in press, AGU Monograph on The History and Dynamics of Global Plate Motions, ed. M. Richards, 1999

The Quest for Self-Consistent Generation of Plate Tectonics in Mantle Convection Models

Paul J. Tackley

Department of Earth and Space Sciences, University of California, Los Angeles

Plate tectonics and mantle convection are different aspects of the same, coupled system, yet mantle convection calculations do not exhibit plate tectonic behavior unless it is imposed by the modeler. This paper explores the quest for self-consistent incorporation of plate tectonics into mantle convection models, and presents new results and parameterizations. Simulations of convection with buoyant, deformable continents but no oceanic plates show that supercontinent cycling is a possible, though not required, aspect of the dynamics, and that the presence of even non-episodic continents has a profound effect on the spectrum of mantle heterogeneity. Ductile shear localization may be the key process for generating weak plate boundaries. Possible localization mechanisms can be described using time-dependent evolution equations, allowing evolving lithospheric strength heterogeneity with long-term memory. Steady-state versions of these rheologies are used to demonstrate the formation of weak transform boundaries in a 2-D model, and convergent, divergent and strike-slip boundaries in 3-D models. However, the present models are far from realistic; future directions and some philosophical issues are discussed.

INTRODUCTION

Plate tectonics, with attendant continental motions, is arguably the most important manifestation of mantle convection. However, despite the sophistication and scale of present-day mantle convection models (e.g., [*Balachandar et al.*, 1996; *Bunge et al.*, 1997; *Tackley et al.*, 1994]), they do not exhibit this fundamental mode of behavior unless constrained to do so by the modeler. Plate-like behavior does not emerge from the type of viscous flow formulation which has commonly been used to model mantle convection, because it does not include the modes of lithospheric behavior responsible for strain localization, weakening and thus plate boundary formation. The quest for self-consistent plate tectonics, in which plates arise naturally from the material description rather than being imposed by the modeler, is regarded as a type of "Holy Grail" by mantle convection modelers. We are now at an exciting time, when mantle convection modelers are beginning to incorporate more realistic lithospheric rheologies into their models, and the first "self-consistent" models are emerging [*Bercovici*, 1996; *Tackley*, 1998a; *Trompert and Hansen*, 1998], although such models are still in their infancy and far from realistic. The transition from models with imposed or kinematic plates to models with self-consistent plates is somewhat analagous to (although not as fundamental as) the recent transition in geodynamo modeling between kinematic dynamo models and fully self-consistent dynamo models [*Glatzmaier and Roberts*, 1995; *Kuang and Bloxham*, 1997]. Questions about the nature and evolution of the terrestrial planets Earth, Venus and Mars, including possible transitions between plate tectonic and non-plate-tectonic regimes [*Davies*, 1995; *Moresi and Solomatov*, 1998; *Turcotte*, 1993], could then be answered in a unified and self-consistent framework rather than piecemeal changing of boundary conditions, etc.

This paper aims to assess the current state of research in the field. One component of plate tectonics is continents, so the paper begins by reviewing models which include continents but not oceanic plates, and presents some new 3-D models to investigate whether supercontinent cycles are an inherent characteristic of the dynamics. This is followed by an examination of the problems inherent in incorporating plate tectonics in convection models, a review of our (presently incomplete) knowledge of lithospheric rheology and mechanisms responsible for localization, a presentation of some simple parameterization and models to illustrate how this may lead to plate boundary formation, and finally a discussion of fundamental issues and problems that must be addressed in future research. In keeping with the spirit of this volume a large fraction of the discussion is review and synthesis, although the majority of the presented results are new and previously unpublished, as is much of the parameterization. This paper does not attempt to cover brittle processes in any detail, but rather focuses on ductile processes.

There has often been a tendency to regard plates and mantle convection as somehow separate and distinct entities, with their own sets of driving and dissipative forces, and usually with one 'driving' the other. However, in reality plates and convection are *one coupled system*, since to first order the plates (particularly oceanic plates [*Davies*, 1989]) are simply the cold upper thermal boundary layer of the convection, with the addition of some compositional complexity. Thus, the forces acting on/in the plates are *the same* as the forces acting on/in the mantle, the dominant driving forces being the negative buoyancy associated with subducted oceanic slabs. To put this succinctly: the plates *are* the convection, and this view provides the starting point for plate modeling in this paper.

Different Timescales, Different Modeling Approaches

Memory plays an important role in the evolution of plate tectonics, since 'new' plate boundaries usually form along long-lived 'pre-existing' zones of weakness, and plate boundaries, once formed, may persist for hundreds of millions of years, possibly changing from one form to another [*Scotese*, 1991; *Scotese et al.*, 1988]. The present state of the lithosphere is thus related to history spanning billions of years.

It is clear that a self-consistent treatment of plate tectonics must include such memory in the form of evolving lithospheric heterogeneity, a point which is discussed in some depth later in this paper. However, even a perfect treatment will not be useful for the purposes of performing transient (i.e., of order 100 million years or less) or instantaneous flow calculations related to a specific present-day or historical plate configuration, since the lithospheric heterogeneity will not have time to develop in such calculations. Thus, for transient or instantaneous calculations it is necessary to insert the plate boundaries and other strength heterogeneity by hand, and much has been learned by so doing, e.g., [*Ricard and Vigny*, 1989; *Zhong and Gurnis*, 1995; *Zhong and Gurnis*, 1996].

Self-consistent plate treatments are most applicable to long-term calculations, which run for billions of years (simulated time)– long enough that the character of the final result is independent of initial conditions. The goal is not to reproduce the exact configuration of plates, mantle heterogeneity, etc in the present-day Earth, but to obtain a "planet X" which resembles Earth in its statistical characteristics (e.g., spectrum of mantle heterogeneity, range of plate sizes).

Which type (timescale) of model is most valuable? The answer is that instantaneous, transient and long-term modeling approaches are complementary, giving different information, and thus all three are necessary and useful. A successful description of Earth's material properties and behavior must be able to satisfy constraints from all three types of modeling approaches. For example, a model which can produce an Earth-like planet X in a long-term simulation but does not work in modeling specific events in real Earth's history is clearly an incorrect model. Similarly, a model which successfully reproduces an event in Earth's history or the present pattern of geoid anomalies is clearly incomplete if, when run for billions of years, it produces a planet X that looks like Venus or Mars rather than Earth. These different approaches are all necessary and complementary.

Goals and Philosophy

The goal of self-consistent generation of plate tectonics is to identify the correct physical description which allows plates to form and evolve naturally in the convective system, with nothing being imposed by the modeler. After billions of years of evolution from an arbitrary initial condition the system should resemble, in a statistical way, Earth. The same material description should apply everywhere in the domain (except for differences due to compositional variations), with different modes of behavior coming in to play as a function of the temperature, composition, pressure, stress, and history.

This paper reviews and extends two aspects of plate tectonic modeling that are linked by the common thread of attempting to introduce more realistic rheology into mantle convection models, thus bringing them closer to the goal of self-consistency. The first aspect is continents: their treatment as compositionally- and rheologically-distinct material and the resulting behavior, while the second and main aspect is the formation of plate boundaries by strain localization and weakening. These two aspects are treated individually both because previous studies have done so, and in order to simplify the models for better understanding. A full model should clearly incorporate both aspects since they interact– for example, there is much debate about the relative roles of localized vs. distributed deformation in the deformation of continents at plate boundaries [*England and McKenzie*, 1982; *Houseman and England*, 1993; *Kong et al.*, 1997].

CONTINENTS AND SUPERCONTINENTS

Clearly, there are two types of lithosphere: oceanic and continental, and a plate may contain one or both types. A common modeling strategy is to treat only the continental parts, examining the dynamics of the coupled mantle-continents system without attempting a full treatment of plate tectonics (e.g., [*Gurnis*, 1988; *Zhong and Gurnis*, 1993]). Implicit in ignoring oceanic plates is the assumption that they passively adapt to the behavior inherent in the coupled continent-mantle system. This assumption is undoubtedly imperfect, but it is good experimental procedure to isolate and investigate one component of the system at a time. Pragmatic concerns also mandate this approach: we do not yet know how to include plate tectonics in models, as discussed later. In any case, such modeling can be performed using presently-available approaches, and can yield some important insights.

'Continents' contain buoyant, non-subductible continental crust with an average thickness of 45 km [*Mooney et al.*, 1998], and continental lithosphere. The lithosphere is of quite heterogeneous thickness, with very thin, tectonically-active regions and old, stable cratonic regions with 'roots' that may extend 300 km into the mantle [*Jordan*, 1975]. Continental mantle lithosphere is generally thought to be more highly depleted in crustal components (Ca and Al rich minerals) than oceanic lithosphere, rendering it buoyant [*Abbott et al.*, 1997; *Durrheim and Mooney*, 1994; *Pollack*, 1986]. This probably means that it is also more depleted with water, rendering it more highly viscous by as much as 2-3 orders of magnitude [*Hirth and Kohlstedt*, 1996; *Karato et al.*, 1986]. The hypothesized continental tectosphere [*Jordan*, 1978] may consist of such buoyant, viscous material.

The general modeling procedure is to insert continents, then let them move around self-consistently for long times. The inclusion of continents into mantle convection models was pioneered by [*Gurnis*, 1988; *Gurnis and Torsvik*, 1994; *Gurnis and Zhong*, 1991; *Zhong and Gurnis*, 1993], who included continents as rigid 'rafts' in two-dimensional geometry. Their models identified some basic cyclicity in continental motions with one or two continents. A similar rigid-block approach was used to model continents in 3-D [*Trubitsyn and Rykov*, 1995]. In 2-D, this approach was extended to include a parameterization of oceanic lithosphere, with rules governing the behavior and continental breakup [*Lowman and Jarvis*, 1993; *Lowman and Jarvis*, 1995; *Lowman and Jarvis*, 1996; *Lowman and Jarvis*, 1999].

The above models treat continents as idealized rigid blocks. [*Lenardic and Kaula*, 1995; *Lenardic and Kaula*, 1996] pioneered the treatment of continents in a more self-consistent manner: as fully dynamic buoyant, compositionally-distinct material, which satisfies the same equations of motion as the rest of the domain, closer to what they are in reality. Their models included continents using lithospheric and crustal material, in a two-dimensional domain with moderately temperature-dependent viscosity– variable enough to give the lithosphere some stiffness, but not enough to allow it to form a rigid lid (for a discussion of these see later). They found that the heat flow out of continents is relatively constant, and that the surface heat flux out of the continents scaled differently with convective vigor (Ra) than the heat flux out of oceanic regions, which helps to explain the Archean heat flux paradox [*Lenardic*, 1997].

New 3-D Results

Here, new results pertaining to the effect of fully dynamic continents in three-dimensional mantle convection are presented. The goal of these is to explore how buoyant, rheologically-distinct continent-like masses interact with and modify the 3-D thermal convection.

Continents. As with the models of [*Lenardic and Kaula*, 1996], the continents are represented as buoyant, compositionally distinct material which is self-consistently governed by the same equations as the mantle. Thus, they can deform, break up or aggregate without any additional 'rules' being necessary. The representation is different from that of [*Lenardic and Kaula*, 1996] in that we choose to make viscosity dependent on composition only, rather than including temperature-dependence, and that we treat only the continental lithosphere, not crust. These are simply idealizations made to simplify the system to its most basic level. Due to the lack of temperature-dependent viscosity, the continental material is made more viscous than the rest of the mantle. Pragmatically this is necessary so that the continents move around as somewhat contiguous units, as observed; realistically it is justified because the continental lithosphere has a higher viscosity than the mantle beneath (largely because it is colder but also because cratonic regions may be depleted in water, as discussed above), and deep continental cratonic roots have a higher viscosity than sub-oceanic mantle at the same depth (e.g., 100 km), for the same reasons. In summary, the models include high-viscosity continental blocks embedded in an otherwise constant-viscosity domain. As such, they should be regarded as a logical step towards self-consistency from 'rigid block' models, in which continents are modeled as rigid blocks in an otherwise constant-viscosity domain. However, they are not yet fully self-consistent since the continental material is inserted by the modeler rather than arising through mantle melting and differentiation, oceanic plates are not present, and rheology leading to localized deformation is not included.

