7.02 Physics of Mantle Convection

Y Ricard, Ecole Normale Supérieure de Lyon, Université de Lyon-1, CNRS, Lyon, France

7.02.1 Introduction

In many textbooks of fluid dynamics, and for most students, the word 'fluid' refers to one of the states of matter, either liquid or gaseous, in contrast to the solid state. This definition is much too restrictive. In fact, the definition of a fluid rests in its tendency to deform irrecoverably. Basically, any material that appears as elastic or nondeformable, with a crystalline structure (i.e., belonging to the solid state) or with a disordered structure (e.g., a glass, which from a thermodynamic point of view belongs to the liquid state), can be deformed when subjected to stresses for a long enough time.

The characteristic time constant of the geologic processes related to mantle convection, typically 10 My (3×10^{14} s), is so long that the mantle, although stronger than steel and able to transmit seismic shear waves, can be treated as a fluid. Similarly, ice, which is the solid form of water, is able to flow from mountain tops to valleys in the form of glaciers. A formalism that was developed for ordinary liquids or gases can therefore be used in order to study the inside of planets. It is not the equations themselves, but their parameters (viscosity, conductivity, spatial dimensions, etc.) that characterize their applicability to mantle dynamics.

Most materials can therefore behave like elastic solids on very short time constants and like liquids at long times. The characteristic time that controls the appropriate rheological behavior is the ratio between dynamic viscosity, η, and elasticity (shear modulus), μ_r, called the Maxwell time τ_M (Maxwell, 1831–79),

$$\tau_M = \frac{\eta}{\mu_r} \quad [1]$$

The rheological transition in some materials like silicon putty occurs in only a few minutes; a silicon ball can bounce on the floor, but it turns into a puddle when left on a table for tens of minutes. The transition time is of the order of a few hundred to a few thousand years for the mantle (see Section 7.02.3.2). Phenomena of a shorter duration than this time will experience the mantle as an elastic solid, while most tectonic processes will experience the mantle as an irreversibly deformable fluid. Surface loading of the Earth by glaciation and deglaciation involves times of a few thousand years for which elastic aspects cannot be totally neglected with respect to viscous aspects.

Although the word 'convection' is often reserved for flows driven by internal buoyancy anomalies of thermal origin, in this chapter, we will more generally use 'convection' for any motion of a fluid driven by internal or external forcing. Convection can be kinematically forced by boundary conditions or induced by density variations. The former is 'forced convection' and the latter is 'free convection,' which can be of compositional or thermal origin. We will, however, mostly focus

on the aspects of thermal convection (or Rayleigh–Bénard convection (Rayleigh, 1842–1919; Bénard, 1874–1939)) when the fluid motion is driven by thermal anomalies and discuss several common approximations that are made in this case. We know, however, that many aspects of mantle convection can be more complex and involve compositional and petrologic density anomalies or multiphase physics. We will therefore review some of these complexities.

The physics of fluid behavior, like the physics of elastic media, is based on the general continuum hypothesis. This hypothesis requires that quantities like density, temperature, and velocity are defined everywhere, continuously, and at 'points' or infinitesimal volumes that have a statistically meaningful number of molecules so that the quantity represents an average independent of microscopic molecular fluctuations. This hypothesis seems natural for ordinary fluids at the laboratory scale. We will adopt the same hypothesis for the mantle, although we know that it is heterogeneous at various scales and made of compositionally distinct mineral grains.

7.02.2 Conservation Equations

The basic equations of this section can be found in more detail in many classical textbooks (Batchelor, 1967; Landau and Lifchitz, 1980). We will only emphasize the aspects that are pertinent to Earth's and terrestrial mantles.

7.02.2.1 General Expression of Conservation Equations

Let us consider a fluid transported by the velocity field $\boldsymbol{\nu}$, a function of position $\mathbf{x}$ and time t. There are two classical approaches to describe the physics in this deformable medium. Any variable A in a flow (like a scalar such as temperature or a vector like velocity) can be considered as a simple function of position and time, $A(\mathbf{x}, t)$, in a way very similar to the specification of an electromagnetic field. This description is the Eulerian point of view (Euler, 1707–83). The second point of view is traditionally attributed to Lagrange (Lagrange, 1736–1813) and considers that a material element of the flow is identified by its unique initial position $\mathbf{x}_0$ and thereafter follows a trajectory specific to that element $\mathbf{x}(\mathbf{x}_0, t)$. An observer following this trajectory would naturally choose the variable $A(\mathbf{x}(\mathbf{x}_0, t), t)$. The same variable A seen by an Eulerian or a Lagrangian observer would have very different time derivatives. According to Euler, the time derivative would simply be the rate of change seen by an observer at a fixed position, that is, the partial derivative $\partial/\partial t$. According to Lagrange, the time derivative, noted with D, would be the rate of change seen by an observer riding on the material element

$$\frac{\mathrm{D}A}{\mathrm{D}t} = \left(\frac{\mathrm{d}A(\mathbf{x}(\mathbf{x}_0,\ t),t)}{\mathrm{d}t}\right)_{x_0} = \sum_{i=1,3} \frac{\partial A}{\partial x_i}\frac{\partial x_i}{\partial t} + \frac{\partial A}{\partial t} \qquad [2]$$

where x_i are the components of $\mathbf{x}$. Since $\mathbf{x}$ is the position of a material element of the flow, its partial time derivative is simply the flow velocity $\mathbf{v}$. The Lagrangian derivative is also sometimes called material derivative, total derivative, or substantial derivative.

The previous relation was written for a scalar field A, but it could easily be applied to a vector field $\mathbf{A}$. The Eulerian and Lagrangian time derivatives are thus related by the symbolic relation

$$\frac{\mathrm{D}}{\mathrm{D}t} = \frac{\partial}{\partial t} + (\mathbf{v}\cdot\boldsymbol{\nabla}) \qquad [3]$$

In Cartesian coordinates (Descartes, 1596–1650), the operator $(\mathbf{v}\cdot\boldsymbol{\nabla})$ is the symbolic scalar $(v_1\partial/\partial x_1 + v_2\partial/\partial x_2 + v_3\partial/\partial x_3)$ and a convenient mnemonic is to interpret it as the scalar product of the velocity field with the gradient operator $(\partial/\partial x_1, \partial/\partial x_2, \partial/\partial x_3)$ denoted with the symbol $\boldsymbol{\nabla}$ called nabla. The operator $(\mathbf{v}\cdot\boldsymbol{\nabla})$ can be applied to a scalar or a vector. Notice however the limitations of this mnemonic as for any vectors $\mathbf{A}$ and $\mathbf{B}$, $(\mathbf{v}\cdot\mathbf{B})\mathbf{A}$ is parallel to $\mathbf{A}$, while $(\mathbf{v}\cdot\boldsymbol{\nabla})\mathbf{A}$ is parallel neither to $\mathbf{A}$ nor to $\mathbf{v}$.

In a purely homogeneous fluid, the flow lines are not visible and the mechanical properties are independent of the original position of fluid particles. In this case, the Eulerian perspective seems natural. On the other hand, a physicist describing elastic media can easily draw marks on the surface of deformable objects and flow trajectories become perceptible for him or her. After elastic deformation, the stresses are also dependent on the initial equilibrium state. The Lagrangian perspective is therefore more appropriate. We will mostly adopt the Eulerian perspective for the description of the mantle. However, when we discuss deformation of heterogeneities embedded in and stirred by the convective mantle, the Lagrangian point of view will be more meaningful (see Section 7.02.5.1.7).

A starting point for describing the physics of a continuum is the conservation equations. Consider a scalar- or a vector-extensive variable A (i.e., mass, momentum, energy, entropy, and number of moles) with a density per unit volume ρ_A (with units of A per unit volume) and a virtual but fixed volume Ω enclosed by the open surface $\sum$. This virtual volume is freely crossed by the flow. The temporal change of the net quantity of A inside Ω is

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \rho_A\ \mathrm{d}V = \int_\Omega \frac{\partial \rho_A}{\partial t}\ \mathrm{d}V \qquad [4]$$

Since Ω is fixed, the derivative of the integral is the integral of the partial time derivative.

The total quantity of the extensive variable A in a volume Ω can be related to a local production, H_A (with units of A per unit volume and unit time), and to the transport (influx or efflux) of A across the interface. This transport can be either a macroscopic advective transport by the flow or a more indirect transport, for example, microscopic diffusion. Let us call $\mathbf{J}_A$ the total flux of A per unit surface area. The conservation of A can be expressed in integral form as

$$\int_\Omega \frac{\partial \rho_A}{\partial t}\ \mathrm{d}V = -\int_\Sigma \mathbf{J}_A\cdot\ \mathrm{d}\mathbf{S} + \int_\Omega H_A\ \mathrm{d}V \qquad [5]$$

where the infinitesimal surface element vector $\mathrm{d}\mathbf{S}$ is oriented with the outward unit normal; hence, the minus sign associates outward flux with a sink of quantity A. Equation [5] is the general form of any conservation equation. When the volume Ω, surface $\sum$, and flux $\mathbf{J}_A$ are regular enough (in mathematical terms when the volume is compact, the surface piecewise

smooth, and the flux continuously differentiable), we can make use of the divergence theorem

$$\int_{\Sigma} \mathbf{J}_A \cdot d\mathbf{S} = \int_{\Omega} \nabla \cdot \mathbf{J}_A \, dV \qquad [6]$$

to transform the surface integral into a volume integral. The divergence of the vector $\mathbf{J}$ with Cartesian coordinates (J_1, J_2, J_3) is the scalar $\nabla \cdot \mathbf{J} = \partial J_1/\partial x_1 + \partial J_2/\partial x_2 + \partial J_3/\partial x_3$, scalar product of the symbolic operator ∇ by the real vector $\mathbf{J}$ (although notice the difference between the scalar $\nabla \cdot \mathbf{J}$ and the operator $\mathbf{J} \cdot \nabla$). Since the integral equation [5] is valid for any virtual volume Ω, we can deduce that the general differential form of the conservation equation is

$$\frac{\partial \rho_A}{\partial t} + \nabla \cdot \mathbf{J}_A = H_A \qquad [7]$$

A similar expression can be used for a vector quantity $\boldsymbol{\rho}_A$ with a tensor flux $\underline{\mathbf{J}}_A$ and a vector source term $\mathbf{H}_A$. In this case, the divergence of the second-order tensor with components J_{ij} is a vector with Cartesian components $\Sigma_{j=1,3} \partial J_{ij}/\partial x_j$ (our convention is that the derivatives in the divergence affect the second index of the tensor; other authors (e.g., Malvern, 1969) use a different convention).

We now apply this formalism to various physical quantities. Three quantities are strictly conserved: the mass, the momentum, and the energy. This means that they can only change in a volume Ω by influx or efflux across the surface $\sum$. We must identify the corresponding fluxes, but no source terms should be present (in fact, in classical mechanics, the radioactivity appears as a source of energy) (see also Section 7.02.2.5.3). One very important physical quantity is not conserved – the entropy – but the second law of thermodynamics ensures the positivity of the associated sources.

7.02.2.2 Mass Conservation

Mass flux of material with density ρ (mass per unit volume) is

$$\mathbf{J}_\rho = \rho \mathbf{v} \qquad [8]$$

Using either the Eulerian or the Lagrangian time derivatives, mass conservation becomes

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \qquad [9]$$

or

$$\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} = 0 \qquad [10]$$

In an incompressible fluid, particles have constant density, and so in the particle frame of reference, the Lagrangian observer does not see any density variation and $D\rho/Dt = 0$. In this case, mass conservation takes the simple form $\nabla \cdot \mathrm{v} = 0$, which is commonly called the continuity equation.

Using mass conservation, a few identities can be derived that are very useful for transforming an equation of conservation for a quantity per unit mass to a quantity per unit volume. For example, for any specific scalar quantity A (i.e., quantity per unit mass),

$$\frac{\partial (\rho A)}{\partial t} + \nabla \cdot (\rho A \mathbf{v}) = \rho \frac{DA}{Dt} \qquad [11]$$

and for any specific vector field $\mathbf{A}$,

$$\frac{\partial (\rho \mathbf{A})}{\partial t} + \nabla \cdot (\rho \mathbf{A} \otimes V) = \rho \frac{D\mathbf{A}}{Dt} \qquad [12]$$

where $\mathbf{A} \otimes \mathbf{v}$ is a dyadic tensor of components $A_i v_j$ (notice that $\mathbf{v} \otimes \mathbf{A} \neq \mathbf{A} \otimes \mathbf{v}$).

7.02.2.3 Momentum Conservation

7.02.2.3.1 General momentum conservation

The rate of change of momentum of material in a volume Ω is related to the body forces acting in the volume Ω and to the surface forces acting on its surface $\sum$, that is, Newton's second law (Newton, 1642–1727). The total momentum is

$$\int_{\Omega} \rho \mathbf{v} \, dV \qquad [13]$$

and its variations are due to

- advective transport of momentum across the surface $\sum$,
- forces acting on this surface,
- internal body forces.

The momentum conservation of an open, fixed volume can therefore be expressed in integral form as

$$\int_{\Omega} \frac{\partial (\rho \mathbf{v})}{\partial t} dV = -\int_{\Sigma} \rho \mathbf{v} (\mathbf{v} \cdot d\mathbf{S}) + \int_{\Sigma} \underline{\boldsymbol{\sigma}} \cdot d\mathbf{S} + \int_{\Omega} \mathbf{f} \cdot dV \qquad [14]$$

The stress tensor $\underline{\boldsymbol{\sigma}}$ corresponds to the traction applied on the surface, per unit area (see Figure 1). Our convention is that σ_{ij} is the i-component of the force per unit area across a plane normal to the j-direction. The term $\mathbf{f}$ represents the sum of all body forces per unit volume and in particular the gravitational forces $\rho \mathbf{g}$ (we will not consider electromagnetic forces that are described in vol. 8).

Using the divergence theorem (for the first term of the right side, $\rho \mathbf{v} (\mathbf{v} \cdot d\mathbf{S})$ can also be written as $\rho (\mathbf{v} \otimes \mathbf{v}) \cdot d\mathbf{S}$) and the equality [12], the differential form of momentum conservation becomes

$$\rho \frac{D\mathbf{v}}{Dt} = \nabla \cdot \underline{\boldsymbol{\sigma}} + \mathbf{f} \qquad [15]$$

It is common to divide the total stress tensor into a thermodynamic pressure – P$\underline{\mathbf{I}}$ where $\underline{\mathbf{I}}$ is the identity stress tensor and a

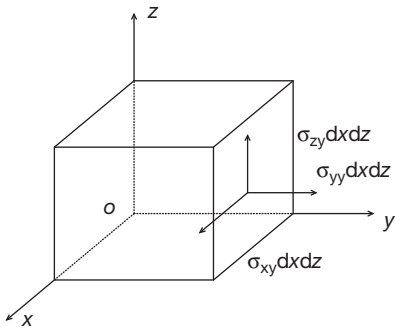

Figure 1 The force per unit area applied on a surface directed by the normal vector $\mathbf{n}$ is by definition $\underline{\boldsymbol{\sigma}} \cdot \mathbf{n}$. The component of this force along the unit vector $\mathbf{e}_i$, therefore $\mathbf{e}_i \cdot \underline{\boldsymbol{\sigma}} \cdot \mathbf{n}$. For example, $\sigma_{zy} dx dz$ is the component acting in the oz direction of force applied on the surface $dxdz$ with normal in the oy direction. Here again, the order of the indexes for the stress tensor is a convention but in agreement with our definition of the divergence of a tensor.

velocity-dependent stress $\underline{\boldsymbol{\tau}}$. The relationship between the tensor $\underline{\boldsymbol{\tau}}$ and the velocity field will be discussed in Section 7.02.3.2. Without motion, the total stress tensor is thus isotropic and equal to the usual pressure. In most geophysical literature, it has been assumed that the velocity-dependent tensor has no isotropic component, that is, it is traceless $\mathrm{tr}(\underline{\boldsymbol{\tau}}) = 0$. In this case, the average isotropic total stress $-\mathrm{tr}(\underline{\boldsymbol{\sigma}})/3$ is the thermodynamic pressure P and the velocity-dependent stress tensor $\underline{\boldsymbol{\tau}}$ is the deviatoric stress tensor (see Section 7.02.3.2 for more details).

As $\boldsymbol{\nabla}\cdot(P\underline{\mathbf{I}}) = \boldsymbol{\nabla}P$, momentum conservation [14], in terms of pressure and deviatoric stresses, is

$$\rho\frac{\mathrm{D}\mathbf{v}}{\mathrm{D}t} = -\boldsymbol{\nabla}P + \boldsymbol{\nabla}\cdot\underline{\boldsymbol{\tau}} + \mathbf{f} \qquad [16]$$

This equation is called the Navier–Stokes equation (Navier, 1785–1836; Stokes, 1819–1903) when the stress tensor is linearly related to the strain rate tensor and the fluid is incompressible (see Section 7.02.3.2).

7.02.2.3.2 Inertia and non-Galilean forces

In almost all studies of mantle dynamics, the fact that the Earth is rotating is simply neglected. It is however worth discussing this point. Let us define a reference frame of vectors $\mathbf{e}_i$ attached to the solid Earth. These vectors rotate with the Earth, such that with respect to a Galilean frame, their rate of change is

$$\frac{\mathrm{d}\mathbf{e}_i}{\mathrm{d}t} = \boldsymbol{\omega}\times\mathbf{e}_i \qquad [17]$$

where $\boldsymbol{\omega}$ is the angular velocity of Earth's rotation. A point of the Earth, $\mathbf{x} = \sum_i x_i\mathbf{e}_i$, has a velocity in the Galilean frame

$$\left[\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\right]_{\mathrm{Gal.}} = \sum_{i=1,3}\left[\left(\frac{\mathrm{d}x_i}{\mathrm{d}t}\right)\mathbf{e}_i + x_i\left(\frac{\mathrm{d}\mathbf{e}_i}{\mathrm{d}t}\right)\right] = \mathbf{v}_{\mathrm{Earth}} + \boldsymbol{\omega}\times\mathbf{x} \qquad [18]$$

where $\mathbf{v}_{\mathrm{Earth}}$ is the velocity in the Earth's frame, and, by repeating the Galilean time derivate, to obtain the acceleration $\boldsymbol{\gamma}_{\mathrm{Gal}}$ in a Galilean frame (Galileo, 1564–1642),

$$\boldsymbol{\gamma}_{\mathrm{Gal}} = \boldsymbol{\gamma}_{\mathrm{Earth}} + 2\boldsymbol{\omega}\times\mathbf{v}_{\mathrm{Earth}} + \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{x}) + \frac{\mathrm{d}\boldsymbol{\omega}}{\mathrm{d}t}\times\mathbf{x} \qquad [19]$$

In this well-known expression, one recognizes on the right side the acceleration in the non-Galilean Earth's reference frame and the Coriolis (Coriolis, 1792–1843), centrifugal, and Poincaré accelerations (Poincaré, 1854–1912).

To quantify the importance of the first three acceleration terms (neglecting the Poincaré term), let us consider a characteristic length scale (the Earth's radius, $a = 6371$ km) and mantle velocity (the maximum plate tectonic speed, $U = 10\ \mathrm{cm\ year^{-1}}$) with which we can compare the various acceleration terms. One immediately gets

$$\frac{\text{Inertia}}{\text{Coriolis}} = \frac{U}{2\omega a} = \frac{1}{2.9\times10^{11}} \qquad [20]$$

$$\frac{\text{Coriolis}}{\text{Gravitational force}} = \frac{2\omega U}{g} = \frac{1}{2.1\times10^{13}} \qquad [21]$$

$$\frac{\text{Centrifugal}}{\text{Gravitational force}} = \frac{\omega^2 a}{g} = \frac{1}{291} \qquad [22]$$

Thus, the inertial term is much smaller than the Coriolis term (this ratio is also known as the Rossby number (Rossby, 1898–1957)), which is itself negligible relative to gravitational force. Even if we argue that a more meaningful comparison would be between the whole Coriolis force $2\rho\omega U$ and the lateral variations of the gravitational force $\delta\rho g$ rather than the total gravity force ρg, inertia and Coriolis accelerations still play a negligible role in mantle dynamics (assuming that the mantle velocity U is related to $\delta\rho g$ by a viscous relation $U \propto \delta\rho g a^2/\eta$ where η is the mantle viscosity, the ratio $2\rho\omega U/(\delta\rho g)$ is the inverse of the Ekman number (Ekman, 1874–1954)). Neglecting inertia means that forces are instantaneously in balance and that changes in kinetic energy are negligible since inertia is the time derivative of the kinetic energy. We can perform a simple numerical estimate of the mantle kinetic energy. The kinetic energy of a lithospheric plate (a square of size 2000 km, thickness 100 km, velocity 5 cm $\mathrm{year^{-1}}$, and density 3000 kg $\mathrm{m^{-3}}$) is 1.51 kJ, which is comparable to that of a medium-sized car (2000 kg) driven at only 4.42 km $\mathrm{h^{-1}}$ or of a man on a bike (say, total about 100 kg) going at a leisurely 20 km $\mathrm{h^{-1}}$.

The centrifugal term is also quite small but not as small (i.e., it is 1/291 of gravitational force). It induces several effects. The centrifugal effect causes the Earth's flattening with an equatorial bulge of 21 km (1/300 of the Earth's radius), although this is a static phenomenon that has no interactions with convective dynamics. The gravitational coupling between the equatorial bulge and other celestial bodies perturbs their Keplerian orbits (Kepler, 1571–1630) and leads to precessions and nutations of their orbits as well as Earth's spin. Finally, the bulge causes the whole planet to rotate along an equatorial axis in order to keep the maximum principal inertial axis coincident with its rotational axis (Ricard et al., 1993b; Spada et al., 1992a). This rotational equilibrium of the Earth will not be discussed here (see e.g., Chandrasekhar (1969) for the static equilibrium shape of a rotating planet and Munk and MacDonald (1960) for the dynamics of a deformable rotating body). The Poincaré acceleration is itself negligible compared to the centrifugal term as the changes of Earth's rotation occur over times much larger than a day. Although Poincaré acceleration has no effect on the mantle, it may affect the dynamics of the fluid core (Vanyo et al., 1995).

We neglect all the acceleration terms in the following, but we should remember that in addition to the convective motion of a nonrotating planet, a rotation of the planet with respect to an equatorial axis is possible. This motion documented by paleomagnetism is called true polar wander (Besse and Courtillot, 1991).

7.02.2.3.3 Angular momentum conservation

The angular momentum per unit mass $\mathbf{J} = \mathbf{x}\times\mathbf{v}$ also obeys a conservation law. This law can be obtained in two different ways. First, as we did for mass and momentum conservation, we can express the balance of angular momentum in integral form. In the absence of intrinsic body couples (i.e., we do not consider the case where distributed body couples exist in addition of distributed body forces), the angular momentum variations are due to

- advective transport of angular momentum across the surface $\sum$,
- torque of forces acting on this surface,
- torque of internal body forces.

The resulting balance is therefore

$$\int_\Omega \frac{\partial(\rho\mathbf{J})}{\partial t}\,\mathrm{d}V = -\int_\Sigma \rho\mathbf{J}(\mathbf{v}\cdot\mathrm{d}\mathbf{S}) + \int_\Sigma \mathbf{x}\times(\underline{\boldsymbol{\sigma}}\cdot\mathrm{d}\mathbf{S}) + \int_\Omega \mathbf{x}\times\mathbf{f}\;\mathrm{d}V \quad [23]$$

The only difficulty to transform this integral form into a local equation is with the integral involving the stress tensor (but is readily solved using tensor or index notation; see Malvern, 1969). After some algebra, eqn [23] becomes

$$\rho\frac{\mathrm{D}\mathbf{J}}{\mathrm{D}t} = \mathbf{x}\times\nabla\cdot\underline{\boldsymbol{\sigma}} + \mathbf{x}\times\mathbf{f} + \mathcal{T} \quad [24]$$

where the torque $\mathcal{T}$ is the vector $(\mathcal{T}_{zy}-\mathcal{T}_{yz}, \mathcal{T}_{xz}-\mathcal{T}_{zx}, \mathcal{T}_{yx}-\mathcal{T}_{xy})$. A second expression can be obtained by the vectorial multiplication of the momentum equation [15] by $\mathbf{x}$. Since

$$\mathbf{x}\times\frac{\mathrm{D}\mathbf{v}}{\mathrm{D}t} = \frac{\mathrm{D}\mathbf{J}}{\mathrm{D}t} - \frac{\partial\mathbf{x}}{\partial t}\times\mathbf{v} = \frac{\mathrm{D}\mathbf{J}}{\mathrm{D}t} - \mathbf{v}\times\mathbf{v} = \frac{\mathrm{D}\mathbf{J}}{\mathrm{D}t} \quad [25]$$

we get,

$$\rho\frac{\mathrm{D}\mathbf{J}}{\mathrm{D}t} = \mathbf{x}\times\nabla\cdot\underline{\boldsymbol{\sigma}} + \mathbf{x}\times\mathbf{f} \quad [26]$$

which differs from eqn [24] by the absence of the torque $\mathcal{T}$. This proves that in the absence of singular sources of angular momentum (distributed body couples), the stress (either $\underline{\boldsymbol{\sigma}}$ or $\underline{\boldsymbol{\tau}}$) must be represented by a symmetrical tensor,

$$\underline{\boldsymbol{\sigma}} = \underline{\boldsymbol{\sigma}}^t, \quad \underline{\boldsymbol{\tau}} = \underline{\boldsymbol{\tau}}^t \quad [27]$$

where $[\;]^t$ denotes tensor transposition.

7.02.2.4 Energy Conservation

7.02.2.4.1 First law and internal energy

The total energy per unit mass of a fluid is the sum of its internal energy, $\mathcal{U}$, and its kinetic energy (this approach implies that the work of the various forces is separately taken into account; another approach that we use in Section 7.02.2.5.3 adds to the total energy the various possible potential energies and ignores forces). In the fixed volume Ω, the total energy is thus

$$\int_\Omega \rho\left(\mathcal{U} + \frac{v^2}{2}\right)\mathrm{d}V \quad [28]$$

A change of this energy content can be caused by

- advection of energy across the boundary $\sum$ by macroscopic flow;
- transfer of energy through the same surface without mass transport, by say diffusion or conduction;
- work of body forces;
- work of surface forces;
- volumetrically distributed radioactive heat production.

Using the divergence theorem, the balance of energy can therefore be written as

$$\frac{\partial}{\partial t}\left(\rho\left(\mathcal{U}+\frac{v^2}{2}\right)\right) = -\nabla\cdot\left(\rho\left(\mathcal{U}+\frac{v^2}{2}\right)\mathbf{v} + \mathbf{q} + P\mathbf{v} - \underline{\boldsymbol{\tau}}\cdot\mathbf{v}\right) + \mathbf{f}\cdot\mathbf{v} + \rho H \quad [29]$$

where $\mathbf{q}$ is the diffusive flux, H is the rate of energy production per unit mass, and the stresses are divided into thermodynamic pressure- and velocity-dependent stresses.

This expression can be developed and simplified by using eqn [11] and the equations of mass and momentum conservation [9] and [16] to reach the form

$$\rho\frac{\mathrm{D}\mathcal{U}}{\mathrm{D}t} = -\nabla\cdot\mathbf{q} - P\nabla\cdot\mathbf{v} + \underline{\boldsymbol{\tau}}:\nabla\mathbf{v} + \rho H \quad [30]$$

The viscous dissipation term $\underline{\boldsymbol{\tau}}:\nabla\mathbf{v}$ is the contraction of the two tensors $\underline{\boldsymbol{\tau}}$ and $\nabla\mathbf{v}$ (of components $\partial v_i/\partial x_j$). Its expression is $\Sigma_{ij}\tau_{ij}\partial v_i/\partial x_j$.

7.02.2.4.2 State variables

The internal energy can be expressed in terms of the more usual thermodynamic state variables, namely, temperature, pressure, and volume. We use volume to follow the classical thermodynamic approach, but since we apply thermodynamics to points in a continuous medium, the volume V is in fact the volume per unit mass or $1/\rho$. We use the first law of thermodynamics that states that during an infinitesimal process, the variation of internal energy is the sum of the heat δQ and work δW exchanged. Although irreversible processes occur in the fluid, we assume that we can adopt the hypothesis of local thermodynamic equilibrium.

The increments in heat and work are not exact differentials: these increments depend not only on the initial and final stages but also on the entire process of energy exchange. Using either a T–V or a T–P formulation, we can write

$$\delta Q = C_V\mathrm{d}T + l\mathrm{d}V = C_p\mathrm{d}T + h\mathrm{d}P \quad [31]$$

where C_P an C_V are the heat capacities per unit mass at constant pressure and volume, respectively, and h and l are two other calorimetric coefficients necessary to account for heat exchange at constant temperature. For fluids, the reversible exchange of work is only due to the work of pressure forces:

$$\delta W = -P\mathrm{d}V \quad [32]$$

which implies that only the pressure term corresponds to an energy that can be stored and recovered without loss when the change in volume is reversed. In contrast, the stresses related to the velocity will ultimately appear in the dissipative, irrecoverable term of viscous dissipation: this point will be further considered in the Section 7.02.3.2 about rheology.

Thermodynamics states that the total variations of energy, $\mathrm{d}\mathcal{U} = \delta Q + \delta W$; enthalpy, $\mathrm{d}\mathcal{H} = \mathrm{d}\mathcal{U} + \mathrm{d}(PV)$; or entropy, $\mathrm{d}\mathcal{S} = \delta Q/T$, are exact differentials and $\mathcal{U}, \mathcal{H}$, and $\mathcal{S}$ are potentials. This means that the net change in energy (or enthalpy or entropy) between an initial state and a final state depends only on the initial and final states themselves and not on the intermediate stages. This implies mathematically that the second partial derivatives of these potentials with respect to any pair of variables are independent of the order of differentiation. Using these rules, a large number of relations can be derived among the thermodynamic coefficients and their derivatives. These are called the Maxwell relations and are discussed in most thermodynamics textbooks (e.g., Poirier, 1991). We can in particular derive the values of l (starting from $\mathrm{d}\mathcal{U}$ and $\mathrm{d}\mathcal{S}$ in the T–V formulation) and h (starting from $\mathrm{d}\mathcal{H}$ and $\mathrm{d}\mathcal{S}$ in the T–P formulation),

$$l = \alpha T K_T, \quad \text{and} \quad h = -\frac{\alpha T}{\rho} \quad [33]$$

In these expressions for l and h, we introduced the thermal expansivity α and the isothermal incompressibility K_T,

$$\alpha = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_P = -\frac{1}{\rho}\left(\frac{\partial \rho}{\partial T}\right)_P$$
$$K_T = -V\left(\frac{\partial P}{\partial V}\right)_T = \rho\left(\frac{\partial P}{\partial \rho}\right)_T \quad [34]$$

The thermodynamic laws and differentials apply to a closed deformable volume $\Omega(t)$, which corresponds to the perspective of Lagrange. We can therefore interpret the differential symbols 'd' of the thermodynamic definitions [31] and [32] as Lagrangian derivatives 'D.'

Therefore, in total, when the expressions for l and h are taken into account, [33], and when the differential symbols are interpreted as Lagrangian derivatives, the change of internal energy, $d\mathcal{U} = \delta Q + \delta W$, can be recast as

$$\frac{D\mathcal{U}}{Dt} = C_V\frac{DT}{Dt} + (\alpha T K_T - P)\frac{\nabla\cdot\mathbf{v}}{\rho} \quad [35]$$

or

$$\frac{D\mathcal{U}}{Dt} = C_P\frac{DT}{Dt} - \frac{\alpha T}{\rho}\frac{DP}{Dt} - P\frac{\nabla\cdot\mathbf{v}}{\rho} \quad [36]$$

In these equations, we also have replaced the volume variation using mass conservation [9]

$$\frac{DV}{Dt} = \frac{D(1/\rho)}{Dt} = -\frac{1}{\rho^2}\frac{D\rho}{Dt} = \frac{\nabla\cdot\mathbf{v}}{\rho} \quad [37]$$

7.02.2.4.3 Temperature

We can now employ either thermodynamic relation [35] or [36] in our conservation equation deduced from fluid mechanics, [30], to express the conservation of energy in terms of temperature variations:

$$\rho C_P\frac{DT}{Dt} = -\nabla\cdot\mathrm{q} + \alpha T\frac{DP}{Dt} + \underline{\boldsymbol{\tau}}:\nabla\mathbf{v} + \rho H$$

$$\rho C_V\frac{DT}{Dt} = -\nabla\cdot\mathrm{q} - \alpha T K_T\nabla\cdot\mathbf{v} + \underline{\boldsymbol{\tau}}:\nabla\mathbf{v} + \rho H \quad [38]$$

Apart from diffusion, three sources of temperature variations appear on the right side of these equations. The last term ρH is the source of radioactive heat production. This term is of prime importance for the mantle, mostly heated by the decay of radioactive elements like uranium (^{235}U and ^{238}U), thorium (^{236}Th), and potassium (^{40}K). All together, these nuclides generate about 20×10^{12} W (McDonough and Sun, 1995). Although this number may seem large, it is in fact very small. Since the Earth now has about 7×10^9 inhabitants, the total natural radioactivity of the Earth is only $\approx$3 kW/person, not enough to run the appliances of a standard kitchen in a developed country. It is amazing that this ridiculously small energy source drives plate tectonics, raises mountains, and produces a magnetic field. In addition to the present-day radioactivity, extinct radionuclides, like that of ^{36}Al (with a half-life of 0.73 My), have played an important role in the initial stage of planet formation (Lee et al., 1976).

The viscous dissipation term $\underline{\boldsymbol{\tau}}:\nabla\mathbf{v}$ converts mechanical energy into heat. This term explains the classical Joule experiment (equivalence between work and heat, Joule, 1818–89) in which the potential energy of a load (measured in joules) drives a propeller in a fluid and dissipates the mechanical energy as thermal energy (measured in calories).

The remaining source term, containing the thermodynamic coefficients (α or αK_T in eqn [38]), vanishes when the fluid is incompressible (e.g., when $\alpha = 0$ or when $\nabla\cdot\mathbf{v} = 0$). This term is related to adiabatic compression and will be discussed in Section 7.02.4.2.2.

7.02.2.4.4 Second law and entropy

We now consider the second law of thermodynamics and entropy conservation. Assuming local thermodynamic equilibrium, we have $d\mathcal{U} = T d\mathcal{S} - P dV$. Using the equation of conservation for the internal energy $\mathcal{U}$, [30], and expressing the volume change in terms of velocity divergence, [37], we obtain

$$\rho T\frac{D\mathcal{S}}{Dt} = -\nabla\cdot\mathbf{q} + \underline{\boldsymbol{\tau}}:\nabla\mathbf{v} + \rho H \quad [39]$$

To identify the entropy sources, we can express this equation in the form of a conservation equation (see eqn [7]),

$$\frac{\partial(\rho\mathcal{S})}{\partial t} = -\nabla\cdot\left(\rho\mathcal{S}\mathbf{v} + \frac{\mathbf{q}}{T}\right) - \frac{1}{T^2}\mathbf{q}\cdot\nabla T + \frac{1}{T}\underline{\boldsymbol{\tau}}:\nabla\mathbf{v} + \frac{1}{T}\rho H \quad [40]$$

The physical meaning of this equation is therefore that the change of entropy is related to a flux of advected and diffused entropy, $\rho\mathcal{S}\mathbf{v}$ and $\mathbf{q}/T$, and to three entropy production terms, including from radiogenic heating.

A brief introduction to the general principles of non-equilibrium thermodynamics will be given in Section 7.02.5.1.4. Here, we simply state that the second law requires that in all situations, the total entropy production is positive. When different entropy production terms involve factors of different tensor orders (tensors, vectors, or scalars), they must separately be positive. This is called the Curie principle (Curie, 1859–1906) (see e.g., de Groot and Mazur, 1984; Woods, 1975). It implies that

$$-\mathbf{q}\cdot\nabla T \geq 0 \quad \text{and} \quad \underline{\boldsymbol{\tau}}:\nabla\mathbf{v} \geq 0 \quad [41]$$

The usual Fourier law (Fourier, 1768–1830) with a positive thermal conductivity $k > 0$,

$$\mathbf{q} = -k\nabla T \quad [42]$$

satisfies the second law.

When the conductivity k is uniform, the thermal diffusion term of the energy equation $-\nabla\cdot\mathbf{q}$ becomes $k\nabla^2 T$ where $\nabla^2 = \nabla\cdot\nabla$ is the scalar Laplacian operator (Laplace, 1749–1827). Instead of a thermal conductivity, a thermal diffusivity κ can be introduced:

$$\kappa = \frac{k}{\rho C_P} \quad [43]$$

(although in principle, isobaric and isochoric thermal diffusivities should be defined). In situations with uniform conductivity, without motion and radioactivity sources, the energy equation [38] becomes the standard diffusion equation

$$\frac{\partial T}{\partial t} = \kappa\nabla^2 T \quad [44]$$

The relation between stress and velocity will be discussed in detail in Section 7.02.3.2. We will show that the relationship

$$\underline{\boldsymbol{\tau}} = 2\eta\left(\underline{\dot{\boldsymbol{\varepsilon}}} - \frac{1}{3}\nabla\cdot\mathbf{v}\underline{\mathbf{I}}\right) \quad [45]$$

is appropriate for the mantle, where η is dynamic viscosity and strain rate tensor $\underline{\dot{\boldsymbol{\varepsilon}}}$ is defined by

$$\underline{\dot{\boldsymbol{\varepsilon}}} = \frac{1}{2}\left([\nabla \mathbf{v}] + [\nabla \mathbf{v}]^t\right) \quad [46]$$

The positivity of the entropy source [41] implies that

$$\underline{\boldsymbol{\tau}} : \underline{\dot{\boldsymbol{\varepsilon}}} = \eta \underline{\dot{\boldsymbol{\varepsilon}}} : \underline{\dot{\boldsymbol{\varepsilon}}} \geq 0 \quad [47]$$

and the dynamic viscosity is therefore a positive quantity. Using this relation and assuming η uniform, the divergence of the stress tensor that appears in the momentum conservation equation has the simple form

$$\nabla \cdot \underline{\boldsymbol{\tau}} = \eta \nabla^2 \mathbf{v} + \frac{\eta}{3} \nabla (\nabla \cdot \mathbf{v}) \quad [48]$$

where the vectorial Laplacian $\nabla^2 \mathbf{v}$ is $\nabla(\nabla \cdot \mathbf{v}) - \nabla \times (\nabla \times \mathbf{v})$. This relationship suggests a meaningful interpretation of the viscosity. The momentum equation [16] can be written as

$$\frac{\partial \mathbf{v}}{\partial t} = \frac{\eta}{\rho} \nabla^2 \mathbf{v} + \text{other terms} \ldots \quad [49]$$

forgetting the other terms, a comparison with the thermal diffusion equation [44] shows that the kinematic viscosity ν, defined by

$$\nu = \frac{\eta}{\rho} \quad [50]$$

should rather be called the momentum diffusivity; it plays the same role with respect to the velocity as thermal diffusivity does with respect to temperature.

7.02.2.5 Gravitational Forces

7.02.2.5.1 Poisson's equation

In this chapter, the only force is the gravitational body force. The Earth's gravity is the sum of this gravitational body force and the centrifugal force already discussed in Section 7.02.2.3.2. The gravitational force per unit mass is the gradient of the gravitational potential ψ, a solution of Poisson's equation (Poisson, 1781–1840), that is,

$$\mathbf{g} = -\nabla \psi, \quad \text{and} \quad \nabla^2 \psi = 4\pi G \rho \quad [51]$$

where G is the gravitational constant. In the force term that appears in the momentum equation [16], $\mathbf{f} = \rho \mathbf{g}$, the gravitational force per unit mass should be in agreement with the distribution of masses: the Earth should be self-gravitating.

7.02.2.5.2 Self-gravitation

When dealing with fluid dynamics at the laboratory scale, the gravitational force can be considered as constant and uniform. The gravitational force is related to the entire distribution of mass in the Earth (and the universe) and is practically independent of the local changes in density in the experimental environment. Therefore, at the laboratory scale, it is reasonable to ignore Poisson's equation and to assume that $\mathbf{g}$ is a uniform and constant reference gravitational field.

Inside a planet, the density can be divided into an average depth-dependent density, $\rho_0(r)$; the source of the reference depth-dependent gravitational field, $\mathbf{g}_0(r)$; and a density perturbation $\delta\rho$, the source of a gravitational perturbation $\delta\mathbf{g}$. The force term $\mathbf{f}$ is therefore to first order $\rho_0 \mathbf{g}_0 + \delta\rho \mathbf{g}_0 + \rho_0 \delta\mathbf{g}$; it is tempting to assume that each term in this expression is much larger than the next one and hopefully that only the first two terms are of importance (neglecting the second term would suppress any feedback between density perturbations and flow). Practically, this assumption would imply consideration of the total density anomalies but only the depth-dependent gravitational field. Solving Poisson's equation to compute the perturbed gravitational field would thus be avoided. We can test the previously mentioned idea and show that unfortunately, the third term, $\rho_0 \delta\mathbf{g}$, may be of the same order as the second one (Panasyuk et al., 1996; Ricard et al., 1984; Richards and Hager, 1984). To perform this exercise, we have to introduce the spherical harmonic functions $Y_{lm}(\theta, \phi)$. These functions of latitude θ and longitude ϕ oscillate on a sphere of radius a just like 2-D sinusoidal functions on a plane. Each harmonic function changes sign $l-m$ times from North to South Pole and m times over the same angle (180°) around the equator. The degree l can thus be interpreted as corresponding to a wavelength of order $2\pi a/l$. Spherical harmonics constitute a basis for functions defined on the sphere and are also eigenfunctions of the angular part of Laplace's equation that facilitates the solution of Poisson's equation.

