The History and Dynamics of Global Plate Motions, GEOPHYSICAL MONOGRAPH 121,
M. Richards, R. Gordon and R. van der Hilst, eds., American Geophysical Union, pp5–46, 2000

The Relation Between Mantle Dynamics and Plate Tectonics: A Primer

David Bercovici[1], Yanick Ricard

Laboratoire des Sciences de la Terre, Ecole Normale Supérieure de Lyon, France

Mark A. Richards

Department of Geology and Geophysics, University of California, Berkeley

Abstract. We present an overview of the relation between mantle dynamics and plate tectonics, adopting the perspective that the plates are the surface manifestation, i.e., the top thermal boundary layer, of mantle convection. We review how simple convection pertains to plate formation, regarding the aspect ratio of convection cells; the forces that drive convection; and how internal heating and temperature-dependent viscosity affect convection. We examine how well basic convection explains plate tectonics, arguing that basic plate forces, slab pull and ridge push, are convective forces; that sea-floor structure is characteristic of thermal boundary layers; that slab-like downwellings are common in simple convective flow; and that slab and plume fluxes agree with models of internally heated convection. Temperature-dependent viscosity, or an internal resistive boundary (e.g., a viscosity jump and/or phase transition at 660km depth) can also lead to large, plate sized convection cells. Finally, we survey the aspects of plate tectonics that are poorly explained by simple convection theory, and the progress being made in accounting for them. We examine non-convective plate forces; dynamic topography; the deviations of seafloor structure from that of a thermal boundary layer; and abrupt plate-motion changes. Plate-like strength distributions and plate boundary formation are addressed by considering complex lithospheric rheological mechanisms. We examine the formation of convergent, divergent and strike-slip margins, which are all uniquely enigmatic. Strike-slip shear, which is highly significant in plate motions but extremely weak or entirely absent in simple viscous convection, is given ample discussion. Many of the problems of plate boundary formation remain unanswered, and thus a great deal of work remains in understanding the relation between plate tectonics and mantle convection.

1. INTRODUCTION

In the late 1930's, following the introduction of Alfred Wegener's theory of Continental Drift [Wegener, 1924], several driving mechanisms for the Earth's apparent surface motions were proposed. While Wegener himself favored tidal and *pole fleeing* (i.e., centrifugal) forces, Arthur Holmes and others hypothesized that thermal convection in the Earth's mantle provided the necessary force to drive continental motions [Holmes, 1931; see Hallam, 1987]. Even with the spurning and demise of the theory Continental Drift, and it's resurrection and revision 30 years later in the form of "Plate Tectonics" [e.g., Morgan, 1968], mantle convection is still widely believed to be the engine of surface motions, and, indeed, for many good reasons as we shall see in this review [see also review by Oxburgh and Turcotte, 1978]. However, there is still no complete physical theory which predicts how plate tectonics in its entirety is driven, or, more appropriately, caused by mantle convection.

There is little doubt that the direct energy source for plate tectonics and all its attendant features (mountain building, earthquakes, volcanoes, etc.) is the release of the man-

[1] Permanently at Department of Geology and Geophysics, School of Ocean and Earth Science and Technology, University of Hawaii

tle's gravitational potential energy through convective overturn (and of course that radiogenic heating and core cooling continue to replenish the mantle's gravitational potential energy). However, the precise picture of how plate motions are caused by convection is far from complete. With the recognition of the importance of slab pull, and that subducting slabs are essentially cold downwellings, it is becoming more widely accepted that the plates are an integral part of mantle convection, or more to the point they *are* mantle convection [e.g., Davies and Richards, 1992]. However, the idea that the plates arise from or are generated by mantle convection yields new complications and questions. Perhaps most difficult of these questions concerns the property of the mantle-lithosphere system which permits a form of convective flow that is unlike most forms of thermal convection in fluids (save perhaps lava lakes; see Duffield, 1972), i.e., a convection which looks like plate tectonics at the surface.

In this paper we review progress on the problem of how plate tectonics itself is explained by mantle convection theory. We begin by surveying the aspects of basic convection in simple fluid systems which are applicable to plate tectonics. We next discuss those features of plate tectonics which are reasonably well explained by basic convection theory. We then examine the numerous features of plate tectonics which are poorly (or not at all) explained by simple convection, and, in that regard, we discuss the possible new phyiscs which is needed to improve (and in some cases radically alter) present theories of mantle convection. This paper is not only a review and tutorial on the aspects of convection relevant to plate tectonics; it is also an exposition of how little we know about the causes of tectonic motions, and how much work remains in the important problem of unifying the theories of plate tectonics and mantle convection.

2. BASIC CONVECTION

Basic thermal convection in fluids is perhaps the most fundamental paradigm of self-organization in nonlinear systems [see Nicolis, 1995]. Such self-organization is characterized by the forcing of simple homogeneous yet perturbable or mobile systems (i.e., whose particles can move) far from equilibrium (e.g. by heating or imposition of chemical disequilibrium); in many cases, the systems can develop complex patterns and oscillatory or chaotic temporal behavior. In basic thermal convection, for example, a layer of fluid is heated uniformly and subsequently develops organized polygonal patterns and cells of cold and hot thermals. Indeed, the formation of tectonic plates is invariably a form of convective self-organization.

Although the phenomenon of convection was first recognized by Count Rumford [1870] and James Thomson [1882] (the brother of William Thomson, Lord Kelvin), the first systematic experimental study of basic convection was carried out by Henri Bénard on thin layers of spermaceti (whale fat) and paraffin [Bénard, 1900, 1901]. Bénard's experiments yielded striking images of honey-combed patterns and vertical cellular structure. After John William Strutt, Lord Rayleigh, failed to theoretically explain the experiments with hydrodynamic stability theory [Rayleigh, 1916], it was eventually deduced that Bénard had witnessed not thermal convection but (because his thin layers of spermaceti were exposed to air) surface-tension driven convection, also known as Marangoni convection [Pearson, 1958; see Berg et al., 1966]. When the thermal convection experiments were carried out correctly, they were found to be very well predicted by Rayleigh's theory [see Chandrasekhar, 1961, ch.2, §18]. Nevertheless, convection in a thin layer of fluid heated along its base is still called Rayleigh-Bénard (and often just Bénard) convection.

2.1. Convective instability and the Rayleigh number

Rayleigh's stability analysis of convection predicted the conditions necessary for the onset of convection, as well as the expected size of convection cells (relative to the layer thickness). Without delving into the mathematics of this analysis, suffice it to say that at the heart of the theory is a search for the infinitesimal thermal disturbance or *perturbation* most likely to trigger convective overturn in a fluid layer that is gravitationally unstable (i.e., hotter, and thus less dense, on the bottom than the top). This most destabilizing perturbation is often called the least-stable or most-unstable mode; its prediction and analysis is fundamental to the study of pattern-selection in convection, as well as other nonlinear systems.

Depending on the nature of the binding top and bottom surfaces of the layer (i.e., whether they are freely mobile, or adjacent to immobile walls), the most unstable mode leads to convection cells which have a width approximately equal to (though as much as 50% larger than) the depth of the fluid layer. This result is quite relevant to plate tectonics since it is the first basic step in predicting the size of convection cells, and thus the size of plates.

The perturbation which leads to this form of convection is deemed most unstable because, of all the possible perturbations, it will induce overturn with the least amount of heating. This heating however is measured relative to the properties of the system and as such defines the Rayleigh number. For example, say the layer is heated by basal heating only; i.e., the bottom boundary is kept isothermal at temperature ΔT hotter than the top boundary (which is also kept isothermal). The vigor of convection caused by this temperature drop ΔT depends on the properties of the system, in particular the fluid's density ρ, thermal expansivity α, dynamic viscosity μ, thermal diffusivity κ, the layer's thickness d, and the gravitational field strength g (acting normal to the layer and downward). These properties reflect how much the system facilitates convection (e.g., larger ρ, g, and α allow more buoyancy) or impedes convection (e.g., larger μ and κ imply that the fluid more readily resists motion or diffuses thermal anomalies away). The combination of these properties in the ratio $\frac{\mu\kappa}{\rho g \alpha d^3}$ gives a temperature that ΔT must exceed in order to cause convection. The Rayleigh number is the

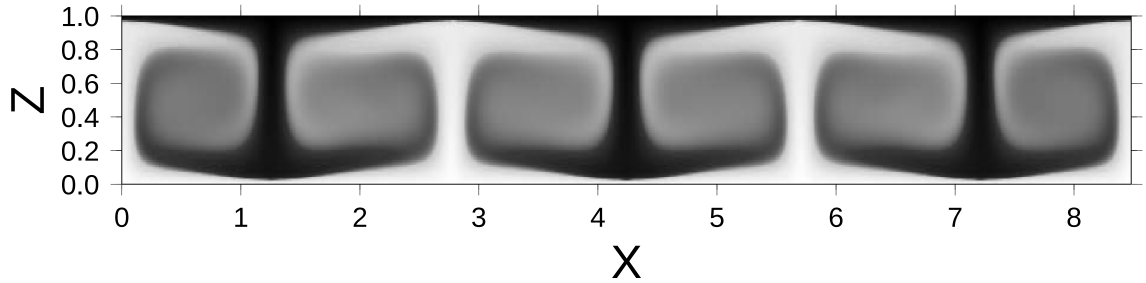

Figure 1. A vertical cross section of the temperature field of a numerical simulation of two-dimensional plane-layer, basally heated, isoviscous (Bénard), convection, with free-slip top and bottom boundaries and $Ra = 10^5$. Black represents cold fluid, light gray is hot fluid. The temperature field shows symmetric convection cells, upwellings, downwellings and thermal boundary layers thickening in the direction of motion (at the top and bottom of the layer, in between the upwellings and downwellings).

dimensionless ratio of these temperatures

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

and therefore indicates how well the heating represented by ΔT will drive convection in this system.

Stability analysis predicts that the onset of convection, triggered by the least stable mode, will occur when Ra is of the order of 1000; this varies between $Ra \approx 600$ (for the most readily induced convection with freely-slipping – called *free-slip*– top and bottom boundaries) to $Ra \approx 2000$ (with the more resistive rigid or *no-slip* top and bottom boundaries). This Ra for the onset of convection is usually referred to as the *minimum critical Rayleigh number*, or Ra_{cr}. The analysis of the onset of convection and the evaluation of Ra_{cr} with various added complications (e.g., rotation, magnetic fields, spherical geometry) is elegantly covered in the classic treatise by Chandrasekhar [1961].

In the end, the amount that the Rayleigh number of a convecting fluid exceeds Ra_{cr} is a measure of its convective vigor. While Ra_{cr} for any fluid layer is of order 10^3, the present Rayleigh number of Earth's mantle is between 10^7 and 10^9. Thus, even though the mantle convects very slowly, it does so over great distances (thousands of kilometers) and against an extreme viscosity ($10^{21} Pa\ s$, as compared to, say, $10^2 Pa\ s$ for cold molasses), and is therefore actually undergoing highly vigorous convection.

2.2. Vertical structure of simple convection

2.2.1. Symmetry Once convection initiates and becomes fully developed it is called *finite amplitude* convection, as opposed to the infinitesimal convection discussed above. The simplest form of such convection involves two-dimensional flow of constant-viscosity fluid in a plane layer that is heated on the bottom and cooled on the top (but whose upper and lower boundaries are otherwise mechanically identical, i.e. both free-slip or both no-slip). Such convection is a highly symmetric system. This symmetry is prescribed by the fact that when convection reaches its final state (i.e., a steady state or time-averaged steady state) the top boundary is cooling the fluid by the same amount that the bottom boundary is heating the fluid, otherwise the layer would heat up or cool down and is thus not in its final state. Thus, cold currents sinking from the top boundary will generally be equal and opposite in temperature and velocity to the hot currents emanating from the bottom boundary (Fig. 1). The vertical cross-section of such 2-D steady convection therefore shows sequences of mirror image pairs of counter-rotating convection cells; each cell's vertical wall is either a hot upwelling or cold downwelling current and the upwelling and downwellings are themselves symmetric mirror images of each other through a 180^o rotation (Fig. 1). The breaking of this symmetry by various effects is vital to understanding the plate tectonic style of mantle convection.

2.2.2. Thermal boundary layers When basally heated convection is reasonably vigorous (i.e. has moderate to high Rayleigh number Ra) it tends to stir the fluid until most of it is as homogeneous as possible, i.e., most of the fluid is at or near the average temperature of the top and bottom boundaries, thereby mitigating a gravitationally unstable situation for the greater part of the layer (Fig. 2). However, directly adjacent to the top and bottom boundaries the fluid temperature must make a transition from the largely homogeneous interior value to the colder (at the top) or hotter (at the bottom) boundary value (again see Fig. 2). These transition regions are called the *thermal boundary layers*. In effect, convection confines the gravitationally unstable parts of the fluid to these relatively thin boundary layers, keeping most of the rest of the fluid gravitationally stable. The Earth's lithosphere and tectonic plates are essentially the horizontal thermal boundary layer along the top surface of the Earth's convecting mantle, thus we will discuss thermal boundary layers in some detail and refer to them repeatedly.

Because of their large vertical temperature gradients, the thermal boundary layers control, via thermal conduction, the influx (through the bottom) and efflux (through the top) of heat. If the fluid is more vigorously stirred by convection then more of the fluid will be homogeneous, and thus the boundary layers will become thinner; although this leads to yet more gravitationally stable material it also causes sharper thermal gradients in the boundary layers and thus even larger heat fluxes in to and/or out of the layer. Therefore, the more vigorous the convective stirring, the greater the heatflow through the layer. (For basally heated convection in which

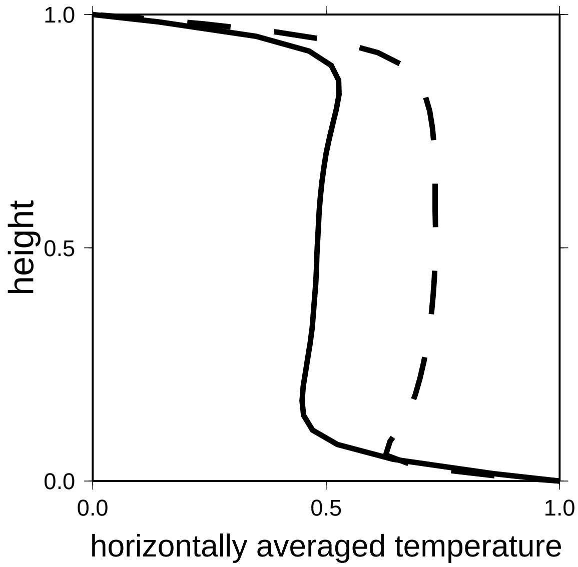

Figure 2. Temperature profiles (i.e., horizontally averaged temperature versus depth) for a basally heated, plane layer of fluid undergoing thermal convection when its viscosity is constant (solid curve) and temperature-dependent (dashed). The profiles show that most of the temperature change across the fluid occurs in relatively narrow *thermal boundary layers* near the top and bottom surfaces. In between the two boundary layers, most of the fluid is stably stratified or (if very well mixed) homogeneous. The fluid with temperature-dependent viscosity develops a stiffer upper thermal boundary layer which acts as a heat plug (i.e., it reduces convection's ability to eliminate heat), causing most of the rest of the fluid to heat up to a larger average temperature. (After Tackley [1996a].)

the top and bottom boundaries are isothermal, the heatflow through the layer is, in theory, unbounded; i.e., convection can be so vigorous as to give the fluid layer the appearance of being a material with infinite thermal conductivity. However, the same is not true for convection with only internal heating (see below) since the heatflow is limited by the net rate of heat production in the layer.)

As mentioned, the thermal boundary layers are where the gravitationally unstable material is confined, and thus fluid in these layers must eventually either sink (if in the cold top boundary layer) or rise (if in the hot bottom boundary layer), thereby feeding the vertical convective currents, i.e., downwellings and upwellings, respectively. The feeding of vertical currents induces motion of the boundary layers toward convergent zones (e.g., a downwelling for the top boundary layer) and away from divergent zones (e.g., over an upwelling). Such flow naturally causes the boundary layers to thicken in the direction of motion (see Fig. 1). For example, in the top cold thermal boundary layer, fluid from hotter depths newly arrived at the divergent zone heats and thus thins the boundary layer; but as the fluid moves toward the convergent zone it cools against the surface, and the boundary layer gradually thickens before eventually growing heavy enough to sink into the downwelling. We will revisit boundary layer thickening again when discussing subduction zones and seafloor topography.

2.2.3. The size of convection cells As mentioned previously, stability theory predicts that convection initiates with cells that are approximately as wide as the layer is thick. For the most part, this also holds true in fully developed basic convection. One of the things determining the lateral extent of a single convection cell is the length of the horizontal boundary layer currents. For example, as material in the top boundary layer current flows laterally it loses heat and buoyancy (being adjacent to the cold surface) and thus can travel only a certain distance before it becomes so cold and negatively buoyant that it sinks and feeds the cold downwelling. In the end, the lateral distance that material in the top (or bottom) boundary layer current can traverse before sinking (or rising) essentially determines the size of convection cells, and by inference (in a naive sense) the size of the tectonic plates themselves.

2.2.4. Thermal boundary layer forces What force makes the thermal boundary layer currents flow laterally? While it may seem intuitively obvious that boundary layer fluid feeding a vertical current must flow horizontally (such as fluid going toward a drain), the forces behind these currents are of significance in regard to plate forces. It should be clearly understood that buoyancy does not drive the boundary layer currents directly; buoyancy only acts vertically while boundary layer currents move horizontally (buoyancy or gravity does eventually deflect these currents back into the convecting layer, but it cannot drive their lateral flow). Indeed, horizontal pressure gradients are the primary driving force for thermal boundary layers. For example, when hot upwelling fluid impinges on the top surface, it is forced to move horizontally away from a high pressure region which is centered above the upwelling itself; the high pressure region results from a force exerted by the surface on the fluid to stop the vertical motion of the upwelling thermal. As the top cold thermal boundary layer moves away from an upwelling to its own downwelling it thickens, gets heavier and acts to pull away from the surface; this induces a suction effect and thus, because of the boundary layer's growing weight, increasingly lower pressures in the direction of motion, eventually culminating in a concentrated low pressure zone where the downwelling separates from the surface. Thus the horizontal boundary layer current flows from the induced high pressure over the upwelling to the low pressure over the downwelling, i.e., it flows down the pressure gradient. Invariably plate driving forces are related to these pressure highs and lows, i.e., pressure gradient effects. Ridge push is the gradual pressure gradient going from the pressure high at a ridge outward; slab pull is effectively due to a concentrated pressure low caused by slabs pulling away from the surface (and the concept of a slab stress-guide simply means that the pressure low is kept concentrated thus leading to sharp pressure gradients). We will discuss plate forces in greater detail later.

Though in this paper we are mainly concerned with the top boundary layer current of the mantle (i.e., the plates), suffice it to say that identical dynamics occurs at the bottom boundary layer current: the cold downwelling induces a high pressure region when it hits the bottom boundary, thereby forcing fluid away from it, etc.

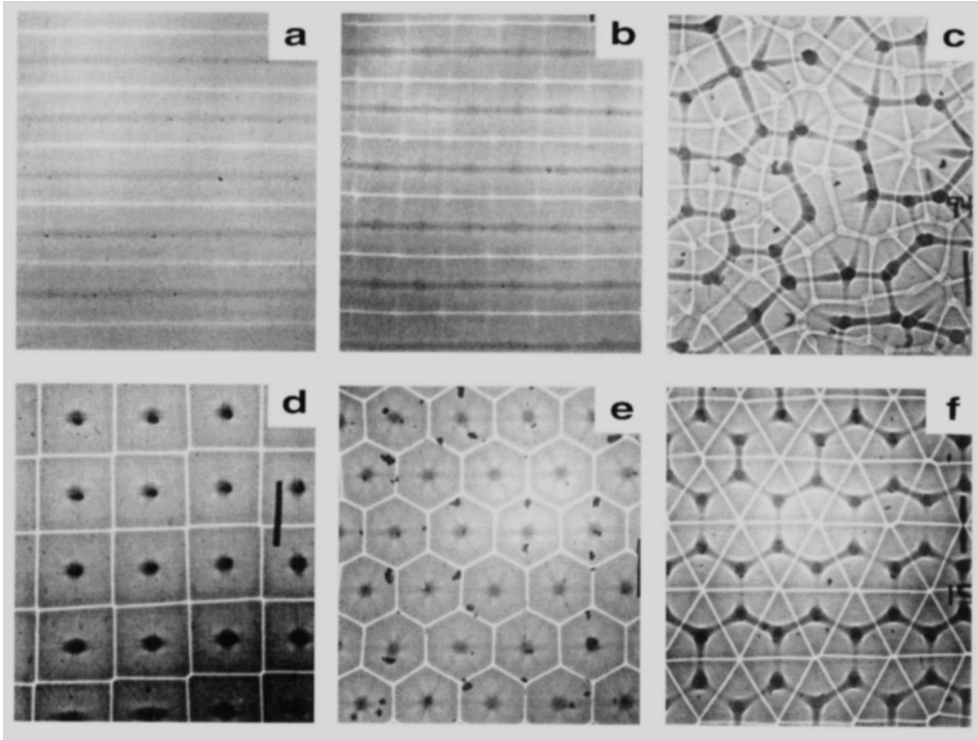

Figure 3. Laboratory experiments on convection in viscous liquids in a plane layer, bounded between two horizontal rigid glass plates, and heated from below. All images shown employ the shadowgraph techniqe (see text for discussion) and thus dark regions are relatively hot, while light regions are colder. The roll pattern (a) is also called 2-D convection; this pattern becomes unstable to the bimodal pattern (b) at moderate Rayleigh numbers, which eventually gives way to the spoke pattern (c) at yet higher Rayleigh numbers (around $Ra = 10^5$). The roll, bimodal and spoke pattern are typical for convection in isoviscous fluids, although they also occur in fluids with temperature-dependent viscosity (the photographs shown for figures (a)-(c) are in fact for weakly temperature-dependent viscosity fluids). More unique to convection in temperature-dependent viscosity fluids are squares (d), hexagons (e) and even triangles (f). (After White [1988].)

