Chapter 7

Convection and complexity

. . . if your theory is found to be against the second law of thermodynamics, I can give you no hope; there is nothing for it but to collapse in deepest humiliation.

Eddington

Contrary to current textbooks . . . the observed world does not proceed from lower to higher "degrees of disorder", since when all gravitationally-induced phenomena are taken into account the emerging result indicates a net decrease in the "degrees of disorder", a greater "degree of structuring" . . . classical equilibrium thermodynamics . . . has to be completed by a theory of 'creation of gravitationally-induced structures' . . .

Gal-or

Overview

In 1900 Henri Bénard heated whale oil in a pan and noted a system of hexagonal convection cells. Lord Rayleigh in 1916 analyzed this in terms of the instability of a fluid heated from below. Since that time Rayleigh–Bénard convection has been taken as the classic example of thermal convection, and the hexagonal planform has been considered to be typical of convective patterns at the onset of thermal convection. Fifty years went by before it was realized that Bénard's patterns were actually driven from above, by surface tension, not from below by an unstable thermal boundary layer. Experiments showed the same style of convection when the fluid was heated from above, cooled from below or when performed in the absence of gravity. This confirmed the top-down surface-driven nature of the convection which is now called Marangoni or Bénard–Marangoni convection.

Although it is not generally recognized as such, mantle convection is a branch of the newly renamed *science of complexity*. Plate tectonics may be a *self-driven far-from-equilibrium system* that organizes itself by dissipation in and between the plates, the mantle being a passive provider of energy and material. Far-from-equilibrium systems, particularly those in a gravity field, can locally evolve toward a high degree of order. Plate tectonics was once regarded as passive motion of plates on top of mantle convection cells but it now appears that continents and plate tectonics organize the flow in the mantle. But mantle convection and plate tectonics involve more than geometry and space-filling considerations. The mantle is a heat engine, controlled by the laws of thermodynamics. One can go just so far without physics. Conservation of mass and energy are involved, as are balancing of forces. Although the mantle behaves as a fluid, mineral physics principles and classical solid-state physics are needed to understand this fluid. The effect of pressure suppresses the role of the lower thermal boundary layer at the core–mantle–boundary (CMB) interface.

The slow uplift of the surface of the Earth in response to the removal of an ice cap or drainage of a large lake changes the shape of the Earth and the geoid; this 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 mantle convection has some unusual characteristics. The container has spherical geometry. The 'fluid' has stress-, pressure- and temperature-dependent properties. It is cooled from above and from within (slabs) and heated from within (radioactivity) and from below (cooling of the core and crystallization of the inner core). The boundary conditions and heat sources change with time. Melting and phase changes contribute to the buoyancy and provide additional heat sources and sinks. Mantle convection is driven partly by plate motions and partly by chemical buoyancy. 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 cooling plates may well organize and drive mantle convection, as well as themselves. A mantle with continents on top will convect differently from one with no continents.

The theory of convection in the mantle cannot be decoupled from the theories of solids and petrology. 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 fluid-dynamic calculations or laboratory experiments. There are, however, some general characteristics of convection that transcend these details. Technical details of normal or classical thermal convection can be found in `textbooks on mantle convection`. Plate tectonics and mantle motions, however, are far from normal thermal convection.

Geochemists consider convection and stirring to be equivalent. They use `convecting mantle` as convenient shorthand for what they consider to be the `homogenous upper mantle`. The underlying assumption is that midocean ridge basalts, known for their chemical homogeneity, must come from a well-stirred mantle reservoir.

Generalities

`SOFFE systems are extraordinarily sensitive to boundary` and initial conditions. The corollary is that small differences between computer or laboratory simulations, or between them and the mantle, can completely change the outcome. The effect of pressure suppresses the role of the lower thermal boundary layer (TBL) at the core–mantle–boundary (CMB) interface. The state of stress in the lithosphere defines the plates, plate boundaries and locations of midplate volcanism. Fluctuations in stress, due to changing boundary conditions, are responsible for global plate reorganizations and evolution of volcanic chains. In `Rayleigh–Bénard convection`, by contrast, temperature fluctuations are the important parameters. In plume theory, plates break where heated or uplifted by hot buoyant upwellings. Ironically, the fluid flows in the `experiments by Bénard`, which motivated the `Rayleigh theory of thermal convection`, were driven by surface tension, i.e. stresses at the surface.

Computer simulations of mantle convection have not yet included a self-consistent thermodynamic treatment of the effect of temperature, pressure, melting and volume on the physical and thermal properties; understanding of the 'exterior' problem (the surface boundary condition) is in its infancy. Plate tectonics itself is implicated in the surface boundary condition. Sphericity, pressure and the distribution of radioactivity break the symmetry of the problem and the top and bottom boundary conditions play quite different roles than in the simple calculations and cartoons of mantle dynamics and geochemical reservoirs. Conventional (Rayleigh–Bénard) convection theory may have little to do with plate tectonics. The research opportunities are enormous.

The history of ideas

Convection can be driven by bottom heating, top or side cooling, and by motions of the boundaries. Although the role of the surface boundary layer and *slab-pull* are now well understood and the latter is generally accepted as the prime mover in plate tectonics, there is a widespread perception that active hot upwellings from deep in the interior of the planet, independent of plate tectonics, are responsible for 'extraordinary' events such as plate reorganization, continental break-up, extensive magmatism and events far away from current plate boundaries. Active upwellings from deep in the mantle are viewed as controlling some aspects of surface tectonics and volcanism, including reorganization, implying that the mantle is not passive. This is called the `plume mode of mantle convection`. This has been modeled by the injection of hot fluids into the base of a tank of motionless fluid.

