Physics of multiscale convection in Earth’s mantle: Onset of sublithospheric convection

Jun Korenaga1
Department of Earth and Planetary Science, University of California, Berkeley, California, USA

Thomas H. Jordan
Department of Earth Sciences, University of Southern California, Los Angeles, California, USA

Received 13 January 2002; revised 20 February 2003; accepted 18 March 2003; published 9 July 2003.

[1] We investigate the physics of multiscale convection in Earth’s mantle, characterized by the coexistence of large-scale mantle circulation associated with plate tectonics and small-scale sublithospheric convection. In this study, conditions for the existence of small-scale convection beneath oceanic lithosphere are investigated by deriving a scaling law for the onset of convection in a fluid whose surface is instantaneously cooled. We employ two dimensional finite element convection modeling to solve this intrinsically time dependent problem for the Rayleigh number of $10^5$–$10^7$ and with a range of temperature-dependent viscosity. Two different forms of temperature dependency, the Arrhenius law and its linear exponential approximation, are used. We present a new scaling analysis, on the basis of the concept of a differential Rayleigh number, to derive a general scaling law covering from constant viscosity to strongly variable viscosity. Compared to previous studies, our scaling law predicts significantly shorter onset time when applied to Earth’s mantle. Possible reasons for this discrepancy are also discussed.

INDEX TERMS: 3040 Marine Geology and Geophysics: Plate tectonics (8150, 8155, 8157, 8158); 8120 Tectonophysics: Dynamics of lithosphere and mantle—general; 8130 Tectonophysics: Heat generation and transport; 8180 Tectonophysics: Tomography; KEYWORDS: scaling analysis, temperature-dependent viscosity, differential Rayleigh number, small-scale convection, oceanic lithosphere, upper mantle tomography


1. Introduction

[2] Heat flux through lithosphere can lead to convective instability in asthenospheric mantle. Such mantle convection is usually referred to as “small-scale convection,” because its spatial scale is expected to be small (i.e., several hundred kilometers) compared to global mantle circulation associated with plate tectonics. This coexistence of two different scales in mantle convection was first studied quantitatively by Richter [1973]. In the presence of large-scale vertical shear corresponding to plate motion, small-scale convection tends to organize itself to minimize convective interference with the large-scale field, forming longitudinal convection rolls, or “Richter rolls” [Richter, 1973; Richter and Parsons, 1975]. Since Richter’s original work, and especially after Parsons and McKenzie [1978] proposed small-scale convection as a physical mechanism for seafloor flattening observed at old ocean basins [e.g., Parsons and Sclater, 1977], the dynamics of small-scale convection beneath oceanic lithosphere has been studied by a number of geophysicists [e.g., Houseman and McKenzie, 1982; Fleitout and Yuen, 1984; Buck and Parmentier, 1986; Davies, 1988; Davaille and Jaupart, 1994; Dumoulin et al., 1999; Solomatov and Moresi, 2000]. Recently, Conrad and Hager [1999b] pointed out that small-scale convection may also play an important role in the thermal evolution of Earth. Because energy dissipation owing to plate bending at subduction zones is significant in the gross energetics of mantle convection [Conrad and Hager, 1999a], the efficiency of plate tectonics is very sensitive to the thickness of lithosphere, which can be potentially controlled by small-scale convection.

[3] For many years, however, geophysical inference for the presence of small-scale convection has been limited to surface observables such as gravity and geoid anomalies, and the physical interpretation of these signals has been controversial. Wessel et al. [1996] provides a concise summary for this issue. Here we simply point out that two different wavelengths have been identified for gravity or geoid undulations, and that, while short-wavelength ($\sim 200–300$ km) signals [e.g., Haxby and Weissel, 1986] may have shallow, crustal origins [e.g., Winterer and Sandwell, 1987; Sandwell et al., 1995], a convective origin remains plausible for long-wavelength ($\sim 1000$ km) signals [e.g., Cazenave et al., 1992].

---

1 Now at Department of Geology and Geophysics, Yale University, New Haven, Connecticut, USA.
[4] By constructing a new kind of seismic tomography using frequency-dependent travel times and ScS reverberations, Katzman et al. [1998] successfully imaged the fine-scale seismic structure of the Pacific upper mantle for the Tonga-Hawaii corridor. They showed that this particular 2-D cross section of the upper mantle is characterized by strong seismic anomalies with the lateral dimension of about 700 kilometers. The Tonga-Hawaii corridor is nearly perpendicular to the absolute motion of the Pacific plate, and these small-scale seismic anomalies seem to correlate with long-wavelength geoid and topography, which are elongated in parallel with the plate motion. Katzman et al. [1998] thus speculated that their tomography may indicate the presence of Richter rolls. Chen et al. [2000] extended this tomographic method to 3-D, and by combining several source-receiver paths in the southwestern Pacific, they showed that the 3-D seismic structure of the Pacific upper mantle is more complicated and less symmetric than one would expect for Richter rolls, though strong regional-scale seismic anomalies are still consistently observed.