Let us consider a density anomaly $\delta_\rho = \sigma\delta(r-a)Y_{lm}(\theta, \phi)$ at the surface of a sphere of radius a and uniform density ρ_0, where $\delta(r-a)$ is the Dirac delta function (Dirac (1902–84); σ has unit of kg m^{-2}). Inside the planet, this mass distribution generates the radial gravitational perturbation field of

$$\delta g = 4\pi G \sigma \frac{l}{2l+1} \left(\frac{r}{a}\right)^{l-1} Y_{lm}(\theta, \phi) \quad [52]$$

We can compare the terms $\rho_0 \langle \delta g \rangle$ and $g_0 \langle \delta\rho \rangle$, both averaged over the planet radius. For a uniform planet, the surface gravitational force per unit mass is $g_0 = 4\pi G \rho_0 a/3$. Since $\langle \delta_\rho \rangle = \sigma Y_{lm}(\theta, \phi)/a$, we get

$$\frac{\rho_0 \langle \delta g \rangle}{g_0 \langle \delta\rho \rangle} = \frac{3}{2l+1} \quad [53]$$

This estimate is certainly crude and a precise computation taking into account a distributed density distribution could be done. However, this rule of thumb would remain valid. At low degree, the effect of self-gravitation $\rho_0 \delta\mathbf{g}$ is about 50% of the direct effect $\delta\rho \mathbf{g}_0$ and reaches 10% of it only near $l \approx 15$. Self-gravitation has been taken into account in various models intended to explain Earth's gravity field from mantle density anomalies (see also Ribe, this volume, and Forte, vol. 3). Some spherical convection codes account for self-gravitation (e.g., Zhong et al., 2008); the codes that do not include this effect lead to an inaccurate computation of the gravity field.

7.02.2.5.3 Conservative forms of momentum and energy equations

In the general remarks on conservation laws in Section 7.02.2.1, we wrote that conserved quantities like mass, momentum, and energy can only be transported but do not have production terms (contrary to entropy). However, in the momentum conservation [16] and in the energy conservation [29], two terms, $\mathbf{f} = \rho \mathbf{g}$ and $\mathbf{fv} = \rho \mathbf{g} \cdot \mathbf{v}$, appear as sources (we also said that the radioactive term ρH appears because classical physics does not identify mass as energy and vice versa. A negligible mass sink of order $-\rho H/c^2$, where c is the speed of light, should, moreover, be present in the mass conservation).

It is interesting to check that our equations can be recast into an exact conservative form. An advantage of writing equations in conservative form is that it is appropriate to treat with global balances, interfaces, and boundaries (see Section 7.02.2.6). We can obtain conservative equations by using Poisson's relation [51] and performing some algebra (using $\rho=(\nabla^2\psi)/(4\pi G)$ and $\mathbf{g}\cdot\boldsymbol{\nabla}\mathbf{g}=\boldsymbol{\nabla}\mathbf{g}\cdot\mathbf{g}$, since $\mathbf{g}=-\boldsymbol{\nabla}\psi$):

$$\rho\mathbf{g}=-\frac{\boldsymbol{\nabla}\cdot\mathbf{g}}{4\pi G}\mathbf{g}=-\frac{1}{4\pi G}\boldsymbol{\nabla}\cdot\left(\mathbf{g}\otimes\mathbf{g}-\frac{1}{2}g^2\underline{\mathbf{I}}\right) \quad [54]$$

$$\begin{aligned}\rho\mathbf{g}\cdot\mathbf{v}&=-\rho\mathbf{v}\cdot\boldsymbol{\nabla}\psi\\&=-\rho\left(\frac{\mathrm{D}\psi}{\mathrm{D}t}-\frac{\partial\psi}{\partial t}\right)\\&=-\rho\frac{\mathrm{D}\psi}{\mathrm{D}t}-\frac{\boldsymbol{\nabla}\cdot\mathbf{g}}{4\pi G}\frac{\partial\psi}{\partial t}\\&=-\rho\frac{\mathrm{D}\psi}{\mathrm{D}t}-\frac{1}{4\pi G}\boldsymbol{\nabla}\cdot\left(\mathbf{g}\frac{\partial\psi}{\partial t}\right)-\frac{1}{8\pi G}\frac{\partial g^2}{\partial t}\end{aligned} \quad [55]$$

If we substitute these two expressions in the momentum and the energy conservation equations, [16] and [29], we obtain the conservative forms

$$\frac{\partial(\rho\mathbf{v})}{\partial t}=-\boldsymbol{\nabla}\cdot\left(\rho\mathbf{v}\otimes\mathbf{v}+P\underline{\mathbf{I}}-\underline{\boldsymbol{\tau}}+\frac{1}{4\pi G}\mathbf{g}\otimes\mathbf{g}-\frac{1}{8\pi G}g^2\underline{\mathbf{I}}\right) \quad [56]$$

$$\begin{aligned}&\frac{\partial}{\partial t}\left(\rho\left(\mathcal{U}+\psi+\frac{v^2}{2}\right)+\frac{g^2}{8\pi G}\right)\\&=-\boldsymbol{\nabla}\cdot\left(\rho\mathbf{v}\left(\mathcal{U}+\psi+\frac{v^2}{2}\right)+\mathbf{q}+P\mathbf{v}-\underline{\boldsymbol{\tau}}\cdot\mathbf{v}+\frac{\mathbf{g}}{4\pi G}\frac{\partial\psi}{\partial t}\right)+\rho H\end{aligned} \quad [57]$$

When the gravitational force is time-independent, a potential ψ can simply be added to the kinetic and internal energies to replace the work of gravitational forces. When gravitational force and its potential are time-dependent (due to mass redistribution during convection, segregation of elements, etc.), two new terms must be added, a gravitational energy proportional to g^2 and a gravitational flux proportional to $\mathbf{g}\partial\psi/\partial t$ (this is equivalent to the magnetic energy proportional to B^2, where $\mathbf{B}$ is the magnetic induction, in tesla (Tesla, 1856–1943), and to the Poynting vector of magnetohydrodynamics (Poynting, 1852–1914)).

In a permanent or in a statistically steady regime, the time-dependent terms of energy equation [57] can be neglected and the equation can then be integrated over the volume of the Earth. The natural assumption is that the Earth's surface velocities are perpendicular to the Earth's surface normal vector and that the surface is either stress-free or with no horizontal velocity (we exclude the case of convection forced by imposing a nonzero surface velocity). Using the divergence theorem to transform the volume integral of the divergence back to a surface integral of flux, most terms in eqn [57] cancel and all that remains is

$$\int_\Sigma\mathbf{q}\cdot\mathrm{d}\mathbf{S}=\int_\Omega\rho H\,\mathrm{d}V \quad [58]$$

The surface flux in a statistically steady regime is simply the total radiogenic heat production.

It may seem surprising at first that viscous dissipation that is everywhere positive does not appear in this balance. To understand this point, we can directly integrate the energy equation written in terms of temperature [38],

$$\int_\Omega\rho C_P\frac{\mathrm{D}T}{\mathrm{D}t}\mathrm{d}V=-\int_\Sigma\mathbf{q}\cdot\mathrm{d}\mathbf{S}+\int_\Omega\left(\alpha T\frac{\mathrm{D}P}{\mathrm{D}t}+\underline{\boldsymbol{\tau}}:\boldsymbol{\nabla}\mathbf{v}\right)\mathrm{d}V+\int_\Omega\rho H\mathrm{d}V \quad [59]$$

On the right side, the first and last terms cancel each other out by eqn [58]. To simplify the left side, notice that the enthalpy $\mathcal{H}=\mathcal{U}+P/\rho$ obeys (see eqn [36])

$$\rho\frac{\mathrm{D}\mathcal{H}}{\mathrm{D}t}=\rho C_P\frac{\mathrm{D}T}{\mathrm{D}t}+(1-\alpha T)\frac{\mathrm{D}P}{\mathrm{D}t} \quad [60]$$

According to eqn [11], $\rho\mathrm{D}\mathcal{H}/\mathrm{D}t=\partial(\rho\mathcal{H})/\partial t+\boldsymbol{\nabla}(\rho\mathbf{v}\mathcal{H})$. The last term cancels when integrated on Earth's volume (using the divergence theorem), and the first term cancels in a statistically steady regime. Therefore, the time-average volume integral of $\rho\mathrm{D}\mathcal{H}/\mathrm{D}t$ is zero and the volume integral of $\rho C_p\mathrm{D}T/\mathrm{D}t$ must be opposite to that of $(1-\alpha T)\mathrm{D}P/\mathrm{D}t$. The time average of the energy equation [59] implies therefore (Alboussiere and Ricard, 2013)

$$\int_\Omega\left(\frac{\mathrm{D}P}{\mathrm{D}t}+\underline{\boldsymbol{\tau}}:\boldsymbol{\nabla}\mathbf{v}\right)\mathrm{d}V=0 \quad [61]$$

The total heat production due to dissipation is balanced by the work due to compression and expansion over the convective cycle (notice that if we had assumed a constant C_P, the left side of eqn [59] would have been zero and we would have found an αT term in front of $\mathrm{D}P/\mathrm{D}t$ (Hewitt et al., 1975), which is incorrect). This balance is global not local. Dissipation occurs mostly near the boundary layers of the convection and compressional work is done along the downwellings and upwellings of the flow. The relation [61] can also be derived by multiplying the Navier–Stoke equation [16] by $\mathbf{v}$ and performing a volume integration.

7.02.2.6 Boundary and Interface Conditions

7.02.2.6.1 General method

A boundary condition is a special case of an interface condition when certain properties are taken as known on one side of the interface. Sometimes, the properties are explicitly known (e.g., the three velocity components are zero on a no-slip surface), but often, an interface condition simply expresses the continuity of a conserved quantity. To obtain the continuity conditions for a quantity A, the general method is to start from the conservation equation of A in its integral form (see eqn [5]). We choose a cylindrical volume Ω (a pillbox) of infinitely small radius R where the top and bottom surfaces are located at a distance $\pm\epsilon$ from a discontinuity surface (see Figure 2). We choose two Cartesian axes Ox and Oy; we call $\mathbf{n}$ the upward

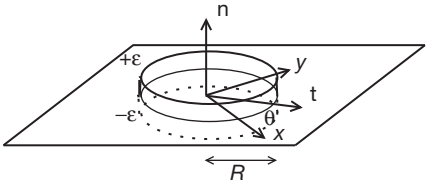

Figure 2 The pillbox volume used to derive the interface and boundary conditions.

unit vector normal to the interface and $\mathbf{t}$ a radial unit vector normal to the cylindrical side of the pillbox; and θ is the angle between $\mathbf{t}$ and Ox. If we now make the volume $\Omega(\epsilon)$ shrink to zero by decreasing ϵ at constant R, the volume integrals of the time-dependent and the source terms will also go to zero (unless the source term contains explicit surface terms like in the case of surface tension, but this is irrelevant at mantle scales). We simply get

$$\lim_{\epsilon\to 0}\int_{\Sigma(\epsilon)} \mathbf{J}_A\cdot d\mathbf{S} = 0 \qquad [62]$$

(the demonstration is here written for the vector flux $\mathbf{J}_A$ of the scalar A but is easily extended to tensor flux of a vectorial quantity $\mathbf{A}$). This integral can also be written as

$$\pi R^2[\mathbf{J}_A]\cdot\mathbf{n} + R\int_{-\epsilon}^{+\epsilon}\int_0^{2\pi}\mathbf{J}_A\cdot\mathbf{t}\,dz\,d\theta \qquad [63]$$

where $[X]$ is the jump of X across the interface, sometimes noted X^+–X^-. In most cases, the second term goes to zero with ϵ because the components of $\mathbf{J}_A$ are bounded or is exactly zero when the flux is not a function of θ (since the double integral becomes the product of an integral in z times $\int_0^{2\pi}\mathbf{t}\,d\theta = 0$). In these cases, the boundary condition for A becomes

$$[\mathbf{J}_A]\cdot\mathbf{n} = 0 \qquad [64]$$

At an interface, the normal flux of A must therefore be continuous. However, in some cases, for example, when $\mathbf{J}_A$ contains a z-derivative, the second term may not cancel and this happens in the case of boundaries associated with phase changes (see Section 7.02.2.6.3).

7.02.2.6.2 Interface conditions in the 1-D case and for bounded variables

Using the mass, momentum, energy, and entropy conservations in their conservative forms in eqns [9], [56], [57], and [40] and assuming for now that no variable becomes infinite at an interface, the interface conditions in the reference frame in which the interface is motionless are

$$\begin{aligned}
[\rho\mathbf{v}]\cdot\mathbf{n} &= 0\\
[\underline{\boldsymbol{\tau}}]\cdot\mathbf{n} - [P]\mathbf{n} &= 0\\
[\rho\mathbf{v}\mathcal{U} + \mathbf{q} - \underline{\boldsymbol{\tau}}\cdot\mathbf{v} + P\mathbf{v}]\cdot\mathbf{n} &= 0\\
\left[\rho\mathbf{v}\mathcal{S} + \frac{\mathbf{q}}{T}\right]\cdot\mathbf{n} &= 0
\end{aligned} \qquad [65]$$

(the gravitational force per unit mass and its potential are continuous). In these equations, we neglected the inertia and the kinetic energy terms in the second and third equations of eqn [65] as appropriate for the mantle. When these terms are accounted for (adding $[-\rho\mathbf{v}\otimes\mathbf{v}]\cdot\mathbf{n}$ to the second equation and $[-\rho\mathbf{v}\upsilon^2/2]\cdot\mathbf{n}$ to the third), these equations are known as the Hugoniot–Rankine conditions (Hugoniot, 1851–87; Rankine, 1820–72).

At the surface of a fluid, and on any impermeable interfaces where $\mathbf{v}\cdot\mathbf{n} = 0$ and $[\mathbf{v}] = 0$, the general jump conditions [65] without inertia imply that the heat flux $[\mathrm{q}]\cdot\mathbf{n}$, the entropy flux $[\mathbf{q}/\mathrm{T}]\cdot\mathbf{n}$ (and therefore the temperature T), and the stress components $[\underline{\boldsymbol{\tau}}]\cdot\mathbf{n} - [P]\mathbf{n}$ are continuous. In 3-D, four boundary conditions are necessary on a surface to solve for the three components of velocity and for the temperature. The temperature (or the heat flux) can be imposed, and, for the velocity and stress, either free-slip boundary conditions $\mathbf{v}\cdot\mathbf{n} = 0$, which is the first condition of eqn [65] and $(\underline{\boldsymbol{\tau}}\cdot\mathbf{n})\times\mathbf{n} = 0$, or no-slip boundary conditions (the horizontal surface velocity $\mathbf{v}$ is imposed) are generally used.

Inside the mantle, minerals undergo several phase transitions at depth and at least two of them, the olivine $\rightleftharpoons$ wadsleyite and the ringwoodite $\rightleftharpoons$ perovskite + ferropericlase transitions around 410 and 660 km depth, respectively, are sharp enough to be modeled by discontinuities. Conditions [65] have been used in several convection models (assuming therefore that all variables remain bounded).

When the kinetic energy is neglected, and the viscous stresses are much smaller than the pressure term, which are two approximations valid for the mantle, the last two boundary conditions are, assuming continuity of temperature,

$$\begin{aligned}
\left[\rho\mathbf{v}\left(\mathcal{U} + \frac{P}{\rho}\right)\right]\cdot\mathbf{n} + [\mathbf{q}]\cdot\mathbf{n} &= 0\\
[\rho\mathbf{v}\mathcal{S}]\cdot\mathbf{n} + \frac{1}{T}[\mathbf{q}]\cdot\mathbf{n} &= 0
\end{aligned} \qquad [66]$$

The diffusive flux $\mathbf{q}$ can be eliminated from these two equations. Since $p\mathbf{v}$ is continuous, and remembering that $\mathcal{U} + P/p$ is the enthalpy $\mathcal{H}$, we simply recognize the Clapeyron condition, which is latent heat release,

$$\Delta\mathcal{H} = T\Delta\mathcal{S} \qquad [67]$$

where the enthalpy and entropy jumps, $[\mathcal{H}]$ and $[\mathcal{S}]$, were replaced by their more traditional notations, $\Delta\mathcal{H}$ and $\Delta\mathcal{S}$. The heat flux is discontinuous across an interface,

$$\Delta\mathcal{H}\rho\mathbf{v}\cdot\mathbf{n} + [\mathbf{q}]\cdot\mathbf{n} = 0 \qquad [68]$$

and the discontinuity amounts to the enthalpy released by the mass flux that has undergone a chemical reaction or a phase change.

7.02.2.6.3 Phase change interfaces and stress continuity?

A mathematical dilemma arises in using the conditions [65] across a phase change interface that does not seem to have been considered except by Corrieu et al. (1995): a phase change induces a stress discontinuity at an interface.

We illustrate the mathematical difficulty by discussing the stress continuity in the second condition of eqn [65]. The problem arises from the term in $\partial\upsilon_z/\partial_z$ present in the rheological law [45] that becomes infinite when the material is forced to change its density discontinuously. To enforce the change in shape that occurs locally, the normal horizontal stresses have to become infinite, and therefore, their contributions to the force equilibrium of a pillbox do not vanish when the pillbox height is decreased.

To derive the appropriate interface condition, we have to consider again eqn [63] where $\mathbf{J}_A$ is substituted by $\underline{\sigma}$. The interface is considered as laterally infinite and perfectly flat (without curvature) and the viscosity is uniform. The only terms that may be unbounded on the interface are σ_{xx}, σ_{yy}, and σ_{zz}. Omitting the other stress components, which would make no contribution to the interface condition when ϵ goes to zero, the stress continuity becomes

$$\pi R^2\left[\underline{\boldsymbol{\sigma}}\right]\cdot\mathbf{n}+R\mathbf{e}_x\int_{-\epsilon}^{+\epsilon}\int_0^{2\pi}\sigma_{xx}\cos\theta\,\mathrm{d}z\mathrm{d}\theta+R\mathbf{e}_y\int_{-\epsilon}^{+\epsilon}\int_0^{2\pi}\sigma_{yy}\sin\theta\,\mathrm{d}z\mathrm{d}\theta=0 \quad [69]$$

Since R is small, we can replace the stresses on the cylindrical side of the pillbox by their first-order expansions, for example, $\sigma_{xx}=\sigma_{xx}(0)+(\partial\sigma_{xx}/\partial x)R\cos\theta+(\partial\sigma_{xx}/\partial y)R\sin\theta$. After integration in θ and division by πR^2, one gets

$$\left[\underline{\boldsymbol{\sigma}}\right]\cdot\mathbf{n}+\mathbf{e}_x\frac{\partial}{\partial x}\int_{-\epsilon}^{+\epsilon}\sigma_{xx}\,\mathrm{d}z+\mathbf{e}_y\frac{\partial}{\partial y}\int_{-\epsilon}^{+\epsilon}\sigma_{yy}\,\mathrm{d}z=0 \quad [70]$$

This expression already demonstrates the continuity of σ_{zz}. Using $\sigma_{xx}=\sigma_{zz}+2\eta\partial v_x/\partial x-2\eta\partial v_z/\partial z$, and assuming that the viscosity remains uniform, we see that

$$\lim_{\epsilon\to0}\int_{-\epsilon}^{\epsilon}\sigma_{xx}\,\mathrm{d}z=-2\eta\lim_{\epsilon\to0}\int_{-\epsilon}^{\epsilon}\frac{\partial v_z}{\partial z}\mathrm{d}z=-2\eta[v_z] \quad [71]$$

The same result holds for the σ_{yy} term. Since v_z is discontinuous (because of mass-flux continuity and the density jump), a sudden change in volume implies a discontinuity of the tangential stresses. The stress boundary conditions are thus

$$[\tau_{xz}]-2\eta\frac{\partial}{\partial x}[v_z]=[\tau_{yz}]-2\eta\frac{\partial}{\partial y}[v_z]=[\tau_{zz}-P]=0 \quad [72]$$

in disagreement with eqn [65]. Infinite horizontal stresses and discontinuous shear stresses occur in the presence of an interface with surface tension (Joseph and Renardy, 1993). In the case of a permeable and infinitely thin phase transition, there should also be a surface tension with a dynamic origin, simply related to $[v_z]$ when the viscosity is uniform. A rigorous mathematical demonstration that proves the existence of a normal stress discontinuity in the case of a curved interface in perfect analogy with usual surface tension and that is extended for a variable viscosity is given in Chambat et al. (2013). Notice also that the pressure and the viscous dissipation should also be unbounded. Until a complete clarification of these discontinuity problems is achieved, it seems that the usual continuity conditions across phase changes may not be mathematically consistent.

7.02.2.6.4 Weakly deformable surface of a convective cell

When a no-slip condition is imposed at the surface, both normal and shear stresses are present at the boundary. These stresses, according to the second interface condition [65], must balance the force $-\underline{\boldsymbol{\tau}}\cdot\mathbf{n}+P\mathbf{n}$ exerted by the fluid. This is reasonable for a laboratory experiment with a fluid totally enclosed in a tank whose walls are rigid enough to resist fluid traction. However, in the case of free-slip boundary conditions, it may seem strange that by imposing a zero vertical velocity, a finite normal stress results at the free surface. It is therefore worth discussing this point in more detail.

The natural boundary conditions should be that both the normal and the tangential stresses applied on the free deformable surface, $z=h(x,y,t)$, of a convective fluid are zero:

$$\left(\underline{\boldsymbol{\tau}}\cdot\mathbf{n}-P\mathbf{n}\right)_{\text{on }z=h}=0 \quad [73]$$

(neglecting atmospheric pressure). In this expression, the topography h is unknown and the normal $\mathbf{n}$, computed at the surface of the planet, is $\mathbf{n}=(\mathbf{e}_z-\boldsymbol{\nabla}_H h)/\sqrt{1+|\boldsymbol{\nabla}_H h|^2}$ where $\mathbf{e}_z$ is the unit vector along z, opposite to gravity, and $\boldsymbol{\nabla}_H$ the horizontal component of the gradient operator.

The variation of topography is related to the convective flow and satisfies

$$\frac{\partial h}{\partial t}+\mathbf{v}_H^0\cdot\boldsymbol{\nabla}_H h-v_z^0=0 \quad [74]$$

This equation expresses the fact that a material particle on the surface remains always on it. In this expression, $\mathbf{v}_H^0$ and v_z^0 are the horizontal and vertical velocity components at the surface of the planet. We will see in Section 7.02.4 that lateral pressure and stress variations are always very small compared to the average pressure (this is because in most fluids, and in the mantle, the lateral density variations remain negligible compared to the average density). This implies that the surface topography is not much affected by the internal dynamics and remains close to horizontal at long wavelengths, $|\boldsymbol{\nabla}_H h|\ll1$. Boundary condition [73] and topography advection [74] can therefore be expanded to first order to give

$$\left(\underline{\boldsymbol{\tau}}\cdot\mathbf{e}_z-P\mathbf{e}_z\right)_{\text{on }z=0}\approx-\rho_0 g_0 h\mathbf{e}_z,\quad \frac{\partial h}{\partial t}=v_z^0 \quad [75]$$

where we again make use of the fact that the total stress remains close to hydrostatic, that is, $\mathbf{n}\cdot\underline{\boldsymbol{\tau}}\cdot\mathbf{n}\ll P$ (p_0 and g_0 are the surface values of density and gravity). To first order, the stress boundary condition on a weakly deformable top surface is therefore zero shear stress but with a time-dependent normal stress related to the surface topography and vertical velocity.

The convection equations with these boundary conditions could be solved, but this is not always useful. Since the boundary conditions involve both displacement h and velocity v_z^0, the solution is akin to an eigenvalue problem. It can be shown that for an internal density structure of wavelength, λ, v_z^0 goes to zero in a time of order $\overline{\eta}/\rho_0 g_0\lambda$ where $\overline{\eta}$ is the typical viscosity of the underlying liquid over the depth λ (Richards and Hager, 1984). For the Earth, this time is the characteristic time of postglacial rebound and is typically a few thousand years for wavelengths of a few thousand kilometers (e.g., Spada et al., 1992b).

For convection, where the characteristic times are much longer, it is thus appropriate to assume that the induced topography is in mechanical equilibrium with the internal density structure. A zero normal velocity can therefore be imposed and the resulting normal stress can be used to estimate the topography generated by the convective flow. Internal compositional interfaces can be treated in a similar manner if they are only weakly deformable (i.e., when their intrinsic density jumps are much larger than the thermal density variations). This is the case for the core–mantle boundary (CMB).

For rapid events (e.g., for a localized thermal anomaly impinging the Earth's surface), the time for topographic equilibration becomes comparable to the timescale of internal convective processes. In this case, the precise computation of a history-dependent topography is necessary and the finite elasticity of the lithosphere, the coldest part of the mantle, plays an important role (Zhong et al., 1996). Various studies have also shown that at short wavelength, the condition of free surface differs significantly from the more usual free-slip condition (Crameri et al., 2012; Kaus et al., 2010; Schmeling et al., 2008). In particular, one-sided subductions, as observed in nature and in analogical experiments, are reproduced much

more accurately by numerical codes that include a true deformable topography according to eqn [73].

7.02.3 Thermodynamic and Rheological Properties

Section 7.02.2 on conservation equations is valid for all fluids (although the interface conditions are mostly discussed when inertia and kinetic energy are negligible). The differences between mantle convection and core, oceanic, or atmospheric convection come from the thermodynamic and transport properties of solids that are very different from those of usual fluids. We review some basic general properties of solids in Section 7.02.3 and will be more specific in Section 7.02.6.

7.02.3.1 Equation of State and Solid Properties

The equation of state of any material (EoS) relates its pressure, density, and temperature. The equation of state of a perfect gas, $PV/T=$constant, is well known, but irrelevant for solids. Unfortunately, there is no equation for solids based on a simple and efficient theoretical model. In the mineralogical community, the third-order finite strain Birch–Murnaghan EoS seems highly favored (Birch, 1952). This equation is cumbersome and is essentially empirical. More physical approaches have been used in Vinet et al. (1987), Poirier and Tarantola (1998), and Stacey and Davis (2004), but it seems that for each solid, the EoS has to be obtained experimentally (for more details, see Stixrude, vol. 2).

In the simplest cases, the density varies around ρ_0 measured at temperature T_0 and pressure P_0 as

$$\rho=\rho_0\left(1-\alpha(T-T_0)+\frac{P-P_0}{K_T}\right) \qquad [76]$$

where the thermal expansivity α and incompressibility K_T have been defined in eqn [34]. This expression is a first-order expansion of any EoS. Equation [76] can however be misleading if one forgets that the parameters α and K_T cannot be constant but must be related through Maxwell relations (e.g., their definitions [34] imply that $\partial(\alpha\rho)/\partial P=-\partial(\rho/K_T)/\partial T$).

Equation [76] can be used for a very simple numerical estimate that illustrates an important characteristic of solid Earth geophysics. Typically for silicates $\alpha\approx10^{-5}\mathrm{K}^{-1}$, $K_T\approx10^{11}$ Pa, while temperature variations in the mantle, ΔT, are of a few 1000 K with a pressure increase between the surface and the core, ΔP, of order of 10^{11} Pa. This indicates that the overall density variations due to temperature differences are negligible compared to those due to pressure differences ($\alpha\Delta T\ll1$ but $\Delta P/K_T\approx1$). In planets, to first order, the radial density is only a function of pressure, not of temperature. This is opposite to most liquid or solid laboratory experiments, where the properties are usually controlled by temperature.

A very important quantity in the thermodynamics of solids is the Grüneisen parameter (Grüneisen, 1877–1949)

$$\Gamma=\frac{\sigma K_T}{\rho C_v}=\frac{1}{\rho C_v}\left(\frac{\partial P}{\partial T}\right)_V \qquad [77]$$

The Grüneisen parameter is dimensionless, does not vary much through the mantle (around 1), and can reasonably be considered as independent of the temperature. An empirical law (Anderson, 1979) relates Γ with the density,

$$\Gamma=\Gamma_0\left(\frac{\rho_0}{\rho}\right)^q \qquad [78]$$

where q is around 1. The Grüneisen parameter can also be related to the microscopic vibrational properties of crystals (Stacey, 1977). At high temperature, above the Debye temperature (Debye, 1884–1966), all solids have more or less the same heat capacity at constant volume. This is called the Dulong and Petit rule (Dulong, 1785–1838; Petit, 1791–1820). At high T, each atom vibrates and the thermal vibrational energy is equipartitioned between the three dimensions (degrees of freedom), which leads to $C_{Vm}=3R$ per mole of atoms, independent of the nature of the solid (R is the gas constant). Assuming that the mantle is made of pure forsterite Mg_2SiO_4 that contains seven atoms for a molar mass of 140 g, its heat capacity at constant volume is therefore close to $C_{Vm}=21R=174.56\ \mathrm{J\,K^{-1}\,mol^{-1}}$ or $C_V=1247\ \mathrm{J\,K^{-1}\,kg^{-1}}$. The approximate constancy of C_V and the fact that Γ is only a function of ρ allow us to integrate [77]

$$P=F(\rho)+\alpha_0K_T^0(T-T_0)\left(\frac{\rho}{\rho_0}\right)^{1-q} \qquad [79]$$

where α_0 and K_T^0 are the thermal expansivity and incompressibility at standard conditions and where $F(\rho)$ is a density-dependent integration constant. A rather simple but acceptable choice for the function $F(\rho)$, at least for mantle dynamicists, is the Murnaghan EoS (Murnaghan, 1951) at constant T that allows us to write an EoS for solids of the form

$$P=\frac{K_T^0}{n}\left[\left(\frac{\rho}{\rho_0}\right)^n-1\right]+\alpha_0K_T^0(T-T_0)\left(\frac{\rho}{\rho_0}\right)^{1-q} \qquad [80]$$

This equation could easily be used to derive any thermodynamic property like $\alpha(P,T)$ or $K_T(P,T)$. The exponent n expresses the variations of K_T with pressure at reference temperature T_0 as $K_T(P,T_0)=K_T^0+nP$. The value of n (also sometimes called K_0') is of order of 4 for most mantle materials (Stixrude and Lithgow-Bertelloni, 2005). This equation has been used implicitly in various models of mantle convection (e.g., Bercovici et al., 1989a, 1992; Glatzmaier, 1988). An important consequence of this EoS assuming $q\approx1$ is that αK_T is more or less constant and that

$$K_T\approx K_T^0\left(\frac{\rho}{\rho_0}\right)^n,\quad \alpha\approx\alpha_0\left(\frac{\rho}{\rho_0}\right)^{-n} \qquad [81]$$

In the mantle, the incompressibility increases and the thermal expansion decreases significantly with depth, typically by a factor ≈8. The geophysical consequences are further discussed in Section 7.02.6.5.1.

Two other thermodynamic equalities can also be straightforwardly deduced by chain rules of derivatives and will be used in the following. A relation between the two heat capacities C_P and C_V of the energy equations [38] can be derived from the two expressions for heat increments, eqn [31], and the definition of h, eqn [33],

$$C_P-C_V=\frac{\alpha T}{\rho}\left(\frac{\partial P}{\partial T}\right)_V \qquad [82]$$

The same expressions for heat increments, eqn [31], and the l and h definitions, eqn [33], imply that for an adiabatic transformation (when δQ and $\mathrm{d}\mathcal{S}$ are zero),

$$\left(\frac{\partial P}{\partial T}\right)_{\mathcal{S}} = \frac{\rho C_P}{\alpha T} \quad \text{and} \quad \left(\frac{\partial V}{\partial T}\right)_{\mathcal{S}} = -\frac{C_V}{\alpha K_T T} \qquad [83]$$

Equations [82] and [83] take simpler forms when the Grüneisen parameter [77] and the adiabatic compressibility defined by

$$K_{\mathcal{S}} = \rho\left(\frac{\partial P}{\partial \rho}\right)_{\mathcal{S}} = -\frac{1}{\rho}\left(\frac{\partial P}{\partial T}\right)_{\mathcal{S}} \Big/ \left(\frac{\partial V}{\partial T}\right)_{\mathcal{S}} \qquad [84]$$

are used; they are

$$\frac{C_P}{C_V} = \frac{K_{\mathcal{S}}}{K_T} = 1 + \Gamma\alpha T \qquad [85]$$

Since $\Gamma \approx 1$ and since $\alpha T \ll 1$, the two heat capacities are basically equal. The incompressibility $K_{\mathcal{S}}$ is defined similarly to K_T but at constant entropy. The theory of elastic waves introduces this parameter that can be obtained from seismic observations

$$K_{\mathcal{S}} = \rho\left(\upsilon_p^2 - \frac{4}{3}\upsilon_s^2\right) \qquad [86]$$

where υ_p and υ_s are the P and S wave velocities. We will see later (**Section 7.02.4.2.2**) that a vigorously convective fluid should be close to adiabatic. As this is the case for the mantle, K_S should be the incompressibility measured along the mantle density profile

$$K_{\mathcal{S}} = \rho\left(\frac{\partial P}{\partial \rho}\right)_{\mathcal{S}} = \rho\frac{\|\boldsymbol{\nabla} P\|}{\|\boldsymbol{\nabla}\rho\|} = \frac{\rho^2\|\mathbf{g}\|}{\|\boldsymbol{\nabla}\rho\|} \qquad [87]$$

The relations [86] and [87] connect a seismological observation $K_S(r)/\rho(r)$ to the density gradient of the real Earth $\rho(r)$ $g(r)/\|\boldsymbol{\nabla}\rho(r)\|$. This is the important Bullen hypothesis (Bullen, 1940) used to build the reference density of the Earth (e.g., Dziewonski and Anderson, 1981).

7.02.3.2 Rheology

In **Section 7.02.2.3** on momentum conservation, no assumption is made about the rheology of the fluid, that is, on the relation between the stress tensor and the flow itself. In contrast, the discussion of energy conservation (**Section 7.02.2.4**) relies on the assumption that the pressure-related work is entirely recoverable, eqn [32]; as a consequence, the work of the deviatoric stresses ends up entirely as a dissipative term, hence a source of entropy. In a real fluid, this may be wrong for two reasons: part of the deviatoric stresses may be recoverable and part of the isotropic work may not be recoverable. In the first case, elasticity may be present, and in the second case, bulk viscosity.

7.02.3.2.1 Elasticity

On a very short timescale, the mantle is an elastic solid in which compressional and shear waves propagate (e.g., Kennett, 2001). In an elastic solid, the linear strain tensor,

$$\underline{\boldsymbol{\varepsilon}}^{\mathrm{e}} = \frac{1}{2}\left(\boldsymbol{\nabla}\mathbf{u} + [\boldsymbol{\nabla}\mathbf{u}]^t\right) \qquad [88]$$

in which $\mathbf{u}$ is the displacement vector (this is valid for small deformations, see e.g., Malvern, 1969; Landau and Lifchitz, 2000, for the large deformation case), is linearly related to the stress tensor,

$$\sigma_{ij}^{\mathrm{e}} = \sum_{k,l} A_{ijkl}^{\mathrm{e}} \epsilon_{kl}^{\mathrm{e}} \qquad [89]$$

where A^{e} is the fourth-rank stiffness tensor. Since both the stress and the strain tensors are symmetrical and because of the Maxwell thermodynamic relations for internal energy (including the elastic energy), $\partial^2\mathcal{U}/\partial\epsilon_{ij}\partial\epsilon_{kl} = \partial^2\mathcal{U}/\partial\epsilon_{kl}\partial\epsilon_{ij}$, the elastic tensor is invariant to permutations of i and j, k and l, and ij and kl. This leaves in the most general case of anisotropy 21 independent stiffness coefficients (Malvern, 1969). In crystals, this number decreases with the number of symmetries of the unit cell. For isotropic elastic solids, only two parameters are needed, the incompressibility K and the rigidity μ_{R}, and the elastic behavior satisfies

$$\underline{\boldsymbol{\sigma}}^{\mathrm{e}} = K\mathrm{tr}\left(\underline{\boldsymbol{\varepsilon}}^{\mathrm{e}}\right)\underline{\mathbf{I}} + 2\mu_{\mathrm{R}}\left(\underline{\boldsymbol{\varepsilon}}^{\mathrm{e}} - \frac{1}{3}\mathrm{tr}\left(\underline{\boldsymbol{\varepsilon}}^{\mathrm{e}}\right)\underline{\mathbf{I}}\right) \qquad [90]$$

where $\mathrm{tr}\left(\underline{\boldsymbol{\varepsilon}}^{\mathrm{e}}\right) = \boldsymbol{\nabla}\cdot\mathbf{u}$.

Two remarks can be made on this rapid presentation of elasticity that are more deeply developed in textbooks of mechanics (e.g., Landau and Lifchitz, 2000; Malvern, 1969) or of seismology (e.g., Dahlen and Tromp, 1998). First, the expression [90] assumes that the displacement vector is computed from an initial situation where the solid is perfectly stress-free, that is, $\underline{\boldsymbol{\sigma}}^{\mathrm{e}} = 0$ when $\underline{\boldsymbol{\varepsilon}}^{\mathrm{e}} = 0$. In practical problems, only incremental displacements with respect to an initial prestressed state are known and $\underline{\boldsymbol{\sigma}}^{\mathrm{e}}$ has to be understood as a variation of the stress tensor. Second, temperature variations are associated with changes in elastic stresses and the definition of incompressibility K should take these variations into account. The incompressibility should be $K = K_S$ for rapid adiabatic seismic waves and $K = K_T$ for isothermal variations. The other elastic parameters that are often introduced, Poisson's ratio, Young's modulus (Young, 1773–1829), and Lamé's parameters (Lamé, 1795–1870), are simple functions of incompressibility and rigidity. Since the term proportional to μ_{R} is traceless, equation [90] leads to $\mathrm{tr}\left(\underline{\boldsymbol{\sigma}}^{\mathrm{e}}\right) = 3K\mathrm{tr}\left(\underline{\dot{\boldsymbol{\varepsilon}}}^{\mathrm{e}}\right)$, and the rheology law can also be written in terms of compliance (i.e., getting $\underline{\boldsymbol{\varepsilon}}^{\mathrm{e}}$ as a function of $\underline{\boldsymbol{\sigma}}^{\mathrm{e}}$),

$$\underline{\boldsymbol{\varepsilon}}^{\mathrm{e}} = \frac{1}{9K}\mathrm{tr}\left(\underline{\boldsymbol{\sigma}}^{\mathrm{e}}\right)\underline{\mathbf{I}} + \frac{1}{2\mu_{\mathrm{R}}}\left(\underline{\boldsymbol{\sigma}}^{\mathrm{e}} - \frac{1}{3}\mathrm{tr}\left(\underline{\boldsymbol{\sigma}}^{\mathrm{e}}\right)\underline{\mathbf{I}}\right) \qquad [91]$$

In these equations, the trace of the stress tensor can also be replaced by the pressure definition

$$\mathrm{tr}\left(\underline{\boldsymbol{\sigma}}^{\mathrm{e}}\right) = -3P \qquad [92]$$

The momentum equation [15] remains valid in a purely elastic solid (except that the advective transport is generally neglected, $\mathrm{D}/\mathrm{D}t \approx \partial/\partial t$), but the discussion of energy conservation and thermodynamics is different for elastic and viscous bodies. The work of the elastic stress is entirely recoverable: a deformed elastic body returns to its undeformed shape when the external forces are released. The internal energy change due to the storage of elastic stress is $\delta W = V\underline{\boldsymbol{\sigma}}_{\mathrm{e}} : \mathrm{d}\underline{\boldsymbol{\varepsilon}}^{\mathrm{e}}$ instead of

$\delta W = -P\mathrm{d}V$, and this is provided by the deformation work term $\underline{\boldsymbol{\tau}} : \nabla \mathbf{v}$, which is therefore nondissipative. Thus, for an elastic body, the temperature equations [38] and the entropy equation [39] hold, but with the $\underline{\boldsymbol{\tau}} : \nabla \mathbf{v}$, source term removed.

7.02.3.2.2 Viscous Newtonian rheology

On a very long timescale, it is reasonable to assume that the internal deviatoric stresses become eventually relaxed and dissipated as heat. This is the assumption that we have implicitly made and that is usual in fluid mechanics. Since the dissipative term is $\underline{\boldsymbol{\tau}} : \nabla \mathbf{v} = \underline{\boldsymbol{\tau}} : \underline{\dot{\boldsymbol{\varepsilon}}}^{v}$ and must be positive according to the second law, this suggests a relationship between velocity-related stresses and velocity derivatives such that the total stress tensor has the form

$$\sigma_{ij}^{v} = -P\delta_{ij} + A_{ijkl}^{v}\dot{\varepsilon}_{kl}^{v} \qquad [93]$$

δ_{ij} being the Kronecker symbol (Kronecker, 1823–91). Except for the time derivative, the only formal difference between this expression and eqn [89] is that pressure exists in a motionless fluid but is always associated with deformation in an elastic solid.

Using the same arguments as for the elastic case, the viscous rheology in the isotropic case can therefore be written in terms of stiffness

$$\underline{\boldsymbol{\sigma}}^{v} = \left(-P + \zeta \mathrm{tr}\left(\underline{\dot{\boldsymbol{\varepsilon}}}^{v}\right)\right)\underline{\mathbf{I}} + 2\eta\left(\underline{\dot{\boldsymbol{\varepsilon}}}^{v} - \frac{1}{3}\mathrm{tr}\left(\underline{\dot{\boldsymbol{\varepsilon}}}^{v}\right)\underline{\mathbf{I}}\right) \qquad [94]$$

where $\mathrm{tr}\left(\underline{\dot{\boldsymbol{\varepsilon}}}^{v}\right) = \nabla \cdot \mathbf{v}$. Using $\mathrm{tr}\left(\underline{\boldsymbol{\sigma}}^{v}\right) = 3\left(-P + \zeta \mathrm{tr}\left(\underline{\dot{\boldsymbol{\varepsilon}}}^{v}\right)\right)$, the rheology can also be expressed in terms of compliance

$$\underline{\dot{\boldsymbol{\varepsilon}}}^{v} = \frac{1}{9\zeta}\left(3P + \mathrm{tr}\left(\underline{\boldsymbol{\sigma}}^{v}\right)\right)\underline{\mathbf{I}} + \frac{1}{2\eta}\left(\underline{\boldsymbol{\sigma}}^{v} - \frac{1}{3}\mathrm{tr}\left(\underline{\boldsymbol{\sigma}}^{v^2}\right)\underline{\mathbf{I}}\right) \qquad [95]$$

The two parameters η and ζ are positive according to the second law and are called the shear and bulk viscosities. When they are intrinsic material properties (i.e., independent of the flow itself), the fluid is called linear or Newtonian. The hypothesis of isotropy of the rheology is probably wrong for a mantle composed of highly anisotropic materials (see Karato, 1998), but only a few papers have tried to tackle the problem of anisotropic viscosity (Christensen, 1997a; Muhlhaus et al., 2004; Pouilloux et al., 2007).

