Mantle convection

Encyclopedia of Solid Earth Geophysics, Harsh Gupta (ed.), Springer

David Bercovici
Professor & Chair
Department of Geology & Geophsyics
Yale University
New Haven, Connecticut 06520-8109, USA
E-mail: david.bercovici@yale.edu
Tel: (203) 432-3168
Fax: (203) 432-3134

December 26, 2010

MANTLE CONVECTION

Synonyms

Mantle dynamics. Mantle circulation.

Definition

Mantle convection: Thermal convection in the terrestrial planetary mantles, the rocky layer between crust and core, in which hot material rises, cold material sinks and the induced flow governs plate tectonic and volcanic activity, as well as chemical segregation and cooling of the entire planet.

Mantle convection

Introduction and History

All planetary bodies retain some heat from their early formation but are inexorably cooling to space. Planetary surfaces are therefore cold relative to their hotter interiors, and thus undergo thermal convection wherein cold material is dense and sinks while hot material is light and rises (liquid water near freezing being one of the rare exceptions to this process). Planetary atmospheres, oceans, rocky mantles and metallic liquid cores convect and are subject to unique patterns of circulation in each domain. Silicate mantles however tend to be the most massive and sluggish part of terrestrial planets and therefore govern how planetary interiors evolve and cool to space. (See Fig 1.)

The theory of mantle convection was originally developed to understand the thermal history of the Earth and to provide a driving mechanism for Alfred Wegener's theory of Continental Drift in the 1930s [see *Schubert et al.*, 2001; *Bercovici*, 2007]. Interest in mantle convection waned for decades as Wegener's theory was criticized and apparently discredited. However, the accumulation of sea-floor sounding data during War World II and refinement of paleomagentic techniques paved the way for the discovery of sea-floor spreading [*Hess*, 1962; *Vine and Matthews*, 1963] and the birth of the grand unifying theory of Plate Tectonics in the 1960s; this consequently revived interest in mantle convection as the driving mechanism for plate motions [*Runcorn*, 1962a,b] as well as non-plate-tectonic volcanism such as the possible Hawaiian plume [*Morgan*, 1971]. The success of mantle convection theory in explaining plate velocities, sea-floor subsidence, volcanism, gravity anomalies, etc., lead to its further application to other terrestrial planets such as Venus and Mars, which also sustained unique forms of mantle convection, evident from volcanic activity.

Basics of thermal or free convection

Rayleigh-Bénard Convection

The simplest form of thermal convection is referred to as Bénard convection named after the French experimentalist Henri Bénard who in 1900 performed the first systematic experiments on convection in thin layers of oil (spermacetti) and recognized both the onset of convection from a static conductive state and the regular patterns formed in a convecting layer [*Bénard*, 1900, 1901]. Fifteen years later, the British theoretical physicist and mathematician Lord Rayleigh (William John Strutt), attempted to explain Bénard's results for the onset of convection [*Strutt, John William (Lord Rayleigh)*, 1916] – the delay in communication between them being caused by the First World War. However, the mismatch between theory and experiment was profound, and not resolved until the late 1950s [*Pearson*, 1958] when it was inferred that Bénard's experiments were strongly influenced by surface tension or Marangoni effects not included in Rayleigh's theory (although Bénard himself was aware of these effects). Because Rayleigh's work provided the framework for nearly all thermal convection theory to follow, the simple Bénard convective sys-

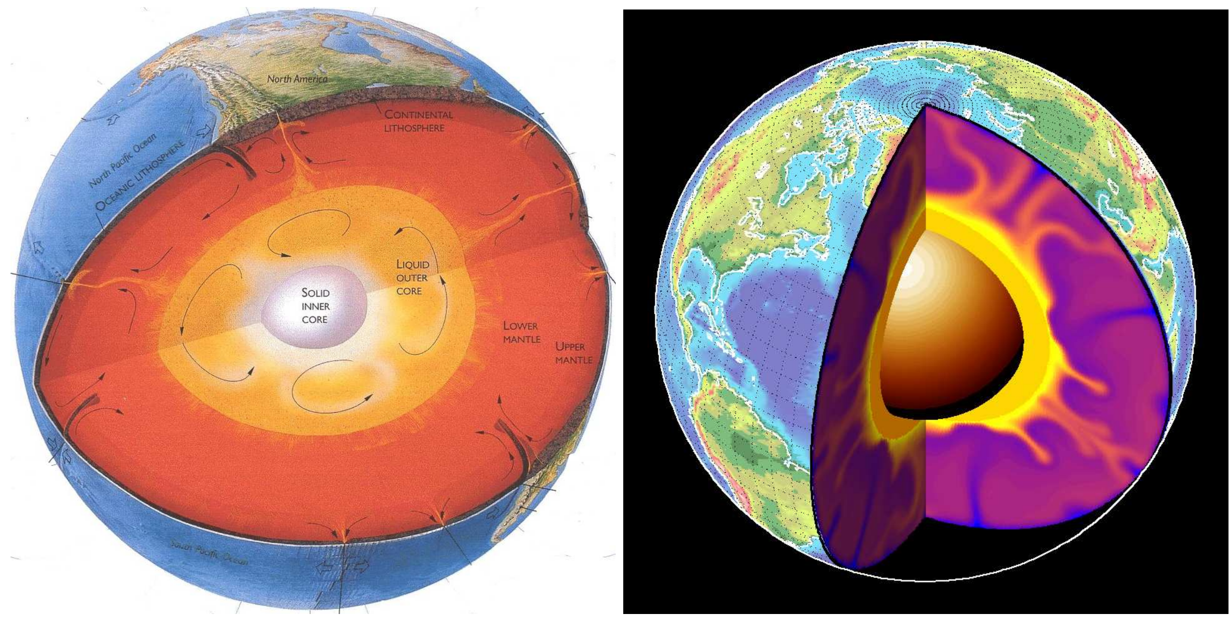

Figure 1: Graphic renditions of cut aways of Earth's structure showing crust, mantle and core (left) and of the convecting mantle (right). The relevant dimensions are that the Earth's average radius is 6371km; the depth of the base of the oceanic crust is about 7km and continental crust about 35km; the base of the lithosphere varies from 0 at mid-ocean ridges to about 100km near subduction zones; the base of the upper mantle is at 410km depth, the Transition Zone sits between 410km and 660km depths; the depth of the base of the mantle (the core-mantle boundary) is 2890km; and the inner core-out core boundary is at a depth of 5150km. Left frame adapted from *Lamb and Sington* [1998]. Right frame, provenance unknown.

tem is also referred to as Rayleigh-Bénard convection.

Although Bénard's experiments were in a metal cavity, Rayleigh-Bénard convection actually refers to Rayleigh's idealized model of a thin fluid layer infinite in all horizontal direction such that the only intrinsic length scale in the system is the layer thickness. The Rayleigh-Bénard system is heated uniformly on the bottom by a heat reservoir held at a fixed temperature (i.e., the bottom boundary is everywhere isothermal) and the top is likewise held at a fixed colder temperature by another reservoir (see Figure 2). If the layer were not fluid, heat would flow from the hot boundary to the cold one by thermal conduction. But since the fluid near the hotter base is (typically) less dense than the fluid near the colder surface, the layer is gravitationally unstable, i.e., less dense material underlies more dense material. To release gravitationaly potential energy and go to a minimum energy state, the layer is induced to turn over.

Convective onset and the Rayleigh number

While the fluid in a Rayleigh-Bénard layer might be *gravitationally* unstable, it is not necessarily *convectively* unstable. Convective overturn of the layer is forced by heating but resisted or damped in two unique ways. Clearly the thermal buoyancy (proportional to density contrast times gravity) of a hot fluid parcel rising from the bottom surface through colder surroundings acts to drive convective overturn. However, viscous drag acts to slow down this parcel, and thermal condcution, or diffusion, acts to erase its hot anomaly (i.e., it loses heat to its colder surroundings). Thus while the fluid layer might be gravitationally unstable, hot parcels rising might move too slowly against viscous drag before being erased by thermal diffusion. Similar argu-

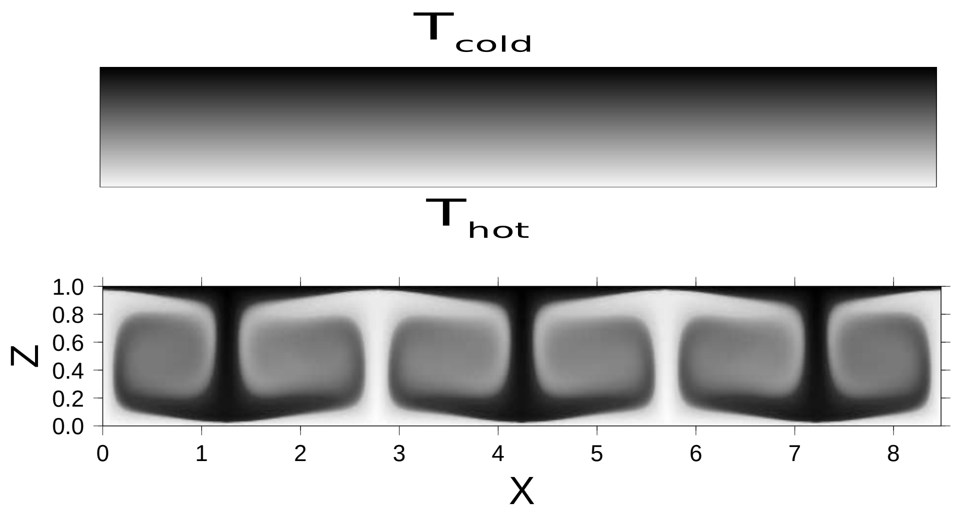

Figure 2: Rayleigh-Benard convection: initially conducting layer (top) and numerical simulation of convection (bottom).

ments can be made for cold material sinking from the top surface through warmer surroundings. The competition between forcing by thermal buoyancy, and damping by viscosity and thermal diffusion, is characterized in dimensionless ratio called the Rayleigh number

$$Ra = \frac{\rho g \alpha \Delta T d^3}{\mu \kappa} \tag{1}$$

where ρ is fluid density, g is gravity, α is thermal expansivity (units of of K^{-1}), ΔT is the difference in temperature between the bottom and top surfaces, d is the layer thickness, μ is fluid viscosity (units of $\mathrm{Pa\ s}$) and κ is fluid thermal diffusivity (units of $\mathrm{m^2 s^{-1}}$).

Even though $\Delta T > 0$ (i.e., heating is from below and causes gravitational instability), Ra still must exceed a certain value, called the critical Rayleigh number Ra_c for convection to occur. For $Ra < Ra_c$ the layer is stable and transports heat by conduction; for $Ra > Ra_c$ the layer will be convectively unstable and transport heat more rapidly via convection. See Figure 3.

Although Ra_c varies depending on the mechanical nature of the horizontal boundaries (whether rigid or a free surface) it is typically of order 1000. This value is easily understood by considering the fate of a hot (or cold) parcel of size a and temperature anomaly ΔT. Dimensional analysis readily shows that the typical ascent rate of the parcel is $\rho g \alpha \Delta T a^2/\mu$ (with units of $\mathrm{m/s}$). However the rate that heat diffuses out of the parcel is κ/a (smaller parcels lose heat faster). The critical state occurs when these two rates are equal; i.e., if the buoyant ascent rate just exceeds the diffusion rate, the parcel should rise without being erased, but if the ascent rate is less than the diffusion rate it will not rise very far before being lost. Therefore the critical state occurs if $\rho g \alpha \Delta T a^3/(\mu\kappa) \approx 1$. Scaling purely by orders of magnitude, a small parcel of fluid can be assumed to be of order 10 times smaller than the entire layer; thus assuming $a \approx d/10$ leads to a critical condition for onset of convection of $\rho g \alpha \Delta T d^3/(\mu\kappa) \approx 1000$.

For the Earth's mantle, the typical average properties from which the Rayleigh number is constructed are $\rho \approx 4000\mathrm{kg/m^3}$, $g = 10\mathrm{m/s^2}$, $\alpha = 3\times10^{-5}\mathrm{K^{-1}}$, $\Delta T \approx 3000\mathrm{K}$, $d = 2900\mathrm{km}$, $\mu = 10^{22}\mathrm{Pa\ s}$ (dominated by the lower mantle), and $\kappa = 10^{-6}\mathrm{m^2/s}$ [see *Schubert et al.*, 2001]. Taken together these lead to a Rayleigh number of approximately 10^7, which is well beyond supercritical; although the mantle viscosity is extremely high, the mantle is also very hot and very large and hence convecting vigorously.

