) *J. Geophys. Res., 74,* 2555–2567.

Gordon, R. G. (1982) The late Maastrichtian paleomagnetic pole of the Pacific Plate, *Geophys. J. R. Astr. Soc., 70,* 129–140.

Gordon, R. G. and C. D. Cape (1981) Cenozoic latitudinal shift of the Hawaiian hotspot and its implications for true polar wander, *Earth Planet. Sci. Lett., 55,* 37–47.

Hager, B. H. (1984) Subducted slabs and the geoid; constraints on mantle rheology and flow, *J. Geophys. Res., 89,* 6003–6015.

Hager, B. H., R. W. Clayton, M. A. Richards, R. P. Comer and A. M. Dziewonski (1985) Lower mantle heterogeneity, dynamic topography and the geoid, *Nature, 313,* 541–545.

Harrison, C. G. A. and T. Lindh (1982) Comparison between the hotspot and geomagnetic field reference frames, *Nature, 300,* 251–252.

Melosh, J. (1975a) *Earth Planet. Sci. Lett., 25,* 322.

Melosh, J. (1975b) *Earth Planet. Sci. Lett., 26,* 353.

Melosh, J. (1980) *Icarus, 44,* 745.

Menard, H. W. and L. M. Dorman (1977) Dependence of depth anomalies upon latitude and plate motion, *J. Geophys. Res., 82,* 5329–5335.

Menard, H. W. and M. K. McNutt (1982) Evidence for and consequences of thermal rejuvenation, *J. Geophys. Res., 87,* 8570–8580.

Morgan, W. J. (1981) Hotspot tracks and the opening of the Atlantic and Indian Oceans. In *The Oceanic Lithosphere* (C. Emiliani, ed.), 443–487, Wiley, New York.

Nakanishi, I. and D. L. Anderson (1984) Measurements of mantle wave velocities and inversion for lateral heterogeneity and anisotropy, Part II: Analysis by the single-station method, *Geophys. J. R. Astron. Soc., 78,* 573–617.

Nataf, H.-C., I. Nakanishi and D. L. Anderson (1986) Measurements of mantle wave velocities and inversion for lateral heterogeneities and anisotropy, Part III: Inversion, *J. Geophys. Res., 91,* 7261–7307.

Parsons, B. and J. G. Sclater (1977) An analysis of the variation of ocean floor bathymetry and heat flow with age, *J. Geophys. Res., 82,* 803–827.

Pollack, H. N. (1980) The heat flow from the earth: *A* review. In *Mechanisms of Continental Drift and Plate Tectonics* (P. A. Davies and S. K. Runcorn, eds.), 183–192, Academic, London.

Pujol, J. and D. M. Fountain (1985) Statistical analysis of the mean heat flow/reduced heat flow relationship for continents and its tectonophysical implications, *J. Geophys. Res., 90,* 11,335–11,344.

Rapp, R. H. (1981) The Earth's gravity field to degree and order 180 using Seaset altimeter data, terrestrial gravity data, and other data, *Report 322, Dept. Geodetic. Sci. and Surv.,* Ohio State University, Columbus.

Richards, M. A. and B. H. Hager (1984) Geoid anomalies in a dynamic Earth, *J. Geophys. Res., 89,* 5987–6002.

Sclater, J., B. Parsons and C. Jaupart (1981) Oceans and continents: similarities and differences in the mechanism of heat transport, *J. Geophys. Res., 86,* 11,535–11,552.

Sclater, J. G., C. Jaupart and D. Galson (1980) The heat flow through oceanic and continental crust and the heat loss of the earth, *Rev. Geophys. Space Phys., 18,* 269–311.

Tanimoto, T. and D. L. Anderson (1984) Mapping convection in the mantle, *Geophys. Res. Lett., 11,* 287–290.

Tanimoto, T. and D. L. Anderson (1985) Lateral heterogeneity and azimuthal anisotropy of the upper mantle: Love and Rayleigh waves 100–250 sec., *J. Geophys. Res., 90,* 1842–1858.

Turcotte, D. L. (1980) On the thermal evolution of the Earth, *Earth Planet. Sci. Lett., 48,* 53–58.

Turcotte, D. L. and J. C. Pflugrath (1985) Thermal structure of the accreting Earth, *J. Geophys. Res., 90,* C541–C544.

Vogt, P. R. (1975) Changes in geomagnetic reversal frequency at times of tectonic change; evidence for coupling between core and upper mantle processes, *Earth Planet. Sci. Lett., 25,* 313–321.

Zhou, H.-W. and R. Clayton (1988) preprint.