2.3. Patterns of convection

The two-dimensional convection described above is a special type of convective flow also known as convection with a roll planform; i.e., the 2D convection cells – when extended into three-dimensions – are infinitely long counter-rotating cylinders or rolls. This pattern of convection is typically unstable at moderately high Rayleigh numbers (i.e., somewhat greater than the critical Ra) at which point convection becomes three dimensional (3D) [Busse and Whitehead, 1971]; i.e., roll-like convection cells break down into more complicated geometrical shapes. The study of convective planforms and pattern selection is a very rich and fundamental field in itself, although a full quantitative discussion is beyond the scope of this paper; the interested reader should read Busse [1978]. However, we will attempt a brief qualitative discussion of convective patterns.

Much of the work on convective planforms was motivated by laboratory experiments in convection in thin layers [e.g., Busse and Whithead, 1971; Whitehead, 1976; Whitehead and Parsons, 1978; see Busse, 1978]. The planforms are most readily (albeit qualitatively) observed through a simple experimental technique called the shadowgraph [Busse and Whitehead, 1971]. Most laboratory observations of convective planforms are done through this method, so it deserves a brief mention. In the laboratory model of convection, the convecting fluid is bounded between two glass plates; the bottom is heated from below by hot water flushing along its outer surface, while the top is cooled by cold water flushed along the top. Collimated light (i.e., with parallel rays) is shown vertically up through the tank. In the convecting fluid, the index of refraction of light is dependent on temperature such that light rays diverge away from hot regions and converge toward colder regions. The projection of the resulting rays on a white screen thus show hot zones as dark shadows, and cold zones as concentrations of brightness.

In the end, the convection patterns observed with experimental (and computational) techniques are quite varied, though we will attempt to provide some order and logic to their presentation. As discussed in the previous section, purely basally-heated, plane-layer isoviscous convection is naturally a very symmetric system. Thus, even when convection is three-dimensional, the upwelling and downwelling currents retain some basic symmetry; e.g., in *bimodal* convection, the upwellings assume the geometry of two adjacent sides of a square, while the downwellings form the other two sides of the square (Fig. 3). However, the convection planform can also become highly irregular as with the *spoke* pattern which shows several nearly linear upwellings joined at a common vertex (and likewise with downwellings) (Fig. 3); even in this irregular planform the upwellings and downwellings are symmetric in that, apart from the sign of their thermal anomalies, their general structure is indistinguishable, i.e., they have the same irregular spoke shapes.

Other patterns of convection are also possible, but are mostly seen with non-constant viscosity fluids. Even in this case, regular polyhedral patterns like squares and hexagons are common (as are irregular patterns like the spoke pattern). One of the basic rules of the formation of regular patterns is that if the fluid layer is to be filled with identical convection cells, then the cells must all fit together exactly (otherwise, a misfit would constitute a cell with an aberrant shape). However, there are only a few geometries that will allow identical cells to fit together exactly, and these are simply infinitely long rolls, rectangles and squares, and hexagons (and triangles). However, even regular cells with, say, hexagonal planform involve signifcant asymmetry between upwellings and downwellings as we will discuss below. Indeed, the breaking of the symmetry of basally heated, isoviscous convection by various effects, most notably internal heating and variable viscosity, is fundamental to understanding the nature of mantle convection and its relation to plate tectonics.

2.4. Influence of internal heating

As we have discussed above, purely basally heated convection in plane layers leads to upwellings and downwellings of equal intensity (equal and oppposite thermal anomalies and velocities). However, one of the defining characteristics of convection in the Earth's mantle is that it is very likley powered by radiogenic heating from the decay of uranium, thorium and potassium distributed throughout the mantle [see Turcotte and Schubert, 1982]. Thus, the mantle is not only heated along the core-mantle boundary by the hotter molten iron outer core, it is also heated throughout its interior. This component of internal heating leads to a very significant breaking of the symmetry between upwellings and downwellings.

The simplest form of convection driven by internal heating is one in which the heat sources are distributed uniformly, the top boundary is kept cold and isothermal, and the bottom boundary is thermally insulated, i.e., no heat passes through it. This is called purely internally heated convection. In this case, the bottom boundary cannot develop any thermal boundary layers on it, as these would carry heat-conducting thermal gradients which are disallowed by the insulating condition of the boundary. Since there is no relatively hot bottom thermal boundary layer to provide concentrations of positive buoyancy then there are no focussed active upwellings emanating from the base of the layer. However, the top boundary is cold and thus does develop a thermal boundary layer whose temperature gradients are responsible for conducting out all of the internally generated heat. Material in this boundary layer is all relatively cold and heavy, and is pulled horizontally (again, via pressure gradients) toward a convergent zone where it eventually sinks to form cold downwellings. The downwellings cool the in-

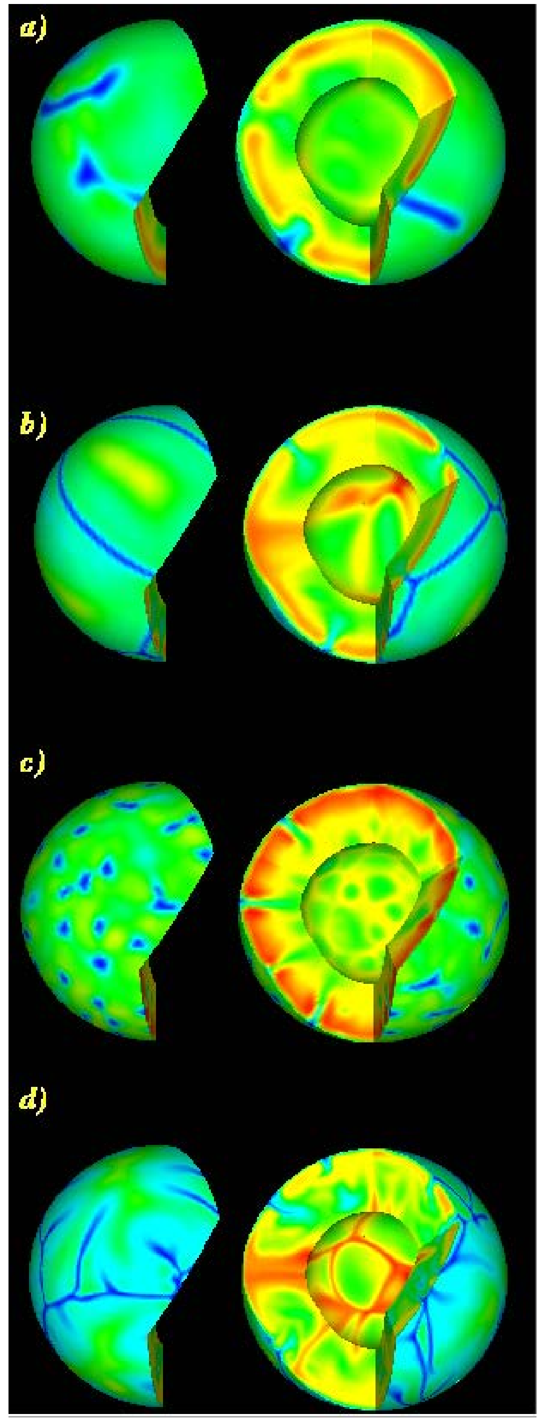

Plate 1. Purely internally heated convection in a spherical-shell model of the mantle with free-slip top and bottom surfaces, at moderate and high Ra, with and without a viscosity jump at 660km depth. Frame (a): Isoviscous mantle at $Ra = 1.6 \times 10^5$ Frame (b): Same as (a), but with the viscosity of the upper mantle (above 660km) reduced by a factor of 30. Frame (c): Same as (a) except the entire mantle viscosity is reduced by 30 (thus Ra is larger by 30). Frame (d): Same as (b), except the viscosity of both layers is reduced by 10. Note that even while an increase in Ra causes more columnar downwellings (from frame (a) to frame (c)), a viscosity jump tends to re-inforce or restore the tendency for sheet-like downwellings (frame (d)). (After Bunge et al. [1996].)

terior of the fluid layer. Thus, in purely internally heated convection, there are only concentrated downwelling currents; to compensate for the resulting downward mass flux, upwelling motion occurs, but it tends to be a broad background of diffuse flow, rising passively in response to the downward injection of cold material (Plate 1). Therefore, since the downwellings descend under their own negative buoyancy they are typically called *active* currents, while the backgrond upwelling is deemed *passive*.

If we now allow for convection with both internal and basal heating (i.e., the bottom boundary is hot and isothermal – instead of insulating – and thus permits the passage of heat) we have a form of convection a bit more complex than with either purely basal heating or purely internal heating. However, the nature of the resulting convection can be understood if one realizes that the bottom thermal boundary layer must conduct in the heat injected through the bottom while the top thermal boundary layer must conduct out both the heat injected through the bottom as well as the heat generated internally. Thus to accomodate this extra heat flux, the top thermal boundary layer develops a larger temperature drop (to sustain a larger thermal gradient) than does the bottom boundary layer. In this way, the top boundary layer has a larger thermal anomaly than the bottom one, leading to larger, more numerous and/or more intense cold downwellings than hot upwellings. Invariably, internal heating breaks the symmetry between upwellings and downwellings by leading to a preponderance of downwellings driving convective flow (Fig. 4). In the case of the Earth, whose net surface heat flux is thought to be 80% or more due to radiogenic sources, the mantle is predominantly heated internally, and only a small amount basally. Thus, one can expect a top thermal boundary layer with a large temperature drop across it feeding downwellings which dominate the overall convective circulation; active upwellings from the heated bottom boundary provide only a small amount of the net outward heat flux and circulatory work. Although we have yet to discuss many of the complexities leading to the plate-tectonic style of mantle convection, this simple picture of internally heated convection is in keeping with the idea that the large scale circulation is driven by downwellings (slabs) fed by an intense thermal boundary layer (the lithosphere and plates), while active upwellings (mantle plumes) are relatively weak and/or few in number [see Bercovici et al., 1989a; Davies and Richards, 1992].

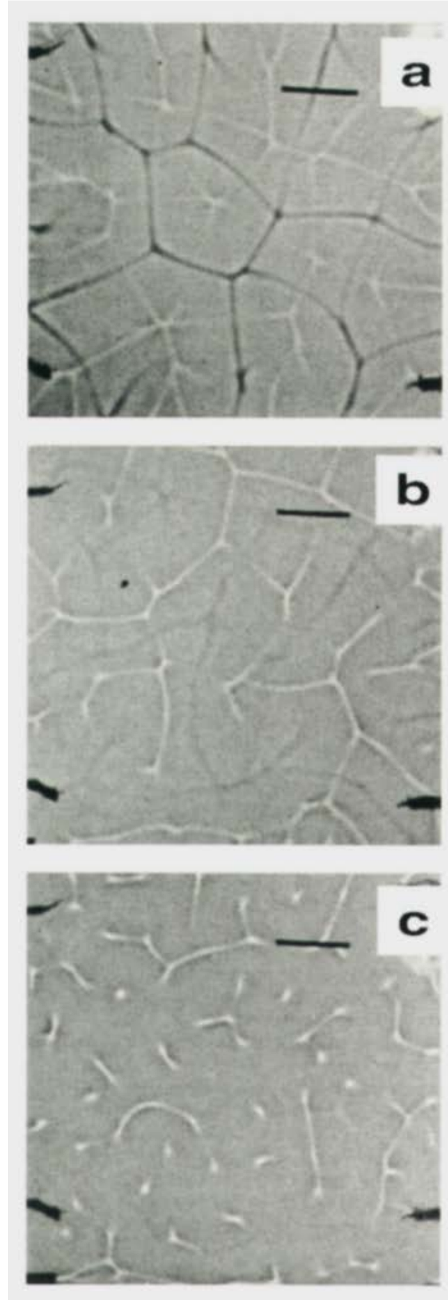

Figure 4. Laboratory experiment for internally heated convection in a plane layer. Images shown employ the shadowgraph technique (see Fig. 3 and text for discussion). Apparent internal heating is accomplished by initiating basally heated convection and then steadily decreasing the temperature of the top and bottom surfaces; this causes the average fluid temperature to be greater than the average temperature of the two boundaries. The fluid thus loses its net heat and the rate of this bulk cooling is a proxy for internal heating. The frames show the convective pattern when the layer is all basally heated (a) and thus upwellings (dark zones) and downwellings (light zones) are comparable; when bulk cooling (internal heating) is stronger than basal heating, the downwellings (light) are dominant (b); when bulk cooling is much stronger than basal heating the upwellings are not distinct enough to register in the shadowgraph (c). (After Weinstein and Olson [1990].)

2.5. Influence of temperature-dependent viscosity

Mantle material is known to have to have a highly temperature dependent viscosity for subsolidus, or solid-state, flow. Whether such subsolidus flow occurs by diffusion creep (deformation through the diffusion of molecules away from compressive stresses toward tensile stresses) or dislocation power-law creep (dislocations in the crystal lattice propagate to relieve compression and tension), the mobility of the molecules under applied stresses depends strongly on thermal activation; i.e., the atom's thermal kinetic energy determines the probablity that it will jump out of a lattice site [Weertman and Weertman, 1975; Evans and Kohlstedt, 1995; Ranalli, 1995]. The viscosity law for silicates therefore contains a quantum-mechanical probablity distribution in the form of the Arrhenius factor $e^{H_a/RT}$ where H_a is the activation enthalpy (basically the height of the energy potential-well of the lattice site out of which the mobilized atom must jump), R is the gas constant, T is temperature, and thus RT represents the average kinetic energy of the atoms in the lattices sites. Because of this factor, a few hundred degree changes in temperature can cause many orders of magnitude changes in viscosity. Moreover, with the inverse dependence on T in the Arrhenius exponent, viscosity is highly sensitive to temperature fluctuations at lower temperatures (i.e., the viscosity versus temperature curve is steepest at low T). Thus, in the coldest region of the mantle, i.e., the lithosphere, viscosity undergoes drastic changes: mantle viscosity goes from $10^{21} Pa\, s$ in the lower part of the upper mantle, to as low as $10^{18} Pa\ s$ in the asthenosphere [see King, 1995] to $10^{25} Pa\ s$ [Beaumont, 1976; Watts et al., 1982] or potentially higher in the lithosphere. Thus the viscosity may change by as much as 7 orders or magnitude in the top 200 hundred kilometers of the mantle. In the end, the effect of temperature-dependent viscosity on mantle convection is to make the top colder thermal boundary – i.e. the lithosphere – much stronger than the rest of the mantle. This phenomenon helps make thermal convection in the mantle plate-like at the surface in some respects, but it can also make convection *less* plate-like in other respects.

A strongly temperature dependent viscosity can break the symmetry between upwellings and downwellings in much the same way as internal heating. This occurs because the top cold thermal boundary layer is mechanically much stronger and less mobile than the hotter bottom thermal boundary layer. The less mobile upper boundary induces something of a heat plug that forces the fluid interior to warm up; this in turn causes the temperature contrast between the fluid interior and the surface to increase, and the contrast between the fluid interior and underlying medium (the core in the Earth's case) to decrease. This effect leads to a larger temperature jump across the top boundary layer (which partially reduces the boundary layer's stiffness by increasing its average temperature), and a smaller one across the bottom boundary layer (see Fig. 2). Thus, the temperature-dependent viscosity can lead to an asymmetry in the thermal anomalies of the top and bottom boundary layers much as we see in the Earth.

Temperature dependent viscosity can also cause a significant change in the lateral extent of convection cells. Because material in the top thermal boundary must cool a great deal and thus for a long time to become negatively buoyant enough to sink against its cold, stiff surroundings, it must travel horizontally a long distance while waiting to cool sufficiently, assuming it travels at a reasonable convective velocity. This can cause the upper thermal boundary layer and thus its convection cell to have excessively large lateral extents relative to the layer depth (i.e., large *aspect ratios*). This effect has been verified in laboratory and numerical experiments (Fig. 5) [e.g., Weinstein and Christensen, 1991; Giannandrea and Christensen, 1993; Tackley, 1996a; Ratcliff et al., 1997] As we will discuss later (§3.6), large aspect ratio convection cells are considered to occur in the Earth, especially if one assumes that the Pacific plate and its subduction zones reflect the dominant convection cell in the mantle. Thus, temperature-dependent viscosity can be used to explain the large aspect convection cells of mantle convection, but with some serious qualifications.

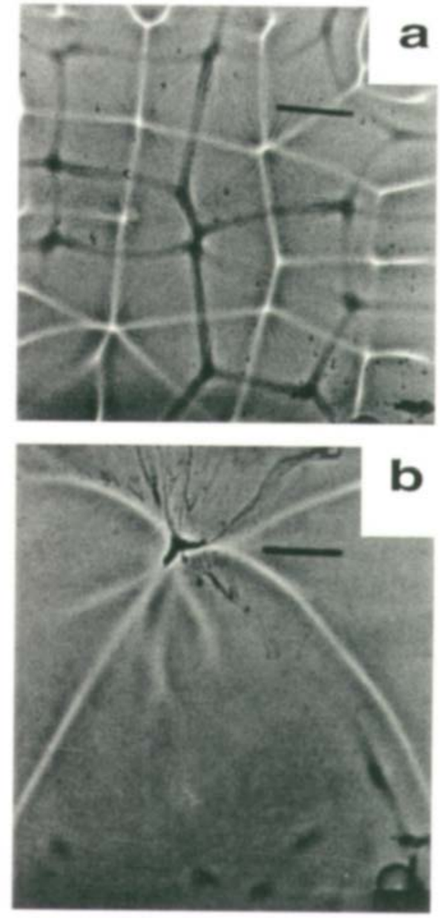

Figure 5. Lab experiment for convection with temperature-dependent viscosity fluid and mobile top thermal boundary layer. The mobilitity of the top boundary layer is facilitated by a free-slip upper boundary, accomplished by inserting a layer of low-viscosity silicon oil between the working fluid (corn syrup) and the top, cold glass plate. Without the oil, the rigidity of the top glass plate tends to help immobilize the cold and stiff top boundary layer. Frame (a) shows, for comparison, the experiment when the top boundary is rigid, yielding a predominantly spoke-like pattern. Frame (b) shows the planform when the oil is used to make the top boundary free-slip, causing a dramatic increase in the convection cell size to what is deemed the *spider* planform. (After Weinstein and Christensen [1991].)

With very strongly temperature dependent viscosity, the top thermal boundary layer can also become completely immobile and the large aspect ratio effect vanishes. The immobile boundary layer happens simply because it is so strong that it cannot move. As a result, the top boundary layer effectively imposes a rigid lid on the rest of the underlying fluid, which then convects much as if it were nearly in isoviscous convection with a no-slip top boundary condition.

Convection then has cells which are again about as wide as they are deep (i.e., with nearly unit aspect ratio). Moreover, in this form of convection the planform can assume various simple geometries such as squares and hexagons (Fig. 3); however, the upwellings and downwellings are not symmetric in that the downwellings form the edges of the hexagons (or squares), and the hot, low-viscosity upwellings tend to rise as pipes or plumes through the center of the hexagons (or squares) [see Busse, 1978], similar in many respects to plumes in the Earth's mantle. Regardless of the pattern of convection, the immobilization of the top thermal boundary layer leads to convection that is unlike the Earth: unit aspect ratio cells and an immobile lithosphere.

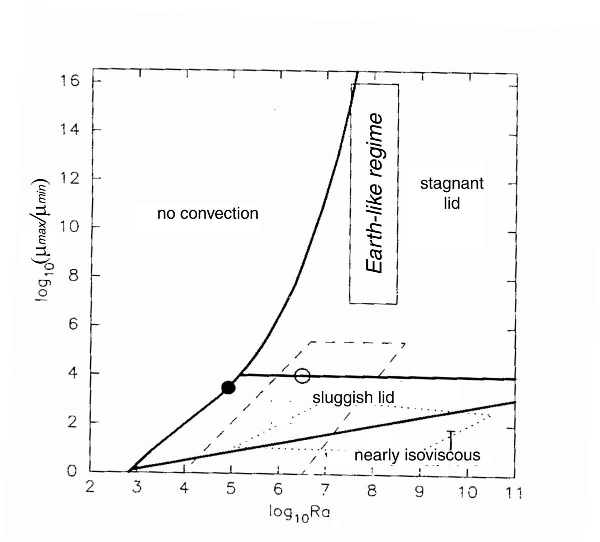

Figure 6. Diagram showing the different convective regimes in "Ra versus viscosity ratio" space for convection in fluid with temperature-dependent viscosity; μ_{max} and μ_{min} are the maximum and minimum allowable viscosities of the fluid, respectively. Dashed and dotted boxes show the regime of various numerical convection experiments. The box with the solid boundary shows the likely regime for the Earth. See text for discussion. (After Solomatov [1995].)

Overall, three different styles or regimes of convection with temperature depedent viscosity were first noted by Christensen [1984a] and elucidated and summarized by Solomatov [1995] with a basic scaling analysis (Fig. 6). These regimes are as follows: 1) for weakly temperature-dependent viscosity, convection tends to appear as if it were nearly isoviscous (unit aspect-ratio cells and a mobile top boundary); 2) with moderately temperature-dependent viscosity, convection develops a sluggish cold top boundary layer that is mobile but with large horizontal dimensions i.e. large aspect-ratio cells; 3) with strongly temperature-dependent viscosity the top boundary layer becomes rigid and immobile and convection assumes much of the appearance (apart from detailed patterns) of isoviscous convection beneath a rigid lid. The size of these regimes depends on the Rayleigh number since vigorous convection is more capable of mobilizing a cold top boundary layer than is gentle convection. These regimes have been deemed the nearly-isoviscous or low-viscosity-contrast regime, the sluggish convection regime, and the stagnant-lid regime. Although the sluggish regime is characterized by large aspect ratio cells suggestive of the Earth's plates, it is most likely that the Earth's Rayleigh number and viscosity variability would put the Earth's mantle in the staganant-lid regime, i.e., if the mantle were a fluid with only temperature dependent viscosity. Indeed, both Venus and Mars, lacking obvious signs of plate tectonics or even universally-mobile lithospheres, appear to operate in the stagnant-lid mode, at least to some extent. However, the very presence of mobile plates on Earth shows that its lithosphere-mantle system has some extremely important effects which mitigate the demobilization of the top thermal boundary layer caused by temperature-dependent viscosity. In fact, it is not at all clear that the extreme Arrhenius-type temperature dependence of mantle or lithosphere viscosity actually occurs from a practical standpoint; i.e., as we shall discuss later, with multiple competing microscopic deformation mechanisms, there is always an effective competition to be the weakest. That is, thermal stiffening caused by one mechanism can be superceded by the onset of new deformation mechanisms at higher stresses. Evidence for this is the fact that stress releases in deep earthquakes do not appear elevated from those of near-surface earthquakes, suggesting that deep, cold and thus apparently stiff slabs do not support inordinately high shear stresses [Tao and O'Connell, 1993].

