

Growth of the hemispheric dichotomy and the cessation of plate tectonics on Mars

A. Lenardic

Department of Earth Science, Rice University, Houston, Texas, USA

F. Nimmo¹

Department of Geological Sciences, University College London, UK

L. Moresi

School of Mathematical Sciences, Monash University, Victoria, Australia

Received 22 August 2003; revised 21 October 2003; accepted 14 November 2003; published 7 February 2004.

[1] Although Mars is currently not tectonically active, it may have experienced plate tectonics early in its history. The southern hemisphere of Mars possesses a thick crust which probably renders the lithosphere positively buoyant. In this paper we present numerical and scaling arguments which show that if the area of positively buoyant lithosphere grows beyond a critical fraction ($\approx 50\%$ for Mars), plate tectonics will stop. Heat transfer through the buoyant lithosphere is inefficient, which causes mean mantle temperatures to increase as the surface area of buoyant lithosphere increases. The resulting reduction in mantle viscosity reduces shear stresses; if these shear stresses drop below the yield strength of the lithosphere, plate motions will cease and the planet will behave as a one-plate system. Thus the end of plate tectonics on Mars is a natural consequence of the growth of the southern highlands. Similar arguments for the Earth predict that it should operate in the plate tectonic regime now but that it may have experienced stagnant lid convection in the past. **INDEX TERMS:** 5475 Planetology: Solid Surface Planets: Tectonics (8149); 5430 Planetology: Solid Surface Planets: Interiors (8147); 5418 Planetology: Solid Surface Planets: Heat flow; 5455 Planetology: Solid Surface Planets: Origin and evolution; **KEYWORDS:** hemispheric dichotomy, mantle convection, thermal evolution, rheology

Citation: Lenardic, A., F. Nimmo, and L. Moresi (2004), Growth of the hemispheric dichotomy and the cessation of plate tectonics on Mars, *J. Geophys. Res.*, 109, E02003, doi:10.1029/2003JE002172.

1. Introduction

[2] An outstanding question regarding early Mars is the origin and time of formation of the hemispheric dichotomy [Zuber, 2001], the global contrast between thin crust in the Northern lowlands, and thick crust in the Southern highlands. One particularly intriguing possibility is that an early episode of plate tectonics was responsible for the development of the lowlands [Sleep, 1994]. The original hypothesis argued that plate tectonics ended when the mantle had cooled sufficiently, or when young, buoyant lithosphere approached a subduction zone. Here, we also assume that plate tectonics operated initially, but demonstrate that growth of a thick, buoyant crust provides an alternative explanation for the cessation of plate tectonics.

[3] In section 2 we discuss the basic observations, and previous work. In section 3, we develop a conceptual model for the effect on the convecting mantle of a chemically

buoyant layer partially covering the surface, and in section 4 we quantify this model. Section 5 compares the scaling arguments derived with numerical simulation results, and section 6 applies the scaling and numerical results to Mars. In section 7 we discuss uncertainties and applications to other planets, particularly Earth.

2. Observations and Previous Work

[4] The composition and thickness of the Martian crust is still poorly known. The southern highlands form $\approx 60\%$ of the planet, and are approximately 2–4 km higher than the northern lowlands. Assuming a mean crustal thickness of 50 km, the thickness contrast between northern lowlands and southern highlands is 20–30 km [Zuber *et al.*, 2000]. The mean thickness is unlikely to be less than 50 km, as this would result in exposed mantle material beneath impact basins [Zuber, 2001]. Gravity and topography studies of one area of the dichotomy yield a mean crustal thickness of 55 ± 20 km [Nimmo, 2002]. Similar studies by Turcotte *et al.* [2002] and Yuan *et al.* [2001] both obtain mean thicknesses of about 100 km for the Hellas basin and the planet, respectively. An upper bound on the mean thickness is

¹On leave at Department of Earth and Space Sciences, University of California, Los Angeles, California, USA.

approximately 100 km, on the basis of arguments relating to viscous topographic relaxation [Zuber *et al.*, 2000; Nimmo and Stevenson, 2001].

[5] Comparisons of Martian gravity and topography suggest that the near-surface crustal density can be as high as 3100 kg m^{-3} or as low as 2200 kg m^{-3} [McGovern *et al.*, 2002; Turcotte *et al.*, 2002; Nimmo, 2002; McKenzie *et al.*, 2002]. The high values are higher than typical terrestrial basaltic values, and are probably due to the iron-rich nature of the Martian mantle [Longhi *et al.*, 1992]. The low values are probably due to the presence of near-surface porosity or ice and are unlikely to represent the bulk crustal density. On the basis of in situ measurements and studies of SNC meteorites, it is likely that $\sim 50\%$ of the total budget of Martian radiogenic materials are in the crust [McLennan, 2001]. This is greater than estimates of the relative proportion of radiogenic materials in the Earth's crust. If correct, this distribution has significant effects on mantle convection and near-surface temperature gradients, which we will discuss below.

[6] SNC studies also allow the age of the crust to be constrained. Isotopic ratios suggest that the bulk of the crust formed within a few 100 Myr of 4.5 Gyr B.P. [Borg *et al.*, 1997; Blichert-Toft *et al.*, 1999; Marty and Marti, 2002]. Geological observations suggest a few km globally averaged subsequent melt production, decreasing with time [Greeley and Schneider, 1991]. Impact cratering studies can be used to date the relative and absolute surface ages of the highlands and lowlands [Tanaka *et al.*, 1992]. Recent estimates, including a large number of apparently buried craters, place the dichotomy as forming during the early Noachian, with only thin (1–2 km) crustal resurfacing taking place subsequently [Frey *et al.*, 2002]. On the basis of a recent cratering calibration [Hartmann and Neukum, 2001], the dichotomy was formed by 3.9 Gyr B.P.

[7] The origin of the dichotomy is uncertain, with both endogenic (e.g., mantle convection [Zhong and Zuber, 2001; McGill and Dimitriou, 1990] or plate tectonics [Sleep, 1994]) and exogenic (e.g., impacts [Wilhelms and Squyres, 1984]) processes being advocated. The current crustal thickness models do not support the impact hypothesis [Zuber, 2001]. Examination of the crater catalog of Barlow [1988] shows that impact crater densities often decline toward the dichotomy boundary. This effect may be due to outward growth of the dichotomy with time, but could also be due to increased erosion toward the lowlands. Identification of wrinkle ridges parallel to the dichotomy boundary imply that its formation included a degree of compression [Watters and Robinson, 1999], perhaps as a result of plate convergence. Conversely, Sleep [1994] argued that the dichotomy resembled a terrestrial passive margin.

[8] Breuer *et al.* [1993] investigated the effect of lateral variations in crustal heat production, and found that large lithospheric thickness variations could result. Zhong and Zuber [2001] showed that a strongly depth-dependent viscosity could result in a degree-1 upwelling which might explain the dichotomy. Wullner and Harder [1998] investigated the effect of lateral variations in crustal heat production in a spherical model with depth-dependent viscosity. They concluded that upwellings were likely to be focused beneath the (enriched) southern highlands, but that lateral variations in mantle heat flux were minor.

[9] Whether or not Mars possessed an early period of plate tectonics is debatable. The first proposal was based on relatively low-resolution geological mapping, and inferred plate tectonics to have existed in the northern lowlands [Sleep, 1994], although subsequent higher-resolution work cast doubt on the details of the interpretation [Pruis and Tanaka, 1995]. Conversely, the linear magnetic features in the southern highlands have been used to argue for plate tectonics in that region [Connerney *et al.*, 1999; Fairen *et al.*, 2002]. In either case, the Frey *et al.* [2002] cratering counts imply that plate tectonics must have ended within 0.7 Gyr of the formation of Mars.

[10] The predictions of the Sleep [1994] model are important because they differ from those presented here. In particular, Sleep [1994] argued that an initial highland-like crust covered the entire planet, and was subsequently subducted in the northern lowlands. The resulting crust was thinner and smooth because of the decrease in mantle temperature and relatively high spreading rate ($\sim 80 \text{ mm/yr}$). Plate tectonics continued until it was stopped by one of two factors: either the mantle cooled sufficiently that melting no longer occurred at mid-ocean ridges, or a spreading center came sufficiently close to a subduction zone that the lithosphere was too buoyant to subduct.

[11] Plate tectonics has been advocated on theoretical grounds for two reasons. Firstly, it is easier to drive a dynamo if plate tectonics operates [Nimmo and Stevenson, 2000], although an initially hot core is a viable alternative [Breuer and Spohn, 2003; Williams and Nimmo, 2003]. Secondly, plate tectonics provides a way of avoiding early massive melting on Mars [Hauck and Phillips, 2002], although it may make subsequent melt generation difficult to achieve [Breuer and Spohn, 2003]. If the first argument is correct, the timing of dynamo cessation provides a constraint on the time at which plate tectonics ended. On the basis of studies of an ancient SNC meteorite, the Martian dynamo existed at 3.9–4.1 Gyr B.P. [Weiss *et al.*, 2002]. Since the large, early Noachian impact basins Hellas and Argyre record no magnetic anomalies [Acuna *et al.*, 1999], the geodynamo failed or weakened before they formed, at around 3.9 Gyr B.P. Thus two lines of evidence suggest that plate tectonics, if it operated, must have ended within 0.7 Gyr of the planet's formation.

[12] The thickness of the mechanical boundary layer (MBL) on Mars at the present day is uncertain. On the basis of elastic thickness measurements, it is probably 200–300 km [McKenzie *et al.*, 2002]. Nimmo and Stevenson [2001] argued that 35 km of basaltic crust would be sufficient to render a 200 km thick lithosphere buoyant and unlikely to subduct. Since the MBL thickness will have been lower in the past, it is highly likely that the southern highland lithosphere (at least) would have been buoyant once the crust reached its present-day thickness. This lithospheric buoyancy is a key assumption in the model we develop below.