[5] In light of this new seismic information on the fine-scale structure of the oceanic upper mantle, our understanding of multiscale convection in Earth’s mantle is still primitive. Whether or not small-scale convection takes place in the first place has not been accurately understood; previous attempts to understand this problem exhibit significant discrepancy. As a starting point for our effort to understand the dynamics of a whole-mantle system that exhibits both large-scale mantle circulation and small-scale convection (Figure 1), therefore, we revisit the condition for the onset of convection with strongly temperature-dependent viscosity. After presenting our problem setting, numerical formulation, and results, the concept of the differential viscosity. After presenting our problem setting, numerical formulation, and results, the concept of the differential Rayleigh number is introduced, with which a nonlinear scaling law is naturally derived. By comparing with numerical results, a general scaling law, which is applicable to a wide range of temperature-dependent viscosity, is obtained for the onset time of convection.

2. Onset of Convection

[6] The onset of convection in a fluid whose surface is instantaneously cooled (Figure 2) is a classical problem in fluid mechanics [e.g., Foster, 1965, 1968]. When a semi-infinite medium of an initially uniform temperature, \( T_0 \), is subject to a sudden decrease in the surface temperature (\( T_s \)), its internal temperature profile, \( T(z, t) \), can be expressed as [e.g., Carslaw and Jaeger, 1959, p. 59]

\[
T(z, t) = T_s + (T_0 - T_s) \text{erf} \left( \frac{z}{2\sqrt{\kappa t}} \right),
\]

where \( \kappa \) is thermal diffusivity, \( t \) is time, and \( z \) is a vertical coordinate originated at the surface. The thermal evolution of oceanic lithosphere is usually discussed on the basis of this equation [e.g., Turcotte and Schubert, 1982]. Because of thermal contraction, this temperature profile creates gravitationally unstable stratification and may eventually lead to convection, after which equation (1) is no longer valid.

[7] The characteristic length scale in the purely heat conduction phase is proportional to \( \sqrt{\kappa t} \). This time-depen-

dent nature of a basic state as well as the presence of another greater length scale (i.e., the system height) makes it impractical to apply conventional approaches in linear stability analysis such as the calculation of critical Rayleigh number or growth rate (see Howard [1966] for an excellent summary on this issue). Thus the onset of convection has to be investigated by employing the following two steps: (1) solve directly the initial-value convection problem, either by numerical simulation or by laboratory experiments, and then (2) derive a scaling law for onset time that is consistent with experimental results.

[8] Earth’s mantle is known to be strongly temperature-dependent [e.g., Weertman, 1970], which presents an additional challenge because the characteristic length scale of convective instability is no longer proportional to \( \sqrt{\kappa t} \). Early computational efforts on this problem include Yuen et al. [1981] (frozen-time analysis), Yuen and Fleitout [1984] (initial-value analysis), and Jaupart and Parsons [1985] (frozen-viscosity analysis). Davaille and Jaupart [1994] were the first to apply the above two-step approach to derive a general scaling law for the onset time of convection with temperature-dependent viscosity, on the basis of their laboratory experiments [Davaille and Jaupart, 1993]. A similar study was also conducted by Choblet and Sotin [2000] on the basis of numerical simulation.

[9] There are, however, several problems in previous studies, and currently available scaling laws, which exhibit significant discrepancy among different studies, are not applicable to Earth’s mantle. In the experiments of Davaille and Jaupart [1993], for example, the top surface was cooled gradually not instantaneously because instantaneous cooling without introducing mechanical disturbance is difficult to achieve in a laboratory. Unfortunately, the timescale of their gradual cooling is comparable to that of convection onset,

Figure 1. Schematic diagram of possible evolution of sublithospheric convection in oceanic mantle.
and hence the accuracy of their data is hard to estimate. In
addition, because laboratory fluids have a limited range of
temperature dependency, laboratory experiments are not ideal
to derive a scaling law that needs to cover a wide
range of temperature dependency. Furthermore, a linear
scaling analysis used by Davaille and Jaupart [1994] and
Chabot and Sotin [2000] does not guarantee accurate
extrapolation of their results to Earth’s mantle. The prob-
temperatures to see if convective instability depends on the form
of temperature dependency. The first form is the Arrhenius
law of temperature-dependent viscosity, which may be
expressed as

\[ \mu(T^*) = \exp \left( \frac{E^*}{T^* + T_{\text{off}}} - \frac{E^*}{1 + T_{\text{off}}} \right). \]

where \( E^* \) is activation energy, \( R \) is the universal gas constant. As a general
measure of temperature dependency, it is useful to define
the temperature derivative of logarithmic viscosity with
internal temperature as [e.g., Morris and Canright, 1984]

\[ \sigma = \frac{d \log \mu}{dT^*} \bigg|_{T^*=1}. \]

For the above Arrhenius law, we have \( \sigma = E^*/(1 + T_{\text{off}})^2 \).
The second viscosity law is the Frank-Kamenetskii
approximation using this parameter \( \sigma \) as

\[ \mu(T^*) = \exp[\sigma(1 - T^*)], \]

which is a linear exponential viscosity. Earth’s mantle
is known to obey the Arrhenius law [e.g., Weertman, 1970; Karato and Wu, 1993]. However, equation (8) has been
widely used in numerical studies of mantle convection
because its limited viscosity variation (for a given \( \sigma \)) makes it computationally less expensive to achieve the accurate
calculation of Stokes flow. These two viscosity laws are very
similar when \( T^* \sim 1 \), but they rapidly diverge as \( T^* \to 0 \).