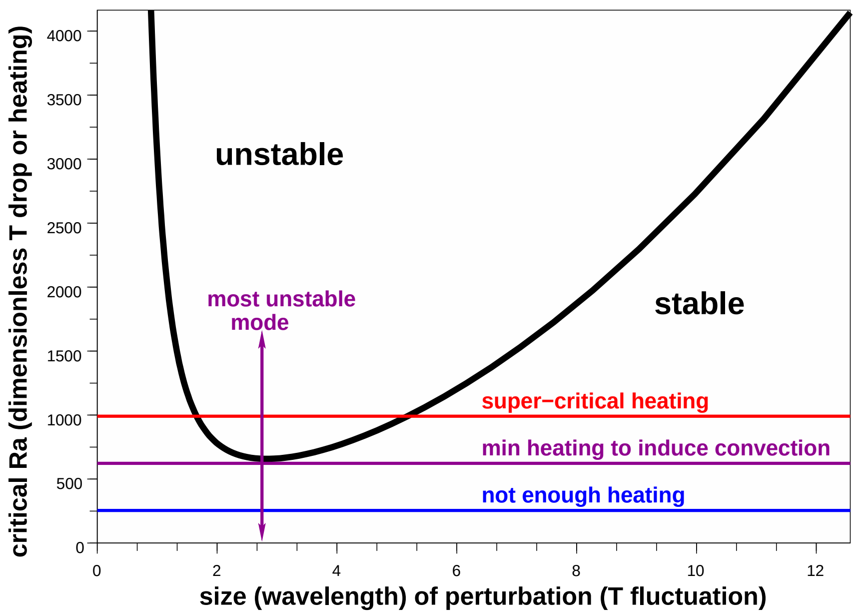

Figure 3: The critical Rayleigh number Ra for the onset of convection is a function of wavelength or size of the thermal perturbation to a static conductive state. For a layer with isothermal and free-slip (shear-stress free) top and bottom boundaries, the relationship is $Ra_{crit} = (k^2 + \pi^2)^3/k^2$ where $k = 2\pi/\lambda$ and λ is wavelength [*Chandrasekhar*, 1961]. Values of Ra above the Ra_{crit} curve are associated with the conductive layer being convectively unstable (perturbations grow), while below the curve the layer is stable (perturbations decay). The minimum in the Ra_{crit} curve occurs at the wavelength of the first perturbation to go unstable as heating and Ra is increased, often called the most unstable mode.

Thermal boundary layers and the Nusselt number

For a Rayleigh number Ra above the critical value Ra_c, convective circulation will mix the fluid layer, and the mixing and homogenization of the fluid will become more effective the more vigorous the convection, i.e., as Ra is further increased. With very large Ra and vigorous mixing most of the layer is largely uniform and isothermal. (In fact, if the layer is deep enough such the pressures are comparable to fluid incompressibility, the fluid layer is not isothermal but adiabatic, wherein even without any heating or cooling, the temperature would drop with fluid expansion on ascent and increase with fluid compression on descent.) Most of the fluid in the Rayleigh-Bénard system is at the mean temperature between the two boundary temperatures. However the temperature still must drop from the well-mixed warm interior to the cold temperature at the top, and to the hotter temperature at the bottom. The narrow regions accomodating these jumps in temperature are called thermal boundary layers (Figure 4).

Thermal boundary layers are of great importance in thermal (and mantle) convection for two reasons. First, most of the buoyancy of the system is bound up in thermal boundary layers since these are where most of the cold material at the top and hot material at the bottom resides and from where hot and cold thermals or currents emanate. Moreover, with regard to convection in the mantle itself, the top cold thermal boundary layer is typically associated with the Earth's lithosphere, the 100km or so thick layer of cold and stiffer mantle rock that is nominally cut up into tectonic plates.

Second, since fluid in these boundary layers

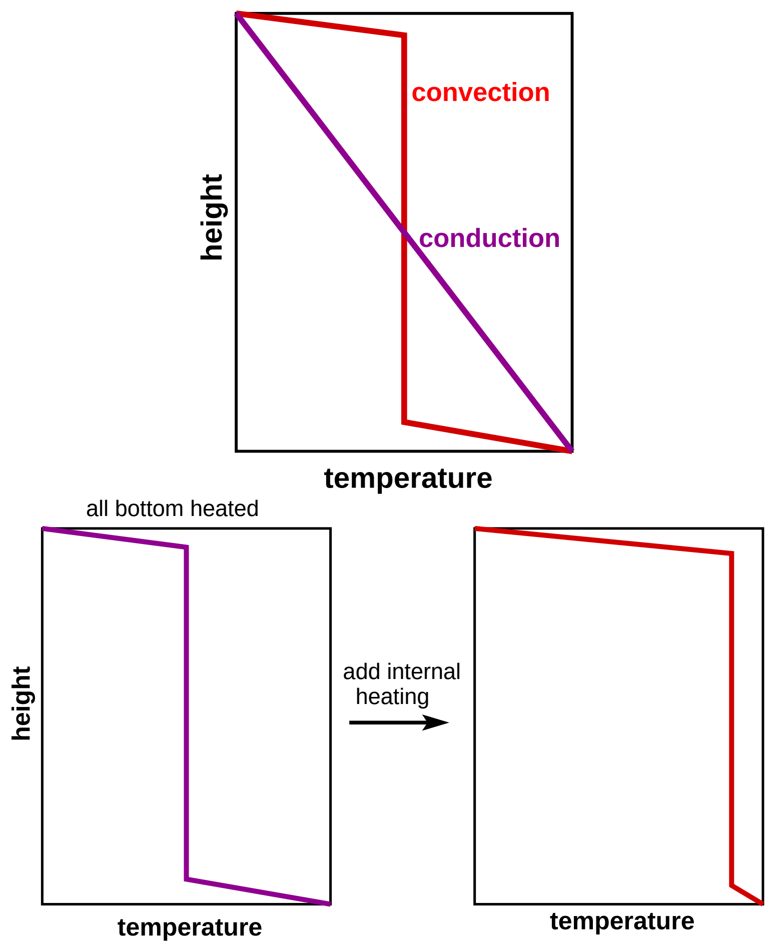

Figure 4: Sketch of temperature profiles, showing how convective mixing homongenizes the conductive mean temperature into a nearly isothermal state (if the fuid is incompressible) with thermal boundary layers connecting it to the cold surface and hot base (top frame). With no internal heating the interior mean temperature is the average of the top and bottom temperatures; the effect of adding internal heating (bottom frames) is to increase the interior mean temperature and thus change the relative size and temperature drop across the top and bottom thermal boundary layers.

is near a horizontal boundary, most of the fluid motion is horizontal and thus heat is only transported vertically across these thin layers by conduction; but since the layers are very thin such conductive transport is rapid. Indeed, the entire cycle of heat transport occurs by heat conducted in rapidly through the bottom boundary layer, after which hot fluid in this layer will, in various spots, become unstable and rise to form a convective upwelling that carries heat out of the boundary layer rapidly across most of the fluid layer and deposits it at the upper boundary, where the heat is then transported out of the system by conduction across the top boundary layer. The eventual heat flow (power output per unit area) out of the well mixed layer is essentially $k\Delta T/\delta$ where k is thermal conductivity (units of $W\ K^{-1}m^{-1}$), $\Delta T/2$ is the temperature drop from the well mixed interior to the surface and we define $\delta/2$ is the thickness of the boundary layer. By comparison, the thermal conduction across a static non-convecting layer is $k\Delta T/d$. The ratio of heat flow in the convectively well mixed layer to the purely conductive layer is thus d/δ, which is called the Nusselt number Nu (named after the German engineer, Wilhelm Nusselt 1882-1957). The relation between Nu and convective vigor parameterized by Ra is important for understanding how convection transports heat and cools off bodies including planets. Convective heat transport is often written as $Nu(k\Delta T/d)$ and in considering this relation *Howard* [1966] argued that vigorous convective heat transport across the bulk of the well-mixed fluid layer is so fast it is not the rate limiting factor in releasing heat (only conduction across the thermal boundary layer is), and thus heatflow should be independent of fluid depth d; this implies that since $Ra \sim d^3$ then $Nu \sim Ra^{1/3}$, which yields a convective heatflow $Nu(k\Delta T/d)$ that is independent of d. In general, since the fluid is conductive for $Ra \leq Ra_c$, one often writes that $Nu = (Ra/Ra_c)^{1/3}$ (although $Nu = 1$ for $Ra < Ra_c$), which is a reasonably accurate relationship born out by simple experiments and computer modeling. This relationship also implies that the ratio of thermal boundary width to fluid layer depth is $\delta/d \sim Ra^{-1/3}$, which shows that the boundary layers become increasingly thin as convective mixing of the layer becomes more vigorous.

The relation of $\delta \sim Ra^{-1/3}$ applies to the horizontally averaged boundary layer thickness. However, boundary layers change with time or distance from their first formation, e.g., where an upwelling impinges on the top boundary. As the fluid in the boundary layer moves from an upwelling to a downwelling it cools and the boundary layer thickens as more material cools next to the cold surface. The thickening depends on the thermal diffusivity κ (with units of m^2/s) and the residence time or age t near the cold boundary (i.e., time since leaving the upwelling). Simple dimensional considerations show that the boundary layer thickness goes as $\sqrt{\kappa t}$; this corresponds to the well known $\sqrt{\mathrm{age}}$ law for subsidence of ocean sea floor with age since formation at mid-ocean ridges, implying that sea-floor gets deeper because of the cooling and thickening lithosphere.

Patterns of convection, structure of upwellings and downwellings: plumes, and slabs

When convection occurs upwellings and downwellings will be separated horizontally by some optimal distance. If they are too close to each other they can induce too much viscous drag on each other and/or lose heat rapidly to each other; if they are too far apart they must roll too much mass between them. The separation distance is also determined by heat transport in the thermal boundary layer between upwellings and downwelling. When hot upwelling fluid reaches the surface it spreads laterally into the thermal boundary layer. As it travels horizontally it cools to the surface and eventually gets cold and sinks into the downwelling; thus the upwelling-downwelling separation is also determined by the distance it takes for fluid to cool off and become heavy enough to sink.

The upwelling-downwelling separation distance or convection cell size is predicted by convective stability theory to approximately equal to the layer depth d (a bit larger at the onset of convection but identically d as Ra becomes very large); i.e., the cell that is either least stable and thus most likely to convect – or equivalently the cell that optimizes gravitational potential energy (and thus heat) release – is usually the cell that is as wide as it is deep.

Viewing a convecting layer from above, the upwelling and downwellings may be separated in various patterns, such as two-dimensional rolls (sheets of upwelling rolling over into sheets of downwelling, and each cell counter-rotating with its neighboring cell). In the Rayleigh-Bénard layer, which is infinite horizontally, no one location of the layer should be different than any other one so ideally the pattern should be a regular repeating tile; as there are only so many polygons that can form a repeating tile, the patterns usually involve convection cells in the shapes of rolls (already mentioned), squares, hexagons and triangles (Figure 5). Of course non-ideality and irregularities can occur due to tiny imperfections for example in the boundaries leading to irregular patterns.

In many instances the 3-D pattern of convection, especially in fluids where hot fluid is less viscous than cold fluid (as is true in many materials, including the mantle) the upwelling is in the form of a cylindrical plume at the center of a canopy of sheet-like downwellings, much like a fountain.

Plumes and Slabs in the mantle: Simple view

The common occurence of sheet-like downwellings and columnar upwellings in simple convection is crudely applicable to mantle convection. Subducting slabs are where tectonic plates sink into the mantle; these are analagous to the cold sheet-like downwellings seen in 3-D convection, although much more complicated by rheology as discussed below. Deep hot narrow upwelling plumes are infered to explain anomalous intraplate volcanism as in Hawaii, as well as the fixity of these volcanic hotspots relative to each other (which suggests deep anchoring in the mantle). These mantle plumes are ostensibly analogous to the pipe-like upwelling in simple 3-D convection, but again more complicated by unique mantle properties. (See “Mantle Plumes” essay by Farnetani and Hofmann.)

While mid-ocean ridges or spreading centers transport material vertically to surface, they are very narrow features and for the most part involve shallow upwelling (inferred from their weak gravity anomalies that suggest shallow isostatic compensation, i.e., they are floating on a shallow buoyant root). Ridges are likely best explained as being pulled and rifted apart passively from a distant force (ostensibly slabs) rather than involving a deep convective upwelling that pries them open.

Energy sources for mantle convection and Earth’s thermo-chemical history

Heatflow coming out of the Earth is measured by heat-flow gauges (measuring conductivity of rocks first and then thermal gradients in boreholes) both in continents and oceans [see *Turcotte and Schubert*, 1982]. The total heat flowing out from beneath the Earth’s surface is approximately 46TW (46 trillion Watts) [*Jaupart et al.*, 2007], which is in fact not a large number given the surface area of the Earth, and is actually tens of thousands of times smaller than the heat absorbed from the Sun. Nevertheless it represents the source of power driving dynamics inside our planet, including mantle convection (and hence tectonic, volcanological and other geological activity) as well as core cooling and flow.

The source of the Earth’s internal heat is a combination of primordial heat, i.e., left over from accretion (gravitational potential energy from formation and collisions), and heat generated by unstable radioactive isotopes, especially the isotopes of uranium ($^{238}\mathrm{U}$), thorium ($^{232}\mathrm{Th}$) and potassium ($^{40}\mathrm{K}$), although the $^{40}\mathrm{K}$ half-life is much shorter than the others and so generated a large heat pulse primarily in the early Earth. Because continental crust is originally formed by partial melting and chemical segregation of early mantle material (indeed the chemical separation allows another energy source in the release of gravitational potential energy, but other than early core formation this is a relatively minor contribution), these radioactive elements tend to be concentrated in crust (i.e., melt more readily dissolves these elements than does solid rock, so they partition toward the melt). Thus the crust itself produces a significant fraction of the net heat output through the surface; removing the crustal component leaves approximately 31TW emanating from the mantle and core [*Schubert et al.*, 2001; *Jaupart et al.*, 2007].