Numerical experiments show that mantle convection is controlled from the top by continents, cooling lithosphere, slabs and plate motions and that plates not only drive and break themselves but can control and reverse convection in the mantle. Studies of the `time dependence in 3D spherical mantle convection with continental drift` show the extreme sensitivity to changes of conditions and give results quite different from simpler simulations. Supercontinents and other large plates generate spatial and temporal temperature variations. The migration of continents, ridges and trenches cause a constantly changing surface boundary condition, and the underlying mantle responds passively. Plates break up and move, and trenches roll back because of forces on the plates and interactions of the lithosphere with the mantle. Density variations in the mantle are, by and large, generated by plate tectonics itself by slab cooling, refertilization of the mantle, continental insulation; these also affect the forces on the plates. Surface plates are constantly evolving and reorganizing although major global reorganizations are infrequent. Plates are mainly under lateral compression although local regions having horizontal least-compressive axes may be the locus of dikes and volcanic chains. The Aegean plate is an example of a 'rigid' plate collapsing, or falling apart, because of changes in stress conditions.

The mantle is generally considered to convect as a single layer (whole mantle convection), or at most two. However, the mantle is more likely to convect in multiple layers as a result of gravitational sorting during accretion, and the density difference between the mantle products of differentiation.

Instabilities

Rayleigh–Taylor (RT) instabilities form when a dense, heavy fluid occurs above a low-density fluid, such as a layer of dense oil placed, carefully, on top of a layer of water. Two plane-parallel layers of immiscible fluid are stable, but the slightest perturbation leads to release of potential energy, as the heavier material moves down under the (effective) gravitational field, and the lighter material is displaced upwards. As the instability develops, downward-moving dimples are quickly magnified into sets of interpenetrating RT fingers or plumes. This process is evident not only in many examples, from boiling water to weather inversions. In mantle geophysics, plumes are often modeled by inserting a light fluid into a tank of a static higher density fluid. This is meant to mimic the instability of a hot basal layer. In the later situation, the instability develops naturally and the density contrast is limited. In the injection experiment, the density contrast is imposed by the experimenter, as is the scale of the upwelling. There is a difference between upwellings of intrinsically hot basal layers and intrinsically light chemical layers. The former case sets up the lateral temperature gradients that are the essence of thermal convection.

Rise of deep diapirs

Delamination and sinking of garnet pyroxenite cumulates, sinking of slabs, and upwelling of mantle at ridges are important geodynamic processes. Diapiric ascent, melt extraction and crystal settling are important processes in igneous petrology. Basic melts apparently separate from magma chambers, or rising diapirs, at depths as great as 90 km and possibly greater. Eclogite sinkers may equilibrate at upper mantle or transition zone depths; they then warm up and rise.

The basic law governing ascent and settling, `Stokes' Law`, expresses a balance between gravitational and viscous forces,

$$V = 2\Delta\rho g R^2/9\eta$$

where R is the radius of a spherical particle or diapir, $\Delta\rho$ is the density contrast, η is the dynamic viscosity and V is the terminal velocity. This equation can be applied to the rising or sinking of blobs through a mantle or a magma chamber, with modifications to take into account non-spherical objects and non-Newtonian viscosity. Additional complications are introduced by turbulence in the magma chamber and finite yield strengths.

Diapirs are usually treated as isolated spheres or cylinders rising adiabatically through a static mantle. Because of the relative slopes of the geotherm, and the melting curve, diapirs become more molten as they rise. At some point, because of the increased viscosity or decreased density contrast, ascent is slowed and cooling, crystallization and crystal settling can occur. The lithosphere serves as a viscosity or strength barrier, and the crust serves as a density barrier. Melt separation can therefore be expected to occur in magma chambers at shallow depths. In a convecting mantle the actual temperatures (adiabatic or subadiabatic) diverge from the melting point as depth increases. In a homogenous mantle, melting can therefore only occur in the upper parts of the rising limbs of convection cells or in thermal boundary layers. The additional buoyancy provided by melting contributes to the buoyancy of the ascending limbs. Although the melts will attempt to rise relative to the adjacent solid matrix, they are embedded in a system that is itself rising and melting further. If broad-scale vertical convection is fast enough, diapirs can melt extensively without fractionating. Fertile, low-melting-point patches, such as eclogite, can melt extensively if surrounded by subsolidus peridotite.

The stresses and temperatures in the vicinity of rising plumes or diapirs are high, and these serve to decrease the mantle viscosity; thus rapid ascent is possible. In order to achieve observed magma temperatures and large degrees of partial melting, allowing for specific and latent heats, melting probably initiates at depths of order 200 km under oceanic ridges and large volcanic provinces, assuming that the mantle is mainly peridotite. The solidus temperature of dry peridotite at this depth is at least 2100 °C. One question is, how fast can material rise between about 200 and 90 km, and is the material at 90 km representative of the deeper mantle source region or has it been fractionated upon ascent? Viscosities in silicates are very stress- and temperature-dependent, and diapirs occur in regions of the mantle that have higher than normal stresses and temperatures. Diapiric emplacement itself is a high-stress process and occurs in regions where mantle convection may have oriented crystals along flow lines. Diapirs may rise rapidly through such low-viscosity material. A 50 km partially molten diapir at a depth of 200 km can rise at a rate of about 40 cm/s. Kimberlites travel an order of magnitude faster still. Crystal settling velocities in magmas are of the order of cm/s. It appears therefore that deep diapirs can rise rapidly enough to entrain both melt and crystals. At depth the melt content and the permeability are low, and melt segregation may be very slow. The fertile low-melting point blobs may be encased in relatively impermeable subsolidus peridotite and can therefore melt extensively as they rise.

