13

Thermal History of the Earth

13.1 Introduction

Mantle convection plays an essential role in determining the evolution of the Earth's temperature through geologic time because it is the primary mechanism by which the Earth transfers heat from its deep interior to its surface. Once the internally generated heat reaches the surface it is transferred to the ocean–atmosphere system by a variety of processes including conduction and hydrothermal circulation through the oceanic crust and is eventually radiated to space. From the perspective of studying the changes in the Earth's interior temperature over geologic time, we can ignore the relatively rapid transport of internal heat through the atmosphere and oceans and assume that all heat delivered to the Earth's surface from below immediately escapes the Earth. The heat lost through the Earth's surface tends to cool the interior, and heat produced within the Earth by the decay of radioactive elements tends to warm it. The thermal evolution of the Earth is a consequence of the competition between internal energy sources producing heat and mantle convection removing it. A quantitative description of the Earth's thermal history is the application of basic energy conservation in a convecting mantle.

While the basic approach to modeling the Earth's thermal history is straightforward, its implementation is a major challenge because of the complexity of a realistic model and available computer resources that limit detailed numerical calculations of three-dimensional, time-dependent convection at the very high Rayleigh numbers applicable to the Earth's present mantle and at the even higher Rayleigh numbers appropriate to the Earth's early mantle. These limitations were discussed in detail in Chapter 10. All the complexities of mantle convection discussed in the previous chapters of this book indicate the severe limitations of any attempt to model the thermal evolution of the Earth. However, despite these limitations, relatively simple models of the Earth's thermal evolution have provided very useful results.

Concepts and results from boundary layer theories of convection and from a large number of numerical and laboratory experiments on convection have been incorporated into Earth thermal history models as a way of accounting for the effects of convective heat transfer across the mantle. This approach is known as parameterized convection and has made possible the study of the thermal evolution of the Earth with essentially analytic models (Sharpe and Peltier, 1978, 1979; Schubert, 1979; Schubert et al., 1979a, b, 1980; Stevenson and Turner, 1979; Turcotte et al., 1979; Davies, 1980; Turcotte, 1980b). The approach uses simple parameterizations between the amount of heat generated in the mantle and the vigor of mantle convection required to extract this heat.

13.2 A Simple Thermal History Model

13.2.1 Initial State

It is now generally accepted that the Earth formed by accretion (Safronov, 1969; Levin, 1972; Greenberg et al., 1978; Wetherill, 1985; Ahrens, 1990), and that upon completion of the accumulation process the Earth was hot and fully differentiated into a mantle and core with the core superliquidus and the mantle near its solidus (Schubert, 1979; Schubert et al., 1979a, b, 1980; Stevenson et al., 1983; Stevenson, 1989b, 1990). The early heat source is gravitational potential energy made available by accretion (Wetherill, 1976, 1985; Weidenschilling, 1976; Safronov, 1978; Kaula, 1979a) and core formation (Birch, 1965; Tozer, 1965b; Flasar and Birch, 1973; Shaw, 1978) contemporaneous with or shortly following accretion (Stevenson, 1981, 1989b, 1990).

The gravitational potential energy per unit mass released upon accretion can be estimated using $E = 3GM/5R$, the gravitational potential energy per unit mass of a constant density body of mass M and radius R (here, G is the universal gravitational constant). For the Earth, $E = 3.75 \times 10^7\,\mathrm{J\,kg^{-1}}$. The equivalent temperature T^* is found using $E = cT^*$, and for $c = 1\,\mathrm{kJ\,kg^{-1}}$, $T^* = 37{,}500\,\mathrm{K}$. The key requirement for core formation during or just after accretion is the retention of a small fraction, say 20%, of the energy of impacting planetesimals by the Earth. The likelihood of this is high if large impactors played a significant role in accretion, since large impacts lead to deep burial of a substantial fraction of the impactor's kinetic energy (Wetherill, 1976, 1985, 1986; Kaula, 1979a; Melosh, 1990). Formation of the Moon may have been one of the consequences of such a large impact with a planetesimal the size of Mars in the late stages of Earth's accretion (Hartmann and Davis, 1975; Cameron and Ward, 1976; papers in Hartmann et al. (1986) discuss the Great Impact Hypothesis for the origin of the Moon; see also Stevenson, 1987) (see the discussion in Chapter 14 on the Moon). The collision of the Earth with a Mars size impactor would release about $7.5 \times 10^6\,\mathrm{J\,kg^{-1}}$ and raise the average temperature of the Earth by 7,500 K if all this energy went into heating the Earth (Melosh, 1990). This energy is enough to have melted, even vaporized, a large part of the Earth. The Moon is supposed to have accreted in orbit around the Earth from terrestrial and impactor material ejected during the cataclysmic collision event. It is generally assumed that the Earth was already differentiated into a core and mantle at the time of the giant impact in order to explain the chemical similarity between the Moon and the Earth's mantle (e.g., Wänke and Dreibus, 1986). Spohn and Schubert (1991) estimate that the Earth would have re-equilibrated (thermally and structurally) on a time scale of 1–10 Myr after the giant impact. This rapid adjustment makes the giant impact event inconsequential for the long-term thermal evolution of the Earth. Though truly cataclysmic at the time, the only trace of the giant impact at the present may be the Moon itself and a slightly altered chemical composition of the Earth's mantle.

The gravitational potential energy released upon core formation is also large, enough to raise the temperature of the whole Earth by 2,000 K (e.g., Birch, 1965; Tozer, 1965b; Flasar and Birch, 1973). Radioactivity could also contribute to heating of the Earth early in its evolution if significant amounts of certain extinct radionuclides, i.e., Aluminum 26, were incorporated into the accreting Earth.

There are many fundamental but unanswered questions about the early evolution of the Earth. For example, the amount of energy available from the sources discussed above would be more than sufficient to melt the entire Earth. However, if the entire mantle were molten it would be expected that solidification would lead to a chemically fractionated planet, a thick

enriched crust, and a depleted mantle. There is no evidence that this occurred (Ringwood, 1990). There may have been a global magma ocean beneath a massive protoatmosphere of water (Abe and Matsui, 1985, 1986, 1988; Matsui and Abe, 1986a, b, c, 1987; Ahrens, 1990), but very rapid subsolidus mantle convection could have maintained most of the mantle at a temperature just below its solidus (Davies, 1990). In addition, convection in the magma ocean may have prevented fractionation by keeping crystals in suspension as the magma ocean solidified (Tonks and Melosh, 1990; Solomatov and Stevenson, 1993a, b). From this point of view, the magma ocean could have been very deep, i.e., a large part of the mantle could have been molten, without fractionation occurring on solidification.

Question 13.1: *Was the Earth's mantle fractionated at the end of accretion?*

On the basis of the above considerations, we adopt a simplified Earth thermal history model consisting of a two-layer Earth with a core and a compositionally homogeneous mantle. The structure is established at time zero (the start of the thermal history) and is unchanged throughout the evolution. The initial thermal state is hot; the core is superliquidus and the mantle is at the solidus. It will be seen that the subsequent thermal evolution of the model Earth consists of an early period of rapid cooling lasting several hundred million years followed by more gradual cooling over most of geologic time. After the period of rapid cooling the subsequent thermal evolution is nearly independent of the initial temperature distribution.

13.2.2 Energy Balance and Surface Heat Flow Parameterization

In this section we derive a thermal history model for the mantle using the assumption that no heat enters the mantle from the core. This assumption, made here to obtain the simplest possible model, is relaxed in the more sophisticated models discussed later in the chapter. From the viewpoint of the energy balance, this assumption leads to a one-layer Earth model.

An integration of the heat equation (e.g., 6.9.13) over the whole mantle gives

$$Mc\frac{\partial T}{\partial t} = MH - Aq \tag{13.2.1}$$

where M is the mass of the mantle, c is the specific heat of the mantle, T is the volume-averaged mantle temperature, H is the average rate of energy release in the mantle per unit mass due to the decay of long-lived radioactive elements (^{238}U, ^{235}U, ^{40}K, and ^{232}Th), A is the outer surface area of the mantle, and q is the average heat flux at the top of the mantle. The integrated energy balance (13.2.1) simply states that the time rate of change of mantle internal thermal energy is balanced by the difference between the heat production rate in the mantle and the rate of heat loss through the surface. In performing this integration we characterize the mantle with a single uniform temperature T and a uniform distribution of radiogenic heat sources. Long-lived radioactivity is an important source of heat for the mantle over geologic time; it is widely accepted that mantle radioactivity is the source of most (e.g., 80%) of the heat flowing through the Earth's surface at present (Turcotte and Schubert, 1982).

Question 13.2: *What fraction of the Earth's surface heat flow can be attributed to radioactive heat generation and what fraction to secular cooling of the Earth?*

We now assume that the specific radiogenic heat production rate H decays with time according to an exponential decay law with a single rate constant λ:

$$H = H_0 e^{-\lambda t} \tag{13.2.2}$$

where H_0 is the specific heat production rate at $t = 0$.

Substitution of (13.2.2) into (13.2.1) gives

$$Mc\frac{\partial T}{\partial t} = MH_0 e^{-\lambda t} - Aq \tag{13.2.3}$$

In order to solve (13.2.3) we require a heat transfer law relating q to the other model variables and parameters, especially the average mantle temperature T. We follow Schubert et al. (1979a, 1980) and specify this dependence in the following parameterized form:

$$q = \frac{k(T - T_s)}{d}\left(\frac{Ra}{Ra_{cr}}\right)^{\beta} \tag{13.2.4}$$

where k is the thermal conductivity of the model mantle, d is the thickness of the mantle, T_s is the surface temperature, Ra is the Rayleigh number given by

$$Ra = \frac{g\alpha(T - T_s)d^3}{\kappa\nu} \tag{13.2.5}$$

Ra_{cr} is the critical value of the Rayleigh number for the onset of convection in the spherical shell, and β is a constant. In (13.2.5), g is the acceleration of gravity in the model mantle, taken to be a constant as is appropriate to the real mantle, α is the assumed constant value of thermal expansivity in the model mantle, κ is the mantle thermal diffusivity also assumed constant, and ν is the kinematic viscosity in the mantle. The viscosity ν is a function of temperature, a dependence that controls the thermal evolution, as elaborated below.

Equation (13.2.4) is the Nu–Ra relation ($Nu = qd/k(T - T_s)$) characteristic of boundary layer theories of convection and of numerous numerical and laboratory experiments on convection as discussed in Section 8.6. A constant of order unity has been tacitly incorporated into Ra_{cr} which typically has a value of order 10^3. The power-law exponent β generally has a value of about 0.3 according to boundary layer theory and a large number of numerical experiments. Boundary layer theory gives $\beta = 1/3$ (Sections 8.6 and 13.5), while experiments give a slightly smaller value of β. The specific form (13.2.4) of the surface heat flow parameterization is suggested by boundary layer theory and experiments on convection of a constant viscosity, Boussinesq fluid in a plane layer heated from below. Its applicability to other situations is surprisingly robust and has been discussed in detail by Schubert et al. (1979a, 1980). The Nu–Ra relations of other heating modes (e.g., internal heating) and geometries (e.g., spherical geometry) can all be written in the form of (13.2.4) by appropriate definitions and identifications of Ra_{cr} and β. The use of (13.2.4) for convection with temperature-dependent viscosity is generally appropriate if T is identified with the characteristic temperature of the convecting part of the fluid, Ra_{cr} and β are properly interpreted, and T_s is chosen as either the surface temperature or the temperature near the

base of any stagnant lid that forms over the convecting system (Schubert et al., 1979a, 1980). For the parameterization of Earth's mantle convection, T_s is properly taken as the surface temperature since the plates are mobile and do not form a stagnant lid over the convecting system.

Criticism of the use of (13.2.4) in Earth thermal history studies has been provided by both Richter and McKenzie (1981) and Christensen (1984c, 1985b) concerning effects of strongly temperature dependent viscosity. The former paper was mainly concerned with the influence of a stagnant upper thermal boundary layer, which does not occur on Earth. The application of (13.2.4) to the Earth or to another planet depends on a proper identification of T_s and T, as discussed above. If there are large viscosity variations within thermal boundary layers that lie within the convecting part of the fluid, e.g., the hot, low-viscosity thermal boundary layer at the bottom of the fluid, then the use of (13.2.4) can be modified appropriately as discussed below. Christensen (1985b) was also concerned with the effects of a sluggish or nearly stagnant lid and on the basis of numerical experiments inferred a very small value of β (about 0.1). This result would make Earth thermal history rather insensitive to Ra or the viscosity of the mantle.

It is now recognized that there are distinct modes of convection in fluids with strongly temperature dependent viscosity – the small viscosity contrast regime, the sluggish-lid regime, and the stagnant-lid regime (Moresi and Solomatov, 1995; Solomatov, 1995; Ratcliff et al., 1997). Separate Nu–Ra parameterizations have been developed for each flow regime (Solomatov, 1995; Reese et al., 1998, 1999). These parameterizations will be discussed in more detail later in this chapter and in Chapter 14 where application is made to the thermal histories of terrestrial planets. The Nu–Ra parameterizations for the sluggish-lid and stagnant-lid convection regimes are relevant to other planets, e.g., Venus, Mars, and the Moon, which lack plate tectonics and are in these convection regimes. Plate tectonics places the Earth in the small viscosity contrast convection regime and (13.2.4) applies. Parameterized convection based on (13.2.4), with a value of β around 0.3, provides the most physically plausible representation of the Earth's thermal evolution (Gurnis, 1989).

13.2.3 Temperature Dependence of Mantle Viscosity and Self-regulation

The strong dependence of mantle viscosity on temperature exerts a controlling influence on the evolution of the mantle. It is consistent with the approximate nature of parameterized convection modeling to assume a Newtonian rheology with a kinematic viscosity ν related to mantle temperature by

$$\nu = \nu_0 \exp\left(\frac{A_0}{T}\right) \tag{13.2.6}$$

where ν_0 and A_0 are constants (e.g., Weertman and Weertman, 1975; Carter, 1976; Poirier, 1985). The parameter A_0 is an activation temperature related to the activation energy E^* of the subsolidus creep deformation by $A_0 = E^*/R$, where R is the universal gas constant, as discussed in Chapter 5. The temperature dependence of mantle viscosity acts as a thermostat regulating the average mantle temperature (Tozer, 1967). Initially, when the Earth is hot, mantle viscosity is low, and extremely vigorous convection rapidly cools the Earth. Later in its evolution, when the Earth is relatively cool, its mantle viscosity is higher and more modest convection cools the planet at a reduced rate. Self-regulation tends to bring the viscosity of the mantle to a value that facilitates efficient removal by convection of the heat generated in the mantle. The temperature of the mantle adjusts to maintain or reach this preferred value

of viscosity. If the mantle is excessively hot to start with, e.g., because of accretional heating and the heat released by core formation, it will rapidly cool to bring its viscosity in line with the value preferred by its internal heat generation. The farther the mantle is from the preferred viscosity, the more rapid is the adjustment. Thus the specific value of the initial temperature $T(0)$ chosen for modeling the thermal history is unimportant. If it is too high, the adjustment by self-regulation rapidly rids the mantle of excess heat. Though not realistic, even an initially cold mantle would heat by radioactivity until the self-regulated viscosity was reached, a process that would have a billion year time scale. Self-regulation indicates that the present state of the convecting mantle has little or no memory of initial conditions, a circumstance which makes thermal evolution models appliable. As mantle radiogenic heat sources decay with time, convection transfers less heat, the preferred mantle viscosity gradually increases, and the mantle undergoes secular cooling. The gradual decrease of mantle temperature with time is a fundamental aspect of mantle evolution and requires that secular cooling contribute to the heat flow through the Earth's surface.

By combining (13.2.3)–(13.2.6) we obtain a single differential equation that contains explicitly the average mantle temperature T in all terms except the heat source:

$$\frac{\partial T}{\partial t} = f_1 e^{-\lambda t} - f_2 (T - T_s)^{1+\beta} \exp\left(\frac{-\beta A_0}{T}\right) \tag{13.2.7}$$

where

$$f_1 = H_0/c \tag{13.2.8}$$

and

$$f_2 = \frac{Ak}{Mcd}\left(\frac{\alpha g d^3}{\kappa \nu_0 Ra_{cr}}\right)^{\beta} \tag{13.2.9}$$

Equation (13.2.7) is solved subject to the initial condition $T = T(0)$ at $t = 0$. The solutions discussed in the next and subsequent sections will demonstrate the self-regulation imposed on mantle evolution by the temperature dependence of viscosity.

13.2.4 Model Results

The numerical integration of (13.2.7)–(13.2.9) subject to the initial condition $T = T(0)$ at $t = 0$ is straightforward. Some results from Schubert et al. (1980) are shown in Figures 13.1 and 13.2. Parameter values for this example are $T(0) = 3{,}273\,\mathrm{K}$, $T_s = 273\,\mathrm{K}$, $\beta = 0.3$, $\lambda = 1.42 \times 10^{-17}\,\mathrm{s}^{-1}$, $A_0 = 7 \times 10^4\,\mathrm{K}$, $f_1 = H_0/c = 4.317 \times 10^{-14}\,\mathrm{K\,s}^{-1}$, $k = 4.18\,\mathrm{W\,m}^{-1}\,\mathrm{K}^{-1}$, $\kappa = 10^{-6}\,\mathrm{m}^2\,\mathrm{s}^{-1}$, $\alpha = 3\times10^{-5}\,\mathrm{K}^{-1}$, $d = 2.8\times10^6\,\mathrm{m}$, $g = 10\,\mathrm{m\,s}^{-2}$, $\nu_0 = 1.65\times10^2\,\mathrm{m}^2\,\mathrm{s}^{-1}$, $Ra_{cr} = 1{,}100$, and $A/Mc = 1.377\times10^{-13}\,\mathrm{m}^2\,\mathrm{K\,J}^{-1}$. These values give $f_2 = 1.91 \times 10^{-14}$ (in SI units).

The kinematic viscosity of the mantle as a function of time is given in Figure 13.1; it increases monotonically from $3.2 \times 10^{11}\,\mathrm{m}^2\,\mathrm{s}^{-1}$ at the start of the model thermal history to $3.4 \times 10^{17}\,\mathrm{m}^2\,\mathrm{s}^{-1}$ after 4.5 Gyr. With $\rho = 3{,}400\,\mathrm{kg\,m}^{-3}$, the latter value of ν gives a viscosity $\mu = 1.2 \times 10^{20}\,\mathrm{Pa\,s}$ in good agreement with inferred values of mantle viscosity. The mantle temperature is also given in Figure 13.1. Temperature decreases monotonically with time but by less than 50% of its initial value because of the very strong temperature dependence of the viscosity. The temperature after 4.5 Gyr, $T = 1{,}950\,\mathrm{K}$, is representative

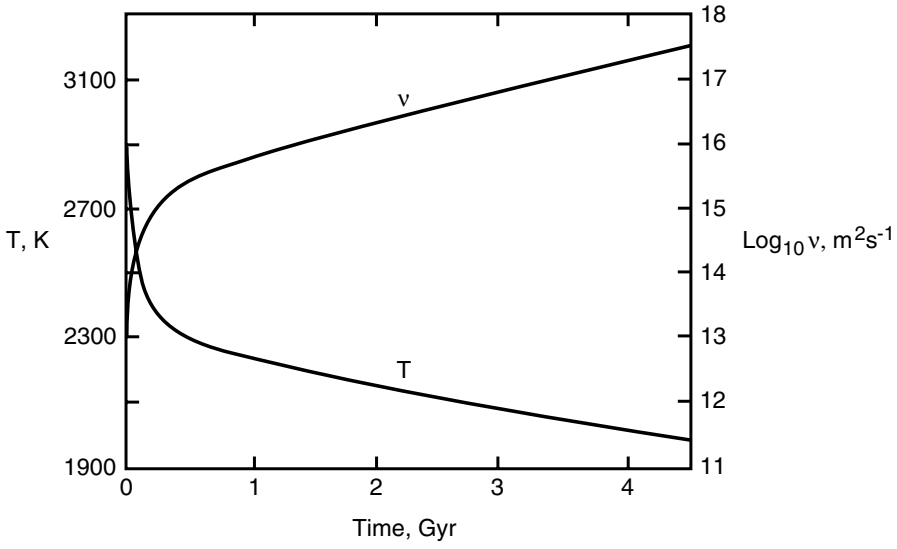


Figure 13.1. Mantle temperature T and kinematic viscosity ν as functions of time in a simple thermal history model of the Earth (after Schubert et al., 1980).