A final word on the effects of temperature dependent viscosity is necessary with regard to the broad concept of "self-regulation" in solid-state convection, both in Earth and other terrestrial planets. Tozer [1972] has elegantly made the case that the internal thermodynamic state of the mantle must be controlled to a large extent by thermal activation of creep. Put crudely, if mantle viscosity is too high for convection to be vigorous enough to remove the heat generated internally, then the mantle will simply heat up until the viscosity is reduced sufficiently. Thus there is a more profound role for temperature-dependent viscosity than just the effects on style or planform discussed here, and considerations of long-term evolution of the plate-mantle system must account for the extreme sensitivity of heat flow to internal temperature via viscosity variability [Davies, 1980; Schubert et al., 1980].

2.6. A brief note on the influence of sphericity

Clearly because the Earth is primarily a sphere, models of convection in thick spherical fluid shells would appear to be most realistic [e.g., Bercovici et al., 1989a; Tackley et al., 1993; Bunge et al., 1996]. However, the effects of sphericity are possibly not so fundamental to the generation of plates. In particular, sphericity tends to break the symmetry between the top and bottom boundary layers, and upwellings and downwellings in the opposite sense of what we think is relevant for the Earth. For purely basally heated spherical shells, conservation of energy prescribes that for a convective solution to be stable (i.e. steady or statistically steady) the total heat input through the bottom boundary must be equal to the heat output through the top. However, because of sphericity, the bottom boundary has significantly less surface area than the top boundary through which heat passes; to compensate for this smaller area, the bottom thermal boundary layer generally has a larger temperature drop (and thus larger thermal gradients) across it than does the top boundary layer. This asymmetry leads to upwellings with larger temperature anomalies and velocities than the downwellings, which is opposite to the asymmetry between upwellings and downwellings thought to exist in the Earth. Thus, effects like internal heating and temperature dependent viscosity are even more important in order to overcome the asymmetry imposed by sphericity, and to give a more Earth-like asymmetry.

2.7. Poloidal and toroidal flow

Convective motion, with its upwelling and downwelling currents, and the associated divergent and convergent zones at the surface and lower boundary, is also called *poloidal* flow. Basic convection in highly viscous fluids essentially has only poloidal flow. However, while a great deal of the Earth's plate motion is also poloidal (speading centers and subduction zone), much of it – with perhaps as much as 50% of the total kinetic energy – also involves strike-slip motion and spin of plates, which is called *toroidal* motion [Hager and O'Connell, 1978, 1979, 1981; Kaula, 1980; Forte and Peltier, 1987; O'Connell et al., 1991; Olson and Bercovici, 1991; Gable et al., 1991; Cadek and Ricard, 1992; Lithgow-Bertelloni et al., 1993; Bercovici and Wessel, 1994; Dumoulin et al., 1998]. The existence of toroidal flow in the Earth's plate-tectonic style of mantle convection is a major quandary for geodynamicists and is at the heart of a unified theory of mantle convection and plate tectonics. Therefore, the reason that toroidal motion does not exist in basic convection deserves some explanation.

Most models of mantle convection treat the mantle as nearly incompressible, i.e., as if it has constant density. In fact, convection models must allow for thermal buoyancy, thus they really treat the mantle as a *Boussinesq* fluid; this means that while density is a function of temperature (and thus actually not constant), the density fluctuations are so small that the fluid is still essentially incompressible *except* when the density fluctuations are acted on by gravity. Thus the fluid acts incompressible, but can still be driven by buoyancy. Moreover, even without thermal density anomalies, the mantle is still not really incompressible; due to increases in pressure its density changes by almost a factor of 2 from the top to the bottom of the mantle (e.g., from the IASP91 reference Earth model [Kennett and Engdah, 1991]). However, mantle flow occurs on a much slower time scale than compressional phenomena – in particular acoustic waves – and thus the mantle is really considered *anelastic* for which the separation of flow into poloidal and toroidal parts still applies [Jarvis and McKenzie, 1980; Glatzmaier, 1988; Bercovici et al., 1992]. However, even the influence of this anelastic component of compressibility is rather small compared to other effects in thermal convection [Bercovici et al., 1992; Bunge et al., 1997].

The incompressibility or Boussinesq condition requires that the rate at which mass is injected into a fixed volume must equal the rate at which the mass is ejected, since no mass can be compressed into the volume if it is incompressible; i.e., what goes in must equal what goes out. Mathematically, when considering infinitesimal fixed volumes – or, equivalently, individual points in space (since a point is essentially an infinitesimal volume) – this condition is called the continuity equation and is written as

$$\nabla \cdot \underline{v} = 0 \tag{1}$$

where $\underline{v}$ is the velocity vector and ∇ is the gradient operator. This equation says that the net divergence of fluid away from or toward a point is zero; i.e., if some fluid diverges away from the point, an equal amount must converge into it in order to make up for the loss of fluid: again, what goes in must balance what goes out. The most general velocity field that automatically satisfies this equation has the form (in Cartesian coordinates)

$$\underline{v} = \nabla \times \nabla \times (\phi \hat{z}) + \nabla \times (\psi \hat{z}) \tag{2}$$

(where $\hat{z}$ is the unit vector in the vertical direction) since $\nabla \cdot \nabla \times$ of any vector is zero. (Note that the velocity field in (2) uses only two independent quantities, namely ϕ and ψ, to account for three independent velocity components v_x, v_y and v_z. This is allowed because equation (1) enforces a dependence of one of the velocity components on the other two and thus there are really only two independent quantitities; i.e., in the parlance of elementary algebra, one equation with three unknowns means that there are really only two unknowns.) The quantity ϕ is called the poloidal potential which represents upwellings, downwellings, surface divergence and convergence and (as we will show below) is typical of convective motion. The variable ψ is the toroidal potential and involves horizontal rotational or vortex-type motion, such as strike-slip motion and spin about a vertical axis; toroidal flow is not typical of convective motion (Fig. 7).

What actually forces the velocity fields in a highly viscous fluid (where fluid particles are always at terminal velocity, i.e., acceleration is negligible) is determined by the balance between buoyancy (i.e., gravitational) forces, pressure gradients and viscous resistance; in simple convection with a constant viscosity fluid, this force balance is expressed as

$$0 = -\nabla P + \mu \nabla^2 \underline{v} - \rho g \hat{z} \tag{3}$$

where P is pressure, μ is viscosity, ρ is density, and g is gravitational attraction ($9.8 m/s^2$). (The viscous resistance term, proportional to μ, in (3) expresses that viscous forces are due to imbalances or gradients in stress, while stress is due to gradients in velocity – i.e., shearing, stretching and squeezing – imposed on a fluid with a certain stiffness or

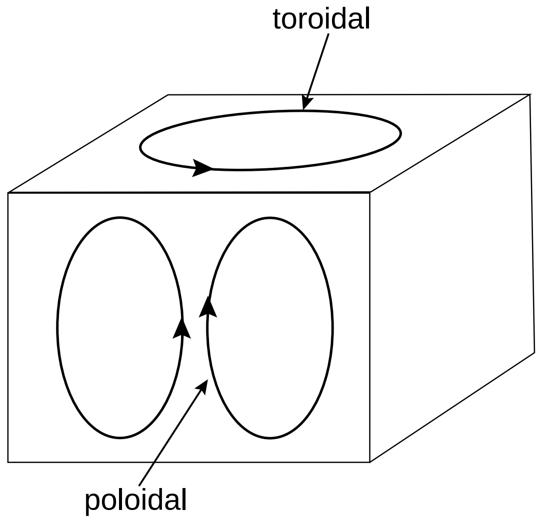

Figure 7. Cartoon illustrating simple flow lines associated with toroidal and poloidal motion.

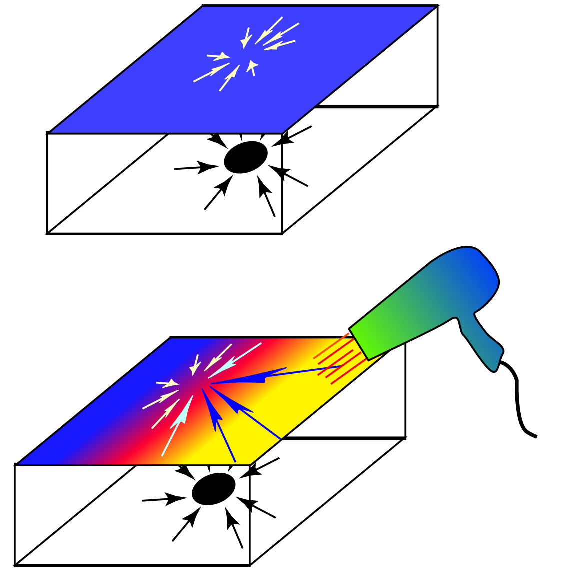

Plate 2. Cartoon illustrating the need for variable viscosity to obtain toroidal flow. Clear viscous fluid is drained through a hole in the bottom of a tank; on the clear fluid floats a uniformly cold congealed oil (e.g., butter or lard), shown in blue. The flow of clear fluid down the drain is predominanty poloidal, i.e., it only involves vertical motion and convergence toward the drain. The overlying oil is forced by the motion of the clear fluid to converge over the drain, and as long as the oil remains at one temperature it converges over the drain symmetrically, thus its flow is also poloidal. However, once one side of the oil is heated up (bottom frame), the warmer and softer side of the oil converges over the drain more readily. The differential motion between the warm, soft oil and the colder, stiffer oil appears as shear or toroidal motion (e.g., compare the light-blue and adjacent yellow flow vectors). Note that toroidal shear concentrates over the largest viscosity gradient in the oil (i.e., in the red region, at the transition between the warmest (yellow) and coldest (blue) oil); moreover, the toroidal motion only occurs because the viscosity gradient is perpendicular to the convergent poloidal flow.

viscosity.) Taking $\hat{z}\cdot\nabla\times\nabla\times$ of (3) leads to

$$\mu\nabla^4\nabla_H^2\phi = -\nabla_H^2(\rho g) \tag{4}$$

where $\nabla_H^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the 2D horizontal Laplacian. Equation (4) shows that the *poloidal* motion (left side of the equation) is directly driven by buoyancy forces (right side). However, if we take $\hat{z}\cdot\nabla\times$ of (3) we only obtain

$$\nabla^2\nabla_H^2\psi = 0 \tag{5}$$

which shows that there are no internal driving forces for toroidal flow. (Note, an identical analysis also exists for spherical geometry [e.g., see Chandrasekhar, 1961; Busse, 1975; Bercovici et al., 1989b].) Thus, for isoviscous convection, toroidal motion does not occur naturally and the only way to generate it is to excite it from the top and/or bottom boundaries; otherwise, toroidal motion does not exist. (The same argument holds even if viscosity μ is a function of height z; see §4.8.2.) To have convective flow drive toroidal motion itself, i.e., from within the medium, requires a forcing term on the right side of (5); as discussed later (§4.8.2), this can only be done if the viscosity varies laterally, i.e., $\mu = \mu(x, y)$ (see Plate 2 for a qualitative explanation). The problem of how to get buoyancy driven poloidal motion to induce toroidal flow, either through boundary conditions or through horizontal viscosity variations, is called the *poloidal-toroidal coupling* problem. This will be discussed in more detail later in the paper (see §4.8).

3. WHERE DOES BASIC CONVECTION THEORY SUCCEED IN EXPLAINING PLATE TECTONICS?

Having reviewed some of the essential aspects of simple viscous convection, we now examine those features of plate tectonics which are reasonably well explained as being characteristic of basic convective flow. Subsequently (§4), we discuss those aspects of plate tectonics which are *poorly* explained by basic convection.

3.1. Convective forces and plate driving forces

3.1.1. Slab pull and convective downwellings If we are to compare simple viscous convection to mantle flow and plate tectonics, then we wish to identify subducting slabs with the downwelling of a cold upper thermal boundary layer. As demonstrated by Forsyth and Uyeda [1975], the correlation between the connectivity of a plate to a slab (i.e., the percent of its perimeter taken by subduction zones) and the plate's velocity argues rather conclusively for the dominance of slab pull as a plate driving force. Therefore, if slabs are simply cold downwellings, then the descent velocity of a downwelling in simple convection should be comparable to that of a slab and by inference the velocity of a plate (in particular a fast slab-connected or *active* plate). (As mentioned earlier (§2.2), the pull of a slab on a plate is in fact a horizontal pressure gradient acting in the cold upper thermal boundary layer and caused by the low pressure associated with a slab pulling away from the surface; invariably the pressure gradient is established so that the boundary layer or plate feeds the slab steadily and thus leads to the appearance that the slab is pulling the plate.) We can estimate the force and velocity of such a cold downwelling with simple scaling analysis, and compare the calculated velocity with tectonic plate velocities.

Generalizing the scaling analysis and boundary layer theory of Davies and Richards [1992] we consider a cold top thermal boundary layer which is L long (from its creation at a divergent zone to its destruction at a convergent zone) and W wide (Fig. 8). As it moves from the divergent to convergent zone, the boundary layer thickens by vertical heat loss which, in the horizontally moving boundary layer, is only due to thermal diffusion.

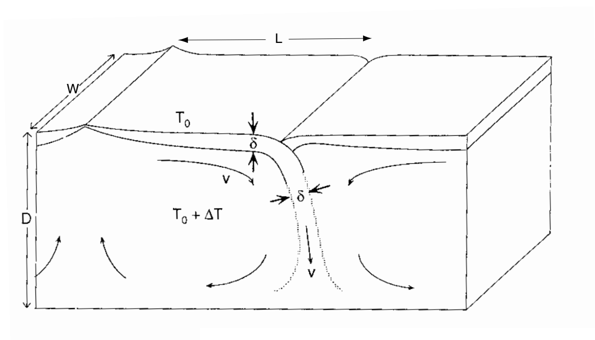

Figure 8. Cartoon illustrating the slab-pull force problem discussed in §3.1.1. For definition of labeled variables refer to text. (Cartoon after Davies and Richards [1992].)

To estimate the boundary layer thickness $\delta(t)$ we apply dimensional homogeneity [see Furbish, 1997]. Assuming that only thermal diffusivity κ (with units of m^2/s) and age t (where $t = 0$ at the divergent zone where the boundary layer is initiated) control the growth of the boundary layer, then we write $\delta(t) = A\kappa^a t^b$ where A is a dimensionless constant (A is dimensionless since we assume no other dimensional properties of the system have any influence on the cooling process). The constants a and b are then determined to match the dimensions of either side of the equation for δ, leading to $a = b = 1/2$. Indeed, a more careful analysis [see Turcotte and Schubert, 1982] shows that $\delta \approx \sqrt{\kappa t}$, within a factor of order unity.

When it reaches the site of downwelling the thickness of the boundary layer is thus $\delta = \sqrt{\kappa L/v}$, where v is the as yet undetermined velocity of the boundary layer. After downwelling, the descending vertical current extends to depth D. Moreover, because the boundary layer involves a temperature contrast of ΔT (from the deeper fluid interior to the surface), then it and the vertical current it feeds have an average thermal anomaly of $\Delta T/2$. The buoyancy force of the downwelling current is therefore

$$F_B = [\rho\alpha\Delta T/2][\delta W D]g \qquad (6)$$

where the first factor in the square brackets represents the density anomaly of the downwelling, and the second factor in brackets is the volume of the downwelling. The drag force on the downwelling current is the viscous stress (which, in this example, is approximately the viscosity times the strain-rate of the fluid as it is sheared vertically between the downwelling and a stagnation point roughly half-way back to the divergent zone) acting on the current times the area of the current on which the stress acts, i.e.,

$$F_D = \left[\mu \frac{v}{L/2}\right][DW] = 2\mu v DW/L \qquad (7)$$

The force balance $F_B - F_D = 0$ yields the downwelling velocity

$$v = \frac{\rho\alpha\Delta T \delta L g}{4\mu} \qquad (8)$$

Substituting $\delta = \sqrt{\kappa L/v}$ to eliminate v we obtain

$$v = \left[\left(\frac{\rho\alpha\Delta T g}{4\mu}\right)^2 L^3 \kappa\right]^{1/3} \qquad (9)$$

As we have seen for most forms of basic convection $L \approx D$ (i.e., convection cells are roughly unit aspect ratio). Thus using $D = 3\times 10^6 m$, $\kappa = 10^{-6} m^2/s$, $\rho = 4000 kg/m^3$, $\alpha = 3\times 10^{-5} K^{-1}$, $\Delta T = 1400K$, $g = 10m/s^2$, $\mu = 10^{22} Pa\ s$ (this viscosity represents the effective whole mantle viscosity which is dominated by the lower mantle) we obtain a velocity of $v \approx 10cm/yr$. (We can adjust less well constrained parameters, e.g., use, within reason, a lower or higher viscosity, or a value of $L > D$; even so, however, the velocity remains of the order of magnitude of $10cm/yr$.) This velocity is indeed the correct order of magnitude for velocities of active (i.e., slab-connected) plates, strongly suggesting that the plate force deemed "slab pull" is well modelled by a simple cold downwelling current in basic convection.

3.1.2. Ridge push is a convective lateral pressure gradient The ridge-push plate force was initially conceived of as an edge force (i.e. acting on the edge of a plate at a midocean ridge) due to the weight of a topographically high ridge [e.g., Forsyth and Uyeda, 1975]. It has since been recognized as a force that is distributed across the area of the plate [Hager and O'Connell, 1981]; since plate area seems to have little correlation with plate velocity [Forsyth and Uyeda, 1975] ridge push is typically assumed secondary to slab pull, (or it is almost exactly cancelled by mantle drag along the base of plates [Hager and O'Connell, 1981]). Indeed, Lithgow-Bertelloni and Richards [1998] estimate that ridge push constitutes only 5-10% the driving force due to subducted slabs. Nevertheless, after slab pull, ridge push may be one of the more significant forces and thus deserves some discussion from the perspective of basic convection.

Ridges are topographically high simply because they are younger and composed of hotter more buoyant material than the surrounding material, which has cooled and grown heavier since moving away from the ridge. The subsidence of material away from a ridge is slow enough that the entire ridge-lithosphere system is essentially in isostatic balance, at least far from subduction zones. If topography is isostatic, then there is a *compensation depth* beneath the lithosphere at which every overlying, infinitely high, column of material with equal cross-sectional area weighs the same; this also means that hydrostatic pressure (i.e., weight per area) is the same at every horizontal position along the compensation level. However, columns that do not extend to the compensation depth do not weigh the same; in particular the column centered on the ridge axis weighs the most since off-axis columns still include too much light material (like water) and not enough heavy material (like lithosphere) to equal the weight of the column of hot asthenosphere beneath the ridge. Thus, for depths less than the isostatic compensation depth, pressure is not horizontally uniform and is in fact highest beneath the ridge axis, thereby causing a pressure gradient pushing outward from the ridge [see Turcotte and Schubert, 1982]; however, this pressure gradient exists everywhere within the lithosphere (i.e., where ever sea floor depth changes laterally), and not as a line-force at the ridge axis.

How does this ridge-force pressure gradient relate to convective forces? As discussed previously (§2.2), the horizontal pressure gradient in the convective thermal boundary layers is what drives the lateral motion of the boundary layers. If we permitted the upper surface to be deformable we would find that pressure highs at divergent zones would push up the surface and the pressure lows at convergent zones would pull down on the surface. Thus the pressure gradient in the thermal boundary layer would be manifested as surface subsidence from a divergent to convergent zone. In fact, as we shall see in the next section, the resulting surface subsidence of a simple convective boundary layer is nearly identical to that observed for oceanic lithosphere. Thus, the ridge-push pressure gradient is essentially indistinguishable from the pressure gradient in a convective thermal boundary layer.

3.2. Structure of ocean basins

Large scale variations in bathymetry, or sea-floor toporaphy, can be related to simple convective processes, in particular the cooling of the top thermal boundary layer. Although we have already discussed the cooling-boundary-layer concept in the previous section, we will re-iterate the arguments in slightly more detail in order to examine the validity of the underlying assumptions.

As mentioned above, when the top thermal boundary layer of a convection cell moves horizontally, it cools by heat loss to the colder surface. In simple convection, the surface is assumed impermeable, i.e., no mass crosses it and thus heat loss out of the top thermal boundary layer is assumed entirely due to thermal diffusion; i.e., since the boundary layer abuts the impermeable surface, vertical heat loss due to ejection of hot mass or ingestion of cold mass across the surface is assumed negligible. While this assumption is reasonable for most of the extent of the thermal boundary layer, we will later see that it causes convective theory to mispredict some details of ocean-basin topography and heat flow

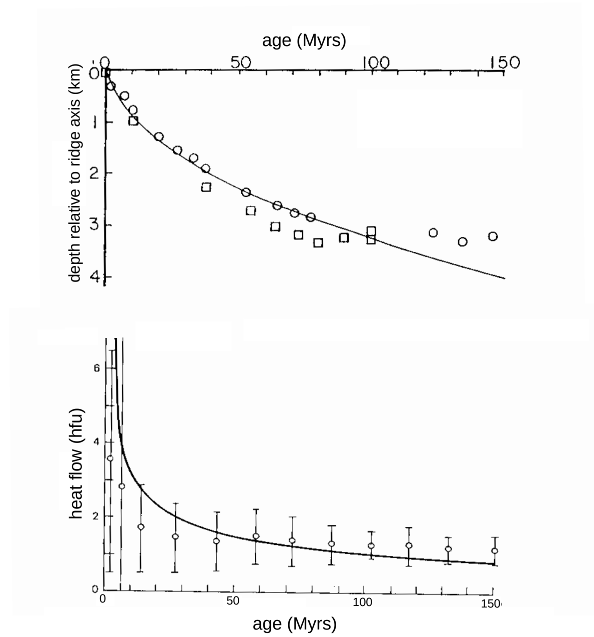

Figure 9. Bathymetry – i.e., seafloor topography – (top frame) and heatflow (bottom frame) versus age for typical ocean floor. Bathymetry is for the North Pacific (circles) and North Atlantic (squares) oceans; heatflow (in $hfu = 41.84 mW/m^2$) shows average global values (circles), and associated error bars. Both frames show the predictions from convective boundary layer theory (solid curves); i.e., a $\sqrt{age}$ law for bathymetry and a $1/\sqrt{age}$ law for heatflow. The theoretical curves fit the observations reasonably well, except near and far from the ridge axis. Adapted from Turcotte and Schubert [1982] after Parsons and Sclater [1977] and Sclater et al., [1980]. See Stein and Stein [1992] for more recent analysis.