A criterion for the onset of convection is similar to
that adopted by Davaille and Jaupart [1993]. We define
the onset time of convection as when the difference between a
horizontally averaged temperature profile and the purely
conductive profile (i.e., equation (1)), which we denote as \( \delta T^* \), exceeds 0.01 at any depth level. Because of the rapid
growth of convective instability, our measurement is not
very sensitive to this particular value of the threshold.

We use the finite element method to solve the coupled
Stokes flow and thermal advection-diffusion equations (2)–
(4). The implementation of our convection code is essen-
tially identical to that of ConMan [King et al., 1990] with
comparable accuracy, which was confirmed by several
benchmark tests on Stokes flow calculations [e.g., Moresi
et al., 1996] and critical Rayleigh number calculations
The computational domain is discretized with uniform 2-D quadrilateral elements. The range of the Rayleigh number is $10^5 – 10^7$, and the length of element side is $1/30$ for $Ra = 10^5$, $1/45$ for $Ra = 10^6$, and $1/60$ for $Ra = 10^7$. By conducting a number of resolution tests using finer meshes, we found that the accurate measurement of onset time is possible with the above mesh resolution, if time stepping is done carefully. In the streamline Petrov-Galerkin formulation of the thermal advection-diffusion problem [Brooks and Hughes, 1982], the largest time step allowed to maintain numerical stability, with uniform element length, is $\Delta t_{\text{diff}} = \frac{1}{2} h^2 \cdot \text{diff}$ for the diffusive limit, and $\Delta t_{\text{ad}} = \frac{h}{u_{\text{max}}} \cdot \text{ad}$ for the advection limit, where the term $u_{\text{max}}$ denotes the largest component in the velocity field. Some fraction of a smaller step is used for actual time stepping, i.e., $\Delta t = a \cdot \min (\Delta t_{\text{diff}}, \Delta t_{\text{ad}})$, where $0 < a < 1$. A common choice for the factor $a$ is 0.5 or 0.25. For finite amplitude convection at supercritical Rayleigh numbers, the diffusive time step is usually much larger than the advective one, so a typical convection code takes large time steps when thermal diffusion is dominant. The onset of convection marks a transition from a diffusion-dominant phase to an advection-dominant phase, and we found that $a = 0.25$ was not sufficiently small to accurately measure this transition. Based on our benchmark tests, we chose to use $a = 0.05$ to secure less than 1% error in measuring onset time. Though a similar effect on accuracy can be automatically obtained by reducing element size, the combination of lower-resolution mesh and finer time stepping is more efficient because of the nonlinear scaling of computational resources required to solve the Stokes flow equation.

[18] The aspect ratio of our convection model is unity. We also tested wider aspect ratios, but compared to the random nature of onset time described below, differences seen in these additional calculations are small. A reflecting boundary condition (i.e., free slip and adiabatic) is applied to the side boundaries. A periodic boundary condition was also tested, with negligible differences. The top and bottom boundaries are rigid. The top and bottom temperatures are fixed at 0 and 1, respectively. The initial internal temperature is set to unity plus small random perturbations with the maximum amplitude, $\varepsilon$. Two values of $\varepsilon$, $10^{-5}$ and $10^{-3}$, are used for the following reasons.

[19] Finite perturbations always exist in a real fluid because of thermal noise. The random nature of thermal noise leads to uncertainty in onset time, the standard deviation of which is typically around 10% [Blair and Quinn, 1969]. Jhaveri and Homsy [1980, 1982] studied this intrinsic uncertainty by solving stochastic differential equations incorporating thermodynamically determined fluctuation. A similar result can be obtained by integrating the regular convection equations starting with a number of different Monte Carlo realizations for initial random perturbations. In order to simulate thermal noise in our calculations, we conducted a number of isoviscous calculations with different maximum amplitudes, for $Ra = 10^5 – 10^7$.