*Parameters***.** We assume a Boussinesq mantle with constant viscosity and other properties, except that continental material is more viscous than non-continental 'mantle' (as discussed above) by a factor of 100. Ideally a higher viscosity contrast would be preferable, but this is limited by numerical difficulties in dealing with the large viscosity jump that occurs at the edge of a continent. An 8x8x1 Cartesian domain with periodic sides boundaries is assumed, with impermeable, free-slip upper and lower boundaries. The upper boundary is isothermal whereas the lower boundary is zero flux. The (temperature-based) Rayleigh number is taken to be 10^6, and the non-dimensional internal heating R is taken to be 20 [*Tackley*, 1996a].

The buoyancy of continental material is constrained by the requirement that continents remain a 'reasonable' thickness. If the buoyancy is too high the material eventually spreads out into a layer covering the entire domain, and if the buoyancy is too low the material accumulates into a big 'blob' extending ~1000 km deep into the mantle. The appropriate chemical buoyancy is that which balances the negative thermal buoyancy due the coldness of the continental material, which physically corresponds to Jordan's isopyctic hypothesis for the continental roots [*Jordan*, 1988; *Lenardic and Moresi*, 1999]. This may not be physically necessary in more 'realistic' calculations because strong temperature-dependent viscosity would give the necessary stiffness to the continents to stop them spreading or pooling. The ratio of chemical to thermal buoyancy is expressed as the buoyancy parameter B:

$$B = \frac{\Delta\rho_{c-m}}{\rho\alpha\Delta T} \tag{1}$$

where $\Delta\rho_{c\text{-}m}$ is the density difference between 'continent' and 'mantle', ρ is the mantle density, α is the thermal expansivity, and ΔT is the temperature difference across the mantle. In these cases B is equal to 2.0.

Numerical Method. The equations are solved using the finite-volume multigrid code STAG3D, detailed elsewhere [*Tackley*, 1994; *Tackley*, 1996a]. Continental material is tracked using clouds of tracer particles, giving zero diffusion. This method works well provided there are at least 10s of tracers per numerical cell [*Christensen and Hofmann*, 1994; *Tackley*, 1998b; *Van Keken et al.*, 1997]. The main numerical problem, which would also apply to a field approach, is entrainment of continents by downwellings at their edges. While some entrainment may be physically realistic, the rate of entrainment is found to be resolution dependent and therefore numerically enhanced. This entrainment limits the practical length of calculations to the time for which entrainment is not large, but this time appears to be long enough for interesting dynamics to occur.

Initial conditions. The initial temperature field is taken from a homogeneous-composition case (i.e., with no continental material) which has reached statistically steady-state. The continents are then inserted and the subsequent development followed. Continents initially have a constant thickness of 300 km and cover 1/3 of the surface area, similar to that on Earth. Two initial configurations are used: one square supercontinent or a 4x4 array of smaller continents.

Results. A number of cases have been run, which will be fully explored in a future paper. Here, two cases are chosen, which are intended to be illustrative rather than definitive. Both are entirely heated from within, differing only in the initial continental configuration.

The case with 16 initial continents is shown in Plate 1 at four different times. The initial convective pattern is typical of internally-heated Boussinesq convection, being dominated by closely-spaced, time-dependent linear downwellings [*Parmentier et al.*, 1994; *Tackley*, 1996b]. When continents are inserted the continental material and also the mantle immediately under it rapidly become cold due to thermal diffusion. However, deeper underneath the continent the mantle gradually heats up because it is not being cooled so effectively by downwellings. Downwellings preferentially occur at the edge of continents. In this case, continents show no tendency to aggregate into a larger unit. Instead, they jiggle around and eventually start breaking up into smaller units due to the small-scale convection associated with internally-heated convection. Continents clearly have a large effect on the thermal structure, particularly at shallow depth. The run was continued until entrainment of continental material became significant.

The case with a single initial supercontinent is a strong contrast to this (Plate 2). Mantle warming under the supercontinent causes long-wavelength stresses which break the continent apart (which apparently happens more rapidly in internally-heated cases than with basally-heated cases [*Lowman and Jarvis*, 1999]), spreading its fragments to the ‘corners’ of the box. Since the box is periodic, this corresponds to a supercontinent at the opposite side of the planet. The new supercontinent then stays there for a while before splitting up again and reforming at the original position. This strong oscillation persists for several cycles (the length of the simulation) and has a periodicity of ~1 Gyr, meaning a new supercontinent aggregation every 500 Myr (based on scaling a nondimensional velocity of 500, which is about twice the rms. surface velocity and half the peak surface velocity, to 5cm/yr), perhaps similar to the timescale observed for Earth’s continents [*Dalziel*, 1997]. It is clear from the rendering of temperature field that this cycle is inducing very long-wavelength heterogeneity into the mantle. There is also much small-scale structure which has the effect of breaking the continental material into small pieces, somewhat unsatisfactorily in terms of modeling Earth-like continents. It is probable that if realistic oceanic plates and lithospheric viscosities were included then these small-scale instabilities would be suppressed. Scaling arguments indicate that the viscosity of continental roots must be at least 1000 times higher than regular upper mantle for their long-term survival [*Lenardic and Moresi*, 1999].

For both cases, the effect on the lengthscales of heterogeneity is profound, as evident from the spectral heterogeneity maps (SHM) illustrated in Figure 1. These are contour plots of the rms. amplitude spectrum of the temperature field as a function of vertical coordinate, z [*Ratcliff et al.*, 1997; *Tackley et al.*, 1994], and are shown for the initial condition (top plot) and for each case. For each case the spectrum is time-averaged (i.e., the spectrum is calculated for 20-30 evenly-spaced frames during the calculation then averaged); individual snapshots may look more extreme. The isochemical initial condition (top frame) shows a fairly broad spectral peak centered at a spatial frequency of ~1.5, corresponding to a wavelength of 2/3 the depth of the box, roughly 2000 km. With multiple continents, the peak is sharper and shifted to roughly 3 times longer wavelength (frequency=0.5, corresponding to a wavelength of 2.0, ≈6000 km), equal to the average spacing of the continents. With a single initial continent the peak is at even longer wavelength, the longest that can be accomodated in this box, a wavelength of 8 (~24000 km). The peak amplitude is much larger with continents present, although this is not visible on the SHM.

Some features in these results are consistent with previous results in 2-D: these include the tendency of continents and mantle immediately below to become cold and for the sub-continental mantle to warm up [*Gurnis and Zhong*, 1991], for small continents to have less dynamical effect than large continents [*Zhong and Gurnis*, 1993], and the cyclicity that can be induced by supercontinent aggregation and dispersal [*Gurnis*, 1988; *Lowman and Jarvis*, 1993]. However, an important new finding is that initial condition does seem to play a role in the dynamics: if the continental material starts off as many small continents it displays no tendency to aggregate and a fairly static pattern is obtained, whereas if the material is somehow collected into a large supercontinent, strong episodicity can result and apparently persists. Basally-heated results (not presented here) display similar trends except that the supercontinent cycle dies out fairly quickly with small continental fragments settling above downwellings. Thus, it appears that supercontinent cycling is not a robust outcome of the dynamics; rather the system must be set up in a special way to induce it.

It is possible that the components of plate tectonics missing from these models, i.e., the oceanic plates, will strongly modulate the dynamics and change this conclusion. The motion of oceanic plates causes large-scale coherent motion and time-dependence which is absent in these results and may act to aggregate the continental material where it would otherwise just sit in the same place. Other factors that must be investigated in future include the effects of depth-dependent viscosity, and phase changes.

SELF-CONSISTENT GENERATION OF PLATE BOUNDARIES

Understanding the generation of weak plate boundaries is the key to understanding plate tectonics. From field and laboratory measurements we have a good idea of the lithospheric and crustal weakening and localization mechanisms which may be important at plate boundaries, however, we do not have the firm quantitative knowledge necessary to write down the relevant equations with certainty. Thus, basic research must focus on investigating the behavior and role of various mechanisms; it is not simply a question of writing a computer code to solve already known equations.

What are the criteria for judging the success of a model in generating plate tectonics? Perhaps the most fundamental is 'plateness', i.e., the ability to generate rigid plates surrounded by weak boundary regions, with most of the deformation taking place in the boundary regions [*Weinstein and Olson*, 1992], although these boundary regions may cover as much as 15% of Earth's surface area [*Gordon and Stein*, 1992]. The plate boundaries should come in three distinct types: spreading centers, single-sided subduction zones, and transform faults. Finally, the time evolution of the plates should bear a resemblance to that of Earth's plates [*Scotese*, 1991], including, for example, sudden changes in plate motion and jumps in plate boundaries [*Bonatti et al.*, 1994].

The following sections examine the basic reasons why plate generation does not arise in mantle convection models, summarizes the mechanisms which may be important at plate boundaries, presents some parameterizations and modeling results, and discusses directions for future work.

The Plate Problem

Mantle convection calculations with rheologies which are appropriate for the bulk of the mantle do not develop plate tectonics. Indeed, while it is clear that temperature-dependent viscosity is important for producing strong plates, this on its own arguably makes the system less Earth-like! Why is this? Plate 3 shows some simple calculations of mantle convection with increasing temperature-dependent viscosity contrast, taken from [*Ratcliff et al.*, 1997]. The system is Boussinesq and heated from below, with temperature-dependent viscosity being the only complexity. With constant viscosity (Plate 3a), upwellings and downwellings are symmetric and linear. With an intermediate viscosity contrast (100-1000) (Plate 3b), the system transitions to a 'sluggish lid' mode of convection, with huge cylindrical downwelling plumes and linear upwellings with plume instabilities [*Ratcliff et al.*, 1995; *Tackley*, 1993; *Trompert and Hansen*, 1996]). As viscosity contrast is increased further, the cold upper boundary layer becomes so stiff that it stops moving altogether and a rigid lid is formed [*Moresi and Solomatov*, 1995; *Ogawa et al.*, 1991; *Ratcliff et al.*, 1996]. This rigid lid is like a single plate covering the entire planet, which may be appropriate for Mars or Venus, but is not the Earth. A short-wavelength pattern of upwelling plumes and sheet-like downwellings transports heat to the base of the rigid lid, which it is conducted through. Non-Newtonian power-law rheologies display a similar trend, although the transitions occur at higher viscosity contrast [*Solomatov*, 1995]. So, temperature-dependent Newtonian or power-law rheologies which are appropriate for the interior of the mantle result in the formation of a rigid lid, not plates.

Despite this, some two-dimensional convection calculations have shown that strain-rate softening, non-Newtonian rheologies can cause weak zones (WZs) and strain rate localization above up- and down-wellings [*Christensen*, 1984; *Cserepes*, 1982], resulting in a rudimentary approximation of plates [*Weinstein*, 1996; *Weinstein and Olson*, 1992]. Why is this? In 2-D there are only convergent and divergent margins, which have concentrated buoyancy/stress sources below them to 'break' the rigid lid and focus deformation. Three-dimensional geometry (3-D) is fundamentally more challenging due to the transform (strike-slip) boundaries. There are no concentrated local buoyancy forces available to drive and localize transform boundaries; indeed, buoyancy in the mantle cannot directly drive them at all if the viscosity is purely depth-dependent [*Ribe*, 1992], as further discussed later. Thus, it is not surprising that the rheologies that produce reasonable 'plates' in 2-D do not work well in 3-D [*Christensen and Harder*, 1991; *Tackley*, 1998a; *Trompert and Hansen*, 1998; *Weinstein*, 1998]. Another point to note is that the experimentally-determined activation energies for Newtonian diffusion creep and powerlaw dislocation creep indicate that the Earth would certainly be in the rigid lid mode, if these were the only processes occuring [*Solomatov*, 1995].