The relative contributions of primordial cooling and radiogenic heating to the mantle (and

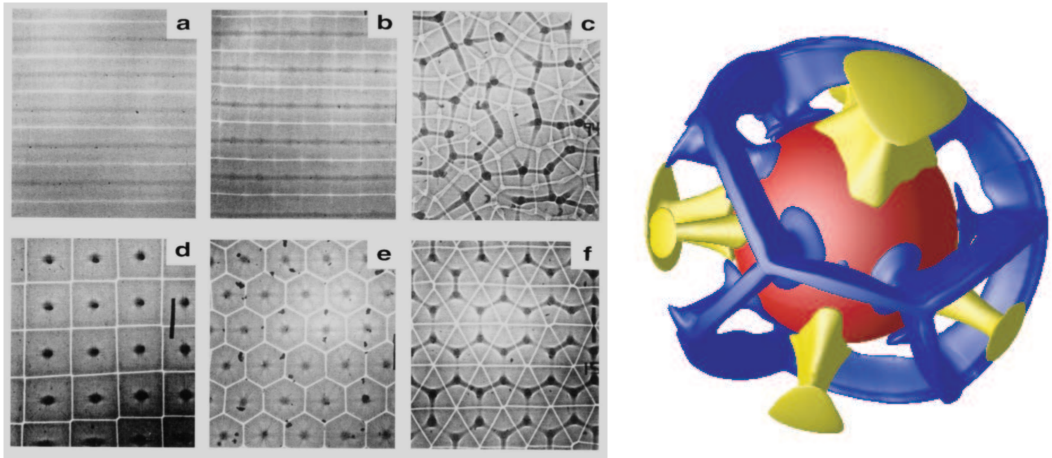

Figure 5: Patterns of convection from laboratory experiments in rectangular tank by *White* [1988] (left), and numerical simulations in a spherical shell by *Zhong et al.* [2000].

core) heat output remains an active area of debate even today and leads to various quandaries. The most direct estimate of the concentraion of radiogenic sources (U, Th, K) is by looking at the concentration in chondritic meteorites, which come for the solar systems' main asteroid belt and because they have been largely unaltered for 4.5Gyrs (i.e., unmelted) are thought to be the same as the original building blocks of the terrestrial planets. The chondritic concentrations of U, Th and K would allow for radiogenic heating to contribute 50% or less of the total heat output [*Korenaga*, 2008]; this is often called the Urey ratio, i.e., the radiogenic heat output to the total output. With radiogenic heating this small, the only way the mantle could be as hot as it is today – while also transporting heat as it does presently – is if it were very hot and nearly molten in the geologically recent past (a few hundred million years); this is geologically untenable since petrological and geochemical analysis demonstrate the presence of solid rock and even liquid water in the early Earth [see *Halliday*, 2001], which therefore ceased to be molten not long after it's formation 4.5 billions years ago. This paradox has led some researchers to assume that the radiogenic sources are super-chondritic, i.e., to allow heat to be produced continously throughout the past rather than by rapid cooling from a molten state. Alternatively, researchers have sought ways to keep chondritic concentrations of U, Th and K by arguing that heat-transport in the past was different than it is today; for example higher temperatures in the past might have allowed for more melting and thus more buoyant crust and/or stronger dehydrated lithosphere that kept the top part of the convecting mantle sluggish or immobile hence bottling up primordial heat for much later release [*Korenaga*, 2008].

Finally, the release of both radiogenic and primordial heat, again termed collectively "internal heating" (i.e., heat coming from the bulk of the fluid) means that the idealized Rayleigh-Bénard model of convection, where heating is only along the base, is inaccurate. The effect of internal heating in addition to "basal heating" is relatively straightforward to understand. While in the Rayleigh-Bénard model the temperature of the well mixed interior is at the average of the two boundary temperatures, the addition of internal heating acts to heat up the well mixed interior to a higher mean temperature. This puts the interior temperature closer to the hot bottom temperature, but further from the cold top temperature; the effect is to create a larger temperature drop across the top thermal boundary layer than across the bottom one; in essence, the top boundary layer must conduct out heat injected through the bottom plus heat generated or released from the interior (Fig 4). These very different boundary layers tend to cause more negatively buoyant and stronger downwelling currents and smaller and weaker upwelling currents; this effect seems to be borne out in the Earth by the presence of many large cold slabs driving plate tectonic motion with large thermal anomalies (of order 500K) relative to fewer upwelling plumes with weaker (200K) thermal anomalies (although this is still somewhat a matter of debate).

Effects of mantle properties

Mantle rheology

The entire mantle of the Earth is potentially a convecting and over-turning fluid. However, the mantle is almost entirely solid (with some small portions of melting near the surface delineated by volcanism, and possibly much smaller areas of melting at depth) and thus flows extremely slowly. (Indeed, the term fluid does not suggest a liquid state but refers to how a medium deforms, as opposed to elastic or brittle deformation; these do not necessarily correlate with states of matter, i.e., gas, liquid, solid, each of which can display, for example, either fluid flow or elastic behavior depending on the nature of the deformation; e.g., atmospheric sound waves are elastic behavior in gas.)

The solid-state viscous flow of the mantle is a complex process and there are various mechanisms that permit such ”irrecoverable” deformation (one of the definitions of viscous flow). A survey of mantle deformation mechanisms or ”rheology” is beyond the scope of this essay [see *Ranalli*, 1995; *Karato*, 2008]. However, in brief the two primary deformation mechanisms are called diffusion and dislocation creep. In any solid-state creep mechanism, mobility depends on the statistical-mechanical probability of a molecule in a crystal lattice leaving the potential well of its lattice site; the potential-well itself is defined by electrostatic or chemical bonds inhibiting escape, and Pauli-exclusion pressure preventing molecules squeezing too closely to each other. Thus mobility depends on the Boltzman distribution measuring the probabilty of having sufficient energy to overcome the lattice potential well barrier, which is often called the activation energy (or allowing instead for pressure variations the activation enthalpy). This probability depends on the (Arrhenius) factor $e^{-E_a/RT}$ where E_a is the activation energy (J/mol), R is the gas constant (J/K/mol) and T is temperature; RT represents the thermal excitation energy of the molecule in the well. As T goes to infinity the probability of escaping the well goes to 1, while as T goes to 0 the probability of escape goes to 0.

Stress imposed on the medium effectively changes the shape of the potential well, such that compressive stress steepens the walls of the well (squeezing molecules closer makes coloumb attraction in the chemical bonds stronger) while tension lowers the walls (separating molecules weakens the bonds); thus the probability of escape is preferred in the direction of tension and away from compression, thus allowing the medium to stretch in the tensile direction by solid-state diffusion of molecules.

Simple diffusive creep works much in this way where differential stress (i.e., non-uniform stress) causes slow diffusion of molecules to allow the entire substance to deform accordingly. However, such deformation occurs at the grain or mineral level inside a rock, with either diffusion through the grains or along the grain boundaries; thus the response depends significantly on grainsize.

Dislocation creep is more complicated. A dislocation can be in the form of a truncated row of molecules in a crystal lattice; shortening of the crystal under compression perpendicular to the row can be accomplished by simply removing that row. Thus differential compressive stress would act to force molecules to diffuse out of that dislocated row into other parts of the lattice. However, stress not only governs the preferential diffusion of molecules (as in diffusion creep) but also the geometry of the dislocations (their spacing and directions), hence the multiple actions of stress are compounded into a nonlinear response.

The viscosities for diffusion and dislocation creep mechanisms can be written as

$$\mu = \begin{cases} Ba^m e^{\frac{E_a}{RT}} & \text{for diffusion creep} \\ A\sigma^{1-n} e^{\frac{E'_a}{RT}} & \text{for dislocation creep} \end{cases} \quad (2)$$

where A and B are proportionality constants, a is grainsize, σ is stress (in fact since stress is a tensor σ is the scalar second invariant of the stress tensor), and m and n are exponents, typically both equal to 3. It should be emphasized that diffusion and dislocation creep occur independently of each other depending on stress and grainsize: for high stress and large grains dislocation creep dominates; for low stress and small grains diffusion creep dominates.

Dislocation creep allows for moderate softening as stress increases; diffusion creep potentially allows for significant softening if stress can reduce grainsize, although mechanisms to allow this are controversial still (see section below on generating plates), and significant hardening via grain growth by standard coarsening of the material (i.e., what happens to all grained materials under the action of grain-surface energy reduction).

The strongest rheological effect is clearly that of temperature; the temperature-dependence of viscosity allows for many orders of magnitude variations in viscosity. For example, while this rheological effect allows subducting slabs to keep their strength and integrity to great depths as they sink, it would make hot upwelling mantle plumes more fluid and, if they have a conduit structure, the plume flow would be relatively rapid, of order 100cm/yr or more. However, this effect is most profound in the cold top thermal boundary layer or lithosphere. If viscosity is strongly temperature dependent, as it is in Earth, the lithosphere can become so stiff that it becomes immobile; in this case convection in the mantle would proceed beneath the lithosphere, which in turn would act like a rigid lid to the mantle [*Solomatov*, 1995]. If mantle rheology obeyed only diffusion or dislocation creep laws, then the lithosphere should be locked and immobile and there should be no plate tectonics. While this scenario might be relevant for Venus and Mars (and Moon and Mercury) which have no plate tectonics, obviously it is missing a vital ingredient to allow plate tectonics on Earth. This paradox underlies the fundamental question and mystery about why Earth has plate tectonics at all and how it is generated on our planet but not others in our solar system [*Bercovici*, 2003].

Compressibility, Melting and Solid phase changes

Pressures deep inside the Earth's mantle are so large they are sizable fractions of rock incompressibility or bulk modulus (e.g., mantle pressures reach 140GPa, or 1.4 million atmospheres, while bulk modulus – which has the same units as pressure – are typically a few to several 100GPa). As downwelling mantle material travels from near the surface to the base of the mantle its density and temperature increase due to compression, called ”adiabatic compression and heating” (and likewise upwelling material undergoes ”adiabatic decompression and cooling”), although these increases are not large (of order several degress celsius). The compression and decompression of circulating material establishes a weak adiabatic temperature and density increase with depth, which has a slight stabilizing effect on convection; however, because the mantle is so viscous the thermal anomalies needed to get it to move – and in particular the temperature variations across the thermal boundary layers – are so large (of order several 100 to 1000 degrees celsius) that the adiabatic variations are small in comparison.

Where compressibility and pressure play a dual important role is in phase changes. First, as hot upwelling mantle material approaches the surface it actually travels along a gradually cooling adiabatic temperature profile. The upwelling does eventually melt when it gets near the surface but not because it gets hotter. Melting occurs because the melting temperature T_m drops with decreasing pressure faster than the upwelling adiabat (in essence, decreasing confining pressure makes it easier for molecules to mobilize into a melt); thus at a certain (usually shallow depth of a few 10s of km to 100km) the upwelling mantle crosses the melting temperature from solid to liquid phase and undergoes melting; however, the mantle is not a single pure substance so in fact only partially melts. Such "pressure release" melting is a shallow process but is vital for chemical segregation of the mantle and development of oceanic and continental crust. In particular, melting is sequential in that the most easily melted material (usually more silica rich material with lower melting temperature) melts first, freezes last and is typically chemically less dense, and thus comes to the surface as lighter crust eventually gathering, after more weathering and reactions into continental crust. (Continental crustal rocks like sandstone and granite have typical densities of $2300\mathrm{kg/m^3}$ and $2700\mathrm{kg/m^3}$, respectively.) The more refractory (harder to melt, silica poor and heavier material) melts last, freezes first and either stays in the mantle or lithosphere or sits in the heavy basaltic oceanic crust. (Oceanic crustal rocks like basalt have densities of $3000\mathrm{kg/m^3}$, while mantle peridotites at near-surface pressure have densities of around $3400\mathrm{kg/m^3}$.)

Extreme pressures with depth can also overcome a mineral's elastic resistance to compression and cause solid-solid phase changes where the minerals change their crystalographic structure to a more compact and incompressible state (but of course their chemistry remains the same). Such mineralogical phase changes have been observed in laboratory experiments in olivine, which is the major component mineral of the upper mantle (at about 60% by weight, the remainder being mostly pyroxene at shallow depths, and garnet at slightly greater depth); moreover, the pressures at which they are predicted to occur have been verified seismologically, wherein the seismic wave speeds and density undergo a jump at the predicted pressures. The first major phase change to occur with depth is from olivine to the same material with a wadsleyite structure, at 410km depth. Wadsleyite changes slightly to a similar ringwoodite structure at 510km depth. The largest phase change occurs from ringwoodite to perovskite/magnesiowüstite at 660km depth.