Since $\mathrm{tr}\left(\underline{\boldsymbol{\sigma}}^{v}\right)/3 = -P + \zeta\, \nabla \cdot \mathbf{v}$, the isotropic average of the total stress is not the pressure term, unless $\zeta \nabla \cdot \mathbf{v} = 0$. Therefore, part of the stress work, $\underline{\boldsymbol{\tau}} : \underline{\dot{\boldsymbol{\varepsilon}}}^{v}$, during isotropic compaction could be dissipated in the form of the heat source $\zeta(\nabla \cdot \mathbf{v})^2$. A density-independent bulk viscosity allows an infinite compression under a finite isotropic stress. The bulk viscosity parameter ζ is generally only introduced to be immediately omitted and we will do the same. However, using eqn [94] with $\zeta = 0$ but keeping $\nabla \cdot \mathbf{v} \neq 0$ is not valid since it would remove all resistance to isotropic compression. We will see that the correct physical behavior is obtained by setting $\zeta = +\infty$: elastic stresses are present and provide the resistance to isotropic compression. A finite bulk viscosity, or some equivalent concept, is however necessary to handle two-phase compaction problems (Bercovici et al., 2001a; McKenzie, 1984) (see **Section 7.02.5.2**).

7.02.3.2.3 Maxwellian viscoelasticity

To account for the fact that the Earth behaves elastically on short time constants and viscously at long times, it is often assumed that under the same stress, the deformation has both elastic and viscous components. By summing the viscous compliance equation [95] with the time derivative of the elastic compliance equation, [91] and in the case of an infinite bulk viscosity ζ, we get

$$\underline{\dot{\boldsymbol{\varepsilon}}} = \frac{1}{9K}\mathrm{tr}\left(\underline{\dot{\boldsymbol{\sigma}}}\right)\underline{\mathbf{I}} + \frac{1}{2\eta}\left(\underline{\boldsymbol{\sigma}} - \frac{1}{3}\mathrm{tr}\left(\underline{\boldsymbol{\sigma}}\right)\underline{\mathbf{I}}\right) + \frac{1}{2\mu_{\mathrm{R}}}\left(\underline{\dot{\boldsymbol{\sigma}}} - \frac{1}{3}\mathrm{tr}\left(\underline{\dot{\boldsymbol{\sigma}}}\right)\underline{\mathbf{I}}\right) \qquad [96]$$

where $\underline{\boldsymbol{\sigma}} = \underline{\boldsymbol{\sigma}}^{v} = \underline{\boldsymbol{\sigma}}^{e}$ and $\underline{\boldsymbol{\varepsilon}} = \underline{\boldsymbol{\varepsilon}}^{v} + \underline{\boldsymbol{\varepsilon}}^{e}$. This time-dependent rheological law is the constitutive law of a linear Maxwell solid.

A few simple illustrations of the behavior of a Maxwellian body will illustrate the physical meaning of equation [96] (see also Ribe, this volume). First, we can consider the case where stress and strain are simple time-dependent sinusoidal functions with frequency ω (i.e., $\underline{\boldsymbol{\sigma}} = \underline{\boldsymbol{\sigma}}_0 \exp(\mathrm{i}\omega t)$ and $\underline{\boldsymbol{\varepsilon}} = \underline{\boldsymbol{\varepsilon}}_0 \exp(i\omega t)$). The solution to this problem can then be used to solve other time-dependent problems by the Fourier or Laplace transforms. Equation [96] becomes

$$\underline{\boldsymbol{\varepsilon}}_0 = \frac{1}{9K}\mathrm{tr}\left(\underline{\boldsymbol{\sigma}}_0\right)\underline{\mathbf{I}} + \frac{1}{2\mu_{\mathrm{R}}}\left(1 - \frac{i}{\omega\tau}\right)\left(\underline{\boldsymbol{\sigma}}_0 - \frac{1}{3}\mathrm{tr}\left(\underline{\boldsymbol{\sigma}}_0\right)\underline{\mathbf{I}}\right) \qquad [97]$$

where $\tau = \eta/\mu_{\mathrm{R}}$ is the Maxwell time, eqn [1]. This equation can be compared to eqn [91] and shows that the solution of a viscoelastic problem is formally equivalent to that of an elastic problem with a complex elastic rigidity. This is called the correspondence principle.

We can also solve the problem of a purely 1-D Maxwellian body (only σ_{zz} and ϵ_{zz} are nonzero), submitted to a sudden load $\sigma_{zz} = \sigma_0 H(t)$ (where H is the Heaviside distribution (Heaviside, 1850–1925)) or to a sudden strain $\epsilon_{zz} = \epsilon_0 H(\mathrm{t})$. The solutions are, for $t \geq 0$,

$$\epsilon_0 = \frac{1}{3k\mu_{\mathrm{R}}}\sigma_0 + \frac{1}{3\eta}\sigma_0 t \qquad [98]$$

and

$$\sigma_0 = 3k\mu_{\mathrm{R}} \exp\left(-k\frac{t}{\tau}\right)\epsilon_0 \qquad [99]$$

respectively, with $k = 3K/(3K + \mu_{\mathrm{R}})$. In both cases, the instantaneous elastic response is followed by a viscous flow. In the first case, the finite elastic deformation is followed by a steady flow. In the second case, the initial elastic stresses are then dissipated by viscous relaxation over a time constant, τ/k. This time constant is different from the Maxwell time constant as both deviatoric and nondeviatoric stresses are present. For mantle material, the time τ/k would however be of the same order as the Maxwell time constant τ (in the midmantle, $K \approx 2\mu_{\mathrm{R}} \approx 200$ GPa).

From equation [96], we can now understand what rheology must be used for a compressible viscous mantle. For phenomena that occur on time constants much larger than the Maxwell time, the deviatoric stresses can only be supported by the viscosity. As a typical viscosity for the deep mantle is in the range 10^{19}–10^{22} Pa s (see **Sections 7.02.4** and **7.02.6**), the appropriate Maxwell times are in the range 30 year–30 kyr, much shorter than those of convection. By contrast, the isotropic stress remains only supported by elasticity in the approximation where the bulk viscosity ζ is infinitely large. The appropriate rheology for mantle convection is therefore given by

$$\underline{\dot{\boldsymbol{\varepsilon}}} = \frac{1}{9K}\mathrm{tr}\left(\underline{\dot{\boldsymbol{\sigma}}}\right)\underline{\mathbf{I}} + \frac{1}{2\eta}\left(\underline{\boldsymbol{\sigma}} - \frac{1}{3}\mathrm{tr}\left(\underline{\boldsymbol{\sigma}}\right)\underline{\mathbf{I}}\right) \qquad [100]$$

This equation is simultaneously a rheology equation for the deviatoric stress and an EoS for the isotropic stress. Using $P=-\mathrm{tr}(\boldsymbol{\sigma})/3$, the stress tensor verifies

$$\underline{\boldsymbol{\sigma}}=-P\underline{\mathbf{I}}+2\eta\left(\underline{\dot{\boldsymbol{\varepsilon}}}-\frac{\dot{P}}{3K}\underline{\mathbf{I}}\right) \quad [101]$$

This equation is intrinsically a viscoelastic equation that can be replaced by a purely viscous equation plus an EoS:

$$\underline{\boldsymbol{\sigma}}=-P\underline{\mathbf{I}}+2\eta\left(\underline{\dot{\boldsymbol{\varepsilon}}}-\frac{1}{3}\mathrm{tr}\left(\underline{\dot{\boldsymbol{\varepsilon}}}\right)\underline{\mathbf{I}}\right) \quad [102]$$

$$\mathrm{tr}\left(\underline{\boldsymbol{\varepsilon}}\right)=\frac{P}{K} \quad [103]$$

Equation [102] is therefore the appropriate limit of eqn [94] for slow deformation, when $\zeta=+\infty$ and when isotropic compaction is resisted by the elastic stresses.

The use of a Maxwellian viscoelastic body to represent the mantle rheology on short timescale remains however rather arbitrary. Instead of summing the elastic and viscous deformations for the same stress tensor, another linear viscoelastic body could be obtained by partitioning the total stress into elastic and viscous components for the same strain rate. Basically, instead of having the elasticity and the viscosity added like a spring and a dashpot in series (Maxwell rheology), this Kelvin–Voigt rheology would connect in parallel a viscous dashpot with an elastic spring (Kelvin, 1824–1907; Voigt, 1850–1919). Of course, further degrees of complexity could be reached by summing Maxwell and Voigt bodies, in series or in parallel. Such models have sometimes been used for the Earth, but the data that could support or dismiss them were scarce (Ross and Schubert, 1986; Yuen et al., 1986). However, the unprecedented precision of the observations of large-scale deformations that followed the magnitude 9 seisms of Indonesia (Aceh earthquake, 2004) and Japan (Tohoku earthquake, 2011) should soon allow to reevaluate and constrain the transient rheology of the mantle (e.g., Trubienko et al., 2013).

7.02.3.2.4 Nonlinear rheologies

Even without elasticity and bulk viscosity, the assumption of a linear Newtonian rheology for the mantle is problematic. The shear viscosity cannot be a direct function of velocity since this would contradict the necessary Galilean invariance of material properties. However, since the shear viscosity is a scalar, it could be any function of the scalar invariants of the stress tensor. There are three invariants of the stress tensor: its trace (related to the pressure), its determinant, and the second invariant $I_2=\left(\underline{\boldsymbol{\tau}}:\underline{\boldsymbol{\tau}}\right)/2$ (where as in eqn [30], the double dots denote tensor contraction). The viscosity could therefore be a function of $\det\left(\underline{\boldsymbol{\tau}}\right)$ and I_2 in addition to the thermodynamic variables P and T and the mineralogical variables, for example, composition, grain size, and fabrics. Because viscosity must be a positive quantity, its stress dependence is assumed to be related to I_2 (which is always positive) and not to $\det\left(\underline{\boldsymbol{\tau}}\right)$ (which can change sign).

The main mechanism of solid-state deformation pertinent to mantle conditions (excluding the brittle and plastic deformations) is either diffusion creep or dislocation creep (see Poirier, 1991). In the first case, finite deformation is obtained by summing the migrations of individual atoms exchanging their positions with crystalline lattice vacancies. In crystals, the average number of lattice vacancies C varies with pressure, P, and temperature, T, according to Boltzmann statistics (Boltzmann, 1844–1906),

$$C\propto\exp\left(-\frac{PV}{RT}\right) \quad [104]$$

(V is the atomic volume and R the gas constant). A mineral is composed of grains of size d with an average concentration of lattice vacancies C_0. Submitted to a deviatoric stress τ, a gradient of vacancies of order $|\nabla C|\propto(C_0/d)(\tau V/RT)$ appears due to the difference in stress regime between the faces in compression and the faces in extension ($\tau V\ll RT$) (see Poirier, 1991; Ranalli, 1995; Schubert et al., 2001; Turcotte and Schubert, 1982). This induces the flux of atoms (number of atoms per unit surface and unit time):

$$J\propto D\frac{C_0}{d}\frac{\tau}{RT} \quad [105]$$

where D is a diffusion coefficient. This flux of atoms goes from the grain faces in compression to the grain faces in extension. Along the direction of maximum compression, each crystal grain shortens by a quantity δd, which corresponds to a total transport of $d^2\delta d/V$ atoms. These atoms can be transported in a time δt by the flux J_V across the grain of section d^2 (with volume diffusion D_V). They can also be transported by grain boundary flux J_B (with grain boundary diffusion D_B) along the grain interfaces through a surface hd (h being the thickness of the grain boundary), according to

$$\frac{d^2\delta d}{V}\approx J_D d^2\delta t,\quad \text{or}\quad \frac{d^2\delta d}{V}\approx J_B dh\delta t \quad [106]$$

As $\dot{\boldsymbol{\varepsilon}}=(\delta d/\delta t)/d$, the previous equations lead to the stress–strain rate relationship:

$$\underline{\dot{\boldsymbol{\varepsilon}}}\propto\frac{C_0V}{d^2RT}\left(D_V+D_B\frac{h}{d}\right)\underline{\boldsymbol{\tau}} \quad [107]$$

These diffusion mechanisms lead to a Newtonian rheology but with a grain size dependence of the viscosity: $\eta\propto d^2$ for Nabarro–Herring creep with diffusion inside the grain (Nabarro, 1916–2006; Herring, 1914–2009) and $\eta\propto d^3$ for Coble grain boundary creep (Coble, 1928–92). The viscosity is also very strongly T-dependent not so much because of the explicit factor T in eqn [107], but because diffusion is a thermally activated process, $D\propto\exp(-E_{dif}/RT)$, where E_{dif} is an activation enthalpy of diffusion. As the diffusion increases with temperature, the viscosity in the diffusion creep regime decreases with temperature.

In the case of dislocation creep, lines or planar imperfections are present in the crystalline lattice, and macroscopic deformation occurs by collective slip motion along these imperfections, called dislocations. Instead of the grain size d for diffusion creep, the mean spacing d_d between dislocations provides the length scale. This distance is experimentally found to decrease with the intensity of the deviatoric stresses and to vary as $\approx 1/\sqrt{I_2}$. Therefore, instead of a diffusion creep with a viscosity in d^n, the resulting rheology is rather in $I_2^{-n/2}$ and is also thermally activated with an activation energy E_{dis}. Dislocation creep leads to a nonlinear regime where the equivalent

viscosity varies with the second invariant with a power $-n/2$, where n is typically of order 2,

$$\underline{\dot{\varepsilon}} \propto I_2^{n/2} \exp(-E_{\text{dis}}/RT)\underline{\tau} \qquad [108]$$

and to a viscosity also strongly decreasing with temperature (this relationship is often written, in short, $\underline{\dot{\varepsilon}} \propto \underline{\tau}^m$ with a stress exponent m of order 3 but $\underline{\tau}^m$ really means $I_2^{(m-1)/2}\underline{\tau}$).

In general, for a given stress and a given temperature, the mechanism with the smallest viscosity (largest strain rate) prevails. Whether linear (grain size-dependent) or nonlinear (stress-dependent), viscosities are also strongly dependent upon temperature, pressure, melt content, water content, mineralogical phase, and oxygen fugacity (e.g., Hirth and Kolhstedt, 1996). In Section 7.02.6.3, we will further discuss the rheological mechanisms appropriate for the Earth.

7.02.4 Physics of Convection

The complex and very general system of equations that we have discussed in Sections 7.02.2 and 7.02.3 can be used to model an infinite number of mantle flow situations. Mantle flow can sometimes be simply modeled as driven by the motion of plates (some examples are discussed in Zhong et al., Parmentier, and King, this volume). It can also be induced by compositional density anomalies (some examples are discussed in Tackley, this volume, and Forte, vol. 3). However, a fundamental cause of motion is due to the interplay between density, pressure, and temperature, and this is called thermal convection.

The phenomenon of thermal convection is common to all fluids (gas, liquid, and creeping solids), and it can be illustrated by simple experiments (see Chapter 7.03). The simplest can be done using water and an experimental setup called the shadowgraph method. Parallel light enters a transparent fluid put in a glass tank and is deflected where there are refractive index gradients due to temperature variations in the fluid. A pattern of bright regions and dark shadows is formed on a screen put on the other side of the tank. From this shadowgraph, the structure of the temperature pattern can be qualitatively assessed (see examples of shadowgraphs in Tritton, 1988).

7.02.4.1 Basic Balance

From a simple thought experiment on thermal convection, we can derive the basic dynamic balance of convection. Let us consider a volume of fluid, Ω, of characteristic size a, in which there is a temperature excess ΔT with respect to the surrounding fluid. The fluid is subject to a gravity $\mathbf{g}$, and it has an average density ρ and thermal expansivity α. The volume Ω, because of its anomalous temperature, experiences an Archimedean force, or buoyancy (Archimedes c.287–212 BC), given by

$$\mathbf{f} = -c_1 a^3 \rho \alpha \Delta T \mathbf{g} \qquad [109]$$

(c_1 is a constant taking into account the shape of Ω, e.g., $c_1 = 4\pi/3$ for a sphere of radius a). If the volume Ω is in a fluid of viscosity η, it will sink or rise with a velocity given by the Stokes law (Stokes, 1819–1903):

$$\mathbf{v}_s = -c_1 c_2 \frac{a^2 \rho \alpha \Delta T \mathbf{g}}{\eta} \qquad [110]$$

(c_2 is a drag coefficient accounting for the shape of Ω, i.e., $c_2 = 1/(6\pi)$ for a sphere).

During its motion, the volume Ω exchanges heat by diffusion with the rest of the fluid and the diffusion equation [44] tells us that a time of order

$$t_d = c_3 \frac{\rho C_p a^2}{k} \qquad [111]$$

is needed before temperature equilibration with its surroundings. During this time, the fluid parcel travels the distance $l = v_s t_d$.

A natural indication of the possibility that the parcel of fluid moves can be obtained by comparing the distance l to the characteristic size a. When $l \gg a$, that is, when the fluid volume can be displaced by several times its size, motion will be possible. On the contrary when $l \ll a$, thermal equilibration will be so rapid that no motion will occur.

The condition $l \gg a$, when v_s and t_d are replaced by the previously mentioned expressions, depends on only one quantity, the Rayleigh number Ra,

$$Ra = \frac{\rho^2 \alpha \Delta T g a^3 C_P}{\eta k} = \frac{\alpha \Delta T g a^3}{\kappa \nu} \qquad [112]$$

in terms of which motion occurs when $Ra \gg 1$ (assuming that $c_1 c_2 c_3 \approx 1$). The Rayleigh number compares the driving mechanism (e.g., the Archimedean buoyancy) to the two resistive mechanisms, the diffusion of heat, represented by κ (see eqn [43]), and the diffusion of momentum, represented by ν (see eqn [50]).

This simple balance suggests that a large nondimensional number Ra favors fluid motion. How large Ra needs to be is a question that we cannot address at this moment, but it will be discussed in Section 7.02.4.4. Convection lifts hot fluid and causes cold fluid to sink (assuming $\alpha > 0$, which is true for most fluids and for the mantle). A convective system will rapidly reach an equilibrium where all thermal heterogeneities are swept up or down (if Ra is large) or thermally equilibrated (if Ra is small), unless a forcing mechanism continuously injects new cold parcels at the top and new hot parcels at the bottom. This can be done by cooling the top surface or heating the bottom one. When a fluid is heated from the side, a lateral temperature anomaly is constantly imposed and the liquid lateral thermal equilibration is prevented. The fluid remains in motion regardless of the amplitude of the imposed temperature anomaly.

7.02.4.2 Two Simple Solutions

7.02.4.2.1 The diffusive solution

Trying to directly and exactly solve the mass, momentum, energy, and Poisson's equations and accounting for a realistic EoS would certainly be a formidable task. This complex system of equations has however two rather obvious but opposite solutions.

A steady and motionless solution is indeed possible. The assumption $\partial/\partial t = 0$ and $\mathbf{v} = 0$ satisfies the mass equation [9], the momentum equation [16] when the pressure is hydrostatic,

$$0 = -\nabla P + \rho \mathbf{g} \qquad [113]$$

and the energy equation [38] when the temperature is diffusive (using the Fourier law [42]),

$$\nabla \cdot (k \nabla T) + \rho H = 0 \qquad [114]$$

Solving analytically for the hydrostatic pressure and the diffusive temperature is trivial when H, k, $\mathbf{g}$, and ρ are uniform. For example, choosing a depth z positive downward, we get

$$P = \rho g z, \quad T = T_0 + \Delta T \frac{z}{h} + \frac{1}{2} \rho H z (h - z) \qquad [115]$$

across a conductive solution with $T = T_0$, $P = 0$ at $z = 0$, and $T = T_0 + \Delta T$ at $z = h$. Computing analytically the conductive solution in the mantle remains feasible but could be quite cumbersome if one introduces a realistic EoS and computes gravity in agreement with the density distribution using Poisson's equation [51]. In Section 7.02.4.4, we will understand why the fluid does not necessarily choose the diffusive solution.

7.02.4.2.2 The adiabatic solution

The previous diffusive solution was obtained for a steady motionless situation. However, the opposite situation where the velocities are very large also corresponds to a rather simple situation. The energy equation [38] can also be written as

$$\rho C_V T \left(\frac{\mathrm{D} \ln T}{\mathrm{D}t} - \Gamma \frac{\mathrm{D} \ln \rho}{\mathrm{D}t} \right) = -\nabla \cdot \mathbf{q} + \underline{\boldsymbol{\tau}} : \nabla \mathbf{v} + \rho H \qquad [116]$$

or

$$\rho C_P T \left(\frac{\mathrm{D} \ln T}{\mathrm{D}t} - \frac{\alpha}{\rho C_P} \frac{\mathrm{D}P}{\mathrm{D}t} \right) = -\nabla \cdot \mathbf{q} + \underline{\boldsymbol{\tau}} : \nabla \mathbf{v} + \rho H \qquad [117]$$

The right sides of these equations were previously shown to be equal to $\rho T \mathrm{D}S/\mathrm{D}t$ in eqn [39]. If we decrease the viscosity in a fluid, the convective velocity increases. The advection terms $\mathbf{v} \cdot \nabla T$, $\mathbf{v} \cdot \nabla \ln \rho$ and $\mathbf{v} \cdot \nabla P$ become then much larger than the time-dependent, diffusion, and radioactive production terms. With a low viscosity, the fluid becomes also unable to sustain large stresses. As a consequence, when convection is vigorous enough, the fluid should evolve toward a situation where

$$\begin{aligned} (\nabla \ln T)_S - \Gamma (\nabla \ln \rho)_S &= 0 \\ (\nabla \ln T)_S - \frac{\alpha}{\rho C_P} (\nabla P)_S &= 0 \end{aligned} \qquad [118]$$

Since such equations imply that the entropy is exactly conserved, $\mathrm{D}S/\mathrm{D}t = 0$, this equilibrium is called the adiabatic equilibrium. We added a subscript $[\ ]_S$ to denote the isentropic state.

Notice that, since the Grüneisen parameter is only density-dependent (see eqn [78]), density and temperature are simply related along the adiabat. For example, if the Grüneisen parameter is a constant, Γ_0 (using $q = 0$ in eqn [78]), the first of the equations [118] implies

$$T = T_0 \left(\frac{\rho}{\rho_0} \right)^{\Gamma 0} \qquad [119]$$

where T_0 and ρ_0 are two reference values. This equation implies that the adiabatic temperature increases by a factor 1.72 (e.g., from 1300 to 2230 K) from the asthenosphere ($\rho_0 \approx 3200\ \mathrm{kg\ m^{-3}}$) to the CMB ($\rho \approx 5500\ \mathrm{kg\ m^{-3}}$), if we assume $\Gamma_0 = 1$.

7.02.4.2.3 Stability of the adiabatic gradient

When a fluid is compressed, it heats up and it cools down when decompressed. This is the same physics that explains why atmospheric temperature decreases with altitude. Of course, this adiabatic effect is vanishingly small in laboratory experiments, but not always in nature.

If a parcel of fluid is rapidly moved up or down along z by a distance a, it changes its temperature adiabatically, by the quantity $a(\mathrm{d}T/\mathrm{d}z)_S$. However, the surrounding fluid will be at the temperature $a(\mathrm{d}T/\mathrm{d}z)$ where $\mathrm{d}T/\mathrm{d}z$ is just the temperature gradient, not necessarily adiabatic, of the fluid at rest. We can define ΔT_{na} as the nonadiabatic temperature $\Delta T_{\mathrm{na}} = a(\mathrm{d}T/\mathrm{d}z - (\mathrm{d}T/\mathrm{d}z)_S)$. The parcel being warmer or colder than the surroundings will rise or sink with a Stokes velocity that, rather than eqn [110], will be of order

$$\mathbf{v}_s = -c_1 c_2 \frac{a^2 \rho \alpha \Delta T_{\mathrm{na}} \mathbf{g}}{\eta} \qquad [120]$$

Since z is depth and $\mathrm{d}z$ is positive along $\mathbf{g}$, the adiabatic gradient is positive and the fluid parcel is locally unstable when the gradient in the surrounding fluid is larger (superadiabatic) than the adiabatic gradient. On the contrary, a subadiabatic gradient is stable with respect to convection. It is therefore not the total temperature difference between the top and the bottom of the fluid that drives motion, but only its nonadiabatic part.

To compare the Stokes velocity with the thermal equilibration time, we need to introduce a modified Rayleigh number

$$Ra = \frac{\alpha \Delta T_{\mathrm{na}} g_0 a^3}{\kappa \nu} \qquad [121]$$

This number is based on the nonadiabatic temperature difference in excess of the adiabatic variation imposed over the height a.

We have shown that inside a convective cell, the thermal gradient should be superadiabatic. Superadiabaticity is the source of convective instability, but vigorous convective stirring involves largely rapid adiabatic vertical motion; so much of convecting fluid is indeed adiabatic, while most of superadiabaticity is bound up in the thermal boundary layers where the vertical motion goes to zero. This mechanism suggests that an adiabatic reference background should not be such a bad assumption for the bulk of a convective fluid.

This adiabaticity hypothesis should, however, not be taken too literally (Jeanloz and Morris, 1987). In most numerical simulations, the resulting averaged geotherm can be far (a few hundred kelvins) from adiabatic (Bunge et al., 2001). First, radioactive heating, dissipation, and diffusion are never totally negligible; second, even if each fluid parcel follows its own adiabatic geotherm, the average geotherm may not correspond to any particular adiabat.

7.02.4.3 Approximate Equations

7.02.4.3.1 Depth-dependent reference profiles

We assume that the thermodynamic state is not far from a hydrostatic adiabat; thus, we choose this state as a reference.

We denote all the reference variables with an overbar, the hydrostatic pressure given by

$$\nabla\overline{P} = \overline{\rho}\,\mathbf{g} \qquad [122]$$

and adiabatic temperature and densities obeying eqn [118]:

$$\begin{aligned}\nabla\overline{T} &= \frac{\overline{\alpha}\,\overline{\mathbf{g}}}{\overline{C_p}}\overline{T}\\ \nabla\overline{\rho} &= \frac{\overline{\alpha}\,\overline{\mathbf{g}}}{\overline{C_p}\Gamma}\overline{\rho}\end{aligned} \qquad [123]$$

where all the parameters are computed along the reference geotherm and where $\overline{\mathbf{g}}$ has been solved using Poisson's equation [51] with the reference density $\overline{\rho}$.

The reference parameters are depth-dependent and usually, even for a simple EoS, cannot be analytically obtained. They can however be computed numerically from eqn [118] assuming that the Grüneisen parameter is only ρ-dependent. Using the EoS [80], all the thermodynamic quantities become functions of depth only, so that the reference profiles can be obtained by quadratures.

7.02.4.3.2 Perturbations of the hydrostatic, adiabatic solution

We can now rewrite the equations of fluid dynamics in terms of perturbations to the reference hydrostatic and adiabatic state (see also Alboussiere and Ricard, 2013; Bercovici et al., 1992; Glatzmaier, 1988; Jarvis and Mckenzie, 1980). The exercise is not easy as all the thermodynamic quantities, α, K_T, C_P, and C_V, are related to pressure, temperature, and density by complex Maxwell relations. Assuming inadvertently the constancy of any thermodynamic quantities may thus lead to inconsistencies. We introduce the perturbations, noted with a prime, and assume that $T=\overline{T}+T'$, $P=\overline{P}+P'$, $\rho=\overline{\rho}+\rho'$. As the reference state is adiabatic, we simply have $\mathcal{S}=\mathcal{S}'$. The variations of density and temperature are related by the following relations:

$$\begin{aligned}\rho' &= \left(\frac{\partial\rho}{\partial S}\right)_P\mathcal{S}' + \left(\frac{\partial\rho}{\partial P}\right)_S P'\\ T' &= \left(\frac{\partial T}{\partial S}\right)_P\mathcal{S}' + \left(\frac{\partial T}{\partial P}\right)_S P'\end{aligned} \qquad [124]$$

The partial derivatives of the previously mentioned expressions can be expressed as functions of the usual thermodynamic quantities. The $\mathcal{S}$ derivatives can be deduced after some algebra starting from the Maxwell relations implied by the enthalpy $d\mathcal{H}=T\mathrm{d}\mathcal{S}+\mathrm{d}P/\rho$ and by the Gibbs potential $\mathrm{d}\mathcal{G}=-\mathcal{S}\mathrm{d}T+dP/\rho$, and using eqns [31] and [33]. The adiabatic derivatives (at constant $\mathcal{S}$) are directly obtained from the reference profiles in eqns [122] and [123]. We finally end up with two equations that we consider valid at first order:

$$\begin{aligned}\rho' &= -\frac{\overline{\alpha\rho}\,\overline{T}}{\overline{C}_P}\mathcal{S}' + \frac{\nabla\overline{\rho}\cdot\mathbf{g}}{\overline{\rho}\|\mathbf{g}\|^2}P'\\ T' &= -\frac{\overline{T}}{\overline{C}_P}\mathcal{S}' + \frac{\overline{\alpha}\overline{T}}{\overline{\rho}\overline{C}_P}P'\end{aligned} \qquad [125]$$

Taking into account the variations of density with entropy and pressure, the momentum conservation [16] becomes

$$\overline{\rho}\frac{\mathrm{D}\mathbf{v}}{\mathrm{D}t} = -\overline{\rho}\nabla\left(\frac{P'}{\overline{\rho}}\right) + \nabla\cdot\underline{\boldsymbol{\tau}} - \frac{\overline{\mathbf{g}\alpha\rho}\,\overline{T}}{\overline{C}_P}\mathcal{S}' \qquad [126]$$

In this equation, the pressure dependence of the density has been collected with the pressure gradient. The energy conservation can take various forms (e.g., eqn [30] or [38]), but here, we use the entropy conservation [39]:

$$\overline{\rho}\,\overline{T}\frac{\mathrm{D}\mathcal{S}'}{\mathrm{D}t} = \nabla\cdot\left(\overline{k}\nabla\left(\overline{T}+T'\right)\right) + \underline{\boldsymbol{\tau}}:\nabla\mathbf{v} + \rho H \qquad [127]$$

7.02.4.3.3 Anelastic approximation

As a principle, the validity of equations cannot depend on the units in which the quantities are expressed. The laws of physics can only relate dimensionless combinations of parameters (e.g., Barenblatt, 1996). This is fundamental in fluid dynamics where a large number of quantities appear in the equations (see also Ribe, this volume). A necessary starting point is therefore to rephrase any fluid dynamics problem involving N dimensional parameters in terms of M dimensionless quantities ($N-M\leq 0$ is the number of independent physical dimensions of the problem).

Equations [126] and [127], completed with the equation of state [125] and of mass conservation, provide a differential system that controls the behavior of a nearly adiabatic and hydrostatic fluid. One cannot perform the nondimensionalization of these equations using the variable reference profiles. We must therefore introduce constant parameters, with indexes $[\,]_0$, corresponding to some typical or mantle-averaged values of the depth-dependent reference values. We introduce, for example, $\alpha_0, K_S^0, \rho_0, \mathrm{g}_0$ or C_p^0. We can use the typical values of these parameters to estimate the different terms of the convection equations.

Without approximation, the mass conservation is

$$\frac{\partial\rho'}{\partial t} + \nabla\cdot((\overline{\rho}+\rho')\mathbf{v}) = 0 \qquad [128]$$

This equation, with the momentum and energy equations, describes not only the fluid convection but also the propagation of sound waves. These waves, with velocity close to $v_\phi=\sqrt{K_S^0/\rho_0}$, travel much faster than the typical velocity of convection, at a velocity comparable to those of the P or S waves. A typical convective velocity is given by the Stokes velocity $v_s=\rho_0\alpha_0\Delta T_{na}g_0a^2/\eta_0$ (see eqn [110]). The Mach number $M=v_s/v_\phi$, ratio of the typical fluid velocity to the sound speed (Mach, 1838–1916) is very small (with the numerical values of **Table 1**, $M=1.5\times10^{-9}$). It should therefore be safe to neglect the time derivative of the density in eqn [128]; this will only filter out the sound propagation that seems a phenomenon decoupled from mantle convection.

In the divergence term of eqn [128], we can also compare $\overline{\rho}$ and ρ'. Density, temperature, and pressure variations are related by eqn [76]:

$$\frac{\rho'}{\overline{\rho}} = -\overline{\alpha}T' + \frac{P'}{K_T} \qquad [129]$$

(a relation that can also be derived from eqn [125], by eliminating $\mathcal{S}'$ and using the relation between K_S and K_T, eqn [85]). The effect of temperature on density is small, of order $\alpha_0\Delta T_{\mathrm{na}}\approx3\times10^{-2}$ (see **Table 1**). In the convecting mantle, P' is comparable to the viscous stresses $P'\approx\eta_0 v_s/a\approx\rho_0\alpha_0\Delta T_{\mathrm{na}}g_0a$ and the density variations due to the pressure fluctuation are thus

Table 1 Typical parameter values for numerical models of mantle convection

		Mantle	*Core*	*Unit*
Size	a	3×10^{6}	3×10^{6}	m
Dynamic viscosity	η_0	10^{21}	10^{-3}	Pa s
Heat capacity	$C_P^0\approx C_V^0$	1000	700	J K^{-1} kg^{-1}
Density	ρ_0	4000	11 000	kg m^{-3}
Heat cond.	k_0	3	50	$Wm^{-1}K^{-1}$
Expansivity	α_0	2×10^{-5}	10^{-5}	K^{-1}
Temperature excess	ΔT_{na}	1500	1?	K
Radioactivity prod.	H	7×10^{-11}	0?	W kg^{-1}
Gravity	g_0	9.8	5	m s^{-2}
Incompressibility	$K_T^0\approx K_S^0$	2×10^{11}	10^{12}	Pa
Kinematic viscosity	$\nu=\eta_0/\rho_0$	2.5×10^{17}	9.1×10^{-8}	m^2 s^{-1}
Thermal diff.	$\kappa=k_0/(\rho_0C_P^0)$	7.5×10^{-7}	6.5×10^{-6}	m^2s^{-1}
	$\in$	3.0×10^{-2}	$1.0\times^{-5}$	
Dissipation number	D_0	0.59	0.21	
Grüneisen parameter	Γ	1	1.3	
Rayleigh	Ra	4.2×10^{7}	2.2×10^{27}?	
Intern. Rayleigh	Ra_H	2.4×10^{10}	0?	
Prandtl	Pr	3.3×10^{23}	1.4×10^{-2}	
Grashof/Reynolds	$Gr=Re=Ra/Pr$	1.3×10^{-16}	1.6×10^{29}?	
Mach	M	1.5×10^{-9}	$\ll1$	
	M^2Pr/Ra or $D_0\epsilon/\Gamma_0$	1.7×10^{-2}	1.6×10^{-6}	

To emphasize the drastic differences between the highly viscous mantle and a real liquid (in which shear waves do not propagate), we added estimates for the core assuming that core convection is so efficient that only 1 K of nonadiabatic temperature difference can be maintained across it. Notice that with only 1 K of temperature difference, the Rayleigh number of the fluid core would already reach 10^{27}! A Stokes velocity is not appropriate to estimate the velocity in the core or its Mach number.

$$\frac{\rho'}{\overline{\rho}}\approx\alpha_0\Delta T_{na}\frac{D_0}{\Gamma_0}\frac{C_p^0}{C_v^0}=M^2\frac{Pr}{Ra}\frac{K_s^0}{K_T^0} \qquad [130]$$

In these estimates, we used either a nondimensional dissipation number

$$D_0=\frac{\alpha_0 g_0 a}{C_p^0} \qquad [131]$$

and a Grüneisen parameter, eqn [77], or the Rayleigh number Ra defined in eqn [121] and the Prandtl, Pr, number:

$$Pr=\frac{\eta_0 C_P^0}{k_0}=\frac{\nu_0}{\kappa_0} \qquad [132]$$

The ratios C_p^0/C_v^0 and K_s^0/K_T^0 are of order 1. The dissipation number compares the natural scale of temperature variations [123] with the layer thickness, a. The dissipation number D_0 is around 0.5 for the Earth's mantle. For any laboratory experiment, this number would be infinitely small; only geophysical or astrophysical problems have large dissipation numbers. The physical meaning of the Rayleigh number as a measure of convective vigor has already been discussed. The Prandtl number (Prandtl, 1975–53) compares the two diffusive processes: namely, the diffusion of momentum and heat. As $D_0\approx0.5$ and $\Gamma_0\approx1$, the density variations due to non hydrostatic pressure variations are smaller but of the same order as those due to temperature variations. There are both of a few percent so that $\rho'\ll\overline{\rho}$. The mass conservation can therefore be approximated in what is called the anelastic approximation (Ogura and Phillips, 1962; Spiegel and Veronis, 1960):

$$\nabla\cdot(\overline{\rho}\mathbf{v})=0 \qquad [133]$$

The anelastic approximation is sometimes called small Mach number approximation. However, the anelastic approximation really requires not only a small M but also a small M^2Pr/Ra (Bercovici et al., 1992). A planet could have a very low Mach number but so large a Prandtl number that the anelastic approximation would not be valid. Similarly, a planet can have a low Reynolds number (creeping convection) with a very large Rayleigh number (chaotic convection).

The differential system eqns [126], [127], [125], and [133] constitute the anelastic approximation of convection that has been used for atmospheric convection as well as for liquid and solid convection, including mantle and core convection. A large difference between gases and liquids or solids is that the former have $\alpha T\approx1$ (and exactly 1 for ideal gas) and the latter have $\alpha T\ll1$. This allows a further simplification of the equation, the anelastic liquid approximation.

7.02.4.3.4 Anelastic liquid approximation

Using the relation between entropy, pressure, and temperature (eqn [125]), and following the same scaling as before, where non hydrostatic pressure and viscous stresses have the same magnitude, yields

$$T'=\frac{\overline{T}}{\overline{C}_P}\mathcal{S}'+\frac{\overline{\alpha}\overline{T}}{\overline{\rho}\overline{C}_P}P'\approx\frac{\overline{T}}{\overline{C}_P}S'+D_0\overline{\alpha}\overline{T}T' \qquad [134]$$

As $D_0\approx0.5$ and $\overline{\alpha}\overline{T}\ll0$, the last term is negligible, and for the mantle or the core, we can identify $\overline{T}\mathcal{S}'$ and $\overline{C}_PT'$. Using $\overline{T}\mathrm{d}\mathcal{S}'=\mathrm{d}(\overline{T}\mathcal{S}')-\mathcal{S}'\mathrm{d}\overline{T}=\mathrm{d}(\overline{C}_PT')-\overline{C}_P(T'/\overline{T})\mathrm{d}\overline{T}$, we can express the anelastic liquid equations by

$$\begin{aligned}&\nabla\cdot(\rho'\mathbf{v})=0\\&\overline{\rho}\frac{\mathrm{D}\mathbf{v}}{\mathrm{D}t}=-\overline{\rho}\nabla\left(\frac{P'}{\overline{\rho}}\right)+\nabla\cdot\underline{\boldsymbol{\tau}}-\overline{\rho}\,\overline{\mathbf{g}}\overline{\alpha}T'\\&\overline{\rho}\frac{\mathrm{D}\overline{C}T'}{\mathrm{D}t}=\overline{\rho}\,\overline{\mathbf{g}}\overline{\alpha}v_gT'+\nabla\cdot\left(\overline{k}\nabla\left(\overline{T}+T'\right)\right)+\underline{\boldsymbol{\tau}}:\nabla\mathbf{v}+\rho H\end{aligned} \qquad [135]$$

where v_g is the velocity along $\overline{\mathbf{g}}$.

In the anelastic liquid approximation, C_P and C_V are equal, as well as K_S and K_T (their relative difference is $\Gamma_0\alpha_0\Delta T_{na}\ll1$

according to eqn [85]); we thus simply label the heat capacity and incompressibility as $\overline{C}$ and $\overline{K}$.