In a chemically stratified mantle, for example residual peridotite over eclogite or fertile peridotite, there is a conductive thermal boundary between the convecting layers. In such a region the thermal gradient is in excess of the melting gradient, and melting is likely to initiate at this depth. Eclogite has a melting temperature about 200 °C below that of dry peridotite and melting of fertile blobs is also likely between TBLs. Partial melting causes a reduction in density, and a Rayleigh–Taylor instability can develop. Material can be lifted out of the eclogite or piclogite – eclogite plus peridotite – layer by such a mechanism and extensive melting occurs during ascent to the shallower mantle. At shallow depths peridotite elevated adiabatically from greater depths can also melt and magma mixing is likely, particularly if the diapir is trapped beneath thick

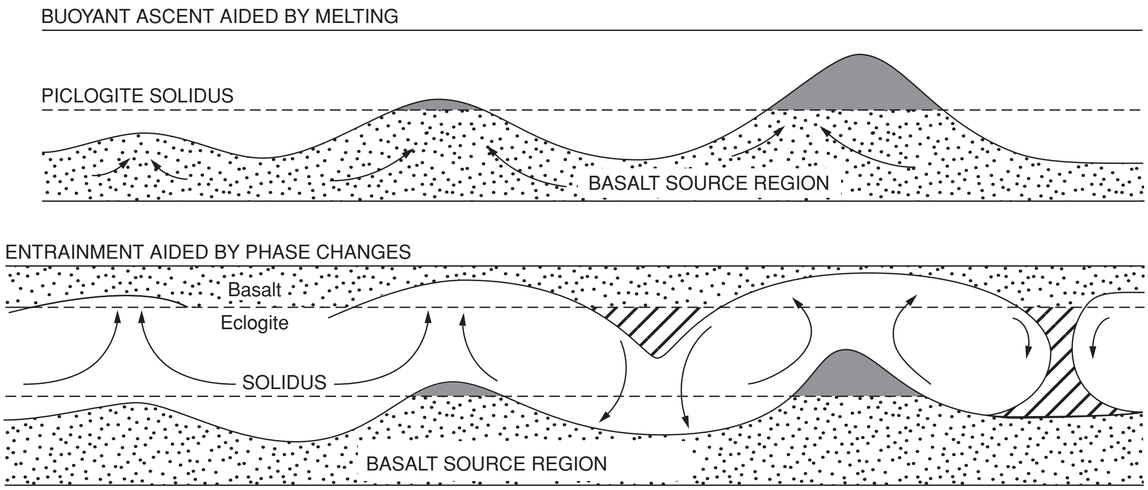

Fig. 7.1 Methods of removing material from a deep, dense layer. In a chemically layered mantle, the interface between layers is highly deformed. This may cause phase changes and melting, and a reduction in density. The deeper layer may also be entrained. If the deep layer is eclogite, from subduction or delamination, it will have a lower melting temperature than ambient mantle and may rise because of melt-induced buoyancy. Adiabatic decompression leads to further melting.

lithosphere. In mid-plate environments, such as Hawaii and other midplate hotspots, melts will cool, fractionate and mix with other melts prior to eruption. Such a mechanism seems capable of explaining the diversity of hotspot (ocean island, continental flood) basalts and midocean ridge magmas.

Material can leave a deep, dense source region by several mechanisms.

(1) Melting in the thermal boundary because of the high thermal gradient compared with melting point gradients.
(2) Melting, or phase changes, due to adiabatic ascent of hotter regions of a layer and crossing of phase boundary.
(3) Entrainment of material by adjacent convecting layers.

Some of these mechanisms are illustrated in Figure 7.1.

The *potential temperature* of the mantle is the temperature of the mantle adiabat if it were to ascend directly to the surface. The potential temperature of the mantle is usually between 1300 and 1400 °C, averaged over large areas. The adiabatic temperature gradient for the solid upper mantle is approximately 0.3–0.5 °C/km. The adiabatic gradient becomes smaller at very high pressures because the thermal expansivities of solids are smaller at high pressures. The adiabatic gradient for liquids (~1 °C/km) is higher than that for solids because the thermal expansivities of liquids are greater than solids.

Melting occurs where the geotherms intersect the mantle solidus. For a dry peridotite mantle, the geotherm (conduction gradient) in a region of high surface heatflow, 100 mW/m^2, intersects the solidus at ~30 km depth. Surface heat flows of 100 mW/m^2 or greater occur only at or near midocean ridges. A geotherm with 40 mW/m^2 surface heatflow, characteristic of the old interiors of continents, never intersects the dry peridotite solidus, implying that partial melting does not occur beneath continental interiors. In suture belts and along arcs there may be low-melting point constituents such as eclogite in the shallow mantle. High heatflow, however, in part, represents intrusion into the plate or thinning of the thermal boundary layer, rather than intrinsically high mantle temperatures.