[13] Our approach is in some ways similar to that of Wullner and Harder [1998], but contains at least three crucial differences. Firstly, our approach assumes a strongly temperature-dependent viscosity. Secondly, we assume that the northern lowlands can participate in mobile-lid tectonics, while the southern highlands are prevented from doing so by the buoyancy of the crust. Thirdly, we do not

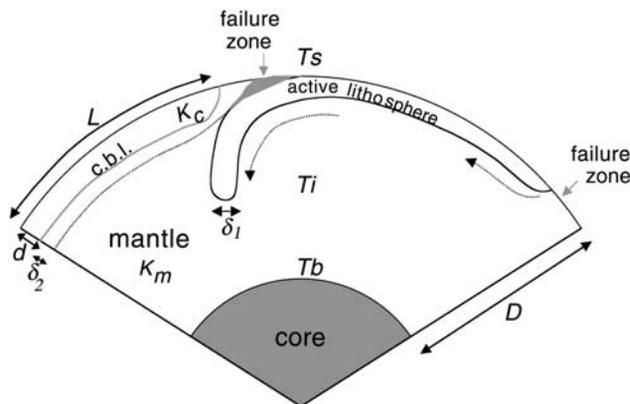


Figure 1. Cartoon of model scenario.

implicitly incorporate additional heat production within the crust, but instead use variable thermal conductivity as a proxy.

[14] In summary, the southern highland crust is 20–30 km thicker than that of the Northern lowlands, and is likely to have resulted in buoyant lithosphere. The growth of the crust and the dichotomy effectively ended within 0.7 Gyr of Mars' formation; plate tectonics, if it existed, also ceased at about the same time. The model we develop below will demonstrate that growth of the buoyant southern highlands is likely to have resulted in the cessation of plate tectonics.

3. Conceptual Model

[15] Our starting assumption is that an active-lid mode of mantle convection operated early in Martian history. Plate tectonics is an example of an active-lid mode in that the majority of the lithosphere, oceanic plates, partake in the convective overturn of the mantle and, as such, contribute significantly to its cooling. Plate tectonics was originally defined in terms of rigid plate interiors and concentrated deformation zones at plate boundaries. More recently it has become clear that diffuse plate boundaries are also a component of terrestrial tectonics [Gordon, 1998]. Although plate tectonics is an active-lid mode of mantle convection, not all active-lid modes need to have a surface manifestation like the plate tectonic system of Earth. One can imagine an active lid mode associated only with diffuse regions of deformation. What is key, in terms of mantle cooling, is that deformation zones form that allow the otherwise cold and strong upper boundary layer of the mantle to overturn. It is this last feature of active-lid convection that is key to our ideas. The exact mode of surface deformation, from highly concentrated to diffuse, will not be integral to our model.

[16] Several studies have explored the conditions required for the generation of an active-lid mode of mantle convection, such as plate tectonics [Bercovici, 1996; Moresi and Solomatov, 1998; Tackley, 1998; Trompert and Hansen, 1998]. These studies share a central idea: Localized lithospheric failure at a critical stress level leads to the formation of weak shear zones that allow otherwise strong lithosphere to participate in mantle convection. The shear zones are model analogues for plate boundaries. If convective stresses do not exceed the yield strength of the lithosphere then the dynamic zones of weakness cannot be maintained and a

stagnant lid, i.e., single plate, mode of convection results [Moresi and Solomatov, 1998]. Within the active lid mode the principal resistance to surface motion comes from the bulk interior viscosity of the mantle. In such a situation the appropriate convective stress scaling depends on the interior viscosity of the mantle [Moresi and Solomatov, 1998]. Thus a dramatic decrease in mantle viscosity over time (e.g., due to an increase in mantle temperature) could potentially cause convective stresses to drop below the level required to dynamically maintain plate boundaries.

[17] Although mantle temperatures normally decay with time, the growth of an insulating layer at the surface can potentially reverse this pattern. We assume that the growth of the southern highlands insulated the Martian mantle by covering a portion of it with a stable lithosphere that did not participate in mantle overturn due to its intrinsic buoyancy. This is akin to ideas of continental insulation on Earth [e.g., Gurnis, 1988]. At a minimum, we assume this chemical lithosphere to be composed of highland crust. We also allow for the possibility that it is composed of crust and a deeper, chemically distinct mantle residuum layer akin to terrestrial cratons [e.g., Boyd *et al.*, 1985]. The growth of the southern highlands, and associated insulation, could have caused a decrease in internal mantle viscosity that offset or exceeded the increase in viscosity associated with planetary cooling. This conjecture is the basis for our conceptual model.

[18] To restate: We consider a plate tectonic-like mode of mantle convection to have operated early in Mars' history at a time when the southern highlands were forming. At some stage the southern highlands grew beyond a critical extent. At this stage, the mantle insulation and resulting higher mantle temperature caused convective stresses to drop below the level required to cause lithospheric failure and the early stage of active-lid mantle convection was, effectively, locked. The next section discusses a quantitative model which is generally applicable. We apply this model specifically to Mars in section 6.

4. Quantitative Model

[19] Figure 1 shows a cartoon to define our notation. We consider the stable portion of the highland lithosphere to be composed of crust or crust plus a chemically buoyant mantle component. We will refer to this stable, chemical boundary layer as CBL. Since we consider this layer to be stable, the minimum thickness of the MBL within the highlands is implicitly assumed to be the same as the CBL thickness. The thickness and lateral extent of highland CBL are denoted by d and L , respectively. Mantle thickness is denoted by D . The surface and the mantle base temperatures are denoted by T_s and T_b , respectively. The average temperatures at the base of the highland CBL and within the mantle are denoted by T_c and T_i , respectively. The thermal conductivities of the highland CBL and the mantle are denoted by K_c and K_m , respectively. The thickness of the active upper thermal boundary layer of the northern lowlands and of the thermal mantle sublayer below the highland CBL are denoted by δ_1 and δ_2 , respectively. The thickness of the northern crust is assumed negligible for purposes related to planetary cooling.

[20] A viscoplastic mantle rheology is assumed [Moresi and Solomatov, 1998]. The rheology law remains on a

temperature-dependent viscous branch for stresses below a specified yield stress, τ_{yield} . Along this branch, diffusion creep is considered to be the deformation mechanism and the viscosity function is given by:

$$\mu_{creep} = A \exp[-ET] \quad (1)$$

where A and E are material parameters. For stresses above a yield stress the flow law switches to a plastic branch. The non-linear, effective viscosity along the plastic deformation branch is given by

$$\mu_{plastic} = \frac{\tau_{yield}}{I} \quad (2)$$

where I is the second strain-rate invariant.

[21] The rheologic formulation above has been used to model the generation of a plate-like mode of convection within a single chemical component mantle [Moresi and Solomatov, 1998]. In an active lid mode of convection, Moresi and Solomatov [1998] observed that the velocity of the upper boundary layer was controlled by the internal viscosity of the mantle, μ_i . For such a case the appropriate stress scale is

$$\tau \sim \frac{\mu_i \kappa}{\delta_1^2} \quad (3)$$

where κ is thermal diffusivity [Moresi and Solomatov, 1998]. On the basis of this form of a convective stress scale, Moresi and Solomatov [1998] developed a theoretical relationship that predicted the transition between active and stagnant lid convection as a function of mantle thermal and rheologic parameters. The criteria matched well with their results from a suite of numerical simulations.

[22] Moresi and Solomatov [1998] also presented a scaling for mantle heat flux, in the active lid regime, as a function of a Rayleigh number, Ra_i , defined using of internal mantle viscosity. The best fit scaling to simulation results was found to be

$$Nu = 0.385 Ra_i^{0.293}, \quad (4)$$

where Nu is the Nusselt number (a measure of non-dimensional heat flux). This is reasonably close to the theoretically expected form of $Nu = a_0 Ra_i^{1/3}$, where a_0 is a geometric scaling constant. To complete the scaling, the average internal mantle temperature must be known. For strong degrees of convective vigor, Moresi and Solomatov [1998] showed that the average internal temperature approached the mean of the surface T_s and base T_b temperatures, which is the theoretically expected value if the active lid system can be considered to behave as an equivalent isoviscous system. This internal temperature would change for the case of total internal heating. Unfortunately, the fraction of bottom heating in early Mars is unclear. However, the requirement of dynamo activity [Nimmo and Stevenson, 2000] and the significant crustal enrichment in radiogenic materials [McLennan, 2001] suggests that bottom heating was important.

[23] The relationships above form a starting point for our analysis. Below we determine the effects of adding a

highland CBL atop the convecting mantle. We first determine an expression for the surface heat flux and internal temperature for a system of the type shown in Figure 1 under the assumption that an active lid mode of convection is occurring in the northern lowlands. We will then use these results together with the stress scaling of Moresi and Solomatov [1998] to determine the conditions under which an active lid mode would no longer be able to operate.

[24] We consider the system of Figure 1 to be a thermal network composed of three heat transfer components. One component is associated with purely conductive heat transfer within the highland CBL. The second component is associated with heat transfer from the convecting mantle into the base of the highland CBL. Mantle heat is transferred across the thermal sublayer (thickness δ_2) that forms below the CBL. Although vertical heat transfer within this boundary layer is conductive its thickness, and thus the efficiency of this heat transfer component, is determined by the vigor of mantle convection. Components one and two, linked in series, form the local heat transfer path of the highlands. This highland path is linked in parallel to the final thermal component associated with mantle heat transfer in the northern lowlands. If an active lid mode of convection holds, internal heat will be transferred to the surface of the lowlands across the active upper thermal boundary layer (thickness δ_1) of the convecting mantle.

[25] We consider the end member cases of a planet completely covered by or devoid of a highland CBL to be associated with effective thermal resistances denoted by R_s and R_n , respectively. Intermediate cases will be associated with an effective network resistance which we will refer to as a total system resistance, R_t . The total system heat flux, q_t , can be expressed as the product of the average temperature drop associated with heat transfer from the interior mantle to the surface, ΔT_t , divided by the effective thermal resistance across the composite upper boundary layer. That is,

$$q_t(T_i) = \frac{\Delta T_t}{R_t(T_i)} \quad (5)$$

Similarly, the heat flux for the highland and lowland end-members can, respectively, be expressed as

$$q_s(T_{is}) = \frac{\Delta T_s}{R_s(T_{is})} \quad (6)$$

and

$$q_n(T_{in}) = \frac{\Delta T_n}{R_n(T_{in})}. \quad (7)$$

Here we are making it explicit that both heat flux and thermal resistance are temperature-dependent; T_{in} and T_{is} are the internal temperatures for each end-member case.