7.02.4.3.5 Nondimensionalization

To nondimensionalize the equations, we perform the change of variables and parameters

$$\begin{aligned} &x,\ v,\ t,\ P',\ T',\ \tau \rightarrow a\widetilde{x}, v_s\widetilde{v},\ \frac{a}{v_s}\widetilde{t},\ \frac{\eta_0 v_s}{a}\widetilde{P}',\ \Delta T_{na}\widetilde{T}',\ \frac{\eta_0 v_S}{a}\widetilde{\tau} \\ &\nabla,\ \overline{\rho},\ \overline{\alpha},\ \overline{C},\ \overline{K},\ \overline{k},\ \overline{g} \rightarrow \frac{1}{a}\widetilde{\nabla},\ \rho_0\widetilde{\rho},\ \alpha_0\widetilde{\alpha},\ C_0\widetilde{C},\ K_0\widetilde{K},\ k_0\widetilde{k},\ g_0\widetilde{g} \end{aligned} \quad [136]$$

where v_s is the Stokes velocity $a^2 g_0\rho_0\alpha_0\ \Delta T_{na}/\eta_0$. We obtain the nondimensionalized fluid anelastic equations:

$$\begin{aligned} &\nabla\cdot(\widetilde{\rho}\mathbf{v}) = 0 \\ &\widetilde{\rho}\frac{Ra}{Pr}\frac{D\mathbf{v}}{Dt} = -\widetilde{\rho}\nabla\left(\frac{P}{\widetilde{\rho}}\right) + \nabla\cdot\underline{\tau} - \widetilde{\rho}\widetilde{\mathbf{g}}\widetilde{\alpha}T \\ &\widetilde{\rho}\frac{D\widetilde{C}\widetilde{T}}{Dt} = \widetilde{\rho}\widetilde{g}\widetilde{\alpha}v_g D_0 T + \frac{1}{Ra}\nabla\cdot\left(\widetilde{k}\nabla\left(\widetilde{T}+T\right)\right) \\ &\qquad + D_0\underline{\tau}:\nabla\mathbf{v} + \frac{1}{Ra}\widetilde{\rho}\frac{\rho_0 Ha^2}{k_0\Delta T_{na}} \end{aligned} \quad [137]$$

For simplicity, we only kept the tilde sign on the reference nondimensionalized profiles, and we omitted the tilde and the primes on all the other quantities. In the momentum equation, the nondimensionalized stress tensor in the Newtonian case (without bulk viscosity) is

$$\underline{\tau} = \widetilde{\eta}\left(\nabla\mathbf{v} + [\nabla\mathbf{v}]^t\right) - \frac{2}{3}\widetilde{\eta}\nabla\cdot\mathbf{v} \quad [138]$$

The formalism is already so heavy that we have not included the self-gravitational term. This term is not negligible at long wavelengths (see Section 7.02.2.5.2). To account for this term, we should have added on the right side of momentum equation of [137] a term $-\widetilde{\rho}\nabla\widetilde{\psi}$. where the perturbed gravitational potential due to the departure of the density from the reference profile satisfies Poisson's equation:

$$\nabla^2\widetilde{\psi} = 3\widetilde{\rho}\frac{\rho_0}{<\rho>}\left(\frac{1}{\widetilde{K}}\frac{D_0}{\Gamma_0}P' - \widetilde{\alpha}T'\right) \quad [139]$$

where $<\rho>$ is the average density of the Earth.

7.02.4.3.6 Dimensionless numbers

The ratio Ra/Pr is also called the Grashof number $Gr = \alpha_0 g_0 \Delta T_{na} a^3/\nu_0^2$ (Grashof, 1826–93). Introducing the Stokes velocity v_S, eqn [110], this number can also be written as $v_S a/\nu_0$ and could be called the Reynolds number, Re, of the flow (Reynolds, 1842–1912). Using one or the other names depends on the quantities that are best known. For example, if the velocity V_0 is a parameter imposed by a boundary condition, using it to perform the nondimensionalization and characterizing the system by the Reynolds number $V_0 a/\nu_0$ would be more natural than using the Grashof number. If a thermal structure is imposed by a velocity boundary condition (e.g., by the thickening of the oceanic lithosphere with age), it would seem natural to introduce a Péclet number $V_0 a/\kappa_0$ (Péclet, 1793–1857), which is nothing more than the Rayleigh number of the flow if the velocity is imposed by the convection.

The Rayleigh and Prandtl numbers can be estimated in different ways. In fact, the only parameter difficult to know is the viscosity. In most textbooks, the value of 10^{21} Pa s, first proposed by Haskell (1937), is given with a unanimity that hides very large uncertainties and most probably a large geographic variability. The mantle viscosity and its depth dependence can be constrained by postglacial rebound, geoid, true polar wander, change of flattening of the Earth, and plate force balance models or extrapolated from laboratory measurements. An increase of viscosity with depth, either one or two orders of magnitudes, is likely, with an asthenosphere significantly less viscous (10^{19} Pa s) at least under oceanic plates and a lower mantle probably around 10^{22} Pa s (see details in Section 7.02.6.1). It is impossible to give justice to all the papers on this subject, but some geodynamic estimates of mantle viscosity can be found in, for example, Peltier (1989), Sabadini and Yuen (1989), Lambeck and Johnston (1998), and Ricard et al. (1993a). Whatever the value of the real viscosity, the ratio Ra/Pr is so small that inertia plays no role in the mantle and the left side of the momentum equation in the anelastic approximation [137] can safely be set to zero (see numerical values in Table 1).

7.02.4.3.7 Boussinesq approximation

As was expected from Section 7.02.2.5.3, where we had shown that the dissipation and the adiabatic terms balance each other in a statistical steady-state regime, these terms are proportional to the same dissipation number D_0. Although D_0 is not so small, most of the physics of mantle convection, except for the additional adiabatic temperature gradient, is captured with models where D_0 is arbitrarily set to zero. The Boussinesq approximation captures most of the mantle dynamics due to the fact that a huge ΔT_{na} is maintained across the top and bottom boundary layers. The other sources and sinks of energy and density anomaly (viscous dissipation and adiabatic heating) are therefore always small (Bercovici et al., 1992; Glatzmaier, 1988; Jarvis and Mckenzie, 1980; Tackley, 1996).

In the Boussinesq approximation, $D_0 = 0$, the reference density and temperature as well as the heat capacity and the thermal expansivity become constants (according to eqn [123], $\nabla\widetilde{T} \approx D_0\widetilde{T}$ and $\nabla\widetilde{\rho} \approx (D_0/\Gamma_0)\widetilde{\rho}$)). The nonadiabatic temperature increase ΔT_{na} becomes simply the total temperature increase ΔT. This approximation is of course excellent for laboratory-scale experiments where effectively $D_0 \ll 1$. The EoS [130] indicates that the density variations due to pressure are D_0/Γ_0 smaller than those due to temperature and are therefore negligible and the fluid dynamics equations become

$$\begin{aligned} &\nabla\cdot\mathbf{v} = 0 \\ &\frac{Ra}{Pr}\frac{D\mathbf{v}}{Dt} = -\nabla P + \nabla\cdot\left(\widetilde{\eta}\left(\nabla\mathbf{v} + [\nabla\mathbf{v}]^t\right)\right) - \widetilde{\mathbf{g}}T \\ &\frac{DT}{Dt} = \frac{1}{Ra}\nabla\cdot\left[\widetilde{k}\nabla T\right] + \frac{1}{Ra}\frac{\rho_0 Ha^2}{k_0\Delta T} \end{aligned} \quad [140]$$

Here again, as in the mantle, $Gr = Ra/Pr \ll 1$, the inertia in the momentum equation [140] can be neglected. The self-gravitational term $-\widetilde{\rho}\widetilde{\nabla}\widetilde{\psi}$ should be added to the momentum equation (second equation of eqn [140]) for large-scale simulations, the gravitational potential being solution of eqn [139] where only the thermal part of the density variation needs to be taken into account.

The nondimensionalized gravity $\widetilde{\mathbf{g}}$, viscosity $\widetilde{\eta}$, and thermal conductivity $\widetilde{k}$ are often assumed constant (e.g., $\widetilde{k} = \widetilde{\eta} = 1$ and $\widetilde{\mathbf{g}} = \widetilde{\mathbf{e}}_g$, a constant unit vector), but they can be variable. The

physical behavior of a large Rayleigh number is obvious in eqn [140]. When $Ra \to \infty$, and in the absence of heat sources, the temperature becomes a purely advected and conserved quantity, $\mathrm{D}T/\mathrm{D}t = 0$.

7.02.4.3.8 Internal heating

In the nondimensionalization, we assumed that the nonadiabatic temperature ΔT_{na} and the radioactive sources are two independent quantities. Of course, in the case where the mantle is only heated from within, the excess temperature is not a free parameter but must result from the heat source and from the flow itself. In the nondimensionalization, we can replace ΔT_{na} by $\rho_0\, Ha^2/k_0$ in such a way that the radioactive heat source of the anelastic or Boussinesq energy equations, [137] and [140], are simply $1/Ra$. This choice requires the introduction of a somewhat different Rayleigh number, the internally heated Rayleigh number (Roberts, 1967),

$$Ra_{\mathrm{H}} = \frac{\alpha_0 H \rho_0^3 g_0 a^5 C_P^0}{\eta_0 k_0^2} \qquad [141]$$

7.02.4.3.9 Change of nondimensionalization

We use a Stokes velocity to nondimensionalize the equations. We have therefore introduced a velocity V_0 of order of 300 m year^{-1} and a time $a/V_0 = 10\,000$ year (see Table 1). This is certainly very fast and short compared with geologic scales. Most physical and geophysical textbooks (e.g., Schubert et al., 2001) use instead a diffusive time $t_{\mathrm{D}} = \rho C_P a^2/k_0$ and velocity a/t_{D}. This is perfectly valid, but Table 1 shows that the diffusive time and velocity amount to $t_{\mathrm{D}} = 400$ byr and $V_{\mathrm{D}} = 7 \times 10^{-6}$ m year^{-1}. These values would on the contrary be very slow and long compared with geologic scales. A nondimensionalization using a diffusive timescale leads to the anelastic equations

$$\begin{aligned}
&\nabla\cdot(\tilde{\rho}\mathbf{v}) = 0\\
&\tilde{\rho}\frac{1}{Pr}\frac{\mathrm{D}\mathbf{v}}{\mathrm{D}t} = -\tilde{\rho}\nabla\left(\frac{P}{\tilde{\rho}}\right) + \nabla\cdot\underline{\boldsymbol{\tau}} - Ra\tilde{\rho}\tilde{\mathbf{g}}\tilde{\alpha}T\\
&\tilde{\rho}\frac{\mathrm{D}\tilde{C}T}{\mathrm{D}t} = \tilde{\rho}\tilde{\mathbf{g}}\tilde{\alpha}D_0\upsilon_{\mathrm{g}}T + \nabla\cdot\left(\tilde{k}\nabla\left(\tilde{T}+T\right)\right) + \frac{D_0}{Ra}\underline{\boldsymbol{\tau}}:\nabla\mathbf{v} + \frac{\rho_0 Ha^2}{k_0\Delta T_{\mathrm{na}}}
\end{aligned} \qquad [142]$$

and the Boussinesq equations

$$\begin{aligned}
&\nabla\cdot(\mathbf{v}) = 0\\
&\frac{1}{Pr}\frac{\mathrm{D}\mathbf{v}}{\mathrm{D}t} = -\nabla P + \nabla\cdot\underline{\boldsymbol{\tau}} - Ra\tilde{\mathbf{g}}T\\
&\frac{\mathrm{D}T}{\mathrm{D}t} = \nabla\cdot\left(\tilde{k}\nabla T\right) + \frac{\rho_0 Ha^2}{k_0\Delta T}
\end{aligned} \qquad [143]$$

Notice that the Ra number appears in different places than in eqns [137] and [140]. Of course, after their appropriate changes of variables, the dimensional solutions are the same.

7.02.4.4 Linear Stability Analysis for Basally Heated Convection

To understand why the static conductive solution is not necessarily the solution chosen by the system, one can test the stability of a static state. This is what physicists call a study of stability analysis (see also Ribe, this volume). It consists of substituting into the basic equations the known static solution plus an infinitely small perturbation and checking whether or not this perturbation amplifies, decreases, or propagates. It is only when perturbations decrease in amplitude that the tested equilibrium is established to be stable.

To illustrate this analysis, we use the Boussinesq approximation, with constant viscosity and conductivity, neglecting inertia and without internal heating. The nondimensionalized equations [140] yield

$$\begin{aligned}
\nabla\cdot\mathbf{v} &= 0\\
-\nabla P + \nabla^2\mathbf{v} - T\mathbf{e}_z &= 0\\
\frac{\partial T}{\partial t} + \mathbf{v}\cdot\nabla T &= \frac{1}{Ra}\nabla^2 T
\end{aligned} \qquad [144]$$

($\mathbf{e}_z$ is the unit vector parallel to $\mathbf{g}$). The steady diffusive nondimensional temperature solution is $T = z$, and we test a solution of the form $T = z + \delta T$. The temperature boundary condition, $T = 0$ on top and $T = 1$ at the bottom, requires that δT vanishes for $z = 0$ and $z = 1$. In static conductive state, the velocity is zero, thus the velocity $\mathbf{v}$ induced by the perturbation δT will be infinitesimally small. In the nonlinear term, we can approximate $\mathbf{v}\cdot\nabla T = \mathbf{v}\cdot\nabla(z + \delta T) \approx \upsilon_z$. With this approximation, the equations are linear and we can find a solution in the form of a plane wave.

For a fluid confined between $z = 0$ and $z = 1$ where the temperature is imposed (i.e., where $\delta T = 0$) and unbounded in the x-direction, a solution, $\delta T = \theta(t)\sin(\pi z)\sin(kx)$, is appropriate and satisfies the boundary conditions. This solution is 2-D, has a single mode in the z-direction, and is periodic in x with wavelength $\lambda = 2\pi/k$. More complex patterns could be tried but the mode we have chosen would destabilize first (see Ribe, this volume). It is then straightforward to deduce that for such a thermal anomaly, the energy equation imposes a vertical velocity

$$\upsilon_z = -\left(\dot{\theta} + \frac{(k^2+\pi^2)}{Ra}\theta\right)\sin(\pi z)\sin(kx) \qquad [145]$$

From mass conservation and assuming top and bottom free-slip conditions, the x-component of the velocity must be

$$\upsilon_x = -\frac{\pi}{k}\left(\dot{\theta} + \frac{(k^2+\pi^2)}{Ra}\theta\right)\cos(\pi z)\cos(kx) \qquad [146]$$

When the velocity and the temperature are introduced in the momentum equations, the pressure eliminated, the time evolution of the temperature perturbation is found

$$\dot{\theta} = \theta\left(\frac{k^2}{(\pi^2+k^2)^2} - \frac{(\pi^2+k^2)}{Ra}\right) \qquad [147]$$

For any wave number k, a small enough Rayleigh number corresponds to a stable solution, $\dot{\theta}/\theta < 0$. When the Rayleigh number is increased, the temperature perturbation with wave number k becomes unstable at the threshold Rayleigh number

$$Ra = \frac{(\pi^2+k^2)^3}{k^2} \qquad [148]$$

This $Ra(k)$ curve is plotted in Figure 3. This curve has a minimum when

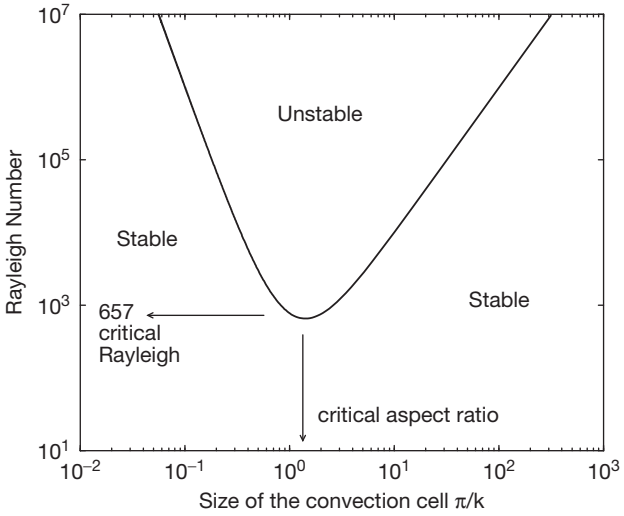

Figure 3 Critical Rayleigh number as a function of the half wavelength π/k (the size of the convection cells). Above this curve, convection occurs with a whole range of unstable wavelengths. Below this curve, the conductive temperature is stable since temperature perturbations of any wavelength decrease. When the Rayleigh number is increased, the first unstable wavelength corresponds to a convection cell of aspect ratio $\sqrt{2}$ and a critical Rayleigh number of 657.

$$k=\frac{\pi}{\sqrt{2}},\quad Ra_c=\frac{27}{4}\pi^4\approx 657 \qquad [149]$$

What can be interpreted as the size of one convective cell is π/k since one wavelength corresponds to two counterrotating cells. The critical cell has an aspect ratio, width over height, of $\sqrt{2}$.

A Rayleigh number of 657 is the critical Rayleigh number for the onset of convection in a layer heated from below with free-slip boundary conditions. As soon as $Ra > Ra_c$, there is a wave number interval over which convection begins. Of course, when convection grows in amplitude, the marginal stability solution becomes less and less pertinent as the assumption that $\mathbf{v}\cdot\nabla\delta T \ll \mathbf{v}\cdot\nabla z$ becomes invalid. The existence of a critical Rayleigh number and its value, here obtained from a marginal stability analysis, have also been obtained in a more general case, where finite amplitude perturbations (instead of infinitesimal ones) are added to the stationary solution. Below the critical Rayleigh number, any finite perturbation can be shown to decay eventually to zero (e.g., Joseph, 1966).

7.02.4.5 Road to Chaos

In Cartesian geometry, when the Rayleigh number reaches its critical value, convection starts and forms rolls. When the Rayleigh number is further increased, complex series of convection patterns can be obtained, first stationary, then periodic, and finally chaotic (see Chapter 7.03). Using the values of Table 1, the critical Rayleigh number of the mantle would be attained for a nonadiabatic temperature difference between the surface and the CMB of only 0.025 K. The mantle Rayleigh number is several orders of magnitude higher than critical and the mantle is in a chaotic state of convection.

Figure 4 shows a stationary convection pattern at $Ra=10^5$ and three snapshots of numerical simulation of convection at higher Rayleigh number. The color scale has been chosen differently in each panel to emphasize the thermal structures that decrease in length scale with Ra. This view is somewhat misleading since all the thermal anomalies become confined in a top cold boundary layer and in a hot bottom one at large Rayleigh numbers. Most of the interior of the cell becomes just isothermal (or adiabatic when anelastic equations are used). The various transitions of convection as the Rayleigh number increases will be discussed in other chapters of this treatise (see Chapters 7.03, 7.04, 7.05).

7.02.5 Introduction to Physics of Multicomponent and Multiphase Flows

The mantle is not a simple homogeneous material. It is made of grains of variable bulk composition and mineralogy and contains fluids, magma, and dissolved volatiles. Discussion of multicomponent and multiphase flows could deal with solids, liquids, or gases; include compressibility or not; and consider elastic, viscous, or more complex rheology. For each combination of these characteristics, a geophysical application is possible. Here, we will restrict the presentation to viscous creep models (i.e., without inertia), where the various components are treated with continuous variables (i.e., each component is implicitly present everywhere). We do not consider approaches where the various components are separated by moving and deformable interfaces. Our presentation excludes cases where the problem is to match properties at macroscopic interfaces between regions of different but homogeneous compositions.

We will focus on two cases. The first will be when all the components are perfectly mixed in variable proportions. This corresponds to the classical chemical approach of multiple components in a solution. This will provide some tools to understand mantle phase transitions and the physics of chemical diffusion and mixing. We will be rather formal and refer to other chapters of this treatise for the applications and illustrations (e.g., Parmentier and Tackley, this volume). Our goal is to explain why and when the advection–diffusion equation can be used in mantle dynamics. The irreversible thermodynamics of multicomponent flows is discussed in various classical books (e.g., de Groot and Mazur, 1984; Haase, 1990). However, as usual with geophysical flows, the mantle has many simplifications and a few complexities that are not necessarily well documented in these classical textbooks.

The second case will be for two-phase flows in which the two phases are separated by physical interfaces that are highly convolved and with spatial characteristics much smaller than the typical size of geodynamic models. This is typically the case where magma can percolate through a compacting matrix (see also Dingwell, vol. 2). This approach was used to model melt extraction and core–mantle interaction (McKenzie, 1984; Scott and Stevenson, 1984). Magma migration has also been treated in a large number of publications where solid and magma are considered as separated in studies of dike propagation through hydraulic fracturing (e.g., Lister and Kerr, 1991) or where fusion is parameterized in some way (e.g., Choblet and Parmentier, 2001; Ito et al., 1999). We do not discuss these latter approaches.

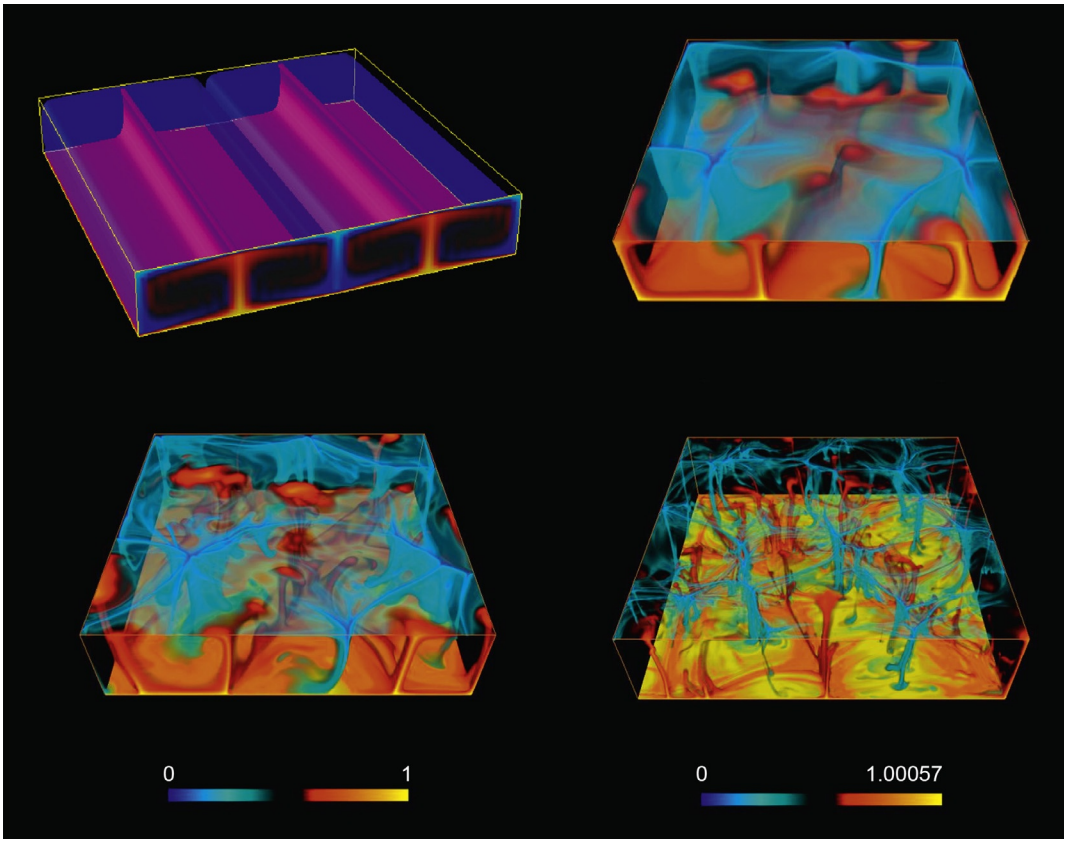

Figure 4 Convection patterns of a fluid heated from below at Rayleigh number 10^5, 10^6, 10^7, and 10^8. The temperature color bars range from 0 (top boundary) to 1 (bottom boundary). The Boussinesq approximation was used (numerical simulations by F. Dubuffet). The increase in Rayleigh number corresponds to a decrease of the boundary layer thicknesses and the width of plumes. Only in the case of the lowest Rayleigh number (top left) is the convection stationary with cells of aspect ratio $\approx \sqrt{2}$ as predicted by marginal stability. For higher Rayleigh number, the patterns are highly time-dependent.

7.02.5.1 Fluid Dynamics of Multicomponent Flows in Solution

7.02.5.1.1 Mass conservation in a multicomponent solution

If we want to study the evolution of major or trace element concentration in the convecting mantle, we can consider the mantle, instead of a homogeneous fluid, as a solution of various components i in volumetric proportions ϕ_i (with $\sum_i \phi_i = 1$) having the densities ρ_i and velocities $\mathbf{v}_i$ (and later, thermal expansivities α_i and heat capacities $C_p^i \ldots$).

Using a mass balance very similar to what we had discussed for a homogeneous fluid, we obtain a mass conservation equation of the form

$$\frac{\partial(\phi_i \rho_i)}{\partial t} + \boldsymbol{\nabla}\cdot(\phi_i \rho_i \mathbf{v}_i) = \Gamma_i \qquad [150]$$

where Γ_i is the rate of mass production of component i. This rate of mass production is zero if no reactions produce the component i.

In the fluid, the average density is

$$\overline{\rho} = \sum_i \phi_i \rho_i \qquad [151]$$

and various average velocities can be defined (weighted by the mass, the volume, the number of moles, etc., of each component i). In this section, we introduce the barycentric velocity, $\mathbf{v}_\mathrm{b}$ (velocity of the center of mass), defined by

$$\mathbf{v}_\mathrm{b} = \frac{\sum \phi_i \rho_i \mathbf{v}_i}{\overline{\rho}} \qquad [152]$$

The average mass conservation can be obtained by summing the equations of component conservation [150],

$$\frac{\partial \overline{\rho}}{\partial t} + \nabla\cdot(\overline{\rho}\mathbf{v}_\mathrm{b}) = 0 \qquad [153]$$

since the sum of the rates of mass production is zero:

$$\sum_i \Gamma_i = 0 \qquad [154]$$

In equation [150], instead of the various component velocities $\mathbf{v}_i$, we can introduce the barycentric velocity $\mathbf{v}_\mathrm{b}$ and the diffusive flux of the component i with respect to this average flow,

$$\frac{\partial(\phi_i \rho_i)}{\partial t} + \boldsymbol{\nabla}\cdot(\phi_i \rho_i \mathbf{v}_\mathrm{b}) = -\boldsymbol{\nabla}\cdot\mathbf{J}_i + \Gamma_i \qquad [155]$$

where we define the diffusive flux, $\mathbf{J}_i$, by

$$\mathbf{J}_i = \phi_i \rho_i (\mathbf{v}_i - \mathbf{v}_\mathrm{b}) \qquad [156]$$

By definition of the barycentric velocity eqn [152], the diffusive flows sum to zero:

$$\sum_i \mathbf{J}_i = 0 \qquad [157]$$

Diffusive transport of mass is nothing more than the advective transfer by component velocities relative to the average barycentric velocity. We will show later in simple cases that the diffusive fluxes are driven by concentration gradients (de Groot and Mazur, 1984; Haase, 1990; Woods, 1975).

If we introduce the mass fraction $C_i = \phi_i \rho_i / \overline{\rho}$ (in kg of i per kg of mixture), we can easily show from eqns [153] and [155] that

$$\overline{\rho} \frac{\mathrm{D}C_i}{\mathrm{D}t} = -\boldsymbol{\nabla}\cdot\mathbf{J}_i + \Gamma_i \qquad [158]$$

where the Lagrangian derivative is defined with the barycentric velocity,

$$\frac{\mathrm{D}}{\mathrm{D}t} = \frac{\partial}{\partial t} + \mathbf{v}_\mathrm{b}\cdot\boldsymbol{\nabla} \qquad [159]$$

7.02.5.1.2 Momentum and energy in a multicomponent solution

In our multicomponent solution, all constituents are present at each point, and they are all locally submitted to the same pressure and stresses. We assume that the viscous stress is simply related to $\mathbf{v}_\mathrm{b}$ and we neglect inertia as appropriate for the mantle. Newton's second law (here, simply the balance of forces) can be applied to the barycenter and implies

$$\boldsymbol{\nabla}\cdot\underline{\boldsymbol{\tau}} - \boldsymbol{\nabla}P + \overline{\rho}\mathbf{g} = 0 \qquad [160]$$

where the only force is due to the (constant) gravity. The momentum equation thus remains identical to that of a fluid with uniform composition and without inertia (eqn [16]).

Since there is only one momentum equation for i components, the $i-1$ other velocity equations will be found by using the constraints of the laws of thermodynamics and in particular the positivity of the entropy source. To derive the energy conservation, we perform the standard balance to account for all the energy exchanges in a volume Ω and across its surface $\sum$. Instead of the one component equation [30], we have to sum up various contributions and we get

$$\sum_i \frac{\partial(\phi_i\rho_i\mathcal{U}_i)}{\partial t} = -\boldsymbol{\nabla}\cdot\left(\sum_i \phi_i\rho_i\mathcal{U}_i\mathbf{v}_i + P\sum_i \phi_i\mathbf{v}_i + \mathbf{q} - \underline{\boldsymbol{\tau}}\cdot\mathbf{v}_\mathrm{b}\right) + \mathbf{g}\cdot\sum_i \phi_i\rho_i\mathbf{v}_i + \overline{\rho}\overline{H} \qquad [161]$$

In this expression, we recognize the temporal changes in energy ($\mathcal{U}_i$ is the internal energy per unit mass of component i and the kinetic energies are neglected), the bulk energy flux, the pressure work, the thermal diffusion, the viscous stress work, the gravity work, and the radioactivity production $\left(\overline{\rho}\overline{H} = \sum_i \phi_i\rho_i H_i\right)$. The various ϕ_i come from the assumption that each component i is present in proportion ϕ_i in both the volume Ω and its surface $\sum$. We assume that thermal diffusion acts equally for each component and that the surface work of the stress tensor is only related to the barycentric velocity.

Using the definition of the barycentric velocity [152], of the diffusive fluxes [156], and of the momentum conservation [160] and using $\sum_i \phi_i = 1$, the energy expression can be simplified to

$$\sum_i \phi_i\rho_i \frac{\mathrm{D}\mathcal{H}_i}{\mathrm{D}t} = -\boldsymbol{\nabla}\cdot\mathbf{q} - \sum_i \mathbf{J}_i\cdot\boldsymbol{\nabla}\mathcal{H}_i + \frac{\mathrm{D}P}{\mathrm{D}t} - \sum_i \Gamma_i\mathcal{H}_i + \underline{\boldsymbol{\tau}} : \nabla\mathbf{v}_\mathrm{b} + \overline{\rho}\overline{H} \qquad [162]$$

where $\mathcal{H}_i$ are the component enthalpies

$$\mathcal{H}_i = \mathcal{U}_i + \frac{P}{\rho_i} \qquad [163]$$

The enthalpy variation for each component i can be expressed as a function of the state variables P and T. According to eqn [60], we can write for each component

$$\rho_i \frac{\mathrm{D}\mathcal{H}_i}{\mathrm{D}t} = \rho_i C_p^i \frac{\mathrm{D}T}{\mathrm{D}t} + (1 - \alpha_i T)\frac{\mathrm{D}P}{\mathrm{D}t} \qquad [164]$$

Finally, the expression for the temperature evolution is

$$\overline{\rho}\overline{C}_p \frac{\mathrm{D}T}{\mathrm{D}t} = -\boldsymbol{\nabla}\cdot\mathbf{q} - \sum_i \mathbf{J}_i\cdot\boldsymbol{\nabla}\mathcal{H}_i + \overline{\alpha}T\frac{\mathrm{D}P}{\mathrm{D}t} - \sum_i \Gamma_i\mathcal{H}_i + \underline{\boldsymbol{\tau}} : \boldsymbol{\nabla}\mathbf{v}_\mathrm{b} + \overline{\rho}\overline{H} \qquad [165]$$

where the average heat capacity and thermal expansivity are $\overline{C}_p = \sum_i \phi_i\rho_i C_p^i / \overline{\rho}$ and $\overline{\alpha} = \sum_i \phi_i \alpha_i$.

Compared to the homogeneous case [38], two new heat source terms are present, the enthalpy exchange through chemical reactions, $\sum_i \Gamma_i\mathcal{H}_i$, and the enthalpy redistribution by component diffusion, $\sum_i \mathbf{J}_i\cdot\boldsymbol{\nabla}\mathcal{H}_i$.

7.02.5.1.3 Entropy conservation in a multicomponent solution

Entropy conservation is essential for deriving the expressions of the diffusive fluxes. The general expression of entropy conservation [7] is

$$\sum_i \frac{\partial(\phi_i\rho_i\mathcal{S}_i)}{\partial t} = -\boldsymbol{\nabla}\cdot\mathbf{J}_\mathcal{S} + H_\mathcal{S} \qquad [166]$$

where $\mathbf{J}_\mathcal{S}$ and $H_\mathcal{S}$ are the yet unknown entropy flux and source. The entropy of the various components takes into account their specific entropies as well as their configurational entropies or mixing entropies due to the dispersion of the component i in the solution. Introducing the barycentric velocities and the diffusive fluxes, this equation can be recast as

$$\sum_i \phi_i\rho_i \frac{\mathrm{D}\mathcal{S}_i}{\mathrm{D}t} = \boldsymbol{\nabla}\cdot\left(\sum_i \phi_i\rho_i\mathcal{S}_i\mathbf{v}_\mathrm{b} + \sum_i \mathcal{S}_i\mathbf{J}_i - \mathbf{J}_\mathcal{S}\right) + H_\mathcal{S} - \mathcal{S}_i\Gamma_i - \mathbf{J}_i\cdot\boldsymbol{\nabla}\mathcal{S}_i \qquad [167]$$

However, a second expression of the entropy conservation can be obtained from the enthalpy conservation, [162]: using $\mathrm{d}\mathcal{H}_i = T\mathrm{d}\mathcal{S}_i + V_i\mathrm{d}P$, which, in our case, can be expressed as

$$\rho_i \frac{\mathrm{D}\mathcal{H}_i}{\mathrm{D}t} = \rho_i T\frac{\mathrm{D}\mathcal{S}_i}{\mathrm{D}t} + \frac{\mathrm{D}P}{\mathrm{D}t} \qquad [168]$$

we derive

$$\sum_i \phi_i\rho_i T\frac{\mathrm{D}\mathcal{S}_i}{\mathrm{D}t} = -\boldsymbol{\nabla}\cdot\mathbf{q} - \sum_i \mathbf{J}_i\cdot\boldsymbol{\nabla}\mathcal{H}_i - \sum_i \Gamma_i\mathcal{H}_i + \underline{\boldsymbol{\tau}} : \boldsymbol{\nabla}\mathbf{v}_\mathrm{b} + \overline{\rho}\overline{H} \qquad [169]$$

A comparison of the two expressions for the entropy conservation, [167] and [169], allows us to identify the total entropy flux

$$\mathbf{J}_S = \frac{\mathbf{q}}{T} + \mathbf{v}_b \sum_i \phi_i \rho_i \mathcal{S}_i + \sum_i \mathcal{S}_i \mathbf{J}_i \qquad [170]$$

and the entropy sources

$$TH_S = -\left(\frac{\mathbf{q}}{T} + \sum_i \mathcal{S}_i \mathbf{J}_i\right) \cdot \nabla T - \sum_i \mathbf{J}_i \cdot \nabla \mu_i - \sum_i \Gamma_i \mu_i + \underline{\boldsymbol{\tau}} : \nabla \mathbf{v}_b + \overline{\rho} \overline{H} \qquad [171]$$

where we introduced the chemical potentials $\mu_i = \mathcal{H}_i - T\mathcal{S}_i$. The total entropy flux, eqn [170], is related to thermal diffusion and to advection and chemical diffusion of component entropies.

In eqn [171], the gradients of chemical potential and temperature are not independent as the chemical potential gradients implicitly include the temperature gradient, so that alternative expressions can be found. For example, we can introduce the gradient of μ at constant temperature $\nabla_T \mu$ as

$$\nabla_T \mu_i = \nabla \mu_i + \mathcal{S}_i \nabla T \qquad [172]$$

which leads to

$$TH_S = -\frac{1}{T} \mathbf{q} \cdot \nabla \mathrm{T} - \sum_i \mathbf{J}_i \cdot \nabla_T \mu_i - \sum_i \Gamma_i \mu_i + \underline{\boldsymbol{\tau}} : \nabla \mathbf{v}_b + \overline{\rho} \overline{H} \qquad [173]$$

This last equation has the advantage of separating the temperature contribution, ∇T, from the compositional contribution, $\nabla_T \mu_i$ ($\nabla_T \mu_i$ varies mostly with composition as composition can change over very short distances; however, this term is also related to pressure variations) (see de Groot and Mazur, 1984).

7.02.5.1.4 Advection–diffusion equation and reaction rates

Among the entropy sources, only terms involving similar tensorial ranks can be coupled in an isotropic medium, according to the Curie principle. The positivity of the entropy production [173] imposes three conditions, coupling tensors, vectors, and scalars:

$$\underline{\boldsymbol{\tau}} : \nabla \mathbf{v}_b \geq 0, \quad -\mathbf{q} \cdot \frac{\nabla T}{T} - \sum_i \mathbf{J}_i \cdot \nabla_T \mu_i \geq 0 \quad -\sum_i \Gamma_i \mu_i \geq 0 \qquad [174]$$

The first term relates tensors and we have already discussed its implications for the rheology in **Section 7.02.3.2**.

The second term relates vectors and we assume, in agreement with the general principle of nonequilibrium thermodynamics (de Groot and Mazur, 1984), that a phenomenological matrix M relates the thermodynamic fluxes $\mathbf{J} = \mathbf{J}_1 \ldots \mathbf{J}_i \ldots \mathbf{q}$ to the thermodynamic forces $\mathbf{X} = -\nabla_T \mu_1 \ldots -\nabla_T \mu_i \ldots -\nabla T/T$:

$$\begin{pmatrix} \mathbf{J}_1 \\ \mathbf{J}_2 \\ \ldots \\ \mathbf{q} \end{pmatrix} = - \begin{pmatrix} m_{11} & m_{12} & \ldots & m_{1q} \\ m_{21} & m_{22} & \ldots & m_{2q} \\ \ldots & \ldots & \ldots & \ldots \\ m_{q1} & m_{q2} & \ldots & m_{qq} \end{pmatrix} \begin{pmatrix} \nabla_T \mu_1 \\ \nabla_T \mu_2 \\ \ldots \\ \nabla T/T \end{pmatrix} \qquad [175]$$

This linear relationship implies that the term of vectorial rank (with superscript υ), in the entropy source, $TH_S^{(\upsilon)}$, appears as

$$TH_S^{(\upsilon)} = \mathbf{X}^t M \mathbf{X} = \mathbf{X}^t \frac{M + M^t}{2} \mathbf{X} \qquad [176]$$

According to the second law of thermodynamics, the symmetrical part of the matrix M, $(M + M^t)/2$, must be positive definite, that is, the right-hand side of equation [176] must be positive for any vectors $\mathbf{X}$.

At microscopic scale, a process and its reverse occur at the same rate. A consequence, known as the Onsager reciprocal relations, is the existence of symmetry or antisymmetry between m_{ij} and m_{ji} (Onsager, 1903–76). A general discussion can be found in, for example, de Groot and Mazur (1984) or Woods (1975). When the forces are independent of the velocities, the matrix M must be symmetrical, $m_{ij} = m_{ji}$.

In the general case, the transport of heat by concentration gradients (the Dufour effect (Dufour, 1832–92)) or the transport of concentration by temperature gradients (the Soret effect (Soret, 1827–90)) is possible. In many situations, these cross effects are small and we will assume that the matrix M does not couple thermal and compositional effects (the last row and column of M are zero except for $m_{qq}/T = k$, the thermal conductivity).

Even without coupling between thermal and compositional effects, chemical diffusion in a multicomponent system remains difficult to discuss in the most general case (the positive definiteness of a symmetrical matrix is not a very strong constraint). We therefore restrict our study to a simple two-component system where

$$\begin{pmatrix} \mathbf{J}_1 \\ \mathbf{J}_2 \end{pmatrix} = - \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} \nabla_T \mu_1 \\ \nabla_T \mu_2 \end{pmatrix} \qquad [177]$$

For such a simple case, the sum of the fluxes must cancel (see eqn [157]), and since the Onsager relations impose the symmetry of the matrix, the coefficients m_{ij} must verify

$$m_{11} + m_{21} = m_{12} + m_{22} = m_{12} - m_{21} = 0 \qquad [178]$$

Only one coefficient, for example, $m = m_{11}$, can be freely chosen, and the fluxes can be written as

$$\begin{aligned} \mathbf{J}_1 &= m \nabla_T (\mu_2 - \mu_1) \\ \mathbf{J}_2 &= m \nabla_T (\mu_1 - \mu_2) \end{aligned} \qquad [179]$$

and the second law requires $m > 0$. If the component 1 is in small quantity (the solute) and the component 2 is in large quantity (the solvent, with $\nabla_T(\mu_2) = 0$), we can easily track the evolution of solute concentration C_1. Its chemical diffusion flux is $\mathbf{J}_1 = -m \nabla_T \mu_1$, and according to eqn [158], its concentration

$$\overline{\rho} \left(\frac{\partial C}{\partial t} + \mathbf{v}_b \cdot \nabla C\right) = \nabla \cdot (m \nabla_T \mu) \qquad [180]$$

where the subscript 1 has been omitted and the source term Γ_1 neglected.

For a solute, the chemical potential is a standard chemical potential μ_0 plus a mixing term expressing the entropy gain (configurational entropy associated with the increased disorder) made by dispersing the solute into the solvent, of the form $RT \log a(C)$ (for crystalline solids, the activity $a(C)$ of the mixing term can be complex since it depends on the number and multiplicity of crystallographic sites (Spear, 1993), but we

just need to know that it is related to C). In a domain where the average density and the standard potential remain uniform, the advection–diffusion equation is obtained:

$$\frac{\partial C}{\partial t} + \mathbf{v}_b \cdot \nabla C = \nabla \cdot (D \nabla C) \qquad [181]$$

with a diffusion coefficient $D = m/\overline{\rho}(\partial\mu/\partial C)$ most likely T-dependent. The negative linear relationship between chemical diffusion flux $\mathbf{J}$ (now in mol m^{-2} s^{-1}) and concentration gradient $\mathbf{J} = -D\nabla C$ (with D in m^2 s^{-1}) is called the first Fick's law (Fick, 1829–1901).