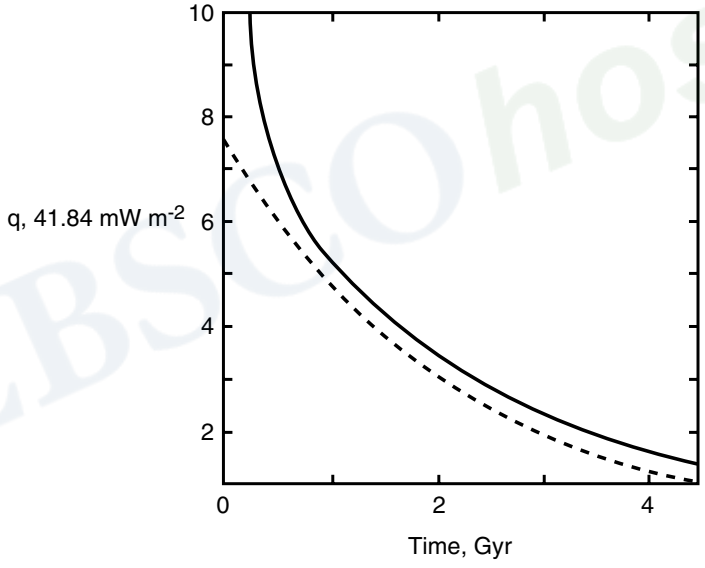


Figure 13.2. Mean surface heat flux q (solid curve) and total internal heat production per unit surface area (dashed curve) versus time for the thermal evolution calculation of Figure 13.1 (after Schubert et al., 1980).

of present temperatures in the mantle. The large drop in temperature and the enormous increase in viscosity during the first few hundred million years of model thermal evolution is a consequence of the self-regulation discussed above. The model mantle rapidly adjusts by early vigorous convection to a viscosity (temperature) that is higher (lower) than the viscosity (temperature) of its initial state. At the end of this early adjustment phase, the model mantle has gotten rid of most of its initial excess heat (Figure 13.2), and it has come into a state in which temperature and viscosity have adjusted to the convective removal of the remaining "primordial" heat and the energy produced by radioactivity. During the rest of geologic time the model mantle undergoes a more gradual secular cooling, with an attendant viscosity increase. The surface heat flow (solid line in Figure 13.2) declines throughout most

of the evolution, tracking the decay in the total radiogenic heat production per unit surface area (dashed line in Figure 13.2), but always remaining in excess of the internal heat release. The predicted surface heat flow after 4.5 Gyr is $q = 60\,\mathrm{mW\,m^{-2}}$, in reasonable agreement with the present average value ($q = 72\,\mathrm{mW\,m^{-2}}$) for the heat flow from the mantle (Earth's surface heat flow with the crustal component removed, see Sections 4.1.3 and 4.1.5 wherein the mean surface heat flow of $87\,\mathrm{mW\,m^{-2}}$ is reduced by 17%, the contribution of heat production in the continental crust to yield $72\,\mathrm{mW\,m^{-2}}$ for the mean mantle heat flux). The difference between the surface heat flow and the decay of radiogenic heat is due to the loss of primordial heat (or heat produced earlier by previous radioactive decay). The loss of primordial heat, or secular cooling, contributes 25% of the surface heat flow in the model of Figure 13.2. This difference will be discussed in some detail in the next section.

An analytic solution to (13.2.7)–(13.2.9) can be found for the early phase of rapid adjustment to the self-regulated state. Since the adjustment period lasts only a few hundred million years, λt is smaller than about 0.13 (for $t = 300\,\mathrm{Myr}$) and $\exp(-\lambda t) \approx 1$. At $t = 0$, the ratio of the second term to the first term in (13.2.7) is about 25 : 1. Thus we can neglect the first term on the right of (13.2.7) and solve

$$\frac{\partial T}{\partial t} \approx -f_2 (T - T_s)^{1+\beta} \exp\left(\frac{-\beta A_0}{T}\right) \tag{13.2.10}$$

subject to $T = T(0)$ at $t = 0$. The analytic solution to (13.2.10), valid approximately during the early rapid adjustment period, is given by the simple quadrature

$$-f_2 t = \int_{T(0)}^{T} \frac{ds \exp(\beta A_0 / s)}{(s - T_s)^{1+\beta}} \tag{13.2.11}$$

The solution given by (13.2.11) neglects radiogenic heating during the early period of rapid cooling. The dependence of temperature on time for the first 200 Myr of Earth's history is shown in Figure 13.3 for the same parameter values as used above in the example of

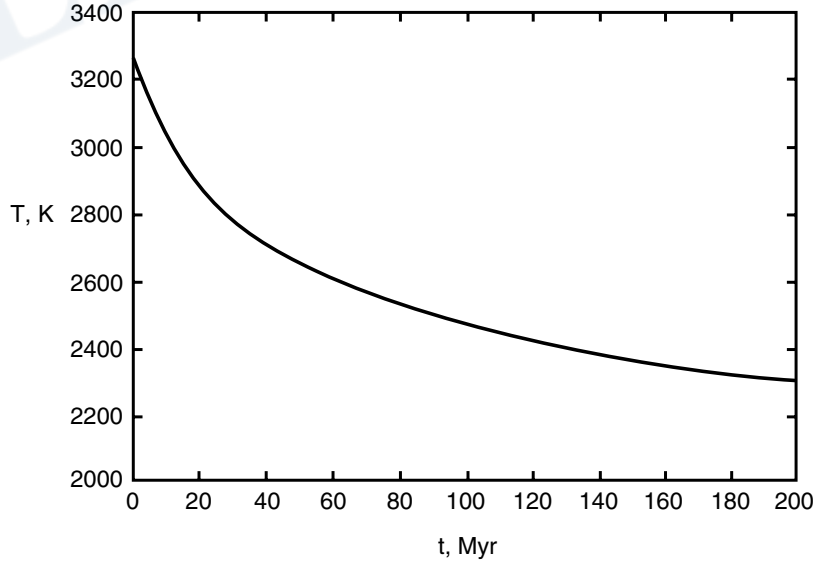


Figure 13.3. The decrease of mantle temperature with time during the early phase of vigorous convection and rapid cooling from (13.2.11). Adjustment to the self-regulated state and loss of most of the primordial heat occurs in only about 100 Myr.

Figures 13.1 and 13.2, namely $T_s = 273\,\mathrm{K}$, $\beta A_0 = 2.1 \times 10^4\,\mathrm{K}$, $T(0) = 3{,}273\,\mathrm{K}$, and $f_2 = 1.91 \times 10^{-14}$ (SI units). The temperature decreases by about 1,000 K in only 200 Myr. By about $t = 200\,\mathrm{Myr}$, the second term on the right of (13.2.7) is comparable to the first term (the radiogenic heating term), as can be seen by comparing Figures 13.1 and 13.3, and the approximation used in obtaining (13.2.11) is no longer valid.

13.2.5 Surface Heat Flow, Internal Heating, and Secular Cooling

Prior to the discovery of radioactivity, the heat flow through the Earth's surface was attributed to the cooling of the Earth's interior. Lord Kelvin used this hypothesis to estimate the age of the Earth as described in Section 4.1. After the discovery of radioactivity but before the widespread acceptance of mantle convection, it was realized that much of the geothermal heat loss had its origin in the decay of radioactive isotopes. Nevertheless, it was also believed that a substantial fraction (say, 25%) of the surface heat flow was due to secular cooling of the Earth (Holmes, 1916; Slichter, 1941). Unfortunately, this idea lost favor upon acceptance of convection as the mode of heat transfer in the deep mantle since it was thought that convection would be so efficient as to establish a balance between radiogenic heat production in the mantle and surface heat flow (Tozer, 1965a; Turcotte and Oxburgh, 1972b). After Lord Kelvin, who attributed 100% of the Earth's heat loss to secular cooling, opinions in the geophysical community underwent a complete reversal regarding the significance of whole-Earth cooling. A century later it was considered to contribute negligibly to the surface geothermal heat flow. The proposed equality of internal heat production and surface heat loss was used as a basis for estimating the abundances of uranium, thorium, and potassium in the Earth and Moon from measurements of surface heat flow (e.g., Langseth et al., 1976).

The idea that vigorous convection in the Earth's mantle established a balance between radiogenic heat production and surface heat flow was generally accepted throughout the 1970s. The use of parameterized convection to study the thermal evolution of the Earth was instrumental in re-establishing that secular cooling contributed importantly to Earth's surface heat flow even with efficient mantle convection. The parameterized convection models of Sharpe and Peltier (1978, 1979) showed that cooling of the Earth by mantle convection could account for Earth's surface heat flow even in the absence of any radiogenic heating in the mantle. Schubert (1979), Schubert et al. (1979a, b, 1980), Stevenson and Turner (1979), Turcotte et al. (1979), Davies (1980), Turcotte (1980b), and Peltier and Jarvis (1982) included mantle radiogenic heat production in their parameterized convection models which, as seen in Figure 13.2 and explained further below, yielded a contribution of secular cooling to surface heat flow as a natural consequence of the cooling Earth model.

The idea that secular cooling contributes significantly to the heat flow at the Earth's surface has already been seen in the model thermal history results of Figure 13.2 (after Schubert et al., 1980). The main reason is the secular decline in radioactive heat sources. There is no difficulty in having a close balance between internal heat production and surface heat flow in a convecting system with steady internal heat sources. Indeed, energy conservation requires this, if internal heating is the only source of energy for the system. However, when the internal heat sources decay with time, as is the case for radiogenic heat sources in the Earth's mantle, the surface heat loss and convection must also decline with time, and the system must cool. The secular decline in internal thermal energy must, by energy conservation, contribute to the flow of heat through the surface. No matter how efficiently convection transports heat through the mantle, the decay with time in the rate of internal heat production insures that secular cooling contributes to surface heat loss. The analyses

of Schubert et al. (1980), Davies (1980), and Stacey (1980) show that the magnitude of this contribution is substantial; about 25% of the Earth's surface heat flow is due to cooling of the Earth. This conclusion is a robust result, drawn from numerous calculations with wide variations in the values of parameters entering the thermal history models (Schubert et al., 1980).

The inequality between surface heat flow and interior heat production is expressed in terms of the Urey ratio,

$$Ur = \frac{MH}{Aq} \tag{13.2.12}$$

the ratio of the heat production term to the heat loss term on the right of (13.2.1). A Urey ratio less than unity implies a net loss of heat and a temperature decrease in the mantle, given by the following relation:

$$\frac{\partial T}{\partial t} = \frac{-Aq}{Mc}(1 - Ur) \tag{13.2.13}$$

The present value of the Urey ratio is 0.75, according to the results in Figure 13.2. This value, and estimates of the other quantities on the right of (13.2.13), give the mantle cooling rate. Substitution of $Ur = 0.75$, $q = 72\,\mathrm{mW\,m^{-2}}$ (mean mantle heat flux, see Section 13.2.4), and $A/Mc = 1.38 \times 10^{-13}\,\mathrm{m^2\,K\,J^{-1}}$ into (13.2.13) gives a present mantle cooling rate of about $80\,\mathrm{K\,Gyr^{-1}}$ $(-\partial T/\partial t)$. The sensitivity of the Earth's Urey ratio to different assumptions about the mix of radiogenic elements in the mantle has been explored in the parameterized thermal history calculations of Jackson and Pollack (1984).

It should be stressed that mantle cooling is inevitable because of convection. Even the assumption of equality between mantle heat loss and heat production leads to an estimate of the mantle cooling rate in accord with the above. With the assumption $MH = Aq$, (13.2.4) and (13.2.5) yield the following expression for mantle temperature in terms of mantle heat production:

$$T - T_s = \frac{\rho H d^2}{k}\left(\frac{M}{\rho A d}\right)^{1/1+\beta}\left(\frac{k\kappa\nu Ra_{cr}}{\alpha g \rho H d^5}\right)^{\beta/1+\beta} \tag{13.2.14}$$

For $\beta = 1/3$, (13.2.14) gives the dimensionless temperature $(T - T_s)/(\rho H d^2/k)$ of an internally heated convecting fluid directly proportional to $Ra_H^{-1/4}$, where Ra_H is the Rayleigh number for internal heating (7.4.6):

$$Ra_H = \frac{\alpha g \rho H d^5}{k\kappa\nu} \tag{13.2.15}$$

Equation (13.2.14) provides an alternative form of parameterization for thermal history models (Turcotte et al., 1979; Turcotte, 1980b; Cook and Turcotte, 1981).

An equation for $\partial T/\partial t$ can be obtained by differentiating (13.2.14) with respect to time, noting that both H and ν are functions of t. With the help of (13.2.2) and (13.2.6), we obtain

$$\frac{\partial T}{\partial t} = \frac{-\lambda(T - T_s)}{(1+\beta)}\left\{1 + \frac{\beta A_0}{(1+\beta)}\frac{(T - T_s)}{T^2}\right\}^{-1} \tag{13.2.16}$$

The second term in the parenthesis on the right of (13.2.16) is about 10 times the first term ($\beta A_0 = 2.1 \times 10^4\,\mathrm{K}$, $\beta = 0.3$, $T \approx 2{,}000\,\mathrm{K}$) and (13.2.16) can be reduced to the simple

equation

$$\frac{\partial T}{\partial t} = \frac{-\lambda T^2}{\beta A_0} \tag{13.2.17}$$

Equation (13.2.17) gives a cooling rate dependent only on mantle temperature, the radioactive decay constant, the activation temperature of the mantle viscosity, and the power-law exponent in the Nu–Ra relation, all reasonably well known parameters. With $\lambda = 1.42 \times 10^{-17}\,\mathrm{s}^{-1}$, $\beta = 0.3$, $T = 2{,}500\,\mathrm{K}$, and $A_0 = 7 \times 10^4\,\mathrm{K}$ we find $\partial T/\partial t \approx -135\,\mathrm{K\,Gyr}^{-1}$ from (13.2.17), in reasonable agreement with the above estimate of the mantle cooling rate. Our theoretical estimates of the Earth's secular cooling rate are in agreement with the $\approx 100\,\mathrm{K\,Gyr}^{-1}$ cooling rates derived by considering the liquidus temperatures and formation mechanisms of Archean komatiites in relation to similar properties of present basalts (Sleep, 1979).

13.2.6 Volatile Dependence of Mantle Viscosity and Self-regulation

Mantle viscosity is not only a strong function of temperature, but it also depends sensitively on the mantle volatile content as well (Jackson and Pollack, 1987; McGovern and Schubert, 1989; Hirth and Kohlstedt, 1996). Dissolved volatiles in the mantle tend to lower the creep activation energy and thus reduce the viscosity at a given temperature. A loss of volatiles from the mantle (degassing or outgassing) would stiffen the mantle, requiring an increase in mantle temperature to maintain the requisite vigor of convection. Conversely, volatile recharging of the mantle (regassing) by tectonic processes such as subduction, overthrusting, and delamination would soften the mantle, requiring a decrease in mantle temperature to maintain the requisite convective vigor. The dependence of mantle viscosity on both temperature and volatile content produces a strong coupling between the evolution of the mantle and the atmosphere–hydrosphere system (Schubert et al., 1989b).

The effects of a volatile-dependent mantle viscosity on thermal evolution can be quantified with a simple extension of our elementary model, along the lines suggested by McGovern and Schubert (1989). The available data on the reduction of the activation temperature for solid-state creep by dissolved volatiles such as water can be represented by

$$A_0 = \alpha_1 + \alpha_2 f \tag{13.2.18}$$

where f is the volatile mass fraction and α_1 and α_2 are empirical constants (α_2 is negative so that mantle viscosity decreases with increasing f). The variable mass fraction of volatiles f adds an additional dependent variable in the model. To represent the physical processes of degassing and regassing that determine the volatile content of the mantle, a parameterization is required.

The rate of mantle degassing $\partial M_v^d \partial t$ (M_v is the mass of volatiles in the mantle and d indicates degassing) can be expressed as

$$\frac{\partial M_v^d}{\partial t} = \rho_m f d_m S \tag{13.2.19}$$

where ρ_m is the mantle density, d_m is the average depth from which volatiles are released from the mantle (assuming complete outgassing to this depth), and S is the area spreading rate for the Earth's mid-ocean ridges. The parameter d_m can be thought of as an "equivalent depth," combining the actual depth of melting with an efficiency factor for the release of

volatiles. Regassing is assumed to take place through subduction. Similar to degassing, the rate of mantle regassing $\partial M_v^r/\partial t$ (r indicates regassing) can be expressed as

$$\frac{\partial M_v^r}{\partial t} = f_c \rho_c d_c \chi_r S \tag{13.2.20}$$

where f_c is the mass fraction of volatiles in the basaltic oceanic crust, ρ_c is the density of the crust, d_c is the average crustal thickness, and χ_r is an efficiency factor representing the fraction of volatiles that actually enters the deep mantle instead of returning to the surface through arc volcanism. The value of d_c can be chosen to reflect the added contribution of a subducted sediment layer.