[26] For the full network, we will need to express ΔT_t in terms of the end-members ΔT_s and ΔT_n . To do so, we assume that the average internal temperature of the full network, T_i , will be a volume weighted average of T_{is} and T_{in} . We define A_s and A_n as, respectively, the nondimensional surface areas of the highlands and lowlands at a given time ($A_s + A_n = 1$). We anticipate the fact that our numerical

testing will be in a Cartesian domain so that the ratios of relative surface areas and mantle volumes will be equal. This leads to the expression

$$T_i = A_s T_{is} + A_n T_{in}. \quad (8)$$

An implicit assumption, contained within equation (8), is that efficient thermal mixing occurs within the convecting Martian mantle.

[27] We must also develop an expression for the resistance of the full thermal network. We assume highland and lowland paths to be linked in parallel and the effective resistance of each individual path to decrease as the surface area associated with it increases. We must also account for the fact that thermal resistance is a function of internal temperature. That is, a higher mantle temperature will lead to a lower internal viscosity which will allow for the more efficient overturn of lowland lithosphere. To account for this we note that the average internal viscosity for the full thermal network can be written as

$$\mu_i = A \exp[-E(A_s T_{is} + A_n T_{in})] \quad (9a)$$

while the average internal viscosity for lowland end-member case can be written as

$$\mu_{in} = A \exp[-E T_{in}]. \quad (9b)$$

As the highlands are insulating, μ_i/μ_{in} will be less than one. To determine the effects of this internal viscosity change on the effective resistance of the lowland path we must know the relationship between internal viscosity and heat flux in an active lid regime. The heat flux in the lowland end-member case is given by

$$q_n(T_{in}) = \frac{K_m(T_b - T_s)}{D} Nu \quad (10)$$

where Nu is given by equation (4). The results of *Moresi and Solomatov* [1998] show that in an active lid regime surface heat flux scales inversely with the internal viscosity to the 0.293 power, equation (4). We can use this scaling, together with equations (9a), (9b), and (10), to account for the added change in the resistance of the lowland path due to changes in internal mantle viscosity. Combining the assumptions above leads to an expression for the effective total system resistance given by

$$R_i(T_i)^{-1} = R_s(T_{is})^{-1} A_s + R_n(T_{in})^{-1} A_n (\exp[EA_s(T_{is} - T_{in})])^{0.293}. \quad (11)$$

Equations (5), (6), (7), (8) and (11) lead to an expression for the total system heat flux given by

$$q_i(T_i) = A_s q_s(T_{is}) \left[A_s + A_n \frac{T_{in}}{T_{is}} \right] + A_n q_n(T_{in}) \left[A_s \frac{T_{is}}{T_{in}} + A_n \right] \cdot (\exp[EA_s(T_{is} - T_{in})])^{0.293} \quad (12)$$

where we have set T_s to zero. The terms $[A_s \frac{T_{is}}{T_{in}} + A_n]$ and $(\exp[EA_s(T_{is} - T_{in})])^{0.293}$ represent changes in lowland heat flux due to the effect of the southern highland CBL on the

mean mantle temperature. As the highlands are considered insulating, both terms will be greater than or equal to unity. The first term accounts for the increase in lowland heat flux due to the increased temperature drop across the active mantle boundary layer of the lowlands. It does not rely on a specific rheologic assumption. The second accounts for the increase in lowland heat flux due to the reduction in mean mantle internal viscosity. It does rely on the rheologic assumption of equation (1) together with the assumption that an active lid mode of convection occurs.

[28] We now consider the highland heat flux q_s . For the highland end-member case, the effective Rayleigh number driving mantle convection will depend on the presence of a conducting chemical layer. We can write the effective Rayleigh number as

$$Ra_{\text{eff}} = \frac{\rho_0 g \alpha (T_b - T_c) (D - d)^3}{\mu_{is} \kappa} \quad (13)$$

where ρ_0 is the reference mantle density, g is gravitational acceleration, α is the thermal expansion coefficient, μ_{is} is the bulk internal viscosity, κ is the thermal diffusivity, and D is the depth of the mantle. As Ra_{eff} is not known a priori, a more standard Rayleigh number will be useful and it is defined as

$$Ra_S = \frac{\rho_0 g \alpha \Delta T D^3}{\mu_S \kappa} \quad (14)$$

where $\Delta T = T_b - T_s$ and μ_S is the mantle viscosity assuming the surface temperature value (we could equivalently use a Rayleigh number based on basal viscosity, Ra_b). The two Rayleigh numbers are related by

$$Ra_{\text{eff}} = \frac{(T_b - T_c)}{\Delta T} (1 - d/D)^3 \frac{\mu_S}{\mu_{is}} Ra_S. \quad (15)$$

[29] By expressing the effective driving Rayleigh number in terms of the μ_{is} we are assuming the dominant resistance to sub-highland mantle convection is coming from the internal viscosity of the mantle. This implies that either the mantle temperature at the CBL-mantle interface is sufficiently high that the base of the CBL is within the rheological transition layer or that plastic failure occurs within the cold upper boundary layer of the mantle that forms below the CBL. The latter is consistent with our previous assumption for the northern lowlands.

[30] The effects of heat source enrichment within the highland CBL can be approximated by allowing it to have a lower thermal conductivity than the mantle [Busse, 1978]. We will use this approach, as opposed to directly allowing for a variable distribution of internal heat sources, for simplicity and to maintain consistency with *Moresi and Solomatov* [1998] who assumed a bottom heated mantle. Following this approach, and assuming that thermal equilibrium has been reached, allows us to equate the average heat flux into the base of the highland CBL, q_c , to the average surface heat flux. We can thus express q_c as

$$q_c = K_c \frac{T_c - T_s}{d}. \quad (16)$$

[31] We assume that, on average, a nearly linear thermal gradient holds across the active upper thermal boundary

layer of the mantle that forms below the highlands. Recall that the average thickness of this layer is denoted by δ_2 . The average temperature drop across it is $T_{is} - T_c$. The mantle heat flux, q_m , can now be written as

$$q_m = K_m \frac{T_i - T_c}{\delta_2}. \quad (17)$$

We assume thermal equilibrium is reached so that q_c and q_m are equal (and equal to q_s). This allows us to equate equations (16) and (17) and thus provides an expression for δ_2 given by

$$\delta_2 = \frac{(T_{is} - T_c)K_m d}{(T_c - T_s)K_c}. \quad (18)$$

[32] We introduce a local boundary layer Rayleigh number, Ra_δ , which we assume remains close to a constant, critical value, Ra_{crit} . This follows from ideas introduced by *Howard* [1966] for systems with no CBL. Our definition of Ra_δ will differ from *Howard's* [1966] as we nondimensionalize by the total system depth versus the convecting layer depth (for the no CBL case the two are equivalent). The definitions must be equivalent as the effects of the CBL go to zero. Although the effects of the CBL become negligible as d goes to zero, this limit does not discriminate between potential definitions of Ra_δ . The CBL effects also become negligible as K_c becomes large relative to K_m . If we defined Ra_δ in terms of Ra_S or Ra_b and considered the limit of K_m/K_c going to zero, then the definition would imply that as d increased the boundary layer thickness would be unchanged, i.e., it would not depend on the depth of the convecting layer ($D - d$). This is not consistent with previous results [*Howard*, 1966]. We thus define Ra_δ in terms of Ra_{eff} , which does maintain consistency. The definition is given by

$$Ra_\delta = \frac{\delta_2^3 (T_{is} - T_c)}{D^3 \Delta T} Ra_{eff}. \quad (19)$$

The critical Rayleigh number for the onset of convection is $\approx 10^3$ for a range of boundary conditions [*Sparrow et al.*, 1964]. However, the local boundary layer Rayleigh number can differ from the critical Rayleigh number for convective onset [*Sotin and Labrosse*, 1999]. To allow for this, we consider $Ra_\delta = a_1 10^3$ where a_1 is a scaling constant. The above leads to a second expression for δ_2 given by

$$\delta_2^3 = \frac{a_1 10^3 D^3 \Delta T}{(T_{is} - T_c) Ra_{eff}}. \quad (20)$$

[33] The problem can be closed by expressing T_{is} in terms of T_c and T_b . The result depends on the boundary conditions at the top and bottom of the convecting region. As we have assumed the highland CBL to be long lived and stable, the mechanical condition the mantle feels below it will be effectively rigid. If the Martian core is liquid, however, the boundary condition at the base will be free-slip. Appendix A derives the relationship between T_{is} , T_c and T_b under these conditions and obtains the result:

$$T_{is} = 0.64T_b + 0.36T_c. \quad (21)$$

[34] Equations (15), (18), (20), and (21) together with our rheologic assumption, equation (1), allow us to derive

an expression for T_c in terms of known parameters. The expression can be simplified by nondimensionalizing the system using D as the length scale, K_m as the conductivity scale, and ΔT as the temperature scale (this allows us to set the nondimensional temperature at the system surface and base to zero and one, respectively). We also nondimensionalize the mantle viscosity in terms of the maximum viscosity contrast defined using mantle viscosities that would result for temperatures at the system surface and base. This allows us to set the pre-exponential constant in equation (1) to unity. The final expression is

$$\frac{(1 - T_c)^5}{T_c^3 \exp[-.36ET_c]} = \frac{a_1 5960.46 K_c^3 \exp[-.64E]}{(d - d^2)^3 Ra_S}. \quad (22)$$

Together with equations (16) and (21), this allows us to solve for the surface heat flux $q_s(T_{is})$ and the average internal mantle temperature T_{is} in terms of known parameters and thus completes the analysis for the southern highlands end-member.