(§4.2).

As discussed above, the assumption of diffusive heat loss means that thickening of the boundary layer is controlled by thermal diffusivity κ, with units of m^2/s; thus, again by dimensional homogeneity (see §3.1.1 and Furbish [1997]), the boundary layer thickness $\delta \sim \sqrt{\kappa t}$. As the thermal boundary layer thickens it weighs more and thereby pulls down on the surface with increasing force; if the surface is deformable, it will be deflected downward (until isostasy is established). The surface subsidence thus mirrors the increasing weight of the boundary layer; the weight increases only because δ grows, and thus surface subsidence goes as $\sqrt{t}$. Therefore, sea floor subsidence is predicted by convective theory (more precisely convective boundary layer theory [Turcotte and Oxburgh, 1967]) to increase as the square-root of lithospheric age; this is called the $\sqrt{age}$ law. Moreover, this type of boundary layer theory predicts that the heatflow (units of $Watts/m^2$) out of the ocean floor should obey a $1/\sqrt{age}$ law, which can be inferred by the fact that the heatflow across the boundary layer is mostly conductive and thus goes as $k\Delta T/\delta$ (where k is thermal conductivity and ΔT is the temperature drop across the thickening thermal boundary layer, as also defined in §3.1.1). Profiles, perpendicular to the spreading center, of sea floor bathymetry and heatflow versus age (Fig. 9) show that, to first order, sea floor subsidence and heatflow do indeed follow boundary layer theory (i.e., follow $\sqrt{age}$ and $1/\sqrt{age}$ laws, respectively) [Parsons and Sclater, 1977; Sclater et al., 1980; see also Turcotte and Schubert, 1982, ch.4; and Stein and Stein, 1992], further emphasizing that the oceanic lithosphere is primarily a convective thermal boundary layer. However, at a greater level of detail, the $\sqrt{age}$ laws fail; this will be discussed later in the paper.

3.3. Slab-like downwellings are characteristic of 3D convection

In simple, purely basally heated plane-layer convection, the convective pattern is often characterized by interconnected sheets of downwellings. For example, the hexagonal pattern is characterized by one upwelling at the center of each hexagon, surrounded by a hexagonal arrangement of downwelling walls (Fig. 3). A similar situation occurs for spherical systems [Busse, 1975; Busse and Riahi, 1982, 1988; Bercovici et al., 1989b]. When internal heating of the fluid is included, the upwellings and downwellings can be so imbalanced as to destroy any trace of symmetry; even so, the downwellings can persist as linear, albeit no longer interconnected, sheet-like downwellings at low to moderate Rayleigh numbers [Bercovici et al., 1989a]. The linear downwellings can desist and give way to downwelling cylinders and blobs at high Ra [Glatzmaier, et al., 1990; Bunge et al., 1996]; however, even at high Ra, the linearity of the downwellings is recovered by allowing viscosity to increase with depth, as is thought to occur in the Earth [Bunge et al., 1996] (see §3.6 and Plate 1). Thus, although sheet-like downwellings do not always occur in simple models of thermal convection, they are a very common feature. In that sense, the occurence of slab-like downwellings in mantle convection is characteristic of basic thermal convection. However, as we will discuss later, there are many other characteristics of subducting slabs that are unlike simple convective flow.

3.3.1. Cylindrical upwellings and mantle plumes are also characteristic of 3D convection Although mantle plumes and hotspots are not, strictly speaking, part of plate tectonics, they play an important role in our ability to measure absolute plate motions (e.g., because of assumed hotspot fixity), and in understanding the nature of mantle convection and the relative sizes of its thermal boundary layers.

As with sheet-like downwellings, cylindrical or plume like upwellings are prevalent in simple thermal convection. (It is possible to get different shaped – e.g. sheet-like – upwellings, but they tend to break down into more columnar features before reaching the surface.) However, such upwellings always require a significant bottom thermal boundary layer heated from below by a relatively hot reservoir, such as the Earth's outer core; this boundary layer provides a source region for plumes and its thickness therefore determines their size [Loper and Eltayeb, 1986; Christensen, 1984b; Bercovici and Kelly, 1997]. In basic convection with significant internal heating, the bottom thermal boundary layer typically has a much smaller temperature drop than does the top thermal boundary layer. In this case, upwellings still retain some plume-like quality though they are fewer in number and/or quite weak. Nevertheless, the existence of cylindrical upwelling plumes in the mantle – indirectly inferred from hotspot volcanism, geochemical analyses [see Olson, 1990; Duncan and Richards, 1991; White and McKenzie, 1995] and perhaps most directly in seismic analysis [Nataf and VanDecar, 1993; Wolfe et al., 1997; Chen et al., 1998] – is consistent with basic models of thermal convection. However, it should be understood that even the weak plumes seen in complete 3D models of internally heated convection are an integral part of the convective flow [see Bercovici et al., 1989a]; i.e., while they may not have much influence on the dynamics of the top cold boundary layer, they will position themselves in accordance with the whole convective pattern, typically as far from downwellings as possible [Weinstein and Olson, 1989]. Plumes well integrated into the global convective circulation are very unlikely to establish themselves where they must pass through adverse conditions (e.g., beneath a downwelling or across a strong "mantle wind"), and even if they do, they can only exist as evanescent features [c.f. Steinberger and O'Connell, 1998].

3.4. Relative fluxes of plumes and slabs are in agreement with the mantle's presumed heating mode

The heat flux transported by mantle plumes, relative to that transported by the cooling lithosphere and slabs, is also well predicted by convection models using the Earth's presumed proportion of internal heating. In basic convection, the net plume heat flux is approximately the same as the heating injected through the bottom boundary; i.e., plumes essentially carry only heat input from below [see Davies and Richards, 1992]. The Earth's mantle is thought to be between 80 and 90% heated internally by radiogenic sources [see Turcotte and Schubert, 1982]. Temperatures at the core-mantle boundary are not well known enough [see Boehler, 1996; also discussion by Davies and Richards, 1992] to definitively constrain the basal heat flux (i.e., to determine the temperature drop across the bottom thermal boundary layer, which is possibly the D" seismic layer; see Loper and Lay, 1995). However, the energy output necessary to power the geodynamo (i.e., the Earth's magnetic field) places the heat flux out of the core at approximately 10% of the net terrestrial flux [see Gubbins, 1977; Verhoogen, 1981, ch.4]. Thus, by the convective picture, plumes should transport on the whole between 10-20% of the total heat flux. Plume heat flux can be estimated by calculating the buoyancy flux necessary to keep hotspot swells inflated [Davies, 1988b; Sleep, 1990]. These calculations suggest that plumes transport roughly 10% of the net heat flow which is very consistent with our picture of basic, predominantly internally heated convection.

3.5. Mantle heterogeneity and the history of subduction

So far we can construct a simple view of the connection between plate tectonics and mantle convection in which subduction zones are associated with convective downwellings and plumes with upwellings; whether ridges are also associated with active convective upwellings is less certain, if not doubtful (see §4.7). If heating is largely internal, then downwelling of the upper boundary layer, or lithosphere, at subduction zones would then be expected to dominate the structure of mantle temperature and density heterogeneity. Indeed, this seems to be the case, but the picture is complicated by the fact that plate motions, and hence subduction zones, are ephemeral on timescales of tens to hundreds of million years. Thus one does not expect a perfect correspondence between present-day subduction zones and mantle downwellings, and it is necessary to invoke plate motion history in addressing this issue.

Averaged over the past 120-150 million years, the locations of subduction zones (in a hotspot reference frame) correspond remarkably well to the location of high seismic velocity anomalies (cold material) in the mantle, at least at very long-wavelengths [Richards and Engebretson, 1992], a view supported in some detail by higher-resolution tomographic studies of mantle structure [Grand et al., 1997]. In fact, both the geoid (the Earth's gravitational equipotential surface) [Ricard et al., 1993] and global plate motions [Lithgow-Bertelloni and Richards, 1998] are well predicted using a mantle density model derived from slabs subducted during Cenozoic and Mesozoic times; again, however, there appears to be no evidence of deeply-seated upwellings beneath mid-ocean ridges that would provide plate driving forces of similar magnitude. A detailed discussion of this type of model is given in the paper by Richards et al. in this volume. For the present purpose, we note mainly that the expected relationships among downwelling flow, mantle density heterogeneity, and plate motion (via subduction) have been confirmed observationally, so that this constitutes a clear success of concepts derived from standard convection theory applied to plate motions. This, of course, bears little on the question of how the plates themselves are generated in the system.

3.6. Temperature-dependent viscosity, internal resistive boundaries, and the aspect ratio of convection cells

The Pacific plate, with most of the kinetic energy and subduction zones of the plate tectonic system, can in many ways be considered the top of the Earth's dominant convection cell. As such, it is a very large convection cell with an excessively large aspect ratio (much longer than it is possibly deep). (The size of several other massive plates could also be used to argue for large aspect ratio cells; however, except for the Indo-Australian plate – which in itself may be composed of several other plates [Gordon et al., 1998] – no other super-large plate is considered *active*, i.e. slab-connected and "fast", and therefore does not really qualify as a possible convection cell.) Large aspect-ratio or long-wavelength convection cells in the mantle are also inferred from seismic tomography [Su and Dziewonski, 1992]. Because of the dominance of such long wavelength features, tomography and by inferrence convection are deemed to have a *reddened* spectrum.

Therefore, in many ways mantle convection in the Earth is thought to be characterized by very large aspect ratio convection cells. As discussed previously (§2.5), convection with temperature-dependent viscosity can allow for large aspect ratio cells [e.g., Weinstein and Christensen, 1991], in particular in the *sluggish* convection regime (with moderately temperature-dependent viscosity and a sluggish but mobile upper thermal boundary layer; see Solomatov, 1995). This effect can theoretically allow for a Pacific-sized convection cell; however, if the Earth's mantle were only a temperature-dependent-viscsosity fluid, estimates of the its convective regime places it more in the *stagnant lid* regime where the top thermal boundary layer is essentially frozen [Solomatov, 1995] and thus would be more characterized by unit aspect ratio convection cells [White, 1988].

However, other relatively simple effects may also cause long wavelength convection cells; these effects primarily involve an internal resistive boundary which, while not a feature of the fundamental fluid dynamics of convection, can be readily added to simple convection models to improve their match with the Earth.

As is well known, the mantle is nominally divided into the upper and lower mantle at the 660km seismic discontinuity [see review by Silver et al., 1988]. This discontinuity is most likely a phase change from spinel to perovskite plus magnesiowüstite, [Liu, 1979; Ito and Takahashi, 1989; Ito et al., 1990] and also involves a viscosity jump with depth by most likely a factor of 30-100 [e.g., Ricard et al., 1984; Richards and Hager, 1984; Hager, 1984; King and Masters, 1992; Ricard et al., 1993; Corrieu et al., 1994; see King, 1995]. Although the nature of convective mass transfer across this 660km boundary has been the subject of extensive debate, seismic tomography has given very compelling evidence for transfer of both slabs [Creager and Jordan, 1984; Grand, 1994; Grand et al. 1997; van der Hilst et al., 1991, 1997] and plumes [Shen et al., 1998] across the phase change. However, because of both the phase change and the viscosity jump, this boundary can still provide resistance to downwellings as they attempt to penetrate into the lower mantle. The phase change resists mass transfer because it is inferred to have a negative Clapeyron slope [Ito and Takahashi, 1989; Ito et al., 1990; Wicks and Richards, 1993]; this means that in relatively cold material the phase transition will occur at a higher pressure than in hotter material. Thus a cold downwelling moves the phase transition to greater pressures, causing a downward deflection of the phase boundary that, because of buoyancy, acts to rebound up and thus resist the motion of the downwelling [Schubert, Turcotte and Yuen, 1975; see also Christensen and Yuen, 1985; Machetel and Weber, 1991; Tackley et al., 1993, 1994; Honda et al., 1993; Weinstein, 1993]. A viscosity jump retards the motion of downwellings because of the greater resistance of the material beneath the boundary. In either case, convection tends to develop large-volume downwellings – and thus large wavelength convection cells – in order to gather enough negative buoyancy to penetrate a moderately resistive boundary. This "reddening" of the convective spectrum (thus creating large-aspect ratio, nearly-Pacific size cells) has been well documented in three-dimensional models of spherical convection at reasonably high Rayleigh number for both the phase change [Tackley et al., 1993, 1994] and the viscosity-jump effects [Bunge et al., 1996; see also Tackley, 1996b].

In the end we see that with temperature-dependent viscosity, or an internal boundary resistive to downwellings, basic convection can generate Earth-like, large-wavelength (or long aspect ratio) plate-sized convection cells. However, while these effects move us closer to obtaining plate-like-scales in convection models, none actually generate plates.

4. WHERE DOES BASIC CONVECTION THEORY FAIL IN EXPLAINING PLATE TECTONICS, AND WHAT ARE WE DOING (OR MIGHT WE BE DOING) TO FIX IT?

Although models of basic thermal convection in the Earth's mantle have made significant progress in explaining many features of plate tectonics, there remain a great number of unsolved problems. While we make no claim to present an exhaustive list of such unresolved issues, here we survey some of the large-scale problems remaining as well as progress being made to solve them.

4.1. Plate forces not well explained by basic convection

While slab pull and to a large extent ridge push (really, distributed ridge push) are manifestations of simple convection, other forces are not. Most of the forces unaccounted for by convection are related to some feature of plate tectonics that convection does not (or does not easily) generate. One such force is transform resistance, as this involves strike-slip motion which cannot be readily generated, if generated at all, in simple convection. The lack of excitation of strike-slip motion is a major short-coming of basic convection theory, thus we will defer discussion of this until later (see §4.8). The precise distribution of the *ridge push* force also cannot be predicted in simple convection models since they do not generate narrow passive divergent zones, i.e. thin ridges fed by shallow nonbuoyant upwellings. The lack of formation of passive but focussed upwellings is also a signficant shortcoming of basic convection theory and will be discussed later as well (see §4.7). Finally, collision resistance does not explicitly occur in basic convection since most convection models do not include continents that resist each other's motion at their contact points. Collision resistance must in the end be a manifestation of chemical segregation, i.e., placement of chemically light continental material at the surface which retains a buoyancy that resists being drawn into a downwelling [e.g., Lenardic and Kaula, 1996].

4.2. Structure of ocean basins (deviations from the $\sqrt{age}$ law)

As discussed previously, the classic bathymetry and heatflow profiles (Fig. 9) show a predominant $\sqrt{age}$ dependence. However, bathymetry and especially heatflow deviate significantly from the $\sqrt{age}$ curve near the ridge axis itself and far from the ridge axis. One must recall that the $\sqrt{age}$ prediction assumes that only diffusive transport of heat occurs in the upper thermal boundary layer. This assumption likely becomes invalid at and far from the ridge axis. At the ridge axis, heat transport by magma migration [see Spiegelman, 1996 and references therein] and cooling by hydrothermal circulation are highly significant [see Turcotte and Schubert, 1982, ch.4; Stein and Stein, 1992, and references therein]. It is likely that supression of the bathymetric and, in particular, the heatflow highs at the ridge is due to cooling not predicted by the diffusive thermal boundary layer theory; the most likely candidate for this extra cooling is indeed hydrothermal circulation. Incorporation of realistic or more complete transport phenonema relevant to ridges (e.g., magmatism and/or water ingestion) into convection models is also related to the problem of how to generate focussed but passive ridges (see §4.7).

The flattening of bathymetry, relative to the $\sqrt{age}$ law, far from the ridge axis has been the subject of some concern. It has been proposed that such flattening is the result of extra heating of the lithosphere (causing effective rejuvenation) brought on by secondary small-scale convection [Parsons and McKenzie, 1978], viscous heating [Schubert et al., 1976] and mantle plumes [Davies, 1988c]. This flattening has also been attributed to active, pressure-driven, asthenospheric flow that lifts up the lithosphere as the flow is forced into an increasingly constricted asthenospheric channel (the constriction being due to the cooling, thickening lithosphere) [Phipps Morgan and Smith, 1992]. However, since three-dimensional convection itself differs significantly from the two-dimensional boundary-layer theory from which the $\sqrt{age}$ law is derived [Turcotte and Oxburgh, 1967], such flattening and other deviations in the far-axis bathymetry are perhaps not altogether surprising from a mantle convection perspective.

4.3. Dynamic topography

Convective currents impinging on or separating from the top boundary of the mantle induce vertical stresses that deflect the Earth's surface. In particular, upwellings should be associated with topographic highs and downwellings with depressions. Of course the most obvious topographic signal is not related to these thermal anomalies but to the chemical differences between the thick, light, continental crust and the denser, thinner oceanic crust. However, one of the more distinct features of the Earth's topography, namely the mid-oceanic ridges, is a direct consequence of mantle convection (see §3.2).

What is referred to as *dynamic topography* is what remains of the observed topography when the topography due to shallow, isostatic mass anomalies in the crust and lithosphere (i.e., thickness variations in the crust and lithosphere) are removed. This remaining topography is thus sometimes called non-isostatic topography. In fact, the distinction between isostatic and dynamic or non-isostatic topography is somewhat nebulous, e.g., while ridges and subsiding seafloor are isostatically compensated, they are also related to dynamic convective processes and plate motion. However, even in this case strict removal of isostatic topography is informative. In particular, the residual (i.e, dynamic or non-isostatic) topography of the sea floor (observed topography minus $\sqrt{age}$ subsidence) reveals, in theory, the presence of thermal anomalies below the lithosphere, especially far from the ridge axis since deviations from the $\sqrt{age}$ law near ridges are presumably due to hydrothermal circulation (see §4.2).

The dynamic topography associated with long-wave-length, tomographically inferred, deep-seated heterogeneities should be theoretically quite large, i.e., ± 1000 meters [e.g., Hager and Clayton, 1987] in order to explain undulations in the geoid (the gravitational equipotential surface) of the Earth [e.g., Ricard et al., 1984; Richards and Hager, 1984]. However attempts to detect this dynamic topography have been unsuccessful, or have concluded that it cannot exceed ± 300 meters, which is in contradiction with theroretical expectations [Colin and Fleitout, 1990; Cazenave and Lago, 1991; Kido and Seno, 1994; Le Stunff and Ricard, 1995]. On a more regional scale, Lithgow-Bertelloni and Silver [1998] have concluded that South-East Africa does indeed have anomalously high topography which correlates, as predicted by theory, with both lower-mantle heterogeneity and geoid undulation. However, many other lower-mantle heterogeneities mapped by tomography are not reflected in the surface topography. The theoretical, convective dynamic topography and the observed topography might be reconciled by assuming that part of the vertical stresses of convective currents are balanced by deflection of internal boundaries (e.g., the phase transitions at 400 and 670 km depth) [Thoraval et al., 1995; Le Stunff and Ricard, 1997; Thoraval and Richards, 1997].

Since mantle convection is likely quite vigorous and time-dependent then convectively generated dynamic topography is also a time-dependent phenomenon. This means that two continents should have significantly different relative sea-level histories. Thus, the sedimentological records of continental platforms should provide rigorous constraints on time-dependent, convection based models of dynamic topography [Gurnis, 1990, 1993; Mitrovica et al., 1989]. That sea-level changes possibly reflect large-scale convective processes is still somewhat alien to the conventional sedimentological view that such changes reflect only the global eustatic signal [Vail et al., 1977]. However, Gurnis [1993] has shown that a qualitative agreement exists between possible variations of dynamic topography during the Phanerozoic and the inundation history of continents. Clearly, a quantitative solution to the dynamic topography problem would be facilitated by further interactions between sedimentologists and geodynamicists.

4.4. Changes in plate motion

Plate motions evolve with various time-scales. Some are clearly related to mantle convection, such as those associated with the so-called Wilson cycle [Wilson, 1966], i.e. the periodic formation and breakup of Pangea, approximately every 500 Myrs). Various hypotheses have been proposed to explain the dispersal of supercontinents. It is generally accepted that the presence of a supercontinent tends to insulate and thus, with radiogenic heating, warm the underlying mantle, eventually inducing a hot upwelling which weakens and breaks up the overlying lithosphere [Gurnis, 1988]. The fact that Pangea was surrounded by subduction also facilitates the thermal insulation of that portion of mantle trapped under the continent. Although this mechanism is physically sound, the present-day continents do not suggest that they stand above hotter than normal mantle. On the contrary, their basal heat flux seems very low [e.g., Guillou et al. 1994]. The effect of a super-continent on the mantle, however, may have been different from that of the present-day normal-sized continents.

Plate motion changes that have occurred on shorter time-scales are much more difficult to understand. The plate tectonic history presumably recorded in hotspot tracks consists of long stages of quasi-steady motions separated by abrupt reorganizations. During a stage of quasi-steady motion, our understanding of force balances on the plates, either in terms of plate or boundary forces (slab pull, ridge push, mantle drag, etc.) [e.g. Forsyth and Uyeda, 1975], or in terms of the buoyancy of large-scale, tomographically-inferred mantle heterogeneities [e.g., Ricard and Vigny, 1989; Lithgow-Bertel-loni and Richards, 1995], is certainly one of the obvious quantitative successes of geodynamics. This success is reinforced by the fact that the parameters entering the theory (e.g., mantle viscosity, lithospheric thickness) are in agreement with independent observations (e.g., post-glacial rebound, sea-floor topography). However, the success in predicting the direction of plate velocities is much less obvious, or is based on a tautology. For constant plate motion, the pattern of sea floor age and the location of slabs are obviously such that their associated forces push and pull the plates in the right direction. Thus, it is not surprising that, when starting from the observed positions of ridges and slabs, we predict the correct directions of plate motions. As the seismic tomography is well correlated, at least in the upper mantle, with the distribution of ridges and trenches, the ability to predict the surface motion from mantle heterogeneity simply confirms that surface tectonics and mantle dynamics belong to the same unique convecting system.