Both degassing and regassing rates have been taken to be proportional to the vigor of mantle convection as expressed in the seafloor spreading rate S. This is related to the average age of subduction of oceanic crust τ by

$$S = \frac{A_{ob}(t)}{\tau} \tag{13.2.21}$$

where $A_{ob}(t)$ is the area of the ocean basins at time t. The heat flux through the ocean floor q (the heat flow from mantle convection) is related to τ by (8.6.8)

$$q = \frac{2k(T - T_s)}{(\pi \kappa \tau)^{1/2}} \tag{13.2.22}$$

Combination of (13.2.21) and (13.2.22) gives the seafloor spreading rate as

$$S = \frac{q^2 \pi \kappa A_{ob}(t)}{\{2k(T - T_s)\}^2} \tag{13.2.23}$$

Reymer and Schubert (1984) have proposed an expression for $A_{ob}(t)$ based on the assumption of constant continental freeboard (mean elevation of the continents above sea level) over the last 500 million years:

$$A_{ob}(t) = A_{ob}^* \left[\frac{V_{0a}^*}{V_0} + \frac{V_{0b}^* q^*}{V_0 q(t)} \right]^{-1} \tag{13.2.24}$$

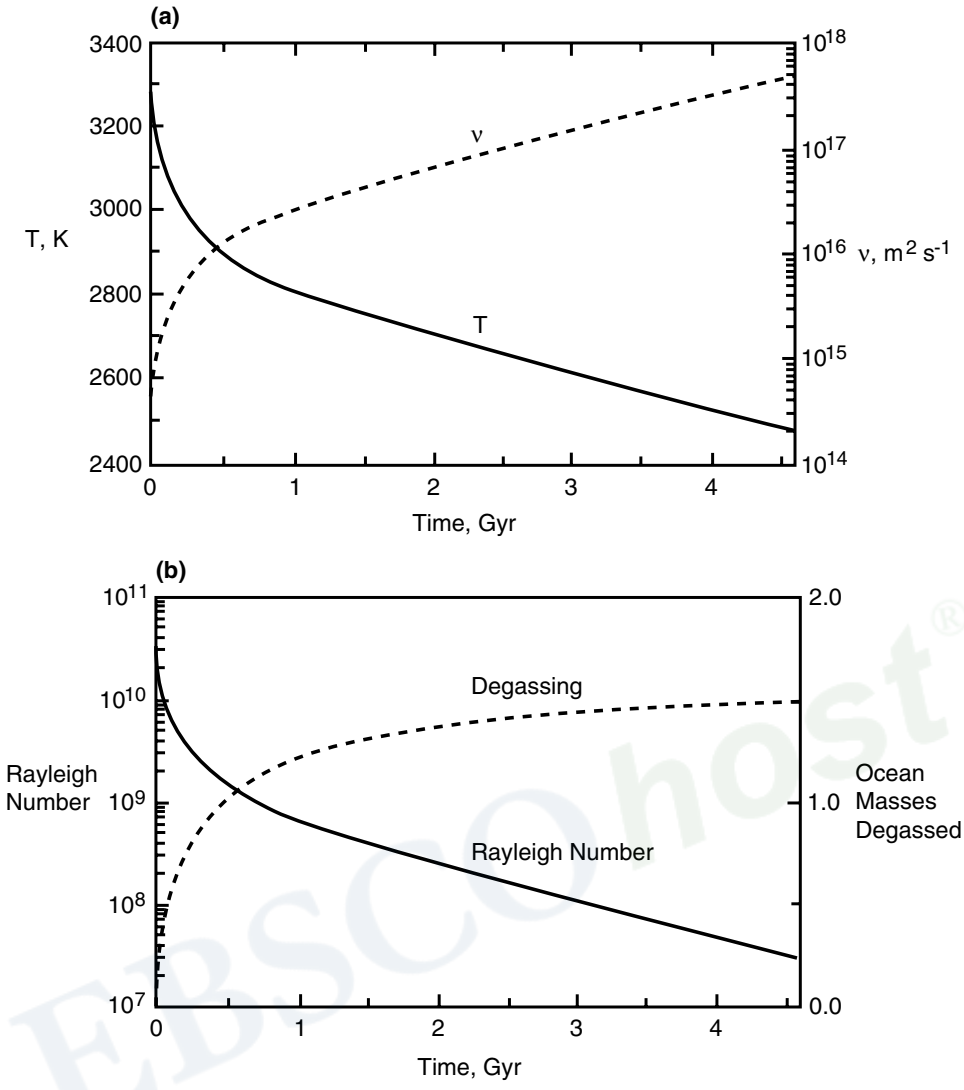
where V_0 is the total volume of water in the oceans (assumed constant in time), V_{0a} is the volume of the ocean basins above the peak ridge height, V_{0b} is the volume of the ocean basins below the peak ridge height, and asterisks denote present values. By combining (13.2.19)–(13.2.24) with the mass balance equation for the volatile content of the mantle,

$$\frac{\partial M_v}{\partial t} = \frac{\partial M_v^r}{\partial t} - \frac{\partial M_v^d}{\partial t} \tag{13.2.25}$$

we obtain

$$\frac{\partial M_v}{\partial t} = (f_c \rho_c d_c \chi_r - \rho_m f d_m) \frac{q^2 \pi \kappa A_{ob}^*}{4k^2 (T - T_s)^2} \left(\frac{V_{0a}^*}{V_0} + \frac{V_{0b}^*}{V_0 q} \right)^{-1} \tag{13.2.26}$$

The addition of (13.2.18) and (13.2.26) to (13.2.3)–(13.2.6) together with an initial condition for M_v extends the simple thermal history model to a mantle with degassing and regassing and a volatile-dependent viscosity.



The results of a typical thermal history calculation with degassing are summarized in Figure 13.4 (after McGovern and Schubert, 1989). The parameter values used in the calculation are listed in Table 13.1. The values of α_1 and α_2 which give the dependence of viscosity on volatile content are based on laboratory data of Chopra and Paterson (1984) for wet dunite. The dependence of mantle viscosity on volatile (water) content has been discussed more recently by Hirth and Kohlstedt (1996). The value of the depth of melting d_m is derived from an estimate of the depth of the basalt eutectic in the Archean (Sleep, 1979). This value is too large to reflect present conditions, but it is intended to model conditions prevalent in the early history of the Earth when the convective vigor was much greater. Since rates of volatile exchange in the model (and presumably in the Earth) are much greater in the early part of the calculation than they are toward the end of the calculation, the value

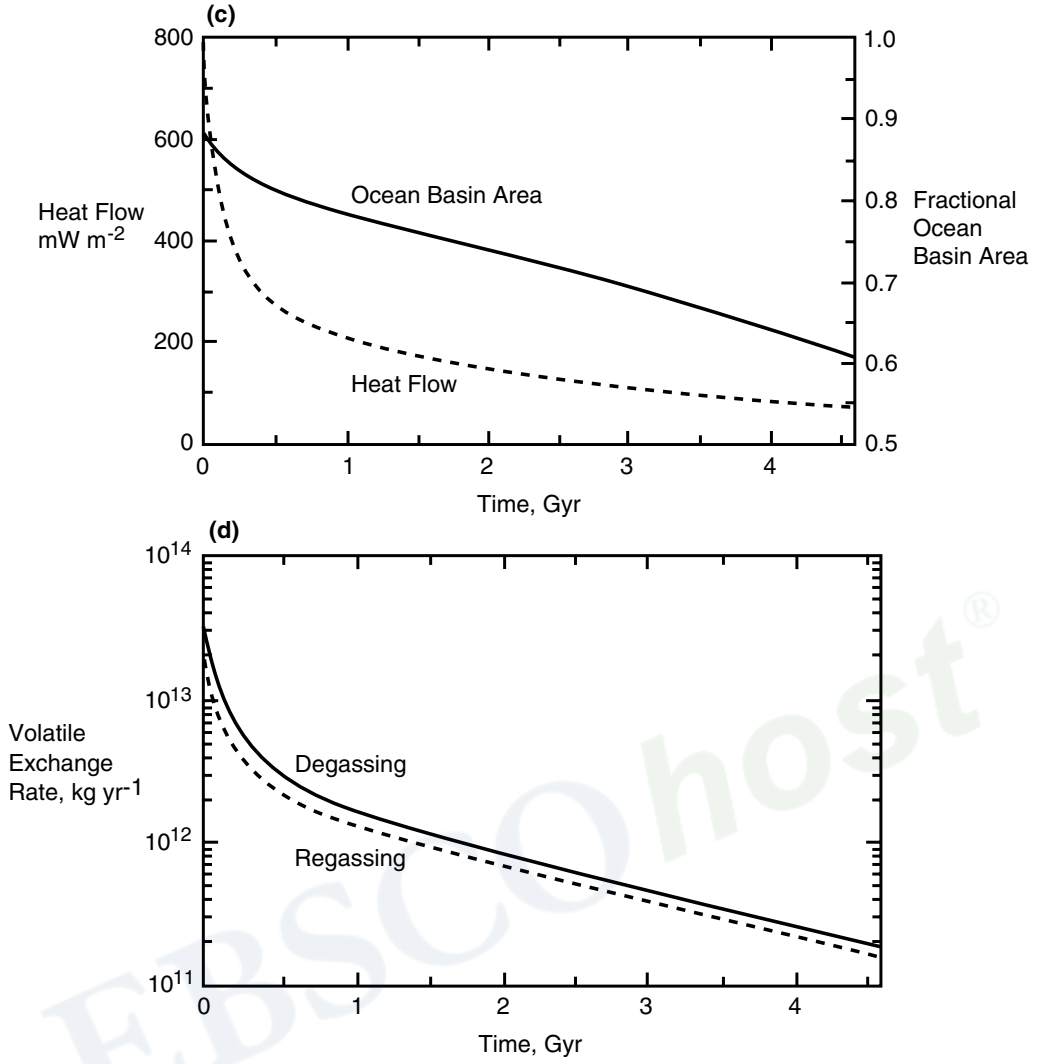


Figure 13.4. (a) Mantle temperature and kinematic viscosity as functions of time for a thermal history model with degassing and a volatile-dependent and temperature-dependent mantle viscosity. (b) Mantle Rayleigh number and amount of outgassing from the mantle (in units of ocean masses) versus time. (c) Heat flow from the mantle and normalized area of the ocean basins as functions of time (normalization is with respect to total surface area of the Earth). (d) Time dependence of mantle degassing and regassing rates (after McGovern and Schubert, 1989).

of d_m should represent conditions early in the Earth's evolution. Ringwood (1966, 1975) has estimated that the mass of dissolved water in the mantle is approximately 3 times that currently in the oceans. With the assumption that the total amount of water in the mantle–hydrosphere–atmosphere system is conserved, we set $n_m = 4$ (n_m is the number of ocean masses in the model mantle at time $t = 0$). The value of $H_0/c = f_1$ is iteratively adjusted so that the heat flow q at $t = 4.6\,\mathrm{Gyr}$ is equal to the present value of about $q^* = 70\,\mathrm{mW\,m^{-2}}$. For the calculation of Figure 13.4, H_0/c turns out to be $3.4 \times 10^{-14}\,\mathrm{K\,s^{-1}}$. It is assumed that the mass of volatiles on the surface is initially zero, i.e., $n_s = 0$ (n_s is the number of ocean masses initially in surface volatile reservoirs).

The degassing history of the mantle (Figure 13.4b) is characterized by an early period of rapid outgassing (more than one ocean mass in the first 500 Myr), followed by a gradual

Table 13.1. ***Parameter Values for a Thermal History Model with Degassing and a Volatile-dependent Mantle Viscosity***

Parameter	Value	Reference
ν_0	$2.21 \times 10^7\ \mathrm{m^2\ s^{-1}}$	Jackson and Pollack (1987)
k	$4.2\ \mathrm{W\,m^{-1}\,K^{-1}}$	Schubert et al. (1980)
		Jackson and Pollack (1987)
g	$9.8\ \mathrm{m\,s^{-2}}$	
λ	$3.4 \times 10^{-10}\ \mathrm{yr^{-1}}$	Jackson and Pollack (1984)
α	$3 \times 10^5\ \mathrm{K^{-1}}$	Schubert et al. (1980)
		Jackson and Pollack (1987)
κ	$10^{-6}\ \mathrm{m^2\ s^{-1}}$	Jackson and Pollack (1987)
R_m	6,271 km	Jackson and Pollack (1987)
R_c	3,471 km	Jackson and Pollack (1987)
T_s	273 K	Jackson and Pollack (1987)
Ra_{cr}	1,100	Jackson and Pollack (1987)
ρc	$4.2\ \mathrm{MJ\,m^3\,K^{-1}}$	Jackson and Pollack (1987)
β	0.3	Jackson and Pollack (1987)
α_1	6.4×10^4 K	McGovern and Schubert (1989)
α_2	-6.1×10^6 K	McGovern and Schubert (1989)
	$(\text{weight fraction})^{-1}$	McGovern and Schubert (1989)
M_{mantle}	4.06×10^{24} kg	Schubert et al. (1980)
d_m	100 km	Sleep (1979)
f_c	0.03	Schubert et al. (1989b)
d_c	5 km	Schubert et al. (1989b)
ρ_c	$2{,}950\ \mathrm{kg\,m^{-3}}$	Turcotte and Schubert (1982)
χ_r	0.8	
M_{ocean}	1.39×10^{21} kg	Walker (1977)
n_m	4.0	Ringwood (1966, 1975)
n_s	0	
A_0^*	$3.1 \times 10^{14}\ \mathrm{m^2}$	Reymer and Schubert (1984)
V_{0a}^*	$7.75 \times 10^{17}\ \mathrm{m^3}$	Reymer and Schubert (1984)
V_{0b}^*	$3.937 \times 10^{17}\ \mathrm{m^3}$	Reymer and Schubert (1984)
V_0	$1.1687 \times 10^{18}\ \mathrm{m^3}$	Reymer and Schubert (1984)
q^*	$70\ \mathrm{mW\,m^{-2}}$	Turcotte and Schubert (1982)

leveling off in the outgassed mass for the remaining 4 Gyr. The change in activation temperature A_0 exhibits similar behavior due to its dependence on mantle volatile content. A comparison of Figures 13.4a and b shows that the time scales for degassing and for rapid initial cooling are approximately the same.

Figure 13.4c shows the area of the Earth's ocean basins A_{ob} (normalized to the total surface area) as a function of time. Although the assumption of constant freeboard is only known to be valid for the last 500 million years (Wise, 1974; Reymer and Schubert, 1984), application of this assumption over the entire thermal history calculation results in a monotonic decrease in ocean basin area (increase in continental area) over geologic time, in qualitative agreement with many crustal growth models (Reymer and Schubert, 1984, Figure 6) (see also Section 13.7). By (13.2.24) and our requirement that the present heat flow q^* matches the measured value, the present value of A_{ob} necessarily agrees with today's area of the ocean basins.

Figure 13.4d shows the mantle degassing and regassing rates as functions of time. The degassing curve of Figure 13.4b is just the integral of the area between these two curves. As implied by Figure 13.4b, these rates start out significantly different, but converge with time.

The Rayleigh number (Figure 13.4b) is very large initially (about 3.5×10^{10}) reflecting the low value of initial kinematic viscosity (about 4×10^{14} m^2 s^{-1}). Mantle convection during the period of early rapid heat loss is indeed vigorous. The Rayleigh number falls by about one and a half orders of magnitude during the first 500 Myr, while viscosity increases by about the same amount. The mantle adjusts to its self-regulated state in about 500 Myr, after which the Rayleigh number decreases with time (while viscosity increases with time) by about another one and a half orders of magnitude over the next 4 Gyr reaching a present value of about 3×10^7 (with a present kinematic viscosity of about 5×10^{17} m^2 s^{-1}). The Rayleigh number decreases approximately exponentially with time and the viscosity increases approximately exponentially with time after the initial period of rapid cooling.

In order to identify the major effects of mantle degassing or regassing on the thermal evolution of the Earth, we compare in Figure 13.5 the results of the model calculation of

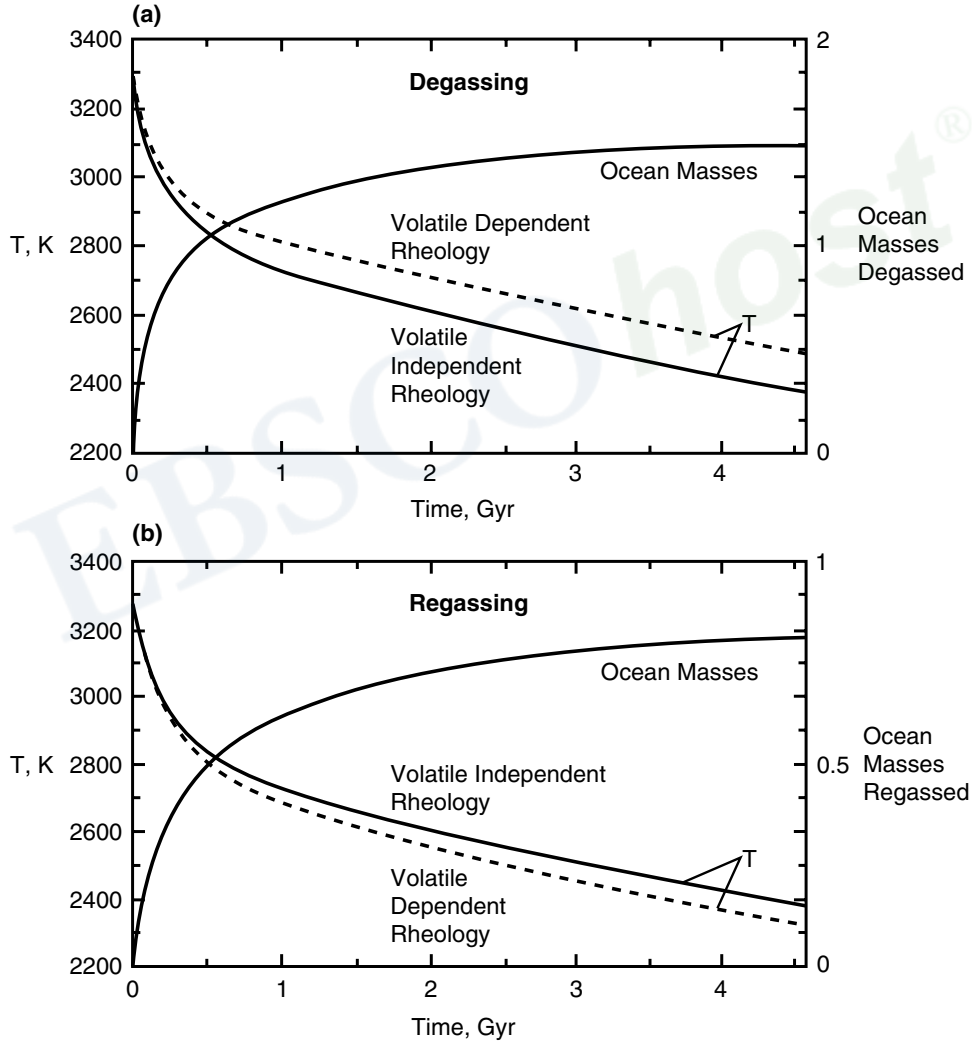


Figure 13.5. (a) The effect of degassing on mantle temperature. Thermal histories for volatile-dependent (dotted line) and volatile-independent (solid line) rheologies are plotted along with the degassing history for the volatile-dependent case. Degassing with volatile dependence raises the present temperature of the mantle. (b) The effect of regassing on mantle temperature. Thermal histories for volatile-dependent (dotted line) and volatile-independent (solid line) rheologies are plotted along with the regassing history for the volatile-dependent case. Regassing with volatile dependence lowers the present temperature of the mantle.

Figure 13.4 with one in which there is no volatile dependence of the rheology (parameters are identical to those in Figure 13.4 except $\alpha_1 = 5.6 \times 10^4$ K and $\alpha_2 = 0$). In addition, Figure 13.5 compares the results of a regassing scenario with volatile-dependent rheology (parameters are identical to those in Figure 13.4 except $d_c = 6$ km, $d_m = 50$ km, $n_m = 4$, $n_s = 2$) to the thermal history with no volatile dependence of viscosity. (In all the cases in Figure 13.5, H_0/c has essentially the same value as it does in the calculation of Figure 13.4.) Outgassing (Figure 13.5a) dries out the interior and tends to increase its viscosity. However, the tendency for devolatilization to increase viscosity is compensated by the effect of temperature on viscosity. Higher viscosity tends to reduce heat flow, allowing heat generated by radiogenic sources to build up and increase mantle temperature. But then the higher temperature tends to reduce viscosity and enhance heat flow. Thus, the mantle adjusts to maintain the required rate of heat loss by increasing temperature, reducing viscosity, and maintaining the level of convective vigor. The net result of degassing is a hotter mantle, but mantle heat flow, viscosity, and convective vigor are essentially the same as in a mantle with volatile-independent rheology. With degassing and a volatile-dependent rheology the mantle is hotter and cools more slowly than it would with a volatile-independent rheology. Regassing (Figure 13.5b) increases the volatile content of the interior and tends to decrease its viscosity. However, as in the degassing case, the tendency for revolatilization to decrease viscosity is compensated by a reduction in mantle temperature so as to maintain viscosity, heat flow, and convective vigor approximately constant. In addition, Figure 13.5b shows the amount of water regassed into the mantle. The evolution, in terms of the amount of cooling and the quantity of water reabsorbed into the mantle, is rapid during the first several hundred million years, becoming more gradual afterwards. In the regassing case about three-quarters of an ocean mass of volatiles (water) is reinjected into the mantle over geologic time, with the bulk of this occurring in the first billion years. The main effect of the volatile-dependent mantle viscosity is a cooler mantle, compared to the case where viscosity depends on temperature only. As in the degassing case, mantle viscosity and heat flow are essentially the same for both the volatile-dependent and volatile-independent viscosities. In both the regassing and degassing scenarios, the time rate of change of temperature eventually tends to the same value for the volatile-dependent and volatile-independent rheologies; during the latter stages of thermal evolution only a constant temperature offset distinguishes the volatile-dependent mantle cooling rate from the volatile-independent one.