Once plate boundaries do form by some mechanism, they play a dominant role in determining the form of the convection. Plate 3d shows a 'rigid lid' calculation where weak plate margins have been inserted at each end of the box, but otherwise the parameters are the same as in Plate 3c. As a result, the system behaves like an oceanic plate, moving from the left of the box to the right of the box, conductively thickening as it goes. Small-scale 'Richter-rolls' [*Richter*, 1973] are observed under the moving plate. Various researchers have investigated the influence of inserted plates on mantle convection, and much has been learned, for example [*Bunge and Richards*, 1996; *Zhong and Gurnis*, 1996]. However, these studies do not attempt to answer more fundamental questions regarding how and why plate boundaries form, what determines the sizes of plates, etc.

Toroidal and Poloidal Flow

The three-dimensional mantle flow velocity field can be divided into poloidal and toroidal components. Poloidal flow is associated with divergence or convergence in a horizontal plane, i.e., spreading centers and subduction zones, while toroidal flow is associated with transform (strike-slip) boundaries. Whereas poloidal flow is driven directly by buoyancy sources in the mantle, toroidal flow is coupled to buoyancy forces only by lateral variations in viscosity, or inhomogenous boundary conditions.

Curiously, the velocity field of Earth's plates has a toroidal component which is a substantial fraction of the poloidal flow [*Hager and O'Connell*, 1978]. Although the term "equipartitioning" is often used, the toroidal component is actually somewhat smaller than the poloidal component, with the amplitude ratio varying between about 0.25 and 0.5 (excluding net rotation) in the last 120 million years [*Lithgow-Bertelloni et al.*, 1993]. Where does the toroidal energy come from? One possibility is that it is generated inside the mantle by laterally varying viscosity; however, calculations have only been able to obtain a small toroidal:poloidal energy ratio this way, of order a few percent [*Balachandar et al.*, 1995; *Christensen and Harder*, 1991]. Another possibility is that it is the coupling of mantle buoyancy sources with the plates themselves which generates toroidal flow. Flow calculations with idealized plate geometries [*Gable et al.*, 1991] indicate that this is indeed the case, and show that the induced toroidal energy is maximum at the surface and decays with depth into the mantle, particularly rapidly if the lower mantle has a higher viscosity than the upper mantle. Thus, a high toroidal:poloidal ratio at the surface does not imply a high ratio deep in the mantle. These flow calculations as well as simple models [*Olson and Bercovici*, 1991] show that the toroidal:poloidal ratio is dependent on the ratio of transform boundary length to convergent/divergent boundary length.

Plate motions seem to organize themselves so as to roughly minimize the toroidal motion for a given plate geometry [*Lithgow-Bertelloni et al.*, 1993]. Thus, the significant fraction of toroidal motion is probably due to the geometries of present-day plates, rather than anything deep in the mantle. Then, of course, the question becomes: why do plates "choose" a geometry which requires a large amount of transform boundary? Toroidal motion does not transport energy, and since energy is dissipated in driving transform boundaries it is not clear why the system would arrange itself to drive them, rather than forming, for example, a system of rolls with purely poloidal motion. One idea is that the system is forced to have transform boundaries because of the sphericity: you can't fit a system of rolls onto a sphere, as you could in a box. Another idea is that continents disrupt the symmetry of the system, causing a more messy and imperfect system of plate boundaries. A more fundamental reason comes from considering dissipation of energy: it can be shown that, contrary to the simple reasoning above, transform boundaries may actually decrease the energy dissipation in the system, and therefore be beneficial [*Bercovici*, 1995a].

It should be noted that while toroidal energy is a strong feature of today's plate motions, it is not *necessary* for plate-like behavior. The most important characteristic of plates is minimal deformation in their interiors with strongly-deforming boundaries, i.e., 'plateness' as mentioned earlier. Plates with only convergent and divergent boundaries can have good plateness. The toroidal:poloidal energy reflects the ratio of strike-slip boundaries to convergent or divergent boundaries, and changes with time [*Lithgow-Bertelloni et al.*, 1993]. Thus, plateness rather than toroidal:poloidal ratio is the best indicator of plate-like behavior.

It should also be noted that, on a large scale, observed plate boundaries tend not to be purely convergent, divergent, or strike-slip, as simple models often assume, but generally contain a mixture of divergence (or convergence) and strike-slip motion. This is clear from the plate divergence and vorticity fields plotted in [*Bercovici and Wessel*, 1994]; for example the only major pure strike-slip boundary is the Pacific:North American boundary in the vicinity of the San Andreas fault. At a local scale, of course, the motion on these mixed boundaries is often accomodated by alternating segments of 'pure' motion, as at mid-ocean spreading centers, with their purely divergent ridge segments and purely strike-slip transform offsets.

Lithospheric Localization Mechanisms

Plate boundaries are zones of lithospheric weakness along which concentrated deformation takes place. The key processes, which are intrinsically linked, are *shear localization* and *weakening*. A number of weakening mechanisms are observed in nature and have been proposed as being important in the formation of plate boundaries. Some of these are applicable to the brittle (seismogenic) zone, while others are applicable to the ductile deformation zone.

A traditional explanation for plate boundaries is that the lithosphere 'breaks' by brittle failure all the way through. However, few material scientists now believe that brittle faults go all the way through the lithosphere. Between the brittly-deforming upper lithosphere (perhaps 10-20 km thick) and the viscously deforming lower lithosphere, lies a zone where ductile or semi-brittle deformation is dominant. Figure 2, taken from a review paper by [*Kohlstedt et al.*, 1995], summarizes the likely strength profile in the oceanic lithosphere. In the brittle regime, lithospheric strength is determined by the frictional stress of faults, proportional to normal stress (hence depth), whereas in the viscous regime, power-law viscous (dislocation) creep prevails. The oceanic lithosphere between about 10 km and 40 km depth is in the regime of ductile or semibrittle deformation, not brittle failure. This is the strongest part of the lithosphere; thus, ductile processes determine the strength of the oceanic lithosphere. To summarize, the emerging picture of plate boundaries is one of brittle failure in the upper ~10-12 km, with ductile shear zones in the middle lithosphere and ductile power-law flow in the lower lithosphere and upper mantle.

While the strongest part of the lithosphere may be deforming ductily, brittle faults may play a role in seeding plate boundaries if brittle failure occurs first (because it occurs at small strains), guiding the locations for ductile shear localization [*Gilbert et al.*, 1994]. Where oceanic lithosphere bends as it goes down a subduction zone, large earthquakes nucleating in the brittle region are observed to propagate deep into the ductile region [*Kanamori*, 1970; *Kikuchi and Kanamori*, 1995]. However, in some cases, ductile shear localization is observed to precede brittle failure [*Jin et al.*, 1998]. It may be that, in general, ductile lithospheric localization is responsible for driving localization in the shallower, brittle layers. [*Batt and Braun*, 1997] showed that such a model is successful in explaining crustal observations from continental compressional orogens in terms of deep single-sided lithospheric subduction.

Localization mechanisms which may be important in brittle and ductile regimes are now discussed.

Brittle Localization

While this paper focuses on ductile localization, it is useful to make some points regarding elasticity and brittle deformation, for completeness and clarification. Firstly, brittle failure occurs at low strains, and secondly, deformation of an elastic material with brittle or plastic failure naturally leads to localization of shear in narrow bands [*Poliakov and Herrmann*, 1994], without any strain weakening being necessary (as it is when deforming a viscous material).

Based on laboratory experiments, the friction on faults is history-dependent, becoming weaker with increasing strain, and stronger with increasing time [*Dieterich*, 1979]. This leads to "velocity-weakening". Thus, presently active faults are much weaker than faults that have been inactive for a long time. This process may be parameterized using an evolving state variable for the memory, leading to rate and state friction laws such as the Ruina-Dieterich and Dieterich-Ruina laws [*Ruina*, 1983]. In these laws, fault strength initially recovers rapidly once slip stops, followed by a more gradual recovery as time progresses. Such recovery may have an analog in ductile weakening mechanisms discussed shortly.

Another, continuum approach to the brittle zone comes from damage theory, in which a state variable representing 'damage' (nominally associated with the opening of microcracks into voids) evolves with strain and causes a reduction in the elastic moduli [*Lyakhovsky et al.*, 1993]. This approach leads to localization in crustal brittle zone and the development of fault networks [*Lyakhovsky et al.*, 1994]. Later another type of 'damage' treatment will be applied to the ductile regime.

Ductile Localization Mechanisms

Ductile shear zones in the crust and upper mantle are observed in the field [*Drury et al.*, 1991; *Jin et al.*, 1998; *Pili et al.*, 1997; *Vissers et al.*, 1995], so it is possible to examine the microstructures to constrain what mechanisms were responsible for the weakening and localization. Upper mantle shear zones are commonly observed in large massifs in orogenic belts [*Drury et al.*, 1991; *Vissers et al.*, 1995], a particularly notable, 170-km long example of which is observed in west Greenland [*Sorensen*, 1983]. When formed at moderate temperatures (i.e., less than about 950°C) and high stresses, these shear zones are narrow (from cm to kilometers width) and contain mylonites, a type of rock characterized by strong grain-size reduction and (commonly) hydration [*Drury et al.*, 1991; *Pili et al.*, 1997]. Shear zones are often characterized by very small grains, hydration, and elevated temperatures. Thus, the possible mechanisms for their formation are thought to be grain-size reduction, volatile ingestion, and viscous heating. These are now examined.

Viscous heating. This was the first proposed method for shear localization. Viscous dissipation increases temperature, which reduces the viscosity, a positive feedback mechanism leading to higher strain rate in the weakened ration, thus more dissipation and even higher temperatures. This heating is balanced by thermal conduction of heat away from the shear zone. Two-dimensional calculations have been used to investigate this [*Fleitout and Froidevaux*, 1980; *Yuen et al.*, 1978] and show that for constant-stress boundary conditions, thermal runaway can occur and the solution diverges to infinite velocities. For more realistic constant velocity boundary conditions, however, the system reaches an equilibrium shear-zone thickness in which conduction is balanced by heat production and the viscosity is about 10^{21}, much lower than typical lithospheric viscosities. There is evidence for this mechanism from observations [*Jin et al.*, 1998; *Obata and Karato*, 1995], and from heat flow measurements along fault zones [*Thatcher and England*, 1998].

Grain size reduction. The most cited mechanism for forming shear zones, this is commonly observed in the field [*Drury et al.*, 1991; *Jaroslow et al.*, 1996; *Jin et al.*, 1998; *Vissers et al.*, 1995]. Grain size reduction occurs when rock is deforming by dislocation creep, in which new, very small grains form at the boundaries between existing grains, and creep is non-Newtonian but independent of grain size. If the average grain size is reduced sufficiently, a transition may occur to diffusion creep (by diffusion of vacancies through the lattice), in which viscosity is a strong function of grain size [*Ranalli*, 1995], and thus viscosity may be reduced by several orders of magnitude, allowing localization. Competing with grain-size reduction is grain growth by annealing, which is highly temperature-dependent. These processes are well quantified through laboratory experiments [*Karato*, 1989; *Karato et al.*, 1980; *Karato et al.*, 1986], in terms of the rate of grain growth, size of recrystallized grains, and domain diagrams. However, the exact evolutionary path of the system is uncertain: dynamic recrystallization does not occur in the diffusion creep regime, so once this region is reached, subsequent evolution must involve both diffusion and dislocation creep. More research is required as to how localization occurs through this mechanism [*Jin et al.*, 1998; *Kameyama et al.*, 1997].