The 410km and 660km phase changes are the two most remarkable and global phase changes in the mantle, and the region between them is called the Transition Zone, since it is where most of the mineralological transitions occur, over a relatively narrow region. The mantle above the Transition Zone is typically identified as the Upper Mantle, although in some papers and books Upper Mantle includes the Transition Zone. Below the Transition Zone is the Lower Mantle and that is universally agreed upon in the literature.

The Transition Zone has anomalous properties due to mixing and transitions in mineral organization; for example it is thought to be able to absorb an order of magnitude more water (per kg) than the mantle both above and below it (although this issue is still somewhat controversial).

Other phase changes are thought to occur with depth, although these are less well resolved, and in some instances do not appear to be global. Recently a new phase change has been inferred in the lowest part of the mantle (the bottom few 100 kilometers), called the perovskite-post-perovskite (or just the post-perovskite) transition [*Murakami et al.*, 2004]. This transition is still an active area of exploration.

The effect of phase transitions on mantle convection has been an active area of research since the 1960s. The discovery of the major phase change at 660km depth coincided with the fact that deep earthquakes along subducting slabs (the Wadati-Benioff Zone) go no deeper than 700km. This seemed to imply that subducting slabs, the mantle's equivalent of cold convective downwellings (see below), did not extend into the lower mantle. Recent seismological studies using tomographic techniques to resolve ostensibly ”cold” and ”hot” areas of the mantle (really seismically fast and slow regions), implied that many slabs might stall or pool temporarily at the 660km boundary but many penetrate into the lower mantle (Fig 6).

Whether the density jump due to a phase change could impede vertical flow has been a key question in studying the interaction of phase changes and convection. In particular, the 660km phase change was inferred mineralogically to be ”endothermic” whereby the entropy of the deeper heavier phase (below 660km) increases, or more simply latent heat is absorbed on going down through the boundary into the denser more compact phase (while unusual in most systems, this is also true of the solid-liquid transition in water). This also means that the dependence of the transition temperature on pressure has a negative slope; thus cold material impinging on the phase boundary causes the phase change to deflect in the cold region to greater pressures; this induces a depression in the boundary that acts to buoyantly rebound upward and oppose the motion of the descending cold material. Thus the endothermic boundary at 660km depth possibly impedes flow across that boundary. (However, the actual 660km phase is complex and at higher temperatures might become exothermic; *Weidner and Wang* [1998].) Computational studies and simulations of mantle flow across this boundary demonstrated that downwellings (i.e., slabs) can indeed be impeded and stalled as they impinge on this boundary, but not permanently or globally; i.e., while some are pooling at the boundary others have gathered up enough “weight” to push through the boundary [*Tackley et al.*, 1993; *Christensen*, 1995]. This picture appears to be in keeping with the picture from seismology that while the phase boundary impedes slab and downwelling flow into the mantle, it is not an impermeable boundary and there is in the end significant exchange between the upper and lower mantle and hence whole mantle convection.

However, while the mineralogical, seismological and geodynamical (fluid mechanical) arguments imply that there is flow between upper and lower mantles, there are numerous data from geochemical analysis of basaltic lavas implying that the mantle is not well stirred on a large scale, i.e., it is possibly layered with poor or non-existent communication between upper and lower mantle.

Structure of mantle convection and mantle mixing

Upwelling mantle reaching the Earth's surface undergoes melting (see above under pressure-release melting) and this melt reaches the surface in two types of volcanic settings: mid-ocean ridges where tectonic plates spread apart and draw mantle up into the opening gap, and ocean-islands or hotspots which are anomalously productive and localized volcanic features not necessarily associated with tectonic activity, Hawaii being the most conspicuous such feature. Melts coming from the mantle in this way are silica poor (relative to more silicic rocks such as granite) and largely basaltic; hence these volcanic regions are said to produce Mid-Ocean Ridge Basalts (MORB) and Ocean-Island Basalts (OIB), respectively. These melts are in effect messengers from the mantle, and their petrological composition, bulk chemistry and trace-element chemistry are extensive areas of research [*Hofmann*, 1997, 2003; *van Keken et al.*, 2002; *Tackley*, 2007]. In the end, while these two basalts nominally come from

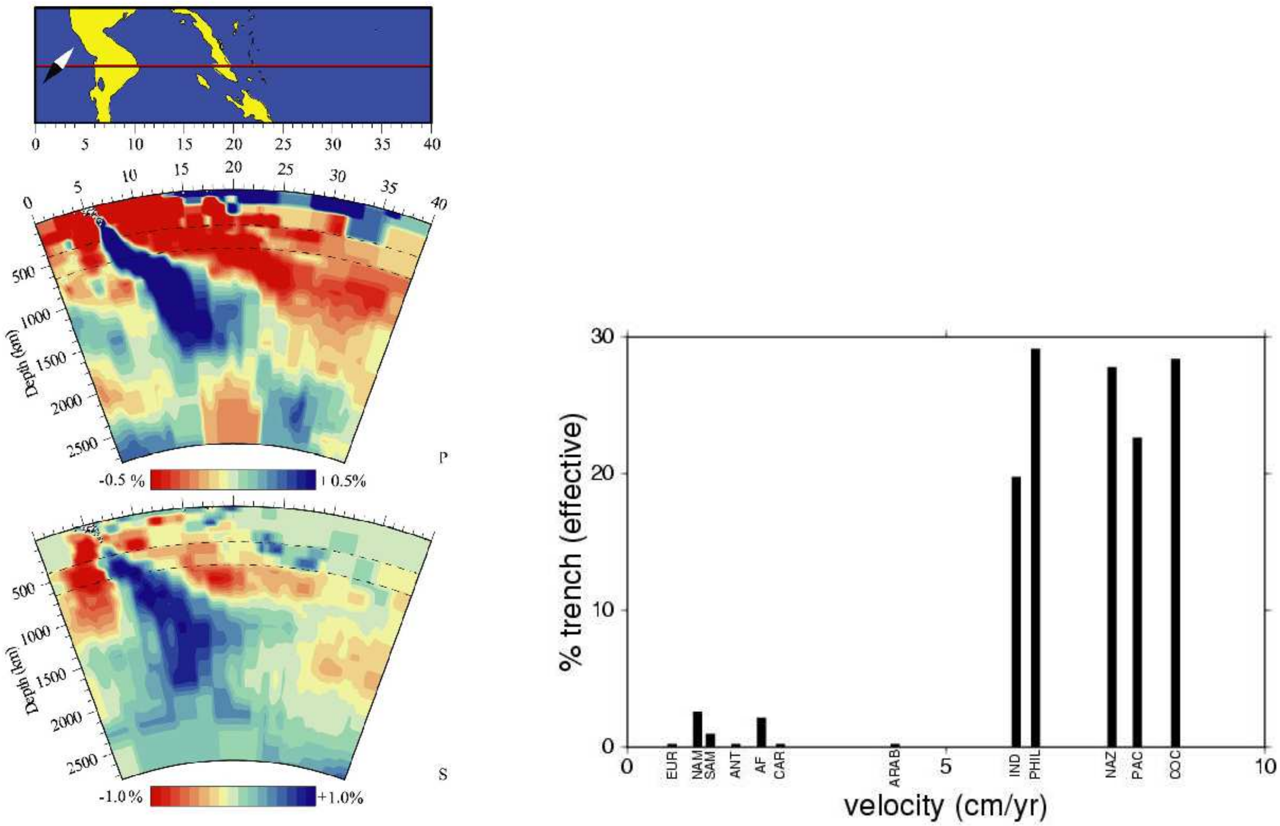

Figure 6: Tomographic image of slabs beneath Mexico (W. Spakman via *van Keken et al.* [2002]) extending into the lower mantle (left); and plate velocity vs trench length following *Forsyth and Uyeda* [1975] showing the fastest plates are connected to slabs (right). Slabs are not only cold mantle downwellings but effectiveley drive plate tectonics

the same mantle, they have distinct features suggesting they come from parts of the mantle that have been isolated from each other for billions of years. That seismology, mineral-physics and fluid-dynamics (geodynamics) argue for a largely well-stirred mantle with whole-layer circulation creates a dichotomy between geophysical and geochemical observations. This paradox has been one of the most fervent areas of debate in mantle dynamics for the last 30-40 years.

Because MORB and OIB are both basalts and thus have similar bulk chemistry, geochemical measurements largely focus on trace elements, in particular incompatible elements, which dissolve more readily in a rock's melt phase than its solid phase; hence during partial melting incompatible elements partition toward the melt. Indeed, the trace-element signature of elements such as uranium, thorium, helium, have demonstrated that MORB and OIB are very distinct. In particular, MORBs appear to be significantly depleted in such trace elements relative to OIB. Since such elements tend to be removed by melting, it implies that MORBs come from a region of the mantle that has already been melted and depleted of trace elements, while OIB come from a region of the mantle that has undergone little previous melting and depletion. This observation implies that MORBs come from an upper mantle that has been cycled repeatedly through the plate tectonic process of mid-ocean ridge melting and separation of crust (and trace elements) from mantle, in essence cleaning the MORB source. In contrast, OIB would appear to come from a part of the mantle that has seen little of this melt processing, and hence would be isolated presumably at depth from the upper mantle and the plate-tectonic circulation.

There are other geochemical observations that argue for separated and isolated regions or reservoirs in the mantle. For example the concentration of radioactive daughter isotopes (e.g., $^{206}\mathrm{Pb}$, which is the final product of the decay of $^{238}\mathrm{U}$) relative to the abundance of that element's dominant and primordial isotope (e.g., $^{204}\mathrm{Pb}$) is a metric for reservoir isolation from surface processing. In particular the relative accumulation of daughter products implies that the rock in which they reside has seen little processing or partial melting that would have cleaned out these elements after they formed. A small relative abundance of daughter isotopes means the sample has been recently processed and cleaned, and thus little time has passed in which to produce new daughter isotopes.

Indeed, many OIBs tend to show distinctly greater relative abundance of daughter products (e.g., the concentration ratio $^{206}\mathrm{Pb}/^{204}\mathrm{Pb}$) than do MORB for many isotopic ratios, implying some isolation of the OIB source. However, OIBs from various islands are also fairly different from each other suggesting that reservoirs isolated from the upper mantle (the presumed MORB source) are possibly also isolated from each other. Moreover, the OIB isotopic ratios have some variation, from low MORB-like values to much higher values, which indicates that there is some mixing of a "young" processed MORB-source like mantle and a more primitive, isolated one.

These isotopic ratios are also more easily interpreted for refractory daughter products since they do not tend to escape the system (also the reason they are used to radiometrically date rocks). Volatile products, especially isotopes of noble gases such as helium and argon, require different intepretations since they can readily escape the mantle and, for helium, escape the Earth. For example, MORB in fact has a high daughter to primordial isotope ratio $^{4}\mathrm{He}/^{3}\mathrm{He}$ relative to many OIBs, which is opposite to the refractory ratios involving, for example, lead isotopes. This is often interpreted as resulting from degassing and loss of primordial helium $^{3}\mathrm{He}$ from the upper mantle through plate tectonic and mid-ocean ridge processing, and the subsequent repopulating of helium with its radiogenic isotope $^{4}\mathrm{He}$ (i.e., α-particles from most large element decay sequences); in contrast an isolated lower mantle or OIB source reservoir would have undergone little loss of primordial helium thus maintaining a smaller isotopic ratio $^{4}\mathrm{He}/^{3}\mathrm{He}$.

The production of the argon isotope $^{40}\mathrm{Ar}$ from the decay of the potassium isotope $^{40}\mathrm{K}$ has two important arguments relative to mantle layering. First, the total amount of original $^{40}\mathrm{K}$ in the Earth can be roughly estimated from chondritic abundances (and other arguments beyond the scope of this essay). However, the amount of $^{40}\mathrm{Ar}$ it should have produced over the age of the Earth is far in excess (by a factor of 2) of the $^{40}\mathrm{Ar}$ in the atmosphere, implying that much of this argon is still buried and isolated. Moreover, one can also estimate from the trace element composition of MORBs themselves that the MORB source region as it stands now would have been lacking primordial $^{40}\mathrm{K}$ and even if the entire mantle were composed of this MORB source region, it would not have been able to produce even the atmospheric levels of $^{40}\mathrm{Ar}$; this implies that the bulk of original $^{40}\mathrm{K}$ was buried in a layer different and more enriched than the MORB-source region, which then produced most of the $^{40}\mathrm{Ar}$, much of which is still buried in this layer.

An often quoted straw-man argument for mantle layering is the heatflow paradox. In this case, it is reasoned that if the entire mantle were composed of MORB source material with its depleted concentration of heat producing elements (U, Th, K), then it would not be able to produce the total heat output through the top of the mantle (about 30TW). This suggests that the the heat producing elements allowing for the mantle heat output must be buried at depth. However, this makes the assumption that most heat output is from radiogenic heating, whereas possibly less than half of it is; if most heat output is from secular cooling (lost of primordial heat) then the heat-flow paradox argument is questionable. A similar but shakier argument is based on the fact that the volcanic flux of the helium isotope $^4\mathrm{He}$, which is produced from heat producing radioactive decay of U and Th, seems to be very low relative to what would be expected given the heat that is emanating from the mantle, implying that $^4\mathrm{He}$ is also buried at depth. Again, if heat ouput is more than 50% from secular cooling and not all from radiogenic heating, then the low flux of helium is to be expected. Even if that were not the case, mechanisms for the flux of helium are not the same as for flux of heat; i.e., heat always escapes to space eventually by convection, conduction even radiation, whereas helium only escapes from the mantle if it passes through narrow melting zones, which is not inevitable, and thus it can be hidden and buried almost anywhere and not only hidden at great depth.