[35] In order to derive the heat-flux for cases intermediate between the two end-members (equation (12)), we need to relate $q_s(T_{is})$ to $q_s(T_i)$. The internal mantle temperature T_i for the full network will be lower than that for the pure highland end-member T_{is} . This effect will increase mantle viscosity which will tend to thicken the mantle sublayer below the highlands. A similar rheologic effect was introduced for the lowlands in equation (11). As equation (22) is a transcendental equation, a simple expression cannot be written for the highland effect but the correction to the highland resistance can be treated using an iterative approach. However, we already know that for the highlands the effect will be weaker than for the lowlands. The reason for this is because the highland path is itself a composite. If the thermal resistance of the highland CBL is much greater than that of the mantle sublayer below, then changes in the thickness of the sublayer will have a negligible effect on the full highland resistance. For the parameter range we are interested in, it can be demonstrated that changes in the mantle sublayer thickness have little effect.

[36] To review: equations (16), (21), and (22) can be used to determine the mantle heat flux $q_s(T_{is})$ and internal temperature T_{is} for the southern highlands end-member case while equations (4) and (10) can be used for the northern lowlands end-member variables $q_n(T_{in})$ and T_{in} . The effects of linking these paths to form the full thermal system are given by equations (8), (11), and (12). Equation (12) can be used to determine the average global mantle heat flux q_t for the full system and also the average local mantle heat flux in the highlands and the lowlands. The latter, together with the internal mantle temperature T_i from equation (8), can be used to determine the thickness of the active upper mantle boundary layer of the lowlands δ_1 . The internal mantle temperature can also be used with equation (1) to determine the bulk internal mantle viscosity, μ_i . Convective stress levels can be solved for using δ_1 and μ_i values together with the stress scaling for active lid convection, equation (3). We can thus address the question of whether the growth of the highlands could shut down a

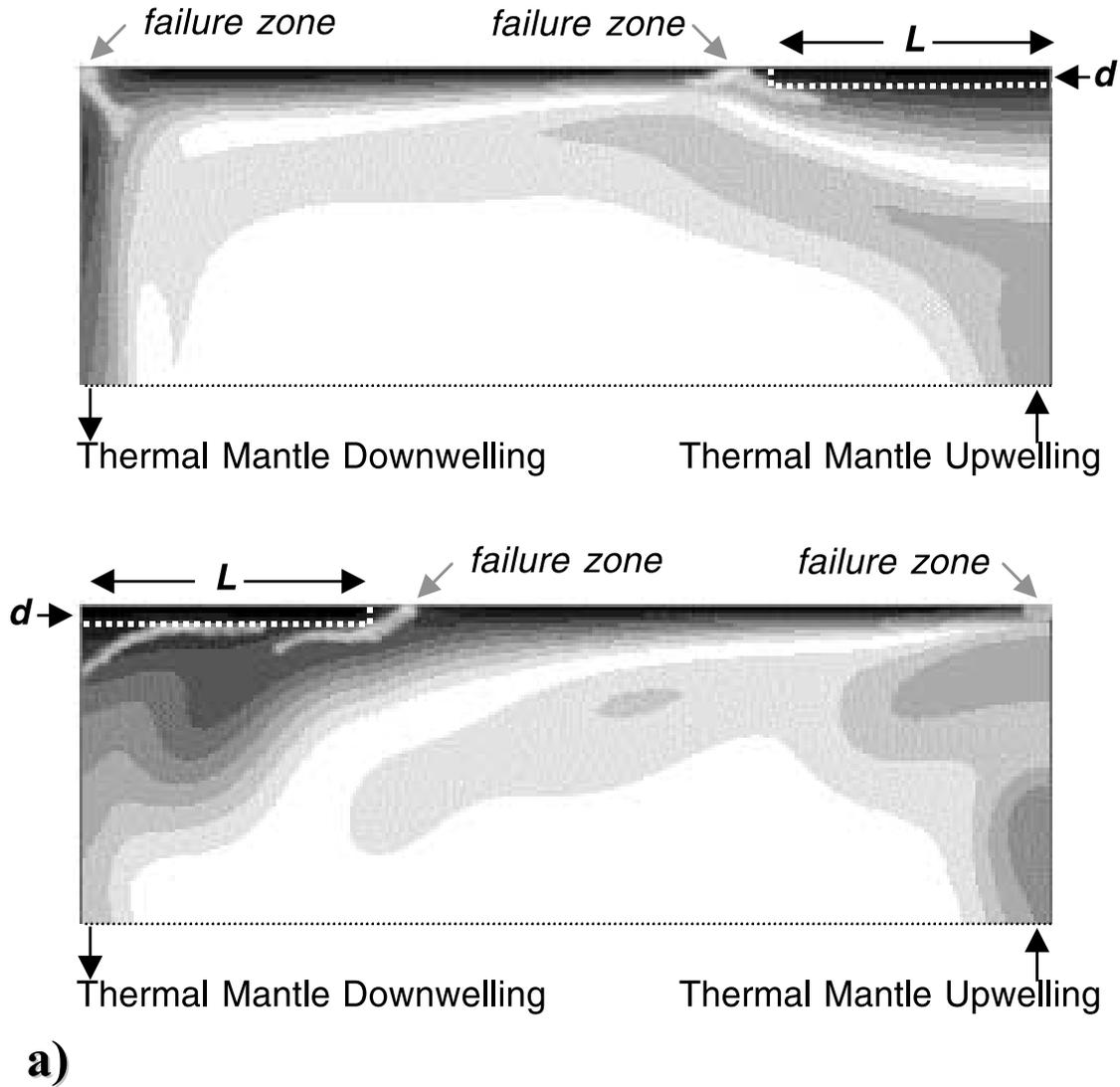


Figure 2. (a) Image plots from numerical simulations. The plots show the upper 1/3 of the full modeling domain. Results from two suites of simulations for global heat flux (b), internal temperature, and the horizontal velocity of lowland lithosphere (c), as a function of highland extent.

plate tectonic-like mode of convection. Before doing this, the theory must be tested to determine its validity.

5. Numerical Testing

[37] To test the ideas of the previous section, we compared our scaling predictions to the results of numerical simulations. The CITCOM finite element code was used [Moresi and Solomatov, 1995]. The simulations are similar to those of Moresi and Solomatov [1998] except that they include a conducting surface layer of variable depth and extent which will represent the stable highland CBL (Figure 2a).

[38] The full non-dimensional system of equations solved by the CITCOM code are given by:

$$\partial_i u_i = 0 \quad (23)$$

$$\partial_j [2\mu(T, \tau_{yield}, D)\epsilon_{ij}] = \partial_i p + Ra_b T \hat{k} \quad (24)$$

$$\partial_i T + u_i \partial_i T = \partial_i^2 T \quad (25)$$

$$\rho = [1 - \alpha(T - T_0)] \quad (26)$$

where

$$Ra_b = \frac{\rho_0 g \alpha \Delta T D^3}{\mu_b \kappa}$$

and u_i is the velocity vector, $\mu(T, C, \tau_{yield}, I)$ is a viscosity function, T is temperature, τ_{yield} is a material yield stress, I is the second invariant of the strain rate tensor, ϵ_{ij} is the strain rate tensor, p is pressure, Ra_b is the basal Rayleigh number, \hat{k} is the vertical unit vector, ρ is the density, α is the coefficient of thermal expansion, g is the gravitational acceleration, ΔT is the temperature drop across the system, D is the system depth, μ_b is the viscosity at the system base, and κ is thermal diffusivity. Subscripts of zero indicate reference values. Equation (23) is the mass conservation equation assuming incompressible material, equation (24) is the momentum conservation equation for creeping flow, equation (25) is the energy conservation equation assuming no internal or frictional heating, and equation (26) is the linearized equation of state. The rheology for the mantle is given by equations (1) and (2) of the previous section.

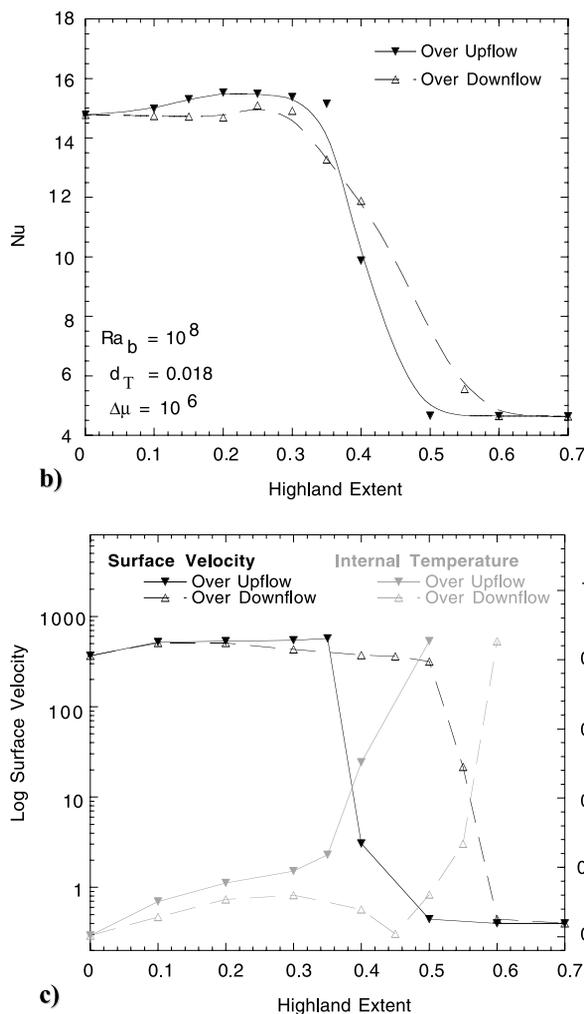


Figure 2. (continued)

[39] Thermal boundary conditions for all the simulations are constant surface and basal temperatures of zero and one, respectively. Mechanical boundary conditions are reflecting, free slip side walls and free slip upper and lower boundaries. Two means of representing the stable southern highland CBL were tested and each led to equivalent results. The first set the viscosity of the highland CBL to a relatively high value while the second prescribed zero velocity conditions in the region of the highland CBL.