Figure 13.6 shows the evolution of the Urey ratio for the degassing case of Figure 13.4. The Urey ratio Ur starts with a relatively low value at $t = 0$ because of the dominance of primordial heat in the initially hot mantle. The ratio quickly reaches a maximum and then slowly and steadily decreases as the mantle volatile exchange rates equilibrate and the value of A_0 approaches its self-regulated value. Because of the decay of the radiogenic heat sources, Ur will tend to zero as $t \to \infty$; Ur is less than unity throughout the entire time.

13.3 More Elaborate Thermal Evolution Models

13.3.1 A Model of Coupled Core–Mantle Thermal Evolution

While the simple models of the previous section are adequate for demonstrating many of the important aspects of the Earth's evolution, more elaborate models are needed if, for example, the evolution of the core is to be included. We now consider a more complex model developed by Stevenson et al. (1983) with a coupled core and mantle. Thermal evolution

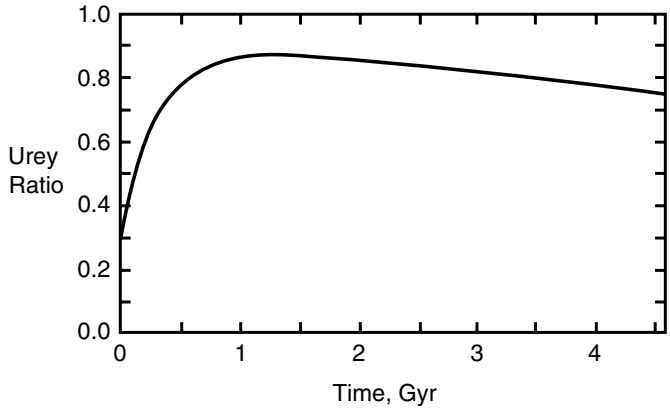


Figure 13.6. Urey ratio versus time in the thermal history calculation of Figure 13.4.

calculations including core–mantle coupling have also been given by Sleep et al. (1988) and by Davies (1993). A core thermal history model has been presented by Buffett et al. (1996). As sketched in Figure 13.7, the model of Stevenson et al. (1983) consists of a spherical shell mantle surrounding a concentric spherical core. The core has radius R_c and density ρ_c and the mantle has outer radius R_p and density ρ_m. There is a solid inner core of radius R_i and a liquid outer core.

Figure 13.7b is a schematic of the radial profile of spherically averaged temperature $T(r)$ for the coupled model. There are thermal boundary layers at the top and bottom of the convecting mantle of thickness δ_s and δ_c, respectively. Temperature is assumed to vary linearly with radius in the boundary layers. The change in temperature across the top cold thermal boundary layer is ΔT_s and the temperature change across the lower hot thermal boundary layer is ΔT_c. Mantle temperature is T_u at the base of the upper thermal boundary layer and T_l at the top of the lower thermal boundary layer. Temperature at the core–mantle boundary is T_{cm}. The surface temperature is T_s. Temperature is assumed to increase adiabatically with depth in the mantle between the values of T_u and T_l in the region outside the boundary layers. The temperature in the fluid outer core is taken to increase adiabatically with depth from T_{cm} to T_{mio}, the liquidus temperature of the core alloy. Nonadiabatic temperature differences or boundary layers are negligible in the convective outer core because of its low viscosity. The dashed curve in Figure 13.7b is the depth or pressure $p(r)$ dependent liquidus temperature of the core alloy $T_m(r)$ (r is the radial distance from the center of the model). In the liquid outer core $T(r) > T_m(r)$ while the reverse is true in the solid inner core.

The solid inner core is assumed to be pure iron while the outer core contains a light alloying element that we take to be sulfur (the model can be trivially modified to deal with other possible light constituents in the core, for example, oxygen). We neglect inner–outer core density differences in computing $p(r)$, but we do take into account the gravitational energy release upon freezing of outer core liquid and growth of the solid inner core, a process which excludes the light alloying element from the solid inner core and concentrates it in the liquid outer core. The gravitational potential energy release upon differentiation of the core is responsible for the convective motions in the outer core that generate the Earth's magnetic field by dynamo action (Braginsky, 1963; Gubbins, 1977a; Loper, 1978a, b; Loper and Roberts, 1979; Stevenson et al., 1983; Glatzmaier and Roberts, 1997). The latent heat release that also occurs with inner core growth provides a thermal drive for convective

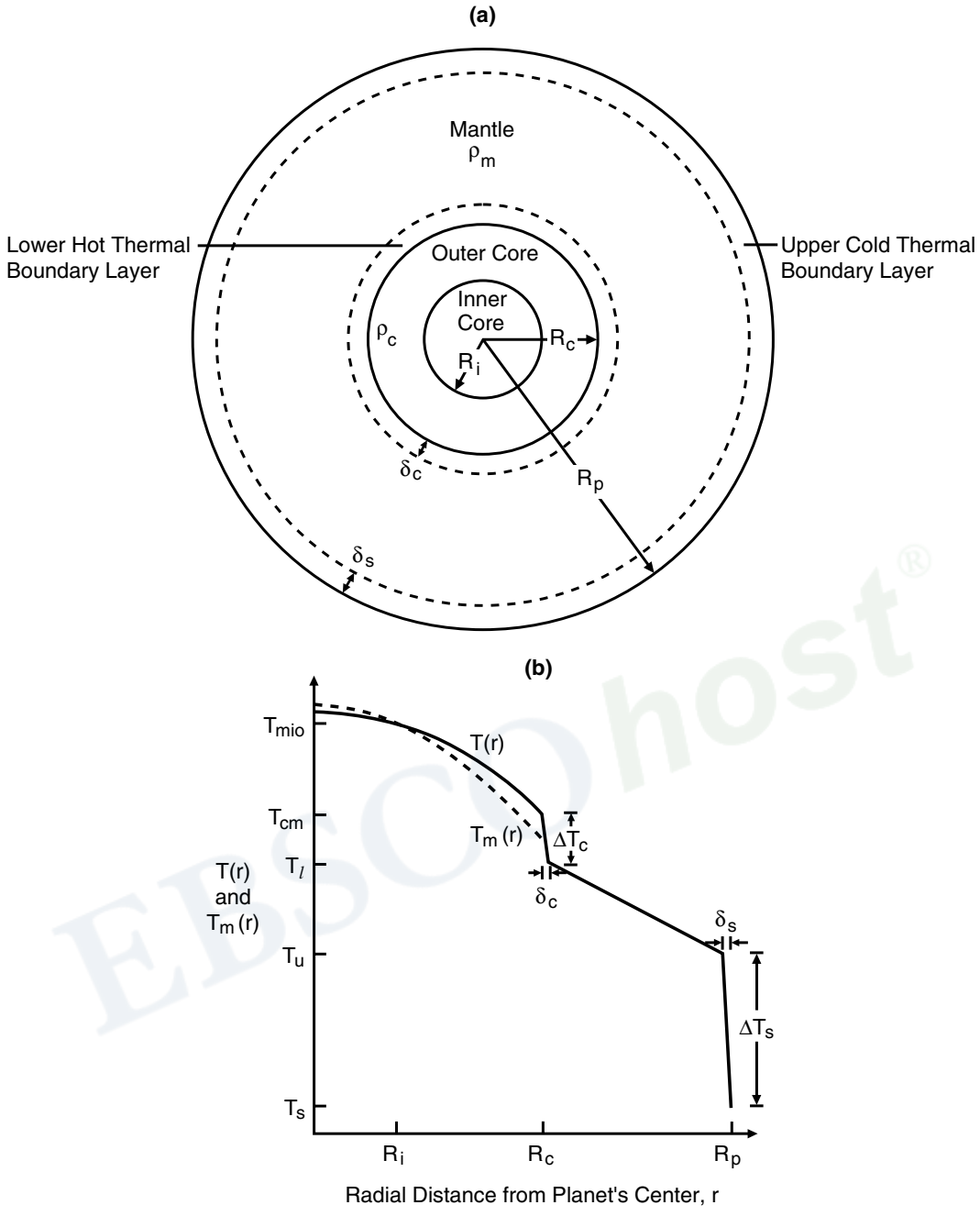


Figure 13.7. (a) Geometry of Earth thermal history model for coupled core–mantle evolution. (b) Schematic of the radial profile of spherically averaged temperature in the coupled core–mantle Earth thermal evolution model.

motions in the liquid outer core (Verhoogen, 1961). Because of the Carnot efficiency factor, a thermally driven dynamo is less thermodynamically efficient than one driven by chemical buoyancy.

The liquidus temperature $T_m(r)$ of the core alloy is expressed as a quadratic in the pressure $p(r)$ (Stevenson et al., 1983):

$$T_m(r) = T_{m0}(1 - 2\chi)(1 + T_{m1}p(r) + T_{m2}p^2(r)) \tag{13.3.1}$$

where T_{m0}, T_{m1}, and T_{m2} are constants, χ is the mass fraction of light alloying constituent in the liquid outer core and it is assumed that $\chi \ll 1$. The temperature along the outer core adiabat $T_c(r)$ is similarly represented by

$$T_c(r) = T_{cm}\left\{\frac{1 + T_{a1}\,p(r) + T_{a2}\,p^2(r)}{1 + T_{a1}\,p_{cm} + T_{a2}\,p_{cm}^2}\right\} \tag{13.3.2}$$

where T_{a1} and T_{a2} are constants and p_{cm} is the pressure at the core–mantle boundary.

The simultaneous solution of (13.3.1) and (13.3.2) gives the pressure p_{io} at the inner core–outer core boundary. The radius of the inner core R_i is then obtained by assuming that the acceleration of gravity in the core is rg/R_c (g is the surface value of gravity):

$$R_i = \left\{\frac{2(p_c - p_{io})R_c}{\rho_c g}\right\}^{1/2} \tag{13.3.3}$$

where p_c is the pressure at the center of the Earth. The mass of the inner core m_{ic} is

$$m_{ic} = \tfrac{4}{3}\pi R_i^3 \rho_c \tag{13.3.4}$$

Initially the core is superliquidus and $R_i = 0$. As the Earth cools, inner core nucleation begins when the liquidus temperature is reached at the center of the Earth. The inner core grows upon further cooling of the Earth. The liquidus temperature of the outer core decreases as the inner core grows and the light alloying constituent is concentrated in the outer core. The decrease in outer core liquidus upon inner core freezing is important in retarding the rate of inner core growth and in preventing complete freezing of the core (not applicable to the Earth but perhaps significant in other planets such as Mercury or Jupiter's moon Ganymede). Conservation of the light constituent mass gives

$$\chi = \frac{\chi_0 R_c^3}{R_c^3 - R_i^3} \tag{13.3.5}$$

where χ_0 is the initial concentration of the light element in the core.

Separate energy balance equations are required for the mantle and core. These are given by

$$\frac{4}{3}\pi\left(R_p^3 - R_c^3\right)\left\{\rho_m H - \rho_m c_m \frac{\partial}{\partial t}\langle T_{\text{mantle}}\rangle\right\} = 4\pi\left\{R_p^2 q_s - R_c^2 q_c\right\} \tag{13.3.6}$$

$$\frac{4}{3}\pi R_c^3\left\{-\rho_c c_c \frac{\partial}{\partial t}\langle T_c\rangle\right\} + (L + E_G)\frac{\partial m_{ic}}{\partial t} = 4\pi R_c^2 q_c \tag{13.3.7}$$

where c_m and c_c are the specific heats of the mantle and core, H is the rate of internal heating per unit mass in the mantle as given by (13.2.2) (it is assumed that there are no radiogenic heat sources in the core), $\langle T_{\text{mantle}}\rangle$ and $\langle T_c\rangle$ are the volume-averaged mantle and core temperatures, and q_s and q_c are the heat fluxes through the surface and core–mantle boundary, respectively. The temperature $\langle T_{\text{mantle}}\rangle$ can be related to T_u by

$$\langle T_{\text{mantle}}\rangle = \eta_m T_u \tag{13.3.8}$$

where η_m is a constant, while $\langle T_c\rangle$ can be similarly related to T_{cm} by

$$\langle T_c\rangle = \eta_c T_{cm} \tag{13.3.9}$$

where η_c is a constant. Use of (13.3.8) and (13.3.9) provides a convenient representation of the quantities $\partial\langle T_{\text{mantle}}\rangle/\partial t$ and $\partial\langle T_c\rangle/\partial t$ in the energy equations in terms of the time rates of change of the single temperatures T_u and T_{cm}. The quantity $\partial m_{ic}/\partial t$ that appears in the core energy balance equation can be related to $\partial T_{cm}/\partial t$ through the use of (13.3.1)–(13.3.4).

The heat fluxes q_s and q_c are given by parameterizations similar to (13.2.4) and (13.2.5). We note that $d(Ra_{cr}/Ra)^{\beta}$ in (13.2.4) is the thermal boundary layer thickness δ and that (13.2.4) is just Fourier's law of heat conduction for the boundary layer $q = k(T - T_s)/\delta$. Accordingly, we can write the expressions for q_s and q_c as (Figure 13.7)

$$q_s = \frac{k\Delta T_s}{\delta_s} = \frac{k(T_u - T_s)}{\delta_s} \tag{13.3.10}$$

$$q_c = \frac{k\Delta T_c}{\delta_c} = \frac{k(T_{cm} - T_l)}{\delta_c} \tag{13.3.11}$$

The thickness of the surface boundary layer δ_s is expressed, using the global Rayleigh number

$$Ra = \frac{g\alpha(\Delta T_s + \Delta T_c)(R_p - R_c)^3}{\nu\kappa} \tag{13.3.12}$$

as

$$\delta_s = (R_p - R_c)\left(\frac{Ra_{cr}}{Ra}\right)^{\beta} \tag{13.3.13}$$

If the mantle were a constant viscosity fluid layer, then the lower thermal boundary layer would have the same thickness as the upper boundary layer $\delta_c = \delta_s$ and δ_c would also be given by (13.3.13). In this case the heat fluxes q_s and q_c would be different only because of differences in the temperature drops ΔT_s and ΔT_c across the boundary layers. However, due to the strongly temperature dependent viscosity of the mantle, it is possible that the lower boundary layer is thinner, on the average, than the upper boundary layer (Daly, 1980; Nataf and Richter, 1982). The lower boundary layer might also be thinned by the ejection of plumes and thermals as a consequence of buoyancy instability enhanced by a reduction in viscosity (Howard, 1966; Richter, 1978; Yuen and Peltier, 1980; Olson et al., 1988). We can account for a reduction in boundary layer thickness at the core–mantle boundary by determining its thickness locally whenever the heat flux from the core is sufficiently large. The experiments of Booker and Stengel (1978) suggest that the local critical Rayleigh number for the breakdown of the boundary layer is

$$Ra_{crb} = \frac{g\alpha\Delta T_c\delta_c^3}{\nu_c\kappa} \approx 2 \times 10^3 \tag{13.3.14}$$

Richter (1978) finds that ν_c should be based on the average temperature within the boundary layer. Hence,

$$\nu_c \equiv \nu_r \exp\left(\frac{A_r}{T_l + \Delta T_c/2}\right) \tag{13.3.15}$$

We use (13.3.14) instead of (13.3.13) to calculate δ_c whenever (13.3.14) gives a smaller thickness. The viscosity ν used in (13.3.13) to get δ_s is given by (13.2.6) with the temperature evaluated at the upper mantle temperature T_u.

The thermal history of this coupled core–mantle model is obtained by integrating with respect to time the energy balance equations and the equation for the rate of inner core growth. The main dependent variables of the model are the upper mantle temperature $T_u(t)$, the core–mantle boundary temperature $T_{cm}(t)$, and the radius of the inner core $R_i(t)$. The boundary and initial conditions for the model are $T(R_p) = T_s$, $T_u(t=0) = T_{u0}$, $T_{cm}(t=0) = T_{cm0}$, $R_i(t=0) = 0$.

13.3.2 Core Evolution and Magnetic Field Generation

The cooling history of the mantle in this coupled core–mantle thermal evolution model is similar to that of the simpler model presented in Section 13.2. Here we focus on the new aspects of the coupled model and describe the thermal evolution of the core and its implications for magnetic field generation by dynamo action. We discuss two models presented by Stevenson et al. (1983), the parameter values for which are given in Table 13.2. The rheological parameters were chosen to give a present mantle kinematic viscosity of about $10^{17}\,\mathrm{m^2\,s^{-1}}$ (Cathles, 1975; Peltier, 1981), and the value of $\rho_m H_0$, together with the chondritic value for λ, gives a present heat flux from the model mantle of about $60\,\mathrm{mW\,m^{-2}}$. Internal heating in this model contributes about 75% of the present surface heat loss, in agreement with the discussion in Section 13.2 of the contribution of secular cooling to the Earth's surface heat flow.

Table 13.2. *Parameter Values for Two Coupled Core–Mantle Thermal Evolution Models of the Earth*

Parameter	Value
Parameters Common to Both Models	
α	$2 \times 10^{-5}\,\mathrm{K^{-1}}$
k	$4\,\mathrm{W\,m^{-1}\,K^{-1}}$
κ	$10^{-6}\,\mathrm{m^2\,s^{-1}}$
$\rho_m c_m = \rho_c c_c$	$4\,\mathrm{MJ\,m^{-3}\,K^{-1}}$
$\rho_m H_0$	$0.17\,\mu\mathrm{W\,m^{-3}}$
λ	$1.38 \times 10^{-17}\,\mathrm{s^{-1}}$
A_0	$5.2 \times 10^4\,\mathrm{K}$
ν_0	$4 \times 10^3\,\mathrm{m^2\,s^{-1}}$
Ra_{cr}	500
β	0.3
R_p	6,371 km
g	$10\,\mathrm{m\,s^{-2}}$
T_s	293 K
T_{m1}	$6.14\,\mathrm{K\,TPa^{-1}}$
T_{m2}	$-4.5\,\mathrm{K\,TPa^{-2}}$
T_{a1}	$3.96\,\mathrm{K\,TPa^{-1}}$
T_{a2}	$-3.3\,\mathrm{K\,TPa^{-2}}$
η_m, η_c	1.3, 1.2
Parameter	**Value, E1, E2**
Parameters Different for the Two Models	
$L + E_G$	$1,2\,\mathrm{MJ\,kg^{-1}}$
T_{m0}	1,950, 1,980 K

The main difference between the models in Table 13.2 is the value for $L + E_G$, the total energy (latent heat plus gravitational energy) released per unit mass on inner core solidification. Model E1 uses $L + E_G = 1\,\mathrm{MJ\,kg^{-1}}$, while model E2 assumes $L + E_G = 2\,\mathrm{MJ\,kg^{-1}}$. The quantity $L + E_G$ is uncertain because of our lack of knowledge of the exact composition of the core and of its thermodynamic properties. In addition, the gravitational energy release E_G depends on core size. Models E1 and E2 also differ in their values of T_{m0} which are chosen to reproduce the correct inner core radius at present. Since the models are constrained to give the present value of R_i they are not sensitive to uncertainties in our precise knowledge of the core melting curve. Melting temperatures in the core are uncertain despite recent experiments to determine the melting point of iron at high pressures (Boehler, 1993, 1994, 1996; Anderson and Ahrens, 1996; Chen and Ahrens, 1996; Anderson and Duba, 1997; Boehler and Ross, 1997) in part because of our lack of knowledge of core composition.