Void-Volatile Weakening. This mechanism was recently proposed by [*Bercovici*, 1998], based on field observations and ductile void growth theory [*Regenauer-Lieb*, 1999]. The basic idea is that voids produced in the deforming rock can then ingest fluid which reduces its viscosity.

It should be noted that these mechanisms are not mutually exclusive; indeed, all three are probably operating to some extent. The important question is which mechanism is dominant, the answer to which could depend on depth (perhaps void-volatile weakening is important at shallow depth, with grain-size reduction dominant at greater depth) and/or tectonic environment. Some researchers favor viscous dissipation as the dominant mechanism at all depths [*Ranalli*, 1995]. This is an area where coupled numerical modeling, field observations and laboratory experiments are necessary.

Evolution Equations

Each weakening mechanism can be described by a variable (e.g., average grain size, temperature, or porosity) which evolves according to some equation, the prototypical form of which is:

$$\frac{dD}{dt} = S(\sigma,\dot{e}) - H(D,T) - \kappa.\ {}^{2}D \qquad (2)$$

where S is the source term, H is the healing (annealing) term and the final term is diffusive. 'D' represents 'damage', which could be either grain size (or inverse grain size if higher D is taken to indicate greater weakening), temperature or porosity.

To complete the parameterization, increasing 'damage' reduces the effective viscosity η_{eff}, as:

$$\eta_{eff} = \eta(D,T,p,\dot{e}) \qquad (3)$$

where p is pressure and $\dot{e}$ is strain rate. These equations are coupled since viscosity determines the relationship between strain rate and stress, which both appear in the source term of (2). Note the distinction between this damage, which applies in the viscous creep regime, and the damage discussed by [*Lyakhovsky et al.*, 1993], which affects elastic moduli.

What form should these terms take? This depends on the mechanism under consideration. The following table summarizes the physical meaning of damage (D), possible forms (proportionalities) of source (S), healing (H), diffusive and viscosity terms for the three mechanisms, taken from [*Govers and Wortel*, 1995; *Kameyama et al.*, 1997; *Karato et al.*, 1980] for grain-size reduction and [*Bercovici*, 1998] for void-volatile weakening.

Weakening Mechanisms

Mech	*D*	*S*	*H*	κ	η_{eff}
VH	T	$\sigma\dot{e}$	0	√	exp(E/T)
GSR	1/d	$\dot{e}D\left(1-Da_r\right)$	D^3.exp(-E/T)	~0	D^{-2} or D^{-3}
V-V	Φ	$\sigma^2/(1-D)$	D.exp(-E/T)?	~0	(1-λD)

E is activation energy, d is grain size, Φ is porosity and a_r is the recrystallized grain size, which is related to stress. The $\sigma\dot{e}$ form of source term is appealing because this is the work being done on the system. A healing term proportional to D leads to exponentially-decaying damage, if there is no damage production.

The dependence of healing rate on temperature for GSR and VV mechanisms allows the cold upper lithosphere to retain weakness for billions of years while the warm mantle adjusts strength instantaneously. For grain-size reduction, for example, the grain growth rate has similar temperature-dependence (activation energy) as diffusion creep rate ($1/\eta$) [*Karato*, 1989].

The viscosity term listed for GSR and VV mechanisms should be multiplied by the 'undamaged' viscosity to get the full viscosity law, for example, for VV weakening [*Bercovici*, 1998],

$$\eta = (1-\lambda D).\eta_u(T,p,\dot{e}), \qquad (4)$$

where η_u is the viscosity of undamaged material and λ is the rate of change of viscosity with damage, ≤1 in the analysis of [*Bercovici*, 1998].

There are several possible meanings of 'weakening': viscosity reduction, plastic yield stress reduction, or elastic modulus reduction. The above equations assume that viscous creep is the dominant mechanism. If the mechanism is instead plastic yielding, then weakening may occur in the form of reduction of the plastic yield stress, rather than reduction of viscosity. This possibility is implicitly included in the parameterization if η_u is already an effective viscosity due to plastic yielding.

Strain Weakening to Strain-Rate Weakening

Clearly, the above descriptions satisfy the requirement for having memory. However, including all three mechanisms in their time-dependent form would lead to a rather complicated and unwieldy model, so for initial investigations some simplification is appropriate. One simplification is to remove the time-dependence by deriving a 'steady-state' version of the weakening, i.e., deriving the stress vs. strain rate relationship for constant strain rate. This can be obtained by setting dD/dt=0 and solving the coupled evolution and viscosity equations for stress as a function of strain rate and other parameters. Examples of this for thermal weakening and void-volatile weakening have been derived in [*Bercovici*, 1996; *Bercovici*, 1998], and show a very characteristic form: stress increases with strain rate up to some critical point, beyond which stress *decreases* with increasing strain rate, a behavior which we may call strain rate weakening (SRW), and has something of an analog in the brittle zone with velocity weakening on faults. To summarize, when strain weakening is combined with time-dependent healing, for the viscous dissipation or void-volatile mechanisms, the result is strain rate weakening. Curiously, SRW does not arise when this steady-state analysis is applied to a simple parameterization of grain-size reduction in the absence of any other mechanism [*L.G.J. Montesi*, *personal communication*, 1998]: some additional effect such as slow cooling [*Braun et al.*, 1999] or viscous dissipation [*Kameyama et al.*, 1997] must also be present. An example of SRW is now derived (most applicable to the void-volatile mechanism- for other mechanisms see the above-referenced papers), followed by a discussion of timescales and its applicability to model lithospheric localization.

To show how strain rate weakening arises, consider the simple scalar damage evolution equation:

$$\frac{dD}{dt} = a\sigma\dot{e} - bD, \tag{5}$$

where

$$\sigma = \eta_{eff}\dot{e} = (1 - D)\eta_u(T, p)\dot{e} \tag{6}$$

η_u, the viscosity for 'undamaged' material, is here taken to be Newtonian. To derive the stress for steady-state deformation (constant strain-rate), the time derivative in (5) is set to zero and (6) substituted into (5):

$$0 = a\eta_u(1 - D)\dot{e}^2 - bD \tag{7}$$

which leads to a solution for D of:

$$D = \frac{a\eta_u\dot{e}^2}{a\eta_u\dot{e}^2 + b}. \tag{8}$$

Finally, the stress is:

$$\sigma = \eta_{eff}\dot{e} = \eta_u(1 - D)\dot{e} = \frac{\eta_u\dot{e}}{1 + \frac{a}{b}\eta_u\dot{e}^2} \tag{9}$$

which is similar to expressions derived by [*Bercovici*, 1998]. Note the asymptotic limits: the usual Newtonian law at low strain rates but stress reduction at high strain rates:

$$\sigma \rightarrow \begin{cases} \eta_u\dot{e} & \text{as } \dot{e} \rightarrow 0 \\ b/a\dot{e} & \text{as } \dot{e} \rightarrow \infty \end{cases} \tag{10}$$

This strain-rate weakening is compared to stress:strain-rate relationships for other well-known rheologies in Figure 3. All rheologies have been normalized such that σ=1 for $\dot{e}$=1, as follows:

$$\sigma = \eta_{eff}\dot{e} = \begin{cases} \dot{e} \\ \dot{e}^{\frac{1}{3}} \\ \min(\dot{e}, 1) \\ 2\dfrac{\dot{e}}{1 + \dot{e}^2} \end{cases} \tag{11}$$

where these apply to Newtonian, power-law (n=3), plastic yielding and strain-rate weakening respectively. From the lower plot of η_{eff} vs. strain rate, it is apparent that for all except Newtonian creep, the viscosity decreases with increasing strain rate. Indeed, from this plot one is not aware of any fundamental difference between the rheologies, except that SRW has greater strain-rate softening. The fundamental difference between the rheologies is, however, clearly apparent from the stress vs. strain rate curves: for n=3 the stress always increases for increasing strain rate, for plastic yielding the stress remains constant for increasing strain rate – only for SRW does the stress actually decrease for increasing strain rate, and as will be shown later, this is the key to obtaining localization and weak plate boundaries.

How do the critical stress and strain rate scale with various parameters? One can derive these by differentiating the stress vs. strain-rate relationship above, finding that:

$$\dot{e}_{\max} = \sqrt{\frac{b}{\eta_u a}} \quad ; \quad \sigma_{\max} = \sqrt{\frac{\eta_u b}{a}} \tag{12}$$

So, higher η_u leads to higher σ_{max} but lower $\dot{e}_{max}$. If a and b were constant, this would lead to a σ_{max} which is maximum at the top of the lithosphere and decreases with increasing depth, the opposite trend to the usual depth-dependent yield stress used to parameterize brittle processes. In general, however, b is temperature-dependent, such that it increases with increasing temperature. If b has the inverse temperature-dependence as viscosity, which may be approximately the case for grain growth, then the product $\eta_u b$ is constant and σ_{max} is independent of temperature. Thus σ_{max} constant and σ_{max} decreasing with depth may be regarded as reasonable limiting cases.

At this point is must be emphasized that σ_{max} is not synonymous with yield stress, since yield stress is the instantaneous strength of the material, whereas σ_{max} is a maximum stress arising from a dynamic balance of 'damage' production and healing.

Timescales. Strain-rate weakening can be regarded as the steady-state condition of the system provided it has enough time to adjust to the deformation regime. What is a realistic timescale for this? There are two timescales involved: one for weakening and one for healing. The time to reach equilibrium is the longer of these. The healing timescale may range from extremely large at the top of the lithosphere, perhaps billions of years, to very small in the mantle. Weakening requires a certain strain which may be of order 10s of percent. The weakening timescale is the time required for this total strain to occur, which depends on velocities and lengthscales. Since velocities and lengthscales change as the system evolves, this timescale is difficult to determine without full modeling. Initial conditions also play a huge role: for example, if the lithosphere were initially homogenous and strong it would take a very long time for substantial strain to occur, whereas if there is ‘preexisting’ weakness such as a brittle fault in the upper lithosphere, the required localized strains could occur much more rapidly. In any case, based on healing timescale alone it is unlikely that the ‘steady-state’ is ever reached in the top part of the lithosphere, but it may be an appropriate description for the middle and lower lithosphere.

Validity. Given that memory plays a large role in the evolution of plate tectonics, and that the ‘steady-state’ rheological condition may take an extremely long time to reach in the shallow lithosphere (these two points are essentially the same thing), under what conditions, if any, is it valid to use a SRW rheology in numerical modeling? One possibility is for deeper regions of the lithosphere, where timescales are shorter, perhaps millions of years or less. However, the key reason for testing models with SRW is not realism but that the parameter space in which localization occurs can be rapidly identified, i.e., localization under SRW is a *prerequisite* for localization using the full time-dependent treatment of strain-weakening and healing. If strain-rate weakening does not result in localization for a given set of parameters, then the full time-dependent treatment won’t either. Since these processes are difficult and costly to treat numerically, investigating the basic physics using SRW models saves a lot of computer and human time, compared to using time-dependent models. SRW should not be regarded as a realistic lithospheric rheology *per se* but rather as a necessary step on the path to realistic time-dependent rheologies.

Flow associated with spreading centers, and the bending of the slab at subduction zones are two locations where memory is probably unimportant, and it is valid to use an 'instantaneous' rheological description. Indeed, while oceanic plates in general appear to be very strong and can support loads such as several kilometer high seamounts, at subduction zones these same oceanic plates deform with an 'effective' viscosity which is similar to that of the upper mantle [e.g., *Gurnis, M., S. Zhong and J. Toth, On the competing roles of fault reactivation and brittle failure in generating plate tectonics from mantle convection, this volume*].