[40] Although we assume that the southern highlands formed during the early operation of a plate tectonic-like mode of convection, the exact manner by which the crust formed is not an element of our scaling theory. As discussed in section 2, the highlands may have formed over either an upwelling or a downwelling. Numerical simulation suites were therefore performed that imposed a southern highland CBL either above a zone of mantle upflow or downflow (Figure 2a).

[41] The simulations of Figure 2a impose a stable CBL of variable depth and extent atop a convecting temperature dependent viscoplastic mantle. The non-dimensional extent of the CBL will be referred to as ϵ . For each simulation suite, the full convection problem is first solved assuming no

highland CBL present ($\epsilon = 0$). The yield stress, for the parameters of the suite, is set to 66% of the transition value determined by *Moresi and Solomatov* [1998]. This ensures that a plate-like mode of convection can operate without an imposed dichotomy. The thermal field from the $\epsilon = 0$ case is then used as the initial thermal condition for a case that imposes a highland CBL with a fixed depth and a lateral extent of $\epsilon = 0.05$. This case is run to a statistically steady-state and is then used as the initial condition for the next simulation of the suite in which the extent of the dichotomy is progressively increased by 0.05. For each simulation within a given suite the average heat flux, internal mantle temperature, and horizontal velocity in the mobile part of the lid are plotted against ϵ (Figures 2b and 2c). The fixed parameters for the simulation suites of Figure 2 are the yield stress, as discussed above, the basal Rayleigh number, Ra_b , the top to bottom viscosity ratio, $\Delta\mu$, and the thermal thickness of the CBL, d_T . This last quantity is defined as dK_m/DK_c .

[42] Figures 2b and 2c show that a transition does occur as ϵ exceeds a critical value. In terms of mantle heat loss, a high and a low heat flux regime can exist, separated by a transitional regime (Figure 2b). The high and low heat flux regimes are associated with an active or a stagnant lowland lithosphere, respectively. The horizontal surface velocity within the lowlands shows most clearly that an active to stagnant lid transition occurs as ϵ exceeds ≈ 0.4 or 0.5, depending on whether highland growth is assumed to occur over a zone of mantle upwelling or downwelling (Figure 2c). As ϵ increases, the mantle temperature increases (Figure 2c), due to the insulation effect. Initially, this increase in temperature causes an increase in the velocity of the lowland lithosphere, due to the reduction in mantle viscosity (Figure 2c). As the mantle viscosity drops further, however, the convective stresses drop below the yield stress, at which point a transition to stagnant lid convection occurs. In the stagnant lid regime, the lithosphere velocity is low, and the entire lithosphere is thus effectively insulating, causing the more rapid rise in internal mantle temperature (Figure 2c) and the fall in surface heat flux (Figure 2b).

[43] The transitional regime between the active and stagnant end-members was found to occur over a narrower ϵ band if the highlands form over a mantle upwelling. For those cases, the transitional regime was associated with an episodic mode of convection akin to that described by *Moresi and Solomatov* [1998]. Long episodes of stagnant lid convection, during which the entire upper thermal boundary layer thickened, were punctuated by rapid boundary layer overturn events, i.e., short bursts of active lid convection. The “over the downwelling” cases went through two modes of time-dependence as they transitioned from an active to a stagnant lid. Figure 2a (bottom) shows the first time-dependent mode. The boundary layer below the stable highlands experienced episodes of thickening followed by instability. The time scale associated with the instabilities was much shorter than the long time scale associated with the episodic mode mapped by *Moresi and Solomatov* [1998]. The short time scale instability mode occurred from ϵ values of 0.35 to 0.45 and slightly enhanced interior cooling (Figure 2c). After this, the downwelling cases also entered into a long time scale transitional mode [*Moresi and Solomatov*, 1998] followed by a stagnant lid mode.

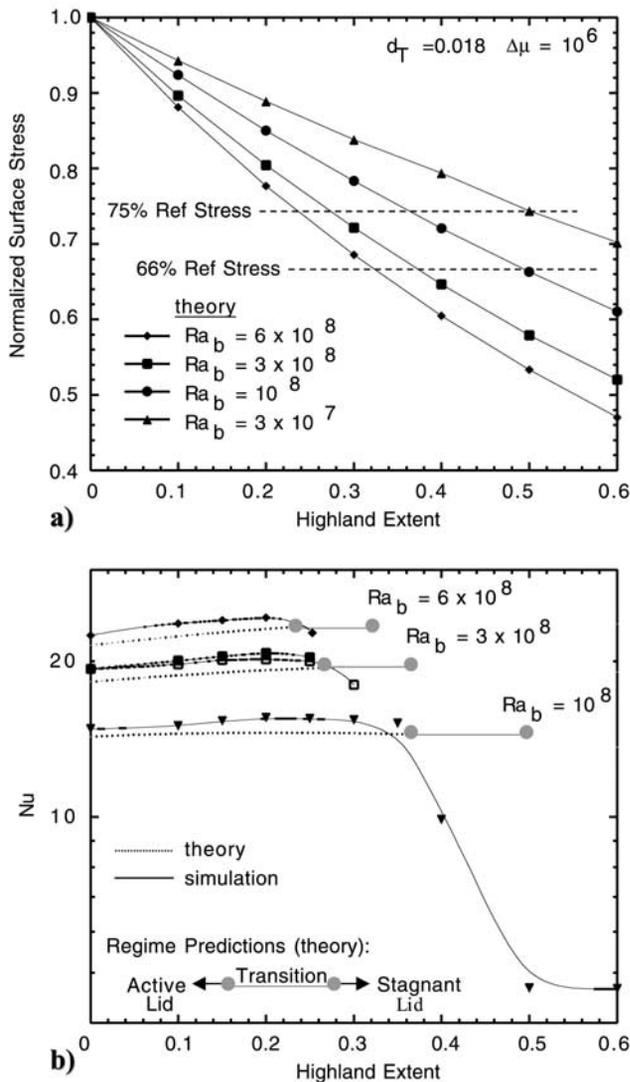


Figure 3. (a) Scaling theory predictions for the effects of the dichotomy on convective mantle stress. Stresses for each set of cases are normalized by the stress level that would result with no dichotomy assuming all other parameters to be equal. (b) Numerical simulation results compared to theoretical predictions. For each Ra_b set of models the yield stress is set to be lower than that needed to allow for plate like behavior in the case of no dichotomy. The yield stress is set to 66% of the active to stagnant lid transition value determined by *Moresi and Solomatov* [1998]. The predicted stress fall, due to dichotomy growth, from Figure 3a gives the theoretical point for each set of experiments at which the transition should start. The simulation results for the two higher Ra_b sets are shown up to the point that an active lid holds. For each set, adding more dichotomy beyond those points puts them in the transitional regime.

[44] Figure 3 compares the results from several suites of numerical models to predictions based on the scaling theory of the previous section. The yield stress for each numerical suite is set to 66% of the value required to cause transition to a stagnant lid convection mode when $\epsilon = 0$ [*Moresi and Solomatov*, 1998]. The yield stress was held fixed for any suite and ϵ then increased as described for Figure 2. Our

theory was used to predict the drop in convective stress due to increased ϵ (Figure 3a). For each suite the reference stress is the convective stress at $\epsilon = 0$. The point at which increased ϵ is predicted to cause convective stress to drop to 66% of its reference value, for any suite, is shown by a dashed horizontal line on Figure 3a. This is the point at which the theory predicts that a stagnant lid regime should hold for any suite. We consider this the point at which the transition from an active to a stagnant lid is complete. Also shown is the point at which convective stress is predicted to drop to 75% of the reference value for any suite. This is motivated by Figure 2 which shows that the transition in convective behavior, with increased ϵ , can be broad. For the simulations of Figure 2, the point at which the transition initiates, based on a pronounced decrease in heat flux, was found to be at a stress value of $\approx 75\%$ times the reference stress. This initiation stress showed a mild dependence on the position of the lowlands relative to mantle upwellings or downwellings.

[45] Figure 3b compares the scaling-derived Nusselt numbers with those derived from the numerical models. Except for the $Ra_b = 10^8$ case, the numerical results are shown up to the point at which any given suite had just entered into the transitional regime as shown by the rollover of the $Nu - \epsilon$ curves. Also shown are theoretical curves for total system heat flux using equation (12). The curves are shown up to the point at which the transitional regime is predicted to initiate on the basis of Figure 3a. The predicted position and width of the transitional regime, for any suite, is determined from Figure 3a using the 75% stress drop as the initiation point and the 66% stress drop as the completion point. To the left of the transition zone an active lid mode is predicted. The theoretical curves match the numerical curves fairly well. The theoretical curves demonstrate that for high Ra the presence of highlands can actually increase global heat flux slightly if active lid convection occurs. The simulations confirm that the presence of highlands can increase global heat flux. This somewhat counterintuitive result occurs because the highland insulation lowers internal mantle viscosity which allows for the more efficient overturn of active lowland lithosphere, and counteracts the reduction in local heat flux through the highland lid.

[46] Numerical suites that varied d_T from 0.009–0.11 and $\Delta\mu$ from $1-10^6$ were also performed and compared to theoretical predictions with results comparable to those shown in Figure 3b. The reasonable match between theory and simulations gives us confidence that the main physical assumptions of the theory are valid and that the theory can provide reasonable first-order predictions.