Finally, the production of continental crust also argues for an isolated layer in the mantle. Continental crust represents an accumulated history of mantle melting, segregation of lighter components and removal of incompatible elements to the surface. If the continental crust were removed uniformly from a whole mantle made of primordial ”bulk silicate earth” (i.e., an Earth derived from chondritic material and only segregated into mantle and core), then the concentration of incompatible elements would not have been reduced enough to produce a mantle made of MORB source material (i.e., removal of continental crust would not have depleted the whole mantle enough to make MORB source). However, if the crust were removed from 1/3 to 1/2 of the mantle, that resulting depleted portion would very closesly match MORB source composition. This argues that the continental crust segregated only from the upper portion of the mantle, not the whole mantle, and thus there remains a deeper unsegregated mantle at depth [see *van Keken et al.*, 2002].

Although there are numerous geochemical arguments for a layered mantle with an isolated and undepleted mantle at depth, they largely conflict with geophysical evidence for whole mantle convection. Mineral physics experiments suggest that the 660km phase change boundary might provide an impediment to mantle fow but not an impermeable barrier. Seismic tomography consistently shows subducting slabs and apparently cold downwellings extending into the lower mantle [*van der Hilst et al.*, 1997; *Grand et al.*, 1997]. Recent high resolution images of a hot upwelling mantle plume beneath Hawaii [*Wolfe et al.*, 2009] as well as seismic images of other plumes [*Montelli et al.*, 2004] also suggest vertical upward transport across the 660km boundary. Finally, geodynamical arguments against separated layers suggest that if a lower mantle held most of the mantles heat producing elements, it would be implausibly hot, and by heating up the bottom of the upper mantle it would generate much bigger mantle plumes than would be observed [*Tackley*, 2002].

The contradiction between geochemical and geophysical inference of layered vs whole mantle convection has been and largely remains an unsolved problem. Attempts to reconcile these observations have been numerous. A reasonably popular approach has been to allow that the 660km boundary is not a barrier to mantle flow, but that the barrier exists at greater depth. There are seismically observable layers at the bottom of the mantle (the D” layer), which could store enriched material, although these are also so thin they could possibly overheat (depending on the amount of radioactive heat sources stored there). As a comprosmise, it has been argued that the enriched mantle exists in an approximately 1000km thick layer at the base of the mantle [*Kellogg et al.*, 1999], although this layer has never been seismologically observed (see Fig

7). More recently, chemical heterogeneity gathered into piles on the core-mantle boundary and below upwelling zones [*Jellinek and Manga*, 2002, 2004] has been suggested by convection models [*McNamara and Zhong*, 2005] with support from joint seismology-gravity analyses [e.g., *Ishii and Tromp*, 1999].

Other mechanisms for reconciling geochemical and geophysical observations, but not invoking layering, have been recently proposed as well. These mostly involving differential melting. For example, one model considers the whole mantle as a plum-pudding mix of enriched and volatile (water) enriched plums in a depleted, drier and harder to melt pudding. The pressure-release melting in mantle plumes is stopped at higher pressures at the base of 100km thick lithosphere, so this could involve mostly melting of easily melted enriched and volatile-rich components; melts making it all the way to the surface at ridges would undergo more pressure drop and thus could also melt the depleted mantle component, resulting in MORB that seems depleted relative to OIB [*Ito and Mahoney*, 2005a,b]. Another model exploits the fact that Transition Zone minerals seem to be able to absorb water more readily than material above it (and below it). A Transition-Zone with a little water will be dry relative to its solubility or water storage capacity. However upwelling material passing through the transition zone would carry this slightly damp material into the upper mantle at the 410km boundary, and since the upper mantle olivine has poor water solubility it would be closer to saturation and likely melt. A little melting at 410km depth of the broad upwelling mantle (forced upward by the downward flux of slabs) passing through and out of the transition zone would cause it to be stripped of incompatible elements as it flows into the upper mantle leaving a depleted MORB source region; because of the high pressures and high compressibility of melt, the partial melt that has cleaned this upwelling mantle would be dense and remain behind, eventually to be entrained by slab driven downwelling back into the lower mantle. Upwelling mantle plumes on the other hand would go through the transition-zone too fast to become hydrated and thus would undergo little melting and filtering at 410km depths, leaving largely enriched OIB. This model, called the Transition-Zone Water Filter [*Bercovici and Karato*, 2003a; *Karato et al.*, 2006; *Leahy and Bercovici*, 2010] (see Fig. 7), predicts that the 410km should be the site of melting, and this has been born out in various seismological studies [e.g., *Revenaugh and Sipkin*, 1994; *Song et al.*, 2004; *Tauzin et al.*, 2010]; however, the theory is still controversial given poor knowledge of melt properties and their solubilities of incompatible elements at these depths and pressures, so it remains an active subject of investigation.

Mantle convection and the generation of plate tectonics

The oldest problem in mantle convection

The link between plate tectonics and mantle convection is one of the oldest and most challenging problems in the history of geodynamics. The original theories of mantle convection put forward by *Holmes* [1931] were developed in the context of explaining continental drift as articulated by *Wegener* [1924]. Although the later theory of plate tectonics is the grand-unifying principle of geology, it is a kinematic theory in that it describes surface motions but not their cause. Mantle convection is widely accepted to be the engine for plate motions since it is a fundamental mechanism for exploiting the energy sources of the Earth's interior, i.e., loss of primordial and radiogenic heat. It is now generally regarded that the plates themselves are a feature of mantle convection in that they are the mobile upper thermal boundary layer of convective cells that thicken as they cool in their migration away from ridges until they are heavy enough to sink along subduction zones.

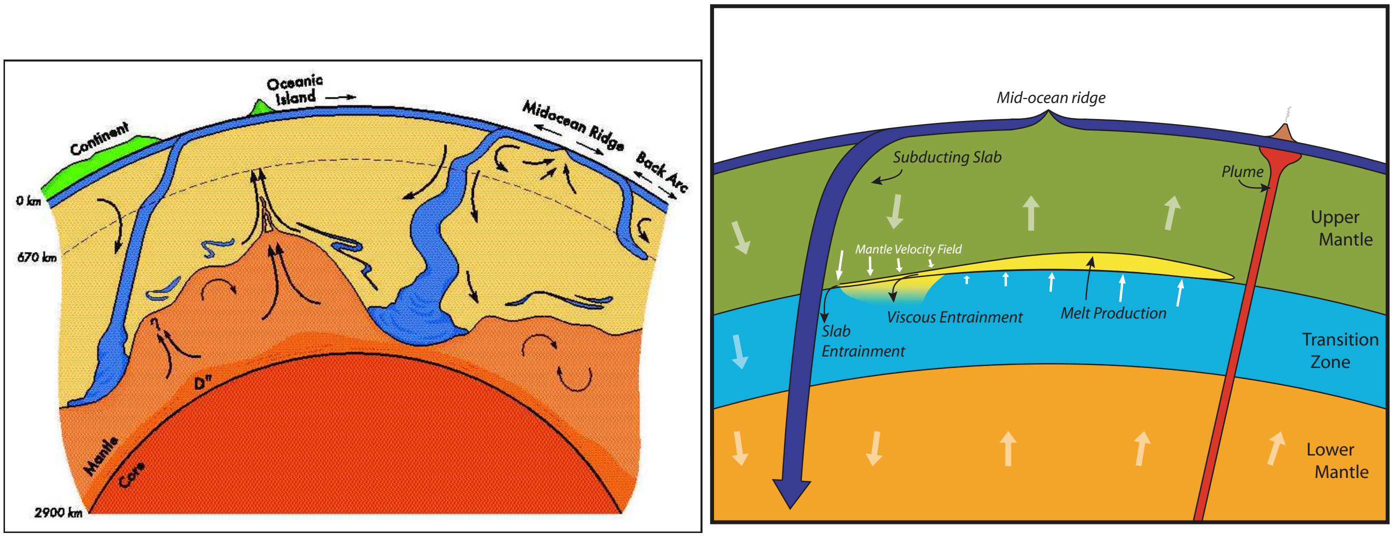

Figure 7: Two end-member mantle mixing models. In the abyssal layered model (left) the source for enriched ocean-island basalt (OIB) is in a deep primordial layer, while the source for depleted mid-ocean ridge basalt (MORB) is in the upper recycled mantle (after *Kellogg et al.* [1999]). In differential melting models, such as the water-filter model (right), the MORB and OIB sources undergo different styles of melting, but the mantle still undergoes whole-mantle circulation; see text for further explanation (after *Bercovici and Karato* [2003a]; *Leahy and Bercovici* [2007]).

One of the major accomplishments of mantle dynamics theory is that convective fluid velocities, calculated in any number of ways, consistently predict the measured scales of plate tectonic velocities, i.e., between 1 and 10cm/yr. This was in fact inferred even in the 1930s from both gravity and heat-flow measurements [*Pekeris*, 1935; *Hales*, 1936] and is well known by the force balance on sinking slabs [e.g., *Davies and Richards*, 1992], as well global heat extraction from mid-mantle cooling by slabs [*Bercovici*, 2003]. Moreover, plate motions are well correlated with the presence of slab forcing, in particular that tectonic plates with a significant portion of subduction all have velocities roughly an order of magnitude faster than plates without substantial subduction zone [*Forsyth and Uyeda*, 1975]; see Figure 6. Thus, because cold sinking slabs seem to be the major expression of convection and major drivers of plate motions implies that the plates are convection.

The plate generation problem

While convective and slab driving forces for plate tectonics are important, understanding how plates self-consistently arise (or why they do or don't arise) from planetary convection has been a major goal in geodynamics. Up until the early 1990s it was believed that since plate-like motion of the lithosphere was essentially discontinuous it could not be predicted or reproduced by fluid dynamical convection theories. However, in the last 15 years or so there has been major progress with fluid dynamical models yielding very plate like motion by incorporating more sophisticated rheological weakening mechanisms, such as brittle/plastic yielding or damage mechanics.

However, even so, there is still no comprehensive theory of how the plates and mantle are related, and in particular how plate tectonics self-consistently arises from a convecting mantle. Both a clue and frustration is that the Earth appears to be the only known terrestrial planet that has plate tectonics in addition to liquid water as well as life, which are all either causative (i.e., necessary conditions for each other) or coincidental. While our planet supports plate tectonics, our ostensible twin, Venus, does not; this remains a leading-order quandary in Earth sciences and to solve it one must understand how and why plate tectonics is generated at all. Of course geoscience has until recently only sampled the few planets of our own solar system and thus the data is sparse; with the advent of extra-solar planet discovery, this sparseness should be mitigated and perhaps other planets will be discovered with plate like mantle circulation from which we will learn more about how our own planet works [e.g., *Valencia et al.*, 2007].

There are of course various qualitative and philosophical questions about how and why plate tectonics forms, evolves and exists, but for these to be predicted or reproduced in a physical theory, one requires quantifiable questions. In short, what are the metrics of plate generation? Two fundamental features of plate-like motion exist for instantaneous motions: these are *plateness* [*Weinstein and Olson*, 1992] and toroidal motion [e.g., *Hager and O'Connell*, 1979; *O'Connell et al.*, 1991; *Bercovici et al.*, 2000]. Plateness is essentially the extent to which surface motion and the strength of the lithosphere is like that for nearly rigid blocks separated by narrow weak boundaries. Toroidal flow is characterized by strike-slip motion and plate spin (Figure 8). Toroidal motion has no direct driving force in buoyant convection (which drives only vertical and divergent motion), however, it has as much energy in the present-day plate tectonic velocity field as the buoyantly driven motion (called poloidal motion). The globally averaged toroidal motion is dependent on the lithosphere's reference frame (e.g., hostpot frame), and the field in general changes through time [*Cadek and Ricard*, 1992; *Lithgow-Bertelloni et al.*, 1993]; however, the toroidal field is a quantifiable and significant feature of global plate motions. Such measurable quantities as plateness and toroidal flow are important for testing the predictions of plate generation theories. Both phenomena rely on reasonably strong nonlinear rheological feedback effects to permit large strength variations for high plateness (i.e., rapidly deforming zones are weak, slowly deforming ones are strong), as well as coupling of buoyantly driven flow to vertical torques that drive toroidal spin and shear [*Bercovici*, 2003].