When a component is present in two domains separated by a compositional interface, its standard chemical potential μ_0 is generally discontinuous. In this case, the gradient of the chemical potential at constant T, $\nabla_T \mu$ is a mathematical distribution that contains a term $\nabla_T \mu_0$, infinite on the compositional interface. This discontinuity drives an infinitely fast diffusion of the solute component across the interface until the equilibrium and the continuity of the total chemical potential $\mu_0 + RT \log a(C)$ are restored. The concentration ratio of C (or partition coefficient of C) must therefore verify

$$\frac{a(C)^+}{a(C)^-} = \exp\left(-\frac{\mu_0^+ - \mu_0^-}{RT}\right) \qquad [182]$$

where $\mu_0^+ - \mu_0^- = [\mu_0]$ is the jump of standard potential across the interface. This equation corresponds to the general rule of chemical equilibrium.

The last entropy source in eqn [174], $\sum_i \Gamma_i \mu_i$, relates two scalars (production rates and chemical potentials). In a mixture of i components involving k stable atomic species, the conservation of these atomic species implies that only $r = i - k$ linearly independent reactions exist. Let n_i^j be the stoichiometric coefficient of the component i in the jth $= (1 \ldots r)$ chemical reaction with reaction rate, Γ_j. We can express Γ_i as

$$\Gamma_i = \sum_{j=1\ldots r} n_i^j \Gamma_j \qquad [183]$$

and the second law imposes

$$-\sum_{j=1\ldots r} \Gamma_j \sum_i n_i^j \mu_i \geq 0 \qquad [184]$$

The positivity of the entropy source is satisfied if the kinetic rates of the jth $= 1 \ldots r$ chemical reaction are proportional to their chemical affinities, $-\Delta G_j = -\sum_i n_i^j \mu_i$, with positive reaction rate factors $\mathcal{R}_j$:

$$\Gamma_j = -\mathcal{R}_j \Delta G_j \qquad [185]$$

In the case of exact thermodynamic equilibrium, $\Delta G_i = 0$, the reaction rate vanishes and the second law is of course satisfied.

Chemical reaction rates are however very rarely simply proportional to the affinities and the $\mathcal{R}_j$ are likely some complex, but positive, functions of P, T and concentrations C_i. Indeed, in the context of chemical reactions, the kinetics of reactions are often approximated by phenomenological laws based on the law of mass action rather than on affinities (e.g., the reaction rates are rather related to $\exp(-\Delta G_j/RT) - 1$ than to $-\Delta G_j$ (see Lasaga, 1998; Rudge et al., 2011)).

In the same way as we defined the affinity $-\Delta G_j$ of the reaction j, we can define its enthalpy $\Delta\mathcal{H}_j = \sum_i n_i^j \mathcal{H}_i$ (see de Groot and Mazur, 1984). The enthalpy exchange term of the energy equation [165] $\sum_i \Gamma_i \mathcal{H}_i$ can also be written as $\sum_j \Gamma_j \Delta\mathcal{H}_j$, which represents the products of the reaction rates and the reaction enthalpies. Various phase changes take place in the mantle, most notably at 410 and 660 km depth. Their effects on mantle convection have been studied by various authors and will be discussed in Section 7.02.6.6.

7.02.5.1.5 Conservation properties of the advection–diffusion equation

We now make the hypothesis that the evolution of concentration of a solute in the convective fluid is controlled by the advection–diffusion equation [181] and that this solute is not involved in any chemical reaction, $\Gamma = 0$. For simplicity, we assume that the barycentric flow is incompressible ($\nabla \cdot \mathbf{v}_b = 0$, C can therefore be a concentration per unit volume or per unit mass) and the diffusion coefficient D is a constant. The fluid and the solute cannot escape the domain Ω; the normal velocity and normal diffusive flux are thus zero on the boundaries of the domain, that is, $\mathbf{v}_b \cdot \mathbf{n} = 0$ and $D\nabla C \cdot \mathrm{n}_b = 0$ on the surface $\sum$ with normal vector $\mathbf{n}$.

First, it is obvious that when integrated over the total domain Ω and with the divergence theorem, the advection–diffusion equation [181] implies

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega C \mathrm{d}V = -\int_\Sigma (C\mathbf{v}_b - D\nabla C) \cdot \mathrm{d}\mathbf{S} = 0 \qquad [186]$$

The initial heterogeneity does not disappear, it is just redistributed through time.

To understand how the heterogeneity is redistributed, we can express the evolution of the concentration variance. The variance is related to $C^2 - \overline{C}^2$ where $\overline{C}$ is the average concentration. We can get an evolution equation for C^2 by multiplying eqn [181] by $2C$ and we get after some algebra

$$\frac{\mathrm{D}C^2}{\mathrm{D}t} = D\nabla^2 C^2 - 2D|\nabla C|^2 \qquad [187]$$

This expression when integrated over the closed volume Ω implies that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega C^2 \mathrm{d}V = -2D\int_\Omega |\nabla C|^2 \mathrm{d}V \qquad [188]$$

Since the right side is always negative, the variance must continuously decrease until $|\nabla \mathrm{C}| = 0$, which corresponds to a state of complete homogenization.

If the initial concentration is $C = 1$ in some domain and $C = 0$ elsewhere, a perfect homogenization is achieved after a time t if the concentration is everywhere $C = \overline{C}$, the average concentration. Since the average mixing rate is proportional to the diffusion D, eqn [188], we note, however, that a nondiffusive flow does not homogenize at all. A nondiffusive flow just stirs the heterogeneities. When there is no diffusion, the initial heterogeneity is stirred and stretched, but the local concentrations remain, for all time, either $C = 1$ or $C = 0$, but never an intermediate value (see Figure 5)

In the case of the Earth's mantle, the solid-state diffusion coefficients are all very low (see Table 2), and many studies have totally neglected chemical diffusion. We see, at face value, that these models are not really homogenizing, only stirring the heterogeneities. Without diffusion, a chemical

heterogeneity (e.g., a piece of subducted oceanic crust) will forever remain the same petrologic heterogeneity, and only its shape will change.

Since the mixing rate is related to the compositional gradient [188], we should discuss the evolution of this gradient. We multiply eqn [181] by the operator $2\nabla C\cdot\nabla$ to obtain

$$\frac{\mathrm{D}|\nabla C|^2}{\mathrm{D}t} = -2\nabla C\cdot\underline{\dot{\boldsymbol{\varepsilon}}}\cdot\nabla C + 2D\left[\nabla\cdot\left(\nabla^2 C\nabla C\right) - \left(\nabla^2 C\right)^2\right] \quad [189]$$

which can be integrated as

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega |\nabla C|^2\,\mathrm{d}V = -2\int_\Omega \nabla C\cdot\underline{\dot{\boldsymbol{\varepsilon}}}\cdot\nabla C\,\mathrm{d}V - 2D\int_\Omega \left(\nabla^2 C\right)^2\mathrm{d}V \quad [190]$$

The rate of gradient production is related to the flow properties through the strain rate tensor $\underline{\dot{\boldsymbol{\varepsilon}}}$ and to the diffusion. The diffusion term is negative and decreases the sharpness of compositional gradients.

The term related to the flow properties through the strain tensor (first term of the right side of eqn [190]) could in principle be either positive or negative. However, as time evolves, this term must become positive. The strain rate tensor has locally three principal axes and three principal strain rates, the sum of them being zero since the flow is incompressible. The stretched heterogeneities become elongated along the direction of the maximum principal strain rate and the concentration gradients reorient themselves along the minimum, and negative, principal strain rate. The term under the first integral on the right side of eqn [190] is thus of the order of $+|\dot{\varepsilon}_{\min}||\nabla C|^2$ ($\dot{\varepsilon}_{\min}$ is the local, negative eigenvalue of the strain rate tensor). Stirring is thus the source of production of concentration gradient (see **Figure 5**).

We can now understand the interplay between advection and diffusion. Even when the diffusion coefficient D is vanishingly small in eqn [190], the stirring of the flow by convection will enhance the concentration gradients until the average diffusion term, proportional to the square of concentration gradients, will become large enough (see eqn [188]), for a rapid decrease of the concentration variance. We illustrate this behavior in the next two subsections by choosing a simple expression for the strain rate and computing the evolution of concentration through time.

7.02.5.1.6 Laminar and turbulent stirring

The efficiency of mixing, mostly controlled by stirring, is therefore related to the ability of the flow to rapidly reduce the thickness of heterogeneities (Olson et al., 1984). In this section, we set aside diffusion and discuss the stirring properties of a flow (see also Tackley, this volume). Let us consider a vertical piece of heterogeneity of width $2d_0$ and height $2l_0$ ($l_0 \gg d_0$) in a simple shear flow $v_x = \dot{\boldsymbol{\varepsilon}}z$, with constant $\dot{\boldsymbol{\varepsilon}}$. Its top and bottom ends are at $(0, l_0)$ and $(0, -l_0)$ and they will be advected to $(\dot{\boldsymbol{\varepsilon}}tl_0, l_0)$, $(-\dot{\boldsymbol{\varepsilon}}tl_0, -l_0)$ after a time t. As the heterogeneity length increases as $2l_0\sqrt{1+\dot{\boldsymbol{\varepsilon}}^2t^2}$, mass conservation implies that its half width $d(t)$ decreases as

$$d(t) = \frac{d_0}{\sqrt{1+\dot{\boldsymbol{\varepsilon}}^2t^2}} \quad [191]$$

On the contrary, in a pure shear flow, $v_z = \mathrm{d}z/\mathrm{d}t = z\dot{\boldsymbol{\varepsilon}}$ with constant $\dot{\boldsymbol{\varepsilon}}$, the length of the heterogeneity would increase as $l = l_0\exp(\dot{\boldsymbol{\varepsilon}}t)$ and its width would shrink as

$$d(t) = d_0\exp(-\dot{\boldsymbol{\varepsilon}}t) \quad [192]$$

Table 2 Homogenization times for helium and uranium assuming a heterogeneity of initial thickness $2d_0 = 7$ km and a strain rate of 5×10^{-16} s^{-1}

	Uranium	*Helium*
D (m^2 s^{-1})	10^{-19}	10^{-13}
t_0 (Ma)	3.88×10^{13}	3.88×10^{7}
t_L (Ma)	3.60×10^{5}	3.60×10^{3}
t_T (Ma)	1575	699

The diffusion coefficients for U and He are from Hofmann and Hart (1978) and Trull and Kurtz (1993).

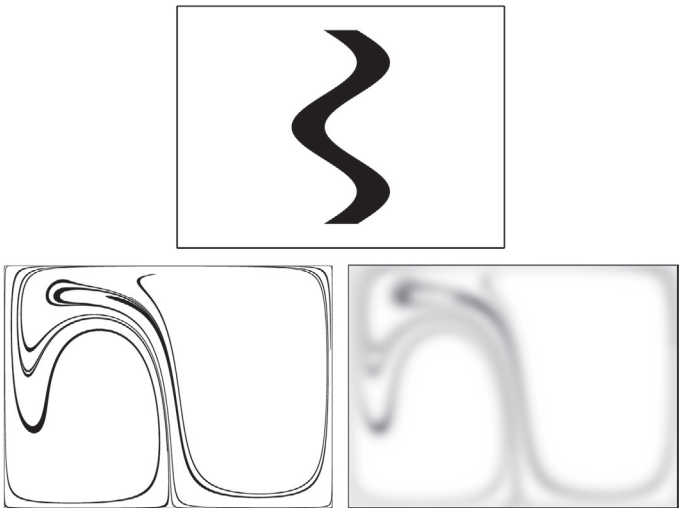

Figure 5 An initial heterogeneity (top) is introduced at $t = 0$ into a time-dependent convection cell. Without diffusion, $D = 0$ (bottom left), the heterogeneity is stirred by convection and then stretched on the form of thin ribbons. However, the variance of the heterogeneity concentration remains constant. It is only with diffusion, $D \neq 0$ (bottom right), that a real homogenization occurs with a decrease of the heterogeneity variance.

Such a flow is said to induce turbulent stirring. This is unfortunate terminology because turbulent stirring can occur in a creeping flow with $Re=0$. Mantle convection is not turbulent but it generates turbulent stirring.

Chaotic mixing flows have globally turbulent stirring properties and the qualitative idea that highly time-dependent convection with high Rayleigh number mixes more efficiently than low Rayleigh number convection is often true (Coltice and Schmalzl, 2006; Schmalzl et al., 1996). However, steady 3-D flows can also induce turbulent mixing. This surprising phenomenon called Lagrangian chaos is well illustrated for some theoretical flows (Dombre et al., 1986) and for various simple flows (Ottino, 1989; Toussaint et al., 2000). For example, in a steady flow under an oceanic ridge offset by a transform fault, the mixing is turbulent (Ferrachat and Ricard, 1998).

7.02.5.1.7 Diffusion in Lagrangian coordinates

In Section 7.02.5.1.5, we discussed the mixing properties from a Eulerian viewpoint. We can also understand the interplay between diffusion and stretching (stirring) by adopting a Lagrangian viewpoint (Kellogg and Turcotte, 1987; Ricard and Coltice, 2004), that is, by solving the advection–diffusion equations in a coordinate frame that follows the deformation.

Let us consider a strip of thickness $2d_0$ with an initial concentration C_0 embedded in an infinite matrix of concentration C_∞. In the absence of motion, the solution of the advection–diffusion equation [181] can be expressed using the error function, and the time-dependent concentration $C(x, t)$ is given by

$$\frac{C(x,t)-C_\infty}{C_0-C_\infty}=\frac{1}{2}\left[\mathrm{erf}\left(\frac{d_0-x}{2\sqrt{Dt}}\right)+\mathrm{erf}\left(\frac{d_0+x}{2\sqrt{Dt}}\right)\right] \quad [193]$$

where x is a coordinate perpendicular to the strip and is zero at its center.

The concentration at the center of the strip ($x=0$) is

$$\frac{C(0,t)-C_\infty}{C_0-C_\infty}=\mathrm{erf}\left(\frac{d_0}{2\sqrt{Dt}}\right) \quad [194]$$

and the concentration decreases by a factor of about 2 in the diffusive time

$$t_0\approx\frac{d_0^2}{D} \quad [195]$$

($\mathrm{erf}(1/2)$ is not far from 1/2). The time needed to homogenize a 7 km thick piece of oceanic crust introduced into a motionless mantle is extremely long (see Table 2); even the relatively mobile helium would be frozen in place since the Earth formed as it would only have migrated around 50 cm.

However, this idea of a $\sqrt{t}$ diffusion is faulty since the flow stirs the heterogeneity and increases compositional gradients (see eqn [190]), which in turn accelerates the mixing process (e.g., eqn [188]). Assuming that the problem remains two-dimensional enough so that diffusion only occurs perpendicular to the deforming heterogeneity, let $d(t)$ with time derivative $\dot{d}(t)$ be half the thickness of the strip containing a chemical heterogeneity. The velocity perpendicular to the strip would locally be at first order

$$v_x=x\frac{\dot{d}(t)}{d(t)} \quad [196]$$

(each side of the stripe, at $x=\pm d(t)$ moves at $\pm\,\dot{d}(t)$).

We can choose as a new space variable $\tilde{x}=xd_0/d(t)$, in such a way that the Lagrangian coordinate $\tilde{x}$ will vary between the fixed values $-d_0$ and d_0. The diffusion equation becomes

$$\left(\frac{\partial C}{\partial t}\right)_{\tilde{x}}=D\left(\frac{d_0}{d(t)}\right)^2\frac{\partial^2 C}{\partial\tilde{x}^2} \quad [197]$$

where the partial time derivative is now computed at constant $\tilde{x}$. We see that the advection–diffusion equation has been turned into a pure diffusive equation where the diffusivity D has been replaced by $D(d_0/d(t))^2$. This equivalent diffusivity is larger than D and increases with time as $d(t)$ decreases.

To solve analytically equation [197], it is appropriate to rescale the time variable by defining $\tilde{t}=F(t)$ with

$$F(t)=\int_0^t\left(\frac{d_0}{d(u)}\right)^2\mathrm{d}u \quad [198]$$

and the resulting advection–diffusion equation in Lagrangian coordinates becomes the simple diffusion equation with constant diffusivity where t and x are replaced by $\tilde{t}$ and $\tilde{x}$. Its solution can be deduced from eqn [193]. For example, the concentration at the center of the deformable strip varies like

$$\frac{C(0,t)-C_\infty}{C_0-C_\infty}=\mathrm{erf}\left(\frac{d_0}{2\sqrt{DF(t)}}\right) \quad [199]$$

and the concentration diminishes in amplitude by a factor of 2 after a time t that satisfies

$$F(t)\approx\frac{d_0^2}{D} \quad [200]$$

To perform a numerical application, let us consider that the flow is either a simple shear, eqn [191], or a pure shear deformation, eqn [192]. Computing $F(t)$ from equation [198] is straightforward and, assuming $\dot{\boldsymbol{\varepsilon}}t\gg 1$, we get from eqn [200] the homogenization times

$$t_\mathrm{L}\approx\frac{3^{1/3}d_0^{2/3}}{\dot{\boldsymbol{\varepsilon}}^{2/3}D^{1/3}} \quad [201]$$

and

$$t_\mathrm{T}\approx\frac{1}{\dot{\boldsymbol{\varepsilon}}}\log\frac{d_0^2\dot{\boldsymbol{\varepsilon}}}{D} \quad [202]$$

for the laminar and turbulent case, respectively. For the same oceanic crust of initial thickness 7 km, we get homogenization times of about 0.70 byr for He and 1.57 byr for U if we use the pure shear mechanism and assume rather arbitrarily that $\dot{\boldsymbol{\varepsilon}}=5\times10^{-16}\ \mathrm{s}^{-1}$ (this corresponds to a typical plate velocity of 7 cm year^{-1} over a plate length of 5000 km). Although He and U have diffusion coefficients 6 orders of magnitude apart, their residence times in a piece of subducted oceanic crust only differ by a factor ≈ 2 (see Table 2).

The use of tracers to simulate the evolution of chemical properties in the mantle is our best method since solid-state diffusion is too slow to be efficiently accounted for in a numerical simulations (e.g., Tackley and Xie, 2002; van Keken et al., 2002). However, by using tracers, we do not necessarily take

into account that some of them may represent points that have been stretched so much that their initial concentration anomalies have completely diffused into the background. In other words, even if diffusion seems negligible, diffusion will erase all heterogeneities after a finite time that is mostly controlled by the stirring properties of the flow.

7.02.5.2 Fluid Dynamics of Two-Phase Flows

Up to now, in all of Section 7.02.5.1, all components were mixed in a single phase. However, another important geophysical application occurs when the multicomponents belong to different phases. This case can be illustrated with the dynamics of partial melt in a deformable compacting matrix. Partial melts are obviously present under ridges and hot spots, but they may also be present in the middle and deep mantle (Bercovici and Karato, 2003; Williams and Garnero, 1996) and they were certainly more frequent when the Earth was younger (Labrosse et al., 2007). We discuss the situation where two phases, fluid and matrix, can interact. In contrast to Section 7.02.5.1, where the proportion and velocity of each component in solution were defined everywhere at a microscopic level, in a partial melt aggregate, the local velocity at a microscopic level is either the velocity of a matrix grain, $\tilde{\mathbf{v}}_\mathrm{m}$, or the interstitial velocity of the melt, $\tilde{\mathbf{v}}_\mathrm{f}$.

We assume that the two phases are individually homogeneous and incompressible (various extensions of this theory are possible, for compressible phases (e.g., Bercovici and Michaut, 2010; Michaut et al., 2013)), when one phase is void (e.g., Bercovici and Ricard, 2003; Ricard and Bercovici, 2003) and when one phase is only manifested by its surface energy (e.g., Hier-Majumder et al., 2006). They have densities ρ_f and ρ_m and Newtonian rheologies with viscosities η_f and η_m. They are isotropically mixed and connected. Their volume fractions are ϕ (the porosity) and $1-\phi$. The rate of magma melting or freezing is $\Delta\Gamma$ (in kg m^{-3} s^{-1}). Although the two phases have very different physical properties, we will require the equations to be material-invariant until we need to use numerical values. This means that swapping f and m, ϕ and $1-\phi$, and $\Delta\Gamma$ and $-\Delta\Gamma$ must leave the equations unchanged. This rule is both a physical requirement and a strong guidance in establishing the general equations (Bercovici et al., 2001a).

We make the hypothesis that there is a mesoscopic size of volume δV, which includes enough grains and interstitial fluid that averaged and continuous quantities can be defined. Classical fluid dynamics also has its implicit averaging volume δV that must contain enough atoms that quantum effects are negligible, but what is needed here is a much larger volume. This averaging approach remains meaningful because the geophysical macroscopic phenomenon that we want to understand (say, melting under ridges) has characteristic sizes large compared to those of the averaging volume (say, a few cm^3) (Bear, 1988).

To do the averaging, we define at microscopic level a function θ that takes the value 1 in the interstitial fluid and the value 0 in the matrix grain. Mathematically, this function is rather a distribution and it has a very convolved topology. From it, we can define first the porosity (volume fraction of fluid) ϕ and then the fluid and matrix averaged velocities, $\mathbf{v}_\mathrm{f}$ and $\mathbf{v}_\mathrm{m}$ (Bercovici et al., 2001a), by

$$\phi = \frac{1}{\delta V}\int_{\delta V} \theta \,\mathrm{d}V \qquad [203]$$

$$\phi\mathbf{v}_\mathrm{f} = \frac{1}{\delta V}\int_{\delta V} \theta\tilde{\mathbf{v}}_\mathrm{f}\,\mathrm{d}V, \quad (1-\phi)\mathbf{v}_\mathrm{m} = \frac{1}{\delta V}\int_{\delta V} (1-\theta)\tilde{\mathbf{v}}_\mathrm{m}\,\mathrm{d}V \qquad [204]$$

7.02.5.2.1 Mass conservation for matrix and fluid

Having defined the average quantities, the derivation of the two mass conservation equations is fairly standard (Bercovici et al., 2001a; McKenzie, 1984). They are

$$\frac{\partial\phi}{\partial t} + \boldsymbol{\nabla}\cdot[\phi\mathbf{v}_\mathrm{f}] = \frac{\Delta\Gamma}{\rho_\mathrm{f}} \qquad [205]$$

$$-\frac{\partial\phi}{\partial t} + \boldsymbol{\nabla}\cdot[(1-\phi)\mathbf{v}_\mathrm{m}] = -\frac{\Delta\Gamma}{\rho_\mathrm{m}} \qquad [206]$$

We get the same equations as in eqn [150] except that we refer to ϕ, $1-\phi$, and $\Delta\Gamma$ instead of ϕ_1, ϕ_2, and Γ_1. When averaged, the mass conservation equations of two separated phase take the same form as the mass conservation equations of the two components in a solution.

We define an average and a difference quantity for any general variable, q, by

$$\overline{q} = \phi q_\mathrm{f} + (1-\phi)q_\mathrm{m}, \quad \Delta q = q_\mathrm{m} - q_\mathrm{f} \qquad [207]$$

The velocity $\overline{\mathbf{v}}$ is volume-averaged and is different from the barycentric velocity (eqn [152]),

$$\mathbf{v}_\mathrm{b} = (\phi\rho_\mathrm{f}\mathbf{v}_\mathrm{f} + (1-\phi)\rho_\mathrm{m}\mathbf{v}_\mathrm{m})/\overline{\rho} = \overline{v} + \phi(1-\phi)\Delta v\Delta\rho/\overline{\rho}$$

By combining the fluid and matrix mass conservation equations, we get the total mass conservation equation

$$\frac{\partial\overline{\rho}}{\partial t} + \boldsymbol{\nabla}\cdot(\overline{\rho}\mathrm{v}_\mathrm{b}) = 0 \qquad [208]$$

(as before, eqn [153]) and the time rate of change in volume during melting

$$\boldsymbol{\nabla}\cdot\overline{\mathbf{v}} = \Delta\Gamma\frac{\Delta\rho}{\rho_\mathrm{f}\rho_\mathrm{m}} \qquad [209]$$

7.02.5.2.2 Momentum conservation of matrix and fluid

Total momentum conservation, that is, the balance of the forces applied to the mixture, is

$$\boldsymbol{\nabla}\cdot\underline{\overline{\boldsymbol{\tau}}} - \boldsymbol{\nabla}\overline{P} + \overline{\rho}\mathbf{g} = 0 \qquad [210]$$

We have considered that the only force is due to gravity, although surface tension between the two phases could also be introduced (Bercovici and Ricard, 2003; Bercovici et al., 2001a). In this equation, $\overline{P}$, $\underline{\overline{\boldsymbol{\tau}}}$, and $\overline{\rho}$ are the average pressure, stress, and density, respectively. The equation is not surprising and looks identical to its counterpart for a multicomponent solution [160]. However, the average pressure and stresses $\overline{P} = \phi P_\mathrm{f} + (1-\phi)P_\mathrm{m}$ and $\underline{\overline{\boldsymbol{\tau}}} = \phi\underline{\boldsymbol{\tau}}_\mathrm{f} + (1-\phi)\underline{\boldsymbol{\tau}}_\mathrm{m}$ are now the sum of two separate contributions, from two separate phases having most likely very different rheologies and different pressures. Hypothesizing that the two phases are subject to the same pressure does not rest on any physical justification. It certainly cannot hold if surface tension is present as the Young–Laplace condition implies in the static case, a pressure difference between phases related to the curvature of their interface. We will show later that the pressure difference between phases ΔP controls the rate of porosity change.

We split the total momentum equation into two equations one for the fluid and one for the matrix:

$$\nabla[\phi P_f] + \phi\rho_f \mathbf{g} + \nabla.\left[\phi\underline{\overline{\tau}}_f\right] + \mathbf{h}_f = 0 \quad [211]$$

$$-\nabla[(1-\phi)P_m] + (1-\phi)\rho_m \mathbf{g} + \nabla.\left[(1-\phi)\underline{\overline{\tau}}_m\right] + \mathbf{h}_m = 0 \quad [212]$$

where $\mathbf{h}_f$ and $\mathbf{h}_m$ satisfy $\mathbf{h}_f + \mathbf{h}_m = 0$ and represent the interaction forces acting on the fluid and on the matrix, across the interfaces separating the two phases. Because of the complexities of the interfaces, these two interaction forces must be parameterized in some way.

The simplest contribution to the interfacial forces that preserves Galilean invariance is a Darcy-like term $c\Delta\mathbf{v} = c(\mathbf{v}_m - \mathbf{v}_f)$ (Drew and Segel, 1971; McKenzie, 1984) (Darcy, 1803–58). The interaction coefficient c is related to permeability that is itself a function of porosity (Bear, 1988). A symmetrical form compatible with the usual Darcy term is (see Bercovici et al., 2001a)

$$c = \frac{\eta_f \eta_m}{k_0\left[\eta_f(1-\phi)^{n-2} + \eta_m \phi^{n-2}\right]} \quad [213]$$

where the permeability of the form $k_0\phi^n$ was used (this is often called the Kozeny–Carman law with $n \approx 2$–3 (Carman, 1939)). Assuming $n = 2$ and $\eta_m \gg \eta_f$, the interaction coefficient becomes a constant, $c = \eta_f/k_0$.

In the absence of gravity and when the pressures are uniform and equal, no motion should occur even in the presence of nonuniform porosity. In this situation where $\Delta\mathbf{v} = \underline{\overline{\tau}}_f = \underline{\overline{\tau}}_m = 0$ and where P is uniform, $P = P_f = P_m$, the force balances are $-P\nabla\phi + \mathbf{h}_f = -P\nabla(1-\phi) + \mathbf{h}_m = 0$. Therefore, the interface forces $\mathbf{h}_f$ and $\mathbf{h}_m$ must also include $P\nabla\phi$ and $P\nabla(1-\phi)$ when the two pressures are equal. This led Bercovici and Ricard (2003) to write the interaction terms as

$$\begin{aligned} \mathbf{h}_f &= c\Delta\mathbf{v} + P_f\nabla\phi + \omega\Delta P\nabla\phi \\ \mathbf{h}_m &= -c\Delta\mathbf{v} + P_m\nabla(1-\phi) + (1-\omega)\Delta P\nabla\phi \end{aligned} \quad [214]$$

with $0 \le \omega \le 1$. These expressions verify $\mathbf{h}_f + \mathbf{h}_m = 0$, are Galilean- and material-invariant, and allow equilibrium of a mixture with nonuniform porosity but uniform and equal pressures (see also McKenzie, 1984).

At microscopic level, the matrix–melt interfaces are not sharp discontinuities but correspond to layers (called 'selvage' layers) of disorganized atom distributions. The coefficient $0 < \omega < 1$ controls the partitioning of the pressure jump (and potentially of the surface tension) between the two phases (Bercovici and Ricard, 2003) and represents the fraction of the volume-averaged surface force exerted on the fluid phase. The exact value of ω is related to the microscopic behavior of the two phases (molecular bond strengths and thickness of the interfacial selvage layers) and measures the extent to which the microscopic interface layer is embedded in one phase more than the other. The only general physical constraints that we have are that ω must be zero when the fluid phase disappears (when $\phi = 0$) and when the fluid phase becomes unable to sustain stresses (when $\eta_f = 0$). A symmetrical form like

$$\omega = \frac{\phi\eta_f}{\phi\eta_f + (1-\phi)\eta_m} \quad [215]$$

satisfies these conditions. For mantle conditions, $\eta_f \ll \eta_m$ and hence, $\omega \approx 0$.

To summarize, general expressions for the equations of fluid and matrix momentum conservation are (Bercovici and Ricard, 2003)

$$-\phi[\nabla P_f - \rho_f \mathbf{g}] + \nabla.\left[\phi\underline{\boldsymbol{\tau}}_f\right] + c\Delta\mathbf{v} + \omega\Delta P\nabla\phi = 0 \quad [216]$$

$$-(1-\phi)[\nabla P_m - \rho_m \mathbf{g}] + \nabla.\left[(1-\phi)\underline{\boldsymbol{\tau}}_m\right] - c\Delta\mathbf{v} + (1-\omega)\Delta P\nabla\phi = 0 \quad [217]$$

The relationship between stress and velocities does not include an explicit bulk viscosity term (Bercovici et al., 2001a), and for each phase j, the deviatoric stress is simply

$$\underline{\boldsymbol{\tau}}_j = \eta_j\left(\nabla\mathbf{v}_j + \left[\nabla\mathbf{v}_j\right]^t - \frac{2}{3}\nabla\cdot\mathbf{v}_j\underline{\mathbf{I}}\right) \quad [218]$$

where j stands for f or m. There is no difference in constitutive relations for the isolated component and the component in the mixture.

7.02.5.2.3 Energy conservation for two-phase flows

In the case where surface energy and entropy exist on interfaces, the conservation of energy deserves more care (Bercovici and Ricard, 2003; Bercovici et al., 2001a; Hier-Majumder et al., 2006; Sramek et al., 2007). Otherwise, the global conservation is straightforward and can be expressed by the following equation where the left side represents the temporal change of internal energy content in a fixed control volume and the right side represents the different contributions to this change, namely, internal heat sources, loss of energy due to diffusion, advection of energy, and rate of work of both surface and body forces,

$$\begin{aligned} \frac{\partial}{\partial t}[\phi\rho_f\mathcal{U}_f + (1-\phi)\rho_m\mathcal{U}_m] &= \overline{\rho}\overline{H} - \nabla\cdot\mathbf{q} - \nabla\cdot[\phi\rho_f\mathcal{U}_f\mathbf{v}_f \\ &+ (1-\phi)\rho_m\mathcal{U}_m\mathbf{v}_m] + \nabla\cdot\left[-\phi P_f\mathbf{v}_f - (1-\phi)P_m\mathbf{v}_m + \phi\mathbf{v}_f\cdot\underline{\boldsymbol{\tau}}_f\right. \\ &+ \left.(1-\phi)\mathbf{v}_m\cdot\underline{\boldsymbol{\tau}}_m\right] + \phi\mathbf{v}_f\cdot\rho_f\mathbf{g} + (1-\phi)\mathbf{v}_m\cdot\rho_m\mathbf{g} \end{aligned} \quad [219]$$

The last equation is manipulated in the standard way using the mass and momentum equations. Because the two phases are incompressible, their internal energies are simply $d\mathcal{U}_f = C_f dT$ and $d\mathcal{U}_m = C_m dT$. After some algebra, we get

$$\begin{aligned} \phi\rho_f C_f\frac{D_f T}{Dt} + (1-\phi)\rho_m C_m\frac{D_m T}{Dt} &= -\nabla\cdot\mathbf{q} - \Delta P\frac{D_\omega\phi}{Dt} \\ &+ \Delta\mathcal{H}\Delta\Gamma + \Psi + \overline{\rho}\overline{H} \end{aligned} \quad [220]$$

where Ψ is the rate of deformational work

$$\Psi = \phi\nabla\mathbf{v}_f : \underline{\boldsymbol{\tau}}_f + (1-\phi)\nabla\mathbf{v}_m : \boldsymbol{\tau}_m + c(\Delta v)^2 \quad [221]$$

It contains the dissipation terms of each phase plus a term related to the friction between the two phases. The fundamental derivatives are defined by

$$\frac{D_j}{Dt} = \frac{\partial}{\partial t} + \mathbf{v}_j\cdot\nabla \quad [222]$$

where $\mathbf{v}_j$ is to be substituted with the appropriate velocity $\mathbf{v}_f$, $\mathbf{v}_m$, or $\mathbf{v}_\omega$ with

$$\mathbf{v}_\omega = \omega\mathbf{v}_f + (1-\omega)\mathbf{v}_m \quad [223]$$

In contrast to Section 7.02.5.1, it would not make much sense to try to keep the equations in terms of an average velocity like $\mathbf{v}_b$ plus some minor diffusion terms. Here, the two components may have very different velocities (e.g., a weak downward compaction of the matrix with a fast upward migration of magma), and we have to define various material derivatives.

Since ω represents a partitioning of pressure jump, it is not surprising to find the velocity $\mathbf{v}_\omega$ (included in D_ω/Dt) in the work term related to this pressure jump. Associated with this partitioning factor, we can also introduce interface values, q^ω, that we will use later. Any quantity $\Delta q = q_m - q_f$ can also be written as $(q_m - q^\omega) - (q_f - q^\omega)$. When the property jump is embedded entirely in the matrix ($\omega = 0$), there should be no jump within the fluid and we must have $q^\omega = q_f$. Reciprocally, when $\omega = 1$, we should have $q^\omega = q_m$. This prompts us to define interface values by

$$q^\omega = (1-\omega)q_f + \omega q_m \qquad [224]$$

Notice in the expressions of the interface velocity $\mathbf{v}_\omega$, eqn [223], and interface value q^ω, eqn [224], that ω and $1-\omega$ are interchanged.

The right side of eqn [220] contains two new expressions in addition to the usual terms (heat production, diffusion, and deformational work). The first term includes the changes in porosity $D_\omega\phi/Dt$ times the difference in pressures between phases, ΔP. The other term contains the difference in the specific enthalpies $\Delta\mathcal{H} = \mathcal{H}_m - \mathcal{H}_f$ where the enthalpy of phase j is defined by $\mathcal{H}_j = \mathcal{U}_j + P_j/\rho_j$. A similar term was found for components reacting in a solution [162].

7.02.5.2.4 Entropy production and phenomenological laws

Entropy conservation is needed to constrain the pressure jump between phases and the melting rate. Starting from entropy conservation and following Sramek et al. (2007)

$$\frac{\partial}{\partial t}[\phi\rho_f\mathcal{S}_f + (1-\phi)\rho_m\mathcal{S}_m] = -\nabla\cdot\mathbf{J}_\mathcal{S} + H_\mathcal{S} \qquad [225]$$

where $\mathbf{J}_\mathcal{S}$ is the total entropy flux and $H_\mathcal{S}$ is the internal entropy production, we compare the energy and the entropy equations [220] and [225] taking into account that, for each incompressible phase, $d\mathcal{S}_j = C_j\, dT/T = d\mathcal{U}_j/T$. After some algebra, one gets

$$\mathbf{J}_\mathcal{S} = \phi\rho_f\mathcal{S}_f\mathbf{v}_f + (1-\phi)\rho_m\mathcal{S}_m\mathbf{v}_m + \frac{\mathbf{q}}{T} \qquad [226]$$

$$TH_\mathcal{S} = -\frac{1}{T}\mathbf{q}\cdot\nabla T - \Delta P\frac{D_\omega\phi}{Dt} + \Delta\mu\Delta\Gamma + \Psi + \overline{\rho}\overline{H} \qquad [227]$$

where we have introduced the difference in chemical potentials between the two phases

$$\Delta\mu = \Delta\mathcal{H} - T\Delta S \qquad [228]$$

and $\Delta S = \mathcal{S}_m - \mathcal{S}_f$ is the change in specific entropies.

Following the standard procedure of nonequilibrium thermodynamics, we choose $\mathbf{q} = -k\nabla T$ and we assume that there is a linear relationship between the two thermodynamic fluxes, $D_\omega\phi/Dt$ and $\Delta\Gamma$, and the two thermodynamic forces $-\Delta P$ and $\Delta\mu$ since they have the same tensorial rank. We write

$$\begin{pmatrix} D_\omega\phi/Dt \\ \Delta\Gamma \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}\begin{pmatrix} -\Delta P \\ \Delta\mu \end{pmatrix} \qquad [229]$$

The matrix of phenomenological coefficients m_{ij} is positive definite and symmetrical by Onsager's theorem, $m_{12} = m_{21}$ (see Sramek et al., 2007). For a 2×2 matrix, it is rather simple to show that the positivity implies $m_{22} > 0$, $m_{11} > 0$, and $m_{11}m_{22} - m_{12}^2 > 0$ (positivity of the determinant).

The form of the phenomenological coefficients m_{ij} can be constrained through thought experiments. First, using mass conservations [205] and [206] and the definitions of $\mathbf{v}_\omega$ and ρ^ω, [223] and [224], we can combine equations [229] to get

$$\Delta P = -\frac{m_{22}}{m_{11}m_{22} - m_{12}^2}\left[(1-\omega)(1-\phi)\nabla\cdot\mathbf{v}_m - \omega\phi\nabla\cdot\mathbf{v}_f + \left(\frac{\rho^\omega}{\rho_f\rho_m} - \frac{m_{12}}{m_{22}}\right)\Delta\Gamma\right] \qquad [230]$$

In the limiting case where the two phases have the same density $\rho_f = \rho_m = \rho^\omega$, melting can occur with no motion, $\mathbf{v}_m = \mathbf{v}_f = 0$, and eqn [230] should therefore predict the equality of pressure between phases, $\Delta P = 0$. In this case, we must choose $m_{12}/m_{22} = 1/\rho_f = 1/\rho_m$. Let us consider now a situation of homogeneous isotropic melting where the melt has such a low viscosity that it cannot sustain viscous stresses and cannot interact with the solid by Darcy terms. For such an inviscid melt, $\omega = 0$, $\rho^\omega = \rho_f$ and $\mathbf{v}_\omega = \mathbf{v}_m$. In this case, since the melt can escape instantaneously, the matrix should not dilate, $\nabla\mathbf{v}_m = 0$, and thus, the two pressures should also be the same, $\Delta P = 0$. In this situation, all the terms in eqn [230] are 0 except for the term proportional to $\Delta\Gamma$. Thus, in the general case,

$$\frac{m_{12}}{m_{22}} = \frac{\rho^\omega}{\rho_f\rho_m} \qquad [231]$$

Using this condition and introducing two positive coefficients, $\zeta = m_{22}/(m_{11}m_{22} - m_{12}^2)$, and $\mathcal{R} = m_{22}$, we can recast eqn [229] as

$$\Delta P = -\zeta\left[\frac{D_\omega\phi}{Dt} - \frac{\rho^\omega}{\rho_f\rho_m}\Delta\Gamma\right] \qquad [232]$$

$$\Delta\Gamma = \mathcal{R}\left[\Delta\mu - \frac{\rho^\omega}{\rho_f\rho_m}\Delta P\right] \qquad [233]$$

The first equation establishes a general relation controlling the pressure drop between phases. The coefficient ζ that links the pressure jump between the two phases to the porosity changes in excess of the melting rate is in fact equivalent to a bulk viscosity as introduced in Section 7.02.3.2 (see also the summary in Section 7.02.5.2.5). The physical requirement that the two-phase mixture should have the incompressible properties of either the matrix or the fluid when $\phi = 0$ or $\phi = 1$ imposes a porosity dependence to ζ with $\lim_{\phi\to 0}\zeta(\phi) = \lim_{\phi\to 1}\zeta(\phi) = +\infty$. Simple micromechanical models (e.g., Bercovici et al., 2001a; Nye, 1953; Schmeling, 2000; Simpson et al., 2010b) allow us to estimate the bulk viscosity as

$$\zeta = K_0\frac{(\eta_f + \eta_m)}{\phi(1-\phi)} \qquad [234]$$

The dimensionless constant K_0 accounts for grain/pore geometry and is of $O(1)$.

A more general, but more hypothetical, interpretation of the entropy positivity could argue that some deformational work, Ψ, might affect the pressure drop of eqn [232]. This hypothesis led to a two-phase damage theory developed in Bercovici and Ricard (2003, 2005), Bercovici et al. (2001a,b), and Ricard and Bercovici (2003) (see also Bercovici et al., this volume). Here, we assume that the system remains close enough to mechanical equilibrium that damage does not occur.

The second equation [233] controls the kinetics of the melting/freezing and, by consequence, defines the equilibrium condition. In the case of mechanical equilibrium, when there is no pressure drop between the two phases, the melting rate cancels when there is equality of the chemical potentials of the two single phases. In case of mechanical disequilibrium, ($\Delta P \neq 0$), the chemical equilibrium does not occur when the two chemical potentials are equal. We define a new effective chemical potential,

$$\mu_i^* = \mathcal{U}_i + \frac{P^\omega}{\rho_i} - TS_i \qquad [235]$$

where i stands for f or m, and write the kinetic equation [233],

$$\Delta\Gamma = \mathcal{R}\Delta\mu^* \qquad [236]$$

Chemical equilibrium imposes the equality of the effective potentials on the interface, at the pressure P^ω at which the phase change effectively occurs.