Figure 3. Stochastic nature of onset time. Evolution of 30 Monte Carlo realizations for initial temperature perturbations is shown in terms of deviation from conducting temperature profile, for two different maximum amplitudes, $\varepsilon = 10^{-5}$ (solid) and $\varepsilon = 10^{-3}$ (gray). (a) $Ra = 10^6$ and (b) $Ra = 10^7$. Constant viscosity is used for all calculations here.
(Figure 3). For a range of the initial perturbation amplitude, the standard deviation of onset time is consistently around 10%, in agreement with the previous studies. With \( \varepsilon = 10^{-5} \), we are able to match our onset times with the scaling law derived by Howard [1966], which is based on laboratory experiments with isoviscous fluid. Because larger perturbations than thermal noise are likely to be present in Earth’s mantle, we also use \( \varepsilon = 10^{-3} \) to test the sensitivity of our scaling laws to the amplitude of initial perturbations. Note that, for the first few steps in calculations, deviation from the reference conduction profile is quite large (\( \sim 0.05 \)) because the adopted mesh resolution is insufficient to resolve a thin thermal boundary layer resulting from instantaneous cooling. The deviation, however, quickly decreases below 0.01 as the boundary layer grows, showing that this initial inaccuracy does not affect succeeding computation. By ignoring the first few inaccurate steps, therefore, our criterion for onset time can be safely applied.

4. Results

[29] We conducted total 104 runs of the instantaneous cooling model, with \( E = 10 - 50 \) kJ mol\(^{-1} \) (10 - 30 for the Arrhenius law) for \( Ra = 10^5 \), \( E = 10 - 120 \) kJ mol\(^{-1} \) for \( Ra = 10^6 \), and \( E = 20 - 200 \) kJ mol\(^{-1} \) for \( Ra = 10^7 \), for the two viscosity laws and the two amplitudes for initial temperature perturbation. We set \( \Delta T = 1300 \) K. The upper limit of the activation energy is chosen so that convective instability sets in before thermal conduction is affected by the bottom boundary. Though the finite domain effect still exists for convective flow, especially for low Rayleigh number and high activation energy, it may be rather appropriate in terms of geophysical applications. Convective instability in asthenosphere is likely to feel some ‘bottom’ because of potential viscosity layering in the mantle. To ensure numerical stability, the largest normalized viscosity is limited to \( 10^5 \).

[21] An example is shown in Figure 4, which corresponds to the Arrhenius viscosity case of \( Ra = 10^6 \), \( E = 40 \) kJ mol\(^{-1} \), and \( \varepsilon = 10^{-3} \). Some ambiguity is apparent regarding the definition of onset time. If we chose to use some kinetic measure to define the onset, such as kinetic energy or maximum velocity, we would have much earlier onset. Instead of \( \tau_c \) and \( \tau_r \), we use some kinetic measure to define the onset, such as kinetic energy or maximum velocity, we would have much earlier onset. Instead of \( \tau_c = 0.0162 \), for example, one may pick \( \tau_c = 0.0085 \) at which kinetic energy starts to grow exponentially (Figure 4e). Such choice, however, would make it difficult to compare our results with laboratory experiments as well as to apply our scaling law to geophysical observations.

[22] Our measurements of onset time are summarized in Figure 5. Onset times, \( \tau_c \), are normalized by the local timescale for boundary layer instabilities, \( \tau_\text{sh} = Ra d^{-2/3} \), and they are plotted as a function of \( \sigma \). The following features are clearly observed: (1) onset times with different Rayleigh numbers, but the same viscosity law and the same perturbation amplitude, collapse reasonably well on a single trend after normalization, (2) onset times are consistently shorter with larger initial temperature perturbation, and (3) onset times are systematically shorter for the Frank-Kamenetskii approximation, except in the limit of constant viscosity. It is also shown that our measurements are significantly different from the scaling laws of Davaille and Jaupart [1994] and Choblet and Sotin [2000]. Possible reasons for this discrepancy will be discussed later.

5. Scaling Analysis

[21] Convection with strongly temperature-dependent viscosity is characterized by an almost rigid lid and a nearly isoviscous convecting interior [e.g., White, 1988]. Existing scaling analyses divide the thermal boundary layer into a rigid lid and a mobile basal region, and investigate the convective instability of the latter [e.g., Davaille and Jaupart, 1994]. The linear exponential approximation has been employed in this division to simplify a stability analysis. We do not follow this common procedure for the following reasons. The transition from the mobile part to the rigid lid is gradual, so the simple binary treatment is unlikely to provide an accurate length scale for the mobile layer, especially when a given viscosity law is not linearly exponential. In addition, an e-fold viscosity contrast in the mobile part is usually neglected, which can further degrade the accuracy of scaling. The systematic difference in onset time between the Arrhenius law and the Frank-Kamenetskii approximation (Figure 5) cannot be explained by the conventional approach, which does not distinguish these two viscosity laws. Furthermore, previous studies incorporated a scaling relationship valid for a steady state convection regime to the analysis of onset time, which is a highly transient phenomena. Mixing up different dynamical regimes should also be avoided.

[21] We still follow, however, a guiding principle proposed by [Howard, 1966] and assume that convection takes place when some kind of local Rayleigh number reaches a critical value, i.e.,

\[
Ra_c(\tau_\text{sh}) = R_c.
\]  

This is a simple yet physical view of convective instability. The tricky part is how to define \( Ra_c \) for strongly temperature-dependent viscosity, on which we will concentrate in the following.