Recent progress

In the last decade, plate generation models have become increasingly sophisticated, in concert with further expansion and accessibility of high-performance computing. Instantaneous plate-like behavior has been achieved with convection models employing various forms of plastic yield and self-weakening criteria. Incorporation of these laws has lead to the prediction of reasonable toroidal and poloidal flow [*Bercovici*, 1995] (Figure 9) and by including the rheological effects of melting at ridges have attained localized passive spreading zones [*Tackley*, 2000] (Figure 10). Most recently these models have been extended to three-dimensional spherical models, creating the first global models of plate generation from mantle convection [*van Heck and Tackley*, 2008; *Foley and Becker*, 2009] (Figure 11).

However, models that use plastic or instantaneous self-weakening rheologies only allow weak zones to exist while being deformed, thus cannot correctly model dormant weak zones (e.g., sutures and inactive fracture zones). Rheological mechanisms that allow weakening to persist over time have also been studied. While thermal weakening is a well understood mechanism, thermal anomalies diffuse too fast, and it is highly unlikely that, for example, sutures are thermal remnants. Weakening by hydration or as a secondary phase (i.e, by providing pore pressure) is a strong candidate as well, although the mechanism for ingesting water to depth is problematic, and requires mechanisms such as cracking enhanced by thermal stresses [*Korenaga*, 2007]. Damage in the form of mi-

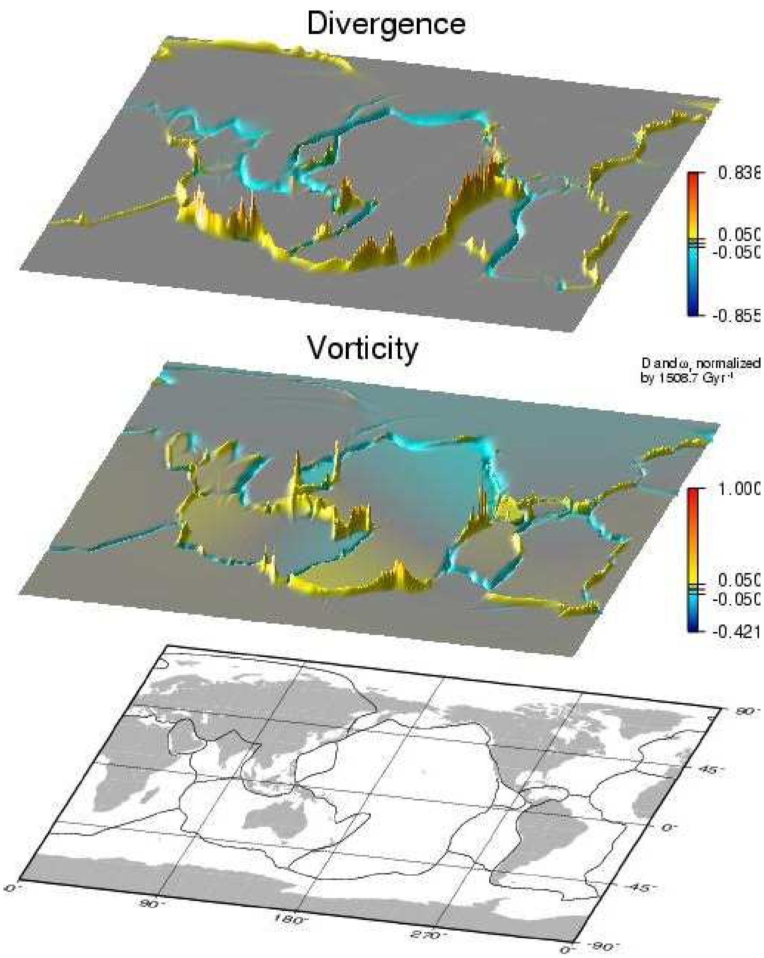

Figure 8: The horizontal divergence (which measures spreading and convergence rates) and vertical vorticity (which measures angular velocity and horizontal shear) of present day plate motions [from *Dumoulin et al.*, 1998], which are associated with poloidal and toroidal flow, respectively. Poloidal flow is equivalent to basic convective motion driven directly by thermal buoyancy. Although present in Earth's plate-mantle system, toroidal flow does not arise naturally in basic viscous convection, but requires the coupling of convective motion with nonlinear rheological effects.

crocracks [*Bercovici*, 1998] and grain-size reduction [e.g., *Karato et al.*, 1980; *Braun et al.*, 1999; *Montési and Hirth*, 2003] are also strong candidate weakening mechanisms because of their longevity and evidence in the form of mylonites [*Jin et al.*, 1998]. The need for shear-localization at significant depth makes grain-size weakening particularly appealing and has proven to be successful at creating plate-like mantle flows [*Bercovici and Ricard*, 2005; *Landuyt et al.*, 2008; *Landuyt and Bercovici*, 2009b] (Figure 12). However, grainsize reduction and weakening tend to occur in exclusive areas of deformation space (i.e., weakening occurs during diffusion creep while reduction occurs in dislocation creep; see *De Bresser et al.* [2001]) and require mixing of mechanisms in physical or grainsize distribution space [*Bercovici and Karato*, 2003b; *Ricard and Bercovici*, 2009].

Use of both plastic/brittle-yielding and damage theories of plate generation have been used to elucidate the planetary dichotomy between Earth and Venus and the causal link between climate, liquid water and plate generation. Earth

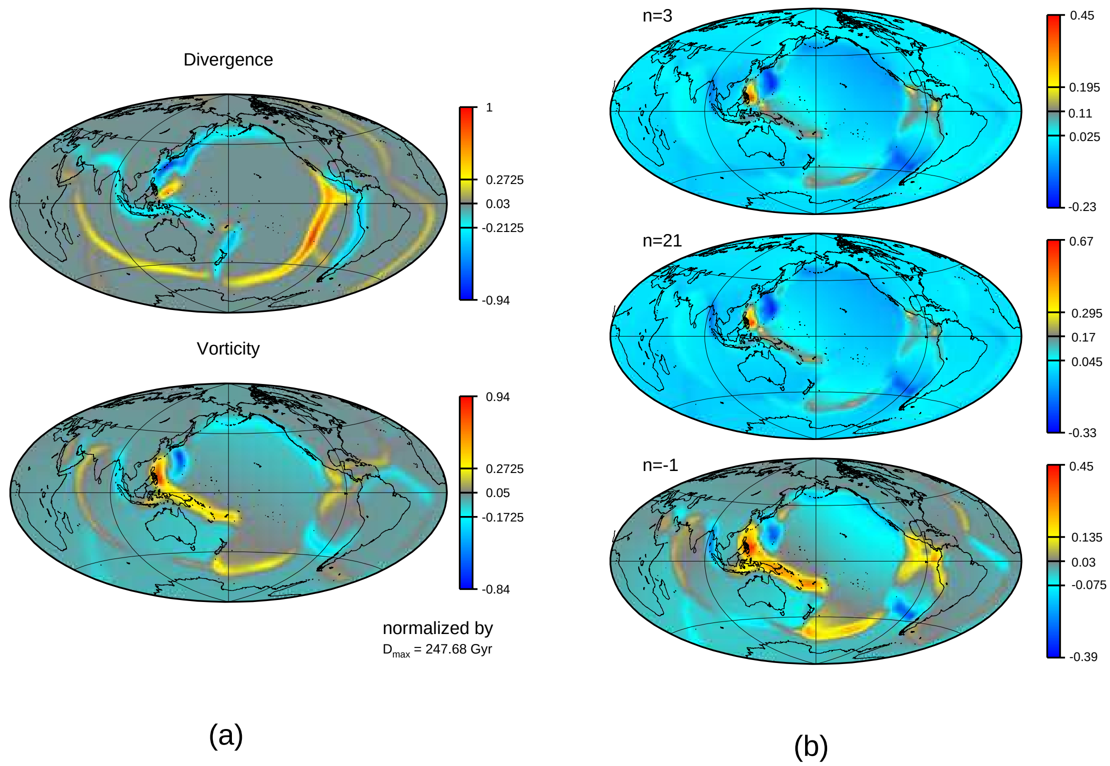

Figure 9: The source-sink model of lithospheric motion uses present day plate motions to drive flow with the divergent field (upper left) in order to model and recover the known vorticity field (lower left) using non-Newtonian flow calculations (right column). Simple power-law or pseudo-plastic rheologies (typical of dislocation creep) do not recover the vorticity field well (top two, upper right), while velocity-weakening or shear-localizing rheologies (where stress decreases with increased strain-rate) recover it very well (lower right). From *Bercovici et al.* [2000] after *Bercovici* [1995]. American Geophysical Union.

and Venus are ostensible twins but Earth has plate tectonics and Venus does not. This is usually attributed to lack of water on Venus which would otherwise lubricate plate motions; however Earth's lithosphere is likely to be no more hydrated than Venus's because of melting dehydration at ridges [*Hirth and Kohlstedt*, 1996]. Recent studies have hypothesized that the role of water in maintaining plate motion is not to lubricate plates, but to be an agent for the carbon cycle, which thus allows for a temperate climate on Earth. A cool surface on Earth, according to one hypothesis [*Lenardic et al.*, 2008] causes a larger temperature drop across the lithosphere than would occur on Venus (because of temperature-dependent viscosity), and thus lithospheric buoyant stresses are large enough on Earth to induce plastic/brittle failure, but not on Venus. An alternative hypothesis [*Landuyt and Bercovici*, 2009a] states that plate-like motion depends on the competition between damage and healing (where, for example, if damage is due to grain-size reduction, healing is due to normal grain growth), where a high damage-to-healing ratio promotes plate-like motion, while a lower ratio yields stagnant-lid behavior. A cooler surface temperature inhibits healing while a hot surface promotes healing, thus leading to plate like behavior on a planet like Earth with a temperate climate but not on a planet like Venus. Both hypotheses emphasize that water dictates conditions for plate generation by its modulation of climate and not on direct strength reduction.

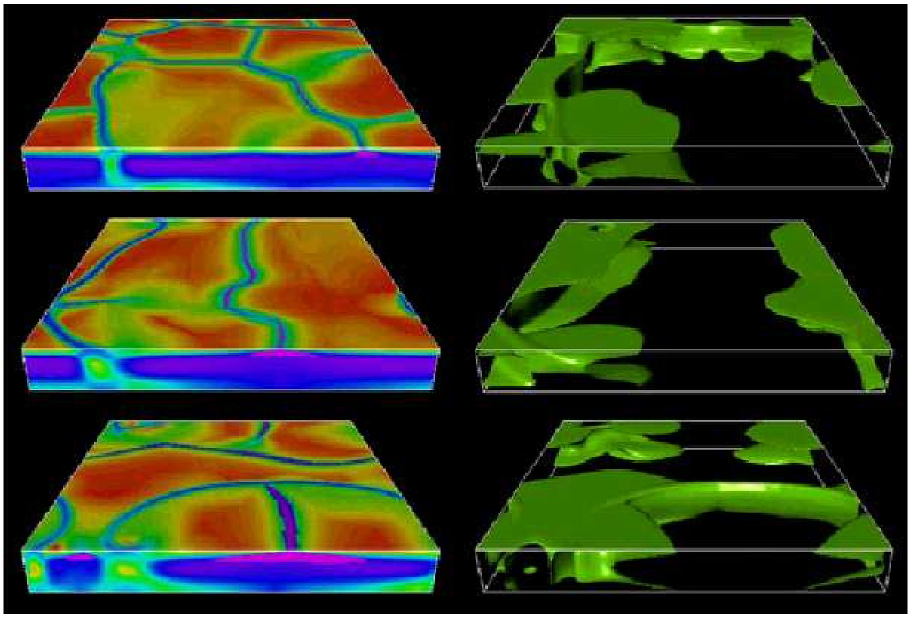

Figure 10: A simulation of plate generation over mantle convection. The plate rheology is visco-plastic and the viscosity reduction associated with melting is parameterized into the model, leading to exceptional plate-like behavior and apparent passive spreading (i.e., narrow spreading centers not associated with any deep upwelling). The right panels show surfaces of constant temperature, which here are dominated by cold downwellings; the left panels show the viscosity field (red being high viscosity and blue low viscosity). Different rows show different times in the simulation. After *Tackley* [2000]. American Geophysical Union.