We can induce melting at low temperatures if we flux the mantle with basalt, eclogite, CO_2 or water, and plate tectonics does all of these things. If portions of the mantle upwell rapidly, either passively in response to spreading at ridges or displacement by sinking slabs, decompressional partial melting occurs. Decompressional melting occurs at greater depths for higher mantle temperatures or lower melting temperatures, allowing higher degrees of partial melting. Any form of adiabatic decompression can give rise to partial melting if the solidus is crossed. Partial melting can be initiated when the plate is thinned by extension or by delamination of the lower crust or removal of the lower part of the TBL. Partial melting can thus occur by an increase in temperature, by adiabatic decompression, or by the depression of the solidus by the presence of eclogite, CO_2 or water. These processes can occur in a variety of tectonic environments, depending upon the background thermal state. Regions of high heat flow, such as midocean ridges and highly extended continental-crustal regions, are characterized by geotherms with steep temperature gradients such that the base of the thermal boundary layer lies well within the partial melting P-T field. In this case, partial melting occurs where the mantle adiabat intersects the solidus. Regions characterized by low heat flow, such as the stable interiors of continents, are not prone to melting as cold geotherms never intersect the solidus of dry peridotite. Partial melting of the mantle wedge overlying subducting slabs occurs because the peridotite solidus has been depressed by the addition of volatiles. Delaminated lower continental crust also lowers the melting point of the mantle that it sinks into.

Forces

Plate tectonic and convective motions represent a balance between driving forces and resisting forces. Buoyancy is the main driving force but there are a variety of resisting forces. Negative buoyancy is primarily created by cooling at the surface. Positive buoyancy is created by heating and melting.

The creation of new plates at ridges, the subsequent cooling of these plates, and their ultimate subduction at trenches introduce forces that drive and break up the plates. They also introduce chemical and thermal inhomogeneities into the mantle. Plate forces such as *ridge push* – a misnomer for the pulling force created by a cooling plate – *slab pull* and *trench suction* are basically gravitational forces generated by cooling plates. They are resisted by transform fault, bending and tearing resistance, mantle viscosity and *bottom drag*. If convection currents dragged plates around, the bottom drag force would be the most important. However, there is no evidence that this is a strong force, and even its sign is unknown (driving or resisting drag). The thermal and density variations introduced into the mantle by subduction also generate forces on the plates.

The pull of subducted slabs – the slab-pull force – is thought to be the main driver of the motions of Earth's tectonic plates and the motions beneath the plates. A slab mechanically attached to a subducting plate can exert a direct pull on the plate; a detached slab may drive a plate by causing a flow in the mantle that exerts a shear traction on the base of the plate. A cold slab can also set up thermal gradients that exert forces on the plates. Slab pull forces may account for about half of the total driving force on plates. Slabs in the deeper mantle are supported by viscous mantle forces and they may reach density equilibration.

Mantle convection may also be driven primarily by descent of dense slabs of subducted oceanic lithosphere. Slab forces cause both subducting and overriding plates to move toward subduction zones and they are also responsible for the migration of trenches and ridges. Ridges and trenches are stationary in the mantle reference frame only in very idealized symmetric cases. Cooling plates also exert forces that cause the plates to move away from ridges and toward subduction zones. One can alternatively think of mantle convection as the passive response to plate-tectonic stresses and thermal gradients created by plate tectonics and lithospheric architecture. Although both the plates and the underlying mantle are parts of the same convecting system it is useful to think of where most of the buoyancy and dissipation in the system resides. In 'normal' convection most of the energy comes from outside the system (bottom heating) and leaves at the top, and the buoyancy (via thermal expansion) and dissipation (viscosity) are distributed internally. In the mantle, much of the energy is provided from within (radioactivity) and much of the buoyancy and dissipation occurs in the surface layer; the upper and lower TBLs are not symmetric;

melting and the redistribution of radioactive elements are important.

Because of the high viscosity of the deep mantle the warm regions are semi-permanent compared to features in the upper mantle. The large viscosity contrast means that the various layers are more likely, on average, to be thermally coupled than shear coupled. From a tomographic point of view, this means that some mantle structures may appear to be continuous even if the mantle is stratified.

Energy

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. One can also view mantle convection as the result of cooling of the surface layer, which then sinks and displaces warmer material upward. In this view, the mantle is passive. 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 base of convecting systems warm up and can also become unstable, generating hot upwellings or plumes. Adiabatically ascending hot upwellings, either passive (responding to slab sinking) or active (plumes), are likely to cross the solidus in the upper mantle, thereby buffering the temperature rise and magnifying the buoyancy. Because of the divergence, with pressure, of melting curves and the adiabat, deep buoyant plumes are subsolidus.

In the simplest model of convection in a homogenous 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 convection depends on the Rayleigh number and the boundary conditions (BC). At moderate Rayleigh number and constant uniform BC three-dimensional patterns result that resemble hexagons or spokes in plan view. Upwellings and downwellings can be different in shape. At very high Rayleigh number the fluid can become turbulent or chaotic. Whole mantle convection with constant properties and no pressure effects would be characterized by a high Rayleigh number and chaotic, well-stirred convection. Layered convection with pressure dependent properties can have very low Rayleigh numbers and an unmixed, heterogenous, mantle.

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 homogenous Newtonian fluid, heated from below, 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. The deformation also introduces dissipation, a role played by internal viscosity 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.

The core–mantle boundary region

The TBL at the base of the mantle generates a potentially unstable situation. The effects of pressure increase the thermal conductivity, decrease the thermal expansion and increase the viscosity. This means that conductive heat transfer from below is more efficient than at the surface, that temperature increases have little effect on density and that any convection will be sluggish. Although the amount of heat coming out of the core may be appreciable, it is certainly less (~10%) than that flowing through the surface. The net result is that lower-mantle upwellings take a long time to develop and they must be very large in order to accumulate enough buoyancy to overcome viscous resistance. The spatial and temporal scale of core–mantle–boundary instabilities are orders of magnitude larger than those at the surface. This physics is not captured in laboratory simulations or calculations that adopt the popular Boussinesq approximation. Pressure also makes it easier to irreversibly chemically stratify the mantle. A small intrinsic density difference due to subtle changes in chemistry can keep a deep layer trapped since it requires such large temperature increases to make it buoyant. Layered-mantle convection is the likely outcome.