The initial concentration of light constituent in the core χ_0 is taken to be 0.1, consistent with sulfur being the light element (Ahrens, 1979). However, the exact identification of the light constituent in the core is not essential in the model and other possibilities such as silicon and oxygen (Ringwood, 1977b; Poirier, 1994a, b) can be accommodated by adjusting the numerical coefficient of χ in (13.3.1). The parameters T_{a1} and T_{a2} for the core adiabat are based on Stacey's (1977b) value for the Grüneisen parameter γ. The choice of core and mantle adiabats following Stacey (1977b) determines the values of η_m and η_c. Core–mantle boundary layer thickness δ_c is calculated using (13.3.14).

Question 13.3: *What is the major light alloying element in the Earth's core?*

The core evolution according to models E1 and E2 is shown in Figure 13.8 in terms of the time dependence of the heat flux from the core. Heat flow from the core initially decreases very rapidly with time during the period when early vigorous mantle convection removes heat quickly from the core. Inner core solidification begins at $t \approx 2.7$ Gyr and 2.3 Gyr in models E1 and E2, respectively, when the core has cooled sufficiently that the core adiabat drops to the core melting temperature at the center of the Earth. Core freezing occurs later in model E1 since it has a lower core melting temperature than model E2 (see the values of T_{m0} in Table 13.2). The present inner core radii in models E1 and E2 are 1,234 km and 1,207 km, respectively; the present inner core radius in model E1 is larger than in model E2 despite the later onset of inner core freezeout in model E1 because twice the mass can be solidified in model E1 for every unit of $L + E_G$ removed from the core by mantle convection.

Were it not for inner core freezing, the monotonic decrease in core heat flux q_c would continue through geologic time (dashed curve in Figure 13.8) and eventually q_c would fall below the value necessary to supply the conductive heat flow along the core adiabat (estimated at $15\,\mathrm{mW\,m^{-2}}$, horizontal dash-dot line in Figure 13.8). Thermal convection in the core is not possible if q_c falls below the heat flux conducted along the core adiabat; a thermally driven dynamo would also not be possible were q_c to drop below the conductive heat flux along the adiabat. Figure 13.8 shows that thermal convection and thermal forcing of a core dynamo would have ceased at about 3.2 Gyr in the Earth models E1 and E2 if not for inner core solidification. However, the cores in the models do begin to solidify at about 2.5 Gyr and the decrease in core heat flow with time is arrested by this event. Once core freezing begins, the release of latent heat and gravitational energy contributes to the heat flow from the core

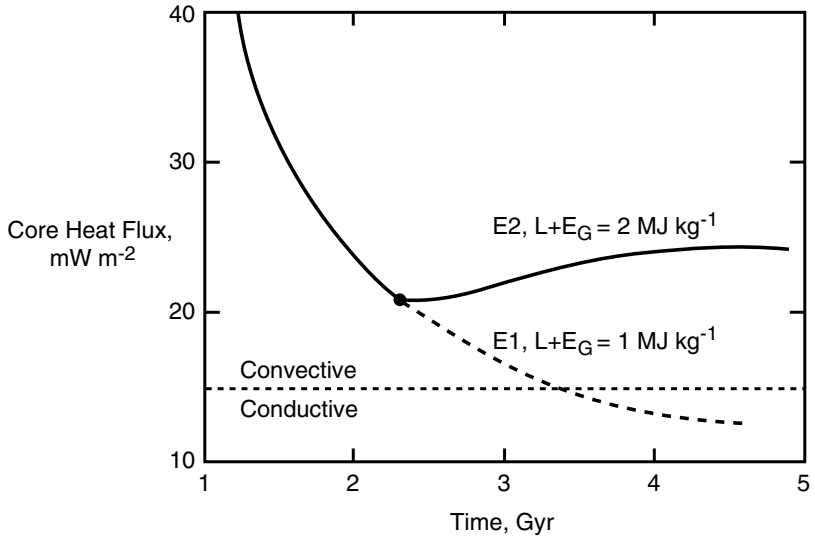


Figure 13.8. The heat flux from the core q_c versus time in a coupled core–mantle thermal history model (after Stevenson et al., 1983) for two values of $L+E_G$, the total energy liberated per unit mass upon core solidification. The horizontal dotted line gives the value of conductive heat flow along the core adiabat. For q_c above this value the core is thermally convecting while for lower values of q_c thermal convection is not possible. Onset of inner core solidification occurs at the filled circles where q_c undergoes an abrupt change in variation with time. The dashed curve indicates the thermal evolution without inner core freezing.

which is maintained above the conductive heat flow along the core adiabat for the rest of geologic time (Figure 13.8). Core heat flow tends toward a plateau at late times, depending on the particular value of $L + E_G$. Convection in the outer core is driven both thermally and compositionally subsequent to inner core freezing, an important implication of these models for the maintenance of core convection and dynamo generation of the geomagnetic field. Gravitational energy release may be more important in driving the dynamo than latent heat release since the mechanical energy is almost entirely available for dynamo generation (Gubbins, 1977a). Other models of the evolution of the Earth's core and dynamo action within it suggest that thermal convection and compositional convection are both important in the generation of the Earth's magnetic field (Braginsky and Roberts, 1995; Buffett et al., 1996).

Question 13.4: *What is the dominant energy source for driving convection and dynamo action in the Earth's outer core?*

One aspect of our model that is more important for other terrestrial planets than for the Earth (see the next chapter for a discussion of the thermal histories of other planets) is the dependence of the core melting temperature on the concentration of the light alloying element. The core melting temperature decreases with increasing concentration of the light constituent. Since the light element in the core is excluded from the solidifying inner core, its concentration in the liquid outer core increases with time as the inner core freezes. The melting temperature of the outer core accordingly decreases with time, thereby retarding inner core growth. Inner core growth rates in models E1 and E2 at present are 0.25 and $0.20\,\mathrm{m\,Myr^{-1}}$, respectively.

The dependence of the core melting temperature on the minor constituent concentration can have important consequences during core formation. If core differentiation occurs contemporaneous with accretion of the Earth, then increasingly lighter material would segregate into the core as the Earth grows and the melting temperature of the iron alloy material increases with pressure in the growing Earth. This could result in a compositionally stratified core with lighter material on top of heavier material and a form of layered double diffusive convection in the core (Stevenson, 1998). The core evolution model discussed above would require modification to account for this style of core convection.

Present core heat flow values in models E1 and E2 are 18.6 and 24.4 mW m^{-2}, respectively. At the surface of the model Earth, these heat flows would be about 5.6 and 7.3 mW m^{-2} based on $q \propto R^{-2}$ and $R_c = 3{,}485$ km in these models. These values of heat flow from the core are in qualitative accord with estimates of the heat advected by mantle plumes (Davies, 1988b; Sleep, 1990), assuming that all the heat lost from the core is transported through the mantle by advection in plumes.

The core thermal history predicted by these models has interesting implications for the Earth's magnetic field. The model shows the onset of inner core freezing relatively late in the Earth's thermal history, about 2 Gyr ago. Since the Earth's magnetic field is at least 3.5 Gyr old (McElhinny and Senanayake, 1980), the mode of powering the dynamo may have changed during the Earth's evolution. Early in the Earth's thermal history, the magnetic field was probably powered by thermal convection with the heat derived from secular cooling of the fluid core. After initiation of inner core growth, the dominant source of energy for the dynamo became gravitational energy release upon concentration of the light element into the liquid outer core. Latent heat release also contributes to the maintenance of the dynamo, but with diminished effectiveness compared with gravitational energy release because of the Carnot efficiency factor associated with any purely thermal energy source (Gubbins, 1977a). The energy released by gravitational and latent heat over the entire time of inner core growth in models E1 and E2 exceeds the energy made available by secular cooling of the outer core during this time interval by about a factor of 6. The energy release rate on core freezing in both models amounts to several terawatts, the level of power estimated to be necessary to drive the dynamo (Gubbins et al., 1979). The models thus indicate that while inner core solidification can power the dynamo, secular cooling by itself cannot.

An estimate of the Earth's magnetic field strength through geologic time can be derived from the model by equating the energy available for dynamo generation to the Ohmic dissipation rate Φ (Stevenson et al., 1983):

$$\Phi = E_G \frac{dm_{ic}}{dt} + \eta \left(L \frac{dm_{ic}}{dt} - \frac{dE_{th}}{dt} - 4\pi R_c^2 q_{ac} \right) \tag{13.3.16}$$

where η is the Carnot efficiency factor (≈ 0.6 if $E_G > L$), dE_{th}/dt is the rate of change of core thermal energy, and q_{ac} is the heat flow conducted along the core adiabat. Since Ohmic dissipation scales as the square of the current or field, Φ provides an estimate of a nominal nondimensional magnetic field strength $H_m(t)$ through

$$H_m(t) \equiv \{\Phi(t)/\Phi(4.5\,\text{Gyr})\}^{1/2} \tag{13.3.17}$$

Figure 13.9 from Stevenson et al. (1983) shows the nominal field strength based on (13.3.17) as a function of time for models E1 and E2. Caution must be used in interpreting this nominal field strength since changes can occur in either the toroidal or poloidal part of the Earth's magnetic field with or without changes in the other part. Nevertheless, the model result

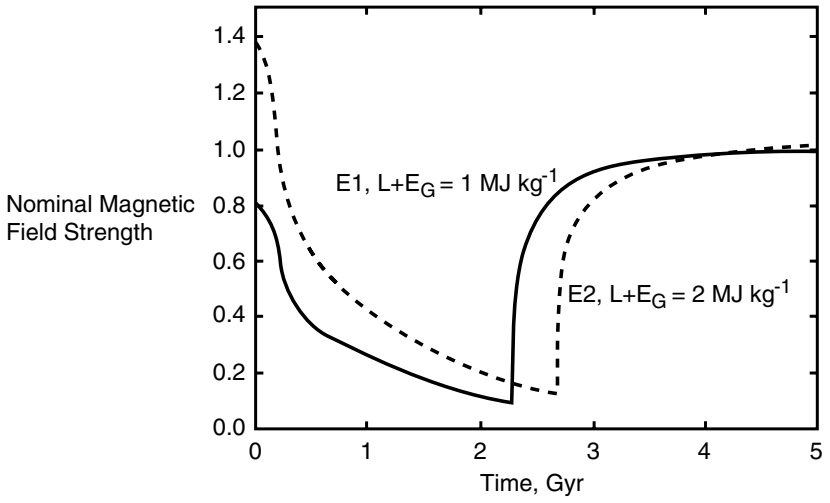


Figure 13.9. Nominal magnetic field strength from (13.3.16) and (13.3.17) versus time for coupled core–mantle thermal evolution models E1 and E2. Magnetic field generation is thermally driven early in the evolution and predominantly gravitationally driven late in the evolution. The abrupt increases in field strength at about 2.5 Gyr mark the switch from a thermally driven dynamo to a gravitationally powered dynamo with the onset of inner core freezing.

gives some indication of possible changes in the Earth's magnetic dipole moment through geologic time. The magnetic field strength declines with time during the first two billion years of evolution as the thermally driven core dynamo decays. Prior to the onset of inner core freezing, magnetic field strengths are low. Upon inner core freezing, the source of energy for the dynamo changes to predominantly gravitational and there is a rapid rise to present magnetic field strengths within about 500 Myr. At present, the paleomagnetic evidence neither supports nor refutes this scenario (Merrill and McElhinny, 1983).

Question 13.5: *Has the Earth always had a geodynamo?*

13.4 Two-layer Mantle Convection and Thermal Evolution

Another generalization of the simple one-layer thermal history model allows us to explore how temperature in the Earth would have evolved in time with separate upper and lower mantle convection systems. This style of mantle convection requires that some component of the density change near 660 km depth be due to a difference in composition between the upper and lower mantle or that a large increase in viscosity between the upper and lower mantle occurs at the 660 km seismic discontinuity. The relative merits of whole-mantle convection versus layered mantle convection are discussed in other chapters. Here we simply explore the consequences of the different styles of convection for Earth thermal history. We will follow the layered mantle convection thermal history model of Spohn and Schubert (1982a). McKenzie and Richter (1981) and Richter (1985) have also analyzed layered mantle convection thermal history models and Christensen (1981) has obtained numerical solutions of convection in a chemically layered mantle. Honda (1995) has studied a parameterized thermal history model in which the mantle, initially in a state of layered convection, undergoes a transition to whole-mantle or one-layer convection. The coupled core–mantle thermal history model of the previous section is similar to the layered mantle

model of this section in the sense that both are coupled two-layer models. Some of the general results we arrive at with the layered mantle model can also be obtained with the coupled core–mantle model.

The model is sketched in Figure 13.10. It consists of two concentric spherical shells surrounding a spherical core. The outer spherical shell coincides with the upper mantle and the inner spherical shell corresponds to the lower mantle. There are thermal boundary layers at the top and bottom of both spherical shells as appropriate for vigorously convecting layers heated partly from below. Accordingly, there are two thermal boundary layers immediately

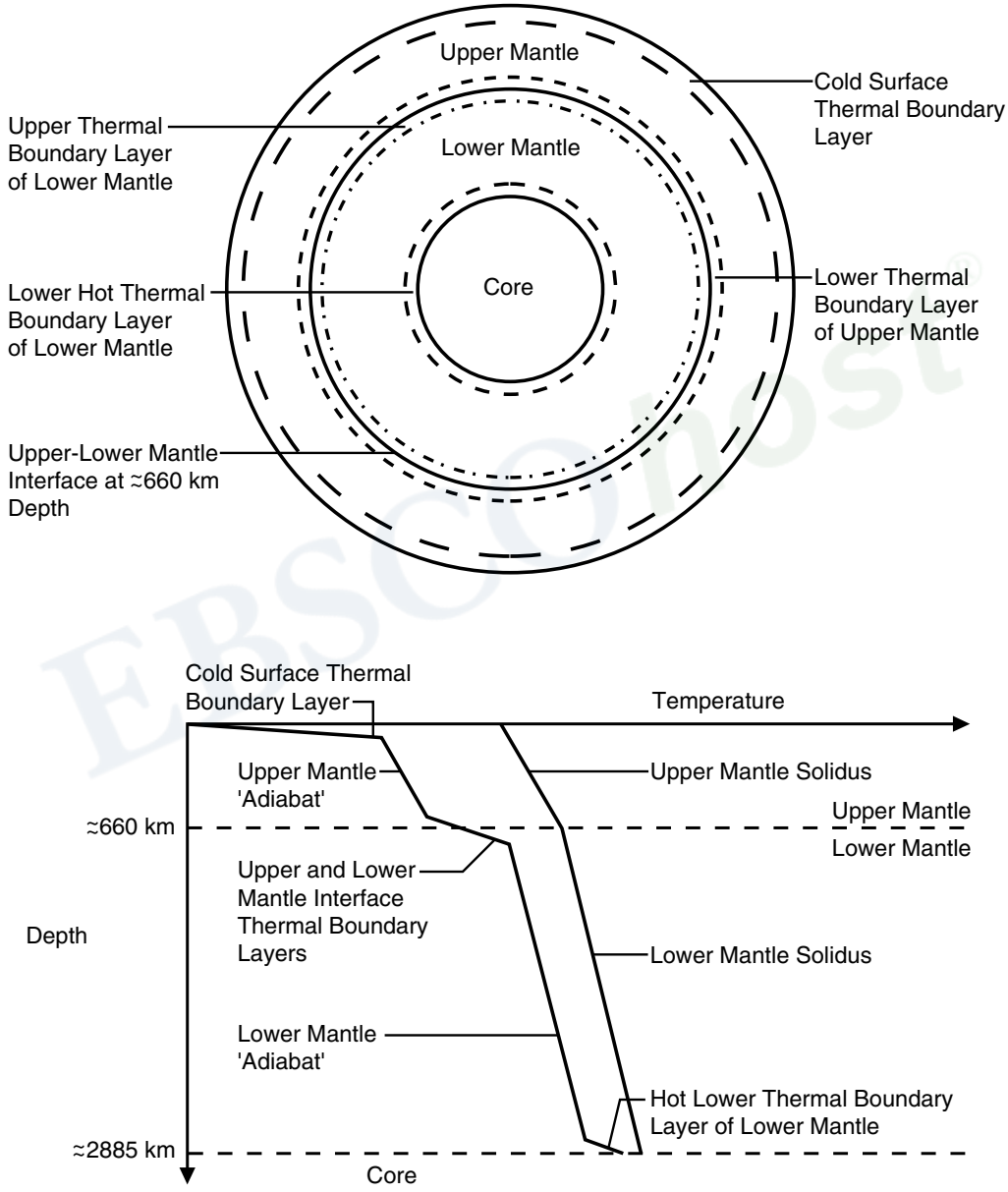


Figure 13.10. Sketch of the two-layer mantle convection thermal history model and the temperature distribution in the model. There are thermal boundary layers on both sides of the upper–lower mantle interface and thermal boundary layers at the surface and the core–mantle interface. The adiabatic temperature profiles in the upper and lower mantle are assumed to parallel the increase in temperature with depth of the upper and lower mantle solidus temperature profiles.

adjacent to the interface between the shells, the lower thermal boundary layer of the upper mantle and the upper thermal boundary layer of the lower mantle. Heat transfer across the boundary layers is by thermal conduction, while outside the boundary layers heat transfer is mainly by advection. The interface boundary layers are therefore additional sources of thermal resistance that are not present in a whole mantle model of convection. For this reason, two-layer mantle convection is less efficient than single-layer or whole-mantle convection, and results in higher mantle temperatures.

Temperature is taken to be a linear function of depth within each boundary layer. Temperature is also assumed to vary linearly with radius in the adiabatic interiors of the shells; the same assumption is made for the solidus temperature of the mantle rocks. For simplicity, the depth profiles of interior temperature are assumed to be parallel to the depth profiles of the mantle solidus. We allow the upper and lower mantle solidus temperatures to have different radial gradients, but it is assumed that solidus temperature is continuous at the upper–lower mantle interface. An interior temperature that is a fixed fraction of the solidus temperature is consistent with isoviscous upper and lower mantles (Weertman, 1970). The ratio of mantle temperature to mantle solidus temperature, known as the homologous temperature, can take different values in the upper mantle and lower mantle of the model. The upper and lower mantle homologous temperatures are functions only of time in the model. Mantle viscosity in the model is proportional to the exponential of the inverse homologous temperature, similar to the dependence of viscosity on actual temperature in (13.2.6).

The model contains internal radiogenic heat sources that are distributed uniformly in the upper and lower mantle shells. The heat source densities in the upper and lower shells of the model are generally unequal and decay with time according to the simple exponential decay law in (13.2.2). Separate energy balance equations govern the cooling histories of the two shells and the core. The shell energy balance equations are identical to (13.3.6) while the core energy balance equation is given by (13.3.7) with $L + E_G = 0$.

Heat fluxes across the thermal boundary layers are given by Fourier's law of heat conduction as in (13.3.10) and (13.3.11) and boundary layer thicknesses are specified as in (13.3.13) with the Rayleigh number of each shell defined in terms of the nonadiabatic temperature rise across each shell and the geometric, thermal, and rheological properties of each shell.