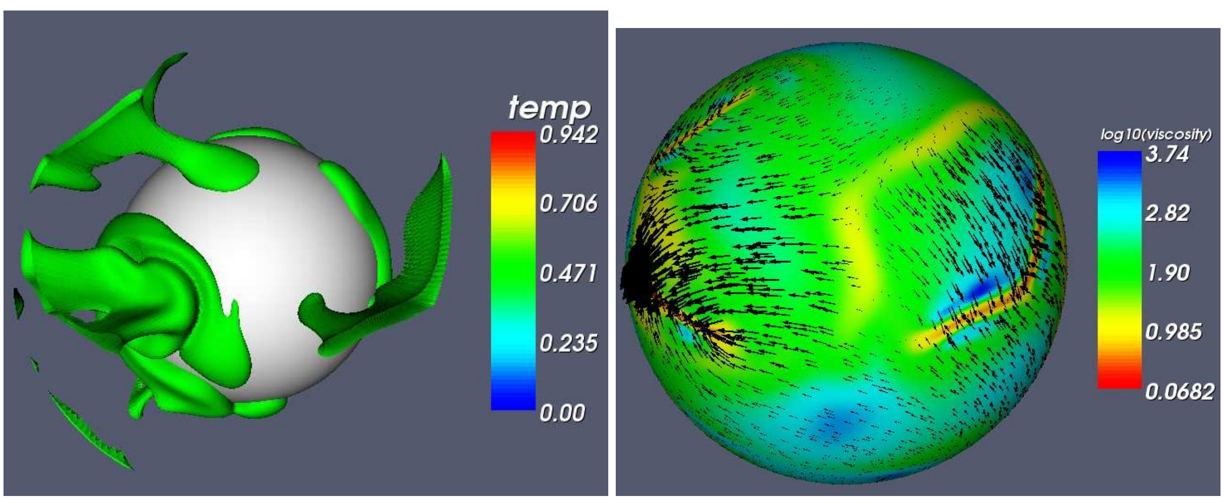

Figure 11: Three-dimensional spherical shell convection with a plastic-type lithospheric rheology showing plate-like behavior. Left panel shows isothermal surfaces and in particular cold downwellings. The right panel shows surface viscosity with velocity vectors super-imposed. Note the passive divergent and rheological weak zone forming mid-way between the two major downwelling regions. Adapted from *Foley and Becker* [2009]. American Geophysical Union.

Finally, subduction initiation continues to be an extremely challenging issue in geodynamics [*Stern*, 2004; *King*, 2007]. The strength of thick cold pre-subduction lithosphere is such that it

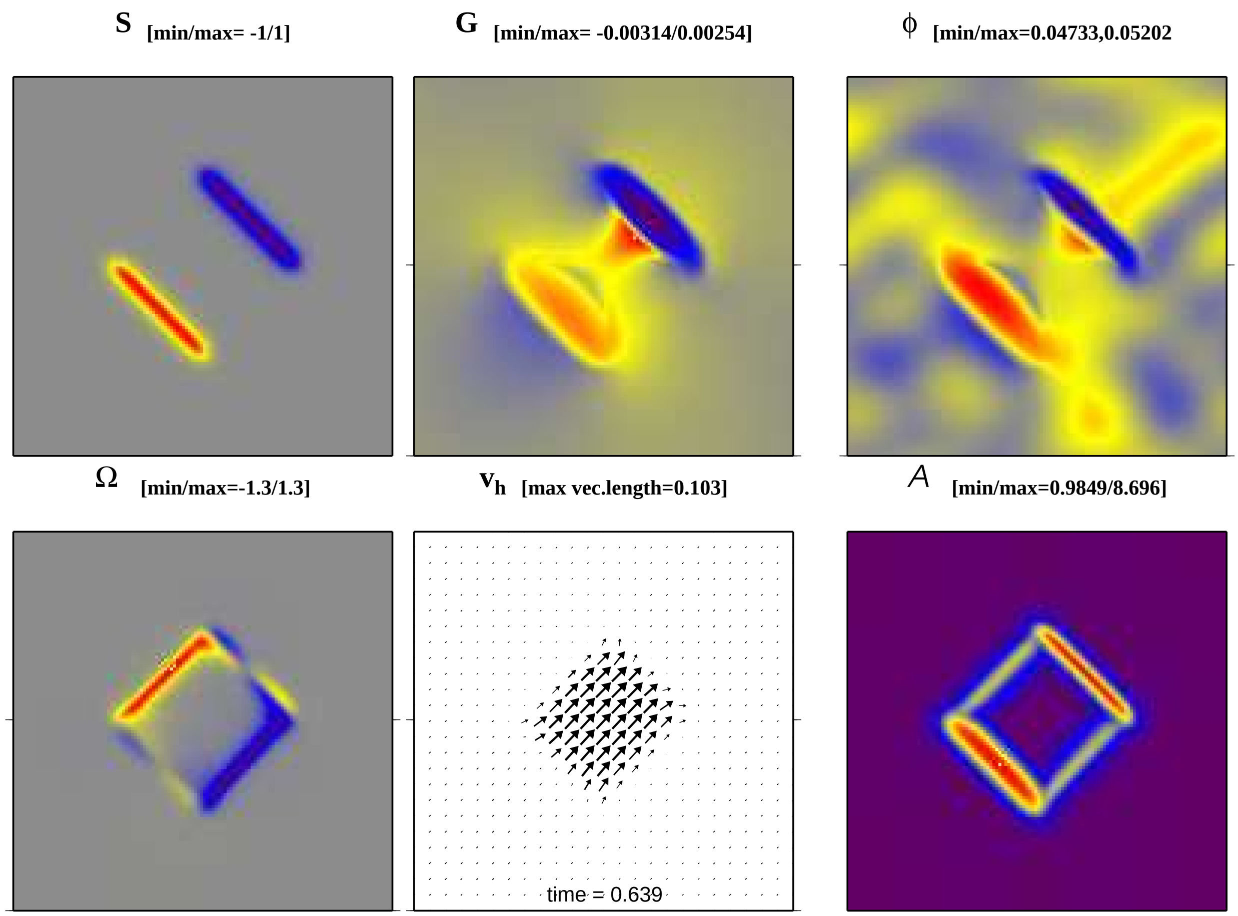

Figure 12: A simple source-sink model of shallow flow with a two-phase and grainsize-reducing damage mechanism. Damage per se involves transfer of deformational work to the creation of surface energy on interfaces by void and/or grain boundary generation in the continuum. In the case shown, all damage is focussed on grainsize reduction. The panel meanings are indicated by symbols where S is the imposed divergence rate (i.e. the source-sink field) that drives flow; G is the dilation rate due to void formation; ϕ is void volume fraction; Ω is vertical vorticity or rate of strike-slip shear; $\mathbf{v_h}$ is horizontal velocity; and $\mathcal{A}$ is the "fineness" or inverse grainsize. This particular calculation shows that fineness-generating, or grainsize reducing, damage is very effective at creating localized fault-like strike-slip zones in vorticity Ω, and solid-body like translation in the velocity field $\mathbf{v_h}$. Adapted from *Bercovici and Ricard* [2005]. American Geophysical Union.

should never go unstable and sink, at least not on geological time scales (or cosmological ones either). Thus how and why subduction zones form remain enigmatic. Mechanisms range from weakening by rifting [*Kemp and Stevenson*, 1996; *Schubert and Zhang*, 1997], sediment loading and water injection [*Regenauer-Lieb et al.*, 2001], and re-activation of pre-existing fault-zones [*Toth and Gurnis*, 1998; *Hall et al.*, 2003], all of which have some observational motivation, although fault re-activitation might be the most compelling [e.g. *Lebrun et al.*, 2003]. Another unresolved enigma concerns the age of subduction zones. The convective picture of plate motions would have plates subduct when they get old, cold and heavy. However, the seafloor age distribution implies that subduction rate is independent of plate age, such that the age of plates at subduction zones is distrubted from age nearly 0 (i.e, subducting ridges), to the oldest ages of roughly 200Myrs [see *Becker et al.*, 2009].

Summary

Mantle convection is the central theory for how the Earth works, i.e., what drives geological motions as well as the cooling history and chemical evolution of the planet. The theory is based on the simple notion that the Earth (and other terrestrial planets) must cool to space and in so doing release heat and gravitational potential energy through thermal convection. Thermal convection theory itself is a well established physical theory rooted in classical mechanics and fluid dynamics. Much of the physics of basic convection goes far in describing circulation and structure in the Earth's mantle, for example the flow velocities, establishment of thermal boundary layers, and even the prevalence of slab-like downwellings and plume-like upwellings. However, numerous quandaries and paradoxes persist because much remains to be understood regarding the many complexities of the exotic convecting "fluid" in the mantle. How the Earth appears to be stirred by deep subducting slabs, but still appear to be unmixed when it comes up at either mid-ocean ridges or ocean-islands, remains a major mystery, much of which is due to our incomplete understanding of the processes of melting, chemical segregation and mixing in the mantle. How and why the Earth's mantle convects in the form of plate tectonics at all and unlike other terrestrial planets remains one of the biggest questions in geoscience; this problem is undoubtedly related to the mantle's exotic rheology in which flow depends on multiple properties including temperature, stress, chemistry and mineral grainsize. Even more than 100 years since Kelvin's controversial dating of the Earth's age by cooling, in fact the cooling history of the Earth remains poorly understood because of incomplete knowledge of radiogenic heat sources as well as the complex physics of mantle flow. Thus while mantle convection remains one of the grand unifying physical theories of how the Earth works, many of the major questions and mysteries about the mantle remain unsolved and are thus ripe for discovery by future generations of Earth scientists.

Bibliography

Becker, T., C. Conrad, B. Buffett, and R. M¨üller, Past and present seafloor age distributions and the temporal evolution of plate tectonic heat transport, *Earth Planet. Sci. Lett.*, *278*, 233–242, 2009.

Bénard, H., Les tourbillons cellulaires dans une nappe liquide, *Revue générale des Sciences pures et appliquées*, *11*, 1261–1271 and 1309–1328, 1900.

Bénard, H., Les tourbillons cellulaires dans une nappe liquide transportant de la chaleur par convection en régime permanent, *Annales de Chimie et de Physique*, *23*, 62–144, 1901.

Bercovici, D., A source-sink model of the generation of plate tectonics from non-newtonian mantle flow, *J. Geophys. Res.*, *100*, 2013–2030, 1995.

Bercovici, D., Generation of plate tectonics from lithosphere-mantle flow and void-volatile self-lubrication, *Earth Planet. Sci. Lett.*, *154*, 139–151, 1998.

Bercovici, D., The generation of plate tectonics from mantle convection, *Earth Planet. Sci. Lett.*, *205*, 107–121, 2003.

Bercovici, D., Mantle dynamics, past, present and future: An overview, in *Treatise on Geophysics*, vol. 7, Mantle Dynamics, edited by c. e. D. Bercovici (G. Schubert, pp. 1–30, Elsevier, New York, 2007.

Bercovici, D., and S. Karato, Whole mantle convection and the transition-zone water filter, *Nature*, *425*, 39–44, 2003a.

Bercovici, D., and S. Karato, Theoretical analysis of shear localization in the lithosphere, in *Reviews in Mineralogy and Geochemistry: Plastic Deformation of Minerals and Rocks*, vol. 51, edited by S. Karato and H. Wenk, chap. 13, pp. 387–420, Min. Soc. Am., Washington, DC, 2003b.

Bercovici, D., and Y. Ricard, Tectonic plate generation and two-phase damage: void growth versus grainsize reduction, *J. Geophys. Res.*, *110*, B03,401, doi:10.1029/2004JB003,181, 2005.

Bercovici, D., Y. Ricard, and M. Richards, The relation between mantle dynamics and plate tectonics: A primer, in *History and Dynamics of Global Plate Motions, Geophys. Monogr. Ser.*, vol. 121, edited by M. A. Richards, R. Gordon, and R. van der Hilst, pp. 5–46, Am. Geophys. Union, Washington, DC, 2000.

Braun, J., J. Chery, A. Poliakov, D. Mainprice, A. Vauchez, A. Tomassi, and M. Daignieres, A simple parameterization of strain localization in the ductile regime due to grain size reduction: A case study for olivine, *J. Geophys. Res.*, *104*, 25,167–25,181, 1999.

Cadek, O., and Y. Ricard, Toroidal/poloidal energy partitioning and global lithospheric rotation during cenozoic time, *Earth Planet. Sci. Lett.*, *109*, 621–632, 1992.

Chandrasekhar, S., *Hydrodynamic and Hydromagnetic Stability*, Oxford University Press, New York, 1961.

Christensen, U., Effects of phase transitions on mantle convection, *Annu. Rev. Earth Planet. Sci.*, *23*, 65–88, 1995.

Davies, G., and M. Richards, Mantle convection, *J. Geology*, *100*, 151–206, 1992.

De Bresser, J., J. ter Heege, and C. Spiers, Grain size reduction by dynamic recrystallization: can it result in major rheological weakening?, *Intl. J. Earth Sci.*, *90*, 28–45, 2001.

Dumoulin, C., D. Bercovici, and P. Wessel, A continuous plate-tectonic model using geophysical data to estimate plate margin widths, with a seismicity based example, *Geophys. J. Int.*, *133*, 379–389, 1998.

Foley, B., and T. Becker, Generation of plate-like behavior and mantle heterogeneity from a spherical, visco-plastic convection model, *Geochem., Geophys., Geosys.*, *10*, q08001, doi:10.1029/2009GC002378, 2009.

Forsyth, D., and S. Uyeda, On the relative importance of the driving forces of plate motion, *Geophys. J. R. Astr. Soc.*, *43*, 163–200, 1975.

Grand, S., R. van der Hilst, and S. Widiyantoro, Global seismic tomography: A snapshot of convection in the earth, *GSA Today*, *7*, 1–7, 1997.

Hager, B., and R. O'Connell, Kinematic models of large-scale flow in the earth's mantle, *J. Geophys. Res.*, *84*, 1031–1048, 1979.

Hales, A., Convection currents in the Earth, *Monthly Notices Roy. Astron. Soc., Geophys. Suppl.*, *3*, 372–379, 1936.