[25] The use of a similarity variable, \( \eta = z^*/(2\sqrt{\tau_\text{sh}}) \), can considerably simplify our analysis. Given the cubic dependency of \( Ra \) on the length scale, the time dependence of the local Rayleigh number may be separated out as

\[
Ra(\tau^*) = Ra_0 \left( 2\sqrt{\tau^*} \right)^3,
\]  

where \( Ra_0 \) is defined in terms of the similarity variable, so the onset time of convection can be expressed as

\[
t^* = \frac{1}{4} \left( \frac{R_c}{Ra_0} \right)^{2/3}.
\]  

[26] The Rayleigh number is fundamentally a macroscopic parameter because of its strong dependence on the system height (equation (5)). Because of this, there is no universal prescription for how to define an effective Rayleigh number in the case of variable viscosity. If there is some sort of differential form of the Rayleigh number, which
can be calculated point-wise, however, it would be straightforward to obtain a domain-wide, effective Rayleigh number simply by integrating it. Our approach here is based on this idea, and we will demonstrate that it is a powerful way to derive the systematics of onset time. First of all, the concept of ‘available buoyancy’ [Conrad and Molnar, 1999] is useful to identify the extent of the mobile thermal boundary layer. In our problem, available buoyancy may be defined as

\[ B(\eta) = \int_0^\infty \frac{1}{\mu^*} \frac{dT^*}{d\eta^T} d\eta^T. \] (14)

Figure 4. Example of numerical solutions for the onset of convection. The Arrhenius viscosity case with \( Ra = 10^6 \), \( E = 40 \text{ kJ mol}^{-1} \), and \( \varepsilon = 10^{-5} \). Snapshots of temperature and velocity fields are shown at (a) \( t^* = 0.01235 \), (b) \( t^* = 0.01620 \), (c) \( t^* = 0.01708 \), and (d) \( t^* = 0.02222 \). Contour interval is 0.1. Velocity arrows are normalized by maximum velocity, which is denoted at every snapshot. Also shown are (e) kinetic energy, (f) deviation from conducting temperature profile, and (g) maximum upwelling (solid) and downwelling (dotted) velocities.
For example, the topmost part of the thermal boundary layer, which is the coldest and thus densest, may not contribute to convective instability if its viscosity is very high owing to temperature-dependent viscosity, and the available buoyancy automatically takes this effect into account. The inflection point of available buoyancy is denoted as \( \eta_0 \) where \( B' = 0 \) and \( B''' > 0 \) (Figure 6a). The ‘origin’ of positive buoyancy, \( \eta_0 \), is calculated using a tangent at the inflection point as

\[
\eta_0 = \eta_i - \frac{B(\eta_i)}{B'(\eta_i)}
\]

(15)

(see also Figure 6a). We then define the ‘differential Rayleigh number’ as

\[
dRa(\eta) = \frac{4\alpha \rho_0 g (\eta - \eta_0)^3}{\kappa \mu} \frac{dT}{d\eta} d\eta
\]

(16)

The local Rayleigh number is obtained by integrating this differential Rayleigh number as

\[
Ra_i = \int_0^\infty dRa(\eta)
\]

(17)

\[
= Ra \cdot F(\mu^*, T^*)
\]

(18)

where \( F(\mu^*, T^*) \) is a functional that depends on a viscosity function and a temperature profile as

\[
F(\mu^*, T^*) = \int_0^\infty \frac{4(\mu - \mu_0)^3}{\mu^*} \frac{dT^*}{d\eta} d\eta
\]

(19)

If \( T^* = \text{erf}(\eta) \) (i.e., instantaneous cooling) and viscosity is normalized by its lowest possible value, then we have \( 0 < F(\mu^*, T^*) \leq 4/\sqrt{n} \) (Figure 6b). We note that, by changing the integration interval from \([0, \infty]\) to \([0, 1]\), the functional \( F \) becomes unity for a linear temperature profile and uniform viscosity (i.e., \( T^* = \eta \) and \( \mu^* = 1 \)). Thus we recover \( Ra_i = Ra \) in the classical case of marginal stability.

[27] As clearly seen in equation (18), the functional \( F \) serves as a scaling factor, which takes into account the effect of variable viscosity as well as the nonlinearity of a temperature profile. Its physical meaning may be considered in terms of buoyancy density, \( B(\eta) = (dT^*/d\eta)/\mu^* \), which is a Gaussian-like distribution (exactly so for \( \sigma = 0 \) with \( T^* = \text{erf}(\eta) \)). The zeroth moment of the buoyancy density is \( B(\infty) \), and \( \eta_0 \) is related to the first moment. By approximating the buoyancy density as a uniform distribution, i.e.,

\[
B'(\eta) \sim \begin{cases} 
B'(\eta_0) & \eta_0 \leq \eta \leq \eta_0 + B(\infty)/B'(\eta_0) \\
0 & \text{otherwise}
\end{cases}
\]