Abrupt changes in plate motion, however, are not easily related to convective processes. The most dramatic plate motion change is recorded in the Hawaiian-Emperor bend, dated at 43Ma; this bend suggests a velocity change of a major plate of approximately 45^o during a period no longer than 5 Myrs, as inferred from the sharpness of the bend. It is only because other hotspots were further away from the rotation pole of the Pacific that they did not record such a sharp bend. However all the Pacific hotspots have moved coherently and the various Pacific hotspot chains are indeed in agreement with a sudden change in the Pacific plate velocity [Fleitout & Moriceau, 1992]. Recent papers have challenged this conventional interpretation and suggest that the kink in the hotspot lines reflects, on the contrary, the motion of hostpots [Norton, 1995; Tarduno and Cottrell, 1997; see also Christensen, 1998]. Nevertheless, we believe, for the time being, that an abrupt change in Pacific plate motion is more physically plausible than the simultaneous motion of all the Pacific hotspots and intevening mantle over such short time scales. All convective type plate-driving forces, either those existing in the buoyancy of tomographically-inferred mantle heterogeneity or those associated with the thermal boundary layer (e.g., slabs, thickening of the oceanic lithosphere) can only evolve with convective time-scales (> 100 Myrs, which is based on a mantle overturn at a typical convective velocity; see §3.1.1). Therefore, changing convective motions in less than ~ 5 Myrs seems physically impossible. Perhaps the only way to explain short time-scale changes would be to invoke mechanisms that pertain to fracture or some other rapid rheological response (not controlled by the convective temperature field), or to changes in plate boundaries and geometries (which are possibly also controlled by rheology; see below). Plate reorganizations due to 1) the anihilation of a subduction boundary by continental collision, 2) loss of a ridge and/or trench by subduction of a ridge, or 3) the initiation of a new subduction zone, [e.g. Hilde et al., 1977] possibly occur on relatively rapid time scales, although such mechanisms have never been quantified or formulated rigorously. An attempt to quantitatively test the idea that the change in the Pacific plate motion was related to the Himalayan collision has simply shown that our present understanding of plate driving forces fails to explain plate velocity changes [Richards and Lithgow-Bertelloni, 1996]. Offsets in plate age and thickness along already-weak oceanic transform boundaries may provide excellent sites for initiation of subduction [e.g., Stern and Bloomer, 1992 and references therein], yielding a possible mechanism for rapid global plate motion changes. This, coupled with the notion of transform faults as long-lived "motion guides" (see §4.8.4), emphasizes the possible key role of transform boundaries in the plate-mantle system. This clearly points to the necessity of future studies to understand the interactions between the long time-scale convective processes, and the short time-scale effects, such as those likely to occur in the lithosphere's rheological response, e.g., fracture, fault sliding and various strain-localization mechanisms (see below).

4.5. Plate-like strength distributions: weak boundaries and strong interiors

One of the predominant plate-like characteristics of the Earth's top thermal boundary layer is its strength or viscosity distribution. This distribution is marked by most of the boundary layer, in particular that which is undergoing horizontal motion, as being strong or of high viscosity; these strong regions are separated by narrow zones of intense deformation (i.e., the boundaries) which are necessarily weak to permit the observed levels of strain. The cause for such a strength distribution is very much at the heart of understanding how plate tectonics arises from convection. and is thus naturally the subject of much debate.

4.5.1. Do plates form because the lithosphere breaks? It is often assumed that plates and their strength distribution form because the entire oceanic lithosphere is brittle and breaks under convective stresses. Thus, once the lithosphere is broken it is not easily mended; the broken edges remain – more or less permanently – as sites of weakness and low-frictional sliding [e.g., Davies and Richards, 1992]. However, the depth of brittle failure in oceanic lithosphere has undergone considerable revision in the rheological literature and is no longer thought to be a simple and deep transition to ductile motion. Indeed, brittle failure and frictional sliding cease at quite shallow depths, on the order of 10km [Kohlstedt et al., 1995], and give way to semi-ductile/semi-brittle behavior which involves ductile cracking and void growth (involving noncontiguous, distributed populations of voids and microcracks; Kohlstedt et al., 1995; Evans and Kohlstedt, 1995) and a variety of complex rheological phenomena. This semi-ductile behavior, and not fracture and frictional sliding, persists for most of the thickness of the lithosphere, before giving way to ductile/viscous flow at greater depths. The model of the broken lithosphere is therefore no longer valid. The lithospheric deformation mechanisms leading to plates likely contain both brittle discontinuous and ductile continuous (as well as other highly complex) processes.

4.5.2. Continuum models and plateness Given the above considerations, many convection models seeking to generate plates employ a continuum approach; i.e., they do not break the top cold boundary layer into pieces, but instead keep it whole while employing rheologies that allow parts of it to become weaker than other parts. In doing so, they try to obtain the property of a plate-like strength (or viscosity) distribution for their top thermal boundary layer. This property is quantifiable and has been called *plateness* [Weinstein and Olson, 1992]. A boundary layer with broad, strong, slowly deforming regions (representing strong plate interiors) separated by narrow zones of weak, rapidly deforming material (representing plate boundaries) have high plateness; clearly convection in the Earth has a top thermal boundary layer with high plateness. A boundary layer with no strength variations, as occurs in basic convection with constant or only depth-dependent viscosity, has zero plateness.

Since plateness involves lateral variations in strength or viscosity, convection models can obviously only generate sufficient plateness if they employ a fully variable viscosity law. Generally viscosity is variable by virtue of being a function of the thermodynamic state (e.g. temperature and pressure), state of stress and/or composition of the medium.

If viscosity is temperature-dependent, the top boundary

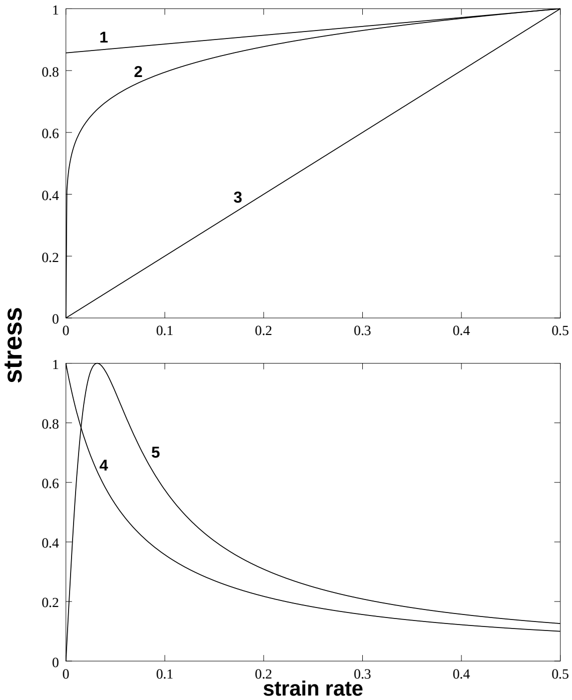

Figure 10. Constitutive (stress versus strain-rate) laws for (1) plastic, (2) non-Newtonian power law, (3) Newtonian, (4) brittle stick-slip and (5) self-lubricating (also called continuous stick-slip) rheologies.

layer will tend to be weakest at the divergent zone (where the boundary layer is thinnest) which approaches plate-like behavior. However, the viscosity will be highest over the cold convergent zone, tending to immobilize it and eventually the entire boundary layer, which is of course one of the root causes for the subduction initiation problem (see §4.6); in the end, this also leads to a thermal boundary layer with little plateness. Therefore, it is a reasonable assumption that plate-like strength distributions – or high plateness – require a lithospheric rheology that permits boundaries to be weak, regardless of temperature.

At high stresses mantle rocks undergo non-Newtonian creep wherin deformation responds nonlinearly to the applied stress; i.e., the stress–strain-rate constitutive law is nonlinear through a power-law relation such that

$$\dot{e} \sim \sigma^n \tag{10}$$

where $\dot{e}$ is strain-rate, σ is stress, and n is the *power-law index* (see Fig. 10); for mantle rocks, $n = 3$, typically [Weertman and Weertman, 1975; Evans and Kohlstedt, 1995]. The effective viscosity is

$$\mu \sim \sigma/\dot{e} \sim \dot{e}^{\frac{1}{n}-1} \tag{11}$$

which means that for $n > 1$ viscosity decreases with increasing strain-rate. The power-law rheology causes regions of the material that are rapidly deformed to become softer, while slowly-deforming regions become relatively stiff. This rheology not only yields reasonable plateness, but also leads to a feedback effect: as the deformed parts of the fluid soften they become more readily deformed, causing further strain to concentrate there, thereby inducing further softening, etc. (Comparable effects can be caused with other rheologies such as Bingham plastics and bi-viscous laws; however, these are just further mathematical models for essentially the same strain-softening effect represented by a power-law rheology.) Such non-Newtonian rheologies have been incorporated into numerous 2-D convection models for $n = 3$ and have found little plate like behavior [e.g., Parmentier et al., 1976; Christensen, 1984c]. Other models, concentrating on plate formation, placed a thin non-Newtonian "lithospheric" fluid layer atop a thicker convecting layer, or mathematically confined the non-Newtonian behavior to a near-surface layer [Weinstein and Olson, 1992; Weinstein, 1996; Moresi and Solomatov, 1998; see also Schmeling and Jacoby, 1981] (Fig. 11). In these latter models, the non-Newtonian effect does indeed yield an upper layer with reasonably high *plateness* for basally heated convection: the layer becomes weaker over both divergent zones (i.e. over upwellings) and convergent zones (downwellings), and is relatively immobile and strong in between the these two zones. Moreover, as noted by King et al. [1992], 2D models with non-Newtonian rheology are often indistinguishable in some respects from those with imposed plate geometries [e.g., Olson and Corcos, 1980; Davies, 1988a] and lithospheric weak zones [e.g., King and Hager, 1990; Zhong and Gurnis, 1995a]. However, to obtain sufficient plateness with the non-Newtonian models, it is necessary to use a power-law index n considerably higher than inferred from rheological experiments (typically $n > 7$ is needed). Thus, it seems that a general strain-weakening type rheology is a plausible and simple solution to generating high *plateness*, i.e. weak boundaries and strong plate interiors, in a convection model. However, when the convective flow is driven by internally heated convection, which has little or no concentrated upwellings, then the divergent zones in the lithospheric layer tend to be broad and diffuse, i.e., not at all like narrow, passively spreading ridges [Weinstein and Olson, 1992] (Fig. 11).

It is undoubtedly naive to assert that plate-like behavior simply requires the generation of weak boundaries; in fact each type of boundary, i.e., convergent, divergent and strike-slip, develops under very different deformational and thermal environments. Obtaining sufficient plateness is invariably related to the the nature of how the different boundaries form. As each type of boundary is uniquely enigmatic, they warrant individual discussion. These will be the focus of the following sections.

4.6. Convergent margins: initiation and asymmetry of subduction

The downwelling currents in simple convection are for the most part very symmetric; i.e., one downwelling is composed of two cold thermal boundary layers converging on each other (Plate 1). In the Earth, the downwellings associated with plate motions – i.e., subducting slabs – are highly asymmetric, i.e., only one side of the convergent zone – only

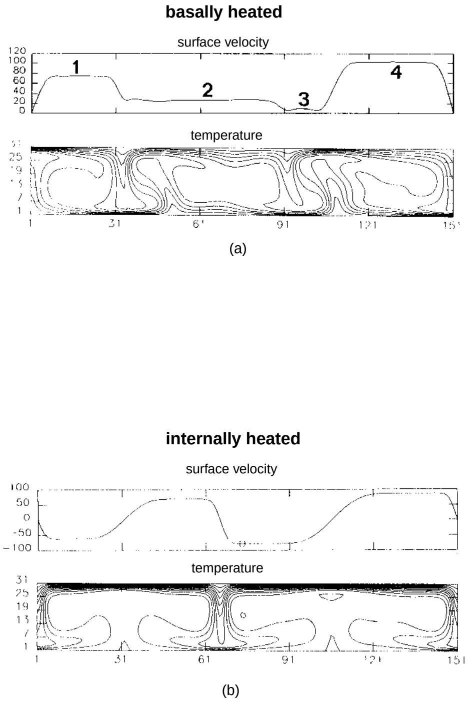

Figure 11. A lithospheric layer with non-Newtonian (power-law) rheology is driven by underlying 2D convection for both basal (a) and internal (b) heating. For both cases the convecting layers have $Ra = 50,000$ and the overlying lithosphere has a rheology with a very large power-law index of $n = 15$. The bottom of segments of (a) and (b) show a cross section of a temperature field in the convecting layer. Above each of these is shown the horizontal velocity for the non-Newtonian lithospheric layer. The step-like transitions in velocities suggest plate-like motion (high plateness) since a segment of nearly constant velocity is moving as a contiguous block with little internal deformation (the plate-like segments are numbered in the top frame). The step-like character of the velocity field is much more distinct for the basally heated convection where there are both active upwellings and downwellings to define narrow divergent and convergent margins. The internal heating has broad passive upwellings which cause the overlying divergent zones to have smooth, wide and thus unplate-like changes in velocity. (After Weinstein and Olson [1992].)

one plate – subducts. The cause for this asymmetry remains one of the greater mysteries in geodynamics. A common hypothesis is that the asymmetry is caused by chemically light continents counteracting the negative buoyancy of one of the boundary layers approaching the convergent zone [e.g., see Lenardic and Kaula, 1996]; however this hypothesis is not universely valid since there are subduction zones entirely in the ocean.

The asymmetry may also reflect an inequality of pressure on either side of the subducting slab if it deviates slightly out of vertical; i.e., a more acute corner flow on one side of the subduction zone induces lower pressure than on the opposite side, thereby causing the slab to be torqued to the more acute side, i.e., to enhance its obliquity [see Turcotte and Schubert, 1982]. The oblique downwelling may then act to impede the motion and descent of the boundary layer approaching the acute angle, since it would have considerable resistance to making a greater than 90^o downward turn, eventually leading to asymmetric convergence. This effect is most significant for slabs that are effectively rigid relative to the surrounding mantle, in order that they act as stiff paddles while being lifted up; thus if it occurs, the effect would theoretically be visible in basic convection calculations with strongly temperature-dependent viscosity. However, if the viscosity is so temperature-dependent that it induces such strong slabs then it will also likely place convection itself into the *stagnant* lid regime; i.e., to make the slabs strong enough will also lock up the lithosphere (see §2.5 and §3.6). To then adjust convection models to mitigate this extra problem requires proper initiation of subduction from cold, thick and strong lithosphere [see Mueller and Phillips, 1991; Kemp and Stevenson, 1996; Schubert and Zhang, 1997]. Models with imposed weak zones in the lithosphere show that cold rigid lithosphere subducts readily and can even assume a fairly oblique slab-like angle [Gurnis and Hager, 1988; King and Hager, 1990; Zhong and Gurnis, 1995a]. However, it is important to note that, based on seismically inferred deformation of slabs, it is not clear whether actual slabs are excessively stiff, or even stronger than the surrounding mantle [Tao and O'Connell, 1993].

The generation of the weak zone which permits subduction is itself enigmatic. Recent models have suggested that the necessary weak zone occurs as the result of faulting and rifting [Kemp and Stevenson, 1996; Schubert and Zhang, 1997]. However, brittle failure of the entire lithosphere at its thickest point is problematic (see §4.5.1). Other weakening mechanisms at the subduction zones may be necessary [e.g., King and Hager, 1990], although the nature of these mechanisms is still unclear, showing that the problem of downwelling asymmetry and subduction initiation remains one of the more elusive yet fruitful fields of geodynamical research.

4.7. Divergent margins: ridges, and narrow, passive upwellings

Basic convection predicts that upwellings occur as either columnar plumes rising actively under their own buoyancy, or as a very broad background of upwelling ascending passively in response to the downward flux of concentrated cold thermals. No where in basic convection theory or modelling does there occur concentrated but shallow and passive upwellings (i.e., which rise in response to lithospheric spreading motion, not because of any buoyancy of their own) analagous to ridges [Lachenbruch, 1976; see Bercovici et al., 1989a]. That all mid-ocean ridges involve passive upwelling is not necessarily universal [Lithgow-Bertelloni and Silver, 1998]. However, the fastest and arguably the most signficant ridge, the East Pacific Rise, is almost entirely devoid of a gravity anomaly, implying shallow, isostatic support and thus that there is no deep upwelling current lifting it up [see Davies, 1988b]. (Although the interpretation of gravity by itself is nonunique, the lack of a signficant free-air gravity anomaly over a topographic feature suggests that the gravity field of the topographic mass excess is being cancelled by the field of a nearby mass deficit; this deficit is most readily associated with a buoyant and shallow isostatic root on which the topographic feature is floating.) Moreover, 90% of the structure of the Earth's geoid (the equipotenial surface) can be explained by the gravitational attraction of the mass anomalies due only to slabs, implying that other mass anomalies, e.g., those due to active buoyant upwellings, are much less significant [Ricard et al., 1993].

The cause and initiation of narrow ridges is easily as enigmatic as the problems with subduction. The orientation of ridges more or less mirrors the subduction zones that they eventually feed, thus they may initiate as a strain localization, such as a self-focussing necking instability [e.g., Ricard et al., 1986], or quite simply an effective tear in the lithosphere. This would suggest that the stresses in the lithosphere due to the pull of slabs is guided considerable distances such that they can be concentrated on particular regions (otherwise, continuous release and diffusion of stresses becomes manifested as distributed – and thus very unridge-like – deformation). A cold lithosphere, strong by virtue of its temperature-dependent viscosity, makes a plausible stress guide. Strain-localization may be due to a variety of nonlinear feedback mechanisms, e.g., necking in a non-Newtonian strain-softening medium, or by weakening due to the presence of magma (induced by pressure release melting within the passive upwelling) or ingestion of volatiles (e.g., water, which both provides lubrication and facilitates melting by reducing the the melting temperature; see Hirth and Kohlstedt [1996]). That strain-localization occurs because of brittle failure across the entire lithosphere is, again, probably not likely (see §4.5.1). Finally, while stress guides and strain-localization are viable mechanisms for ridge formation, the physics that determines the distances to which the stresses are guided before causing a ridge – distances which also determine plate size – is not clear and remains a fruitful area of research.

4.8. Strike-slip margins: generation of toroidal motion

While the upwellings, downwellings, divergent and convergent zones in basic convection may or may not appear exactly in the plate-tectonic form of convection seen on the Earth, they at least do exist in both forms of convection. However, as mentioned earlier in this paper, while strike-slip motion (i.e., toroidal flow, which also includes plate spin) is very significant in the Earth's plate-tectonic style of convection – comprising possibly as much as 50% of the plates net kinetic energy, depending on the choice of reference frame [Hager and O'Connell, 1978, 1979, 1981; O'Connell et al, 1991] (see Plate 3) – it does not exist at all in basic convection with simple rheologies (i.e., with constant or depth dependent viscosity) and is very weak in convection with slightly more complex rheologies (e.g., with temperature-dependent viscosity). Thus, the correction to basic convection theory necessary to obtain strike-slip motion is nontrivial; this is because it is not merely a matter of adjusting the form of the existing convective (poloidal) flow field, but entails obtaining an entirely new flow field that would not otherwise exist.

A very basic but important aspect of toroidal motion is that it cannot occur in two-dimensional (2-D) models of convection (i.e. models with only x and z coordinates). Since toroidal flow occurs as vortex type motion in the $x-y$ plane (see Fig. 7), and since convection is driven by buoyancy forces which point in the z direction, a flow model with convection and toroidal motion needs all three directions; thus only 3-D convection models can obtain toroidal motion. Two-dimensional models of lithospheric motion in the $x-y$ plane are able to generate toroidal motion [Bercovici, 1993, 1995a], as discussed below, however these do not explicitly involve thermal convection. Another important criterion for plate-like toroidal motion is that it be manifest in narrow but intense bands of strike-slip shear which is what deformation on a fault would look like.

4.8.1. Imposing plates or faults One way of inducing toroidal motion in mantle flow is to apply plate motion on the surface of an isoviscous mantle; the imposed strike-slip motion at the surface forces toroidal motion in the fluid, or mathematically speaking, it provides a boundary condition to (5) that yields non-null values of ψ [e.g., Hager and O'Connell, 1978, 1979; Bunge and Richards, 1996]. This method of plate-driven mantle flow has yielded a variety of successful predictions, e.g., the observation that the flow lines of the forced motion correlates with slab dip [Hager and O'Connell, 1978]. A possibly more self-consistent approach is to drive the plate motion with buoyancy driven (poloidal) flow; the tractions on the base of a plate are first provided by the poloidal flow and then the resulting plate motion subsequently drives toroidal flow. This method can be used to infer the mantle's density and viscosity structure by solving for the internal mass anomalies and viscosity stratification that gives the observed plate motions, as well as the observed geoid and topography [Hager and O'Connell, 1981; Ricard et al., 1988, 1989, 1993; Ricard and Vigny, 1989; Ricard and Bai, 1991; Ricard and Froidevaux, 1991; Vigny and Ricard, 1991; Lithgow-Bertelloni and Richards, 1998]. It is also possible to use this technique to drive plates and toroidal motion with thermal convection [Gable et al., 1991]; models employing this method give insight into how toroidal and convective/poloidal flow co-evolve and co-exist. However, methods employing prescribed plate geometries cannot address the problem of how plates are generated.