Hall, C. E., M. Gurnis, M. Sdrolias, L. L. Lavier, and R. D. Mueller, Catastrophic initiation of subduction following forced convergence across fracture zones, *Earth and Planetary Science Letters*, *212*, 15–30, 2003.

Halliday, A., In the beginning..., *Nature*, *409*, 144=145, 2001.

Hess, H., History of ocean basins, in *Petrologic Studies - A Volume in Honor of A.F. Buddington*, edited by A. Engeln, H. James, and B. Leonard, pp. 599–620, Geol. Soc. America, New York, 1962.

Hirth, G., and D. Kohlstedt, Water in the oceanic upper mantle: implications for rheology, melt extraction and the evolution of the lithosphere, *Earth Planet. Sci. Lett.*, *144*, 93–108, 1996.

Hofmann, A., Mantle geochemistry: the message from oceanic volcanism, *Nature*, *385*, 219–228, 1997.

Hofmann, A., Sampling mantle heterogeneity through oceanic basalts: Isotopes and trace elements, in *Treatise on Geochemistry: The Mantle and Core*, edited by R. C. H. Holland and c. e. K.K. Turekian, pp. 61–101, Elsevier, 2003.

Holmes, A., Radioactivity and earth movements, *Geol. Soc. Glasgow Trans.*, *18*, 559–606, 1931.

Howard, L. N., Convection at high Rayleigh number, in *Proceedings of the Eleventh International Congress of Applied Mechanics*, edited by H. Gortler, pp. 1109–1115, Springer-Verlag, New York, 1966.

Ishii, M., and J. Tromp, Normal-mode and free-air gravity constraints on lateral variations in velocity and density of earth's mantle, *Science*, *285*, 1231 – 1236, 1999.

Ito, G., and J. Mahoney, Flow and melting of a heterogeneous mantle: 1. method and importance to the geochemistry of ocean island and mid-ocean ridge basalts, *Earth Planet. Sci. Lett.*, *230*, 29–46, 2005a.

Ito, G., and J. Mahoney, Flow and melting of a heterogeneous mantle: 2. implications for a chemically nonlayered mantle, *Earth Planet. Sci. Lett.*, *230*, 47–63, 2005b.

Jaupart, C., S. Labrosse, and J.-C. Mareschal, Temperatures, heat and energy in the mantle of the earth, in *Treatise on Geophysics*, vol. 7, Mantle Dynamics, edited by c. e. D. Bercovici (G. Schubert, pp. 253–303, Elsevier, New York, 2007.

Jellinek, A., and M. Manga, The influence of a chemical boundary layer on the fixity and lifetime of mantle plumes, *Nature*, *418*, 760–763, 2002.

Jellinek, A., and M. Manga, Links between long-lived hotspots, mantle plumes, D" and plate tectonics, *Rev. Geophys.*, *42*, RG3002, 10.1029/2003RG000,144, 2004.

Jin, D., S. Karato, and M. Obata, Mechanisms of shear localization in the continental lithosphere: Inference from the deformation microstructures of peridotites from the ivrea zone, northwestern italy, *J. Struct. Geol.*, *20*, 195–209, 1998.

Karato, S., *Deformation of Earth Materials: An Introduction to the Rheology of Solid Earth*, Cambridge Univ. Press, 2008.

Karato, S., M. Toriumi, and T. Fujii, Dynamic recrystallization of olivine single crystals during high temperature creep, *Geophysical Research Letters*, *7*, 649–652, 1980.

Karato, S., D. Bercovici, G. Leahy, G. Richard, and Z. Jing, The transition zone water filter model for global material circulation: where do we stand?, in *Earth's Deep Water Cycle, AGU Monograph Series*, vol. 168, edited by S. Jacobsen and S. van der Lee, pp. 289–313, American Geophysical Union, 2006.

Kellogg, L., B. Hager, and R. van der Hilst, Compositional stratification in the deep mantle, *Science*, *283*, 1881–1884, 1999.

Kemp, D., and D. Stevenson, A tensile flexural model for the initiation of subduction, *Geophys. J. Int.*, *125*, 73–94, 1996.

King, S., Mantle downwellings and the fate of subducting slabs: Constraints from seismology, geoid, topography, geochemistry, and petrology, in *Treatise on Geophysics*, vol. 7, Mantle Dynamics, edited by c. e. D. Bercovici (G. Schubert, pp. 325–370, Elsevier, New York, 2007.

Korenaga, J., Thermal cracking and the deep hydration of oceanic lithosphere: A key to the generation of plate tectonics?, *J. Geophys. Res.*, *112*, doi:10.1029/2006JB004502, 2007.

Korenaga, J., Urey ratio and the structure and evolution of earth's mantle, *Rev. Geophys.*, *46*, rG2007, doi:10.1029/2007RG000241, 2008.

Lamb, S., and D. Sington, *Earth Story*, Princeton Univ. Press, Princeton, New Jersey, 1998.

Landuyt, W., and D. Bercovici, Variations in planetary convective via the effect of climate on damage, *Earth Planet. Sci. Lett.*, *277*, 29–37, 2009a.

Landuyt, W., and D. Bercovici, Formation and structure of lithospheric shear zones with damage, *Phys. Earth Planet. Int.*, doi:10.1016/j.pepi.2009.03.005, 2009b.

Landuyt, W., D. Bercovici, and Y. Ricard, Plate generation and two-phase damage theory in a model of mantle convection, *Geophys. J. Int.*, *174*, 1065–1080, 2008.

Leahy, G., and D. Bercovici, On the dynamics of a hydrous melt layer above the transition zone, *J. Geophys. Res.*, *112*, doi:10.1029/2006JB004,631, 2007.

Leahy, G., and D. Bercovici, Reactive infiltration of hydrous melt above the mantle transition zone, *J. Geophys. Res.*, *115*, b08406, doi:10.1029/2009JB006757, 2010.

Lebrun, J.-F., G. Lamarche, and J.-Y. Collot, Subduction initiation at a strike-slip plate boundary: The cenozoic pacific-australian plate boundary, south of new zealand, *J. Geophys. Res.*, *108*, 2453, doi:10.1029/2002JB002041, 2003.

Lenardic, A., M. Jellinek, and L.-N. Moresi, A climate change induced transition in the tectonic style of a terrestrial planet, *Earth Planet. Sci. Lett.*, *271*, 34–42, 2008.

Lithgow-Bertelloni, C., M. Richards, Y. Ricard, R. O'Connell, and D. Engebretson, Toroidal-poloidal partitioning of plate motions since 120 ma, *Geophys. Res. Lett.*, *20*, 375–378, 1993.

McNamara, A., and S. Zhong, Thermochemical structures beneath africa and the pacific ocean, *Nature*, *437*, 1136–1139, 2005.

Montelli, R., G. Nolet, F. A. Dahlen, G. Masters, E. R. Engdahl, and S.-H. Hung, Finite-frequency tomography reveals a variety of plumes in the mantle, *Science*, *303*, 338–343, 2004.

Montési, L., and G. Hirth, Grain size evolution and the rheology of ductile shear zones: From laboratory experiments to postseismic creep, *Earth Planet. Sci. Lett.*, *211*, 97–110, 2003.

Morgan, W., Convection plumes in the lower mantle, *Nature*, *230*, 42–43, 1971.

Murakami, M., K. Hirose, K. Kawamura, N. Sata, and Y. Ohishi, Post-perovskite phase transition in $MgSiO_3$, *Science*, *304*, 855–858, 2004.

O'Connell, R., C. Gable, and B. Hager, Toroidal-poloidal partitioning of lithospheric plate motion, in *Glacial Isostasy, Sea Level and Mantle Rheology*, edited by R. S. et al., pp. 535–551, Kluwer Adademic, Norwell, Mass., 1991.

Pearson, J., On convection cells induced by surface tension, *J. Fluid Mech.*, *4*, 489–500, 1958.

Pekeris, C., Thermal convection in the interior of the earth, *Monthly Notices Roy. Astron. Soc., Geophys. Suppl.*, *3*, 343–367, 1935.

Ranalli, G., *Rheology of the Earth*, Chapman and Hall Publishers, London, 1995.

Regenauer-Lieb, K., D. Yuen, and J. Branlund, The initiation of subduction: Criticality by addition of water?, *Science*, *294*, 578–580, 2001.

Revenaugh, J., and S. Sipkin, Seismic evidence for silicate melt atop the 410-km mantle discontinuity, *Nature*, *369*, 474–476, 1994.

Ricard, Y., and D. Bercovici, A continuum theory of grain size evolution and damage, *J. Geophys. Res.*, *114*, B01,204, doi:10.1029/2007JB005,491, 2009.

Runcorn, S., Towards a theory of continental drift, *Nature*, *193*, 311–314, 1962a.

Runcorn, S., Convection currents in the earth's mantle, *Nature*, *195*, 1248–1249, 1962b.

Schubert, G., and K. Zhang, Foundering of the lithosphere at the onset of subduction, *Geophys. Res. Lett.*, *24*, 1527–1529, 1997.

Schubert, G., D. Turcotte, and P. Olson, *Mantle Convection in the Earth and Planets*, Cambridge Univ. Press, 2001.

Solomatov, V., Scaling of temperature dependent and stress dependent viscosity convection, *Phys. Fluids*, *7*, 266–274, 1995.

Song, T., D. V. Helmberger, and S. P. Grand, Low velocity zone atop the 410 seismic discontinuity in the northwestern united states, *Nature*, *427*, 530–533, 2004.

Stern, R. J., Subduction initiation: spontaneous and induced, *Earth and Planetary Science Letters*, *226*, 275–292, 2004.

Strutt, John William (Lord Rayleigh), On convective currents in a horizontal layer of fluid when the higher temperature is on the under side, *Phil. Mag.*, *32*, *529–546*, 1916.

Tackley, P., Self-consistent generation of tectonic plates in time-dependent, three-dimensional mantle convection simulations, 2. strain weakening and asthenosphere, *Geochem. Geophys. Geosystems (*G^3*)*, *1*, 2000GC000,043, 2000.

Tackley, P., Strong heterogeneity caused by deep mantle layering, *Geochemistry, Geophysics, Geosystems (*G^3*)*, *3*, 10.1029/2001GC000,167, 2002.

Tackley, P., Mantle geochemical dynamics, in *Treatise on Geophysics*, vol. 7, Mantle Dynamics, edited by c. e. D. Bercovici (G. Schubert, pp. 437–505, Elsevier, New York, 2007.

Tackley, P., D. Stevenson, G. A. Glatzmaier, and G. Schubert, Effects of an endothermic phase transition at 670 km depth in a spherical model of convection in the earth's mantle, *Nature*, *361*, 699–704, 1993.

Tauzin, B., E. Debayle, and G. Wittlinger, Seismic evidence for a global low-velocity layer within the earth's upper mantle, *Nature Geoscience*, *3*, 718–721, doi:10.1038/ngeo969, 2010.

Toth, G., and M. Gurnis, Dynamics of subduction initiation at preexisting fault zones, *J. Geophys. Res.*, *103*, 18,053–18,067, 1998.

Turcotte, D., and G. Schubert, *Geodynamics*, John Wiley & Sons, New York, 1982.

Valencia, D., R. O'Connell, and D. Sasselov, Inevitability of plate tectonics on super-earths, *Astrophys. J.*, *670*, L45–L48, 2007.

van der Hilst, R., S. Widiyantoro, and E. Engdahl, Evidence for deep mantle circulation from global tomography, *Nature*, *386*, 578–584, 1997.

van Heck, H., and P. Tackley, Planforms of self-consistently generated plates in 3d spherical geometry, *Geophys. Res. Lett.*, *35*, L19,312, doi:10.1029/2008GL035,190, 2008.

van Keken, P., E. Hauri, and C. Ballentine, Mantle mixing: The generation, preservation, and destruction of chemical heterogeneity, *Annu. Rev. Earth Planet. Sci.*, *30*, 493–525, 2002.

Vine, F., and D. Matthews, Magnetic anomalies over oceanic ridges, *Nature*, *199*, 947–949, 1963.

Wegener, A., *The Origin of Continents and Oceans*, Methuen and Co., Ltd., 1924.

Weidner, D., and Y. Wang, Chemical- and clapeyron-induced buoyancy at the 660 km discontinuity, *J. Geophys. Res.*, *103*, 7341–7441, 1998.

Weinstein, S., and P. Olson, Thermal convection with non-newtonian plates, *Geophys. J. Int.*, *111*, 515–530, 1992.

White, D., The planforms and onset of convection with temperature-dependent viscosity, *J. Fluid Mech.*, *191*, 247–286, 1988.

Wolfe, C., S. Solomon, G. Laske, J. Collins, R. Detrick, J. Orcutt, D. Bercovici, and E. Hauri, Mantle shear-wave velocity structure beneath the hawaiian hotspot, *Science*, *326*, 1388–1390, 2009.

Zhong, S., M. Zuber, L. Moresi, and M. Gurnis, Role of temperature-dependent viscosity and surface plates in spherical shell models of mantle convection, *J. Geophys. Res.*, *105*, 11,063–11,082, 2000.