The functional \( F \) is therefore related to the total buoyancy density and the width of the buoyancy distribution. The appearance of the third power is consistent with the cubic dependency of \( Ra \) on the length scale. The asymptotic behavior of \( F \) at large \( \sigma \) can be expressed as

\[
F \sim C \left[ \frac{1 - \exp(-\sigma)}{\sigma} \right]^3
\]

(21)

where \( C \) is a scaling factor of order 1.
For instantaneous cooling with uniform viscosity, we have from equations (13), (18), and (19),

\[ t^*_{c, \sigma=0} = \left( \frac{\sqrt{\pi} R_c^{1/2}}{2^5 \alpha^3} \right). \]  

(22)

On the other hand, Howard [1966] used the slope of a temperature profile at the surface to define the length scale for a thermal boundary layer (i.e., \( \eta = \pi^{1/2}/2 \); see Figure 6a), and he obtained

\[ t^*_{c, \sigma=0} = \frac{1}{\pi} \left( \frac{R_c}{R_{c, \sigma}} \right)^{1/2}. \]  

(23)

Comparing these two expressions, we can see the relation between our \( R_c \) and the conventional critical Rayleigh number as \( R_{c, \sigma} = (\pi^{1/2}/2) R_c \), and the onset time of convection can be expressed as

\[ t^* = \left( \frac{4}{\pi^2 F(\alpha^* , T^*)} \right) \frac{R_{c, \sigma}}{R_{c, \sigma}}. \]  

(24)

The critical Rayleigh number is estimated by fitting equation (24) to our data. Only data with \( R_{c, \sigma} = 10^7 \) are used in this regression because they cover the widest range of temperature dependency, and also because data with lower Rayleigh numbers are more affected by the presence of the bottom boundary. We obtained \( R_{c, \sigma} = 2000 \) for \( \varepsilon = 10^{-5} \) and \( R_{c, \sigma} = 1290 \) for \( \varepsilon = 10^{-3} \). Note that a single critical Rayleigh number can simultaneously handle both the Arrhenius viscosity and the linear exponential viscosity because a different viscosity law has a different value of \( F \) (Figure 6b). Our integral approach provides a more accurate and unified treatment of temperature-dependent viscosity than a conventional local gradient approach, and as a consequence, we are able to derive a general scaling law covering from constant viscosity to strongly variable viscosity. An extension to temperature- and depth-dependent viscosity is presented by Korenaga and Jordan [2002a].

6. Discussion and Conclusion

Significant discrepancy among our scaling law and preexisting laws (Figure 5) urges a critical review of previous studies. We begin with the work of Davaille and Jaupart [1993, 1994]. As already noted, their laboratory experiments do not accurately model instantaneous cooling. Furthermore, though they reported the largest viscosity contrast achieved across the entire thermal boundary layer is around 10^9, the temperature-dependent viscosity of Golden Syrup used in most of their experiments is superexponential (Figure 7a), and the largest \( \sigma \) is only \( \sim 6.4 \). For this limited range of a viscosity contrast, the mismatch between our numerical results and the scaling law of Davaille and Jaupart [1994] may be reasonable (Figure 5a), and a part of this mismatch is most likely due to the use of different form of temperature-dependent viscosity. To investigate this, we calculated the onset time of convection with the viscosity law of Golden Syrup, to compare with the scaling law of Davaille and Jaupart [1994] as well as our scaling law with the two types of viscosity laws used in this study (Figure 7c). Using the critical \( R_{c, \sigma} \) of 1300, our scaling law is fairly consistent with that of Davaille and Jaupart [1994]. A minor remaining difference probably reflects possible errors introduced by the use of gradual cooling in their experiments.

It may be difficult to derive a scaling law from the mixture of experimental ensembles all with different forms of temperature-dependent viscosity. In the experiments of Davaille and Jaupart [1993], different viscosity contrasts were obtained by changing temperature boundary conditions. This resulted in sampling different parts of the Golden Syrup viscosity law (Figure 7a), which varies from almost linear exponential to Arrhenius (Figure 7b). This gradual change in the form of temperature dependency translates into the higher sensitivity of onset time with respect to activation energy or \( \sigma \) (Figure 7c). This becomes more serious for stronger temperature dependency, and their scaling law predicts more than twice as large onset time as ours for \( \sigma \geq 15 \) (Figure 5), which is expected for Earth’s
mantle. We note, however, that it is possible to derive a scaling law similar to ours even from the analysis of Davaille and Jaupart [1994], by removing their final linear approximation (A. Davaille and S. Zaranek, personal communication, 2002), though it still breaks down at the weakly temperature-dependent viscosity regime ($\sigma < \sim 2$) and is unable to distinguish different types of variable viscosity.