[47] As well as testing assumptions, comparison to numerical simulations also highlights some uncertainties. The theory has made no geometric assumption. The free scaling constant of equation (22) was constrained to a value of 0.1 by direct comparison of theory predictions to our 2-D, Cartesian simulations. This implies a low boundary layer Rayleigh number relative to the critical Rayleigh number for convective onset which is in accord with the results of *Sotin and Labrosse* [1999]. The geometric constant will depend on the characteristic wavelength of convective cells. We have not systematically explored wavelength effects so the inherent assumption, in terms of theory application, is that as the dichotomy grows to a critical point, the waveband of mantle convection does not drastically change. The constant

will also depend on the geometry of the domain, Cartesian versus spherical. We have not constrained the constant for spherical cases, although results from other authors suggest that the differences are likely to be small [e.g., Reese *et al.*, 1999]. Spherical versus Cartesian geometry will also result in a lower average internal mantle temperature, all other factors remaining equal. This will shift the reference internal temperature associated with the situation of a planet lacking lateral surface variations. In applying our scalings we assume that the effects of a growing dichotomy on increasing the internal temperature above a reference value are not overly sensitive to system geometry. A minor source of uncertainty is the fact that the transition between active and stagnant lid convection is not sharp. Figure 3b shows that the transition occurs over a range of ϵ of about 0.1 at relatively low Ra_b ; the sharpness of the transition increases as Ra_b increases. For application, the effects of these uncertainties are assumed to be minor relative to those associated with physical parameters, which the theory does constrain. The above should be kept in mind when considering theory application.

6. Application

[48] The theoretical and numerical results may be summarized as follows. The main effect of a partial insulating lid is to increase the mean interior temperature of the mantle. This increased temperature leads to a decreased viscosity, and hence decreased viscous stresses. If these stresses fall below the yield stress of the lithosphere, plate tectonics will stop. The critical value of ϵ decreases with increasing Ra (and increasing Nu). The effect of the lid extent on mean heat flux depends on whether the reduction in mantle viscosity or the insulating properties of the lid are more important. Below, we will apply these results to Mars.

6.1. Mars Parameters

[49] Table 1 gives parameters relevant to conditions early in Martian history. The most important uncertainties for this work are the values of Ra_b , $\Delta\mu$, and d_T . For a buoyant layer 50 km thick (see section 2), d_T is about 0.025 when $K_m/K_c = 1$. If the bulk of the heat producing elements are within the crust, or the buoyant layer is 100 km thick, d_T increases to about 0.05.

[50] The value of E may be obtained by [Solomatov, 1995]

$$E = \frac{Q}{RT_i^2}$$

where Q is the activation energy and R is the gas constant. For likely mantle internal temperatures of 1700–1800 K [Nimmo and Stevenson, 2000; Hauck and Phillips, 2002] and using the parameters given in Table 1, $E \approx 0.01 - 0.02 K^{-1}$ and $\Delta\mu$ is $\sim 10^6 - 10^{14}$.

[51] The initial value of Ra_b depends principally on the mantle viscosity, which is not well known. Nimmo and Stevenson [2000] obtained a starting viscosity of $\sim 10^{16}$ Pa s, while Hauck and Phillips [2002] used $\sim 10^{18}$ Pa s. The present-day viscosity of the Martian mantle is probably comparable to or larger than that of the Earth, $\sim 10^{21}$ Pa s. Using the other parameters in Table 1, the initial value of Ra_b was probably in the range $10^{10} - 10^{12}$, declining to $\sim 10^7$ through time.

Table 1. Parameters Adopted for Mars

Quantity	Value	Units	Reference ^a
E	240–480	kJ mol^{-1}	1, 2
T_s	200	K	1
K_m	3.2	$\text{W m}^{-1} \text{K}^{-1}$	1
D	1950	km	1
ρ_0	3400	kg m^{-3}	1
g	3.7	m s^{-2}	1
α	4×10^{-5}	K^{-1}	1
κ	8×10^{-7}	$\text{m}^2 \text{s}^{-1}$	1

^aReferences: 1, Nimmo and Stevenson [2000]; 2, Hauck and Phillips [2002].

[52] Nimmo and Stevenson [2000] and Spohn *et al.* [2001] obtain values for Nu of 40–50 after 500 Myrs of plate tectonics, and up to 80 at earlier times. While these models did not include a CBL, Figure 3 shows that the effect of such a lid on Nu is quite small, and thus that similar values should apply. Hauck and Phillips [2002] obtained a value of $Nu \approx 20$ because they assumed stagnant lid convection was operating at this time.

6.2. Results

[53] Figure 4 illustrates the relationship between Nu and the critical value of ϵ also shown in Figure 3b. For a constant Nu , an increase in ϵ leads to a transition between active lid and stagnant lid convection, as the stresses decrease. Similarly, for a constant ϵ , an increase in Nu leads to the same transition, for the same reason.

[54] Figure 4 shows that for Nu in the range 40–50, appropriate to early Mars, the critical value of ϵ required to halt plate tectonics is 0.5–0.65, similar to the observed present-day value. A higher value of Nu (i.e., earlier growth of the buoyant lithosphere) would reduce the critical value of ϵ required. Thus, if the insulating southern highlands grew to their current size in the first 0.5 Gyr of Martian history, it is likely that their growth was responsible for the cessation of plate tectonics.

[55] It is important to examine the uncertainties in this conclusion arising from parameter uncertainties. Figure 4 shows that the value of $\Delta\mu$ has a rather small effect on the critical value of ϵ . Figure 5a shows the effect of varying d_T . Larger values of d_T result in a smaller critical value of ϵ , because the insulating effect is more pronounced. This effect is less important at high Nu , because the bulk of the heat is transferred by the lowland mantle. Reducing d_T from 0.056 to 0.014 (and thus spanning the likely Martian values) increases the critical value of ϵ from 0.3 to 0.4 for $Nu = 75$.

[56] Figure 5b shows the effect of the assumed yield stress on the critical value of ϵ . Increasing the yield stress dramatically reduces the critical value of ϵ , because the reduction in stress required to end active lid tectonics is smaller for larger yield stress. Unfortunately, the effective yield stress of the lithosphere is poorly constrained even for the Earth. For the Martian case, we must consider it as effectively a free parameter given the lack of any independent constraints.

7. Discussion and Conclusions

[57] We have argued above that the growth of the bulk of the southern highlands ($\epsilon \approx 0.5$) is likely to have stopped an early episode of plate tectonics. This growth was apparently

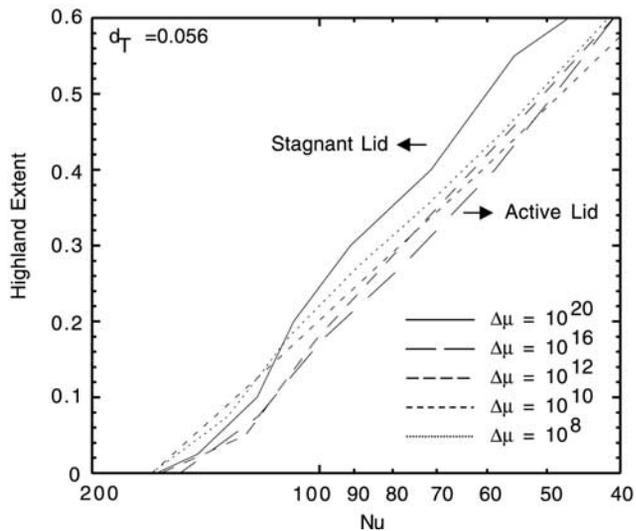


Figure 4. The predicted active to stagnant lid transition curves in terms of global mantle Nusselt number and dichotomy extent for various degrees of temperature dependent mantle viscosity. An Ra point in the planet's history at which plate tectonics initiates is assumed. The yield stress is set to be equal to the convective stress at that point. Planetary cooling is associated with lowering of the Ra values. At each lower Ra point the internal viscosity is increased accordingly which on its own increases convective stress. The extent of the dichotomy which would cause the convective stress to drop below the yield stress at each Ra point is calculated as is the Nusselt number at that point and these values determine the transition curves.

completed within the first 0.7 Gyr of Martian history (see section 2). Mantle cooling takes place over longer time-scales, and leads to a long-term decrease in Nu . Figure 4 shows that a decrease in Nu at constant ϵ ought to lead to a transition from stagnant lid to active lid tectonics, which has obviously not happened on Mars. Mode transitions are an important aspect of planetary thermal evolution in general [Sleep, 2000]; it is therefore of interest to investigate why this particular transition has not taken place.

[58] A potential explanation is that the effective yield stress of the lithosphere has changed with time. For instance, the effective yield stress of Venus is much higher than that of the terrestrial lithosphere due to the absence of water on the former planet [e.g., Kaula, 1994]. This has the effect of increasing fault friction and making relative lithospheric motion harder to achieve. Mars, however, contains abundant water in at least the shallow subsurface. Furthermore, Barnett and Nimmo [2002] argued that faults on Mars need be no stronger than terrestrial faults, suggesting that fault friction is not a likely answer. A more likely possibility is that the effective yield stress increases as the rigidity of the lithosphere increases [Conrad and Hager, 1999]. Older terrains on Mars generally show lower effective elastic thicknesses than more recently loaded terrains, suggesting that the lithospheric rigidity has indeed increased with time [e.g., McGovern et al., 2002]. It is also possible that the yield stress itself does not depend on elastic thickness but that a thicker elastic core effectively lowers the maximum convective stress within the lithosphere. That

is, a thicker elastic core could result in convective stress being distributed over a larger depth range while a thinner elastic lithosphere would promote a stress concentration effect which would increase the peak stress within the lithosphere [e.g., Kuszniir and Bott, 1977].

[59] An alternative possibility is that further crustal growth took place after the cessation of plate tectonics [e.g., Spohn et al., 2001]. This is a reasonable suggestion, as the presence of an insulating lid leads to higher mantle temperatures (Figure 2c) and thus increased melt production. In particular, if the crust in the northern lowlands were sufficiently thickened that the lithosphere there also became buoyant, then plate tectonics would have been unable to initiate subsequently. The post-Noachian cover of the northern plains is probably only 1–2 km thick [Frey et al., 2002], but the amount of underplating is unknown.