2-D Results

Investigations of plate generation using SRW rheologies were pioneered by Bercovici, who developed models in a two-dimensional horizontal sheet representing the lithosphere. (Note that although SRW rheology is often described as 'instantaneous' in the context of these models, in fact it represents the long-term steady-state solution.) Source and sinks (i.e., lines of convergence and divergence) were specified and strike-slip boundaries allowed to develop naturally. Results showed that the ‘plateness’ and localization was poor with powerlaw rheologies, but very good with SRW rheology [*Bercovici*, 1993]. This was later extended to spherical geometry and Earth’s plate configuration [*Bercovici*, 1995b], in which a remarkably good match to Earth's present strike-slip plate boundaries was generated by specifying only the convergent/divergent motion and SRW rheology, indicating that the evolution timescales for Earth may be long enough that SRW rheology is a good approximation. Specifically considering thermal weakening, a transition to plate-like behavior was observed [*Bercovici*, 1996]. Recently, the void-volatile mechanism was introduced [*Bercovici*, 1998].

The above models assumed a sheet lithosphere, so it was not necessary to consider how the rheology changes with depth through the lithosphere and mantle. In order to apply such mechanisms to modeling a more realistic lithosphere and mantle, it is necessary to consider how various quantities change with temperature and pressure, and how the above localization mechanisms should be combined with other rheological complexities such as powerlaw dislocation creep, plastic yielding, and elasticity. In order to gain some insight as to these processes, it is convenient to consider local models of the lithosphere driven by boundary conditions, before attempting full, three-dimensional models covering the entire mantle.

Thus, in this section, we consider localization along a transform boundary by taking a two-dimensional (x,z) plane, where z is vertical, with motion only in the y-direction, i.e., in and out of the plane. The momentum equation thus becomes:

$$(\eta v_{y,x})_{,x} + (\eta v_{y,z})_{,z} = 0 \qquad (13)$$

with v_x=v_z=0. [*Tackley*, 1998a] considered this geometry and showed how localization occurs in a prescribed thickness lithosphere underlain by an asthenosphere separated by a step viscosity jump, and a standard SRW rheology. A narrow shear zone developed through the entire lithosphere, leading to a substantial stress reduction. Here we progress to a more complex and complete rheological description, and consider how this effects the development of a shear zone through the lithosphere. The rheological description is intended to be a plausible, through not unique, way of combining various rheological behaviors.

In the following nondimensional equations, numerical values of parameters are based on the assumption of the following nondimensionalization: η to 10^{19} Pa.s, length to

100 km, and strain rate to 10^{-15} s^{-1}. This leads to nondimensionalizations for other quantities: velocity to 0.315 cm/yr, time to 31.7 Myr, and stress to 10^4 Pa (10^{-2} MPa).

Firstly, a cooling half-space temperature profile is assumed, with nondimensional temperature varying from 0 (300 K) at the surface to 1 (1600 K) at depth:

$$T = erf(z) \qquad (14)$$

The rheological starting point is a standard temperature-dependent powerlaw rheology,

$$\eta_{pl} = \left(\exp\left[37\left(\frac{1}{T+T_{surf}} - \frac{1}{1+T_{surf}}\right)\right]\right)^{\frac{1}{n}} \dot{e}^{\frac{1}{n}-1} \qquad (15)$$

with the nondimensional activation energy of 37.0 corresponding to 400 kJ mol^{-1}. In the calculations presented here, n is taken to be 3 and T_{surf} is 0.23 (300 K).

Next, a plastic, depth-dependent yield stress is specified, and the 'undamaged' viscosity becomes the minimum of the dislocation creep viscosity and the effective viscosity for plastic yielding:

$$\eta_u = \min\left[\eta_{pl}, \frac{\sigma_y(z)}{\dot{e}}\right] \; ; \; \sigma_y(z) = \min\left[z\sigma'_{brit}, \sigma_{duct}\right] \qquad (16)$$

The depth-dependent yield stress is the minimum of the frictional resistance on faults, which is proportional to depth and represented as σ'_{brit}, and a ductile yield stress σ_{duct} to account for ductile, semi-brittle processes in the middle lithosphere. It is uncertain what the pressure-dependence of the latter should be; the present treatment is just an estimate based on Figure 2. Numerical values are σ'_{brit}=600,000 and σ_{duct}=50,000, corresponding to dimensional values of 60 MPa/km and 500 MPa respectively.

Finally, the effective viscosity is assumed to vary linearly with damage D, i.e., equation (4) with λ=1, and a damage evolution equation with dissipative source and exponential (with time) healing is assumed:

$$\frac{dD}{dt} = \sigma : \dot{e} - B(T)D = \eta_u(1-D)\dot{e}^2 - B(T)D \qquad (17)$$

where $B(T)$ is taken as either constant, or temperature-dependent with an arbitrarily chosen law picked to give reasonable variation of healing timescale ($\tau_h=1/B$) at various depths.

$$B(T) = B_0 \exp[20(T-0.5)] \qquad (18)$$

A 2x1 domain is assumed, with nominal depth of 100 km. Side boundaries have a constant v_y of -2 (left) and +2 (right), corresponding to a velocity difference of 1.26 cm/yr across the domain, with top and bottom boundaries stress-free. The system is initialized with some 'seed' damage along a vertical line down the center, then evolved in time according to equation (17), until it reaches a steady-state. The solution is obtained numerically using a finite-difference method on a 128x64 grid. The velocity solution for given viscosity field is obtained using a direct matrix solver. At each time instant a few subiterations on the viscosity field are necessary, due to the nonlinearities in equations (15) and (16). For numerical reasons the maximum viscosity is limited to 10^5. So, the numerical scheme uses explicit damage evolution, not SRW rheology, although the presented results are the same because the system is evolved to a steady-state.

Plate 4 shows results for four different parameter combinations. If the healing rate B is set to an extremely large value such as 10^{20}, then the damage stays close to zero and no localization occurs (bottom case). The lithospheric stress is then 50,000, as expected because this is the ductile yield stress. As can be seen, a plastic yield stress does not cause any localization in the context of a viscous rheology. Elasto-plastic rheologies do exhibit localization, however, and this may change the picture in the upper part of the lithosphere.

For a temperature-dependent B with $B_0=10^4$ (top case), the system forms a narrow, low-viscosity shear zone cutting through most of the lithosphere, with increasingly distributed deformation below this depth. The velocity field (a.) is approximately constant away from the shear zone, as indicated by the strain rate plot (b.). Effective viscosity (c.) is very low in the shear zone, particularly at shallow depth. This is because the healing rate B becomes very small at shallow depth, and the stress in the asymptotic regime is proportional to B (equation 14). Deeper in the shear zone, viscosity increases until the system is no longer in the 'weak' branch of the rheology. The lower lithosphere below the shear zone acts as a stress guide (d.) and thus governs the overall strength of the lithosphere. The maximum stress is 4.8×10^2, two orders of magnitude lower than the case with no weakening.

What effect does B_0 have? In parts e. and f., B_0 is increased by 3 orders of magnitude to 10^7. Due to the more rapid healing, localization is more difficult and the shear zone does not extend as deep. The lower lithosphere again acts as a stress guide, and the maximum stress is 2.0×10^4, only a factor of 2.5 lower than the non-weakening case. So, although the picture does not look so much different, quantitatively there is a lot of difference.

What if B is constant, not temperature-dependent? Parts g. and h. show the result for constant B of 10^4. As shown earlier, constant B implies that the maximum stress for SRW is proportional to $\sqrt{\eta_u}$. thus decreasing with increasing depth. With constant B, the shear zone cuts cleanly through the lithosphere with almost constant internal viscosity, and there is no 'stress guide' through the lower lithosphere. So, in terms of producing a plate boundary, constant healing rate B 'works' better than the arguably more realistic temperature-dependent healing rate.

An issue that must be considered in numerical investigations is the effect of numerical resolution. In these cases, there is no physical limit on the localization process and thus localization occurs down to the scale of the grid. With better resolution the shear zone becomes narrower and with lower internal viscosity. Convergence tests indicate that resolution does not affect the qualitative nature of the result, provided it is not very coarse, and that the maximum stress is weakly dependent on resolution, for example, the maximum stress is about 15% higher on a 64x32 grid. Weakly dependent because overall stress is not determined by the internal viscosity of the shear zone, but rather by the 'stress guide' through the lower lithosphere. It is possible to limit the localization physically by specifying a minimum viscosity [*Tackley*, 1998a], and this should probably be done.

3-D Results

In this section, SRW rheology is applied to three-dimensional flow calculations, with flow driven by thermal buoyancy rather than boundary conditions or imposed sources and sinks. The rheological description is simpler than that of the previous section. Two cases are presented, one using a temperature field from free convection, previously presented in [*Tackley*, 1998a], and the other using a specified temperature field designed to resemble an oceanic plate. Both are instantaneous flow calculations using a 'steady-state' SRW rheology, and thus are not intended to be realistic mantle models, but rather idealized tests designed to elucidate some basic physics.

A major conceptual difference exists between these cases and the preceding analyses in that the system is driven by fixed sources of stress (thermal buoyancy) rather than fixed sources of strain rate (velocity boundary conditions). If the 1-D and 2-D cases were driven by fixed shear stress rather than fixed velocities, solutions would diverge to infinity because stress decreases with increasing strain rate. However, in 3-D the bulk of the mantle, which is not in the SRW regime, limits the maximum velocities that can be obtained: even if the lithosphere tended towards zero strength, flow velocities would saturate and not approach infinity.

The form of SRW equations used are similar to before, but parameterized in terms of maximum stress:

$$\sigma_{ij} = 2\frac{\sigma_{\max}^2\eta}{\sigma_{\max}^2 + \eta^2\dot{e}^2}\dot{e}_{ij} \equiv 2\eta_{eff}\dot{e}_{ij} \tag{19}$$

where

$$\dot{e} = \sqrt{\dot{e}_{ij}\dot{e}_{ij}} \tag{20}$$

The experimental procedure is as follows: (i) A temperature field is obtained, as discussed for the individual models, (ii) The instantaneous velocity-pressure/viscosity solution for the system but with SRW rheology (and either temperature- or depth- dependence) is calculated self-consistently for the entire domain. This involves (ii.a) initializing the viscosity field to a value of 1.0 everywhere, then (ii.b) iterating between the velocity-pressure solution, calculated from the usual mantle convection equations using a finite-volume multigrid technique, both fully described elsewhere [*Tackley*, 1994; *Tackley*, 1996a], and the viscosity solution, calculated from equation (1). Roughly 30-60 iterations are necessary to achieve convergence. The subsequent time evolution of the system is not considered here.

Numerical resolution is generally 32 cells in the vertical direction, with a proportional number in the horizontal directions. Convergence tests indicate that numerical resolution does not make very much difference to the general solution provided it is high enough for strain rates to extend well into the 'weak' branch: with too low resolution the solution decays into a homogeneous rigid lid.

Free convection temperature field. In this case, previously reported in [*Tackley*, 1998a], a temperature field is taken from a Boussinesq, constant-viscosity, 3-D convection calculation in an 8x8x1 box with periodic sides, 100% basal heating and $Ra=10^5$, which has reached statistically steady-state. The lithosphere is taken to be a constant-thickness layer approximately 100 km thick with an undamaged viscosity of 10^4, while the underlying mantle has an undamaged viscosity of 1.0.