Dimensionless scaling relations

The theory of convection is littered with dimensionless numbers named after prominent dead physicists. The importance of these numbers to Earth scientists is that they tell us what kinds of experiments and observations may be relevant to the mantle. Experiments and calculations that are in a different parameter space from the mantle are not realistic. Atmospheric thunderheads and smoke-stack plumes cannot be used as analogs to what might happen in the mantle.

The relative importance of conduction and convection is given by the Peclet number

$$Pe = vL/\kappa$$

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

$$\kappa = K/\rho C_P$$

expressed in terms of conductivity, density and specific heat at constant pressure. The Peclet 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 (L small), 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.

Dynamic similarity depends on two other non-dimensional parameters: the Grashof number, which involves the buoyancy forces and the resisting forces,

$$Gr = g\alpha \Delta T L^3/\nu^2$$

and the Prandtl number ($\mathrm{Pr} = \nu/\kappa$) where ν is the kinematic viscosity. Only when both of these are the same in two geometrically similar situations can the flow patterns be expected to be same. In general, the numbers appropriate for the mantle cannot be duplicated in the laboratory. The Prandtl number for the mantle is essentially infinite, $\sim 10^{23}$. Inertial forces can be ignored in mantle convection.

The Rayleigh number,

$$Ra = \mathrm{Gr\ Pr}$$
$$Ra = g\alpha \Delta T L^3/\nu\kappa$$

is the ratio between thermal driving and viscous dissipative forces, and is proportional to the temperature difference, ΔT, and the cube of the scale of the system. It is a measure of the vigor of convection due to thermally induced density variations, $\alpha\Delta T$, in a fluid of viscosity ν operating in a gravity field g. This is for a uniform fluid layer of thickness d with a superadiabatic temperature difference of $\alpha\Delta T$ maintained between the top and the bottom. If the fluid is heated internally, the $\alpha\Delta T$ term is replacd by 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 depends on the scale of convection as well as the physical properties.

If L is taken to be the depth of the mantle one obtains very large Ra. However, if convection is layered, the scale drops and Ra can become very small. At high pressure, the combination of properties in Ra also drives Ra down.

The Nusselt number gives the relative importance of convective heat transport compared with the total heat flux:

$$\begin{aligned}\mathrm{Nu} &= \text{total heat flux across the layer/conducted heat flux in the absence of convection}\\ &= QL/K\,\Delta T\end{aligned}$$

where Q is the rate of heat transfer per unit unit area, and L and ΔT are the length and temperature difference scales.

The Nusselt number is the ratio of the actual heat flux to the flux that would occur in a purely conducting regime (so it expresses the efficiency of convection for enhancing heat transfer). 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, $\mathrm{Pr} = \nu/\kappa$ varies from about 10^{-2} in liquid metals to 1 for most gases, and slightly more than 1 for liquids such as water and oil. Pr for the mantle is about 10^{24}, which means that the viscous response to a perturbation is instantaneous relative to the thermal response. If one changes a boundary condition, or inserts a crack into a plate, the effect is felt immediately by the whole system. Thermal perturbations, however, take time to be felt. The square root of Pr gives the ratio of the thicknesses of the mechanical boundary layer (MBL) to the thermal boundary layer (TBL). If the container, or the mantle is smaller than this, then convection will organize itself so as to have a small number of upwellings and downwellings. The limiting case is rotation of the whole fluid.

One cannot take one's experience with smoke plumes in the atmosphere or hot plumes in boiling pots of water and apply it to the mantle. One must look at systems with comparable Rayleigh and Prandtl numbers. Narrow plumes are characteristic of high Rayleigh number, low Prandtl number flows. The mantle is the opposite. There are no large velocity gradients in high Pr fluids.

The Reynolds number is defined by $\mathrm{Re} = UL/v$, for a fluid of kinematic viscosity v, flowing with speed U past a body of size L. In aerodynamics, it characterizes similarities between flows with the same Reynolds number. Turbulence at very high Reynolds numbers is expected to be controlled by inertial effects (as viscosity passively smoothes out the smallest scales of motion), as seen for flows around an aircraft or car. The Reynolds number can also be written

$$\mathrm{Re} = \mathrm{Pe}/\mathrm{Pr}$$

For typical plate tectonic rates and dimensions, $\mathrm{Re} \sim 10^{-21}$. For $\mathrm{Re} \ll 1$ inertial effects are negligible, and this is certainly true for the mantle. Re is important in aerodynamics and hydrodynamics but not in mantle dynamics. Inertia can be ignored in plate tectonics. In free thermal convection Re and the velocity are functions of Pr and Ra; they are not independent parameters.

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 gradients can also initiate convection. Convection can be driven in a tank of fluid where the side-walls differ in temperature. There is no critical Rayleigh number in this situation. Lateral temperature gradients can be caused by the presence of continents or variations in lithosphere thickness such as at fracture zones or between oceans and continents. For high Ra and Pr most of the temperature contrast occurs across narrow boundary layers.

In the case of natural convection, velocity is not imposed but is set by buoyancy effects. A central issue is to find a relationship between a temperature difference applied to the system and the corresponding heat flux. Fundamental studies are often concerned with Rayleigh–Bénard convection of a fluid layer heated from below and cooled from above, and where temperature is the only control on density. Natural systems are not so ideal. At very high Ra, the velocity and the transported heat flux are expected to become independent of viscosity and heat conductivity, which is reached in large-scale systems such as the atmosphere.