Using eqns [232] and [233], we can show that the entropy production is indeed positive and given by

$$TH_S = k\frac{1}{T}|\nabla T|^2 + \frac{\Delta P^2}{\zeta} + \mathcal{R}(\Delta\mu^*)^2 + \Psi + \overline{\rho}\overline{H} \qquad [237]$$

Chemical relaxation and bulk compression are associated with entropy sources, in addition to diffusion, viscous dissipation, and radioactive heating (see also Rudge et al., 2011).

7.02.5.2.5 Summary equations

For convenience, we summarize the governing equations when the matrix is much more viscous than the fluid phase ($\eta_f \ll \eta_m$) as typical for melting scenarios, which implies that $\underline{\boldsymbol{\tau}}_f = 0$, $\omega = 0$, $\rho^\omega = \rho_f$, $P^\omega = P^\omega = P_f$, and $\mathbf{v}_\omega = \mathbf{v}_m$.

The mass conservation equations are [205] and [206]. The equation of conservation of momentum for the fluid phase is

$$-\phi\nabla[P_f + \rho_f gz] + c\Delta\mathbf{v} = 0 \qquad [238]$$

(assuming z positive upward). This is the Darcy law with $c = \eta_f\phi^2/k(\phi)$, where k is the permeability (often varying as $k_0\phi^n$ with $n = 2$ or 3 (Simpson et al., 2010b)). The second momentum equation could be the matrix momentum (eqn [212]) or a combined force-difference (or action–reaction) equation:

$$-\nabla[(1-\phi)\Delta P] + (1-\phi)\Delta\rho\mathbf{g} + \nabla\cdot\left[(1-\phi)\underline{\boldsymbol{\tau}}_m\right] - \frac{c\Delta\mathbf{v}}{\phi} = 0 \qquad [239]$$

where the deviatoric stress in the matrix is given by

$$\underline{\boldsymbol{\tau}}_m = \eta_m\left(\nabla\mathbf{v}_m + [\nabla\mathbf{v}_m]^T - \frac{2}{3}\nabla\cdot\mathbf{v}_m\underline{\mathbf{I}}\right) \qquad [240]$$

and the pressure jump between phases, eqn [232], becomes

$$\Delta P = -\zeta(1-\phi)\nabla\cdot\mathbf{v}_m \qquad [241]$$

if an equivalent bulk viscosity is used, or

$$\Delta P = -K_0\eta_m\frac{\nabla\cdot\mathbf{v}_m}{\phi} \qquad [242]$$

from the micromechanical model, eqn [234], of Bercovici et al. (2001a) or of Simpson et al. (2010b).

The action–reaction equation [239] can also be written in a different way, for example, by elimination of $\Delta\mathbf{v}$ taken from the Darcy equilibrium [238],

$$-\nabla P_f + \nabla\cdot\left[(1-\phi)\underline{\boldsymbol{\tau}}_m^*\right] + \overline{\rho}\mathbf{g} = 0 \qquad [243]$$

where $\underline{\boldsymbol{\tau}}_m^*$ includes the ΔP term and is defined by

$$\underline{\boldsymbol{\tau}}_m^* = \eta_m\left(\nabla\mathbf{v}_m + [\nabla\mathbf{v}_m]^T - \frac{2}{3}\nabla\cdot\mathbf{v}_m\mathbf{I}\right) + \zeta(1-\phi)\nabla\cdot\mathbf{v}_m\mathbf{I} \qquad [244]$$

This shows that if the pressure is defined everywhere as the fluid pressure, then it is equivalent to use for the matrix a rheology (see eqn [95]), with a bulk viscosity $(1-\phi)\zeta \approx \zeta$ (McKenzie, 1984). This analogy only holds without surface tension between phases (Bercovici and Ricard, 2003; Bercovici et al., 2001a; Ricard et al., 2001).

The rate of melting is controlled by

$$\Delta\Gamma = \mathcal{R}\left(\Delta\mathcal{U} + P_f\left(\frac{1}{\rho_m} - \frac{1}{\rho_f}\right) - T\Delta S\right) \qquad [245]$$

and the energy equation is

$$\begin{aligned}\rho_f\phi C_f\frac{D_f T}{Dt} + \rho_m(1-\phi)C_m\frac{D_m T}{Dt} - \Delta\mathcal{H}\Delta\Gamma \\ = \overline{\rho}\overline{H} - \nabla\cdot\mathbf{q} + \frac{\Delta\Gamma^2}{\mathcal{R}} + K_0\eta_m\frac{1-\phi}{\phi}(\nabla\cdot\mathbf{v}_m)^2 + \Psi\end{aligned} \qquad [246]$$

where we have assumed the relation [242].

These equations have been used by many authors with various levels of approximation (Simpson et al., 2010a). The most benign have been to replace $1-\phi$ by 1. Most authors have also considered the bulk viscosity ζ as a porosity-independent parameter (e.g., Choblet and Parmentier, 2001; Katz et al., 2004; Kelemen et al., 1997; McKenzie, 1984; Ribe, 1985a,b; Richter and Mckenzie, 1984; Scott and Stevenson, 1984, 1986, 1989; Spiegelman and Kelemen, 2003; Spiegelman et al., 2001; Stevenson, 1989). This overestimates the possibilities of matrix compaction at low porosity. Porosity-dependent parameters have been explicitly accounted for in other papers (e.g., Bercovici et al., 2001a; Connolly and Podladchikov, 1998; Fowler, 1985; Rabinowicz et al., 2002; Ricard et al., 2001; Schmeling, 2000; Sleep, 1988).

For direct application to magma extraction, an earlier work by Spiegelman and McKenzie (1987) imposed matrix motions and solved for the magma trajectories that were found attracted toward ridges and subduction corners. The melting rates were later sometimes imposed (e.g., Turcotte and Morgan, 1992), sometimes solved according to the energy equation, assuming univariant melting (Fowler, 1989; Hewitt and Fowler, 2008; Sramek et al., 2007) or even multiple component melting (Hewitt, 2010; Rudge et al., 2011). Surface tension (between matrix and magma and between matrix grains) has been considered (Bercovici and Ricard, 2003; Bercovici et al., 2001a;

Hier-Majumder et al., 2006; Riley et al., 1990). The situation where the melt is denser than the matrix has been considered by Karato et al. (2006) and by Hernlund and Jellinek (2010) for applications to the mantle transition zone and to the deep mantle, two layers where the presence of magma is suspected from seismological observations (Tauzin et al., 2010; Williams and Garnero, 1996). The interaction between the matrix and the magma can lead to the channelization of the melt (Hewitt and Fowler, 2009; Weatherley and Katz, 2012) and to the development of a fabric of the matrix that favors the migration of the melt to the ridge (Holtzman et al., 2003).

Similar equations have also been used to describe the interaction between iron and silicates near the CMB (Buffett et al., 2000) or during the segregation of planetary cores (Golabek et al., 2008; Ricard et al., 2009; Sramek et al., 2010, 2012), the formation of dendrites (Poirier, 1991), the compaction of lava flows (Massol et al., 2001; Michaut and Bercovici, 2009), and the dynamics in volcanic conduits (Bercovici and Michaut, 2010; Michaut et al., 2009, 2013).

7.02.6 Specifics of Earth's Mantle Convection

In this last section, we discuss various aspects of physics unique to large-scale mantle convection. We are aware of the impossibility to be exhaustive, but most of the important points are more deeply developed in other chapters of the treatise (in particular, we leave the problems of partial melting to Parmentier and to Ito and Van Keken, the connections with geochemistry to Tackley, and the aspects of plate tectonics to Bercovici et al., this volume; see also the books by Schubert et al. (2001) and Davies (1999)).

7.02.6.1 A Negligible Inertia

The most striking difference between mantle convection and most other forms of convection is that inertia is totally negligible. This is because the Prandtl number is much larger than the (already very large) Rayleigh number. This implies that the mantle velocity flow obeys

$$\begin{aligned} \boldsymbol{\nabla}\cdot(\overline{\rho}\mathbf{v}) &= 0 \\ -\boldsymbol{\nabla}P + \boldsymbol{\nabla}\cdot\underline{\boldsymbol{\tau}} + \delta\rho\mathbf{g} + \overline{\rho}\delta\mathbf{g} &= 0 \\ \nabla^2\psi &= 4\pi G\delta\rho \\ \delta\mathbf{g} &= -\boldsymbol{\nabla}\psi \end{aligned} \qquad [247]$$

in agreement with eqn [137]. In this set of equation, we keep the self-gravitation term as appropriate at long wavelengths. If the internal loads $\delta\rho$ are known, the flow can be computed independently of the temperature equation. This time-independent system has been used by many authors to infer the mantle flow properties.

7.02.6.1.1 Dynamic models

The system of eqn [247] can been solved analytically for a depth-dependent viscosity, when variables are expressed on the basis of spherical harmonics (see Hager and Clayton, 1989 and also Ribe, this volume and Forte, vol. 3). Various possible surface observables (geoid height or gravity free-air anomalies, velocity divergence, amplitude of deviatoric stress at the surface, surface dynamic topography, CMB topography, etc.) can be expressed on the basis of spherical harmonics with components O_{lm}. Through eqn [247], they are related to the spherical harmonics components of the internal density variations $\delta\rho_{lm}(r)$ by various degree-dependent Green functions (Green, 1793–1841):

$$O_{lm} = \int G_l^O(r)\delta\rho_{lm}(r)\,\mathrm{d}r \qquad [248]$$

(see Ribe, this volume, for analytic details). The Green functions $G_l^O(r)$ can be computed from the averaged density and viscosity profiles.

Before seismic imaging gave us a proxy of the 3-D density structure of the mantle, various theoretical attempts have tried to connect models of mantle convection to plate velocities (Hager and Oconnell, 1979; Hager and O'Connell, 1981), to Earth's gravity field (or to the geoid, proportional to the potential ψ), to the lithospheric stress regime, or to the topography (Bai et al., 1992; Lago and Rabinowicz, 1984; Parsons and Daly, 1983; Ricard et al., 1984; Richards and Hager, 1984; Runcorn, 1964).

An internal load of negative buoyancy induces a downwelling flow that deflects the Earth's surface, the CMB, and any other internal compositional boundaries, if they exist. The amount of deflection corresponds to the usual isostatic rule for a load close to an interface: the weight of the induced topography cancels to first order the weight of the internal load. The total gravity anomaly resulting from a given internal load is affected by the mass anomalies associated with the flow-induced boundary deflections as well as by the load itself. Due to the deflection of the Earth's surface, the geoid perturbation induced by a dense sinking anomaly is generally negative (e.g., free-air gravity has a minimum above a dense load). However, when the mantle viscosity increases significantly with depth, by 1–2 orders of magnitude, a mass anomaly close to the viscosity increase, induces a larger CMB deformation and a lower surface deformation. The resulting gravity anomaly corresponds to a geoid high. The fact that cold subduction zones correspond to a relative geoid high suggests a factor $\geq$30 viscosity increase around the upper–lower mantle interface (Hager et al., 1985; Lago and Rabinowicz, 1984; Ricard et al., 1984; Richards and Hager, 1984). Shallow anomalies and anomalies near the CMB, being locally compensated, do not contribute to the long-wavelength gravity field. The lithospheric stress field, like the geoid, is affected by mid-mantle density heterogeneities. The surface deflection induced by a deep-seated density anomaly decreases with the depth of this anomaly, but even lower mantle loads should significantly affect the surface topography.

7.02.6.1.2 Mantle flow and postglacial models

As soon as seismic tomography started to image the mantle structures, these seismic velocity anomalies have been used to further constrain the mantle viscosity. The fact that the geoid and seismic velocity anomalies are positively correlated around the transition zone but negatively in the deep mantle heterogeneities suggests a viscosity larger than 10 but not too large (less than 100); otherwise, the mantle would be everywhere

positively correlated with gravity (Hager and Clayton, 1989; Hager et al., 1985; King and Hager, 1994; Ricard et al., 1984; Richards and Hager, 1984) (see **Figure 6**). The same modeling approach, assuming a proportionality between seismic velocity anomalies and density variations, was also used to match the observed plate divergence (Forte and Peltier, 1987), the plate velocities (Ricard and Vigny, 1989; Ricard et al., 1991), and the lithospheric stresses (Bai et al., 1992; Lithgow-Bertelloni and Guynn, 2004). The initial Boussinesq models were extended to account for compressibility (Forte and Peltier, 1991). The instantaneous flow computations that were initially restricted to long-wavelength patterns have recently capitalized on advances in adaptative mesh refinements on parallel computers to reach very fine resolutions, down to kilometer scales (Alisic et al., 2012; Stadler et al., 2010).

Joint inversions of gravity with postglacial rebound were also performed to further constrain the mantle viscosity profile. The viscosity increase required by subduction was initially thought to be too large to reconcile with postglacial rebound (Peltier, 1996). The various approaches (time-dependent for the postglacial models and time-independent for the geoid models) seem to have converged to a standard viscosity profile with a significant increase with depth (Mitrovica and Forte, 1997). Whether this viscosity increase occurs across a discontinuity (at the upper–lower mantle interface or deeper) or as a gradual increase is probably beyond the resolution of these approaches.

Although these dynamic models explain the observed geoid, they require surface topography of order 1 km, induced by mantle convection and called dynamic topography (in contrast to the isostatic topography related to crustal and lithospheric density variations). The direct observation of dynamic topography, however, is not evident in Earth's topography corrected for isostatic crustal contributions (e.g., Colin and Fleitout, 1990; Kido and Seno, 1994; Lestunff and Ricard, 1995; Lithgow-Bertelloni and Silver, 1998). However, in recent years, the role of a deep dynamic support to explain the present-day topography and the flooding and emergence of continent interiors during the geologic times have been documented by various authors (Conrad and Gurnis, 2003; Conrad and Husson, 2009; Flament et al., 2013; Husson and Conrad, 2006; Moucha et al., 2008).

7.02.6.1.3 Time-dependent models

The thermal diffusion in the mantle is so slow that even over 100–200 My, it can be neglected in some long-wavelength global models. Equation [247] can thus be solved by imposing the known paleoplate velocities at the surface and advect the mass anomalies with the flow without explicitly solving the energy equation. This forced-convection approach has shown that the deep mantle structure is mostly inherited from the Cenozoic and Mesozoic plate motion (Lithgow-Bertelloni and Richards, 1998; Richards and Engebretson, 1992). From plate paleoslab reconstructions only, a density model can be obtained that gives a striking fit to the observed geoid and is in relative agreement with long-wavelength tomography (Ricard et al., 1989). This approach was also used to study the hot spot fixity (Richards, 1991; Steinberger and O'Connell, 1998), the sea level changes (Lithgow-Bertelloni and Gurnis, 1997), or the polar wander of the Earth (Richards et al., 1997; Spada et al., 1992a).

Instead of an imposed forced convection, a promising method proposes to infer self-coherently the past mantle structure from the whole set of convection equations, including temperature evolution (Bunge et al., 2003; Ismail-Zadeh et al., 2004; Liu and Gurnis, 2008). Using tools of variational data assimilation developed by meteorologists, the past density heterogeneities can be considered as unknowns to be inverted for in such a way that their time evolution is in agreement with present-day observations (e.g., with tomographic models), paleoreconstructions (the geometry and velocities of plates), or paleogeography (Shephard et al., 2010).

7.02.6.2 A Mantle with Internal Heating

When the top and bottom boundary conditions are the same (i.e., both free-slip or both no-slip), purely basally heated convection in a Cartesian box leads to a perfectly symmetrical system. We could simultaneously reverse the vertical axis and the color scale of **Figure 4** and get temperature patterns that are also convective solutions. The convective fluid has a near adiabatic core and the temperature variations are confined into two boundary layers, a hot bottom layer and a cold top layer. The thicknesses of these two boundary layers and the temperature drops across them are the same. The middepth temperature is simply the average of the top and bottom temperatures. Instabilities develop from the bottom layer (hot rising plumes) and the cold layer (cold downwelling plumes). They have a temperature hotter or colder than the depth-dependent average temperature. They are active structures driven by their intrinsic positive or negative buoyancy. The Earth's mantle has however a large number of characteristics that break the symmetry between upwellings and downwellings.

What is probably the major difference between mantle convection and purely basally heated convection is that the Earth is largely powered by radiogenic heating from the decay of uranium, thorium, and potassium and loss of primordial heat left over from its accretion. Convection purely heated from within is depicted in **Figure 7**. In the extreme case where the fluid is entirely heated from within, the fluid has no hot bottom boundary layer. There are only concentrated downwelling currents sinking from the top cold boundary layer. The downwellings are active as they are moved by their own negative buoyancy. To compensate for the resulting downwelling flow, the background rises passively, that is, without being pushed up by a positive buoyancy (Bercovici et al., 2000). In the case of basal heating, any plume leaving the top or bottom boundary layer travels adiabatically (neglecting diffusion and shear heating). However, in the case of internal heating, while the rapid downwellings remain close to adiabatic, the radioactive decay can accumulate heat within the slow upwellings. This heating is opposite to the adiabatic cooling and the average temperature in an internally heated system remains more homogeneous and with a significant subadiabatic gradient (Parmentier et al., 1994).

The Earth's mantle is however not in such an extreme situation. Some heat flow is extracted across the CMB from the molten iron outer core. This basal heat flux drives active upwellings (hot spots). The ratio of the internal radioactive heat to the total heat extracted at the Earth's surface is called the Urey number (Urey, 1951). Geochemical models of mantle composition (McDonough and Sun, 1995; Rudnick and Fountain, 1995) imply that about 50% of the surface heat flux is due to mantle and core cooling and only 50% or even

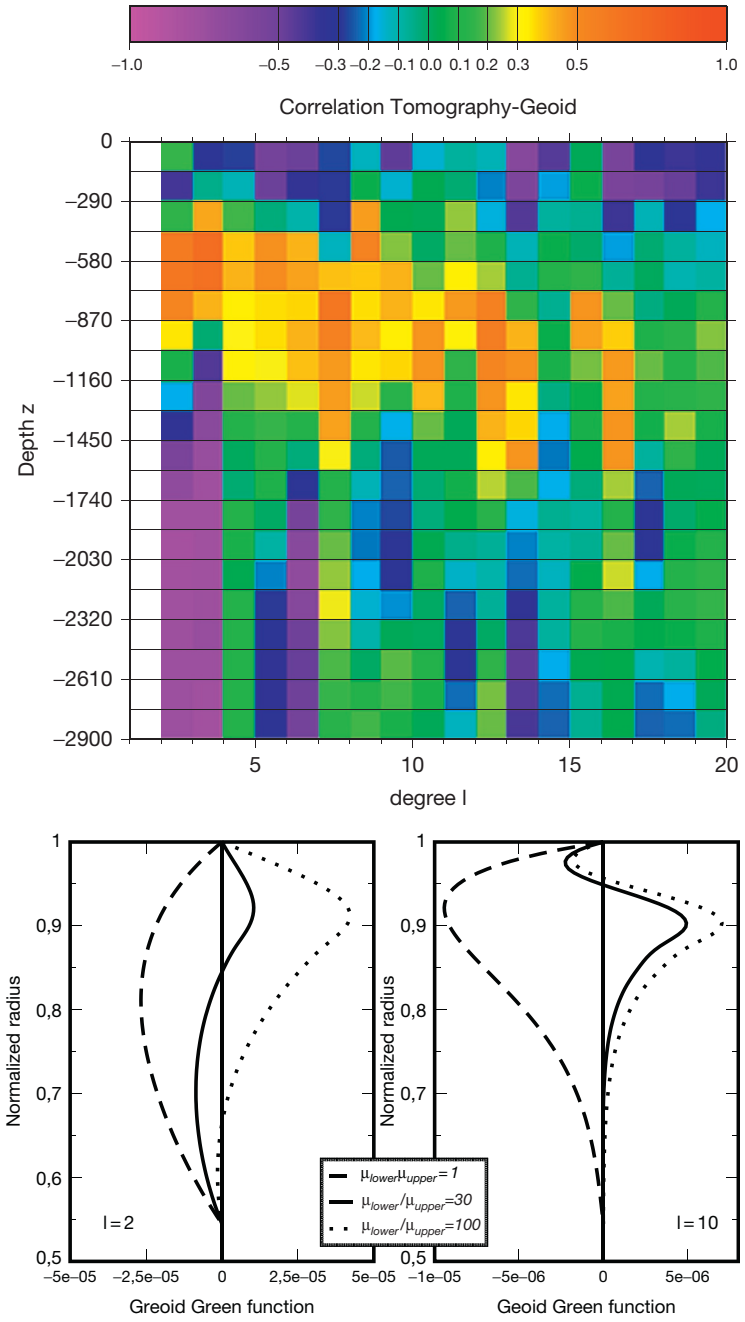

Figure 6 Correlations between gravity and the synthetic tomographic SMEAN model (Becker and Boschi, 2002) as a function of degree *l* and normalized radius (top). The seismic velocities, proxy of the density variations, are positively correlated with gravity around the upper–lower mantle interface (warm colors) but negatively correlated near the surface and in the deep lower mantle (cold colors). Geoid Green functions for degree 2 (bottom left) and degree 10 (bottom right) and three possible viscosity increases between upper mantle and lower mantle. The geoid Green function for a uniform viscosity (dashed line) is everywhere negative, and all the anomalies around the upper–lower mantle would induce a gravity signal opposite to that observed. A too large viscosity increase (a factor 100 for the dotted lines) cannot explain the rather good negative correlation between lower mantle anomalies and the geoid at long wavelength. A moderate increase (a factor 30 for the solid line) leads to the best fit as the sign of the Green functions is everywhere that of the observed density–gravity correlations. The different Green functions are computed for an incompressible mantle, with a lithosphere, ten times more viscous than the upper mantle.

less (Lyubetskaya and Korenaga, 2007) to radioactive decay. Generally, geophysicists have difficulties with these numbers as they seem to imply a too large mantle temperature in the past (Davies, 1980; Schubert et al., 1980). From convection modeling of the Earth's secular cooling, they often favor ratios of the order of 80% radioactive and 20% cooling. A more complex coupling between core and mantle cooling (Labrosse et al., 2007), or the peculiar properties of the lithosphere, may reconcile the thermal history of the Earth with a low radiogenic content (Grigne et al., 2005; Korenaga, 2003) (see Jaupart et al. and Hernlund and McNamara, this volume). The basal heat flux at the CMB represents the core cooling component that is

part of the total cooling rate of the Earth. The secular cooling and the presence of internal sources tend to decrease the thickness of the hot bottom layer compared to that of the cold top layer, increase the active role of downwellings (the subducting slabs), and decrease the number or the strength of the active upwellings (the hot spots).

7.02.6.3 A Complex Rheology

We have shown that the rheological laws of crystalline solids may be linear or nonlinear, depending on temperature, grain size, and stress level. Various deformation mechanisms (grain diffusion, grain boundary diffusion, dislocation creep, grain boundary sliding, etc.) act simultaneously. The equivalent viscosity of each individual mechanism can be written in the form

$$\eta = AI_2^{-\frac{(n-1)}{2}} d^{-m} \exp\frac{E^* + PV^*}{RT} \qquad [249]$$

where E^* and V^* are the activation energy and volume, P and T the pressure and temperature, R the perfect gas constant, d the grain size, m the grain size exponent, I_2 the second stress invariant, and n a stress exponent (Ranalli, 1995; Weertman and Weertman, 1975). The multiplicative factor A varies with water content, melt content, and mineralogy. In general, the composite rheology is dominated by the mechanism leading to the lowest viscosity.

In Figure 8, we plot as a function of temperature, and for various possible grain sizes (0.1 mm, 1 mm, and 1 cm), the stress rate at which the strain rate predicted for the dislocation mechanism and that for diffusion mechanism are the same (see eqns [108] and [107]). The data correspond to dry upper mantle (Karato and Wu, 1993). Low stress and temperature favor diffusion creep, while high stress and high temperature favor dislocation creep. These experiments suggest that in the upper mantle or at least in its shallowest part, nonlinear creep is likely to occur. At greater depth, the decrease in the average deviatoric stress should favor a diffusive regime with a Newtonian viscosity. The observation of rheological parameters at lower mantle conditions is more difficult, but the lower mantle should mostly be in diffusive creep regime except the zones of intense shear around subductions (McNamara et al., 2001).

The lateral variations of viscosity due to each separate parameter, stress exponent, temperature, water content, or grain size can potentially be very large. Surprisingly, attempts to deduce these variations directly from geodynamic observations have not been very successful. Attempts to explain the Earth's gravity from internal loads do not seem to require lateral viscosity variations in the deep mantle (Zhang and

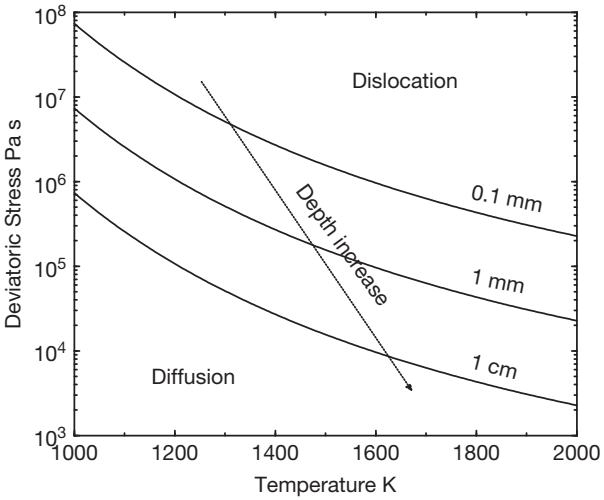

Figure 8 Creep regime map for dry olivine. High deviatoric stress or temperature favors a dislocation mechanism. A decrease in the grain size favors diffusion. In the upper mantle, the stress and temperature conditions tend to bring the creep regime from dislocation to diffusion at depth ($n = 2.5$, $m = 0$, and $E^* = 540$ kJ mol^{-1} for dislocation creep and $n = 0$, $m = 2.5$, and $E^* = 300$ kJ mol^{-1} for diffusion creep).

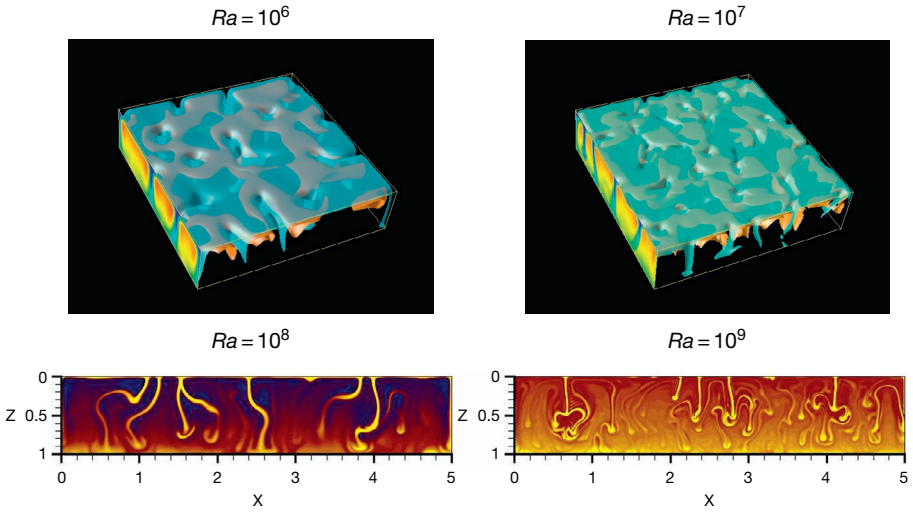

Figure 7 Convection patterns of a fluid entirely heated from inside at Rayleigh number 10^6, 10^7, 10^8, and 10^9 (simulation compliments of Fabien Dubuffet). Cold fingerlike instabilities are sinking from the top boundary layer and spread on the bottom boundary layer. No active upwellings are present (compare with convection patterns for a fluid heated from below, Figure 4).

Christensen, 1993). Near the surface, viscosity variations are present, at least between continental and oceanic lithospheres (Cadek and Fleitout, 2003; Ricard et al., 1991). The gain in fitting the Earth's gravity or postglacial rebound data with 3-D viscosity models remains rather moderate compared to the complexities added to the modeling (Gasperini et al., 2004), and most models of mantle viscosity are restricted to radial profiles (Mitrovica and Forte, 1997). Even the modeling of slabs, their curvatures, and their stress patterns do not really require that they are much stiffer than the surrounding mantle (Tao and Oconnell, 1993).

7.02.6.3.1 Temperature dependence of viscosity

Mantle viscosity is a strong function of temperature and the cold lithosphere seems to have a viscosity of order 10^{25} Pa s (Beaumont, 1978), 4–6 orders of magnitude stiffer than the asthenosphere. The activation energy E^* is typically from 300 to 600 kJ mol^{-1} (Drury and FitzGerald, 1998) the lowest values being for diffusion creep. This implies a factor ≈ 10 in viscosity decrease for a 100 K temperature increase (using $T \approx 1600$ K). The effect of temperature dependence of viscosity on the planform of convection was recognized experimentally using oils or syrups (Booker, 1976; Nataf and Richter, 1982; Richter, 1978; Weinstein and Christensen, 1991). In the case of a strongly temperature-dependent viscosity, the definition of the Rayleigh number is rather arbitrary as the maximum, the minimum, or some average viscosity can be chosen in its definition. Another nondimensional number (e.g., the ratio of viscosity variations, η_{max}/η_{min}) must be known to characterize the convection.

Not surprisingly, two extreme regimes are found. For a viscosity ratio lower than about 100, the convection pattern remains quite similar to convection with uniform viscosity. On the other hand, if the viscosity of the cold boundary layer (the lithosphere) is more than 3000 times that of the underlying asthenosphere, the surface becomes stagnant (Solomatov, 1995). Below this immobile lid, the flow resembles convection below a rigid top surface (Davaille and Jaupart, 1993). In between, when the viscosity ratios are in the range 100–3000, the lithosphere deforms slowly, and in this sluggish regime, the convection cells have large aspect ratios.

Convection with temperature-dependent viscosity has been investigated by various authors (Christensen, 1984b; Kameyama and Ogawa, 2000; Parmentier et al., 1976; Tackley et al., 1993; Trompert and Hansen, 1998b). Since the top boundary layer is stiffer than the bottom boundary layer, the top boundary layer is also thicker than the bottom one. This impedes surface heat removal, reduces the heat flux across the bottom boundary layer, and raises the average mantle temperature. Convection patterns computed with T-dependent viscosity remain however quite far from Earth-like convection. The major difference is that when the T dependence is too strong, the surface freezes and becomes immobile, while on the real Earth, the lithosphere is highly viscous but broken into tectonic plates separated by weak boundaries. Without mechanisms other than a simple T dependence of viscosity, the Earth would be in a stagnant-lid regime (see Bercovici et al., this volume).

Various modelers have thus tried to use T-dependent rheologies but have imposed a platelike surface velocity. This has been very useful to understand the initiation of subduction (Toth and Gurnis, 1998), the interaction of slabs with the phase changes in the transition zone (Christensen, 1996, 1997b), and the relationship between subduction and gravity (King and Hager, 1994). These numerical experiments, mostly intended to model slabs, compare satisfactorily with laboratory experiments (Guilloufrottier et al., 1995; Kincaid and Olson, 1987).

To conclude this brief section on temperature dependence of viscosity, we discuss the general concept of self-regulation of planetary interiors (Tozer, 1972). If a planet were convecting too vigorously, it would lose more heat than radioactively produced. It would therefore cool down until the viscosity is large enough to reduce the heat transfer. In contrast, a planet convecting too slowly would not extract its radioactive energy and would heat up until the viscosity is reduced sufficiently to increase the surface heat flow (see also Jaupart et al. this volume). The internal temperature of planets is mostly controlled to the activation energy (or rather enthalpy) of the viscosity (assuming that planets have similar amount of heat sources). To first order, large and small terrestrial rocky planets probably have the same internal temperatures.

7.02.6.3.2 Depth dependence of viscosity

The activation volume V^* of the viscosity [249] is typically around 10^{-5} m^3 mol^{-1}. Extrapolating to CMB conditions, this suggests a large viscosity increase throughout the mantle. However, measurements of viscosity at both high T and P conditions are very difficult (see also Weidner, vol. 2). The viscosity increase by a factor 30–100 suggested by geodynamics (see **Section 7.02.6.1**) is probably as robust a constraint as could be deduced from mineral physics experiments.

The effect of a depth-dependent viscosity on the planform of convection has been studied by, for example, Gurnis and Davies (1986), Cserepes (1993), and Dubuffet et al. (2000). At least two important geodynamic observations can be explained by an increase of viscosity with depth. One is the relative stability of hot spots. A sluggish lower mantle where convection is decreased in intensity by a larger viscosity (and also by a smaller expansivity and a potentially larger thermal conductivity as discussed in **Section 7.02.6.5.3**) favors the relative hot spot fixity (Richards, 1991; Steinberger and O'Connell, 1998). A second consequence is a depth dependence of the wavelengths of the thermal heterogeneities. A viscosity increase (together with the existence of plates and continents that impose their own wavelengths (see **Section 7.02.6.7**)) induces the existence of large-scale thermal anomalies at depth (Bunge and Richards, 1996). A slab crossing a factor 30–100 viscosity increase should thicken by a factor of order 3–5 (Gurnis and Hager, 1988; Ribe et al., 2007). This thickening is observed in tomographic models (van der Hilst et al., 1997) and can be inferred from geoid modeling (Ricard et al., 1993a). When the viscosity increases with pressure, the self-regulation effect causes the mantle to heat up. The temperature dependence of the viscosity then mitigates the effects of the pressure dependence (Tackley et al., 2013).

7.02.6.3.3 Stress dependence of viscosity

Starting from Parmentier et al. (1976), the effect of a stress-dependent viscosity has been studied by Christensen (1984a), Malevsky and Yuen (1992), van Keken et al. (1992), and Larsen

et al. (1993), assuming either entirely nonlinear or composite rheologies (where deformation is accommodated by both linear and nonlinear mechanisms). At moderate Rayleigh number, the effect of a nonlinear rheology is not very significant. In fact, the nonlinearity in the rheology is somewhat opposed to the temperature dependence of the rheology. As shown by Christensen (1984a), a T-dependent, nonlinear rheology with an exponent $n \approx 3$ leads to convection cells rather similar to what would be obtained with a linear rheology and an activation energy divided by $\approx n$. Convection with both nonlinear and T-dependent rheology looks more isoviscous than convection with only stress-dependent or only T-dependent rheologies.

At large Rayleigh number, however, nonlinear convection becomes more unstable (Malevsky and Yuen, 1992), and the combination of nonlinear rheology, T-dependent rheology, and viscous dissipation can accelerate the rising velocity of hot plumes by more than an order of magnitude (Larsen and Yuen, 1997).

7.02.6.3.4 Grain size dependence of viscosity

In the diffusion regime, the viscosity [249] can be strongly variable as controlled by the grain size d to a power m of order 3 (Karato et al., 1986). The difficulty to estimate the rheological conditions of the mantle is that the grain size d of eqn [249] has its own dynamics with various feedbacks with the deformation itself. In the diffusive regime, the grains are slowly coarsening: the slight overpressure of small grains due to their surface tension drives a diffusional flow toward the larger grains (e.g., Hillert, 1965). In contrast, grains undergoing dislocation creep can go through recrystallization and therefore can reduce in size. These two opposite mechanisms should drive the mantle material in a situation where diffusion and dislocation creeps have roughly the same efficiency (De Bresser et al., 2001; Rozel et al., 2011). The existence of an equilibrium grain size related to the average deviatoric stress is observed in laboratory experiments (Karato, 2008; Vanderwal et al., 1993) and inferred from thermodynamic considerations (Austin and Evans, 2007; Bercovici and Ricard, 2005; Ricard and Bercovici, 2009).

There is a potential feedback interaction between deformation, grain size reduction by dynamic recrystallization, viscosity reduction, and further localization (Braun et al., 1999; Jaroslow et al., 1996; Kameyama et al., 1997). However, in its simplest version, this localizing mechanism is problematic as it implies simultaneously dynamic recrystallization and diffusion creep that ostensibly occur in exclusive domains of deformation space. It has however been proposed that the presence of minor mineralogical phases can both hinder the grain coarsening (Hiraga et al., 2010) and ease the recrystallization, thus providing a consistent mechanism for localization (Bercovici and Ricard, 2012, 2014; Warren and Hirth, 2006). The effect is potentially important in the mantle and even more important in the lithosphere. More details are given in Bercovici et al. this volume.

7.02.6.4 Importance of Sphericity

An obvious difference between the convection planform in a planet and that of in an experimental tank is due to the sphericity of the former. In the case of purely basally heated convection, the same heat flux (in a statistical sense) has to be transported through the bottom boundary layer and the top boundary layer. However, as the CMB surface is about four times smaller than the top surface, this implies a four times larger thermal gradient through the bottom boundary layer than across the lithosphere. A large gradient across the bottom boundary layer reinforces the upwelling hot instabilities with respect to the downgoing cold instabilities. Sphericity affects the average temperature and the top and bottom boundary layer thicknesses in a way totally opposite to the effects of internal sources (see **Section 7.02.6.2**) or T-dependent viscosity (see **Section 7.02.6.3**). Although numerically more time-consuming, spherical convection models are more and more common (Bercovici et al., 1989a,b, 1992; Bunge et al., 1997; Choblet et al., 2007; Glatzmaier, 1988; Tackley et al., 1993; Zhong et al., 2000). In the last years, they reached a spectacular resolution (down to length scales of 1 km (Alisic et al., 2012; Stadler et al., 2010)) and robustness (handling up to 19 orders of magnitude variations in viscosity (Tackley, 2008)).

7.02.6.5 Other Depth-Dependent Parameters

7.02.6.5.1 Thermal expansivity variations

The thermal expansivity varies with depth, as predicted by the EoS [80], from which we can easily deduce that

$$\alpha = \frac{\alpha_0}{(\rho/\rho_0)^{n-1+q} + \alpha_0 (T - T_0)} \qquad [250]$$

where $n-1+q \simeq 3$.

It decreases with both temperature and density and thus with depth. The expansivity varies from $\approx 4 \times 10^{-5}\ \mathrm{K}^{-1}$ near the surface to $\approx 8 \times 10^{-6}\ \mathrm{K}^{-1}$ near the CMB (Chopelas and Boehler, 1992). This diminishes the buoyancy forces and slows down deep mantle convection (Hansen et al., 1993). Like the increase of viscosity with depth, a depth-dependent thermal expansivity broadens the thermal structures of the lower mantle and suppresses some hot instabilities at the CMB. On the other hand, hot instabilities gain buoyancy as they rise in the mantle, which favors their relative lateral stationarity. In addition to its average depth dependence, the temperature dependence of the expansivity also affects the buoyancy of slabs (Schmeling et al., 2003).

7.02.6.5.2 Increase in average density with depth

To take into account compressibility and the depth dependence of density, the Boussinesq approximation has been replaced by the anelastic approximation in several studies. Such investigations have been carried out by Jarvis and Mckenzie (1980), Glatzmaier (1988), and Bercovici et al. (1992) and since extended to higher Rayleigh numbers (e.g., Balachandar et al., 1992, 1993; Zhang and Yuen, 1996).

One of the difficulties with compressible fluids is that the local criterion for instability (see **Section 7.02.4.2.3**) is related to the adiabatic gradient. Depending on assumptions about the curvature of the reference geotherm with depth (the slope of the adiabatic gradient), part of the fluid can be unstable while the other part is stable. Assuming a uniform adiabatic gradient does not favor the preferential destabilization of either the upper or the lower mantle. On the other hand, assuming that the reference temperature increases exponentially with depth (i.e., taking eqn [123] with constant

parameters) would lead to an easier destabilization of the top of the mantle than of its bottom as a much larger heat flux would be carried along the lower mantle adiabat. In the real Earth, the adiabatic gradient (in K km^{-1}) should decrease with depth (due to the decrease in expansivity $\overline{\alpha}$ with depth insufficiently balanced by the density increase; see eqn [123]). Since less heat can be carried out along the deep mantle adiabat, compressibility should favor the destabilization of the deep mantle.

Compressible convection models generally predict downgoing sheets and cylindrical upwellings reminiscent of slabs and hot spots (e.g., Zhang and Yuen, 1996). Viscous dissipation is positive (as an entropy-related source) but maximum just below the cold boundary layer and just above the hot boundary layer, where rising or sinking instabilities interact with the layered structures of the boundary layers. On the contrary, the adiabatic source heats the downwellings and cools the upwellings. On average, it reaches a maximum absolute value in the midmantle. Locally, viscous dissipation and adiabatic heatings can be larger than radiogenic heat production, although integrated other the whole mantle and averaged over time, the adiabatic and dissipative sources cancel out (see eqn [61]).

7.02.6.5.3 Thermal conductivity variations

The thermal conductivity of a solid is due to two different effects. First, a hot material produces blackbody radiation that can be absorbed by neighboring atoms. This radiative transport of heat is probably a minor component since the mean free path of photons in mantle materials is very small. Second, phonons, which are collective vibrations of atoms, are excited and can dissipate their energies by interacting with other phonons, with defects and with grain boundaries. The free paths of phonons being larger, they are the main contributors to the thermal conductivity.

According to Hofmeister (1999), thermal conductivity should increase with depth by a factor $\approx$2–3. The recent observations of phase transitions in the bottom of the lower mantle should also be associated with another conductivity increase (Badro et al., 2004); this is one more effect (with the viscosity increase and the thermal expansivity decrease) that should decrease the deep mantle convective vigor. Thermal conductivity increase with depth also broadens the thermal anomalies and thins the bottom boundary layer (Dubuffet et al., 1999).