The maximum stress $\sigma_{\max}$ is taken to be 10^5. Solutions are not strongly dependent on the exact value, except that a homogeneous rigid lid is obtained if it is too high. Lower values lead to broader weak zones but the same basic pattern. The chosen value is just below that which would cause a rigid lid ($\sim 2(\eta\dot{e})_{\max}$ for Newtonian rheology).

The temperature field (Plate 5a) has a pattern of upwellings and downwellings characteristic of Boussinesq, constant-properties, basally-heated convection at this *Ra*. Examination of the lithospheric viscosity and velocity fields (Plate 5b) shows a pattern of strong plates separated by narrow weak zones (WZ). WZs form above the upwellings and downwellings (as expected from the stress concentration there), but more importantly, also connect the up/downwellings together. The WZs above upwellings and downwellings have a width comparable to the width of the convective feature, which is expected because the region of high stress associated with the feature has this characteristic size. These WZ are associated with strong convergence and divergence (Plate 5c), while the interconnecting zones localize down to a size which is limited by the minimum viscosity cutoff, and are associated with high vertical vorticity. Some of the plates have strong rotation which also shows up in the vorticity. While the plate boundaries appear to be well separated into divergent and strike slip in this figure, they generally include both strike-slip and convergent/divergent motion.

A visco-plastic rheology is much less effective in producing plate-like behavior (Plate 5d). WZs occur above up- and down-wellings but not in-between. Regions of high vorticity now cluster around the convergent/divergent regions, i.e., where the lid is already weakened, compatible with previous results of convection with temperature-dependent viscosity [*Balachandar et al.*, 1995; *Balachandar et al.*, 1996].

To summarize, a simple SRW rheology is very effective in producing strong plates separated by narrow WZ in a 3-D flow calculation, whereas a simple visco-plastic rheology is not. This assumption of a constant-thickness lithosphere was made for numerical reasons but constitutes a lapse of self-consistency which is corrected in the next case.

Oceanic plate and slab. In this experiment, a temperature field is defined based on an oceanic plate and the slab attached to it, then the instantaneous flow is calculated. Half of the box has the cooling plate, while the other half has a constant-thickness lithosphere, so that a transform margin should form between the two regions. The aim of the model is to try and obtain a velocity field which is consistent with the initial assumptions made to construct the temperature field, then advance the system forwards in time. In the present paper, only the initial solution is presented.

The domain is a 1x1x1 box. The initial nondimensional temperature field is defined for a plate traveling in the $+x$ direction as:

$$T(x,y,z) = \begin{cases} \min(T_{plate}(x,z), T_{slab}(x)) & y<0.5 \\ T_{plate\_max} & y>0.5 \end{cases} \quad (21)$$

where

$$T_{plate} = erf\left(\frac{z}{2\sqrt{age}}\right) \quad (22)$$

$$T_{slab} = 0.5 + 0.5\, erf\left(\frac{x-1}{2\sqrt{age_{\max}}}\right) \quad (23)$$

and

$$age = \frac{x}{v_{spread}} \quad (24)$$

$age_{\max}$ is the maximum age (at x=1) and T_{plate_max} is z-dependent temperature at this age. v_{spread} is the spreading velocity, taken to be 90. This temperature field is illustrated in Plate 6a.

The 'undamaged' viscosity is taken to have an Arrhenius temperature-dependence:

$$\eta_u = \exp\left(\frac{23}{T+1} - \frac{23}{1.5}\right) \quad (25)$$

and the maximum stress in the SRW equation (24) is also temperature-dependent, with its variation taken to be mid-way between the two limited cases discussed earlier (constant or proportional to $\sqrt{\eta_u}$), i.e.,

$$\sigma_{\max} = \sigma_{\max 0}\, \eta_u^{\frac{1}{4}} \quad (26)$$

where $\sigma_{\max 0}=3\times10^3$.

The numerical procedure is the same as that for the previous case, except that the solution is initially calculated on a 32x32x32 grid, then refined on a series of grids up to 128x128x32, with uneven z spacing to give much finer spacing in the lithosphere. $Ra=10^5$. Viscosity is cut off at 1000 and 1.0.

The result is shown in Plate 6b&c. Weak zones are formed at the spreading center, subduction zone and transform margin. Vorticity is concentrated in a line along the transform margin. Clearly, the plate margins are not perfect; for example the transform zone bends as it approaches the spreading center and subduction zone. It is thought that this behavior may be due to the large width of the weak zones; if the resolution were adjusted to allow very narrow zones (such as in the 2-D results), the transform 'fault' might be more linear.

DISCUSSION AND CONCLUSIONS

This paper has examined the problem of incorporating plate tectonics into mantle convection models. Models including continents but no oceanic plates showed the presence of supercontinent cycles, given the right initial condition. Ductile weakening mechanisms in the mid-lithosphere are thought to be most important for generating weak plate boundaries. The most likely mechanisms are grain size reduction, void-volatile weakening and viscous dissipation. These may be described using time-dependent evolution equations, giving long-term memory. For initial investigations it is convenient to parameterize them as instantaneous strain-rate weakening. It was demonstrated how SRW leads to localization in a high-resolution local model of lithospheric shear, and in global three-dimensional models with coarser resolution.

However, this line of research is still at an early stage and there is much to be done. This section first discusses some philosophical issues which have been raised concerning the goal of self-consistent plate generation, then discusses future directions and the serious problems that must be overcome.

Objections to Plate Generation

There are a number of problems and pitfalls associated with achieving the goal of self-consistent plate generation, which have led some researchers to question whether it is a worthwhile and valid goal. These are partly based on a misinterpretation of the idealizations made in the initial, very simple models (such as those presented here), which are made for the purposes of testing basic physics, rather than for realism. The assumptions of 'steady-state' rheology and an initially homogeneous lithosphere fall into this category: they are neither thought to be realistic, nor necessary for self-consistent plate generation to 'work'. Here it will be argued that there are no fundamental barriers to reaching this goal, but certainly some serious technical difficulties which may be overcome with sufficient time and effort.

Memory. The first possible problem is that of memory, and the lithospheric heterogeneity that arises from it. The crust and upper lithosphere (at temperatures less than about 600ºC) are riddled with faults and other zones of weakness, which have been built up over billions of years of deformation. "New" plate boundaries are usually a change in style of existing plate boundaries or are formed at pre-existing zones of weakness. In contrast, mantle convection models and the initial plate generation models such as those presented here use an 'instantaneous' viscosity formulation in which a weak zone could, in principal, recover its strength immediately if deformation along it stopped. As shown earlier, this is not a fundamental problem: it is straightforward to include one or more state variables to represent lithospheric weakening mechanisms, as was done in the two-dimensional calculations. The rate of 'healing' of the weakening can be temperature-dependent, so that cold parts of the lithosphere retain a memory for billions of years whereas the mantle adjusts almost instantaneously.

Weakening and localization mechanisms are still important even in the presence of pre-existing heterogeneity, since shear zones and faults that become tectonically inactive do heal at some finite rate; indeed, the warm middle lithosphere, which is the strongest part and thus the most important for determining lithospheric strength, may 'heal' quite rapidly. Thus 'reweakening' and 'relocalization' are important if the region is reactivated at a later date. Isomorphic to the statement that pre-existing weakness controls everything is the following question and answer: Q: Why is the Earth spherical? A: Because it was spherical yesterday- it has pre-existing sphericity- it has always been spherical. While true, some essential physics is being missed by such arguments.

Initial condition. Associated with the issue of memory is the question of initial condition. Whereas the first, simple models assume an initially homogeneous lithosphere, it is clear that the lithosphere was always heterogeneous. As stated already, an initially homogeneous lithosphere is not at all necessary for such models, (indeed, the 2-D models presented here 'seeded' with some initial strength heterogeneity): it is simply an idealization made for the purposes of elucidating some basic physics. However, it is worth considering the likely effect of initial conditions on a simulation that is run for billions of years. Fluid dynamical systems, after sufficient time, generally adopt length scales and characteristics that are independent of initial conditions. While some initial heterogeneity is needed as a 'seed', it does not affect the general character of the final answer. If weak zones are commonplace, the system will self-organize and 'select' those which are closest to the desired breaking-points, allowing it to behave as it 'wants'. Thus, it is most unlikely that the pattern of heterogeneity 4.5 billion years ago determines present-day plate sizes and behavior.

Small-scale details. A common objection is that global models will not be able to reproduce the details of small-scale structure observed at real plate boundaries. It is not clear that this is true, given the rapid improvements in computing power and numerical methods, but if true, the question is: does it matter? It matters if the small-scale structure determines the large scale dynamics. If, instead, the large scale is doing "what it wants" and small-scale details are simply a stochastic reaction to it, then it does not matter.

Future Challenges

Straightforward future directions include using a full time-dependent evolution of material strength and carefully investigating the role of each possible localization mechanism, by numerical models closely coupled to field observations and materials science. This section focuses on some of the most challenging and fundamental scientific and technical issues that need to be addressed.

Small Lengthscales. Structures that occur at plate boundaries, such as shear zones and faults, are very narrow compared to the typical resolution of a global calculation, posing both technical challenges and scientific questions. How can such narrow structures be represented? One possibility is to avoid them by adjusting the parameterization so that localization occurs to a lengthscale which is fairly large. Then it is not clear whether the model is truly representing the important physics: can a description which is appropriate for narrow shear zones be applied to regions 10s of km wide? Too-wide plate boundaries may also be 'leaky': if a shear zone is broad then material can flow in and out of it, and it will have some convergence or divergence.

Multiple lengthscales. Examining plate boundary regions, it is clear that lithospheric weakening must exist on different scale lengths. For example, at subduction zones the dominant deformation is along a narrow fault (or shear zone) but the slab must also be significantly weakened in order to bend around the corner. Deformation at the North America / Pacific margin is distributed over ~100 km, with a large fraction occurring on a single fault, the San Andreas fault [*Lisowski et al.*, 1991]. Thus, more than one mechanism is acting to give these multiple scale lengths.

Elasticity. This paper has barely touched upon the effect of elasticity which is certainly important in the crust and upper lithosphere. Localization in the elastic region, either by brittle failure or plastic flow, may play an important role in generating plate boundaries.

Multiple timescales. A problem with trying to incorporate the elastic region into global calculations is that elastic localization occurs at much smaller timescales than mantle convection. An extreme example is earthquakes: is it necessary to track individual earthquakes on faults? The aim must surely be to develop an implicit parameterization of these processes which allows large timesteps.

Acknowledgments. The author thanks Shun Karato for a thoughtful, detailed review and helpful discussions. This research also benefited from discussions with Greg Hirth and David Bercovici, and reviews by Mark Richards and Ulrich Christensen. Supported by NASA/HPCC Cooperative Agreement No. NCCS5-147 and the David and Lucile Packard Foundation.

REFERENCES

Abbott, D.H., R. Drury, and W.D. Mooney, Continents As Lithological Icebergs - the Importance of Buoyant Lithospheric Roots, *Earth and Planetary Science Letters*, *149* (1-4), 15-27, 1997.

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, *Geophysical Research Letters*, *22* (4), 477-480, 1995.

Balachandar, S., D.A. Yuen, and D.M. Reuteler, High Rayleigh Number Convection At Infinite Prandtl Number With Weakly Temperature-Dependent Viscosity, *Geophysical and Astrophysical Fluid Dynamics*, *83* (1-2), 79-117, 1996.

Batt, G.E., and J. Braun, On the thermomechanical evolution of compressional orogens, *Geophys. J. Int.*, *128*, 364-382, 1997.

Bercovici, D., A Simple-Model Of Plate Generation From Mantle Flow, *Geophysical Journal International*, *114* (3), 635-650, 1993.

Bercovici, D., On the Purpose Of Toroidal Motion In a Convecting Mantle, *Geophysical Research Letters*, *22* (23), 3107-3110, 1995a.