A still further step toward self-consistent generation of plates is to impose faults or narrow weak zones in a lithosphere which either overlies a convecting mantle [Zhong and Gurnis, 1995a,b], or moves under its own pressure gradients [Zhong and Gurnis, 1996; Zhong et al., 1998; see also Gurnis et al., 1999, this volume]. This approach is largely motivated by the fact that many weak zones persist within the lithosphere for long times and are always available to be reactivated by convective forces; this concept is addressed below in §4.8.6 and in more detail in Gurnis et al. [1999, this volume]. In these models, plate-like motion is best generated by low-friction faults, and the direction of motion is neither solely determined by the convective (or plate) forces nor the fault orientation; lithospheric motion instead assumes some

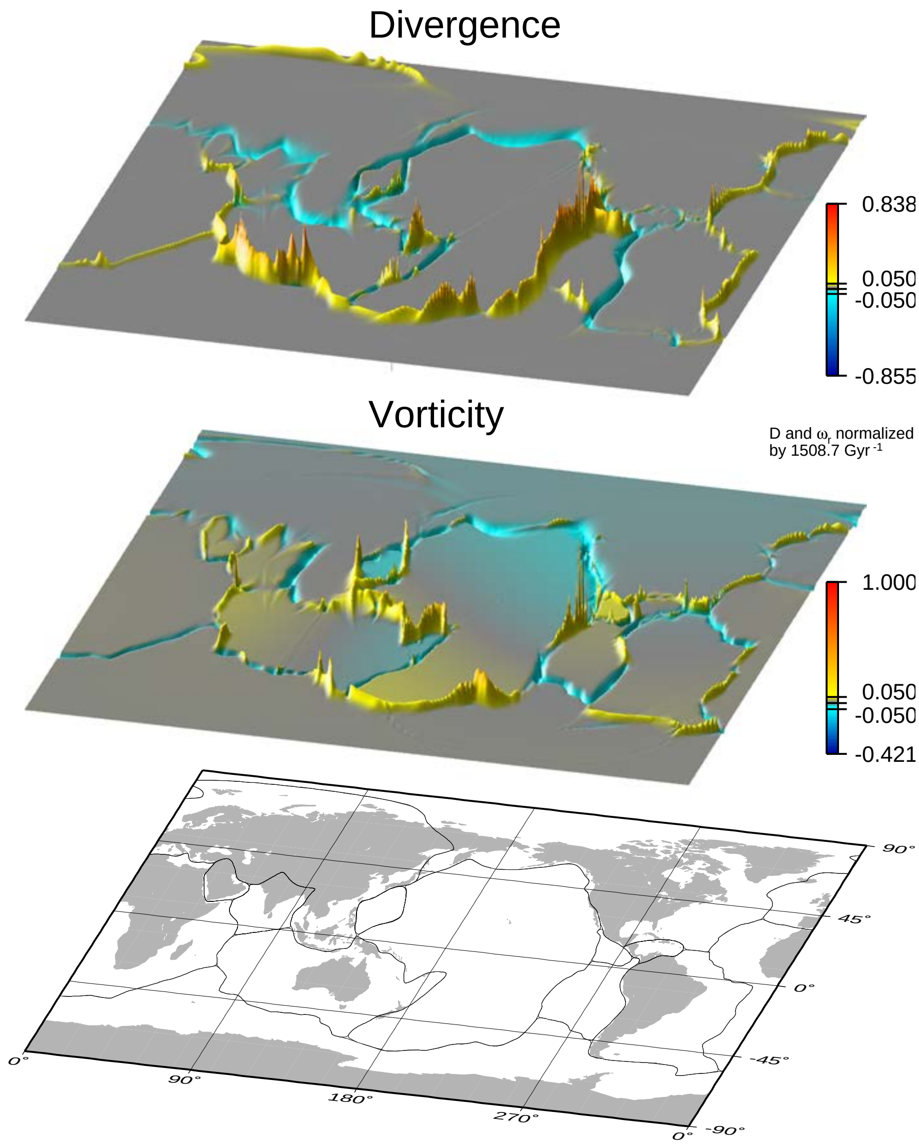

Plate 3. Approximate horizontal divergence (top) and vertical vorticity (middle) fields of the Earth (plate and continental outlines are shown for reference in the bottom frame). Horizontal divergence is $\nabla_h \cdot \underline{v} = \nabla_h^2 \phi$ (where $\underline{v}$ is the surface velocity); it is thus a fine-scaled representation of poloidal flow at the Earth's surface and represents the so-called rate of creation (divergence) and destruction (convergence or negative divergence) of the plates at ridges and subduction zones, respectively. Vertical vorticity is $\hat{z} \cdot \nabla \times \underline{v} = -\nabla_h^2 \psi$; it is thus a fine-scaled representation of toroidal flow and represents the rate of strike-slip shear between plates as well as plate spin (vorticity is essentially a measure of the local angular velocity, i.e., angular velocity of a point in the fluid). The divergence and vorticity fields shown are estimated directly from a model that uses finite-width plate margins based on seismicity distributions [Dumoulin et al., 1998]; for the standard plate model with infinitesimally thin margins, these fields are singularities and thus cannot be reliably estimated.

optimum configuration that takes best advantage of the orientation of both the forces and the faults (Plate 4). Generally, the motion will arrange itself so that strike-slip shear occurs along the low friction faults (see discussion below §4.8.4 on the purpose of toroidal flow) and thus the models generate significant toroidal motion.

However, the methods described above permanently prescribe the plate or fault geometries and thus these methods do not permit the plates or plate boundaries themselves to evolve from the convective flow. As with the 2-D convection models which strive to generate plateness, 3-D models that attempt to generate toroidal motion (as well as plateness) out of convective flow use a continuum approach with variable viscosity [e.g., Kaula, 1980; Christensen and Harder, 1991; Ribe, 1992; Cadek et al., 1993; Balanchandar and Yuen, 1995a,b; Bercovici, 1993, 1995a,b, 1996, 1998; Zhang and Yuen, 1995; Tackley, 1998, 1999 this volume; Weinstein, 1998; Trompert and Hansen, 1998].

4.8.2. Why is variable viscosity necessary for toroidal motion? The reason that the generation of toroidal motion in continuum models requires variable viscosity is somewhat mathematical and thus warrants some theoretical development. As shown earlier (§2.7), if viscosity is constant in viscous buoyancy-driven flows (with homogeneous boundaries), the toroidal motion is zero (see equation (5)). If viscosity μ is variable, then instead of the equation for toroidal

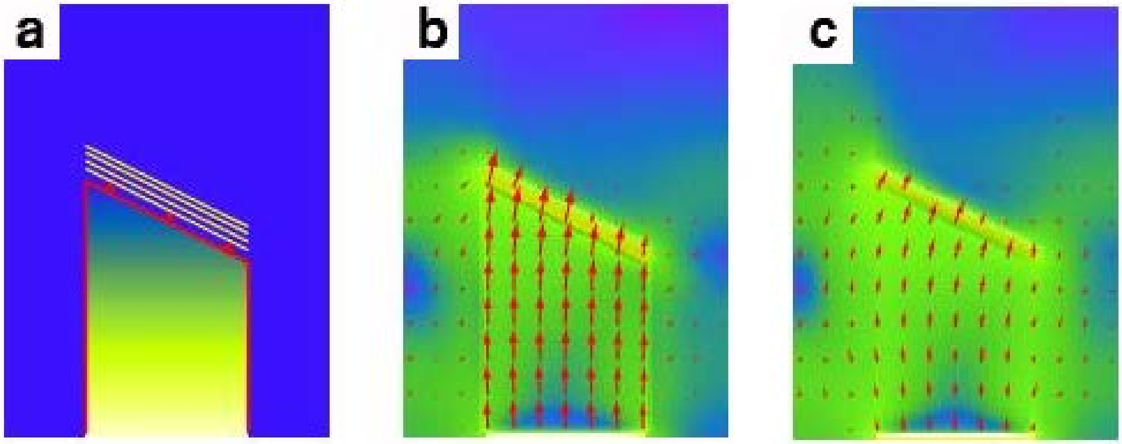

Plate 4. A model of lithospheric flow driven by "ridge push" (i.e., a lateral pressure gradient due to aging and thickening of the lithosphere) and slab pull with both non-Newtonian rheology and weak faults. Frame (a): the color shading shows lithospheric age (proportional to thickening), slab dip (yellow contours indicate slab depth) and imposed weak faults (red lines). Frame (b): arrows show velocity field and color shows viscosity for the case with weak faults. Frame (c): same as middle frame, but for the case without faults. The weak faults provide lubricated tracks that facilitate much more plate-like motion (middle frame) than without the faults (right frame). (After Zhong and Gurnis [1996]; see also Zhong et al. [1998].)

flow being homogeneous as in (5) we have

$$\begin{aligned}\mu\nabla^2\nabla_H^2\psi &= \hat{z}\cdot\nabla\mu\times\nabla^2\underline{v}\\ &\quad+\hat{z}\cdot\nabla\times\left(\nabla\mu\cdot(\nabla\underline{v}+[\nabla\underline{v}]^t)\right)\end{aligned} \tag{12}$$

(where $[..]^t$ implies transpose of a tensor). Thus the toroidal field is forced internally (i.e. within the fluid) if $\nabla\mu\neq 0$; if this condition is not met, then the right side of (12) is zero leading to a null solution for ψ, as in (5). This might lead one to believe that toroidal motion can be generated for any variable viscosity, including one that is dependent only on depth (which might occur because of the natural stratification and compression of mantle rock). However this is not the case. In order to guarantee that the toroidal field obtains some energy from convection, even indirectly, we require that the convective poloidal field ϕ appears on the right side of (12); i.e., the poloidal field gets energy from convective buoyancy (see (4)) and then passes on some of this energy to the toroidal field via viscosity gradients. However, the poloidal field ϕ does not appear at all in (12) if μ is only a function of height z. Therefore, the toroidal field only couples to thermal convection through the poloidal field if $\nabla_H\mu\neq 0$, i.e. if viscosity is at least horizontally variable.

The necessity for laterally variable viscosity can be understood qualitatively as well. If convective forces pull or push the upper thermal boundary layer, and part of this boundary layer is weak, then the weak area tends to get moved or deformed more easily than the neighboring stronger zones. The differential motion between the strong and the weak zones can lead to strike-slip type shear and toroidal motion (see Plate 2).

4.8.3. What variable viscosities give both plateness and toroidal flow? As noted already in this paper, viscosity is variable by virtue of being dependent on the thermodynamic state, state of stress and composition of the fluid. Nevertheless, a temperature-dependent viscosity by itself is probably not capable of generating the recquisite toroidal flow and plateness. Numerous studies of basic 3-D convection with temperature-dependent viscosity yielded little plate-like motion, usually with negligible plateness and toroidal kinetic energies on the order of 10% or less of the total [Christensen and Harder, 1991; Ogawa et al., 1991]. (Recall, toroidal motion on the Earth has possibly as much as 50% of the net kinetic energy of the plates.) The reason these models yielded little toroidal motion is likely due to the upper thermal boundary layer being too cold and stiff to permit any additional motion such as toroidal flow. It may also reflect the fact that since both viscosity and buoyancy are dependent on temperature, the viscosity and poloidal flow fields are in phase, whereas viscosity gradients orthogonal to the poloidal flow are probably necessary to generate toroidal flow (e.g., note in Plate 2 that strike-slip shear concentrates over viscosity gradients only because these gradients are perpendicular to the convergent flow in the upper buoyant layer). Thus, the necessary heterogeneities in viscosity likely arise from other nonlinear rheological mechanisms.

Given the success of creating respectable *plateness* from non-Newtonian power-law or plastic rheologies in 2-D convection models, it stands to reason that similar rheologies would generate significant toroidal motion in 3-D convection models. However, this appears not to be the case. Models using power-law rheologies (even with an excessively large power-law index $n > 3$) or plastic rheologies yield little plate-like toroidal flow (i.e., the strike-slip shear zones tend to be broad and diffuse) [Christensen and Harder, 1991; Bercovici 1993, 1995a; Cadek et al., 1993; Tackley, 1998; c.f. Trompert and Hansen, 1998].

It has been proposed that strain-softening type power-law rheologies are not sufficient to excite plate-like flow and instead that *self-lubricating* rheologies are necessary to generate such motion [Bercovici, 1993, 1995a]. A self-lubrication rheology involves a constitutive law that acts like a normal, highly viscous fluid at low strain-rates, but at larger strain-rates the viscosity weakens so rapidly with increasing strain-rate that the very resistance to flow itself – i.e., the stress – decreases [Whitehead and Gans, 1974]; this rheology is most simply represented by the constitutive law

$$\sigma \sim \frac{\dot{e}}{1 + \dot{e}^2/\gamma^2} \tag{13}$$

where γ is the strain-rate at which the stress σ reaches a maximum and begins to decrease with increasing strain-rate (Fig. 10). When employed in various flow models (which permit motion in the $x - y$ plane and thus toroidal flow) the self-lubricating rheology yields much more plate-like toroidal motion than do the power-law rheologies [Bercovici, 1993, 1995a; Tackley, 1998] (see Plate 5, Plate 6). In addition to generating sufficient toroidal motion, these models with self-lubricating rheologies also generate high plateness [Bercovici, 1993, 1995a; Tackley, 1998] and even suggest formation of relatively narrow passive upwellings [Tackley, 1998], something theretofore never accomplished in a fluid model. However, none of the models which have employed the self-lubricating rheology [Bercovici, 1993, 1995a; Tackley, 1998; see also Tackley, 1999 this volume] explicitly involve thermal convection, but instead prescribe the poloidal or buoyancy fields. Moreover, a technical difficulty with the self-lubricating rheology is that it can cause self-focussing of weak zones to sub-grid scale in numerical calculations [Tackley, 1998], thus the practical implementation of this rheology may require some arbitrary "clipping" of the effective viscosity field at low values. Nevertheless, these studies show that fluid mechanical models can produce plate-like toroidal motion, and that the self-lubricating rheology is reasonably successful at obtaining plate-like flows.

4.8.4. Why does toroidal motion form at all? Continuum models which generate toroidal motion from first principles allow us to address an important question: Why does toroidal motion form in the first place, and what role does it play? Since poloidal motion involves upwellings and downwellings, it can enhance its own energy uptake by increasing the rate at which gravitational potential energy is released; this feedback effect is of course the reason that heated layers go convectively unstable in the first place. However, toroidal flow only dissipates the energy it is given: since it involves purely horizontal motion it does not explicitly enhance gravitational energy release, nor does it, by the same token, transport heat and directly facilitate cooling of the fluid. By all appearances, the toroidal field is superfluous. And yet it is omnipresent in the Earth and is readily generated in fluid models with non-Newtonian (especially self-lubricating) rheologies.

The reason that toroidal motion is generated appears to relate to the thermodynamic efficiency of surface flow. Examination of the total viscous dissipation in the continuum models of Bercovici [1993, 1995a] show that even though the toroidal flow field is added on to the convectively driven poloidal field, it in fact makes the system *less* dissipative than if it were not present [Bercovici 1995b]; i.e., generation of toroidal motion *enhances* the efficiency of surface motion. This result can be simply interpretted to mean that the toroidal field establishes itself in focussed zones of strike-slip shear that act as lubricated motion guides. The bulk of the fluid outside of these zones of intense toroidal motion remain relatively undeformed and thus entail little dissipation. Within these zones deformation may be intense, but the viscosity is so low that the dissipation is small. Moreover, the zones of deformation are ideally so narrow and involve so little volume that they make little contribution to the total dissipation generated throughout the whole medium. This correlates with studies by Zhong and Gurnis [1996] and Zhong et al. [1998] which show that optimum convective flows occur when strike-slip motion concentrates on frictionless faults.

4.8.5. Plateness and toroidal motion are not completely related While variable viscosity is a requirement for both plateness and toroidal flow, it does not mean that they are manifestations of the same phenomenon. For example, mantle flow models with imposed faults may yield significant toroidal motion in the lithosphere if the faults are very weak; however, the plateness may become poor if the viscosity of the underlying asthenosphere is too high [Zhong et al., 1998]. This occurs because highly plate-like flows (i.e., with large plateness and toroidal motion) seem to occur if a weak zone or fault can decouple the lithosphere on one side of the weak zone/fault from the other. However, a viscous asthenosphere transmits stresses (that were normally decoupled across the fault) from one side of the fault to the other, causing the lithosphere to deform, weaken (via non-Newtonian behavior) and reach low plateness in response, even while the overall strike-slip motion remains significant. (A similar but less illustrative decorrelation of plateness and toroidal motion was found by Bercovici [1995a].)

4.8.6. Instantaneous and time-dependent rheologies and "true" self-lubricating mechanisms Up to now we have discussed the use of strain-rate- (or equivalently stress-) dependent rheologies, and amongst such rheologies the most successful one to date at generating plate-like motion appears to be the self-lubricating mechanism. However, there are two problems in this approach so far: 1) as pointed out by Zhong et al. [1998] and Gurnis et al. [1999, this volume], these strain-rate dependent rheologies are *instantaneous rheologies*, even though weak plate boundaries are not instantaneously formed; and 2) the self-lubricating rheology is an ad hoc parameterization of more complex physics. In fact, these two problems and their possible solution are closely related.

The explicitly strain-rate–dependent rheologies are called *instantaneous* because viscosity responds only to the instantaneous strain-rate field. Thus weak zones (what we would like to call plate boundaries) exist only as long as they are being deformed; once deformation stops, the boundaries cease. However, it is known that while new plate boundaries continue to be formed, old inactive ones persist as fabric in the

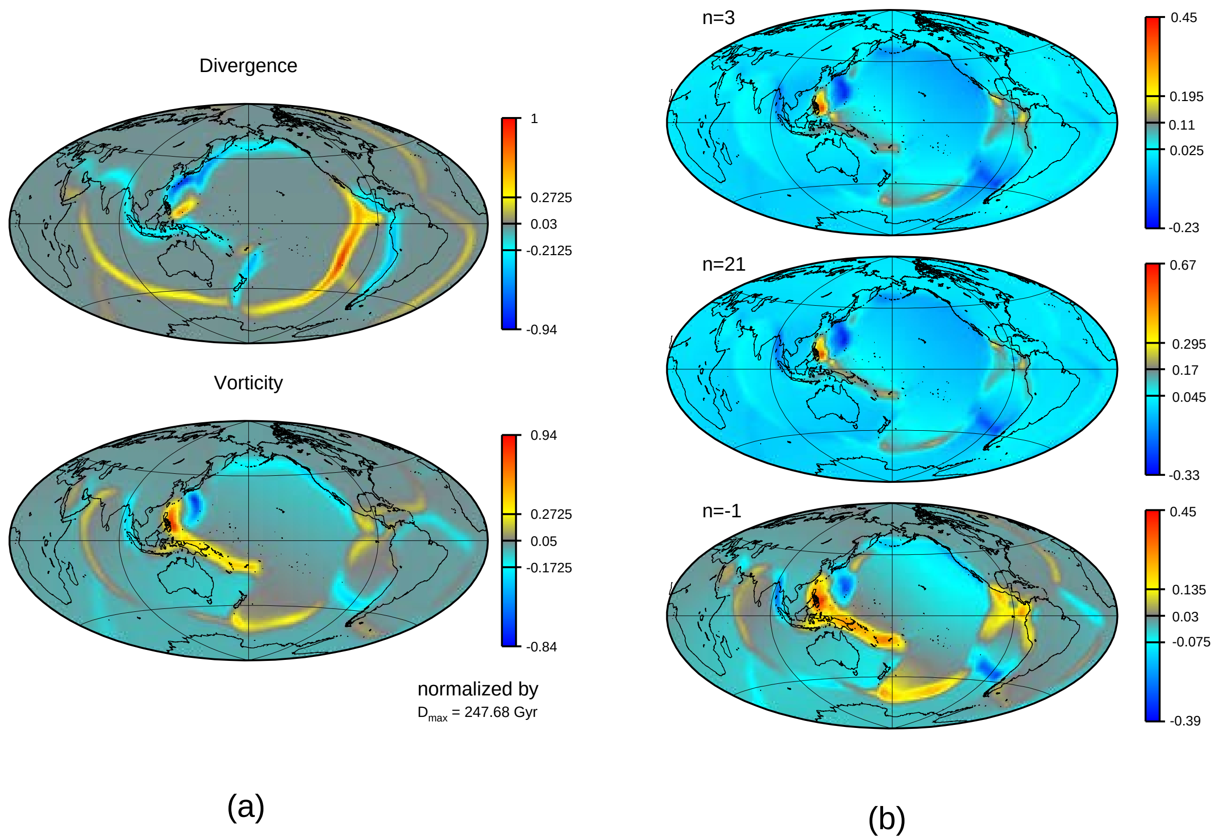

Plate 5. Generation of the Earth's toroidal motion via a source-sink model. The Earth's horizontal divergence (Fig. a – top) (a simpler and smoother version of that shown in Plate 3) is used as a source-sink field to drive flow in a non-Newtonian thin (lithospheric) fluid shell. The divergent zones (yellow and red) are used as sources of material, while the convergent zones (blue) are used as sinks. Ideally, the interaction of the source-sink field and non-Newtonian rheology in the lithospheric fluid shell would recover the Earth's vertical vorticity field (Fig. (a) – bottom) (again, simpler and smoother than that shown in Plate 3). Flow in the models with power-law rheologies (see equation (10)) do not generate Earth-like vorticity (Fig. (b) – top and middle frames) for power-law indices that are mantle-like, $n = 3$, and extreme, $n = 21$. Flow with the continuous stick-slip or self-lubricating rheologies (labelled as $n = -1$; see equation (13)) (Fig. (b) – bottom) generates a much more Earth-like vorticity field. See Bercovici [1995a] for further discussion.

lithospere, and even provide intrinsic weak zones that are preferred sites at which to re-activate deformation. Thus, boundaries have memory: they can remain weak even without being deformed, and can be moved around with the material in which they are embedded. This implies that the weakness is an actual transportable property, like temperature or chemical concentration. Moreover, such weak zones can persist for extremely long times, possibly much longer than the typical convective time scale of 10^8 years (a typical convective overturn time) [see Gurnis et al., 1998, this volume]; this suggests that the property which determines plate boundary weakness obeys a much longer time scale than does convection [see Bercovici, 1998].

That plate boundaries are caused by some weakening material property rather than an instantaneous rheological response in fact correlates with the fundamental origins of the the self-lubricating rheology. The rheological law expressed in (13) is actually a simplified, one-dimensional (1-D), steady-state representation [Whitehead and Gans, 1974; Bercovici, 1993] of the classical self-lubrication mechanism involving the coupling of viscous heating with temperature-dependent viscosity [Schubert and Turcotte, 1972]. In this self-lubrication mechanism, deformation causes frictional heating, which warms and weakens the material; the weak zone is more readily deformed causing a concentration of deformation, and therefore more heating, weakening, etc. In this case, a weak zone corresponds to a temperature anomaly, and thus temperature acts as the transportable time-dependent property whose history can be retained for a finite period of time.