The model equations are integrated forward in time assuming that the mantle is initially at the solidus. Continuity of temperature is applied at interfaces and the surface temperature is held constant at 300 K. Parameter values for the two-layer model of this section are listed in Table 13.3. In addition to the parameters given in Table 13.3, the depth profile of the solidus temperature has the same slope in both upper and lower mantle shells to facilitate comparison with a whole-layer mantle convection model having the same solidus temperature (this whole-layer model has $\nu_0 = 100\,\mathrm{m^2\,s^{-1}}$). The radiogenic heat source densities in the upper and lower mantle shells are determined by requiring the model to have a present surface heat flux of about $60\,\mathrm{mW\,m^{-2}}$ and an approximately isoviscous mantle with a kinematic viscosity of about $10^{17}\,\mathrm{m^2\,s^{-1}}$.

Results of the calculations are presented in Table 13.4 and in Figures 13.11 and 13.12. According to Table 13.4 the concentration of radiogenic heat sources in the lower mantle shell of the two-layer model is only about 1.5% of the average heat source concentration in the entire mantle. We infer that in two-layer convection the lower mantle must be strongly depleted in radiogenic heat sources, because the thermal boundary layers at the upper mantle–lower mantle interface are limited in the amount of heat they can conduct across the interface. The limitation arises because the combined temperature difference across the boundary layers cannot exceed the difference between the solidus temperature and the upper mantle adiabatic

Table 13.3. ***Parameter Values for Two-layer Mantle Convection Thermal Evolution Models of the Earth***[a]

Parameter	Value
ρc	$4.2\,\mathrm{MJ\,m^{-3}\,K^{-1}}$
ρc (core)	$10.9\,\mathrm{MJ\,m^{-3}\,K^{-1}}$
k	$4.2\,\mathrm{W\,m^{-1}\,K^{-1}}$
κ	$10^{-6}\,\mathrm{m^2\,s^{-1}}$
α	$3 \times 10^{-5}\,\mathrm{K^{-1}}$
g	$10\,\mathrm{m\,s^{-2}}$
λ	$1.23 \times 10^{-17}\,\mathrm{s^{-1}}$
R_p	6,371 km
Radius of upper–lower mantle interface	5,671 km
Core radius	3,485 km
ν_0 (upper mantle)	$100\,\mathrm{m^2\,s^{-1}}$
ν_0 (lower mantle)	$6 \times 10^3\,\mathrm{m^2\,s^{-1}}$
β	1/3
Ra_{cr}	10^3
Solidus temperature at surface	1,500 K
Solidus temperature at core–mantle boundary	3,900 K
Activation parameter for the homologous temperature in the viscosity law	30

[a] Unless otherwise stated, upper and lower mantle parameter values are the same.

Table 13.4. ***Characteristics of Two-layer and Whole-layer Mantle Convection Thermal History Models with Similar Values of Surface Heat Flux and Mantle Viscosity***[a]

	Two-layer Mantle Convection Model		Whole-mantle Convection Model
	Upper Mantle	Lower Mantle	
ρH ($\mu\mathrm{W\,m^{-3}}$)	0.44	0.25×10^{-2}	0.13
Surface heat flux ($\mathrm{mW\,m^{-2}}$)	58.6		58.6
ν ($\mathrm{m^2\,s^{-1}}$)	1.4×10^{17}	1.6×10^{17}	1.6×10^{17}
Urey ratio	84.3%		69.4%

[a] All quantities are present values except for ρH, the initial rate of heat production per unit volume.

temperature without melting the lower mantle. Schubert and Spohn (1981) have shown that if the mantle convects in two layers, the lower mantle could not be solid at present if it contained more than about 10% of all mantle radiogenic heat sources.

In comparison, whole-mantle convection is much more efficient at removing heat from the Earth's interior. An initial heat generation rate per unit volume of only $0.13\,\mu\mathrm{W\,m^{-3}}$ is required to balance the average present surface heat flux in the whole-mantle model. This is about 17% less than the mantle average for the two-layer model. Thus, the present ratio of heat generation to heat loss, the Urey ratio, is only 69%, about 15% less than the value for two-layer convection. The whole-mantle convection model removes about 2×10^{30} J

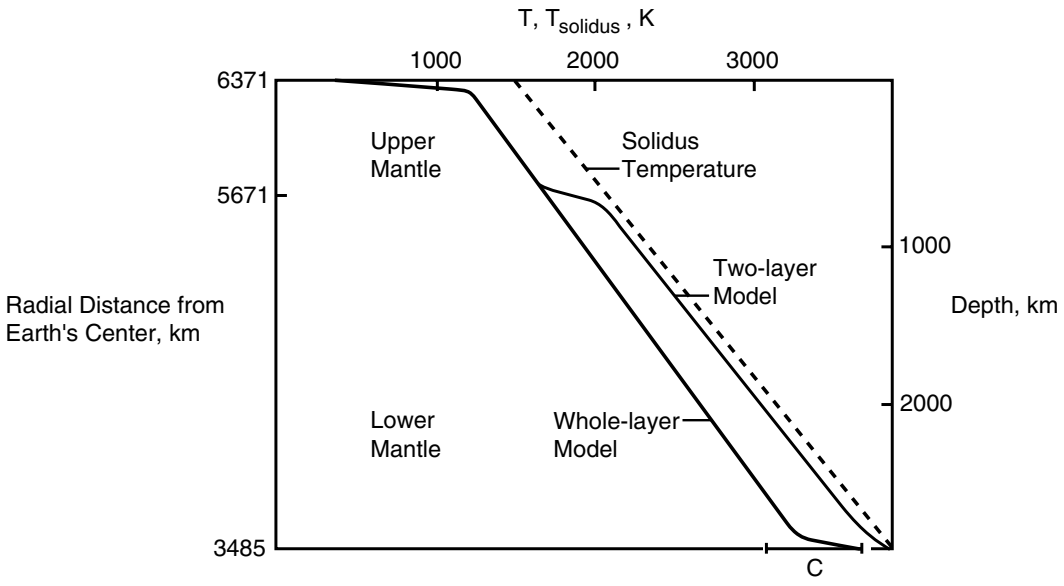


Figure 13.11. Present geotherms (solid curves) in the two-layer and whole-layer mantle convection models of Table 13.4. The mantle solidus temperature (dashed line) is the same in both thermal history models. The two-layer convection model has a hotter lower mantle than does the whole-layer model primarily due to the temperature increases across the interface thermal boundary layers in the two-layer model.

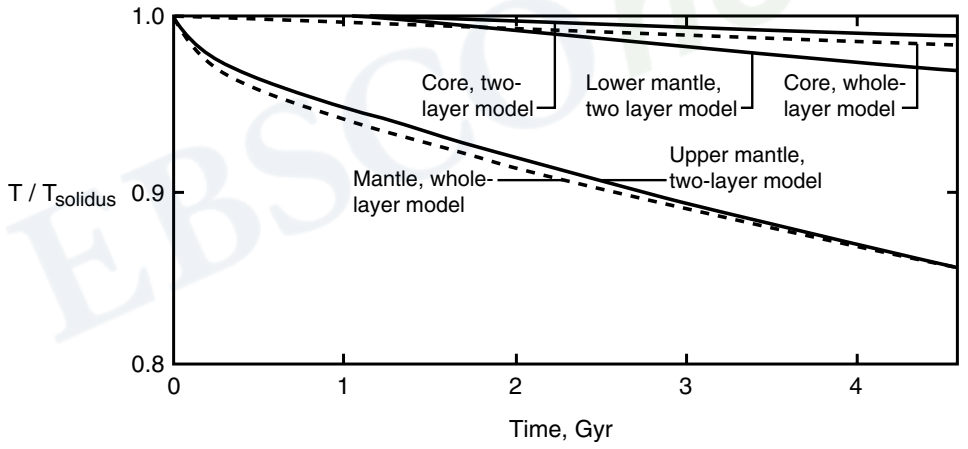


Figure 13.12. Thermal histories in the two-layer and whole-layer mantle convection models. The core temperatures are normalized relative to the value of the solidus temperature at the core–mantle boundary.

of primordial heat over its entire thermal history, while the two-layer model removes only about 6×10^{29} J, about a factor of 3 less. A larger fraction of the Earth's present surface heat flow would have to be attributed to radiogenic heating if the Earth's mantle convects in two or more layers than if it convects as a single layer.

The model results summarized in Table 13.4 have shown that the two-layer mode of mantle convection is less efficient at cooling the Earth than is the whole-layer mode of mantle convection and that the lower mantle must be strongly depleted in radiogenic heat sources if the mantle convects in two layers. Spohn and Schubert (1982a) have shown that these conclusions are robust to variations in the values of upper mantle and lower mantle rheological parameters, depth profiles of solidus temperature including unequal slopes of the

solidus curves in the upper mantle and lower mantle, assumptions about the thicknesses of the lower thermal boundary layers including the complete disappearance of these layers, and choices for the relative slopes of the depth profiles of the upper and lower mantle adiabats and solidus temperatures.

The present geotherms for the two-layer and whole-layer mantle convection thermal history models of Table 13.4 are shown in Figure 13.11, together with the mantle solidus curve. Both models have a surface thermal boundary layer or lithosphere about 80 km thick, with a temperature rise of about 940 K. Temperatures in the upper mantle are essentially the same for both models, approximately 0.80 of the solidus temperature. However, in the two-layer model temperature increases by $\approx$ 480 K across the thermal boundary layers separating the upper and lower mantles. Therefore, temperatures in the lower mantle of the two-layer model are about 360 K higher than those in the whole-layer model. Core temperatures in these models differ by $\approx$ 220 K. The lower mantle temperature rises by $\approx$ 330 K across the lower thermal boundary layer in the whole-mantle model compared with $\approx$ 170 K in the two-layer model. This boundary layer is $\approx$ 140 km thick in the two-layer model but only 80 km thick in the whole-mantle model. The present heat flux from the core is therefore about 3.5 times larger for the whole-mantle model. The core heat flux was even larger earlier in the Earth's thermal history since its time integral is $\approx$ 6 times larger for the whole-mantle model. Figure 13.11 clearly shows that whole-mantle convection not only removes more heat from the core, but it also removes more primordial heat from the mantle. The primordial heat removed after 4.5 Gyr is proportional to the area between the geotherm and the solidus curve. Thus the excess primordial heat removed by whole-mantle convection is proportional to the area between the two geotherms and amounts to about 10^{30} J after correction for differences in radiogenic heat production.

The thermal histories of the models of Table 13.4 are illustrated in Figure 13.12. The thermal evolution of the upper mantle for two-layer convection is very similar to that of the entire mantle for single-layer convection. However, the whole-mantle convection model cools the Earth's interior much more efficiently because of the absence of internal boundary layers. The lower mantle thermal history in the two-layer model is quite similar to the thermal evolutions of the cores in both the one- and two-layer models. The upper mantle of the two-layer system and the entire mantle of the single-layer model cool off very fast initially; the initial high cooling rate decays exponentially with time, and the cooling rate becomes approximately steady after the first billion years.

The cooling histories of upper mantle and whole mantle, and lower mantle and core, reflect fundamental differences between thermally insulated layers and freely cooling layers driven by large nonadiabatic temperature differences. The nonadiabatic temperature difference across the lower mantle increases as the upper mantle cools. Similarly, cooling of the core depends on the development of boundary layers at the core–mantle interface and is restricted by the temperature increase across these boundary layers. Because of the strong temperature dependence of mantle viscosity, most of the total nonadiabatic temperature rise occurs across the surface boundary layer, even late in the thermal history.

Question 13.6: *Did layered mantle convection ever occur in the Earth's thermal history?*

13.5 Scaling Laws for Convection with Strongly Temperature Dependent Viscosity

As already noted in Section (13.2.2), the heat flow–Rayleigh number parameterization (13.2.4) is valid only for convection with constant viscosity. Application of this parameterization to studies of planetary thermal history is therefore limited by the fact that mantle viscosity is strongly temperature dependent. One major effect of the strong temperature dependence of viscosity on thermal convection is the creation of a high-viscosity region near the upper surface where temperatures are relatively cold. The high-viscosity near-surface layer can participate sluggishly in the convection or "freeze up" to form an immobile or stagnant lid, depending on the "strength" of the viscosity variation with temperature and the temperature difference across the layer. Formation of a sluggish or rigid lid in convection with strongly temperature dependent viscosity reduces the efficiency of heat transport across the layer because cold near-surface material cannot be effectively circulated to deeper and hotter parts of the layer. Methods for incorporating this reduction in heat transfer efficiency into (13.2.4) have been mentioned in Section 13.2.2. One way is to interpret T in (13.2.4) and (13.2.5) as the temperature of the efficiently convecting material below the sluggish or rigid lid.

The sluggish or rigid lid of a convecting system with strongly temperature dependent viscosity is the analogue of the lithosphere on the Earth and other planets. In contrast to other planets however, Earth's lithosphere does not act as a globally intact rigid lid. Nonviscous deformation mechanisms (e.g., faulting) allow it to break up into pieces (plates), many of which (oceanic plates) are subductible. The net result is that the Earth's "rigid lid" can be circulated deeply into the hot mantle and mantle convective heat transfer is essentially as efficient as if the mantle were convecting as a constant viscosity fluid. Thus, for the Earth, (13.2.4) suffices to study its thermal evolution. However in the absence of plate tectonics, the other terrestrial planets do possess globally intact lithospheres which must behave as the sluggish or rigid lids of convection with strongly temperature dependent viscosity. Accordingly, thermal history investigations of the other planets (Chapter 14) can benefit from heat flow–Rayleigh number parameterizations specifically formulated to account for the strong temperature dependence of mantle viscosity. We discuss these parameterizations here and in the next chapter.

Parameterization of heat transport by convection in a constant viscosity fluid layer heated from below with isothermal and stress-free top and bottom boundaries requires only two dimensionless parameters, the Nusselt number Nu and the Rayleigh number Ra. When the viscosity μ of the fluid is strongly temperature T dependent, however, an additional dimensionless parameter is needed to characterize $\mu(T)$. Further, since viscosity enters the formula for Ra, it is necessary to specify the temperature at which the viscosity is evaluated in Ra. With temperature-dependent viscosity, the definitions of Ra and the Nu–Ra relation become nonunique, and the literature on convection with strongly temperature dependent viscosity reflects this nonuniqueness. In one approach, the Rayleigh number Ra_0 is defined in terms of the viscosity μ_0 evaluated at the temperature T_0 of the upper surface:

$$Ra_0 \equiv \frac{\rho g \alpha \Delta T d^3}{\kappa \mu\,(T_0)} = \frac{\rho g \alpha \Delta T d^3}{\kappa \mu_0} \tag{13.5.1}$$

where ΔT is the total temperature change across the fluid layer of thickness d

$$\Delta T \equiv T_1 - T_0 \tag{13.5.2}$$

and T_1 is the temperature of the lower surface. The Rayleigh number Ra_1 has also been defined in terms of the viscosity μ_1 evaluated at the temperature T_1 of the lower surface:

$$Ra_1 \equiv \frac{\rho g \alpha \Delta T d^3}{\kappa \mu (T_1)} = \frac{\rho g \alpha \Delta T d^3}{\kappa \mu_1} \tag{13.5.3}$$

A Rayleigh number $Ra_{1/2}$ based on the viscosity $\mu_{1/2}$ evaluated at the average temperature $(T_0 + T_1)/2$ of the upper and lower surfaces is often used:

$$Ra_{1/2} \equiv \frac{\rho g \alpha \Delta T d^3}{\kappa \mu ((T_0 + T_1)/2)} = \frac{\rho g \alpha \Delta T d^3}{\kappa \mu_{1/2}} \tag{13.5.4}$$

Finally, a Rayleigh number Ra_i based on the viscosity μ_i evaluated at the nearly uniform temperature of the actively convecting layer beneath the sluggish or rigid lid is also widely employed:

$$Ra_i \equiv \frac{\rho g \alpha \Delta T d^3}{\kappa \mu (T_i)} = \frac{\rho g \alpha \Delta T d^3}{\kappa \mu_i} \tag{13.5.5}$$

The specific form of $\mu(T)$ adopted in some studies is the Arrhenius law (13.2.6), but a linearized version of this law

$$\frac{\mu}{\mu_{\text{ref}}} = \exp\{-E(T - T_{\text{ref}})\} \tag{13.5.6}$$

is commonly used. In (13.5.6), T_{ref} and μ_{ref} are the reference temperature and viscosity values used in the definition of the Rayleigh number. With (13.5.6), the viscosity ratio across the layer r_μ is given by

$$r_\mu \equiv \frac{\mu_0}{\mu_1} = \exp(E \Delta T) \tag{13.5.7}$$

independent of the choice of T_{ref} and μ_{ref}. The Frank–Kamenetskii parameter θ is

$$\theta = \ln r_\mu = E \Delta T \tag{13.5.8}$$

For the viscosity law (13.5.6) the heat flow (Nusselt number)-Rayleigh number parameterization can be written as

$$Nu = Nu(\theta, Ra) \tag{13.5.9}$$

where Ra in (13.5.9) is one of the Rayleigh numbers defined above. The Rayleigh numbers Ra_1, $Ra_{1/2}$, and Ra_0 are related to each other by

$$Ra_1 = Ra_{1/2}\sqrt{r_\mu} = r_\mu Ra_0 \tag{13.5.10}$$

Before discussing heat flow–Rayleigh number parameterizations appropriate to convection with strongly temperature dependent viscosity, we first describe the nature of this type of convection using the following thought experiment. Imagine that $Ra_{1/2}$ is kept constant at some large value (so that convection occurs) in a sequence of experiments in which

r_μ is increased from 1. At small viscosity contrasts, the cold fluid near the upper boundary is entirely mobile and participates freely in the convective motions. This is the small viscosity contrast regime and it occurs for values of r_μ less than about 10^2. As viscosity contrasts grow larger (r_μ in the range 10^2–10^3), cold fluid near the upper boundary becomes increasingly more viscous and is less able to participate in convective overturning; a sluggish lid of cold viscous fluid develops. This is the sluggish-lid convection regime. When the viscosity contrast is increased still further (r_μ about 10^4), the cold fluid at the upper surface becomes so viscous as to form a stagnant lid, effectively ceasing to participate in the convective motions that occur below it. This is the stagnant-lid mode of convection. In addition to differences in the style of lid deformation, each of these convective regimes has a distinct horizontal planform. The small viscosity contrast convection regime has a horizontal wavelength comparable to the depth of the fluid layer. The sluggish-lid regime of convection has a larger horizontal scale. However, the horizontal scale of the stagnant-lid convection regime is small compared to the fluid layer depth. The morphologies of the upflows and downflows are also different in the different convection regimes. The dominant mechanism by which strongly temperature dependent viscosity influences the planform and morphology of convection is through the depth dependence of horizontally averaged viscosity. This picture summarizes results by Ratcliff et al. (1997) for fully three dimensional convection with strongly temperature dependent viscosity in both spherical shells and Cartesian boxes. Their results are discussed in the context of planetary convection in Chapter 14. The styles of convection in the different regimes are shown in Figures 14.12 and 14.13.