Bercovici, D., A Source-Sink Model Of the Generation Of Plate-Tectonics From Non-Newtonian Mantle Flow, *Journal Of Geophysical Research Solid Earth*, *100* (B2), 2013-2030, 1995b.

Bercovici, D., Plate Generation In a Simple-Model Of Lithosphere-Mantle Flow With Dynamic Self-Lubrication, *Earth and Planetary Science Letters*, *144* (1-2), 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., and P. Wessel, A Continuous Kinematic Model Of Plate-Tectonic Motions, *Geophysical Journal International*, *119* (2), 595-610, 1994.

Bonatti, E., M. Ligi, L. Gasperini, A. Peyve, Y. Raznitsin, and Y.J. Chen, Transform Migration and Vertical Tectonics At the Romanche Fracture-Zone, Equatorial Atlantic, *Journal Of Geophysical Research-Solid Earth*, *99* (B11), 21779-21802, 1994.

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

Bunge, H.P., and M.A. Richards, The Origin Of Large-Scale Structure In Mantle Convection - Effects Of Plate Motions and Viscosity Stratification, *Geophysical Research Letters*, *23* (21), 2987-2990, 1996.

Bunge, H.P., M.A. Richards, and J.R. Baumgardner, A sensitivity study of 3-dimensional spherical mantle convection at 10^8 Rayleigh number - Effects of depth-dependent viscosity, heating mode, and an endothermic phase change, *J. Geophys. Res.*, *102* (B6), 11991-12007, 1997.

Christensen, U., Convection With Pressure-Dependent and Temperature-Dependent Non-Newtonian Rheology, *Geophysical Journal Of the Royal Astronomical Society*, *77* (2), 343-384, 1984.

Christensen, U., and H. Harder, 3-D Convection With Variable Viscosity, *Geophysical Journal International*, *104* (1), 213-226, 1991.

Christensen, U.R., and A.W. Hofmann, Segregation of subducted oceanic crust In the convecting mantle, *J. Geophys. Res.*, *99* (B10), 19867-19884, 1994.

Cserepes, L., Numerical-Studies Of Non-Newtonian Mantle Convection, *Physics Of the Earth and Planetary Interiors*, *30* (1), 49-61, 1982.

Dalziel, I.W.D., Neoproterozoic-Paleozoic Geography and Tectonics - Review, Hypothesis, Environmental Speculation, *Geological Society Of America Bulletin*, *109* (1), 16-42, 1997.

Davies, G.F., Mantle Convection Model With a Dynamic Plate - Topography, Heat-Flow and Gravity-Anomalies, *Geophysical Journal International*, *98* (3), 461-464, 1989.

Davies, G.F., Punctuated Tectonic Evolution Of the Earth, *Earth and Planetary Science Letters*, *136* (3-4), 363-379, 1995.

Dieterich, J.H., Modeling of rock friction 1. Experimental results and constitutive equations, *J. Geophys. Res.*, *84*, 2161-2168, 1979.

Drury, M.R., R.L.M. Vissers, D. Vanderwal, and E.H.H. Strating, Shear Localization In Upper Mantle Peridotites, *Pure and Applied Geophysics*, *137* (4), 439-460, 1991.

Durrheim, R.J., and W.D. Mooney, Evolution of the Precambrian Lithosphere - Seismological and Geochemical Constraints, *Journal of Geophysical Research Solid Earth*, *99* (B8), 15359-15374, 1994.

England, P., and D. McKenzie, A Thin Viscous Sheet Model For Continental Deformation, *Geophysical Journal Of the Royal Astronomical Society*, *70* (2), 295-321, 1982.

Fleitout, L., and C. Froidevaux, Thermal and mechanical evolution of shear zones, *J. Struct. Geol.*, *2*, 159-164, 1980.

Gable, C.W., R.J. Oconnell, and B.J. Travis, Convection In 3 Dimensions With Surface Plates - Generation Of Toroidal Flow, *Journal Of Geophysical Research Solid Earth and Planets*, *96* (B5), 8391-8405, 1991.

Gilbert, L.E., C.H. Scholz, and J. Beavan, Strain Localization Along the San-Andreas Fault - Consequences For Loading Mechanisms, *Journal Of Geophysical Research Solid Earth*, *99* (B12), 23975-23984, 1994.

Glatzmaier, G.A., and P.H. Roberts, A 3-Dimensional Self-Consistent Computer-Simulation Of a Geomagnetic-Field Reversal, *Nature*, *377* (6546), 203-209, 1995.

Gordon, R.G., and S. Stein, Global Tectonics and Space Geodesy, *Science*, *256* (5055), 333-342, 1992.

Govers, R., and M.J.R. Wortel, Extension Of Stable Continental Lithosphere and the Initiation Of Lithospheric Scale Faults, *Tectonics*, *14* (4), 1041-1055, 1995.

Gurnis, M., Large-Scale Mantle Convection and the Aggregation and Dispersal Of Supercontinents, *Nature*, *332* (6166), 695-699, 1988.

Gurnis, M., and T.H. Torsvik, Rapid Drift Of Large Continents During the Late Precambrian and Paleozoic - Paleomagnetic Constraints and Dynamic-Models, *Geology*, *22* (11), 1023-1026, 1994.

Gurnis, M., and S. Zhong, Generation Of Long Wavelength Heterogeneity In the Mantle By the Dynamic Interaction Between Plates and Convection, *Geophysical Research Letters*, *18* (4), 581-584, 1991.

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

Hirth, G., and D.L. Kohlstedt, Water in the Oceanic Upper-Mantle - Implications For Rheology, Melt Extraction and the Evolution of the Lithosphere, *Earth and Planetary Science Letters*, *144* (1-2), 93-108, 1996.

Houseman, G., and P. England, Crustal Thickening Versus Lateral Expulsion In the Indian-Asian Continental Collision, *Journal Of Geophysical Research-Solid Earth*, *98* (B7), 12233-12249, 1993.

Jaroslow, G.E., G. Hirth, and H.J.B. Dick, Abyssal Peridotite Mylonites - Implications For Grain-Size Sensitive Flow and Strain Localization In the Oceanic Lithosphere, *Tectonophysics*, *256* (1-4), 17-37, 1996.

Jin, D.H., S. Karato, and M. Obata, Mechanisms of Shear Localization in the Continental Lithosphere - Inference From the Deformation Microstructures of Peridotites From the Ivrea Zone, Northwestern Italy, *Journal of Structural Geology*, *20* (2-3), 195-209, 1998.

Jordan, T.H., The continental tectosphere, *Rev. Geophys. Space Phys.*, *13* (3), 1-12, 1975.

Jordan, T.H., Composition and development of the continental tectosphere, *Nature*, *274*, 544-548, 1978.

Jordan, T.H., Structure and Formation of the Continental Tectosphere, *J. Petrology, Special Lithospheric Issue*, 11-37, 1988.

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* (20), 2523-2526, 1997.

Kanamori, H., Seismological evidence for a lithospheric normal faulting- the Sanriku earthquake of 1933, *Phys. Earth Planet. Inter.*, *4*, 289-300, 1970.

Karato, S., Grain-Growth Kinetics In Olivine Aggregates, *Tectonophysics*, *168* (4), 255-273, 1989.

Karato, S.-I., M. Toriumi, and T. Fujii, Dynamic recrystallization of olivine single crystals during high-temperature creep, *J. Geophys. Res.*, *7*, 649-652, 1980.

Karato, S.I., M.S. Paterson, and J.D. Fitzgerald, Rheology of synthetic olivine aggregates - Influence of grain-size and water, *J. Geophys. Res.*, *91* (B8), 8151-8176, 1986.

Kikuchi, M., and H. Kanamori, The Shikotan earthquake of October 4, 1994: Lithospheric earthquake, *Geophys. Res. Lett.*, *22* (9), 1025-1028, 1995.

Kohlstedt, D.L., B. Evans, and S.J. Mackwell, Strength Of the Lithosphere - Constraints Imposed By Laboratory Experiments, *Journal Of Geophysical Research Solid Earth*, *100* (B9), 17587-17602, 1995.

Kong, X., Y. Yin, and T.M. Harrison, Evaluating the role of preexisting weaknesses and topographic distributions in the Indo-Asian collision by use of a thin-shell numerical model, *Geology*, *25* (6), 527-530, 1997.

Kuang, W.L., and J. Bloxham, An Earth-Like Numerical Dynamo Model, *Nature*, *389* (6649), 371-374, 1997.

Lenardic, A., On the Heat-Flow Variation From Archean Cratons to Proterozoic Mobile Belts, *Journal Of Geophysical Research Solid Earth*, *102* (B1), 709-721, 1997.

Lenardic, A., and W.M. Kaula, Mantle Dynamics and the Heat-Flow Into the Earths Continents, *Nature*, *378* (6558), 709-711, 1995.

Lenardic, A., and W.M. Kaula, Near-Surface Thermal/Chemical Boundary-Layer Convection At Infinite Prandtl Number - 2-Dimensional Numerical Experiments, *Geophysical Journal International*, *126* (3), 689-711, 1996.

Lenardic, A., and L.-N. Moresi, Some thoughts on the stability of cratonic lithosphere: Effects of buoyancy and viscosity, *J. Geophys. Res.*, *104* (B6), 12747-12758, 1999.

Lisowski, M., J.C. Savage, and W.H. Prescott, The Velocity-Field Along the San-Andreas Fault in Central and Southern California, *Journal of Geophysical Research Solid Earth and Planets*, *96* (B5), 8369-8389, 1991.

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, *Geophysical Research Letters*, *20* (5), 375-378, 1993.

Lowman, J.P., and G.T. Jarvis, Mantle Convection Flow Reversals Due to Continental Collisions, *Geophysical Research Letters*, *20* (19), 2087-2090, 1993.

Lowman, J.P., and G.T. Jarvis, Mantle Convection Models Of Continental Collision and Breakup Incorporating Finite Thickness Plates, *Physics Of the Earth and Planetary Interiors*, *88* (1), 53-68, 1995.

Lowman, J.P., and G.T. Jarvis, Continental Collisions In Wide Aspect Ratio and High Rayleigh Number 2-Dimensional Mantle Convection Models, *Journal Of Geophysical Research Solid Earth*, *101* (B11), 25485-25497, 1996.

Lowman, J.P., and G.T. Jarvis, Effects of mantle heat source distribution on supercontinent stability, *J. Geophys. Res.*, *104* (B6), 12773-12746, 1999.

Lyakhovsky, V., Z. Benavraham, and M. Achmon, The Origin Of the Dead-Sea Rift, *Tectonophysics*, *240* (1-4), 29-43, 1994.

Lyakhovsky, V., Y. Podladchikov, and A. Poliakov, A Rheological Model Of a Fractured Solid, *Tectonophysics*, *226* (1-4), 187-198, 1993.

Mooney, W.D., G. Laske, and T.G. Masters, Crust-5.1 - a Global Crustal Model At 5-Degrees-X-5-Degrees, *Journal of Geophysical Research Solid Earth*, *103* (B1), 727-747, 1998.

Moresi, L., and V. Solomatov, Mantle Convection With a Brittle Lithosphere - Thoughts On the Global Tectonic Styles of the Earth and Venus, *Geophysical Journal International*, *133* (3), 669-682, 1998.

Moresi, L.N., and V.S. Solomatov, Numerical Investigation Of 2d Convection With Extremely Large Viscosity Variations, *Physics Of Fluids*, *7* (9), 2154-2162, 1995.

Obata, M., and S. Karato, Ultramafic Pseudotachylite From the Balmuccia Peridotite, Ivrea Verbano Zone, Northern Italy, *Tectonophysics*, *242* (3-4), 313-328, 1995.