However, when this viscous-heating-based, time-dep-endent self-lubrication mechanism is incorporated into flow models, it is not altogether successful at generating large-scale plate-like motions [Balachandar et al., 1995a,b; Zhang and Yuen, 1995; Bercovici, 1998]. This lack of plate motion likely occurs because temperature-anomalies due to viscous dissipation are generally not as important as the temperature anomalies due to the heating which is powering convection (e.g., heating from the core, or more importantly radiogenic heating). Thus, the temperature drop across the cold upper thermal boundary layer is usually much larger than any temperature anomaly due to viscous heating, causing this boundary layer to continue to be stiff and immobile. In some cases, viscous heating can be significant on small scales in convection, and, with temperature-dependent viscosity, can lead to shear-focussing and modest concen-

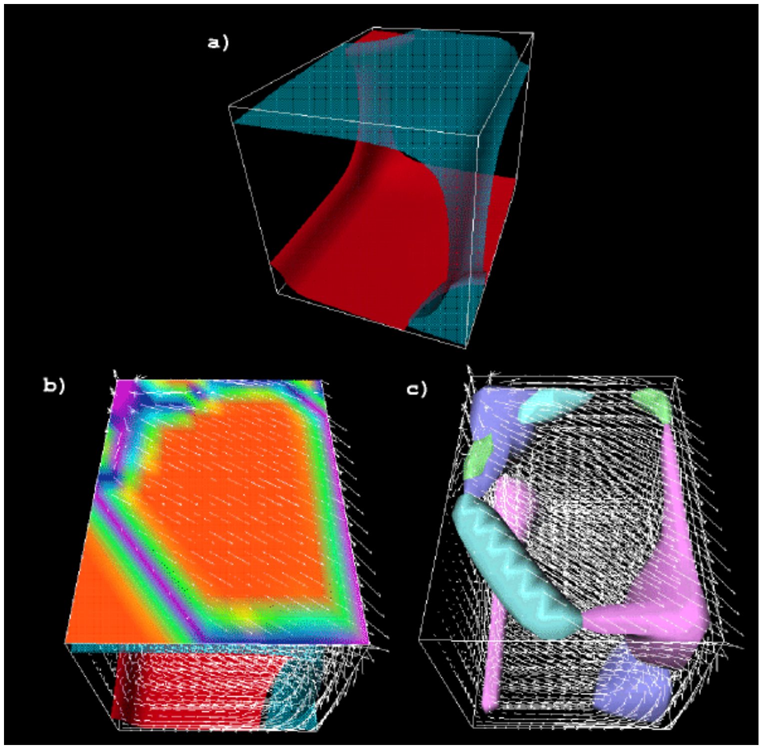

Plate 6. Generation of plate-like motion in lithospheric flow driven from below by a convective-type buoyancy field. Frame (a) shows the temperature field from simple basally heated convection. This is used as a static source of buoyancy to drive flow that then drags an overlying layer of fluid with stick-slip rheology (equation (13)). The sharp changes in viscosity (b) surrounding nearly isoviscous (red) blocks, along with nearly uniform velocity fields in these blocks, is suggestive of highly plate like flow. Frame (c) shows isosurfaces of horizontal divergence (poloidal flow) in the lithospheric layer (pink and purple), and vertical vorticity or toroidal flow (light blue and green). The elongated blue surface corresponds to a long, fault-like concentration of strike-slip shear. (After Tackley [1998]; see also Tackley [1999], this volume.)

trations of toroidal motion on a local scale [Balachandar et al., 1995a,b; Zhang and Yuen, 1995]; however, the surface motions and strength distributions created are not distinctly plate-like and the global toroidal energy is typically very small. Moreover, if the Earth's weak zones were simply temperature anomalies (and were fortuitously not overwhelmed by convective temperature anomalies), they would diffuse away on the order of a few million years, at most a few hundred million years (assuming plate boundaries are $10-100km$ wide and thermal diffusivity is $10^{-6}m^2/s$); i.e., their lifetime would be of the same order or shorter than the convective time scale, which is not observed [Gurnis et al., 1998, this volume]. Finally, if plate boundaries, in particular strike-slip margins, were caused by viscous heating then they should be perceptible with heat-flow measurements; however, this is generally though not to be the case [Lachenbruch and Sass, 1980; c.f., Thatcher and England, 1998].

Self-lubricating mechanisms involving other weakening properties are of course also possible. One well documented shear-localization mechanism involves grain-size dependent rheology and dynamic recrystallization [Karato, 1989]. Although this is a fairly complicated mechanism, the possible self-lubricating effect relies on the fact that at low stresses silicate rheology is largely controlled by diffusion creep where the viscosity depends on grain size (such that viscosity decreases with decreasing grain-size). In the process of diffusion creep, crystals metamorphose in response to tensile or compressive stresses; however, recrystallization and grain growth occurs up to a maximum crystal size which is itself inversely dependent on stress to some power. Thus in a region of higher stress, new crystals form with a smaller size causing the viscosity in the region to diminish. This zone of decreased viscosity represents a weakness on which deformation concentrates, causing higher stresses, thus even smaller recrystallized grain sizes, smaller viscosities, more stress concentration, and so on. (See Karato [1989] and Jin et al. [1998] for more rigorous explanations; see also Kameyama et al., [1997]; and Tackley [1999] this volume.) Shear localization due to this feedback mechanism is well documented in field evidence in continental lithospheric shear zones [e.g., Jin et al. 1998]. Theoretical models of shear zones incorporting this mechanism have yielded strain-localization [Kameyama et al., 1997], although the localization only occurs transiently and requires the addition of viscous heating and thermally activated viscosity. However, the time-scales controlling this mechanism, i.e., for grain growth, are relatively short [Karato, 1989] and thus probably cannot by themselves yield long-lived plate boundaries.

An attractive yet somewhat speculative alternative mechanism involves using water as the weakening agent; this of

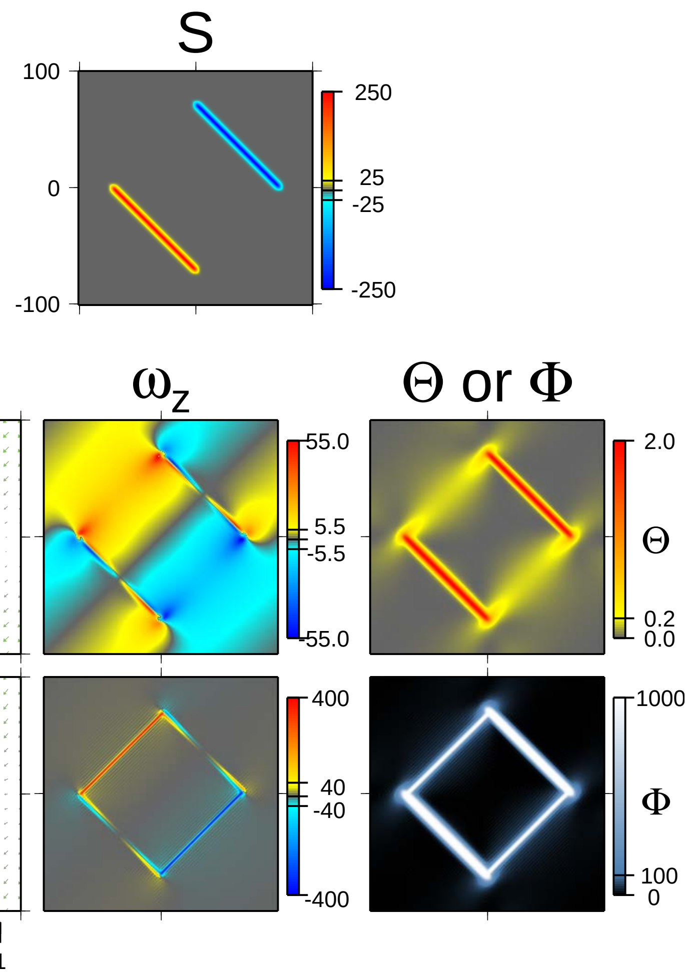

Plate 7. A simple model of a thin lithospheric-type layer of fluid driven by a basic source-sink field S (top row; yellow and red represent a source of material, blues represent sinks). The viscosity of the fluid layer is variable. In one model the viscosity is a function of temperature Θ and heat is generated by viscous dissipation (middle row showing, left to right, dimensionless horizontal velocity $\underline{v}_h$, vertical vorticity or strike-slip shear ω_z [see also Plate 5], and temperature Θ). In the second model viscosity is a function of volatile (e.g., water) content measured in terms of volume fraction of voids or porosity Φ; the voids are generated according to the amount that the fluid is stressed and damaged (bottom row, showing left to right, velocity, vorticity and porosity). The fields S, ω_z, Θ and Φ as shown are multiplied by a factor of 1000. The interaction of the source-sink field and the variable viscosity leads to toroidal flow. Toroidal motion (represented with vertical vorticity) and plate strength or *plateness* (measured by temperature or porosity distribution, where hot or porous is weak, cold and unporous is strong) remains diffuse and unplate-like in the viscous-heating based mechanism even with very temperature-dependent viscosity. However, toroidal motion in the void-volatile model can allow extremely focussed, nearly fault-like concentrations of strike-slip shear or toroidal motion (bottom middle frame) and plate-like strength distributions (bottom right frame). (After Bercovici [1998].)

course recalls the oft-stated hypothesis that plate tectonics only exists with the lubricating effects of water and water-laden sediments (and which also partially explains the lack of plate tectonics on the terrestrial planets devoid of water) [Tozer, 1985; Lenardic and Kaula, 1994, 1996]. Moreover, if the longevity of weak zones generated by ingestion of water are controlled by water's chemical diffusivity in rocks (which is several orders of magnitude slower than thermal diffusivity [Brady, 1995]), then such weak zones would survive for extremely long times, much longer than convective thermal anomalies [Bercovici, 1998]. However, in order for water to induce self-lubrication it must not only weaken the material into which it is ingested, but the water content must be affected by the deformation itself. Only in this way does one obtain the necessary shear-focussing feedback mechanism (similar to the viscous heating mechanism) whereby water content causes weakening, deformation concentrates on the weak zone, the enhanced deformation causes greater ingestion of water which causes further weakening, etc. The physics by which ingestion of water is controlled by rate of deformation is a necessary (albeit poorly understood) link and may occur when voids are created by damaging of the material [see Bercovici, 1998; Regenauer-Lieb, 1998 and references therein]; water (or any volatile) is then ingested into the voids thereby weakening the material (e.g. through pore pressure; Kohlstedt et al., 1995). This idea also draws on the fact that much of the lithosphere exists in a transitional region between brittle and ductile failure involving ductile cracking and void growth [Kohlstedt et al., 1995; Evans and Kohlstedt, 1995]. This *void-volatile* model of self-lubrication has, to date, only been used (with signficant simplifying assumptions) in an idealized mantle flow model;

nevertheless, this model yielded extremely plate-like behavior with narrow, concentrated shear zones highly suggestive of discontinuous strike-slip faults [Bercovici 1998] (Plate 7). However, this simple void-volatile model requires considerably more work and rigorous analysis. Indeed, while water likely plays a role in facilitating the formation and maintenance of plate tectonics, it is unclear over what depths within the lithosphere it is available to cause weakening [e.g., see Hirth and Kohlstedt, 1996].

In the end, it is unlikely that any one shear-focussing or self-lubrication mechanism will be the so-called magic bullet of plate boundary formation and plate generation. A distinct possibility may be that a combination of rheological mechanisms, such as brittle failure at shallow depths, ductile void growth and water/volatile ingestion at intermediate lithospheric depths, and grain-size dependent rheologies at greater lithospheric depths, yields the requisite shear-localizing behavior [see also Tackley, 1999, this issue]. Indeed, in this case, the longevity of plate boundaries may really only reflect the diffusion and healing rates of failed zones and volatile-filled voids at shallow depths; while these long-lived weak zones may not, therefore, penetrate the entire lithosphere, they may serve as initiation sites for shear-localization due to grain-size variations at greater depths. Unfortunately, the depth and material properties of weak zones associated with plate boundaries is poorly constrained. Many of these questions, from what shear-localizing mechanisms dominate within the lithosphere, to the depth of plate boundaries must be addressed observationally (e.g., with detailed, short-wavelength seismic studies).

5. FUTURE DIRECTIONS

Although the theory of plate tectonics has provided one of the greatest predictive tools in the solid-Earth sciences, it is by no means a complete theory. Plate tectonics is a kinematic model in that it describes motions, but it does not involve dynamics, i.e., either the forces or energy behind plate motions; nor does it explain the cause and generation of the plates themselves. In contrast, mantle convection theory describes and incorporates the essential dynamics behind plate motions, but is not yet sophisticated enough to predict from ab initio models the formation of the plates – with weak boundaries, strong plate interiors, asymmetric subuction, passive spreading centers, sudden and sometimes global changes in plate motions, and toroidal motion. A theory of mantle convection which generates plates – in essence a unified theory of mantle dynamics and plate tectonics – remains a lofty goal of geodynamics.

Considerable progress has been made in the last decade on obtaining fluid mantle flow models that yield plate-like behavior. In many ways, much of this success has come about by prescribing lithospheric weak zones and ad hoc rheologies (or rheologies like brittle failure that, while empirically based, are not applicable to the entire lithosphere) that yield more or less the correct behavior. While ad hoc formulations give important clues about what basic stress-strain-rate laws will yield plate-like behavior, they do not elucidate the underlying physics leading to these rheologies or to plate generation [see also Tackley, 1999, this issue]. In this sense, future progress will not necessarily come about by yet more convection modelling, but instead by an examination of the physical and microscopic mechanisms behind shear-localization and the depths and conditions at which these mechanisms occur. The areas of physics and chemistry involving melting and continent formation, the state and fate of water in the mantle, grain-growth, recrystallization, microcracking and void nucleation are but a few examples mentioned here that will play important roles in understanding the dynamics of plate generation. Moreover, observational, in particlar seismic data on the material properties and 3-D structure of plate boundaries (for boundaries of many ages, from ancient and inactive to incipient ones) is of vital importance in constraining dynamic models and microscopic theories of plate boundary formation. As seismic coverage of the Earth continues to improve with the deployment of ocean-bottom seismometers, the amount of such valuable information will increase tremendously.

In the end, it is clear that far from being tried, true and completed models, the theories of plate tectonics and mantle convection are in their youth, requiring years (if not decades) of research, as well as advances in microscopic physics, continuum mechanics, and observational geophysics of which we have yet to conceive.

Acknowledgments. This work was supported by NSF grant EAR-9458405, NASA HPCC Grant NCCS5-147, and the Centre National de la Recherche Scientifique (CNRS) of France. Figure 1 was made by Nicolas Coltice using ConMan (King et al., 1990) and GMT (Wessel and Smith, 1998).

References

Balachandar,S., D.A. Yuen, and D.M. Reuteler, Localization of toroidal motion and shear heating in 3-D high Rayleigh number convection with temperature-dependent viscosity, *Geophys. Res. Lett., 22*, 477-480, 1995a.

Balachandar,S., D.A. Yuen, D.M. Reuteler and G.S. Lauer, Viscous dissipation in 3-dimensional convection with temp-erature-dependent viscosity, *Science, 267*, 1150-1152, 1995b.

Beaumont, C., The evolution of sedimentary basins on a viscoelastic lithosphere, *Geophys. J.R. Astron. Soc., 55* 471–497, 1976.

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 simple model of plate generation from mantle flow, *Geophys. J. Int., 114*, 635–650, 1993.

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

Bercovici, D., On the purpose of toroidal flow in a convecting mantle, *Geophys. Res. Lett., 22*, 3107–3110, 1995b.

Bercovici, D., Plate generation in a simple model of litho-sphere-mantle flow with dynamic self-lubrication, *Earth Planet. Sci. Lett.*, *144*, 41–51, 1996.

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., G. Schubert and G. A. Glatzmaier, Three-dimensional, spherical models of convection in the Earth's mantle, *Science, 244*, 950–955, 1989a.

Bercovici, D., G. Schubert, G. A. Glatzmaier, and A. Zebib, Three-dimensional thermal convection in a spherical shell, *J. Fluid Mech.*, *206*, 75–104, 1989b.

Bercovici, D., G. Schubert and G. A. Glatzmaier, Three-dimensional, infinite Prandtl number, compressible convection in a basally heated spherical shell, *J. Fluid Mech.*, *239*, 683–719, 1992.

Bercovici, D. and P. Wessel, A continuous kinematic model of plate tectonic motions, *Geophys. J. Int.*, *119*, 595–610, 1994.

Bercovici, D. and A. Kelly, Nonlinear initiation of diapirs and plume heads, *Phys. Earth Planet. Int.*, *101*, 119-130, 1997.

Berg, J.C., A. Acrivos, and M. Boudart, Evaporative convection, *Advances Chem. Eng*, *6*, 61-123, 1966.

Boehler, Melting temperature of the Earth's mantle and core: Earth's thermal structure, *Ann. Rev. Earth Planet. Sci.*, *24*, 15–40, 1996.

Brady, J.B., Diffusion data for silicate minerals, glasses, and liquids, in *Mineral Physics and Crystallography: A Handbook of Physical Constants, AGU Reference Shelf 2*, American Geophysical Union, 269–290, 1995.

Bunge, H.-P. and M.A. Richards, The origin of large scale structure in mantle convection: effects of plate motions and viscosity stratification, *Geophys. Res. Lett.*, *23*, 2987–2990, 1996.

Bunge, H.-P., M.A. Richards, and J.R. Baumgardner, Effect of depth-dependent viscosity on the planform of mantle-convection, *Nature*, *379*, 436–438, 1996.

Bunge, H.-P., M.A. Richards, and J.R. Baumgardner, A sensitivity study of 3-D spherical mantle convection at 10^8 Rayleigh number: Effects of depth-dependent viscosity, heating mode, and an endothermic phase change, *J. Geophys. Res.*, *102*, 11991-12007, 1997.

Busse, F. H., Patterns of convection in spherical shells, *J. Fluid Mech.*, *72*, 67–85, 1975.

Busse, F. H., Non-linear properties of thermal convection, *Rep. Prog. Phys.*, *41*, 1929, 1978.

Busse, F. H. and Riahi, N., Patterns of convection in spherical shells. Part 2, *J. Fluid Mech.*, *123*, 283–301, 1982.

Busse, F. H. and Riahi, N., Mixed-mode patterns of bifurcations from spherically symmetric basic states, *Nonlinearity*, *1*, 379–388, 1988.

Busse, F.H. and J.A. Whitehead, Instabilities of convection rolls in high Prandtl number fluid, *J. Fluid Mech.*, *47* 305–320, 1971.

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

Cadek, O., Y. Ricard, Z. Martinec and C. Matyska, Comparison between Newtonian and non-Newtonian flow driven by internal loads, *Geophys. J. Int.*, *112*, 103–114, 1993.

Cazenave, A. and B. Lago, Long wavelength topography, sea floor subsidence and flattening, *Geophys. Res. Lett.*, *19*, 1257-1260, 1991.

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

Christensen, U.R., Heat transport by variable-viscosity convection and implications for the Earth's thermal evolution, *Phys. Earth Planet Int.*, *35*, 264–282, 1984a.

Christensen, U.R., Instability of a hot thermal boundary layer and initiation of thermo-chemical plumes, *Anal. Geophys.*, *2*, 311-314, 1984b.

Christensen, U.R., Convection with pressure- and temper-ature-dependent non-Newtonian rheology, *Geophys. J.R. Astron. Soc, 77*, 343–384, 1984c.

Christensen, U.R., Fixed hotspots gone with the wind, *Nature, 391*, 739-740, 1998.

Christensen, U.R., and H. Harder, Three-dimensional convection with variable viscosity, *Geophys. J. Int.*, *104*, 213-226, 1991.

Christensen, U.R., and D.A. Yuen, Layered convection induced by phase changes, *J. Geophys. Res.*, *90*, 291–300, 1985.

Creager, K.C. and T.H. Jordan, Slab penetration into the lower mantle, *J. Geophys. Res.*, *89* 3031–3049, 1984.

Colin, P., and L. Fleitout, Topography of the ocean floor: thermal convection of the lithosphere and interaction of the deep mantle heterogeneity with the lithosphere, *Geophys. Res. Letters*, *11*, 1961-1964, 1990.

Corrieu, V., Y. Ricard and C. Froidevaux, Radial viscosity of the Earth deduced from mantle tomography, *Phys. Earth Planet. Inter.*, *84*, 3-13, 1994.

Davies, G.F., Thermal histories of convective Earth models and constraints on radiogenic heat production in the Earth, *J. Geophys. Res.*, *85*, 2517–2530, 1980.

Davies, G.F., The role of the lithosphere in mantle convection, *J. Geophys. Res.*, *93*, 10451–10466, 1988a

Davies, G.F., Ocean bathymetry and mantle convection, 1. Large-scale flow and hotspots, *J. Geophys. Res.*, *93*, 10467–10480, 1988b.

Davies, G.F., Ocean bathymetry and mantle convection, 2. Small scale flow, *J. Geophys. Res.*, *93*, 10481–10488, 1988c.

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

Duffield, W.A., A naturally occurring model of global plate tectonics, *J. Geophys. Res.*, 77, 2543–2555, 1972.

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.

Duncan, R.A. and M.A. Richards, Hotspots, mantle plumes, flood basalts and true polar wander, *Rev. Geophys.*, *29*, 31–50, 1991.

Evans, B. and D.L. Kohlstedt, Rheology of rocks, in *Rock Physics and Phase Relations: A Handbook of Physical Constants, AGU Reference Shelf 3*, American Geophysical Union, 148–165, 1995.

Fleitout, L. and C. Moriceau, Short-wavelength geoid, bath-ymetry and the convective pattern beneath the Pacific Ocean, *Geophys. J. Int.*, *110*, 6–28, 1992.

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.

Forte, A.M. and W.R. Peltier, Plate tectonics and aspherical earth structure: The importance of poloidal-toroidal coupling, *J. Geophys. Res.*, *92*, 3645-3679, 1987.

Furbish, D.J., *Fluid Physics in Geology*, Oxford Univ. Press, 1997.

Gable, C.W., R.J. O'Connell, and B.J. Travis, Convection in three dimensions with surface plates: Generation of toroidal flow, *J. Geophys. Res.*, *96*, 8391–8405, 1991.

Giannandrea, E. and U. Christensen, Variable viscosity convection experiments with a stress-free upper boundary and implications for the heat transport in the Earth's mantle, *Phys. Earth. Planet. Int.*, *78*, 139–152, 1993.

Glatzmaier, G.A., Numerical simulations of mantle convection: Time-dependent, three-dimensional, compressible, spherical shell, *Geophys. Astrophys. Fluid Dyn., 43*, 223–264, 1988.

Glatzmaier, G.A., G. Schubert and D. Bercovici, Chaotic subduction-like downflows in a spherical model of convection in the Earth's mantle, *Nature, 347*, 274–277, 1990.

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

Gordon, R.G., C. DeMets, and J.-Y. Royer, Evidence for long-term diffuse deformation of the lithosphere of the equatorial Indian Ocean, *Nature, 395*, 370–374, 1998.

Gubbins, D. Energetics of the Earth's core, *J. Geophys., 43*, 453–464, 1977.