Different heat flow–Rayleigh number parameterizations apply in the separate convection regimes. Solomatov (1995) and Moresi and Solomatov (1995) use the differences in Nu–Ra relationships from numerical calculations of variable viscosity convection in two-dimensional geometry to identify the different convection regimes. Figure 13.12 shows approximate regime boundaries on a plot of $\log_{10} r_\mu$ versus $\log_{10} Ra_1$ (Solomatov, 1995). Points representing the three-dimensional calculations of Ratcliff et al. (1997) for $Ra_{1/2} = 10^5$ and $r_\mu = 1$, 10^3, 10^4, and 10^5 are indicated in the figure. According to Ratcliff et al. (1997) convection should be in the small viscosity contrast regime for $r_\mu = 1$ (point a, region I), in the sluggish-lid regime for $r_\mu = 10^3$ (point b, region II), and in the stagnant-lid regime for $r_\mu = 10^3$ and 10^4 (points c and d, respectively, region III). The locations of the approximate regime boundaries are in accord with the positions of these points. Regime diagrams for non-Newtonian viscous convection with strongly temperature dependent viscosity have been given in Solomatov (1995) and Reese et al. (1998, 1999). These are similar to the regime diagram of Newtonian viscous convection except that the regime boundaries are shifted to much larger temperature-dependent viscosity contrasts.

The heat flow–Rayleigh number parameterization that applies in the small viscosity contrast regime (Figure 13.13, region I) is given by (13.2.4) with T given by T_i and μ given by μ_i (Solomatov, 1995). Simple scaling arguments lead to (13.2.4) with $\beta = 1/3$ (Schubert et al., 1979a; Solomatov, 1995). For example, if the temperature drop across the layer is distributed equally between the upper and lower boundary layers and if these boundary layers each have the same thickness δ (Figure 13.14a), then the heat flow q from Fourier's law of heat conduction gives

$$q \sim \frac{k\Delta T}{\delta} \tag{13.5.11}$$

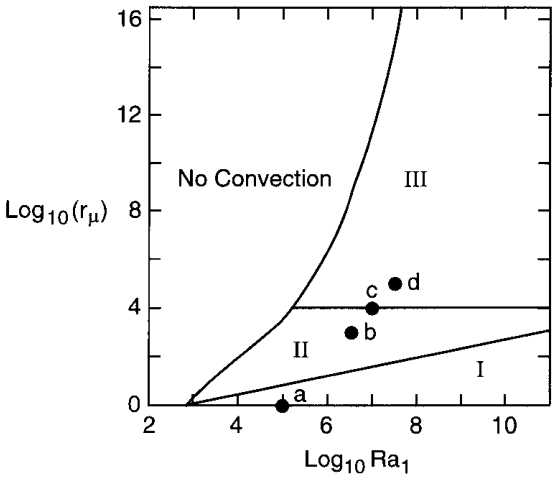


Figure 13.13. A diagram of the approximate boundaries of the small viscosity contrast (I), sluggish-lid (II), and stagnant-lid (III) convection regimes (regime boundaries are redrawn from Solomatov, 1995). The approximate boundary for the onset of convection is also shown. The diagram applies strictly to two-dimensional convection in a layer heated from below with isothermal, stress-free boundaries, but may be qualitatively applied to more general situations. Points a through d represent the three-dimensional convection simulations of Ratcliff et al. (1997) (in a Cartesian box and a spherical shell) for $Ra_{1/2} = 10^5$ and $r_\mu = 1$, 10^3, 10^4, and 10^5 ($Ra_1 = 10^5$, $10^{6.5}$, 10^7, $10^{7.5}$) which were in the small viscosity contrast, sluggish-lid, and stagnant-lid convection regimes, respectively.

Further, if δ scales as

$$\delta \sim \left(\frac{\kappa d}{u}\right)^{1/2} \tag{13.5.12}$$

where u is the magnitude of the horizontal velocity in the upper and lower thermal boundary layers (velocities in the boundary layers are equal in magnitude, Figure 13.14a), then

$$q \sim k\Delta T \left(\frac{u}{\kappa d}\right)^{1/2} \tag{13.5.13}$$

The velocity scale u can be found by equating the integral dissipation rate in the layer to the integral mechanical work done by thermal convection per unit time

$$\mu_i \left(\frac{u}{d}\right)^2 \sim \frac{\alpha g q}{c_p} \tag{13.5.14}$$

or

$$u \sim \left(\frac{\alpha g q d^2}{\mu_i c_p}\right)^{1/2} \tag{13.5.15}$$

Substitution of this velocity scale into (13.5.13) gives

$$q \sim \frac{k\Delta T}{d}\left(\frac{\rho g \alpha \Delta T d^3}{\kappa \mu_i}\right)^{1/3} \tag{13.5.16}$$

or

$$Nu = \frac{q}{(k\Delta T/d)} \sim Ra_i^{1/3} \tag{13.5.17}$$

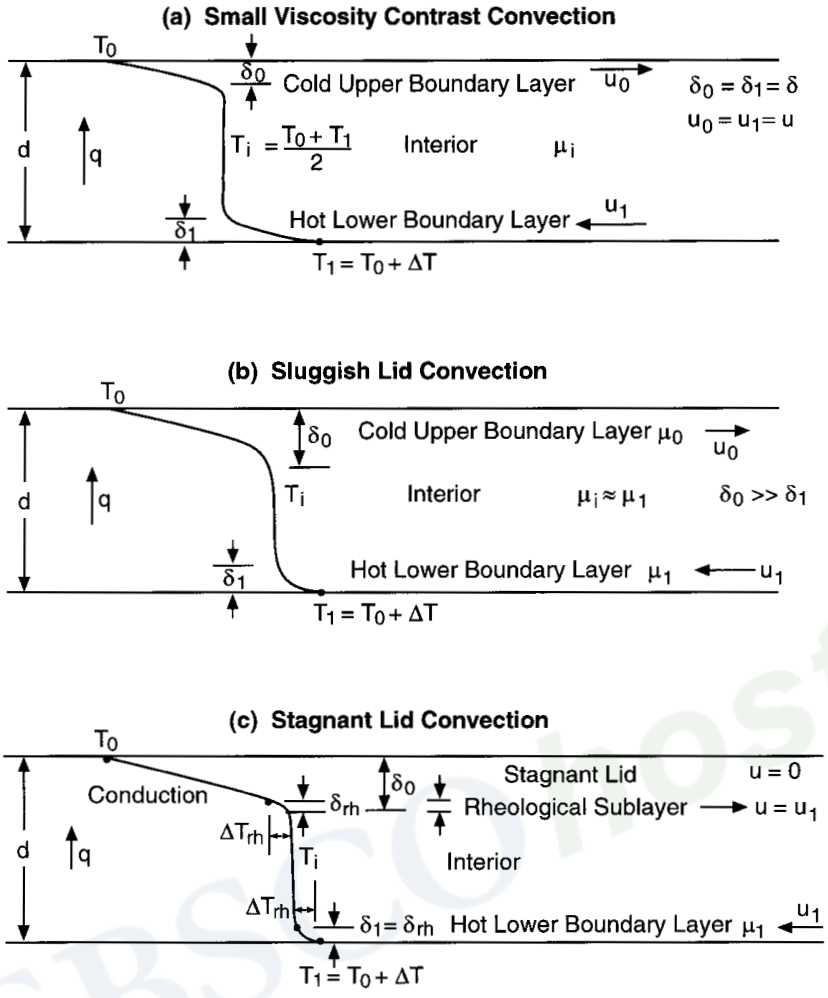


Figure 13.14. Sketch of the different regimes of thermal convection with strongly temperature dependent viscosity; (a) small viscosity: contrast convection regime; (b) sluggish-lid convection regime; (c) stagnant-lid convection regime. As the viscosity contrast across the convecting layer increases, convection undergoes transitions from the small viscosity contrast regime to the sluggish-lid regime and finally to the stagnant-lid regime. The temperature of the convecting interior increases as the viscosity contrast across the layer increases.

In the sluggish-lid convection regime (Figure 13.14b) the scaling proceeds as follows Solomatov (1995). It is assumed that half of the convective rate of doing work $\frac{1}{2}\left(\alpha g q / c_p\right)$ is balanced by the rate of dissipation in the nearly isothermal convecting region beneath the sluggish lid:

$$\mu_i \left(\frac{u_1}{d}\right)^2 \sim \frac{1}{2}\frac{\alpha g q}{c_p} \tag{13.5.18}$$

where u_1 is the horizontal velocity scale in the interior convecting region (similar to the velocity u_1 in the lower hot thermal boundary layer, Figure 13.14b). The other half of the convective work is assumed to overcome dissipation in the cold upper boundary layer:

$$\mu_0 \left(\frac{u_0}{d}\right)^2 \sim \frac{1}{2}\frac{\alpha g q}{c_p} \tag{13.5.19}$$

where subscript zero refers to the cold upper boundary layer. The upper and lower boundary layer thicknesses, δ_0 and δ_1, respectively, are

$$\delta_0 \sim \left(\frac{\kappa d}{u_0}\right)^{1/2} \quad \text{and} \quad \delta_1 \sim \left(\frac{\kappa d}{u_1}\right)^{1/2} \tag{13.5.20}$$

Substitution of (13.5.18) and (13.5.19) into (13.5.20) gives

$$\delta_0 \sim \left(\frac{\alpha g q}{\mu_0 c_p \kappa^2}\right)^{-1/4} \tag{13.5.21}$$

$$\delta_1 \sim \left(\frac{\alpha g q}{\mu_i c_p \kappa^2}\right)^{-1/4} \tag{13.5.22}$$

From Fourier's law of heat conduction and the concept of thermal resistances in series

$$q = \frac{k\Delta T}{d}\left(\frac{\delta_0}{d} + \frac{\delta_1}{d}\right)^{-1} \tag{13.5.23}$$

Substitution of (13.5.23) into (13.5.21) and (13.5.22) together with the assumption $\delta_0 \gg \delta_1$ yields

$$\frac{\delta_0}{d} \sim Ra_0^{-1/3} \tag{13.5.24}$$

and

$$\frac{\delta_1}{d} \sim Ra_i^{-1/4}\left(\frac{\delta_0}{d}\right)^{1/4} = Ra_i^{-1/4} Ra_0^{-1/12} \tag{13.5.25}$$

From (13.5.23) we can write

$$q \sim \frac{(k\Delta T/d)}{(\delta_0/d + \delta_1/d)} \tag{13.5.26}$$

or

$$Nu \sim \frac{1}{(\delta_0/d + \delta_1/d)} \tag{13.5.27}$$

The interior temperature can be determined from the temperature drop across the cold upper boundary layer:

$$T_i - T_0 = \delta_0 \frac{q}{k} = \frac{\delta_0 \Delta T}{\delta_0 + \delta_1} \tag{13.5.28}$$

In the stagnant-lid convection regime (Figure 13.14c) most of the cold upper boundary layer is an immobile conductive region across which almost the entire ΔT occurs (Solomatov, 1995). Convection penetrates only a small distance into the lid given by the rheological length scale δ_{rh} (Morris and Canright, 1984; Fowler, 1985)

$$\delta_{rh} \sim \frac{\delta_0}{E\Delta T} = \frac{\delta_0}{\theta} \tag{13.5.29}$$

and only the small temperature difference across this rheological sublayer ΔT_{rh} is available to drive convection in the underlying region (Figure 13.14c) (Davaille and Jaupart, 1993)

$$\Delta T_{rh} \sim \frac{\Delta T}{\theta} \tag{13.5.30}$$

Convection beneath the stagnant lid is essentially constant viscosity convection driven by the temperature difference ΔT_{rh}. By replacing ΔT with ΔT_{rh} in (13.5.16) we can immediately write

$$q \sim \frac{k\Delta T}{d}\theta^{-4/3} Ra_i^{1/3} \tag{13.5.31}$$

or

$$Nu \sim \theta^{-4/3} Ra_i^{1/3} \tag{13.5.32}$$

Since q is also given by

$$q \sim \frac{k\Delta T}{\delta_0} \tag{13.5.33}$$

we find

$$\frac{\delta_0}{d} \sim \theta^{4/3} Ra_i^{-1/3} \tag{13.5.34}$$

The heat flux q is also given by

$$q \sim \frac{k\Delta T_{rh}}{\delta_1} \tag{13.5.35}$$

which, together with (13.5.30) and (13.5.33), results in

$$\frac{\delta_1}{d} \sim \frac{1}{\theta}\frac{\delta_0}{d} \tag{13.5.36}$$

Heat flow–Rayleigh number parameterizations and scaling relationships for non-Newtonian viscous convection with strongly temperature dependent viscosity have been given by Solomatov (1995) and Reese et al. (1998, 1999). The heat flow–Rayleigh number parameterizations for the different convection regimes can be used to study the thermal histories of other terrestrial planets, as discussed in Chapter 14. Such studies rely on the convection regime appropriate for a planet at each stage in its evolution. The regime

diagram of Figure 13.13 provides the information necessary to determine the style of convection. Transitions between different styles of convection could occur as a planet evolves with perhaps discernible consequences at its surface (see the discussion of Venus' thermal history in the next chapter).

The heat flow–Rayleigh number parameterizations discussed above pertain to convection in fluid layers heated from below. Grasset and Parmentier (1998) have studied convection in volumetrically heated fluid layers with strongly temperature dependent viscosity and Davaille and Jaupart (1993) have studied the closely related problem of the transient cooling of fluids with strongly temperature dependent viscosity. With volumetric heating the appropriate form of the Rayleigh number Ra_0 is

$$Ra_0 = \frac{\alpha g \rho^2 H d^5}{\kappa k \mu_0} \tag{13.5.37}$$

where H is the rate of internal heating per unit mass (see 13.2.15). For Ra_0 sufficiently small, convection occurs in a stagnant-lid regime similar to that of bottom-heated convection: a stagnant conductive lid forms at the surface and convection occurs below the lid (Figure 13.15). The convective region is essentially isothermal with temperature T_i. The temperature difference between the isothermal interior and the bottom of the stagnant conductive lid $T_i - T_c$ (T_c is the temperature at the base of the conductive lid) is a rheological temperature difference given by an equation analogous to (13.5.30) (Davaille and Jaupart, 1993; Grasset and Parmentier, 1998):

$$T_i - T_c = 2.23 \left(\frac{-\mu}{(d\mu/dT)} \right)_{T=T_i} \tag{13.5.38}$$

For the viscosity law (13.5.6), $T_i - T_c$ is $2.23E^{-1}$ and the viscosity ratio $\mu\,(T = T_c)\,/\mu\,(T = T_i)$ is exp(2.23) or nearly a factor of 10. For an Arrhenius viscosity law (13.2.6), $T_i - T_c$ is $2.23\,RT_i^2/E^*$, where E^* is the activation energy.

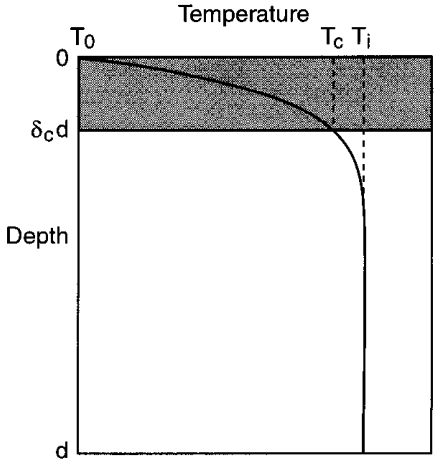


Figure 13.15. Illustration of the nature of stagnant-lid convection in an internally heated fluid layer with strongly temperature dependent viscosity (redrawn from Grasset and Parmentier, 1998).

The heat flow–Rayleigh number parameterization for stagnant-lid, heated from within, convection with strongly temperature dependent viscosity is of the form

$$\frac{T_i - T_c}{H\left(d - \delta_c d\right)^2 / k} = a \left\{ \frac{\alpha g \rho^2 H \left(d - \delta_c d\right)^5}{\kappa k \mu_i} \right\}^{\beta} \tag{13.5.39}$$

where $\delta_c d$ is the thickness of the conducting lid and interior temperature replaces heat flux in the parameterization for fluids heated volumetrically. This is identical to the parameterization for constant viscosity fluids; the appropriate value of β is also the same as it is in constant viscosity internally heated convection.

According to Grasset and Parmentier (1998), constant viscosity parameterized convection laws can be applied to the stagnant-lid convection regime if the quantities in the parameterization are properly identified with those of the convecting region beneath the stagnant conducting lid and if the temperature difference between the convecting region and the base of the lid is given by (13.5.38). Application of this approach to planetary thermal evolution requires that the temperature difference $T_i - T_c$ be calculated from (13.5.38) implying that T_c must evolve with T_i and not be specified a priori at some constant value. In other words, the temperature at the base of the conducting lid T_c evolves with the mantle temperature and T_c is not a constant. Mantle convection evolves so as to maintain a temperature difference below the conducting lid given by the characteristic rheological temperature difference. The conducting lid (lithosphere) will thicken as the mantle cools but the temperature at the base of the lid will also decrease with time. This effect will result in model lithospheres that are thicker than those predicted in models where lid basal temperature is held fixed. An example of this will be given in Chapter 14 when we discuss the thermal history of Mars.

Numerical calculations of steady, two-dimensional convection with strongly temperature dependent viscosity for very large values of Ra_0 show that transitions can occur among the modes of variable viscosity convection. Hansen and Yuen (1993) have found a transition from stagnant-lid convection to a small viscosity contrast regime at Ra_0 around 10^7 for fixed $r_\mu = 10^3$. In this small viscosity contrast regime at high Ra_0, Nusselt numbers are about 100 and surface temperatures are around 1,000 K or more. Such a mode of convection might have been relevant to the early Earth shortly after the end of accretion when a dense water-dominated atmosphere might have produced high surface temperatures through the greenhouse effect (see Section 13.2.1 and references therein for a discussion of an early massive water atmosphere near the end of accretion). This high Ra_0 mode of convection might also be relevant to mantle convection in Venus at present (surface temperature on Venus is about 700 K).

Question 13.7: *Has convection in the Earth's mantle always occurred in the small viscosity contrast regime?*

13.6 Episodicity in the Thermal Evolution of the Earth

A major characteristic of the simple parameterized convection models of the Earth's thermal history presented in this chapter is the monotonic decline in mantle temperature, surface heat flow, and convective vigor. While the Earth has undoubtedly undergone gradual cooling over geologic time, there probably have been periods in Earth history when this secular decline was interrupted by intervals of enhanced convective vigor and surface heat flow. For example,

we know from the geologic record that major flood basalt events have occurred and plate motions have undergone rapid changes.

Numerical models show that mantle convection may be an inherently chaotic phenomenon capable of unpredictable spurts of enhanced or decreased activity. Convection models also demonstrate that the mantle is capable of sudden and perhaps catastrophic movements of material and heat. The most dramatic examples are the so-called avalanches of cold material piled up on the phase change at 660 km depth (Machetel and Weber, 1991; Tackley et al., 1993; Honda et al., 1993b, a; Solheim and Peltier, 1994b), plume or diapir release from the thermal boundary layer at the base of the mantle and plume arrival at the base of the lithosphere, plate tectonic flux variability due to ridge subduction events and continental collisions, and delamination of parts of the lithosphere.

Question 13.8: *Is convection in the Earth's mantle chaotic?*

Question 13.9: *Have avalanches occurred in the Earth's mantle?*

One way of studying nonmonotonicity in Earth thermal history is to incorporate the physical processes responsible for such behavior into numerical models. Another way is to add ad hoc conditions to parameterized convection models that allow them to change behaviors in a priori specified ways. The disadvantage of this approach is that the model behaves in a way that is forced upon it by the externally imposed conditions. We will discuss studies of the Earth's thermal history that use both approaches. Major changes have also occurred in the thermal histories of other planets and Chapter 14 discusses such an event for Venus.