Choblet and Sotin [2000] derived a scaling law for onset time on the basis of 3-D numerical modeling using linear exponential viscosity with $\sigma = 4 – 8$. Their scaling analysis is essentially the same as that of Davaille and Jaupart [1994]. The reason why their scaling law is so different from ours and also from that of Davaille and Jaupart [1994] may be because they run their model with zero initial perturbation and let convective instability grow from numerical error. Though our study is limited to 2-D modeling with a unit aspect ratio, neither two dimensionality nor the effect of walls can explain why our onset times are shorter than those of Choblet and Sotin [2000] because both factors enhance the convective stability of a fluid [e.g., Korenaga and Jordan, 2001].

The use of initial random perturbations to model thermal noise in our numerical calculations is not a perfect approach because thermal noise is continuously generated in a real fluid, but it is still a simple and reasonable approach. Random noise contains all wavelengths. Most unstable components have relatively long wavelengths, and because of this, their initial perturbations decay only very slowly. Their $\epsilon$-fold timescale is comparable to a fraction of unit diffusion time, and this is substantially longer than onset timescale. So, even though real noise maintains shorter wavelength components all the time, our approach should be able to capture the nature of steady state thermal noise in terms of its capability to trigger convective instability. The following observations seem to support this: (1) With the same Rayleigh number and for a range of $\sigma$, onset timescale has a similar order of magnitude. So the amplitude of destabilizing components should be similar at onset for a range of $\sigma$, and (2) the amplitude of destabilizing components should also be similar at onset for the range of $Ra (10^5 – 10^7)$ because, after normalization with local boundary layer timescale, data with the same viscosity law are very similar.

To sum, a new scaling law for the onset of convection was derived on the basis of 2-D numerical calculations with a wide range of temperature-dependent viscosity. Compared to previously known scaling laws, our result suggests significantly short onset time for strong temperature-dependency. There are a number of reasons why our scaling law is more applicable to Earth’s mantle, as discussed above. The new scaling law has been used to infer the viscosity of oceanic mantle [Korenaga and Jordan, 2002b], and the result is shown to be consistent with other geophysical estimates. Our study suggests that small-scale convection is likely to take place beneath oceanic lithosphere, and that further study on the possible evolution of sublithospheric convection is warranted. The onset of convection is just the first step in understanding the physics of multiscale convection in Earth’s mantle (Figure 1). The

Figure 7. (opposite) Temperature-dependent viscosity of Golden Syrup used in the experimental work of Davaille and Jaupart [1993]. (a) Normalized viscosity as a function of normalized temperature is plotted for the ten different experimental conditions employed in their work. (b) Corresponding values of the functional $F$ are plotted as a function of $\sigma$ (solid circles). The cases for the Arrhenius viscosity (dotted) and linear exponential viscosity (solid) are also shown for comparison. (c) Predicted onset time of convection is calculated with $Ra_c$ of 1300, which is chosen to make the prediction to be most consistent with the scaling law of Davaille and Jaupart [1994] (dashed).
subsequent phases of sublithospheric convection will be investigated by J. Korenaga and T. H. Jordan (Physics of multiscale convection in Earth’s mantle: Evolution of sublithospheric convection, submitted to Journal of Geophysical Research, 2003) with a continued emphasis on basic scaling laws. Finally, the role of small-scale convection in the whole-mantle system will be characterized on the basis of 3-D numerical modeling (J. Korenaga and T. H. Jordan, Physics of multiscale convection in Earth’s mantle: Whole-mantle model with a single plate, manuscript in preparation, 2003).

Appendix A: Benchmark Test of Our Finite Element Code

[35] In order to verify our convection code, both for the Stokes flow solver and for the thermal advection-diffusion solver, we choose to calculate numerically the critical Rayleigh number (for marginal stability) and compare it with that from an analytical method [e.g., Zhong and Gurnis, 1993]. The marginal stability of a fluid has been studied by a linear stability analysis (LSA) for constant viscosity [e.g., Chandrasekhar, 1981] as well as for temperature-dependent viscosity [Stengel et al., 1982] (Table A1). If the bounding surfaces are both free slip, for example, the critical Rayleigh number for an isoviscous fluid with infinite horizontal extent is $27\pi^4/4 \approx 658$ at the wave number of $\pi/\sqrt{2}$ [Chandrasekhar, 1981, p. 36]. For temperature-dependent viscosity, we use the exponential viscosity of the form (8), which is different from that used by Stengel et al. [1982] by a factor of $\exp(\sigma/2)$. The critical Rayleigh numbers calculated by Stengel et al. [1982] are thus multiplied by this factor in Table A1. We note that the overall accuracy of the results of Stengel et al. [1982] is reported to be more than three significant figures.