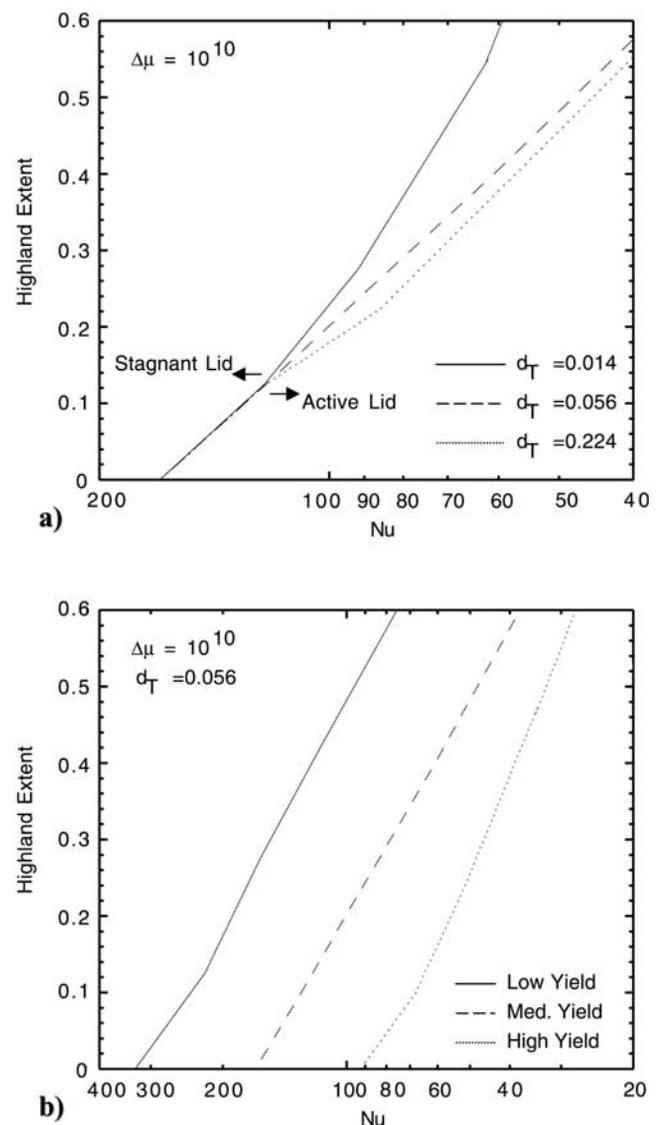


Figure 5. (a) The predicted active to stagnant lid transition curves for various thermal thicknesses of the southern highland CBL. (b) The predicted active to stagnant lid transition curves for different assumptions as to the yield strength of the Martian lithosphere.

[60] The models presented here show that growth of a thick highland crust will lead to an increase in mantle temperature, stopping plate tectonics. Conversely, *Sleep* [1994] posited a reduction in the total area of highland crust (due to subduction) and a decrease in mantle temperature ending plate motion. Future mapping of age variations across the southern highlands should thus provide one way of discriminating between the two mechanisms. Similarly, the mechanisms have differing implications for the subsequent thermal evolution of the planet. In particular, the rate of subsequent melt generation would be greater in the present model than in the *Sleep* [1994] model. Unfortunately, the rate of melt production with time on Mars is very poorly constrained [*Greeley and Schneid*, 1991], although there have been attempts to model the observations [*Hauck and Phillips*, 2002; *Breuer and Spohn*, 2003].

[61] *Nimmo and Stevenson* [2000] demonstrated that an early episode of plate tectonics was capable of driving a Martian dynamo, and that a cessation in plate tectonics would halt the dynamo. An unsatisfactory aspect of their model was that the end of plate tectonics was imposed by fiat; here we have demonstrated that growth of an insulating lid, beyond a critical extent, naturally shuts off plate tectonics. Thus the end of the dynamo may be indirectly linked to the growth of the southern highlands.

[62] We can thus summarize a potential tectonic history for Mars as follows. Mars initially underwent active lid tectonics, driving a geodynamo and producing the southern highlands through some mixture of melt generation and crustal accretion. The growth of this area of thickened crust produced an area of buoyant lithosphere, which increased the mean temperature of the mantle. The consequent reduction in stress led to the cessation of plate tectonics. The mantle subsequently cooled slowly in the stagnant lid regime, but plate tectonics never re-initiated, probably because of the increased rigidity or increased buoyancy of the lithosphere.

[63] While we have focused on Mars, it is also of interest to consider the effect of an insulating lid on the other terrestrial planets. For Earth, the present-day Nu is about 40, the value of ϵ is about 0.35, and d_T is in the range 0.01–0.03. Comparison with Figure 5a shows that Earth is well within the active lid regime, which is reassuring. However, this diagram does also suggest that earlier in its history, when $Nu > 80$ –90, the Earth may well have operated in the stagnant lid regime. Although oceanic crust has been identified in Archean ophiolites [e.g., *Moore*, 2002], it has never been clear whether or not plate tectonics operated during this interval [e.g., *Hamilton*, 1998]. The transition from stagnant to active lid convection on Earth may have been possible because the rigidity of the terrestrial lithosphere is a factor of 10–100 lower than the present-day rigidity of Mars.

[64] The crust of Venus is basaltic, perhaps 30 km thick on average, and relatively uniform across the planet [*Kaula*, 1994]. It is probably insufficiently buoyant to prevent plate tectonics on its own. However, as noted above, the effective yield strength of the Venusian lithosphere is so high, due to the absence of water, that the convective stresses may never have been high enough to have initiated plate tectonics.

[65] We have demonstrated both numerically and analytically that the growth of an insulating lid will lead to a

cessation in plate tectonics, because the resulting increase in mantle temperature causes the convective stresses to drop below the critical yield stress. We have applied our model to Mars, and shown that for reasonable parameter choices an insulating lid extent of $\approx 50\%$, similar to that observed, was sufficient to end plate tectonics early in Martian history. In theory, plate tectonics could have subsequently re-started as the mantle cooled and convective stresses increased. This did not occur, probably because the effective yield stress or buoyancy of the lithosphere increased through time. For the case of Earth, our models predict that plate tectonics should be operating at the present day, but that a stagnant lid regime may have operated in the past, when convective stresses were lower.

Appendix A

[66] We wish to establish the relationship between internal temperature T_i , bottom temperature T_b and the temperature at the base of the CBL, T_c , given that the mechanical boundary conditions are free-slip at the base and rigid at the top. We proceed by considering the limiting case of CBL thickness going to zero while the boundary conditions are unchanged. The temperature drops across the upper and lower boundary layers of the mantle are denoted by ΔT_u and ΔT_l , respectively. The thickness of the upper and lower boundary layers are denoted by δ_u and δ_l . The boundary layer Rayleigh numbers for the upper and lower boundary layers are denoted by Ra_{δ_u} and Ra_{δ_l} . We again consider boundary layer thickness to scale with the boundary layer Rayleigh number to the $-1/3$ power [*Howard*, 1966]. We also assume that the heat flow rate across the upper and lower boundary layers is equal. Thus

$$\frac{\Delta T_u}{\Delta T_l} = b_1 \left[\frac{Ra_{\delta_u}}{Ra_{\delta_l}} \right]^{1/3} \quad (\text{A1})$$

where b_1 is a scaling constant. If we follow *Howard's* [1966] general assumption that the boundary layer Rayleigh number remains near a critical value and further follow his more specific assumption that this value is the same as the critical Rayleigh number for the onset of convection then the expression becomes

$$\frac{\Delta T_u}{\Delta T_l} = b_1 \left[\frac{1707}{657} \right]^{1/3} \quad (\text{A2})$$

where we have used the critical Rayleigh numbers for convective onset for rigid and free boundaries [*Sparrow et al.*, 1964]. The scaling constant should be near unity. If we consider the system to be nondimensionalized such that the nondimensional surface temperature is zero and the base temperature is one, then this would predict a bulk internal nondimensional temperature of 0.58. For comparison, a value of 0.5 is predicted for vertically symmetric mechanical boundary conditions.

[67] We ran a series of numerical simulations in the no CBL limit with a rigid upper and a free slip lower boundary to test the above. For an Ra range of 10^4 – 10^9 the bulk internal temperature approached a nondimensional value of 0.64 for the higher Ra cases (i.e., $Ra = 10^6$ – 10^9). This

suggests that the scaling constant b_1 is 1.3. We adopt this value and assume that boundary layer asymmetry described by equation (A2) will continue to hold for the active upper and lower boundary layers in the situation where a stable CBL is present above the convecting mantle. This allows us to write an expression for the bulk internal temperature (equation (23) in the main text).

[68] **Acknowledgments.** Supported by NASA-MDAP Grant NAG5-12166 and the Royal Society. Thanks to Shijie Zhong and Norm Sleep for constructive reviews that led to an improved paper.