The Grüneisen parameter can be regarded as a nondimensional incompressibility

$$\gamma = \alpha K_s/\rho C_p = \alpha K_T/\rho C_s$$

It is important in compressible flow calculations. This effect is different from the effect of compression on physical and thermal properties.

The density scale height in a convecting layer is

$$h_d = \delta z/\delta \ln \rho = \gamma C_p/\alpha g = K_s/\rho g$$

where z is the radial (vertical) coordinate.

The dimensionless dissipation number, Di, is

$$\mathrm{Di} = hg\alpha/C_p$$

where h is the thickness of the convecting layer. If h is the thickness of the mantle Di is about 0.5. One hesitates to assign someone's name to this number, even a dead physicist. The dissipation number divided by the Grüneisen ratio is the ratio of the thickness of the convecting layer to the density-scale height or $h_d g\rho/K_s$. When Di is large, the assumption of incompressible flow is not valid. Nevertheless, the incompressible mass conservation equation is usually adopted in mantle convection studies. Compression also changes the physical properties of the mantle, the Rayleigh number and the possibility of chemical stratification.

The buoyancy ratio is

$$B = \Delta\rho_c/\rho\alpha\Delta T$$

where $\Delta\rho_c$ is the intrinsic chemical density contrast between layers. When this is small we have purely thermal convection but when it is large the dense components can no longer be entrained and chemical layering results. Pressure serves to decrease α and therefore to stabilize chemically stratified convection. In discussions of layered mantle convection B is the most important parameter.

Rayleigh-like numbers

The first order questions of mantle dynamics include:

(1) Why does Earth have plate tectonics?
(2) What controls the onset of plate tectonics, the number, shape and sizes of the plates, the locations of plate boundaries and the onset of plate reorganization?
(3) What is the organizing principle for plate tectonics; is it driven or organized from the top or by the mantle? What, if anything, is *minimized*?

Surprisingly, these are not the questions being addressed by mantle geodynamicists or computer simulations.

Marangoni or Bénard–Marangoni convection is controlled by a dimensionless number,

$$M = \sigma\Delta TD/\rho\nu\kappa$$

where σ is the temperature derivative of the surface tension, S, ΔT is the temperature difference, D is the layer depth, and ρ, ν and κ are the density, kinematic viscosity and thermal diffusivity of the layer. Fluid is drawn up at warm regions of the surface and flows toward cell boundaries where it returns to the interior. Surface tension forces replace thermal buoyancy which appears in the Rayleigh number. Systems with large Rayleigh or Marangoni numbers are far from conductive equilibrium; they are SOFFE systems.

In a fluid cooled from above, even without surface tension, the cold surface layer becomes unstable and drives convection in the underlying fluid when the local Rayleigh number of the thermal boundary layer (TBL) exceeds a critical value. Like Marangoni convection, this type of convection is driven from the top. Cold downwelling plumes are the only active elements; the upwellings are passive, reflecting mass balance rather than thermal instabilities. Plate tectonics, to a large extent, is driven by the unstable surface thermal boundary layer and therefore resembles convection in fluids which are cooled from above.

Pressure decreases α and increases ν and κ so it is hard to generate buoyancy or vigorous small-scale convection at the base of the mantle. In addition, heat flow across the CMB is about an order of magnitude less than at the surface so it takes a long time to build up buoyancy. In contrast to the upper TBL (frequent ejections of narrow dense plumes), the lower TBL is sluggish and does not play an active role in upper mantle convection.

There are additional surface effects. Lithospheric architecture and slabs set up lateral temperature gradients that drive small-scale convection. For example, a newly opening ocean basin juxtaposes cold cratonic temperatures of about 1000 °C at 100 km depth with asthenospheric temperatures of about 1400 °C. This lateral temperature difference, ΔT, sets up convection, the vigor of which is characterized by the Elder number,

$$\mathrm{El} = g\alpha\Delta T L^3/\kappa\nu$$

where L is a characteristic horizontal dimension, e.g. the width of a rift or an ocean basin or the distance between cratons, and ν is now the viscosity of the asthenosphere. This kind of small-scale convection has been called `EDGE, for edge-driven gyres and eddies`. Convective flows driven by this mechanism can reach 15 cm/year and may explain volcanism at the margins of continents and cratons, at oceanic and continental rifts, and along fracture zones and transform faults. Shallow upwellings by this mechanism are intrinsically 3D and may create such features as Iceland and Bermuda.

The role of pressure in mantle convection

Pressure decreases interatomic distances in solids and this has a strong nonlinear effect on such properties as thermal expansion, conductivity and viscosity, all in the direction of making it difficult for small-scale thermal instabilities to form in deep planetary interiors. Convection is sluggish and large-scale at high pressure. The Boussinesq approximation, widely used in geodynamics calculations, assumes that density, or volume (V), is a function of temperature (T) but that all other properties are independent of T, V and pressure (P), even those that are functions of V. This approximation, although thermodynamically (and algebraically) inconsistent, is widely used to analyze laboratory convection and is also used in mantle convection simulations. Sometimes this approximation is supplemented with a depth-dependent viscosity or with T-dependence of parameters other than density. It is preferable to use a thermodynamically self-consistent approach. To first order, the properties of solids depend on interatomic distances, or lattice volumetric strain, and to second order on what causes the strain (T, P composition, crystal structure). This is the basis of `Birch's Law`, the `seismic equation of state`, various laws of `corresponding states` and the `quasiharmonic approximation`.