[36] Our numerical approach is the following. On the linear, conductive temperature profile, $T^* = 1 - z^*$, we add an initial temperature perturbation as

$$\delta T^*(r^* = 0) = A \cos(2\pi r^*/L_c) \sin(\pi z^*),$$

(A1)

where $A = 0.01$ and $L_c$ is the critical wavelength for a given viscosity law. The width of a box is $L_c/2$ and the reflecting boundary condition is applied for side boundaries. We then run a convection model with a range of Rayleigh number for a sufficiently long time (i.e., comparable to the diffusion timescale) and monitor the kinetic energy defined as

$$E = \int |\mathbf{v}|^2 dV, \quad (A2)$$

Figure A1. Numerical determination of critical Rayleigh number. Example for F-F, $\sigma = 6$, and $48 \times 48$ mesh. (a) Evolution of kinetic energy for a range of Rayleigh number. (b) Linear regression based on the growth rate, which is calculated for the period of $0.4 < r^* < 0.6$. 

<table>
<thead>
<tr>
<th>Type</th>
<th>$\sigma$</th>
<th>$L_c$</th>
<th>$Ra_c$(LSA)</th>
<th>$Ra_c$(FEM)$^b$</th>
<th>$Ra_c$(FEM)$^b$</th>
<th>$Ra_c$(FEM)$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-F</td>
<td>0</td>
<td>2.828</td>
<td>657.5</td>
<td>6.59 x 10^3(0.2)</td>
<td>6.58 x 10^3(0.1)</td>
<td>6.58 x 10^3(0.0)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.93</td>
<td>1.98 x 10^3</td>
<td>1.99 x 10^3(0.3)</td>
<td>1.98 x 10^3(0.2)</td>
<td>1.98 x 10^3(0.1)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3.21</td>
<td>6.91 x 10^3</td>
<td>6.92 x 10^3(0.2)</td>
<td>6.91 x 10^3(0.1)</td>
<td>6.90 x 10^3(0.1)</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>3.16</td>
<td>2.53 x 10^3</td>
<td>2.54 x 10^3(0.5)</td>
<td>2.53 x 10^3(0.1)</td>
<td>2.53 x 10^3(0.0)</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>2.24</td>
<td>8.40 x 10^3</td>
<td>8.47 x 10^3(0.9)</td>
<td>8.41 x 10^3(0.2)</td>
<td>8.39 x 10^3(0.1)</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>1.63</td>
<td>2.09 x 10^3</td>
<td>2.12 x 10^3(1.3)</td>
<td>2.09 x 10^3(0.1)</td>
<td>2.08 x 10^3(0.3)</td>
</tr>
<tr>
<td>R-R</td>
<td>0</td>
<td>2.016</td>
<td>1708</td>
<td>1.72 x 10^5(0.5)</td>
<td>1.71 x 10^5(0.2)</td>
<td>1.71 x 10^5(0.1)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.03</td>
<td>4.88 x 10^3</td>
<td>4.91 x 10^3(0.6)</td>
<td>4.89 x 10^3(0.3)</td>
<td>4.89 x 10^3(0.2)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2.02</td>
<td>1.48 x 10^4</td>
<td>1.49 x 10^4(0.6)</td>
<td>1.48 x 10^4(0.1)</td>
<td>1.48 x 10^4(0.0)</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1.96</td>
<td>4.37 x 10^4</td>
<td>4.41 x 10^4(0.9)</td>
<td>4.38 x 10^4(0.2)</td>
<td>4.37 x 10^4(0.1)</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>1.78</td>
<td>1.17 x 10^5</td>
<td>1.19 x 10^5(1.8)</td>
<td>1.18 x 10^5(0.5)</td>
<td>1.17 x 10^5(0.1)</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>1.58</td>
<td>2.73 x 10^5</td>
<td>2.81 x 10^5(2.8)</td>
<td>2.76 x 10^5(0.5)</td>
<td>2.74 x 10^5(0.3)</td>
</tr>
</tbody>
</table>

$^a$F-F denotes free-slip boundaries for both top and bottom surfaces; R-R denotes rigid surface boundaries.

$^b$Values in parentheses are deviation from LSA in percent.
from the entire domain of the model box. When a given Rayleigh number is below the critical value, the kinetic energy generated by initial perturbation should decrease and eventually becomes zero, and when it is above critical, it should increase. The critical Rayleigh number is therefore determined such that the growth rate, $dE/dt$, is zero for the system with that Rayleigh number (Figure A1).

We tested three different resolutions (with uniform quadrilateral elements), $32 \times 32$, $48 \times 48$, and $64 \times 64$, for $\sigma$ ranging from 0 (constant viscosity) to 10. A good agreement with the linear stability analyses is achieved as summarized in Table A1. For the combination of mesh resolution and the range of $\sigma$ employed in this study, numerical error is generally less than 1%.

[37] Acknowledgments. This work was sponsored by the U.S. National Science Foundation under grant EAR-0049044. We thank Rafi Katzman, Li Zhao, and Liqian Chen, whose work was the primary source of inspiration for our study. Mark Jellinek, Michael Manga, Anne Davaille, Claude Jaupart, and Sarah Zaranek provided helpful reviews on the earlier version of the manuscript. We also thank four anonymous official reviewers, whose input was helpful to improve the clarity of the manuscript.

References


Cazenave, A., S. Houry, B. Lago, and K. Dominh, Geosat-derived geoid