Ogawa, M., G. Schubert, and A. Zebib, Numerical Simulations Of 3-Dimensional Thermal-Convection In a Fluid With Strongly Temperature-Dependent Viscosity, *Journal Of Fluid Mechanics*, *233*, 299-328, 1991.

Olson, P., and D. Bercovici, On the Equipartition Of Kinetic-Energy In Plate-Tectonics, *Geophysical Research Letters*, *18* (9), 1751-1754, 1991.

Parmentier, E.M., C. Sotin, and B.J. Travis, Turbulent 3-D thermal convection in an infinite Prandtl number, volumetrically heated fluid - Implications for mantle dynamics, *Geophys. J. Int.*, *116* (2), 241-251, 1994.

Pili, E., Y. Ricard, J.M. Lardeaux, and S.M.F. Sheppard, Lithospheric Shear Zones and Mantle-Crust Connections, *Tectonophysics*, *280* (1-2), 15-29, 1997.

Poliakov, A.N.B., and H.J. Herrmann, Self-Organized Criticality Of Plastic Shear Bands In Rocks, *Geophysical Research Letters*, *21* (19), 2143-2146, 1994.

Pollack, H.N., Cratonization and Thermal Evolution Of the Mantle, *Earth and Planetary Science Letters*, *80* (1-2), 175-182, 1986.

Ranalli, G., *Rheology of the Earth*, 413 pp., Chapman and Hall, London, 1995.

Ratcliff, J.T., G. Schubert, and A. Zebib, 3-Dimensional Variable Viscosity Convection Of an Infinite Prandtl Number Boussinesq Fluid In a Spherical-Shell, *Geophysical Research Letters*, *22* (16), 2227-2230, 1995.

Ratcliff, J.T., G. Schubert, and A. Zebib, Effects Of Temperature-Dependent Viscosity On Thermal-Convection In a Spherical-Shell, *Physica D*, *97* (1-3), 242-252, 1996.

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

Regenauer-Lieb, K., Dilatant plasticity applied to Alpine collision: ductile void growth in the intraplate area beneath the Eifel volcanic field, *J. Geodyn. (UK)*, *27* (1), 1-21, 1999.

Ribe, N.M., The Dynamics Of Thin Shells With Variable Viscosity and the Origin Of Toroidal Flow In the Mantle, *Geophysical Journal International*, *110* (3), 537-552, 1992.

Ricard, Y., and C. Vigny, Mantle Dynamics With Induced Plate-Tectonics, *Journal Of Geophysical Research-Solid Earth and Planets*, *94* (B12), 17543-17559, 1989.

Richter, F.M., Convection and the large-scale circulation of the mantle, *J. Geophys. Res.*, *78* (35), 8735-8745, 1973.

Ruina, A., Slip Instability and State Variable Friction Laws, *Journal of Geophysical Research*, *88* (B12), 359-370, 1983.

Scotese, C.R., Jurassic and Cretaceous Plate Tectonic Reconstructions, *Palaeogeography Palaeoclimatology Palaeoecology*, *87* (1-4), 493-501, 1991.

Scotese, C.R., L.M. Gahagan, and R.L. Larson, Plate Tectonic Reconstructions of the Cretaceous and Cenozoic Ocean Basins, *Tectonophysics*, *155* (1-4), 27-48, 1988.

Solomatov, V.S., Scaling of temperature-dependent and stress-dependent viscosity convection, *Phys. Fluids*, *7* (2), 266-274, 1995.

Sorensen, K., Growth and Dynamics Of the Nordre-Stromfjord Shear Zone, *Journal Of Geophysical Research*, *88* (B4), 3419-3437, 1983.

Tackley, P.J., Effects of strongly temperature-dependent viscosity on time-dependent, 3-dimensional models of mantle convection, *Geophys. Res. Lett.*, *20* (20), 2187-2190, 1993.

Tackley, P.J., Three-dimensional models of mantle convection: Influence of phase transitions and temperature-dependent viscosity, Ph.D. Thesis, California Institute of Technology, Pasadena, 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, 1998a.

Tackley, P.J., Three-dimensional simulations of mantle convection with a thermochemical CMB boundary layer: D"?, in *The Core-Mantle Boundary Region*, edited by M. Gurnis, M.E. Wysession, E. Knittle, and B.A. Buffett, pp. 231-253, American Geophysical Union, 1998b.

Tackley, P.J., D.J. Stevenson, G.A. Glatzmaier, and G. Schubert, Effects of multiple phase transitions in a 3-dimensional spherical model of convection in Earth's mantle, *J. Geophys. Res.*, *99* (B8), 15877-15901, 1994.

Thatcher, W., and P.C. England, Ductile Shear Zones Beneath Strike-Slip Faults - Implications For the Thermomechanics Of the San-Andreas Fault Zone, *Journal Of Geophysical Research Solid Earth*, *103* (B1), 891-905, 1998.

Trompert, R., and U. Hansen, Mantle Convection Simulations With Rheologies That Generate Plate-Like Behavior, *Nature*, *395* (6703), 686-689, 1998.

Trompert, R.A., and U. Hansen, The Application Of a Finite-Volume Multigrid Method to 3-Dimensional Flow Problems In a Highly Viscous-Fluid With a Variable Viscosity, *Geophysical and Astrophysical Fluid Dynamics*, *83* (3-4), 261-291, 1996.

Trubitsyn, V.P., and V.V. Rykov, A 3-D Numerical-Model Of the Wilson Cycle, *Journal Of Geodynamics*, *20* (1), 63-75, 1995.

Turcotte, D.L., An Episodic Hypothesis For Venusian Tectonics, *Journal Of Geophysical Research Planets*, *98* (E9), 17061-17068, 1993.

Van Keken, P.E., S.D. King, H. Schmeling, U.R. Christensen, D. Neumeister, and M.P. Doin, A comparison of methods for the modeling of thermochemical convection, *Journal Of Geophysical Research Solid Earth*, *102* (B10), 22477-22495, 1997.

Vissers, R.L.M., M.R. Drury, E.H.H. Strating, C.J. Spiers, and D. Vanderwal, Mantle Shear Zones and Their Effect On Lithosphere Strength During Continental Breakup, *Tectonophysics*, *249* (3-4), 155-171, 1995.

Weinstein, S.A., Thermal-Convection In a Cylindrical Annulus With a Non-Newtonian Outer Surface, *Pure and Applied Geophysics*, *146* (3-4), 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 P.L. Olson, Thermal-Convection With Non-Newtonian Plates, *Geophysical Journal International*, *111* (3), 515-530, 1992.

Yuen, D.A., L. Fleitout, G. Schubert, and C. Froidevaux, Shear deformation zones along major transform faults and subducting slabs, *Geophys. J. R. astr. Soc.*, *54*, 93-119, 1978.

Zhong, S.J., and M. Gurnis, Dynamic Feedback Between a Continent-Like Raft and Thermal-Convection, *Journal Of Geophysical Research-Solid Earth*, *98* (B7), 12219-12232, 1993.

Zhong, S.J., and M. Gurnis, Mantle Convection With Plates and Mobile, Faulted Plate Margins, *Science*, *267* (5199), 838-843, 1995.

Zhong, S.J., and M. Gurnis, Interaction Of Weak Faults and Non-Newtonian Rheology Produces Plate-Tectonics In a 3d Model Of Mantle Flow, *Nature*, *383* (6597), 245-247, 1996.

Paul J. Tackley, Department of Earth and Space Sciences, University of California, Los Angeles, 595 Charles Young Drive East, Los Angeles, CA 90095-1567

FIGURE CAPTIONS

Figure 1. Spectral heterogeneity maps (i.e., contour plots of the spatial frequency amplitude spectrum as a function of vertical coordinate) for the continental simulations. Top: initial condition for both cases (single snapshot); Center: Time-averaged SHM for the case with multiple initial continents; Bottom: Time-averaged SHM for the case with a single initial continent. Z is the vertical coordinate ranging from 0 at the bottom of the box to 1 at the top of the box. The spatial frequency is defined such at 1.0 means a horizontal wavelength of 1.0, i.e., equal to the depth of the box. Contours are scaled to the maximum amplitude of each case, in 10 equal steps.

Figure 2. Strength envelope for 60 Ma old oceanic lithosphere (from [*Kohlstedt et al.*, 1995], Figure 9). The solid line in the top part of the lithosphere represents frictional sliding on brittle faults, the solid lines in the bottom part of the figure represent powerlaw viscous creep, and the dotted line connecting these represents ductile or semibrittle flow. The transition from brittle to ductile/semibrittle is known as the brittle-ductile transition (BDT) and the transition from that to viscous creep is known as the brittle-plastic transition (BPT).

Figure 3. Stress vs. strain rate and η_{eff} vs. strain rate for different rheologies, normalized so that they all pass through the point (1,1). Newtonian, powerlaw with n=3, Newtonian+plastic yield stress, and strain-rate-weakening (SRW).

Plate 1. Mantle convection simulation with 16 initial continents at four times during the simulation, corresponding roughly to 0, 0.8, 1.5 and 2.3 billion years. The left plots show the continental material (isocontour C=0.5), and the right plots show downwellings, contouring where the temperature is 0.1 colder than the horizontal average (i.e., isocontour T_{resid}=-0.1)

Plate 2. Mantle convection simulation with a single initial supercontinent, at four times during the simulation, corresponding to 0, 425, 950, and 1420 million years. The left plots show the continental material (isocontour C=0.9), and the right plots show downwellings, contouring where the temperature is 0.1 colder than the horizontal average (i.e., isocontour T_{resid}=-0.1)

Plate 3. The three regimes of convection with temperature-dependent rheology, plus the effect of adding weak plate boundaries. Plotted are isocontours of residual temperature T_{resid}, with hot contours in red and cold contours in green. a. constant viscosity, T_{resid}=±0.1, b. viscosity contrast of 1000, T_{resid}=0.1, c. viscosity contrast of 10^5, T_{resid}=0.05, d. viscosity contrast of 10^5 plus low-viscosity zones imposed at each end of the box, T_{resid}=0.1.

Plate 4. Calculations of lithospheric shear localization with a composite, SRW rheology, in a 2-D (x,z) plane with motion in the y-direction, i.e., in and out of the plane. a. y-velocity, b. strain rate, c. effective viscosity, and d. stress, for the reference case with temperature-dependent healing and B_0=10^4. e. η_{eff} and f. stress for a case with temperature-dependent healing and B_0=10^7. g. η_{eff} and h. stress for a case with a constant healing coefficient B=10^4. i. η_{eff} and j. stress for a case with effectively no strain-rate weakening (obtained by setting B=10^{20}).

Plate 5. Plate formation by SRW and the failure of visco-plastic yield stress rheology in a three-dimensional instantaneous calculation with constant-thickness lithosphere, from [*Tackley*, 1998]. (a) Residual temperature isocontours 0.15 (red) and -0.15 (blue). (b) Lithospheric viscosity and velocity vectors for SRW rheology. Maximum velocity in the domain is 780. (c) Isocontours of horizontal divergence (±50, green and blue) and vertical vorticity (±25, yellow and mauve) and for SRW rheology. (d) as (b) for visco-plastic rheology. (e) As (c) for visco-plastic rheology. Horizontal divergence (±40, green and blue) and vertical vorticity (±10, yellow and mauve).

Plate 6. The formation of plate boundaries in a 3D instantaneous flow calculation with imposed oceanic temperature structure, and temperature-and SRW rheology. a. temperature field: isocountour T=0.8, and two slices, b. logarithm of viscosity, one grid level down from the surface, with variation from 1.0 to 10^3. For rendering, T was subsampled every 2 points and velocity every 4 points. c. Vorticity (yellow, -10.0), divergence (green, 10.0) and convergence (purple, 10.0).