7.02.6.6 Thermochemical Convection

7.02.6.6.1 Density variations in the mantle

Except in Section 7.02.5, a simple negative relationship was assumed between density variations and temperature variations, through the thermal expansivity, $\Delta\rho = -\alpha\rho\Delta T$. However, in the mantle, several sources of density anomalies are present (see also, Stixrude, vol. 1). The density in the mantle varies with the temperature T for a given mineralogical composition, or phase content, symbolized by the symbol ϕ (e.g., for a given proportion of oxides and perovskite in the lower mantle). The mineralogy for a given bulk elemental composition χ (e.g., the proportion of Mg, Fe, and O atoms) evolves with pressure and temperature to maintain the Gibbs energy minimum. The variations of density in the mantle at a given pressure have potentially three contributions that can be summarized as

$$\Delta\rho = \left(\frac{\partial\rho}{\partial T}\right)_{\phi}\Delta T + \left(\frac{\partial\rho}{\partial\phi}\right)_{T}\left(\frac{\partial\phi}{\partial T}\right)_{\chi}\Delta T + \left(\frac{\partial\rho}{\partial\phi}\right)_{T}\left(\frac{\partial\phi}{\partial\chi}\right)_{T}\Delta\chi \quad [251]$$

The first term on the right side is the intrinsic thermal effect computed assuming a fixed mineralogy; we have already discussed this term. The second term is a thermochemical effect. The density is a function of the mineralogical composition controlled at uniform pressure and elemental composition, by the temperature variations. This effect is responsible for a rise in the 410 km deep interface and it deepens the 660 km interface in the presence of cold downwellings (Irifune and Ringwood, 1987). The last term is the intrinsic chemical effect (related to variations of the mineralogy due to changes in the elemental composition at constant temperature). The three contributions have very similar amplitudes and none of them is negligible (Ricard et al., 2005).

7.02.6.6.2 Phase changes

The effect of the second term has been rather well studied (Christensen, 1996; Christensen and Yuen, 1984; Machetel and Weber, 1991; Peltier and Solheim, 1992; Schubert et al., 1975; Tackley, 1995; Tackley et al., 1993). Phase changes in cold downgoing slabs occur at shallower depth in the case of exothermic phase changes and at greater depth for endothermic phase changes (the ringwoodite to oxides plus perovskite phase change at 660 km depth is endothermic; all the important other phase changes of the transition zone are exothermic). These sources of anomalies and their signs are related to the Clapeyron slope of the phase transitions (Schubert et al., 1975). The existence of latent heat release during phase change (see eqn [165]) is a secondary and minor effect. The recent discovery of a phase transformation in the deep lower mantle (Murakami et al., 2004) (the postperovskite phase) suggests that part of the complexities of the D″ layer is related to the interaction between a phase change and the hot boundary layer of the mantle (Nakagawa and Tackley, 2006) (see also Irifune, vol. 2).

The fact that below the normal 660 km depth interface there is a region where slabs remain in a low-density upper mantle phase instead of being transformed into the dense lower mantle phase is potentially a strong impediment to slab penetration. The idea that this effect induces a layering of convection at 660 km or a situation where layered convection is punctuated by large 'avalanche' events dates back to Ringwood and Irifune (1988) and was supported by numerical simulations in the 1990s (e.g., Honda et al., 1993; Machetel and Weber, 1991; Tackley, 1995). It seems however that the importance of this potential effect has been reduced in recent simulations with more realistic Clapeyron slopes, phase diagrams (taking into account both the pyroxene and the garnet phases), thermodynamic reference values (the phase change effect has to be compared with thermal effects and thus an accurate choice for the thermal expansivity is necessary), and viscosity profiles.

7.02.6.6.3 Abyssal layers

In the last years, various complexities were discovered in the deep lower mantle that was previously considered as rather

homogeneous. At small scale, a laterally intermittent layer at the base of D″ (ULVZ, ultralow-velocity zone), with a maximum thickness near 40 km and a strong decrease of the compressional wave velocity, is most simply explained as the result of partial melt at this depth (Williams and Garnero, 1996). A pair of seismic discontinuities observed in some fast (cold) regions of D″ could be the result of a double-crossing of the postperovskite phase boundary by the geotherm at two different depths (Hernlund et al., 2005). Two deep slow velocity anomalies under West Pacific and Africa (roughly underneath the two maxima of the geoid) have unusual seismic properties. They have an anomalously large ratio of compressional to shear velocity ratio, v_p/v_s (Masters et al., 2000), and an anticorrelation between density and seismic velocities (Ishii and Tromp, 1999) and between bulk and shear velocities (Kennett and Gorbatov, 2004) (the bulk velocity is defined as $\sqrt{K_s/\rho} = \sqrt{v_p^2 - (4/3)v_s^2}$). These anomalous regions have very sharp boundaries (Ni et al., 2002) and depending on the authors have been named megaplumes, thermochemical piles, or LLSVP (large low-shear-velocity province, according to Garnero and McNamara (2008)). Notice, however, that as these LLSVPs only cover part of the CMB surface, which is itself four times smaller than the Earth's surface, and as they only extend up to a few hundred kilometers, their total volume is not so large but only of order three times larger that of the continental crustal volume (Hernlund and Houser, 2008).

These observations of the deep mantle heterogeneity cannot easily be explained by temperature variations. They seem to require lateral variations of Fe or Si contents in the mantle (Bolton and Masters, 2001; Saltzer et al., 2004), although some authors interpret these observations in the framework of pure thermal models (e.g., Davies et al., 2012). The LLSVP should be intrinsically denser to resist entrainment by convection. The stability of deep dense layers has been discussed by various authors (Davaille, 1999; Kellogg et al., 1999; LeBars and Davaille, 2002; Samuel and Farnetani, 2003; Tackley and Xie, 2002). These compositional pyramids may anchor the hot spots (Davaille et al., 2002; Jellinek and Manga, 2002). The presence of a petrologically dense component of the source of hot spots also seems necessary to explain their excess temperature (Farnetani, 1997). These abyssal heterogeneities help to bridge the gap between geochemical observations and convection modeling (Coltice and Ricard, 1999; Coltice et al., 2011; van Keken et al., 2002).

Deep mantle heterogeneities may be a consequence of plate tectonics or be relics of the early evolution of our planet. A large well-documented elemental differentiation is between the oceanic crust (poor in Mg and rich in Al and Si) and the mantle. The oceanic crust in its high pressure eclogitic facies is ≈5% denser that the average mantle density in most of the mantle except in the shallowest 100 km of the lower mantle where it is lighter (Irifune and Ringwood, 1993).

In the deepest mantle, it is not yet totally clear whether the eclogite remains denser, neutrally buoyant, or even slightly lighter than the average mantle (e.g., Ricolleau et al., 2004). Thermochemical simulations starting with the pioneering paper of Christensen and Hofmann (1994) show the possibility of a partial segregation of oceanic crust during subduction, forming pyramidal piles on the CMB. These results have been confirmed (e.g., Davies, 2002; Nakagawa and Tackley, 2006; Tackley, 2000b). Not only present-day subductions can generate compositional anomalies in the mantle. Geochemists have often argued for a deep layer of primitive material. A possible origin of the LLSVP could be related to the formation of the core in the Hadean time. The presence of patches of melt on top of the CMB (Williams and Garnero, 1996) suggests that in the past, a much thicker deep magma ocean may have existed in between a deep solid mantle and the liquid core (Labrosse et al., 2007). The fractionation during the freezing of this magma could have resulted in a progressive enrichment of the residue (the LLSVP) and of the melt (the ULVZ) in dense components. This model therefore proposes a scenario for the formation of petrologic anomalies in the deep mantle (Coltice et al., 2011). This subject of mantle dynamics near the CMB is covered in greater detail by Hernlund and McNamara, this volume.

7.02.6.7 A Complex Lithosphere: Plates and Continents

The lithosphere is part of the convective circulation and plate tectonics and mantle convection cannot be separated. The fact that the cold lithosphere is much more viscous and concentrates most of the mass heterogeneities of the mantle makes it behaving to some extent like a membrane on top of a less viscous fluid. This suggests some analogy between mantle convection and what is called Marangoni convection (Marangoni, 1840–1925). Marangoni convection (Nield, 1964) is controlled by temperature-dependent surface tension on top of thin layers of fluids.

The Earth's mantle is certainly not controlled by surface tension and Marangoni convection, strictly speaking, has nothing to do with mantle convection. However, the equations of thermal convection with internal heating and with a highly viscous lithosphere can be shown to be mathematically related (through a change of variables) to those of Marangoni convection (Lemery et al., 2000). There are large differences between mantle convection and surface-driven convection, for example, because the presence of subducting slabs throughout the mantle affects the flow differently than what surface tractions can do. However, this analogy has sometimes been advocated as a 'top-down' alternative to mantle convection (Anderson, 2001), which is really not that different from what has been known for 30 years using boundary layer theory. Classically, the interpretation of plate cooling in terms of ridge-push force (Turcotte and Schubert, 1982) or the analysis of tectonic stresses using thin sheet approximations (England and Mckenzie, 1982) belongs to the same top-down approach.

Due to the complexities of the lithosphere properties, the boundary condition at the surface of the Earth is far from being a uniform free-slip condition. Both continents and tectonic plates impose their own wavelengths and specific boundary conditions on the underlying convecting asthenosphere. Of course, the position of the continents and the number and shape of the plates are themselves consequences of mantle convection. The plates obviously organize the large-scale flow in the mantle (Hager and Oconnell, 1979; Ricard and Vigny, 1989). They impose a complex boundary condition where the angular velocity is piecewise-constant. The continents with their reduced heat flow (Jaupart and Mareschal, 1999) also impose a laterally variable heat flux boundary condition.

Convection models with continents have been studied numerically (Grigné and Labrosse, 2001; Gurnis and Hager, 1988) and experimentally (Guillou and Jaupart, 1995). Continents with their thick lithosphere tend to increase the thickness of the top boundary layer and the temperature below them (see **Figure 9**). Hot rising currents are predicted under continents and downwellings are localized along continental edges. The existence of a thick and stable continental root must be due to a chemically lighter and more viscous subcontinental lithosphere (Doin et al., 1997). The ratio of the heat flux extracted across continents compared to that extracted across oceans increases with the Rayleigh number. This suggests that the continental geotherms were not much different in the past when the radiogenic sources were larger; it is mostly the oceanic heat flux that was larger (Lenardic, 1998). Simulating organized plates self-consistently coupled with a convective mantle has been a very difficult quest. The attempts to generate plates using T-dependent or simple nonlinear rheologies have failed. Although in 2-D some successes can be obtained in localizing deformation in platelike domains (Schmeling and Jacoby, 1981; Weinstein, 1996; Weinstein and Olson, 1992), they are obtained with stress exponents (e.g., $n \geq 7$) that are larger than what can be expected from laboratory experiments ($n \approx 3$). The problems are however worst in 3-D. Generally, these early models do not predict the important shear motions between plates that we observe (Christensen and Harder, 1991; Ogawa et al., 1991).

Some authors have tried to mimic the presence of plates by imposing platelike surface boundary conditions. These studies have been performed in 2-D and 3-D (Gable et al., 1991; King et al., 1992; Monnereau and Quéré, 2001; Ricard and Vigny, 1989). Although they have confirmed the profound effect of plates on the wavelengths of convection, on its time dependence and on the surface heat flux, these approaches cannot predict the evolution of surface plate geometry. **Figure 10** illustrates the organizing effect of plates in spherical, internally heated compressible convection with depth-dependent viscosity (Bunge and Richards, 1996). To obtain a self-consistent generation of surface plates, more complex rheologies that include brittle failure, strain softening, and damage mechanisms must be introduced (e.g., Auth et al., 2003; Bercovici, 1993, 1995; Moresi and Solomatov, 1998). The existence of plates seems also to require the existence of a weak sublithospheric asthenosphere (Richards et al., 2001). In the last ten years, the first successes in computing 3-D models that spontaneously organize their top boundary layer into plates have been reached using viscoplastic rheologies (Foley and Becker, 2009; Stein et al., 2004; Tackley, 1998, 2000c,d,e; Trompert and Hansen, 1998a; van Heck and Tackley, 2008). This important breakthrough made by the modelers allows convection simulations that display increasingly realistic seafloor spreading and continental drift, with plates, continents, ridges and one-sided subductions (Coltice et al., 2012; Rolf and Tackley, 2011; Rolf et al., 2012) (see **Figure 11**).

The Earth's plate boundaries keep the memory of their weakness over geologic times (Gurnis et al., 2000). This implies that the rheological properties cannot be a simple time-independent function of stress or temperature but have a long-term memory. The rheologies that have been used to predict plates in convective models remain empirical and their interpretation in terms of microscopic behavior and damage theory remains largely to be done (Bercovici and Ricard, 2005,

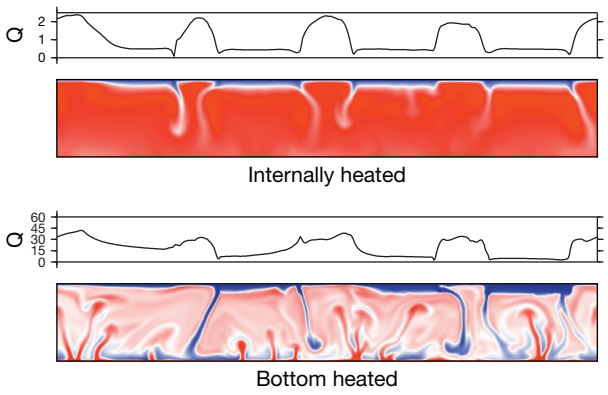

Figure 9 Convection patterns in the presence of four continents. The total aspect ratio is 7; the continents are defined by a viscosity increase by a factor 10^3 over the depth 1/10. The viscosity is otherwise constant. The Rayleigh number based on the total temperature drop (bottom panels) or on the internal radioactivity (top panels) is 10^7. The downwellings are localized near the continent margins. A large difference in heat flux is predicted between oceans and continents. In the case of bottom heating, hot spots tend to be preferentially anchored below continents where they bring an excess heat. This tends to reduce the surface heat flux variations.

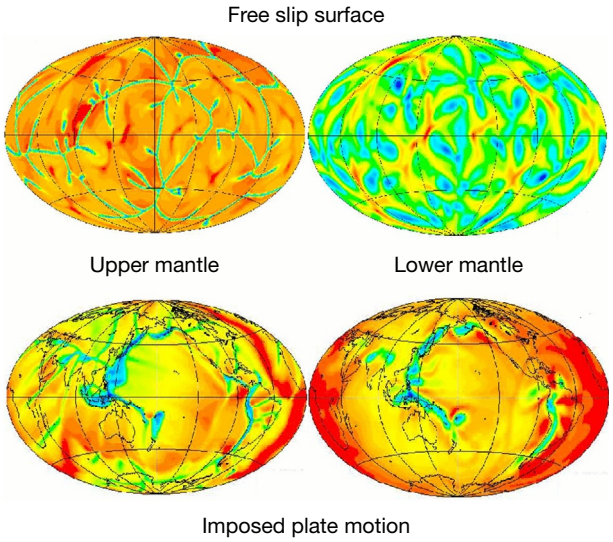

Figure 10 This figure depicts spherical compressible internally heated convection models where the viscosity increases with depth (simulations by Peter Bunge). In the first row, a uniform free-slip condition on top has been used. In the second row, the present-day observed plate motion is imposed at the surface. The left column shows the temperature field in the middle of the upper mantle and the right column in the middle of the lower mantle. This figure summarizes various points discussed in the text: the presence of linear cold downwellings, the absence of active upwellings in the absence of basal heating, and the enlargement of thermal structure in the more viscous lower mantle (top row). Although the modeling is not self-consistent (i.e., the presence of plates and the constancy of plate velocities are totally arbitrary), it is clear that the presence of plates can change radically the convection patterns (compare top and bottom rows).

2012, 2013, 2014). Reviews on the rapid progress and the limitations of self-coherent convection models can be found in Bercovici et al. (2000), Tackley (2000a), and Bercovici (2003) and are discussed in Bercovici et al., this volume.

7.02.6.8 Super-Earths

The study of the other planets of the solar system and their differences and similarities with the Earth is a very active research field described in the vol. 10 of the treatise. To the four terrestrial planets and the various dwarf planets and satellites of the solar system, many terrestrial extrasolar planets have been added in the last decade (e.g., Valencia et al., 2007b). A subset of these planets, approximately one to ten times as massive as Earth, have been termed super-Earths, because they are presumably analogous to our own planet. One can speculate that plate tectonics is a favorable condition for the existence of life. Plate tectonics buffers the climate by controlling the atmospheric carbon dioxide level with erosion, weathering, and subduction (Walker et al., 1981). Plate tectonics may also be necessary for the existence of life by providing a source of thermodynamic disequilibrium through continuous recycling of the surface (e.g., Southam and Westall, vol. 10). The question of whether super-Earths are more or less favorable to plate tectonics has therefore been debated.

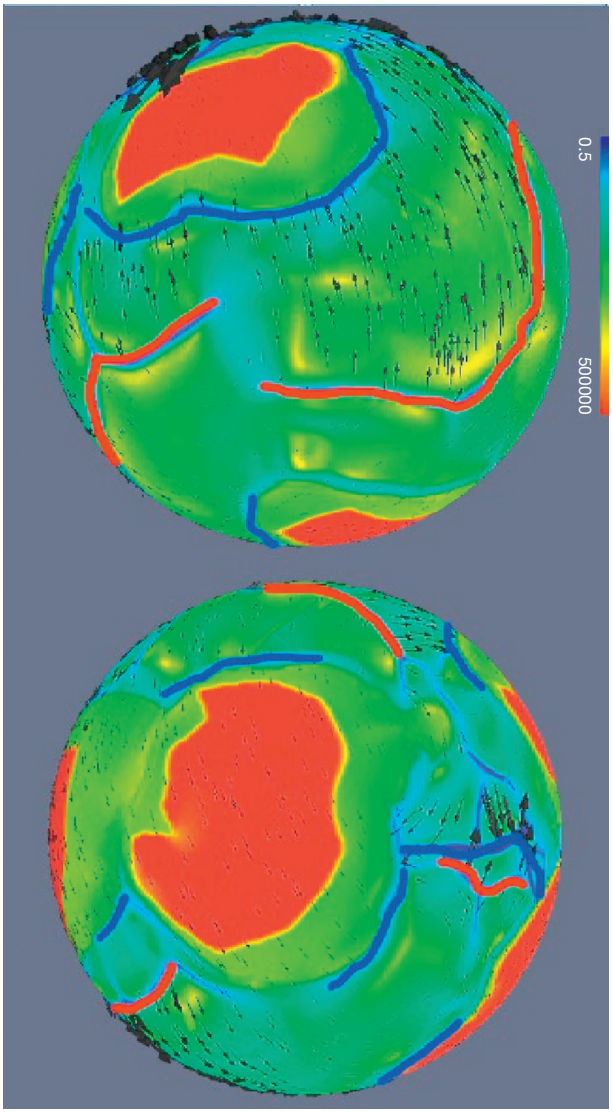

Figure 11 Viscosity and velocity fields at the surface of a 3-D spherical convection solution with platelike behavior and continental lithosphere. The red color is for the highest viscosity in the nucleus of continental lithosphere, surrounded by dark green weaker continental lithosphere. The light green is the viscosity of unstrained oceanic lithosphere, while blue represents low-viscosity regions where deformation localizes. The major divergent zones are highlighted by red curves, and the major convergent zones by blue curves. The surface velocity is close to that of rigid plates moving relative to each other. Courtesy of T. Rolf, N. Coltice and PJ. Tackley.

The topic is both controversial and somewhat academic. First, we do not have a decisive argument to explain why plate tectonics exists on Earth and not on other planets of the solar system (and particularly not on the Earth's sister planet, Venus). Second, the direct observations of the presence or absence of plate tectonics on super-Earths may also not happen in a close future. The debate on the propensity on plate tectonics on other planets illustrates our limits in the understanding of our own planet. The first papers on the subject had reached opposing conclusions about the likelihood of plate tectonics on super-Earths, inevitable for Valencia et al. (2007a) while less probable for O'Neill and Lenardic (2007).

The following studies, with all variables properly scaled, have concluded that plate tectonics is as favorable or more favorable on larger terrestrial planets than on Earth (Foley et al., 2012; Korenaga, 2010; Valencia and O'Connell, 2009; van Heck and Tackley, 2011). The stress level imposed by a more vigorous convection on larger planets (containing more radioactive elements and keeping its initial accretion heat a longer time) is higher. Combined with a thinner and weaker lithosphere, the internal stresses are more likely to induce the yielding of this lithosphere and to produce plate tectonic convection. However, other effects, namely, the surface temperature of the planet (Bercovici and Ricard, 2014; Foley et al., 2012) and the presence of water (Korenaga, 2010), have been shown equally or more important than the internal temperature for reaching plate tectonic convection. A hot temperature favors the healing of the lithosphere and makes plate tectonics less likely. Water lubricates the faults and makes plate tectonics more likely. This suggests that the climate of a planet controlling both the surface temperature and the water content may affect or even cause plate tectonics (Landuyt and Bercovici, 2009; Lenardic et al., 2008). The age of the planet since its accretion seems also an important parameter to consider (Lenardic and Crowley, 2012).

Acknowledgment

This chapter benefited from a very careful check of the equations by Frédéric Chambat and from the detailed and constructive reviews by Neil Ribe, Gerry Schubert, and Dave Bercovici. I also thank Ondrej Sramek, William Landuyt, Nicolas Coltice, Fabien Dubuffet, Thierry Alboussière, and others for their comments and suggestions.

References

Alboussiere T and Ricard Y (2013) Reflections on dissipation associated with thermal convection. *Journal of Fluid Mechanics* 725. http://dx.doi.org/10.1017/jfm.2013.241.

Alisic L, Gurnis M, Stadler G, Burstedde C, and Ghattas O (2012) Multi-scale dynamics and rheology of mantle flow with plates. *Journal of Geophysical Research* 117. http://dx.doi.org/10.1029/2012JB009234.

Anderson OL (1979) Evidence supporting the approximation $\gamma\rho = \text{const}$ for the Grüneisen parameter of the earth lower mantle. *Journal of Geophysical Research* 84: 3537–3542.

Anderson DL (2001) Geophysics – Top-down tectonics? *Science* 293: 2016–2018.

Austin NJ and Evans B (2007) Paleowattmeters: A scaling relation for dynamically recrystallized grain size. *Geology* 35(4): 343–346. http://dx.doi.org/10.1130/G23244A.1.

Auth C, Bercovici D, and Christensen UR (2003) Two-dimensional convection with a self-lubricating, simple-damage rheology. *Geophysical Journal International* 154: 783–800.

Badro J, Rueff JP, Vanko G, et al. (2004) Electronic transitions in perovskite: Possible nonconvecting layers in the lower mantle. *Science* 305: 383–386.

Bai WM, Vigny C, Ricard Y, and Froidevaux C (1992) On the origin of deviatoric stresses in the lithosphere. *Journal of Geophysical Research* 97: 11,729–11,737.

Balachandar S, Yuen DA, and Reuteler D (1992) Time-dependent 3-dimensional compressible convection with depth-dependent properties. *Geophysical Research Letters* 19: 2247–2250.

Balachandar S, Yuen DA, and Reuteler D (1993) Viscous and adiabatic heating effects in 3-dimensional compressible convection at infinite prandtl number. *Physics of Fluids A: Fluid Dynamics* 5: 2938–2945.

Barenblatt GI (1996) *Scaling, Self-similarity, and Intermediate Asymptotics*. Cambridge University Press.

Batchelor GK (1967) *An Introduction to Fluid Dynamics*. Cambridge University Press.

Bear J (1988) *Dynamics of Fluids in Porous Media*. New York: Dover publishers.

Beaumont C (1978) Evolution of sedimentary basins on a viscoelastic lithosphere – Theory and examples. *Geophysical Journal of the Royal Astronomical Society* 55: 471–497.

Becker TW and Boschi L (2002) A comparison of tomographic and geodynamic mantle models. *Geochemistry, Geophysics, Geosystems* 3: 1003. http://dx.doi.org/10.1029/2001GC000168.

Bercovici D (1993) A simple-model of plate generation from mantle flow. *Geophysical Journal International* 114: 635–650.

Bercovici D (1995) A source-sink model of the generation of plate-tectonics from non-Newtonian mantle flow. *Journal of Geophysical Research* 100: 2013–2030.

Bercovici D (2003) The generation of plate tectonics from mantle convection. *Earth and Planetary Science Letters* 205: 107–121.

Bercovici D and Karato S (2003) Whole-mantle convection and the transition-zone water filter. *Nature* 425: 39–44.

Bercovici D and Michaut C (2010) Two-phase dynamics of volcanic eruptions: Compaction, compression and the conditions for choking. *Geophysical Journal International* 182(2): 843–864. http://dx.doi.org/10.1111/j.1365-246X.2010.04674.x.

Bercovici D and Ricard Y (2003) Energetics of a two-phase model of lithospheric damage, shear localization and plate-boundary formation. *Geophysical Journal International* 152: 581–596.

Bercovici D and Ricard Y (2005) Tectonic plate generation and two-phase damage: Void growth versus grain size reduction. *Journal of Geophysical Research* 110.

Bercovici D and Ricard Y (2012) Mechanisms for the generation of plate tectonics by two-phase grain-damage and pinning. *Physics of the Earth and Planetary Interiors* 202–203: 27–55. http://dx.doi.org/10.1016/j.pepi.2012.05.003.

Bercovici D and Ricard Y (2013) Generation of plate tectonics with two-phase grain-damage and pinning: Source-sink model and toroidal flow. *Earth and Planetary Science Letters* 365: 275–288. http://dx.doi.org/10.1016/j.epsl.2013.02.002.

Bercovici D and Ricard Y (2014) Plate tectonics, damage and inheritance. *Nature* 508: 513–516.

Bercovici D, Ricard Y, and Richards MA (2000) The relation between mantle dynamics and plate tectonics: A primer. In: Richards MA, Gordon R, and VanderHilst R (eds.) *The History and Dynamics of Global Plate Motions*. *Geophysical Monograph Series*, vol. 21, pp. 5–46. Washington, DC: American Geophysical Union.

Bercovici D, Ricard Y, and Schubert G (2001a) A two-phase model of compaction and damage 1. General theory. *Journal of Geophysical Research* 106: 8887–8906.

Bercovici D, Ricard Y, and Schubert G (2001b) A two-phase model for compaction and damage 3. Applications to shear localization and plate boundary formation. *Journal of Geophysical Research* 106: 8925–8939.

Bercovici D, Schubert G, and Glatzmaier GA (1989a) 3-dimensional spherical models of convection in the earth's mantle. *Science* 244: 950–955.

Bercovici D, Schubert G, and Glatzmaier GA (1992) 3-dimensional convection of an infinite-Prandtl-number compressible fluid in a basally heated spherical-shell. *Journal of Fluid Mechanics* 239: 683–719.

Bercovici D, Schubert G, Glatzmaier GA, and Zebib A (1989b) 3-dimensional thermal-convection in a spherical-shell. *Journal of Fluid Mechanics* 206: 75–104.

Besse J and Courtillot V (1991) Revised and synthetic apparent polar wander paths of the African, Eurasian, North America and Indian plates, and true polar wander since 200 Ma. *Journal of Geophysical Research* 96: 4029–4050.

Birch F (1952) Elasticity and constitution of the Earth's interior. *Journal of Geophysical Research* 57: 227–286.

Bolton H and Masters G (2001) Travel times of P and S from the global digital seismic networks: Implications for the relative variation of P and S velocity in the mantle. *Journal of Geophysical Research* 106: 13,527–13,540.

Booker JR (1976) Thermal-convection with strongly temperature-dependent viscosity. *Journal of Fluid Mechanics* 76: 741–754.

Braun J, Chery J, Poliakov A, et al. (1999) A simple parameterization of strain localization in the ductile regime due to grain size reduction: A case study for olivine. *Journal of Geophysical Research* 104: 25,167–25,181.

Buffett BA, Garnero EJ, and Jeanloz R (2000) Sediments at the top of Earth's core. *Science* 290: 1338–1342.

Bullen KE (1940) The problem of the Earth's density variation. *Bulletin of the Seismological Society of America* 30: 235–250.

Bunge H, Hagelberg C, and Travis B (2003) Mantle circulation models with variational data assimilation: inferring past mantle flow and structure from plate motion histories and seismic tomography. *Geophysical Journal International* 152(2): 280–301. http://dx.doi.org/10.1046/j.1365-246X.2003.01823.x.

Bunge HP, Ricard Y, and Matas J (2001) Non-adiabaticity in mantle convection. *Geophysical Research Letters* 28: 879–882.

Bunge HP and Richards MA (1996) The origin of large scale structure in mantle convection: Effects of plate motions and viscosity stratification. *Geophysical Research Letters* 23: 2987–2990.

Bunge HP, Richards MA, and Baumgardner JR (1997) A sensitivity study of three-dimensional spherical mantle convection at 10(8) Rayleigh number: Effects of depth-dependent viscosity, heating mode, and an endothermic phase change. *Journal of Geophysical Research* 102: 11991–12007.

Cadek O and Fleitout L (2003) Effect of lateral viscosity variations in the top 300 km on the geoid and dynamic topography. *Geophysical Journal International* 152: 566–580.

Carman P (1939) Permeability of saturated sands, soils and clays. *Journal of Agricultural Science* 29(Part 2): 262–273.

Chambat F, Benzoni-Gavage S, and Ricard Y (2014) Jump conditions and dynamic surface tension at permeable interfaces such as the inner core boundary. *Comptes Rendus Geoscience* 346: 110–118.

Chandrasekhar S (1969) *Ellipsoidal Figures of Equilibrium*. Haven: Yale Press.

Choblet G, Cadek O, Couturier F, and Dumoulin C (2007) OEDIPUS: A new tool to study the dynamics of planetary interiors. *Geophysical Journal International* 170(1): 9–30. http://dx.doi.org/10.1111/j.1365-246X.2007.03419.x.

Choblet G and Parmentier EM (2001) Mantle upwelling and melting beneath slow spreading centers: effects of variable rheology and melt productivity. *Earth and Planetary Science Letters* 184: 589–604.

Chopelas A and Boehler R (1992) Thermal expansivity in the lower mantle. *Geophysical Research Letters* 19: 1983–1986.

Christensen U (1984a) Convection with pressure-dependent and temperature-dependent non-Newtonian rheology. *Geophysical Journal of the Royal Astronomical Society* 77: 343–384.

Christensen UR (1984b) Heat-transport by variable viscosity convection and implications for the Earth's thermal evolution. *Physics of the Earth and Planetary Interiors* 35: 264–282.

Christensen UR (1996) The influence of trench migration on slab penetration into the lower mantle. *Earth and Planetary Science Letters* 140: 27–39.

Christensen UR (1997a) Some geodynamical effects of anisotropic viscosity. *Geophysical Journal of the Royal Astronomical Society* 91: 711–736.

Christensen UR (1997b) Influence of chemical buoyancy on the dynamics of slabs in the transition zone. *Journal of Geophysical Research* 102: 22,435–22,443.

Christensen UR and Harder H (1991) 3-d convection with variable viscosity. *Geophysical Journal International* 104: 213–226.

Christensen UR and Hofmann A (1994) Segregation of subducted oceanic crust in the convecting mantle. *Journal of Geophysical Research* 99: 19,867–19,884.

Christensen UR and Yuen DA (1984) The interaction of a subducting lithospheric slab with a chemical or phase-boundary. *Journal of Geophysical Research* 89: 4389–4402.

Colin P and Fleitout L (1990) Topography of the ocean-floor – Thermal evolution of the lithosphere and interaction of deep mantle heterogeneities with the lithosphere. *Geophysical Research Letters* 17: 1961–1964.

Coltice N, Moreira M, Hernlund J, and Labrosse S (2011) Crystallization of a basal magma ocean recorded by Helium and Neon. *Earth and Planetary Science Letters* 308(1–2): 193199. http://dx.doi.org/10.1016/j.epsl.2011.05.045.

Coltice N and Ricard Y (1999) Geochemical observations and one layer mantle convection. *Earth and Planetary Science Letters* 174: 125–137.

Coltice N, Rolf T, Tackley PJ, and Labrosse S (2012) Dynamic causes of the relation between area and age of the ocean floor. *Science* 336(6079): 335–338. http://dx.doi.org/10.1126/science.1219120.

Coltice N and Schmalzl J (2006) Mixing times in the mantle of the early Earth derived from 2-D and 3-D numerical simulations of convection. *Geophysical Research Letters* 33(23): L23304. http://dx.doi.org/10.1029/2006GL027707.

Connolly JAD and Podladchikov YY (1998) Compaction-driven fluid flow in viscoelastic rock. *Geodinamica Acta* 11: 55–84.

Conrad C and Gurnis M (2003) Seismic tomography, surface uplift, and the breakup of Gondwanaland: Integrating mantle convection backwards in time. *Geochemistry, Geophysics, Geosystems* 4. http://dx.doi.org/10.1029/2001GC000299.

Conrad CP and Husson L (2009) Influence of dynamic topography on sea level and its rate of change. *Lithosphere* 1(2): 110–120. http://dx.doi.org/10.1130/L32.1.

Corrieu V, Thoraval C, and Ricard Y (1995) Mantle dynamics and geoid Green functions. *Geophysical Journal International* 120: 516–523.

Crameri F, Tackley PJ, Meilick I, Gerya TV, and Kaus BJP (2012) A free plate surface and weak oceanic crust produce single-sided subduction on Earth. *Geophysical Research Letters* 39. http://dx.doi.org/10.1029/2011GL050046.

Cserepes L (1993) Effect of depth-dependent viscosity on the pattern of mantle convection. *Geophysical Research Letters* 20: 2091–2094.

Dahlen FA and Tromp J (1998) *Theoretical Global Seismology*. Princeton, NJ: Princeton University Press.

Davaille A (1999) Two-layer thermal convection in viscous fluids. *Journal of Fluid Mechanics* 379: 223–253.

Davaille A, Girard F, and LeBars M (2002) How to anchor hotspots in a convecting mantle? *Earth and Planetary Science Letters* 203: 621–634.

Davaille A and Jaupart C (1993) Transient high-rayleigh-number thermal-convection with large viscosity variations. *Journal of Fluid Mechanics* 253: 141–166.

Davies GF (1980) Thermal histories of convective earth models and constraints on radiogenic heat-production in the earth. *Journal of Geophysical Research* 85: 2517–2530.

Davies GF (1999) *Dynamic Earth Plates, Plumes and Mantle Convection*, p 458. Cambridge, UK: Cambridge University Press.

Davies GF (2002) Stirring geochemistry in mantle convection models with stiff plates and slabs. *Geochimica et Cosmochimica Acta* 66: 3125–3142.

Davies DR, Goes S, Davies JH, et al. (2012) Reconciling dynamic and seismic models of Earth's lower mantle: The dominant role of thermal heterogeneity. *Earth and Planetary Science Letters* 353: 253–269. http://dx.doi.org/10.1016/j.epsl.2012.08.016.

De Bresser J, Ter Heege J, and Spiers C (2001) Grain size reduction by dynamic recrystallization: can it result in major theological weakening? *International Journal of Earth Sciences* 90(1): 28–45.

de Groot SR and Mazur P (1984) *Non-Equilibrium Thermodynamics*. New York: Dover Publications.

Doin MP, Fleitout L, and Christensen U (1997) Mantle convection and stability of depleted and undepleted continental lithosphere. *Journal of Geophysical Research* 102: 2771–2787.

Dombre T, Frisch U, Greene JM, et al. (1986) Chaotic streamlines in the abc flows. *Journal of Fluid Mechanics* 167: 353–391.

Drew DA and Segel LA (1971) Averaged equations for 2-phase flows. *Studies in Applied Mathematics* 50: 205–220.

Drury MR and FitzGerald JD (1998) *The Earth's Mantle: Composition, Structure, and Evolution*. *Mantle Rheology: Insights from Laboratory Studies of Deformation and Phase Transition*. Cambridge, UK: Cambridge University Press.

Dubuffet F, Rabinowicz M, and Monnereau M (2000) Multiple scales in mantle convection. *Earth and Planetary Science Letters* 178: 351–366.

Dubuffet F, Yuen DA, and Rabinowicz M (1999) Effects of a realistic mantle thermal conductivity on the patterns of 3-D convection. *Earth and Planetary Science Letters* 171: 401–409.

Dziewonski AM and Anderson D (1981) Preliminary reference earth model. *Physics of the Earth and Planetary Interiors* 25: 297–356.

England P and Mckenzie D (1982) A thin viscous sheet model for continental deformation. *Geophysical Journal of the Royal Astronomical Society* 70: 295–321.

Farnetani CG (1997) Excess temperature of mantle plumes: The role of chemical stratification across D″ *Geophysical Research Letters* 24: 1583–1586.

Ferrachat S and Ricard Y (1998) Regular vs, chaotic mantle mixing. *Earth and Planetary Science Letters* 155: 75–86.

Flament N, Gurnis M, and Müller RD (2013) A review of observations and models of dynamic topography. *Lithosphere* http://dx.doi.org/10.1130/L245.1.

Foley BJ and Becker TW (2009) Generation of plate-like behavior and mantle heterogeneity from a spherical, viscoplastic convection model. *Geochemistry, Geophysics, Geosystems* 10. http://dx.doi.org/10.1029/2009GC002378.

Foley BJ, Bercovici D, and Landuyt W (2012) The conditions for plate tectonics on super-Earths: Inferences from convection models with damage. *Earth and Planetary Science Letters* 331: 281–290. http://dx.doi.org/10.1016/j.epsl.2012.03.028.

Forte AM and Peltier WR (1987) Plate tectonics and aspherical earth structure – The importance of poloidal-toroidal coupling. *Journal of Geophysical Research* 92: 3645–3679.

Forte AM and Peltier WR (1991) Viscous flow models of global geophysical observables.1. forward problems. *Journal of Geophysical Research* 96: 20,131–20,159.

Fowler AC (1985) A mathematical model of magma transport in the asthenosphere. *Geophysical and Astrophysical Fluid Dynamics* 33: 63–96.

Fowler AC (1989) Generation and creep of magma in the earth. *SIAM Journal on Applied Mathematics* 49: 231–245.

Gable CW, Oconnell RJ, and Travis BJ (1991) Convection in 3 dimensions with surface plates – generation of toroidal flow. *Journal of Geophysical Research* 96: 8391–8405.

Garnero EJ and McNamara AK (2008) Structure and dynamics of Earth's lower mantle. *Science* 320(5876): 626–628. http://dx.doi.org/10.1126/science.1148028.

Gasperini P, Forno GD, and Boschi E (2004) Linear or non-linear rheology in the Earth's mantle: The prevalence of power-law creep in the postglacial isostatic readjustment of laurentia. *Geophysical Journal International* 157: 1297–1302.

Glatzmaier GA (1988) Numerical simulation of mantle convection: Time-dependent, three dimensional, compressible, spherical shell. *Geophysical and Astrophysical Fluid Dynamics* 43: 223–264.

Golabek GJ, Schmeling H, and Tackley PJ (2008) Earth's core formation aided by flow channelling instabilities induced by iron diapirs. *Earth and Planetary Science Letters* 271(1–4): 24–33. http://dx.doi.org/10.1016/j.epsl.2008.02.033.

Grigné C and Labrosse S (2001) Effects of continents on Earth cooling: Thermal blanketing and depletion in radioactive elements. *Geophysical Research Letters* 28: 2707–2710.

Grigne C, Labrosse S, and Tackley PJ (2005) Convective heat transfer as a function of wavelength: Implications for the cooling of the earth. *Journal of Geophysical Research, Solid Earth* 110.

Guillou L and Jaupart C (1995) On the effect of continents on mantle convection. *Journal of Geophysical Research* 100: 24,217–24,238.

Guilloufrottier L, Buttles J, and Olson P (1995) Laboratory experiments on the structure of subducted lithosphere. *Earth and Planetary Science Letters* 133: 19–34.

Gurnis M and Davies GF (1986) Numerical study of high Rayleigh number convection in a medium with depth-dependent viscosity. *Geophysical Journal of the Royal Astronomical Society* 85: 523–541.

Gurnis M and Hager BH (1988) Controls of the structure of subducted slabs. *Nature* 335: 317–321.

Gurnis M, Moresi L, and Mueller RD (2000) The history and dynamics of global plate motions. In: Richards MA, Gordon R, and vanderHilst R (eds.) *Models of mantle Convection Incorporating Plate Tectonics: The Australian Region since the Cretaceous*, pp. 211–238. Washington, DC: American Geophysical Union.

Haase R (1990) *Thermodynamics of Irreversible Processes*. New York: Dover Publications.

Hager BH and Clayton RW (1989) Constraints on the structure of mantle convection using seismic observation, flow models and the geoid. In: Peltier WR (ed.) *Mantle Convection, Plate Tectonics and Global Dynamics*. New York: Gordon and Breach Science Publishers.

Hager BH, Clayton RW, Richards MA, Comer RP, and Dziewonski AM (1985) Lower mantle heterogeneity, dynamic topography and the geoid. *Nature* 313: 541–546.

Hager BH and O'Connell RJ (1981) A simple global model of plate tectonics and mantle convection. *Journal of Geophysical Research* 86: 4843–4867.

Hager BH and Oconnell RJ (1979) Kinematic models of large-scale flow in the Earth's mantle. *Journal of Geophysical Research* 84: 1031–1048.

Hansen U, Yuen DA, Kroening SE, and Larsen TB (1993) Dynamic consequences of depth-dependent thermal expansivity and viscosity on mantle circulations and thermal structure. *Physics of the Earth and Planetary Interiors* 77: 205–223.

Haskell NA (1937) The viscosity of the asthenosphere. *American Journal of Science* 33: 22–28.

Hernlund JW and Houser C (2008) The statistical distribution of seismic velocities in Earth's deep mantle. *Earth and Planetary Science Letters* 265(3–4): 423–437. http://dx.doi.org/10.1016/j.epsl.2007.10.042.