References

- Acuna, M. H., et al. (1999), Global distribution of crustal magnetization discovered by the Mars Global Surveyor MAG/ER experiment, *Science*, *284*, 790–793.
- Barlow, N. G. (1988), Crater size-frequency distributions and a revised Martian relative chronology, *Icarus*, *75*, 285–305.
- Barnett, D. N., and F. Nimmo (2002), Strength of faults on Mars from MOLA topography, *Icarus*, *157*, 34–42.
- Bercovici, D. (1996), Plate generation in a simple model of lithosphere-mantle flow with dynamic self-lubrication, *Earth Planet. Sci. Lett.*, *144*, 41–51.
- Blichert-Toft, J., J. D. Gleason, P. Telouk, and F. Albarede (1999), The Lu-Hf isotope geochemistry of shergottites and the evolution of the Martian mantle-crust system, *Earth Planet. Sci. Lett.*, *173*, 25–39.
- Borg, L. E., L. E. Nyquist, L. A. Taylor, H. Wiesmann, and C. Y. Shih (1997), Constraints on Martian differentiation processes from Rb-Sr and Sm-Nd isotopic analyses of the basaltic shergottite QUE 94201, *Geochim. Cosmochim. Acta*, *61*, 4915–4931.
- Boyd, F. R., J. J. Gurney, and S. H. Richardson (1985), Evidence for a 150–200 km thick Archean lithosphere from diamond inclusion thermobarometry, *Nature*, *315*, 387–389.
- Breuer, D., and T. Spohn (2003), Early plate tectonics versus single-plate tectonics on Mars: Evidence from magnetic field history and crust evolution, *J. Geophys. Res.*, *108*(E7), 5072, doi:10.1029/2002JE001999.
- Breuer, D., T. Spohn, and U. Wullner (1993), Mantle differentiation and the crustal dichotomy of Mars, *Planet Space Sci.*, *41*, 269–283.
- Busse, F. H. (1978), A model of time-periodic mantle flow, *Geophys. J. R. Astron. Soc.*, *52*, 1–12.
- Connerney, J. E. P., M. H. Acuna, P. J. Wasilewski, N. F. Ness, H. Reme, C. Mazelle, D. Vignes, R. P. Lin, D. L. Mitchell, and P. A. Cloutier (1999), Magnetic lineations in the ancient crust of Mars, *Science*, *284*, 794–798.
- Conrad, C. P., and B. H. Hager (1999), Effects of plate bending and fault strength at subduction zones on plate dynamics, *J. Geophys. Res.*, *104*, 17,551–17,571.
- Faïren, A. G., J. Ruiz, and F. Anguita (2002), An origin for the linear magnetic anomalies on Mars through accretion of terranes: Implications for dynamo timing, *Icarus*, *160*, 220–223.
- Frey, H. V., J. H. Roark, K. M. Shockey, E. L. Frey, and S. E. H. Sakimoto (2002), Ancient lowlands on Mars, *Geophys. Res. Lett.*, *29*(10), 1384, doi:10.1029/2001GL013832.
- Gordon, R. G. (1998), The plate tectonic approximation: Plate nonrigidity, diffuse plate boundaries, and global plate reconstructions, *Annu. Rev. Earth Planet. Sci.*, *26*, 615–642.
- Greeley, R., and B. D. Schneider (1991), Magma generation on Mars—Amounts, rates, and comparisons with Earth, Moon, and Venus, *Science*, *254*, 996–998.
- Gurnis, M. (1988), Large-scale mantle convection and the aggregation and dispersal of supercontinents, *Nature*, *332*, 695–699.
- Hamilton, W. B. (1998), Archean magmatism and deformation were not products of plate tectonics, *Precambrian Res.*, *91*, 143–179.
- Hartmann, W. K., and G. Neukum (2001), Cratering chronology and the evolution of Mars, *Space Sci. Rev.*, *96*, 165–194.
- Hauck, S. A., II, and R. J. Phillips (2002), Thermal and crustal evolution of Mars, *J. Geophys. Res.*, *107*(E7), 5052, doi:10.1029/2001JE001801.
- Howard, L. N. (1966), Convection at high Rayleigh number, in *Applied Mechanics, Proceedings of the 11th Congress of Applied Mechanics, Munich (Germany)*, edited by H. Gortler, pp. 1109–1115, Springer-Verlag, New York.
- Kaula, W. M. (1994), The tectonics of Venus, *Philos. Trans. R. Soc. London, Ser. A*, *349*, 345–355.
- Kuszniir, N. J., and M. H. P. Bott (1977), Stress concentration in the upper lithosphere caused by underlying visco-elastic creep, *Tectonophysics*, *43*, 247–256.
- Longhi, J., E. Knittle, J. R. Holloway, and H. Wanke (1992), The bulk composition, mineralogy and internal structure of Mars, in *Mars*, edited by H. H. Kieffer et al., pp. 345–382, Univ. of Ariz. Press, Tucson.
- Marty, B., and K. Marti (2002), Signatures of early differentiation of Mars, *Earth Planet. Sci. Lett.*, *196*, 251–263.
- McGill, G. E., and A. M. Dimitriou (1990), Origin of the Martian global dichotomy by crustal thinning in the late Noachian or early Hesperian, *Geology*, *28*, 391–394.
- McGovern, P. J., S. C. Solomon, D. E. Smith, M. T. Zuber, M. Simons, M. A. Wieczorek, R. J. Phillips, G. A. Neumann, O. Aharonson, and J. W. Head (2002), Localized gravity/topography admittance and correlation spectra on Mars: Implications for regional and global evolution, *J. Geophys. Res.*, *107*(E12), 5136, doi:10.1029/2002JE001854.
- McKenzie, D., D. N. Barnett, and D. N. Yuan (2002), The relationship between Martian gravity and topography, *Earth Planet. Sci. Lett.*, *195*, 1–16.
- McLennan, S. M. (2001), Crustal heat production and the thermal evolution of Mars, *Geophys. Res. Lett.*, *28*, 4019–4022.
- Moores, E. M. (2002), Pre-1 Ga (pre-Rodinian) ophiolites: Their tectonic and environmental implications, *Geol. Soc. Am. Bull.*, *114*, 80–95.
- Moresi, L.-N., and V. S. Solomatov (1995), Numerical investigations of 2D convection with extremely large viscosity variations, *Phys. Fluids*, *7*, 2154–2162.
- Moresi, L.-N., and V. S. Solomatov (1998), Mantle convection with a brittle lithosphere: Thoughts on the global styles of the Earth and Venus, *Geophys. J. Int.*, *133*, 669–682.
- Nimmo, F. (2002), Admittance estimates of mean crustal thickness and density at the Martian hemispheric dichotomy, *J. Geophys. Res.*, *107*(E11), 5117, doi:10.1029/2000JE001488.
- Nimmo, F., and D. Stevenson (2000), The influence of early plate tectonics on the thermal evolution and magnetic field of Mars, *J. Geophys. Res.*, *105*, 11,969–11,979.
- Nimmo, F., and D. Stevenson (2001), Estimates of Martian crustal thickness from viscous relaxation of topography, *J. Geophys. Res.*, *106*, 5085–5098.
- Pruis, M. J., and K. L. Tanaka (1995), The Martian northern plains did not result from plate tectonics, *Lunar Planet. Sci.*, *XXVI*, 1147–1148.
- Reese, C. C., V. S. Solomatov, J. R. Baumgardner, and W. S. Yang (1999), Stagnant lid convection in a spherical shell, *Phys. Earth Planet. Inter.*, *116*, 1–7.
- Sleep, N. H. (1994), Martian plate tectonics, *J. Geophys. Res.*, *99*, 5639–5655.
- Sleep, N. H. (2000), Evolution of the mode of convection within terrestrial planets, *J. Geophys. Res.*, *105*(E7), 17,563–17,578.
- Solomatov, V. S. (1995), Scaling of temperature- and stress-dependent viscosity convection, *Phys. Fluids*, *7*, 266–274.
- Sotin, C., and S. Labrosse (1999), Three-dimensional thermal convection in an iso-viscous, infinite Prandtl number fluid heated from within and from below: Applications to the transfer of heat through planetary mantles, *Phys. Earth Planet. Inter.*, *112*, 171–190.
- Sparrow, E. M., R. J. Goldstein, and V. K. Jonsson (1964), Thermal instability in a horizontal fluid layer: Effect of boundary condition and non-linear temperature profile, *J. Fluid Mech.*, *18*, 513–528.
- Spohn, T., M. H. Acuna, D. Breuer, M. Golombek, R. Greeley, A. Halliday, E. Hauber, R. Jaumann, and F. Sohl (2001), Geophysical constraints on the evolution of Mars, *Space Sci. Rev.*, *96*, 231–262.
- Tackley, P. J. (1998), Self-consistent generation of tectonic plates in three-dimensional mantle convection, *Earth Planet. Sci. Lett.*, *157*, 9–22.
- Tanaka, K. L., D. H. Scott, and R. Greeley (1992), Global stratigraphy, in *Mars*, edited by H. H. Kieffer et al., pp. 345–382, Univ. of Ariz. Press, Tucson.
- Trompert, R., and U. Hansen (1998), Mantle convection simulations with rheologies that generate plate-like behavior, *Nature*, *395*, 686–689.
- Turcotte, D. L., R. Shcherbakov, B. D. Malamud, and A. B. Kucinskas (2002), Is the Martian crust also the Martian elastic lithosphere?, *J. Geophys. Res.*, *107*(E11), 5091, doi:10.1029/2001JE001594.
- Watters, T. R., and M. S. Robinson (1999), Lobate scarps and the Martian crustal dichotomy, *J. Geophys. Res.*, *104*, 18,981–18,990.
- Weiss, B. P., H. Vali, F. J. Baudenbacher, J. L. Kirschvink, S. T. Stewart, and D. L. Shuster (2002), Records of an ancient Martian magnetic field in ALH84001, *Earth Planet. Sci. Lett.*, *201*, 449–463.
- Wilhelms, D. E., and S. W. Squyres (1984), The Martian hemispheric dichotomy may be due to a giant impact, *Nature*, *309*, 138–140.

- Williams, J.-P., and F. Nimmo (2003), Thermal evolution of the Martian core: Implications for an early dynamo, *Geology*, in press.
- Wullner, U., and H. Harder (1998), Convection underneath a crust inhomogeneously enriched in heat sources: Application to Martian mantle dynamics, *Phys. Earth Planet. Inter.*, *109*, 129–150.
- Yuan, D. N., W. L. Sjogren, A. S. Konopliv, and A. B. Kucinskis (2001), Gravity field of Mars: A 75th degree and order model, *J. Geophys. Res.*, *106*, 23,377–23,401.
- Zhong, S., and M. T. Zuber (2001), Degree-1 mantle convection and the crustal dichotomy on Mars, *Earth Planet. Sci. Lett.*, *189*, 75–84.
- Zuber, M. T. (2001), The crust and mantle of Mars, *Nature*, *412*, 220–227.
- Zuber, M. T., et al. (2000), Internal structure and early thermal evolution of Mars from Mars Global Surveyor topography and gravity, *Science*, *287*, 1788–1793.

A. Lenardic, Department of Earth Science, Rice University, MS 126, P.O. Box 1892, Houston, TX 77251-1892, USA. (adrian@esci.rice.edu)

L. Moresi, School of Mathematical Sciences, Monash University, Building 28, Victoria 3800, Australia. (louis.moresi@monash.edu)

F. Nimmo, Department of Geological Sciences, University College London, Gower Street, London WC1E 6BT, UK. (nimmo@ess.ucla.edu)