As noted above, mantle avalanches have the potential to change the Earth in major ways (Weinstein, 1993). Avalanches could occur on a global scale and cause layered mantle convection to suddenly change to a whole-mantle convection pattern. Less dramatic but still significant would be the occurrence of avalanches on a regional scale in a partially layered mantle. The sudden arrival of cold avalanche material at the core–mantle boundary would displace hot material in the thermal boundary layer and might cause the ejection of massive hot plumes. The enhanced heat flow from the core into the suddenly cold overlying mantle would affect motions in the core and cause changes in the geomagnetic field; the effect might result in changes in the frequency of magnetic field reversals (Larson and Olson, 1991). Hot material from the lower mantle would be rapidly injected into the upper mantle in order to conserve mass in an avalanche with attendant thermal consequences at the surface. The sudden arrival of hot material beneath the plates from the mass-compensating injection of hot lower mantle material into the upper mantle or from massive hot plumes ejected from the core–mantle boundary layer could cause abrupt changes in plate motion and the creation of new plate margins (Brunet and Machetel, 1998; Ratcliff et al., 1998). A global avalanche or flush instability causing a change in mantle convection from the two-layer to the whole-mantle mode has been invoked by Breuer and Spohn (1995) to explain the Archean–Proterozoic transition. The late Archean was a time of profound geologic change and rapid continental growth (Taylor and McLennan, 1995).

The possibility that plumes can dramatically influence plate motions by lubricating the undersides of plates is, of course, not necessarily linked to the occurrence of mantle avalanches. Plumes are a fundamental feature of thermal convection; they occur whenever

there is some heating from below. Plume activity is naturally variable in time-dependent convection and it can be expected that large plumes in a chaotically convecting mantle will aperiodically impinge on the bottom of the lithosphere, spread hot material beneath it, and change plate velocities (Larsen et al., 1996b; Larsen and Yuen, 1997a; Ratcliff et al., 1998).

The tendency of the spinel–perovskite + magnesiowüstite phase transition to promote layering in mantle convection has been found to be stronger at higher Rayleigh number (Christensen and Yuen, 1985; Zhao et al., 1992). Accordingly, mantle convection in the early Earth may have been layered because of more vigorous convection at higher Rayleigh number than at present. If mantle convection is not layered or is only partially layered today then the Earth would have experienced a transition from two-layer mantle convection to whole-mantle convection at some time in its evolution. As discussed above, such a transition might have had a profound impact on the geologic record. The same idea has been applied to Venus by Steinbach and Yuen (1992), who suggested that this transition from layered to whole-mantle convection could have caused the major resurfacing of Venus some 750 Myr ago (see Chapter 14 for a discussion of Venus and its thermal history).

The general decrease in Rayleigh number with time as the Earth evolved could have led to another transition in convective style with thermal and tectonic consequences. At extremely high Rayleigh numbers (greater than about 10^7) thermal convection occurs in a regime of hard turbulence (Hansen et al., 1990, 1992b, a). Convection in this regime involves the rise of disconnected thermals or plumes or diapirs from the lower hot boundary layer. In the regime of soft turbulence that occurs at lower Rayleigh numbers, plumes remain connected to their source in the lower boundary layer. The vigorously convecting mantle in the early Earth could have been in the regime of hard turbulence while the transition to the less time-dependent chaotic state as the vigor of convection waned left the present mantle in the regime of soft turbulence (Yuen et al., 1993).

A parameterized convection approach to study the effects of phase-change-induced layering and associated mantle avalanches on Earth thermal history has been put forward by Davies (1995b). His model assumes that the mantle convects in two layers with ad hoc conditions for the breakdown of layering based on the attainment of either a critical temperature difference between the layers or a sufficiently cold upper mantle temperature. The Earth cools through geologic time in this model, but it can do so while experiencing periods of layered and nonlayered convection separated by major overturns of the mantle. Periods of layered convection in this model occur more often in the early evolution of the Earth and are gradually replaced by whole-mantle convection as the Earth evolves to the present. Episodic overturns could result in major spurts of tectonic activity and continental crustal formation.

Question 13.10: *What dynamical processes in the mantle are responsible for episodicity in the geological record?*

13.7 Continental Crustal Growth and Earth Thermal History

One of the major products of the Earth's thermal evolution is the formation of the continental crust. The continental crust is also the only accessible repository of information about the Earth's thermal state billions of years ago. In order to decipher the record of the Earth's thermal evolution contained in the continental crust, it is necessary to characterize the age distribution of the crust (Figure 13.16) and understand the processes involved in crustal

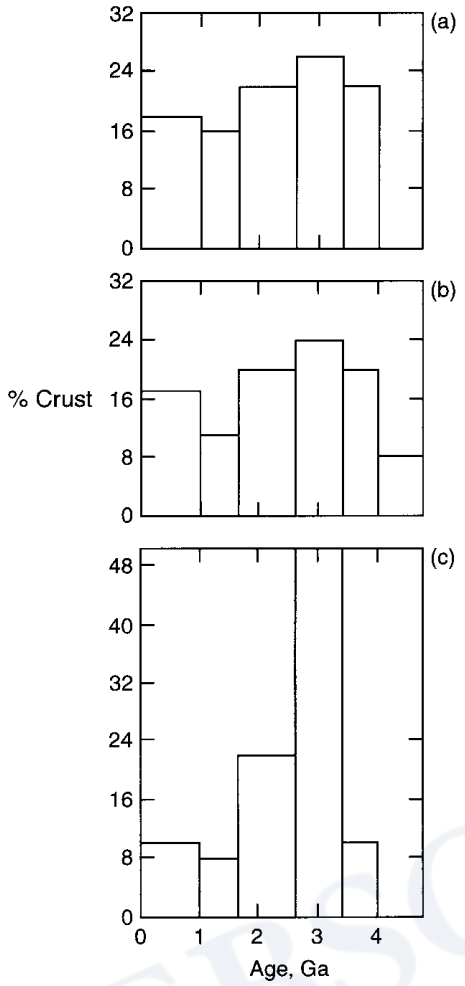


Figure 13.16. Age distribution of the crust based on Nd model ages (a, b, and c) and the freeboard constraint (c). Models (a) and (c) assume no crust at 4 Ga, model (b) assumes 8% crust at 4 Ga. Data are from Taylor and McLennan (1995). The episodic nature of crustal growth is evident particularly in model (c) which shows a strong spurt of crustal growth in the late Archean.

growth and their regulation by mantle convection. Formation of continental crust and its geochemical consequences are discussed in Sections 2.7 and 2.10 and Chapter 12. Here we discuss some aspects of crustal growth and its relation to mantle convection and the Earth's thermal evolution.

One of the major questions about continental crust is whether its growth has been continuous or episodic; related to this is the question of whether the continental crustal volume has increased with time, decreased with time from a maximum volume reached earlier in the Earth's history, remained approximately constant with time following early rapid growth, or oscillated with time.

Question 13.11: *Has continental crustal growth occurred continuously or episodically?*

Question 13.12: *How has the volume of the continents changed through geologic time?*

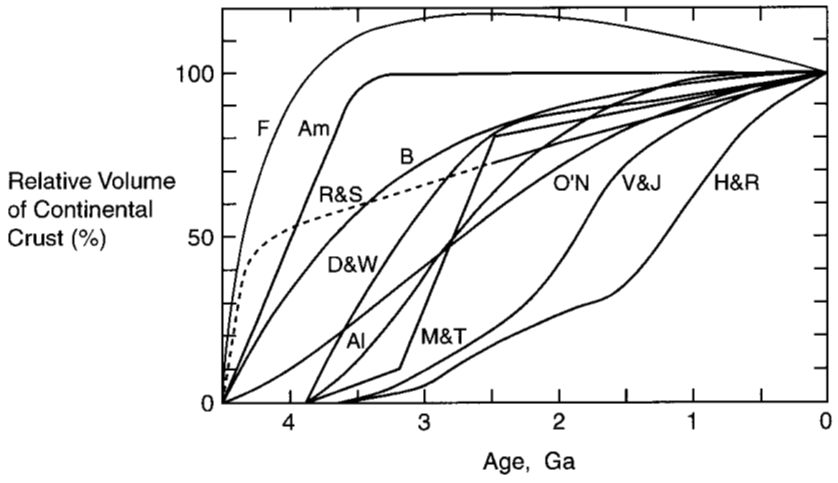


Figure 13.17. Crustal growth curves versus age (after Reymer and Schubert, 1984). The relative crustal volume is with respect to the current volume of continental crust. V & J, Veizer and Jansen (1979); M & T, McLennan and Taylor (1982); H & R, Hurley and Rand (1969); Al, Allègre (1982); O'N, O'Nions et al. (1979); R & S, Reymer and Schubert (1984); D & W, Dewey and Windley (1981); B, Brown (1979); Am, Armstrong (1981); F, Fyfe (1978).

Continental growth (the net gain in the volume or mass of the continental crust) is the net result of processes which both add and subtract material from the continents (Section 2.7). While continental growth is controlled by mantle convection, it is not clear whether enhanced mantle convective activity results in a net gain or loss of continental crust since both the addition processes (island arc and hot spot magmatic activity) and subtraction processes (sediment subduction and delamination) are more active with increased convective vigor. Figure 13.17 shows the variety of proposed continental crustal growth curves. The crustal growth curve of Fyfe (1978) shows rapid crustal growth early in the Earth's history and a decline in crustal volume since 2.5 Ga. The crustal growth curve of Armstrong (1981) shows a constant crustal volume for the last 3.5 Gyr. Reymer and Schubert (1984) proposed early accumulation of about 50% of the continental crust (within a few hundred million years of the end of accretion) followed by more gradual crustal growth over most of geologic history to the present. Many of the proposed crustal growth curves do not feature early rapid crustal growth and some delay crustal growth until 3.8 Ga. While considerable crustal differentiation probably occurred in the early, hot, vigorously convecting Earth, early crustal growth is not assured because the survivability of that crust is an open question.

Note that none of the curves shown in Figure 13.17 show episodic crustal growth. However, there is isotopic evidence that the continental crust has accumulated in spurts of activity (Moorbath, 1978; Patchett et al., 1981; Page et al., 1984; Nelson and DePaolo, 1985; DePaolo et al., 1991; Taylor and McLennan, 1995; Sylvester et al., 1997). Episodic crustal growth is suggested by the crustal age distribution models of Figure 13.16, especially model (c) which has a large spurt of crustal growth in the late Archean. The apparent episodicity in crustal growth suggested by isotopic data could be a consequence of inadequate and incomplete sampling of the crust. On the other hand, the discussion of the previous section emphasized that mantle convection can result in episodic crustal growth through its inherent chaotic behavior and through associated processes such as mantle avalanches (Stein and Hofmann, 1994). O'Nions and Tolstikhin (1996) have argued that ^{36}Ar and ^{40}Ar abundances in the

atmosphere together with estimates of the He and Ar isotopic composition of the upper and lower mantle and estimates of U, K, and Th in the mantle severely limit mass exchange between the upper mantle and lower mantle and the possibility that mantle avalanches could be responsible for episodic crustal growth.

The crustal growth curve derived by Reymer and Schubert (1984) is based on the secular cooling of the Earth and the assumption of constancy of freeboard since the Archean. Freeboard is defined as the average height of the continents above sea level. There is geological evidence that freeboard has been approximately constant since the end of the Archean at 2.5 Ga (Ambrose, 1964; Wise, 1974; Windley, 1977). The constancy of freeboard has been used to argue that no growth of the continents occurred during the Proterozoic and Phanerozoic (Armstrong, 1968, 1981, 1991). However, the secular decline in the heat production of the mantle causes the ocean basins to deepen and the volume of the oceans to increase with time; accordingly, growth of the continental crust is necessary to maintain constant freeboard (Reymer and Schubert, 1984; Schubert and Reymer, 1985; Galer, 1991). The growth of the continents can be quantified by utilizing the principle of isostasy and the relation between the depth of the ocean basins and surface heat flow (Schubert and Reymer, 1985); the resulting crustal growth curve is labeled R & S in Figure 13.17 (see also the discussion in 13.2.6). The constancy of freeboard constraint applies strictly only in the Phanerozoic and Proterozoic; its extrapolation into the Archean is speculative.

Gurnis and Davies (1985) combined the present distribution of crustal ages and crustal growth curves with a parameterized convection Earth thermal history model. They assumed that crustal production is related to the Earth's heat flux through plate velocities and that crustal removal (recycling into the mantle) is related to both heat flux and existing crustal volume. Their work emphasizes the sensitivity of the present crustal age distribution to the nature and variability of crustal production and recycling processes. For example, they show how a late-Archean peak in the crustal age distribution may have resulted from preferential removal of younger crust rather than a peak in production.

Another major question concerning the formation of the continental crust is whether crustal generation in the Archean involved processes similar to those at present. At present, continental crust is produced in island arcs and at hot spots with the former process dominant. However, the relative importance of these two processes may have changed since the Archean. The evolution of Archean and some younger terrains, such as the Arabian–Nubian Shield, may have involved a large amount of hot spot type addition in order to explain the very rapid addition rates that prevailed in these areas (Reymer and Schubert, 1984). Puchtel et al. (1998) propose continental growth by accretion of an oceanic plateau in the Archean.

Differences in the composition of Archean and post-Archean crustal rocks suggest different crustal production mechanisms (Rudnick, 1995; Taylor and McLennan, 1995). Archean tonalities and trondhjemites may have resulted from slab or mantle wedge melting at higher temperatures and lower pressures than occur in the present mantle. At the higher temperatures that prevailed in the Archean, there would have been more rapid subduction of young hot slabs than occurs at present and slabs could have melted before undergoing complete dehydration (Martin, 1986; Defant and Drummond, 1990, 1993; Drummond and Defant, 1990; Abbott et al., 1994; Taylor and McLennan, 1995). At present, subducted old oceanic crust dehydrates, driving fluids into the overlying mantle wedge. These fluids induce further melting of the wedge in a poorly understood process (see Section 2.7.3). The resulting basic magmas pond in the continental crust and generate silicic granitic and andesitic magmas as secondary melts.

Crustal subtraction processes may also have been different in the Archean. The relative importance of sediment subduction and delamination in returning crustal material to the mantle may have been different in the Archean compared to the present. Rudnick (1995) has suggested that delamination of the lower crust may have been an important recycling process in the Archean. Figure 13.18 illustrates the possible differences between Archean

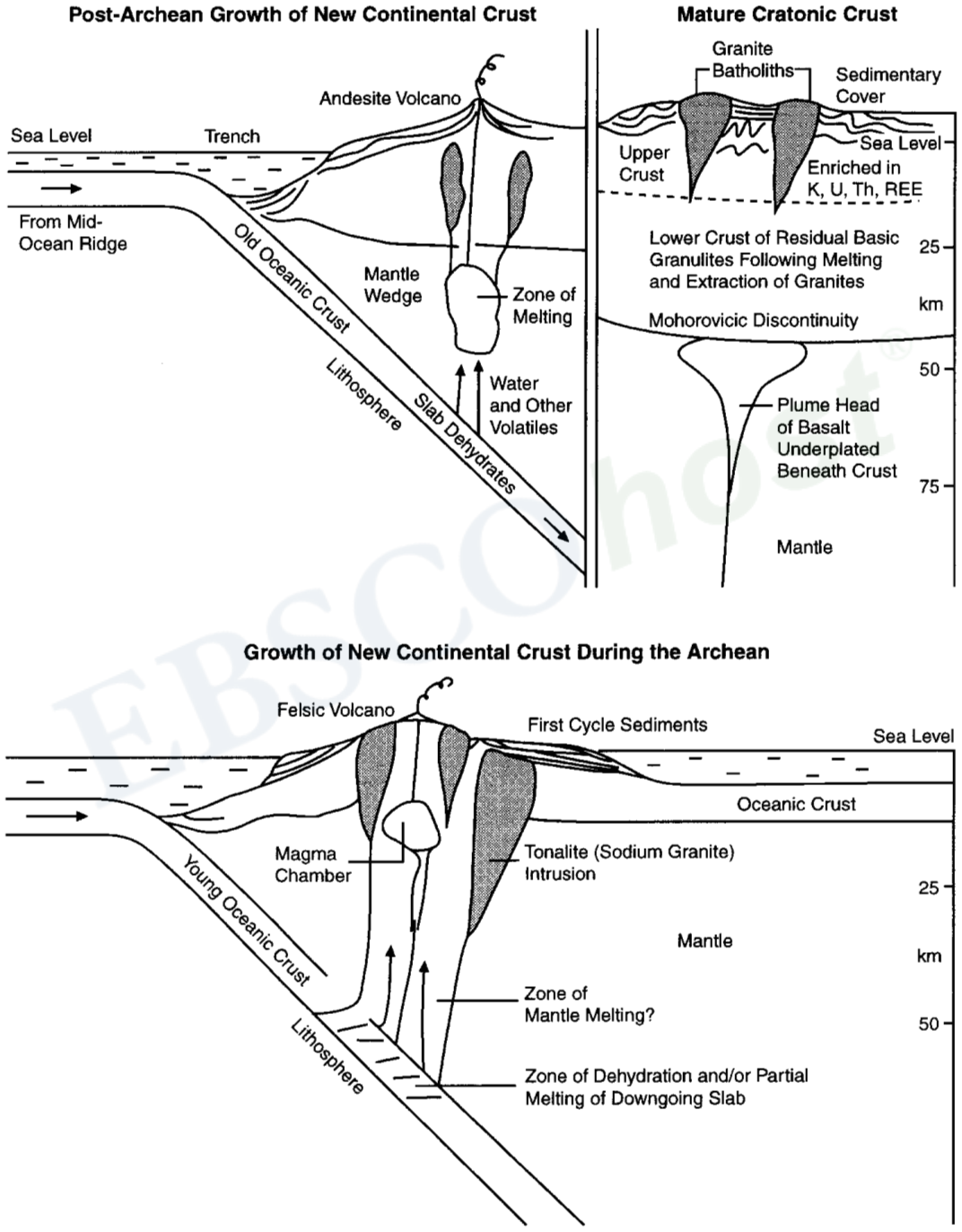


Figure 13.18. Sketch of possible differences between subduction-related crustal growth processes in the Archean and post-Archean (after Taylor and McLennan, 1995).

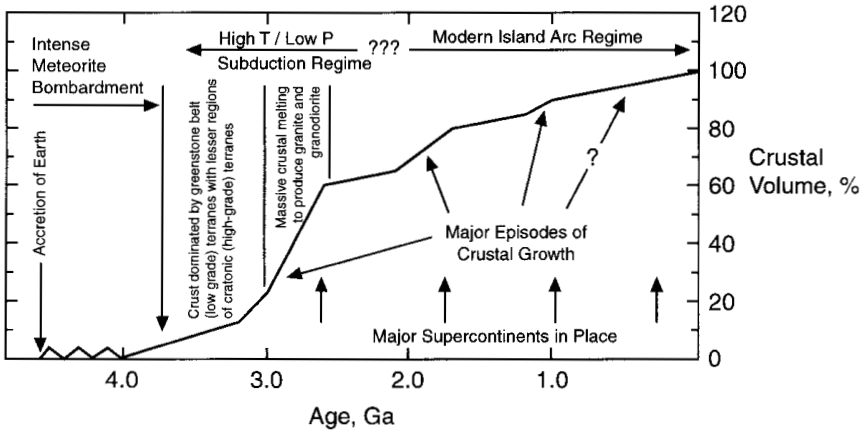


Figure 13.19. Sketch of possible crustal growth through geologic time with periods of enhanced growth indicated. The sketch suggests a possible connection between crustal growth episodes and assembly phases of supercontinents (after Taylor and McLennan, 1995).

and post-Archean crustal formation and Figure 13.19 summarizes the ideas discussed above regarding crustal growth and episodicity over geologic time.

Question 13.13: *What processes acted in the Archean to produce and recycle continental crust?*