Volume as a scaling parameter

As far as physical properties are concerned, the main effects of pressure, temperature and phase changes are via volume changes. The thermal, elastic and rheological properties of solids depend on interatomic distances, or lattice volumetric strain, and are relatively indifferent as to what causes the strain (T, P or crystal structure). Intrinsic temperature effects are those that occur at constant volume. The quasiharmonic approximation is widely used in mineral physics but not in seismology or geodynamics where less physically sound relationships are traditionally used.

A parameter that depends on P, T, phase (ϕ) and composition © (within limits, e.g. constant mean atomic weight) can be expanded as

$$M(P, T, \phi, ©) = M(V) + \varepsilon$$

where ε represents higher-order intrinsic effects at constant molar V. Lattice dynamic parameters and thermodynamic and anharmonic parameters are interrelated via V.

Beyond Boussinesq

The effect of volume changes on thermodynamic properties are determined by dimensionless parameters. Scaling parameters for volume-dependent properties can be written as power laws or as logarithmic volume derivatives about the reference state;

Lattice thermal conductivity	$-\mathrm{d}\ln \kappa_L/\mathrm{d}\ln V \sim 4$
Bulk modulus	$-\mathrm{d}\ln K_T/\mathrm{d}\ln V \sim 4$
Thermal expansivity	$-\mathrm{d}\ln \alpha/\mathrm{d}\ln V \sim -3$

Volume changes in laboratory convection experiments are small, so the changes of thermal parameters associated with volume changes are small, with the possible exception of viscosity. The Boussinesq approximation ignores these effects. The specific volume at the base of the mantle is 64% of that at the top. Compression, composition and phase changes, and to some extent, temperature, are all involved. When the above numbers are multiplied by $\Delta V/V$ the changes in physical properties are non-negligible for the mantle. Although volume scalings such as the Debye theory and the `quasiharmonic approximation` are strongly grounded in classical physics they have not been implemented in mantle convection codes.

Top-down tectonics

One can think of mantle convection as having various origins. The mantle is cooled from above; instability of the cold upper TBL is a *top-down* mechanism, which is basically plate tectonics. Heating the mantle from below is a *bottom-up* or thermal plume mechanism. Lithospheric architecture provides a *sideways* or EDGE mechanism that is lacking in Rayleigh–Bénard convection, or in simple fluids with simple boundary conditions. Internal heating generates time-dependent upwellings, an *inside-out* mechanism. *Delamination* is a *bottoms-off* thermo-chemical mechanism that does not involve the whole outer shell. Cooling of the surface and the motions of plates and plate boundaries, and their effect on the underlying mantle, constitute the main, or large-scale, mode of planetary convection. *Small-scale* convection takes the form of gyres and eddies, rolls and sprouts; these are secondary effects of plate tectonics. *Edge-driven* convection, and stress variations and cracks in the plates are all consequences of plate tectonics and offer explanations of volcanic chains and volcanism that are not at plate boundaries. Lateral variations in temperature, melting temperature, density and fertility of the upper mantle are also consequences of plate tectonics; recycling, continental insulation and slab cooling can explain variations in volcanic output from place to place. So-called *midplate volcanism, melting anomalies* and *hotspots* can be consequences of plate tectonics, and do not require high mantle temperatures or deep fluid dynamic instabilities.

Early views of plate tectonics treated plates as responding passively to mantle convection. Plates and continents drifted about passively on the surface. Ridges were the upwellings and slabs were the downwellings. Narrow hot upwellings were generally held responsible for 'hotspots,' 'hotlines' and 'hotspot tracks'; giant upwellings or plume heads were held responsible for large igneous provinces and continental break-up and for influencing plate motions. These ideas developed from the tacit assumptions that the lithosphere is rigid and uniform, the upper mantle is isothermal, generally subsolidus, and homogenous, and that locations of volcanoes are controlled by mantle temperature and convection, not the stress state of the lithosphere.

It was then recognized that plates could drive themselves and could also organize the underlying mantle convection. Plates and the architecture of the lithosphere provided the template and stress conditions for midplate magmatism and tectonics, phenomena not obviously related to plate tectonics or plate boundaries in the context of rigid plates. Variations in the thickness of the crust and the ages of continents, and the cooling of oceanic plates, set up lateral temperature gradients which can be just as important in driving mantle convection as the non-adiabatic temperature gradients in TBLs. Secular cooling of the Earth maintains a surface TBL; cooling from above can initiate and maintain mantle convection and plate tectonics. Although the mantle is, to some extent, heated from within, and from below, it is basically a system that is driven from the top.

The tectonic plate system can be viewed as an open, far-from-equilibrium, dissipative and self-organizing system that takes matter and energy from the mantle and converts it to mechanical forces (ridge push, slab pull), which drive the plates. Subducting slabs, delamination and cratonic roots cool the mantle and create pressure and temperature gradients that drive mantle convection. The plate system thus acts as a template to organize mantle convection. In contrast, in the conventional view, the lithosphere is simply the surface boundary layer of mantle convection and the mantle is the self-organizing dissipative system. If most of the buoyancy and dissipation – the alternative to mantle viscosity – is provided by the plates while the mantle simply provides heat, gravity, matter, and an entropy dump, then plate tectonics is a candidate for a self-organized system, in contrast to being organized by mantle convection or heat from the core. Stress fluctuations in such a system cause global reorganizations without a causative convective event in the mantle. Changes in stress affect plate permeability and can initiate or turn off fractures, dikes and volcanic chains. The mantle itself need play no active role in plate tectonic 'catastrophes.'