Guillou, L., J.-C. Mareschal, C. Jaupart, C. Gariépy, G. Bienfait, and R. Lapointe, Heat flow, gravity, and structure of the Abitibi belt, Superior Province, Canada: Implications for mantle heat flow, *Earth Planet. Sci. Lett., 122*, 103-123, 1994.

Gurnis, M., Large-scale mantle convection and the aggregation and dispersal of supercontinents, *Nature, 335*, 695–699, 1988.

Gurnis, M., Bounds on global dynamic topography from Phanerozoic flooding of continental platforms, *Nature, 344*, 754-756, 1990.

Gurnis, M., Phanerozoic marine inundation of continents driven by dynamic topography above subducting slabs, *Nature, 364*, 589–593, 1993.

Gurnis, M., and B.H. Hager, Controls of the structure of subducted slabs, *Nature, 335*, 317–321, 1988.

Gurnis, M., S. Zhong, and J. Toth, On the competing roles of fault reactivation and brittle failure in generating plate tectonics from mantle convection, *AGU Monograph on the History and Dynamics of Global Plate Motions*, edited by M. Richards, R. Gordon and R. van der Hilst, American Geophysical Union, 1999.

Hager, B.H. and R.W. Clayton, Constraints on the structure of mantle convection using seismic observations, flow models and the geoid, in *Mantle Convection*, ed. by W.R. Peltier, pp.657–763, Gordon and Breach, Newark, N.J., 1989.

Hager, B.H., and R.J. O'Connell, Subduction zone dip angles and flow driven by plate motion, *Tectonophysics, 50*, 111–133, 1978.

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

Hager, B.H., and R.J. O'Connell, A simple global model of plate dynamics and mantle convection, *J. Geophys. Res., 86*, 4843–4867, 1981.

Hager, B.H., Subducted slabs and the geoid: constraints on mantle rheology and flow, *J. Geophys. Res., 89*, 6003-6015, 1984

Hallam, A., Alfred Wegener and the hypothesis of continental drift, in *Scientific Genius and Creativity: Readings from 'Scientific American'*, W.H. Freeman & Co., pp77-85. (Reprinted from *Scientific American*, February, 1975.) 1987.

Hilde, T.W.C., S. Uyeda, and L. Kroenke, Evolution of the western Pacific and its margins, *Tectonophysics, 38*, 145–165, 1977.

Hirth, G. and D.L. 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.

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

Honda, S., D.A. Yuen, S. Balachandar, and D. Reuteler, Three-dimensional instabilities of mantle convection with multiple phase transitions, *Science, 259*, 1308–1311, 1993.

Ito, E., and E. Takahashi, Postspinel transformation in the system $Mg_2SiO_4 - Fe_2SiO_4$ and some geophysical implications, *J. Geophys. Res., 94*, 10,637–10,646, 1989.

Ito, E., M. Akaogi, L. Topor, and A. Navrotsky, Negative pressure–temperature slopes for reactions forming $MgSiO_3$ perovskite from calorimetry, *Science, 249*, 1275–1278, 1990.

Jarvis, G.T. and D.P. McKenzie, Convection in a compressible fluid with infinite Prandtl number, *J. Fluid Mech., 96* 515–583, 1980.

Jin, D., S.-I. 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.

Kameyama, M., D.A. Yuen and H. Fujimoto, The interaction of viscous heating with grain-size dependent rheology in the formation of localized slip zones, *Geophys. Res. Lett., 24*, 2523–2526, 1997.

Karato, S.-I., Grain growth kinetics in olivine aggregates, *Tectonophysics, 168*, 255–273, 1989.

Kaula, W.M., Material properties for mantle convection consistent with observed surface fields, *J. Geopys. Res., 85*, 7031-7044, 1980.

Kennet, B.L.N., and E.R. Engdahl, Travel times for global earthquake location and phase identification, *Geophys. J. Int., 105*, 429–466, 1991.

Kido M. and T. Seno, Dynamic topography compared with residual depth anomalies in oceans and implications for age-depth curves, *Geophys. Res. Lett., 21*, 717-720, 1994.

King, S. D., The viscosity structure of the mantle, *Reviews of Geophysics (Supplement): U.S. Quadrennial Report to the IUGG 1991-1994*, pp. 11-17, American Geophysical Union, Washington, DC., 1995.

King, S.D., A. Raefsky, and B.H. Hager, ConMan: Vectorizing a finite element code for incompressible two-dimensional convection in the Earth's mantle, *Phys. Earth Planet. Int., 59*, 196–208, 1990.

King, S.D. and B.H. Hager, The relationship between plate velocity and trench viscosity in Newtonian and power-law subduction calculations, *Geophys. Res. Lett., 17*, 2409–2412, 1990.

King, S.D., C.W. Gable and S. Weinstein, Models of convection driven plates: A comparison of methods and results, *Geophys. J. Int., 109*, 481-487, 1991.

King, S.D. and G.T. Masters, An inversion for radial viscosity structure using seismic tomography, *Geophys. Res. Lett., 19*, 1551–1554, 1992.

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

Kohlstedt, D.L., B. Evans and S.J. Mackwell, Strength of the lithosphere: Constraints imposed by laboratory experiments, *J. Geophys. Res., 100*, 17587–17602, 1995.

Lachenbruch, A.H., Dynamics of a passive spreading center, *J. Geophys. Res., 81*, 1883-1902, 1976.

Lachenbruch, A.H. and J.H. Sass, Heat flow and energetics of the San Andreas fault zone, USGS Open File Rep. 80-625, pp.47-146, 1980.

Lenardic, A., and W.M. Kaula, Self-lubricated mantle convection: Two-dimensional models, *Geophys. Res. Lett., 21*, 1707–1710, 1994.

Lenardic, A., and W.M. Kaula, Near surface thermal/chem-ical boundary layer convection at infinite Prandtl number: Two-dimensional numerical experiments, *Geophys. J. Int., 126*, 689–711, 1996.

Le Stunff Y., and Y. Ricard, Topography and geoid due to lithospheric mass anomalies, *Geophys. J. Int., 122*, 982-990, 1995.

Le Stunff, Y., and Y. Ricard, Partial advection of equidensity surfaces: a solution for the dynamic topography problem? *J. Geophys. Res., 102*, 24655-24667, 1997.

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

Lithgow-Bertelloni, C. and M.A. Richards, Cenozoic plate driving forces, *Geophys. Res. Lett., 22*, 1317-1320, 1995.

Lithgow-Bertelloni, C. and P.G. Silver, Dynamic topography, plate driving forces and the African superswell, *Nature, 395*, 269–272, 1998.

Liu, L.-G., in *Earth: Its Origins, Structure and Evolution*, M.W. McElhinny, ed., 177-202, Academic Press, San Diego, CA, 1979.

Loper, D.E. and I.A. Eltayeb, On the stability of D" layer, *Geophys. Astrophys. Fluid Dyn., 36*, 229–255, 1986.

Loper, D.E. and T. Lay, The core-mantle boundary region, *J. Geophys. Res., 100* 6397–6420, 1995.

Machetel and Weber, Intermittent layered convection in a model mantle with an endothermic phase change at 670km, *Nature, 350*, 55-57, 1991.

Mitrovica, J. X., C. Beaumont, and, G. T. Jarvis, Tilting of continental interiors by the dynamical effects of subduction, *Tectonics, 8*, 1079-1094, 1989.

Moresi, L. and V. Solomatov, Mantle convection with a brittle lithosphere: Thoughts on the global tectonic style of the Earth and Venus, *Geophys. J. Int., 133*, 669–682, 1998.

Morgan, W.J., Rises, trenches, great faults and crustal blocks, *J. Geophys. Res., 73*, 1959–1982, 1968.

Mueller, S. and R.J. Phillips, On the initiation of subduction, *J. Geophys. Res., 96* 651–665, 1991

Nataf, H.-C. and J. VanDecar, Seismological detection of a mantle plume? *Nature, 364*, 115–120, 1993.

Nicolis, G., *Introduction to Nonlinear Science*, Cambridge Univ. Press, 1995.

Norton, I.O., The 43Ma nonevent, *Tectonics, 14*, 1080-1094, 1995.

O'Connell, R.J., C.W. Gable, and B.H. Hager, Toroidal-poloidal partitioning of lithospheric plate motion, in *Glac-ial Isostasy, Sea Level and Mantle Rheology*, edited by R. Sabadini et al., pp. 535–551, Kluwer Academic, Norwell, Mass., 1991.

Ogawa, M., G. Schubert and A. Zebib, Numerical simulations of three-dimensional thermal convection in a fluid with strongly temperature-dependent viscosity, *J. Fluid Mech., 233*, 299–328, 1991.

Olson, P. and G.M. Corcos, A boundary layer model for mantle convection with surface plates, *Geophys. J.R. Astr. Soc., 62* 195–219, 1980.

Olson, P., Hotspots, swells and mantle plumes, in *Magma Transport and Storage*, M.P. Ryan, ed., John Wiley, New York, pp33-51, 1990.

Olson, P. and D. Bercovici, On the equipartitioning of kinetic energy in plate tectonics, *Geophys. Res. Lett., 18*, 1751-1754, 1991.

Oxburgh, E.R. and D.L. Turcotte, Mechanisms of continental drift, *Reports on Progress in Physics, 41*, 1249–1312, 1978.

Parmentier, E. M., Turcotte, D. L., & Torrance, K. E., Studies of finite amplitude non-Newtonian thermal convection with applications to convection in the Earth's mantle, *J. Geophys. Res. 81*, 1839–1846, 1976.

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

Parsons, B. and D.P. McKenzie, Mantle convection and the thermal structure of plates, *J. Geophys. Res., 83*, 4485–4496, 1978.

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

Phipps Morgan, J. and W.H.F. Smith, Flattening of the seafloor depth-age curve as a response to asthenospheric flow? *Nature, 359*, 524–527, 1992.

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

Ratcliff, J.T., P.J. Tackley, G. Schubert and A. Zebib, Transitions in thermal convection with strongly variable viscosity, *Phys. Earth. Planet. Int., 102*, 201–212, 1997.

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

Regenauer-Lieb, K., Dilatant plasticity applied to Alpine collision: Ductile void growth in the intraplate area beneath the Eifel volcanic field, *J. Geodynamics, 27*, 1-21, 1998.

Ribe, N.M., The dynamics of thin shells with variable viscosity and the origin of toroidal flow in the mantle, *Geophys. J. Int., 110*, 537–552, 1992.

Ricard, Y., L. Fleitout, and C. Froidevaux, Geoid heights and lithospheric stresses for a dynamic Earth, *Annales Geophysicae, 2*, 267-286, 1984.

Ricard, Y., and C. Froidevaux, Stretching instabilities and lithospheric boudinage, *J. Geophys. Res., 91*, 8314-8324, 1986.

Ricard, Y., C. Froidevaux and L. Fleitout, Global plate motion and the geoid: A physical model, *Geophys. J., 93*, 477-484, 1988.

Ricard, Y., and C. Vigny, Mantle dynamics with induced plate tectonics, *J. Geophys. Res., 94*, 17,543-17,559, 1989.

Ricard, Y., C. Vigny, and C. Froidevaux, Mantle heterogeneities, geoid and plate motion: a Monte Carlo inversion, *J. Geophys. Res., 94*, 13,739-13,754, 1989.

Ricard, Y. and Bai Wuming, Inferring the mantle viscosity and its three dimensional structure from geoid, topography and plate velocities, *Geophys. J., 105*, 561-571, 1991.

Ricard, Y. and C. Froidevaux, Seismic imaging, plate velocites and geoid, the direct and inverse problem, "Glacial isostasy, Sea-level and Mantle rheology", Sabadini and Lambeck eds, Kluwer Academic Publishers, 553-569, 1991.

Ricard Y., M. A. Richards, C. Lithgow-Berteloni and Y. Le Stunf, A geodynamic model of mantle mass heterogeneities, *J. Geophys. Res., 98*, 21895-21909, 1993.

Richards, M.A. and B.H. Hager, Geoid anomalies in a dynamic Earth, *J. Geophys. Res., 89*, 5487–6002, 1984.

Richards, M.A. and C. Lithgow-Bertelloni, Plate motion changes, the Hawaiian-Emperor bend, and the apparent success and failure of geodynamical models, *Earth Planet. Sci. Lett., 137*, 19–27, 1996.

Rumford, Count, Of the propagation of heat in fluids, *Complete Works, 1*, 239, American Academy of Arts and Sciences, Boston, 1870.

Schmeling, S., and Jacoby, W.R., On modeling the lithosphere in mantle convection with non-linear rheology, *J. Geophys., 50*, 89–100, 1981.

Schubert, G., C. Froidevaux and D.A. Yuen, Oceanic lithosphere and asthenosphere: thermal and mechanical structure, *J. Geophys. Res., 81*, 3525–3540, 1976.

Schubert, G., and D.L. Turcotte, One-dimensional model of shallow mantle convection, *J. Geophys. Res., 77*, 945–951, 1972.

Schubert, G., D.L Turcotte, and D.A. Yuen, Role of phase transitions in a dynamic mantle, *Geophys. J.R. Astron. Soc., 42*, 705–735, 1975.

Schubert, G., D. Stevenson, and P. Cassen, Whole planet cooling and the radiogenic heat source contents of the Earth and Moon, *J. Geophys. Res., 85*, 2531-2538, 1980.

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

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

Shen, Y., S.C. Solomon, I.Th. Bjarnason, and C.J. Wolfe, Seismic evidence for a lower mantle origin of the Iceland plume, *Nature, 395*, 62–65, 1998.

Silver, P.G., R.W. Carlson, and P. Olson, Deep slabs, geochemical heterogeneity, and the large-scale structure of mantle convection: an investigation of an enduring paradox, *Ann. Rev. Earth Planet. Sci., 16*, 477–541, 1988.

Sleep, N., Hotspots and mantle plumes: Some phenomenology, *J. Geophys. Res., 95*, 6715–6736, 1990.

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

Spiegelman, M. Geochemical consequences of melt transport in 2-D: The sensitivity of trace elements to mantle dynamics, *Earth Planet. Sci. Lett., 139* 115–132, 1996.

Stein, C.A. and S. Stein, A model for the global variation in oceanic depth and heat flow with lithospheric age, *Nature, 359*, 123–129, 1992.

Steinberger B., and R.J. O'Connell, Advection of plumes in mantle flow; implications for hotspot motion, mantle viscosity and plume distribution, *Geophys. J. Int., 132*, 412-434, 1998.

Stern, R.J., and S.H. Bloomer, Subduction zone infancy: Examples from the Eocene Izu-Bonin-Mariana and Jurassic California arcs, *Geol. Soc. Am. Bull., 104* 1621–1636, 1992.

Su, W.J. and A.M. Dziewonski, On the scale of mantle heterogeneity, *Phys. Earth Planet. Int., 74*, 29–54, 1992.

Tackley, P.J., D.J. 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.

Tackley, P.J., D.J. Stevenson, G. A. Glatzmaier, and G. Schubert, Effects of a multiple phase transitions in a 3-dimensional spherical model of convection in Earth's mantle, *J. Geophys. Res., 99*, 15877–15901, 1994.

Tackley, P.J., Effects of strongly variable viscosity on three-dimensional compressible convection in planetary mantles, *J. Geophys. Res., 101*, 3311-3332, 1996a.

Tackley, P.J., On the ability of phase transitions and viscosity layering to induce long-wavelength heterogeneity in the mantle, *Geophys. Res. Lett., 23*, 1985–1988, 1996b.

Tackley, P.J., Self-consistent generation of tectonic plates in three-dimensional mantle convection, *Earth Planet. Sci. Lett., 157*, 9-22, 1998.

Tackley, P.J., The quest for self-consistent generation of plate tectonics in mantle convection models, *AGU Monograph on the History and Dynamics of Global Plate Motions*, edited by M. Richards, R. Gordon and R. van der Hilst, American Geophysical Union, 1999.

Tao, W.C. and R.J. O'Connell, Deformation of a weak subducted slab and variation of seismicity with depth, *Nature, 361*, 626–628, 1993.

Tarduno J.A. and R.D. Cottrell, Paleomagnetic evidence for motion of the Hawaiian hotspot during formation of the Emperor seamounts, *Earth Planet. Sci. Lett., 153*, 171-180, 1997.

Thatcher, W. and P.C. England, Ductile shear zones beneath strike-slip faults - implications for the thermomechanics of the San Andreas fault zone, *J. Geophys. Res, 103*, 891–905, 1998.

Thomson, J., On a changing tesselated structure in certain liquids, *Proc. Phil. Soc. Glasgow, 13*, 464–468, 1882.

Thoraval, C., P. Machetel, and A. Cazenave, Locally layered convection inferred from dynamic models of the Earth's mantle, *Nature, 375*, 777–780, 1995.

Thoraval, C. and M.A. Richards, The geoid constraint in global geodynamics: Viscosity structure, mantle heterogeneity models, and boundary conditions, *Geophys. J. Int., 131*, 1-8, 1997.

Tozer, D.C., The present thermal state of the terrestrial planets, *Phys. Earth Planet. Int., 6*, 182–197, 1972.

Tozer, D.C., Heat transfer and planetary evolution, *Geophys. Surv., 7*, 213–246, 1985.

Trompert, R. and U. Hansen, Mantle convection simulations with rheologies that generate plate-like behavior, *Nature*, in press, 1998.

Turcotte, D.L. and E.R. Oxburgh, Finite amplitude convection cells and continental drift, *J. Fluid Mech., 28* 29–42, 1967.

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

Vail, P.R., R.M. Mitchum, Jr., and S. Thompson III, Global cycles of relative changes in sea level, *Amer. Assoc. Petroleum Geologists Memoir, 26*, 83-97, 1977.

van der Hilst, R., R. Engdahl, W. Spakman and G. Nolet, Tomographic imaging of subducted lithosphere below northwest Pacific island arcs, *Nature, 353* 37–43, 1991.

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

Verhoogen, J., *Energetics of the Earth*, National Academy of Sciences, 1981.

Vigny, C., Y. Ricard, and C. Froidevaux, The driving mechanism of plate tectonics, *Tectonophysics, 187*, 345-360, 1991.

Watts, A.B., G.D. Karner, and M.S. Steckler, Lithosphere flexure and the evolution of sedimentary basins, *Philos. Trans. R. Soc. London Ser. A, 305*, 249–281, 1982.

Weertman, J. and J.R. Weertman, High temperature creep of rock and mantle viscosity, *Ann. Rev. Earth. Planet. Sci., 3* 293–315, 1975.

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

Weinstein, S.A., Catastrophic overturn of the Earth's mantle driven by multiple phase changes and internal heating, *Geophys. Res Lett., 20*, 101–104, 1993.

Weinstein, S.A., Thermal convection in a cylindrical annulus with a non-Newtonian outer surface, *PAGEOPH, 146*, 551–572, 1996.

Weinstein, S.A., The effect of convection planform on the toroidal-poloidal energy ratio, *Earth Planet. Sci. Lett., 155*, 87-95, 1998.

Weinstein, S.A. and U. Christensen, The effects of a stress-free upper boundary on the convection planform of a fluid with a temperature-dependent viscosity, *Geophys. Res. Lett., 18*, 2035–2038, 1991

Weinstein, S.A. and P. Olson, The proximity of hotspots to convergent and divergent plate boundaries, *Geophys. Res. Lett., 16*, 433-436, 1989.

Weinstein, S.A. and P. Olson, Planforms in thermal convection with internal heat sources at large Rayleigh and Prandtl numbers, *Geophys. Res. Lett., 18*, 239-242, 1990

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

Wessel, P., and W.H.F. Smith, New, improved version of the Generic Mapping Tools Released, *EOS Trans. AGU, 79*, 579, 1998.

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

White, R.S. and D.P. McKenzie, Mantle plumes and flood basalts, *J. Geophys. Res., 100*, 17543–17585, 1995.

Whitehead, J.A., Convection models: Laboratory versus mantle, *Tectonophysics, 35*, 215–228, 1976.

Whitehead, J.A. and B. Parsons, Observations of convection at Rayleigh numbers up to 760,000 in a fluid with a large Prandtl number, *Geophys. Astrophys. Fluid Dyn., 19*, 201–217, 1978.

Whitehead, J.A., and R.F. Gans, A new, theoretically tract-able earthquake model, *Geophys. J. R. Astron. Soc., 39*, 11–28, 1974.

Wicks, C.W., and M.A. Richards, Depression of the seismic 670 km discontinuity beneath the Izu-Bonin and Marianas subduction zones, *Science, 261*, 1424–1427, 1993.

Wilson, J.T., Did the Atlantic close and then re-open? *Nature, 211*, 676-681, 1966.

Wolfe, C.J., I.Th. Bjarnason, J.C. VanDecar and S.C. Solomon, Seismic structure of the Iceland mantle plume, *Nature, 385*, 245–247, 1997.

Zhang, S. and D.A. Yuen, Intense local toroidal motion generated by variable viscosity compressible convection in 3-D spherical shell, *Geophys. Res. Lett., 23* 3135–3138, 1995.

Zhong S., and M. Gurnis, Mantle convection with plates and mobile, faulted plate margins, *Science, 267*, 838–843, 1995a.

Zhong S., and M. Gurnis, Towards a realistic simulation of plate margins in mantle convection, *Geophys. Res. Lett., 22*, 981–984, 1995b.

Zhong S., and M. Gurnis, Interaction of weak faults and non-Newtonian rheology produces plate tectonics in a 3D model of mantle flow, *Nature, 383*, 245–247, 1996.

Zhong, S., M. Gurnis, and L. Moresi, Role of faults, nonlinear rheology, and viscosity structure in generating plates from instantaneous mantle flow models, *J. Geophys. Res., 103*, 15255–15268, 1998.

D. Bercovici, Department of Geology and Geophysics, SOEST, University of Hawaii, Honolulu, HI 96822. (email: dberco@soest.hawaii.edu)

Y. Ricard, Laboratoire des Sciences de la Terre, Ecole Normale Supérieure de Lyon, 46 allée d'Italie, 69364 Lyon, Cedex 07, France. (email: Yanick.Ricard@ens-lyon.fr)

M.A. Richards, Department of Geology and Geophysics, University of California, Berkeley, CA 94720. (email: markr@seismo.berkeley.edu)

This preprint was prepared with AGU's LaTeX macros v5.01, with the extension package 'AGU^{++}' by P. W. Daly, version 1.6b from 1999/08/19.