The traditional view of mantle geodynamics and geochemistry is that magmatism, and phenomena such as continental break-up and plate reorganization, are due to convection currents in the mantle, and the importation of core heat, via plumes, into the upper mantle. Mantle convection by-and-large controls itself, and can experience massive overturns called `mantle avalanches`. But mantle dynamics may be almost entirely a top-down system and it is likely that mantle convection of various scales is controlled by plates and plate tectonics, not vice versa. The surface boundary layer is the active element, the 'convecting mantle' is the passive element. When a plate tectonic and continental template is placed on top of the convecting system, it organizes the convective flow and the plates themselves become the dissipative self-organized system.

Plate driven flow

Marangoni convection is driven by surface tension. Since surface tension is isotropic, the fluid flows radially from regions of low surface tension to the cell boundaries, which are hexagonal in planform, where linear downwellings form. The equivalent surface force in mantle convection is the *ridge-push–slab-pull* gravitational force which has the same units as surface tension. Since plates are not fluids the forces are not isotropic. Plates move from ridge to trench, pulling up material at diverging regions, which are the equivalent of the centers of Bénard–Marangoni hexagons, and inserting cold material at subduction zones. The other difference between Marangoni and plate-driven convection is that plates are held together by lateral compression and fail in lateral extension. Cell boundaries are convergent and elevated and are regions of compressive stress in Marangoni convection.

The plate-tectonic equivalent of the `Marangoni number` can be derived by replacing surface tension by plate forces. I define the plate tectonic or Platonic number

$$\mathrm{Pl} = g\alpha \Delta T L^2 / U\mathcal{D}$$

where L is a characteristic length (e.g. ridge-trench distance) and U is plate velocity. $\mathcal{D}$ is a dissipation function, which accounts for plate deformation, intraplate resistance and mantle viscosity. It is a rheological parameter. The `roles` of plate bending and fault strength at subduction zones may be as important as mantle viscosity in mantle dynamics and in controlling the cooling of the mantle. When the only resisting forces are lithospheric bending we have the dimensionless lithosphere number

$$\mathrm{Pl} = g\alpha\Delta T r^3/\kappa\nu_1$$

where r is the radius of curvature of the bend and ν_1 is the lithospheric viscosity. The thermal evolution of an Earth with strong subduction zones is quite different from one with a completely fluid mantle.

When plate interactions are involved we also need coupling parameters across plates such as transform fault resistance and normal stress. In the plate-tectonic system the plates (and slabs) account for much of both the driving force and dissipation (see Figure 4.8) and in this respect they play the role of the convecting fluid in Rayleigh–Bénard convection (internal sources of both buoyancy and viscous dissipation). Both the buoyancy and dissipative stresses affect the whole system. The plate sizes, shapes and velocities are self-controlled and should be part of the solution rather than input parameters. Even the rheology of the surface material may be self-controlled rather than something you can look up in a handbook. *Foams* and *grains* are examples of *fragile* or *soft* materials that to some extent control their own fates. The behavior of rocks, and probably the lithosphere, is controlled by damage rheology and complex feedback processes involving fracture and friction rather than fluid mechanics.

Layered mantle convection

If pressure is ignored it requires about a 6% increase in *intrinsic density* for a deep mantle layer to be stable against a temperature-induced overturn. Plausible differences in density between the silicate products of accretional differentiation, which are intermediate in density between the crust and the core, are about 1 or 2%, if the variations are due to changes in silicon, aluminum, calcium and magnesium. Changes in iron content can give larger variations. Such density differences have been thought to be too small to stabilize stratification. However, when pressure is taken into account chemical stratification is likely; dense layers become trapped, although the effect on seismic velocities can be slight. These are therefore *stealth* layers or reservoirs which are below the ability of seismic waves to detect, except by special techniques.

The lower parts of the mantle are now at high pressure and low temperature compared with accretional conditions. Chemically distinct dense material, accumulated at the base of the mantle, must become very hot in order to become buoyant, because of the very low thermal expansivity at lower-mantle pressures. Scaling relations also show that only very large features would accumulate enough buoyancy to rise.

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. 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 homogenous 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.

The magnitude of the Rayleigh number is a measure of the vigor of convection and the distance from static equilibrium. Most geodynamic discussions assume Ra to be 10^6 to 10^8. Using parameters appropriate for the base of the mantle yields a value of 4000. If the lower 1000 km of the mantle acts as an isolated layer, because of high intrinsic density, Ra drops to 500. The critical Rayleigh number in a spherical shell is about 10^3. The implication of these results are far-reaching. Instabilities at the base of the mantle must be sluggish and immense not narrow or plume-like. These inferences are consistent with lower mantle tomography and make it more plausible than previously thought for the mantle to be chemically stratified. Deep dense layers need only be a fraction of a percent denser that the overlying layers in order to be trapped. This is a consequence of low α (coefficient of thermal expansion) and the difficulty of creating buoyancy with available-temperature variations and heat sources. The gravitational differentiation of the deep mantle may be irreversible and ancient. The widespread belief that the mantle convects as a unit (whole mantle convection, deep slab penetration) or in accessible layers is based on non-Boussinesq calculations or experiments. A corner-stone of the standard model of mantle geochemistry is that the upper mantle is homogenous and is therefore vigorously convecting. The considerations in this section make this unlikely. If the mantle is heterogenous in chemistry and melting point the geochemical motivation for deep-mantle plumes disappears.