Hernlund JW and Jellinek AM (2010) Dynamics and structure of a stirred partially molten ultralow-velocity zone. *Earth and Planetary Science Letters* 296(1–2): 1–8. http://dx.doi.org/10.1016/j.epsl.2010.04.027.

Hernlund J, Thomas C, and Tackley P (2005) A doubling of the post-perovskite phase boundary and structure of the Earth's lowermost mantle. *Nature* 434(7035): 882–886. http://dx.doi.org/10.1038/nature03472.

Hewitt IJ (2010) Modelling melting rates in upwelling mantle. *Earth and Planetary Science Letters* 300(3–4): 264–274. http://dx.doi.org/10.1016/j.epsl.2010.10.010.

Hewitt IJ and Fowler AC (2008) Partial melting in an upwelling mantle column. *Proceedings of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences* 464(2097): 2467–2491. http://dx.doi.org/10.1098/rspa.2008.0045.

Hewitt IJ and Fowler AC (2009) Melt channelization in ascending mantle. *Journal of Geophysical Research* 114. http://dx.doi.org/10.1029/2008JB006185.

Hewitt JM, McKenzie DP, and Weiss NO (1975) Dissipative heating in convective flows. *Journal of Fluid Mechanics* 68: 721–738.

Hier-Majumder S, Ricard Y, and Bercovici D (2006) Role of grain boundaries in magma migration and storage. *Earth and Planetary Science Letters* 248: 735–749.

Hillert M (1965) On theory of normal and abnormal grain growth. *Acta Metallurgica* 13: 227–235.

Hiraga T, Tachibana C, Ohashi N, and Sano S (2010) Grain growth systematics for forsterite +/- enstatite aggregates: Effect of lithology on grain size in the upper mantle. *Earth and Planetary Science Letters* 291(1–4): 10–20. http://dx.doi.org/10.1016/j.epsl.2009.12.026.

Hirth G and Kolhstedt DL (1996) Water in the oceanic upper mantle: implications for rheology, melt extraction and the evolution of the lithosphere. *Earth and Planetary Science Letters* 144: 93–108.

Hofmann A and Hart S (1978) Assessment of local and regional isotopic equilibrium in mantle. *Earth and Planetary Science Letters* 38(1): 44–62. http://dx.doi.org/10.1016/0012-821X(78)90125-5.

Hofmeister AM (1999) Mantle values of thermal conductivity and the geotherm from phonon lifetimes. *Science* 283: 1699–1706.

Holtzman B, Kohlstedt D, Zimmerman M, et al. (2003) Melt segregation and strain partitioning: Implications for seismic anisotropy and mantle flow. *Science* 301(5637): 1227–1230. http://dx.doi.org/10.1126/science.1087132.

Honda S, Yuen DA, Balachandar S, and Reuteler D (1993) 3-dimensional instabilities of mantle convection with multiple phase-transitions. *Science* 259: 1308–1311.

Husson L and Conrad CP (2006) Tectonic velocities, dynamic topography, and relative sea level. *Geophysical Research Letters* 33(18). http://dx.doi.org/10.1029/2006GL026834.

Irifune T and Ringwood AE (1987) Phase transformations in a harzburgite composition to 26 GPa: implications for dynamical behaviour of the subducting slab. *Earth and Planetary Science Letters* 86: 365–376.

Irifune T and Ringwood A (1993) Phase transformations in subducted oceanic crust and buoyancy relationships at depths of 600–800 km in the mantle. *Earth and Planetary Science Letters* 117: 101–110.

Ishii M and Tromp J (1999) Normal-mode and free-air gravity constraints on lateral variations in velocity and density of Earth's mantle. *Science* 285(5431): 1231–1236. http://dx.doi.org/10.1126/science.285.5431.1231.

Ismail-Zadeh A, Schubert G, Tsepelev I, and Korotkii A (2004) Inverse problem of thermal convection: numerical approach and application to mantle plume restoration. *Physics of the Earth and Planetary Interiors* 145(1–4): 99–114. http://dx.doi.org/10.1016/j.pepi.2004.03.006.

Ito G, Shen Y, Hirth G, and Wolfe CJ (1999) Mantle flow, melting, and dehydration of the Iceland mantle plume. *Earth and Planetary Science Letters* 165: 81–96.

Jaroslow GE, Hirth G, and Dick HJB (1996) Abyssal peridotite mylonites: Implications for grain-size sensitive flow and strain localization in the oceanic lithosphere. *Tectonophysics* 256: 17–37.

Jarvis GT and Mckenzie DP (1980) Convection in a compressible fluid with infinite prandtl number. *Journal of Fluid Mechanics* 96: 515–583.

Jaupart C and Mareschal JC (1999) The thermal structure and thickness of continental roots. *Lithos* 48: 93–114.

Jeanloz R and Morris S (1987) Is the mantle geotherm subadiabatic. *Geophysical Research Letters* 14: 335–338.

Jellinek AM and Manga M (2002) The influence of a chemical boundary layer on the fixity, spacing and lifetime of mantle plumes. *Nature* 418: 760–763.

Joseph D (1966) Nonlinear stability of boussinesq equations by method of energy. *Archive for Rational Mechanics and Analysis* 22(3): 163–184.

Joseph DD and Renardy Y (1993) *Fundamentals of Two-Fluid Dynamics. Part 1: Mathematical Theory and Applications*. New York: Springer-Verlag.

Kameyama M and Ogawa M (2000) Transitions in thermal convection with strongly temperature-dependent viscosity in a wide box. *Earth and Planetary Science Letters* 180: 355–367.

Kameyama M, Yuen DA, and Fujimoto H (1997) The interaction of viscous heating with grain-size dependent rheology in the formation of localized slip zones. *Geophysical Research Letters* 24: 2523–2526.

Karato SI (1998) Seismic anisotropy in the deep mantle, boundary layers and the geometry of mantle convection. *Pure and Applied Geophysics* 151: 565–587.

Karato S (2008) *Deformation of Earth Materials: An Introduction to the Rheology of Solid Earth*. Cambridge, UK: Cambridge University Press.

Karato S, Bercovici D, Leahy G, Richard G, and Jing Z (2006) The transition-zone water filter model for global material circulation: Where do we stand? In: Jacobsen SD and van der Lee S (eds.) *Earth's Deep Water Cycle*. *Geophysical Monograph Series*, vol. 168, pp. 289–313. Washington, DC: AGU.

Karato SI, Paterson MS, and Gerald JDF (1986) Rheology of synthetic olivine aggregates – Influence of grain-size and water. *Journal of Geophysical Research* 91: 8151–8176.

Karato SI and Wu P (1993) Rheology of the upper mantle: A synthesis. *Science* 260: 771–778.

Katz RF, Spiegelman M, and Carbotte SM (2004) Ridge migration, asthenospheric flow and the origin of magmatic segmentation in the global mid-ocean ridge system. *Geophysical Research Letters* 31.

Kaus BJP, Muehlhaus H, and May DA (2010) A stabilization algorithm for geodynamic numerical simulations with a free surface. *Physics of the Earth and Planetary Interiors* 181(1–2): 12–20. http://dx.doi.org/10.1016/j.pepi.2010.04.007.

Kelemen PB, Hirth G, Shimizu N, Spiegelman M, and Dick HJB (1997) A review of melt migration processes in the adiabatically upwelling mantle beneath oceanic spreading ridges. *Philosophical Transactions of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences* 355: 283–318.

Kellogg LH, Hager BH, and van der Hilst RD (1999) Compositional stratification in the deep mantle. *Science* 283: 1881–1884.

Kellogg LH and Turcotte DL (1987) Homogenization of the mantle by convective mixing and diffusion. *Earth and Planetary Science Letters* 81: 371–378.

Kennett BLN (2001) *The Seismic Wavefield*. Cambridge, UK: Cambridge University Press.

Kennett B and Gorbatov A (2004) Seismic heterogeneity in the mantle – Strong shear wave signature of slabs from joint tomography. *Physics of the Earth and Planetary Interiors* 146(1–2): 87100. http://dx.doi.org/10.1016/j.pepi.2003.07.033, Superplume International Workshop, Tokyo Inst. Tech, Tokyo, Japan, Jan 28–31, 2002.

Kido M and Seno T (1994) Dynamic topography compared with residual depth anomalies in oceans and implications for age-depth curves. *Geophysical Research Letters* 21: 717–720.

Kincaid C and Olson P (1987) An experimental-study of subduction and slab migration. *Journal of Geophysical Research* 92: 13,832–13,840.

King SD, Gable CW, and Weinstein SA (1992) Models of convection-driven tectonic plates – A comparison of methods and results. *Geophysical Journal International* 109: 481–487.

King SD and Hager BH (1994) Subducted slabs and the geoid.1. Numerical experiments with temperature-dependent viscosity. *Journal of Geophysical Research* 99: 19,843–19,852.

Korenaga J (2003) Energetics of mantle convection and the fate of fossil heat. *Geophysical Research Letters* 30.

Korenaga J (2010) On the likelihood of plate tectonics on super Earth's: does size matter? *Astrophysical Journal Letters* 725(1): L43–L46. http://dx.doi.org/10.1088/2041-8205/725/1/L43.

Labrosse S, Hernlund JW, and Coltice N (2007) A crystallizing dense magma ocean at the base of the Earth's mantle. *Nature* 450(7171): 866–869. http://dx.doi.org/10.1038/nature06355.

Lago B and Rabinowicz M (1984) Admittance for convection in a layered spherical shell. *Geophysical Journal of the Royal Astronomical Society* 77(2): 461–482.

Lambeck K and Johnston P (1998) *The Earth's mantle: composition, structure, and evolution*. *The viscosity of the mantle*. Cambridge, UK: Cambridge University Press, pp. 461–502.

Landau L and Lifchitz E (1980) *An Introduction to Fluid Dynamics*. Oxford: Pergamon Press.

Landau L and Lifchitz E (2000) *Theory of elasticity*. London: Butterworth-Heinemann.

Landuyt W and Bercovici D (2009) Variations in planetary convection via the effect of climate on damage. *Earth and Planetary Science Letters* 277(1–2): 29–37. http://dx.doi.org/10.1016/j.epsl.2008.09.034.

Larsen TB, Malevsky AV, Yuen DA, and Smedsmo JL (1993) Temperature-dependent Newtonian and non-Newtonian convection – Implications for lithospheric processes. *Geophysical Research Letters* 20: 2595–2598.

Larsen TB and Yuen DA (1997) Fast plumeheads: Temperature-dependent versus non-Newtonian rheology. *Geophysical Research Letters* 24: 1995–1998.

Lasaga A (1998) *Kinetics Theory in the Earth Sciences*. Princeton, NJ: Princeton University Press.

LeBars M and Davaille A (2002) Stability of thermal convection in two superimposed miscible viscous fluids. *Journal of Fluid Mechanics* 471: 339–363.

Lee T, Papanastassiou DA, and Wasserburg GL (1976) Demonstration of ^{26} mg excess in allende and evidence for al^{26}. *Geophysical Research Letters* 3: 41–44.

Lemery C, Ricard Y, and Sommeria J (2000) A model for the emergence of thermal plumes in Rayleigh–Bénard convection at infinite Prandtl number. *Journal of Fluid Mechanics* 414: 225–250.

Lenardic A (1998) On the partitioning of mantle heat loss below oceans and continents over time and its relationship to the Archaean paradox. *Geophysical Journal International* 134: 706–720.

Lenardic A and Crowley JW (2012) On the notion of well-defined tectonic regimes for terrestrial planets in this solar system and others. *Astrophysical Journal* 755(2): 132. http://dx.doi.org/10.1088/0004-637X/755/2/132.

Lenardic A, Jellinek AM, and Moresi LN (2008) A climate change induced transition in the tectonic style of a terrestrial planet. *Earth and Planetary Science Letters* 271: 33–42.

Lestunff Y and Ricard Y (1995) Topography and geoid due to lithospheric mass anomalies. *Geophysical Journal International* 122: 982–990.

Lister JR and Kerr RC (1991) Fluid-mechanical models of crack-propagation and their application to magma transport in dykes. *Journal of Geophysical Research, Solid Earth* 96: 10,049–10,077.

Lithgow-Bertelloni C and Gurnis M (1997) Cenozoic subsidence and uplift of continents from time-varying dynamic topography. *Geology* 25: 735–738.

Lithgow-Bertelloni C and Guynn JH (2004) Origin of the lithospheric stress field. *Journal of Geophysical Research* 109. http://dx.doi.org/10.1029/2003JB002467.

Lithgow-Bertelloni C and Richards MA (1998) Dynamics of cenozoic and mesozoic plate motions. *Reviews of Geophysics* 36: 27–78.

Lithgow-Bertelloni C and Silver PG (1998) Dynamic topography, plate driving forces and the African superswell. *Nature* 395: 269–272.

Liu L and Gurnis M (2008) Simultaneous inversion of mantle properties and initial conditions using an adjoint of mantle convection. *Journal of Geophysical Research* 113(B8). http://dx.doi.org/10.1029/2008JB005594.

Lyubetskaya T and Korenaga J (2007) Chemical composition of Earth's primitive mantle and its variance, 1, method and results. *Journal of Geophysical Research* 112(B3): B03212.

Machetel P and Weber P (1991) Intermittent layered convection in a model mantle with an endothermic phase change at 670 km. *Nature* 350: 55–57.

Malevsky AV and Yuen DA (1992) Strongly chaotic non-Newtonian mantle convection. *Geophysical and Astrophysical Fluid Dynamics* 65: 149–171.

Malvern L (1969) *Introduction to the Mechanics of a Continuum Medium*. *Prentice-Hall Series in Engineering of the Physical Sciences*. Englewood, NJ: Prentice-Hall.

Massol H, Jaupart C, and Pepper DW (2001) Ascent and decompression of viscous vesicular magma in a volcanic conduit. *Journal of Geophysical Research* 106: 16,223–16,240.

Masters G, Laske G, Bolton H, and Dziewonski A (2000) The relative behavior of shear velocity, bulk sound speed, and compressional velocity in the mantle: Implications for chemical and thermal structure. In: Karato S, Forte AM, Liebermann RC, Masters G, and Stixrude L (eds.) *Earth's Deep Interior*. *Geophysical Monograph Series*, vol. 117. Washington, DC: AGU.

McDonough WF and Sun SS (1995) The composition of the earth. *Chemical Geology* 120: 223–253.

McKenzie D (1984) The generation and compaction of partially molten rock. *Journal of Petrology* 25: 713–765.

McNamara AK, Karato SI, and van Keken PE (2001) Localisation of dislocation creep in the lower mantle: Implications for the origin of seismic anisotropy. *Earth and Planetary Science Letters* 191: 85–99.

Michaut C and Bercovici D (2009) A model for the spreading and compaction of two-phase viscous gravity currents. *Journal of Fluid Mechanics* 630: 299–329. http://dx.doi.org/10.1017/S0022112009006612.

Michaut C, Bercovici D, and Sparks RSJ (2009) Ascent and compaction of gas rich magma and the effects of hysteretic permeability. *Earth and Planetary Science Letters* 282(1–4): 258–267. http://dx.doi.org/10.1016/j.epsl.2009.03.026.

Michaut C, Ricard Y, Bercovici D, and Sparks RSJ (2013) Eruption cyclicity at silicic volcanoes potentially caused by magmatic gas waves. *Nature Geoscience* 6(10): 856–860. http://dx.doi.org/10.1038/NGEO1928.

Mitrovica JX and Forte AM (1997) Radial profile of mantle viscosity: Results from the joint inversion of convection and postglacial rebound observables. *Journal of Geophysical Research* 102: 2751–2769.

Monnereau M and Quéré S (2001) Spherical shell models of mantle convection with tectonic plates. *Earth and Planetary Science Letters* 184: 575–588.

Moresi L and Solomatov V (1998) Mantle convection with a brittle lithosphere: Thoughts on the global tectonic styles of the Earth and Venus. *Geophysical Journal International* 133: 669–682.

Moucha R, Forte AM, Mitrovica JX, et al. (2008) Dynamic topography and long-term sea-level variations: There is no such thing as a stable continental platform. *Earth and Planetary Science Letters* 271(1–4): 101108. http://dx.doi.org/10.1016/j.epsl.2008.03.056.

Muhlhaus HB, Moresi L, and Cada M (2004) Emergent anisotropy and flow alignment in viscous rocks. *Pure and Applied Geophysics* 161: 2451–2463.

Munk WH and MacDonald GJF (1960) *The Rotation of the Earth: A Geophysical Discussion*. New York: Cambridge University Press.

Murakami T, Hirose K, Kawamura K, Sata N, and Ohishi Y (2004) Post-perovskite phase transition in $MgSiO_3$. *Science* 304: 855–858.

Murnaghan FD (1951) *Finite Deformation of an Elastic Solid*. New York: Wiley.

Nakagawa T and Tackley PJ (2006) Three-dimensional structures and dynamics in the deep mantle: Effects of post-perovskite phase change and deep mantle layering. *Geophysical Research Letters* 33: L12S11. http://dx.doi.org/1029/2006GL025719.

Nataf HC and Richter FM (1982) Convection experiments in fluids with highly temperature-dependent viscosity and the thermal evolution of the planets. *Physics of the Earth and Planetary Interiors* 29: 320–329.

Ni S, Tan E, Gurnis M, and Helmberger D (2002) Sharp sides to the African superplume. *Science* 296(5574): 1850–1852. http://dx.doi.org/10.1126/science.1070698.

Nield DA (1964) Surface tension and buoyancy effects in cellular convection. *Journal of Fluid Mechanics* 19: 341–352.

Nye JF (1953) The flow law of ice from measurements in glacier tunnels, laboratory experiments and the Jungfraufirn borehole experiment. *Proceedings of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences* 219: 477–489.

O'Neill C and Lenardic A (2007) Geological consequences of super-sized Earths. *Geophysical Research Letters* 34(19). http://dx.doi.org/10.1029/2007GL030598.

Ogawa M, Schubert G, and Zebib A (1991) Numerical simulations of 3-dimensional thermal-convection in a fluid with strongly temperature-dependent viscosity. *Journal of Fluid Mechanics* 233: 299–328.

Ogura Y and Phillips N (1962) Scale analysis of deep and shallow convection in the atmosphere. *Journal of Atmospheric Science* 19(2): 173–179. http://dx.doi.org/10.1175/1520-0469.

Olson P, Yuen DA, and Balsiger D (1984) Mixing of passive heterogeneities by mantle convection. *Journal of Geophysical Research* 89: 425–436.

Ottino JM (1989) *The Kinematics of Mixing: Stretching, Chaos and Transport*. New York: Cambridge University Press.

Panasyuk S, Hager B, and Forte A (1996) Understanding the effects of mantle compressibility on geoid kernels. *Geophysical Journal International* 124: 121–133.

Parmentier EM, Sotin C, and Travis BJ (1994) Turbulent 3-d thermal-convection in an infinite Prandtl number, volumetrically heated fluid – Implications for mantle dynamics. *Geophysical Journal International* 116: 241–251.

Parmentier EM, Turcotte DL, and Torrance KE (1976) Studies of finite-amplitude non-Newtonian thermal convection with application to convection in Earth's mantle. *Journal of Geophysical Research* 81: 1839–1846.

Parsons B and Daly S (1983) The relationship between surface topography, gravity anomaly and the temperature structure of convection. *Journal of Geophysical Research* 88: 1129–1144.

Peltier WR (1989) Mantle convection: Plate tectonics and Global geodynamics. In: *Mantle Viscosity*, pp. 389–478. New York: Gordon and Breach Science Publishers.

Peltier WR (1996) Mantle viscosity and ice-age ice sheet topography. *Science* 273: 1359–1364.

Peltier WR and Solheim LP (1992) Mantle phase transition and layered chaotic convection. *Geophysical Research Letters* 19: 321–324.

Poirier JP (1991) *Introduction to the Physics of the Earth's Interior*. Cambridge: Cambridge University Press.

Poirier JP and Tarantola A (1998) A logarithmic equation of state. *Physics of the Earth and Planetary Interiors* 109: 1–8.

Pouilloux L, Kaminski E, and Labrosse S (2007) Anisotropic rheology of a cubic medium and implications for geological materials. *Geophysical Journal International* 170(2): 876–885. http://dx.doi.org/10.1111/j.1365-246X.2007.03461.x.

Rabinowicz M, Ricard Y, and Gregoire M (2002) Compaction in a mantle with a very small melt concentration: Implications for the generation of carbonatitic and carbonate-bearing high alkaline mafic melt impregnations. *Earth and Planetary Science Letters* 203: 205–220.
Ranalli G (1995) *Rheology of the Earth*. Dordrecht: Kluwer academic publishers.
Ribe NM (1985a) The deformation and compaction of partial molten zones. *Geophysical Journal of the Royal Astronomical Society* 83: 487–501.
Ribe NM (1985b) The generation and composition of partial melts in the Earth's mantle. *Earth and Planetary Science Letters* 73: 361–376.
Ribe NM, Stutzmann E, Ren Y, and van der Hilst R (2007) Buckling instabilities of subducted lithosphere beneath the transition zone. *Earth and Planetary Science Letters* 254(1–2): 173–179. http://dx.doi.org/10.1016/j.epsl.2006.11.028.
Ricard Y and Bercovici D (2003) Two-phase damage theory and crustal rock failure: The theoretical 'void' limit, and the prediction of experimental data. *Geophysical Journal International* 155: 1057–1064.
Ricard Y and Bercovici D (2009) A continuum theory of grain size evolution and damage. *Journal of Geophysical Research* 114(B01204). http://dx.doi.org/10.1029/2007JB005491.
Ricard Y, Bercovici D, and Schubert G (2001) A two-phase model for compaction and damage 2. Applications to compaction, deformation, and the role of interfacial surface tension. *Journal of Geophysical Research* 106: 8907–8924.
Ricard Y and Coltice N (2004) Geophysical and geochemical models of mantle convection: Successes and future challenges. In: *The state of the planet: Frontiers and challenges in geophysics*. *Geophysical Monograph Series*, vol. 150, pp. 59–68. Washington, DC: AGU.
Ricard Y, Doglioni C, and Sabadini R (1991) Differential rotation between lithosphere and mantle – A consequence of lateral mantle viscosity variations. *Journal of Geophysical Research* 96: 8407–8415.
Ricard Y, Fleitout L, and Froidevaux C (1984) Geoid heights and lithospheric stresses for a dynamic Earth. *Annals of Geophysics* 2: 267–286.
Ricard Y, Mattern E, and Matas J (2005) Earth's deep mantle: Structure, composition, and evolution. In: van der Hilst RD, Bass JD, Matas J, and Trampert J (eds.) *Synthetic Tomographic Images of Slabs from Mineral Physics*. *Geophysical Monograph Series*, pp. 285–302. Washington, DC: AGU.
Ricard Y, Richards M, Lithgow-Bertelloni C, and Stunff YL (1993a) A geodynamic model of mantle density heterogeneity. *Journal of Geophysical Research* 98: 21,895–21,909.
Ricard Y, Spada G, and Sabadini R (1993b) Polar wander of a dynamic earth. *Geophysical Journal International* 113: 284–298.
Ricard Y, Sramek O, and Dubuffet F (2009) A multi-phase model of runaway core–mantle segregation in planetary embryos. *Earth and Planetary Science Letters* 284(1–2): 144–150. http://dx.doi.org/10.1016/j.epsl.2009.04.021.
Ricard Y and Vigny C (1989) Mantle dynamics with induced plate-tectonics. *Journal of Geophysical Research* 94: 17,543–17,559.
Ricard Y, Vigny C, and Froidevaux C (1989) Mantle heterogeneities, geoid and plate motion: a monte carlo inversion. *Journal of Geophysical Research* 94: 13,739–13,754.
Richards M (1991) Hotspots and the case for a high viscosity lower mantle. In: Sabadini R, et al. (eds.) *Glacial Isostasy, Sea-Level and Mantle Rheology*, pp. 571–587. Dordrecht: Kluwer Academic Publishers.
Richards MA and Engebretson DC (1992) Large-scale mantle convection and the history of subduction. *Nature* 355: 437–440.
Richards MA and Hager BH (1984) Geoid anomalies in a dynamic earth. *Journal of Geophysical Research* 89: 5987–6002.
Richards MA, Ricard Y, Lithgow-Bertelloni G Spada, and Sabadini R (1997) An explanation for Earth's long-term rotational stability. *Science* 275: 372–375.
Richards MA, Yang WS, Baumgardner JR, and Bunge HP (2001) Role of a low-viscosity zone in stabilizing plate tectonics: Implications for comparative terrestrial planetology. *Geochemistry, Geophysics, Geosystems* 2.
Richter FM (1978) Experiments on the stability of convection rolls in fluids whose viscosity depends on temperature. *Journal of Fluid Mechanics* 89: 553–560.
Richter FM and Mckenzie D (1984) Dynamical models for melt segregation from a deformable matrix. *Journal of Geology* 92: 729–740.
Ricolleau A, Fiquet G, Perillat J, Daniel I, Menguy N, Cardon H, Addad A, Vanni C, and Guignot N (2004) The fate of subducted basaltic crust in the earth's lower mantle: an experimental petrological study. *Eos, Transactions of the American Geophysical Union* 85(47), Fall Meet. Suppl., Abstract U33B-02.
Riley GN, Kohlstedt DL, and Richter FM (1990) Melt migration in a silicate liquid-olivine system – An experimental test of compaction theory. *Geophysical Research Letters* 17: 2101–2104.
Ringwood AE and Irifune T (1988) Nature of the 650-km seismic discontinuity – Implications for mantle dynamics and differentiation. *Nature* 331: 131–136.
Roberts PH (1967) Convection in horizontal layers with internal heat generation: Theory. *Journal of Fluid Mechanics* 30: 33–49.
Rolf T, Coltice N, and Tackley PJ (2012) Linking continental drift, plate tectonics and the thermal state of the Earth's mantle. *Earth and Planetary Science Letters* 351: 134–146. http://dx.doi.org/10.1016/j.epsl.2012.07.011.
Rolf T and Tackley PJ (2011) Focussing of stress by continents in 3D spherical mantle convection with self-consistent plate tectonics. *Geophysical Research Letters* 38. http://dx.doi.org/10.1029/2011GL048677.
Ross M and Schubert G (1986) Tidal dissipation in a viscoelastic planet. *Journal of Geophysical Research* 91(B4): D447–D452. http://dx.doi.org/10.1029/JB091iB04p0D447.
Rozel A, Ricard Y, and Bercovici D (2011) A thermodynamically self-consistent damage equation for grain size evolution during dynamic recrystallization. *Geophysical Journal International* 184(2): 719–728. http://dx.doi.org/10.1111/j.1365-246X.2010.04875.x.
Rudge JF, Bercovici D, and Spiegelman M (2011) Disequilibrium melting of a two phase multicomponent mantle. *Geophysical Journal International* 184(2): 699–718. http://dx.doi.org/10.1111/j.1365-246X.2010.04870.x.
Rudnick RL and Fountain DM (1995) Nature and composition of the continental crust – A lower crustal perspective. *Reviews of Geophysics* 33: 267–309.
Runcorn SK (1964) Satellite gravity measurements and the laminar viscous flow model of the Earth's mantle. *Journal of Geophysical Research* 69: 4389.
Sabadini R and Yuen DA (1989) Mantle stratification and long-term polar wander. *Nature* 339: 373–375.
Saltzer RL, Stutzmann E, and van der Hilst RD (2004) Poisson's ratio in the lower mantle beneath Alaska: Evidence for compositional heterogeneity. *Journal of Geophysical Research* 109.
Samuel H and Farnetani CG (2003) Thermochemical convection and helium concentrations in mantle plumes. *Earth and Planetary Science Letters* 207: 39–56.
Schmalzl J, Houseman GA, and Hansen U (1996) Mixing in vigorous, time-dependent three-dimensional convection and application to Earth's mantle. *Journal of Geophysical Research* 101: 21,847–21,858.
Schmeling H (2000) *Physics and Chemistry of Partially Molten Rocks*. *Partial Melting and Melt Migration in a Convecting Mantle*. Dordrecht: Kluwer Academic Publisher.
Schmeling H and Jacoby WR (1981) On modeling the lithosphere in mantle convection with non-linear rheology. *Journal of Geophysics* 50: 89–100.
Schmeling H, Marquart G, and Ruedas T (2003) Pressure- and temperature-dependent thermal expansivity and the effect on mantle convection and surface observables. *Geophysical Journal International* 154: 224–229.
Schmeling H, et al. (2008) A benchmark comparison of spontaneous subduction models – Towards a free surface. *Physics of the Earth and Planetary Interiors* 171(1–4): 198–223. http://dx.doi.org/10.1016/j.pepi.2008.06.028.
Schubert G, Stevenson D, and Cassen P (1980) Whole planet cooling and the radiogenic heat-source contents of the earth and moon. *Journal of Geophysical Research* 85: 2531–2538.
Schubert G, Turcotte D, and Olson P (2001) *Mantle Convection in the Earth and Planets*. Cambridge, UK: Cambridge University Press.
Schubert G, Yuen DA, and Turcotte DL (1975) Role of phase-transitions in a dynamic mantle. *Geophysical Journal of the Royal Astronomical Society* 42: 705–735.
Scott DR and Stevenson DJ (1984) Magma solitons. *Geophysical Research Letters* 11: 9283–9296.
Scott DR and Stevenson DJ (1986) Magma ascent by porous flow. *Journal of Geophysical Research* 91: 9283–9296.
Scott DR and Stevenson DJ (1989) A self-consistent model of melting, magma migration and buoyancy-driven circulation beneath mid-ocean ridges. *Journal of Geophysical Research* 94: 2973–2988.
Shephard GE, Mueller RD, Liu L, and Gurnis M (2010) Miocene drainage reversal of the Amazon river driven by plate-mantle interaction. *Nature Geoscience* 3(12): 870–875. http://dx.doi.org/10.1038/NGEO1017.
Simpson G, Spiegelman M, and Weinstein MI (2010a) A multiscale model of partial melts: 1. Effective equations. *Journal of Geophysical Research* 115. http://dx.doi.org/10.1029/2009JB006375.
Simpson G, Spiegelman M, and Weinstein MI (2010b) A multiscale model of partial melts: 2. Numerical results. *Journal of Geophysical Research* 115. http://dx.doi.org/10.1029/2009JB006376.
Sleep NH (1988) Tapping of melt by veins and dikes. *Journal of Geophysical Research* 93(B9): 10255–10272. http://dx.doi.org/10.1029/JB093iB09p10255.
Solomatov VS (1995) Scaling of temperature-dependent and stress-dependent viscosity convection. *Physics of Fluids* 7: 266–274.
Spada G, Ricard Y, and Sabadini R (1992a) Excitation of true polar wander by subduction. *Nature* 360: 452–454.

Spada G, Sabadini R, Yuen DA, and Ricard Y (1992b) Effects on postglacial rebound from the hard rheology in the transition zone. *Geophysical Journal International* 109: 683–700.

Spear FS (1993) *Metamorphic Phase Equilibria and Pressure–Temperature–Time Paths*, p. 799. Washington, DC: Mineralogical Society of America.

Spiegel E and Veronis G (1960) On the Boussinesq approximation for a compressible fluid. *Astrophysical Journal* 131(2): 442–447. http://dx.doi.org/10.1086/146849.

Spiegelman M and Kelemen PB (2003) Extreme chemical variability as a consequence of channelized melt transport. *Geochemistry, Geophysics, Geosystems* 4.

Spiegelman M, Kelemen PB, and Aharonov E (2001) Causes and consequences of flow organization during melt transport: The reaction infiltration instability in compactible media. *Journal of Geophysical Research* 106: 2061–2077.

Spiegelman M and McKenzie D (1987) Simple 2-D models for melt extraction at midocean ridges and island arcs. *Earth and Planetary Science Letters* 83(1–4): 137–152. http://dx.doi.org/10.1016/0012-821X(87)90057-4.

Sramek O, Milelli L, Ricard Y, and Labrosse S (2012) Thermal evolution and differentiation of planetesimals and planetary embryos. *Icarus* 217(1): 339–354. http://dx.doi.org/10.1016/j.icarus.2011.11.021.

Sramek O, Ricard Y, and Bercovici D (2007) Simultaneous melting and compaction in deformable two phase media. *Geophysical Journal International* 168(3): 964–982.

Sramek O, Ricard Y, and Dubuffet F (2010) A multiphase model of core formation. *Geophysical Journal International* 181(1): 198–220. http://dx.doi.org/10.1111/j.1365-246X.2010.04528.x.

Stacey FD (1977) Application of thermodynamics to fundamental earth physics. *Geophysical Surveys* 3: 175–204.

Stacey FD and Davis PM (2004) High pressure equations of state with applications to the lower mantle and core. *Physics of the Earth and Planetary Interiors* 142: 137–184.

Stadler G, Gurnis M, Burstedde C, et al. (2010) The dynamics of plate tectonics and mantle flow: From local to global scales. *Science* 329(5995): 1033–1038. http://dx.doi.org/10.1126/science.1191223.

Stein C, Schmalzl J, and Hansen U (2004) The effect of rheological parameters on plate behaviour in a self-consistent model of mantle convection. *Physics of the Earth and Planetary Interiors* 142: 225–255.

Steinberger B and O'Connell RJ (1998) Advection of plumes in mantle flow: Implications for hotspot motion, mantle viscosity and plume distribution. *Geophysical Journal International* 132: 412–434.

Stevenson DJ (1989) Spontaneous small-scale melt segregation in partial melts undergoing deformation. *Geophysical Research Letters* 16: 1067–1070.

Stixrude L and Lithgow-Bertelloni C (2005) Thermodynamics of mantle minerals I. Physical properties. *Geophysical Journal International* 162(2): 610–632. http://dx.doi.org/10.1111/j.1365-246X.2005.02642.x.

Tackley P (1995) On the penetration of an endothermic phase transition by upwellings and downwellings. *Journal of Geophysical Research* 100: 15,477–15,488.

Tackley PJ (1996) Effects of strongly variable viscosity on three-dimensional compressible convection in planetary mantles. *Journal of Geophysical Research* 101: 3311–3332.

Tackley PJ (1998) Self-consistent generation of tectonic plates in three-dimensional mantle convection. *Earth and Planetary Science Letters* 157: 9–22.

Tackley P (2000a) The quest for self-consistent incorporation of plate tectonics in mantle convection. In: Richards MA, Gordon R, and van der Hilst R (eds.) *History and Dynamics of Global Plate Motions*. *AGU Geophysical Monograph Series*, vol. 121, pp. 47–72. Washington, DC: AGU.

Tackley PJ (2000b) Mantle convection and plate tectonics: Toward an integrated physical and chemical theory. *Science* 288: 2002–2007.

Tackley PJ (2000c) Self consistent generation of tectonic plates in time-dependent, three dimensional mantle convection simulations, part 1: Pseudoplastic yielding. *Geochemistry, Geophysics, Geosystems* 1.

Tackley PJ (2000d) Self consistent generation of tectonic plates in time-dependent, three dimensional mantle convection simulations, part 2: Strain weakening and asthenosphere. *Geochemistry, Geophysics, Geosystems* 1.

Tackley PJ (2008) Modelling compressible mantle convection with large viscosity contrasts in a three-dimensional spherical shell using the yin-yang grid. *Physics of the Earth and Planetary Interiors* 171(1–4): 7–18. http://dx.doi.org/10.1016/j.pepi.2008.08.005 (special issue).

Tackley PJ, Ammann M, Brodholt JP, Dobson DP, and Valencia D (2013) Mantle dynamics in super-Earth's: Post-perovskite rheology and self-regulation of viscosity. *Icarus* 225(1): 50–61. http://dx.doi.org/10.1016/j.icarus.2013.03.013.

Tackley PJ, Stevenson DJ, Glatzmaier GA, and Schubert G (1993) Effects of an endothermic phase-transition at 670 km depth in a spherical model of convection in the Earth's mantle. *Nature* 361: 699–704.

Tackley PJ and Xie SX (2002) The thermochemical structure and evolution of Earth's mantle: Constraints and numerical models. *Philosophical Transactions of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences* 360: 2593–2609.

Tao WC and Oconnell RJ (1993) Deformation of a weak subducted slab and variation of seismicity with depth. *Nature* 361: 626–628.

Tauzin B, Debayle E, and Wittlinger G (2010) Seismic evidence for a global low-velocity layer within the Earth's upper mantle. *Nature Geoscience* 3(10): 718–721. http://dx.doi.org/10.1038/NGEO969.

Toth J and Gurnis M (1998) Dynamics of subduction initiation at preexisting fault zones. *Journal of Geophysical Research* 103: 18,053–18,067.

Toussaint V, Carriere P, Scott J, and Gence JN (2000) Spectral decay of a passive scalar in chaotic mixing. *Physics of Fluids* 12: 2834–2844.

Tozer DC (1972) The present thermal state of the terrestrial planets. *Physics of the Earth and Planetary Interiors* 6: 182–197.

Tritton DJ (1988) *Physical Fluid Dynamics*. Oxford: Oxford University Press.

Trompert R and Hansen U (1998a) Mantle convection simulations with rheologies that generate plate-like behaviour. *Nature* 395: 686–689.

Trompert RA and Hansen U (1998b) On the Rayleigh number dependence of convection with a strongly temperature-dependent viscosity. *Physics of Fluids* 10: 351–360.

Trubienko O, Fleitout L, Garaud J-D, and Vigny C (2013) Interpretation of interseismic deformations and the seismic cycle associated with large subduction earthquakes. *Tectonophysics* 589: 126–141. http://dx.doi.org/10.1016/j.tecto.2012.12.027.

Trull T and Kurtz M (1993) Experimental measurements of the He-3 and He-4 mobility in olivine and clinopyroxene at magmatic temperatures. *Geochimica et Cosmochimica Acta* 57(6): 1313–1324. http://dx.doi.org/10.1016/0016-7037(93)90068-8.

Turcotte DL and Morgan JP (1992) Mantle flow and melt generation at mid-ocean ridges. In: Phipps Morgan J, Blackmann DK, and Simpson JM (eds.) *The Physics of Magma Migration and Mantle Flow Beneath a Mid-Ocean Ridge*, pp. 155–182. New York: AGU.

Turcotte DL and Schubert G (1982) *Geodynamics: Applications of Continuum Physics to Geological Problems*, pp. 1–450. New York: Wiley.

Urey HC (1951) The origin and development of the earth and other terrestrial planets. *Geochimica et Cosmochimica Acta* 1: 209–277.

Valencia D and O'Connell RJ (2009) Convection scaling and subduction on Earth and super-Earth's. *Earth and Planetary Science Letters* 286(3–4): 492–502. http://dx.doi.org/10.1016/j.epsl.2009.07.015.

Valencia D, O'Connell RJ, and Sasselov DD (2007a) Inevitability of plate tectonics on super-Earths. *Astrophysical Journal* 670(1, Part 2): L45–L48. http://dx.doi.org/10.1086/524012.

Valencia D, Sasselov DD, and O'Connell RJ (2007b) Radius and structure models of the first super-earth planet. *Astrophysical Journal* 656(1, Part 1): 545–551. http://dx.doi.org/10.1086/509800.

van der Hilst RD, Widiyantoro S, and Engdahl ER (1997) Evidence for deep mantle circulation from global tomography. *Nature* 386: 578–584.

van Heck HJ and Tackley PJ (2008) Planforms of self-consistently generated plates in 3D spherical geometry. *Geophysical Research Letters* 35(19). http://dx.doi.org/10.1029/2008GL035190.

van Heck HJ and Tackley PJ (2011) Plate tectonics on super-Earths: Equally or more likely than on Earth. *Earth and Planetary Science Letters* 310(3–4): 252–261. http://dx.doi.org/10.1016/j.epsl.2011.07.029.

van Keken PE, Hauri EH, and Ballentine CJ (2002) Mantle mixing: The generation, preservation, and destruction of chemical heterogeneity. *Annual Review of Earth and Planetary Sciences* 30: 493–525.

van Keken P, Yuen DA, and Vandenberg A (1992) Pulsating diapiric flows – Consequences of vertical variations in mantle creep laws. *Earth and Planetary Science Letters* 112: 179–194.

Vanderwal D, Chopra P, Drury M, and Fitz Gerald J (1993) Relationships between dynamically recrystallized grain-size and deformation conditions in experimentally deformed olivine rocks. *Geophysical Research Letters* 20(14): 1479–1482. http://dx.doi.org/10.1029/93GL01382.

Vanyo J, Wilde P, Cardin P, and Olson P (1995) Experiments on precessing flows in the earth's liquid core. *Geophysical Journal International* 121: 136–142.

Vinet P, Ferrante J, Rose JH, and Smith JR (1987) Compressibility of solids. *Journal of Geophysical Research* 92: 9319–9325.

Walker J, Hays P, and Kasting J (1981) A negative feedback mechanism for long-term stabilization of Earth's surface temperature. *Journal of Geophysical Research* 86(NC10): 9776–9782. http://dx.doi.org/10.1029/JC086iC10p09776.

Warren JM and Hirth G (2006) Grain size sensitive deformation mechanisms in naturally deformed peridotites. *Earth and Planetary Science Letters* 248(1–2): 438–450. http://dx.doi.org/10.1016/j.epsl.2006.06.006.

Weatherley SM and Katz RF (2012) Melting and channelized magmatic flow in chemically heterogeneous, upwelling mantle. *Geochemistry, Geophysics, Geosystems* 13. http://dx.doi.org/10.1029/2011GC003989.

Weertman J and Weertman JR (1975) High-temperature creep of rock and mantle viscosity. *Annual Review of Earth and Planetary Sciences* 3: 293–315.

Weinstein SA (1996) Thermal convection in a cylindrical annulus with a non-Newtonian outer surface. *Pure and Applied Geophysics* 146: 551–572.

Weinstein SA and Christensen U (1991) Convection planforms in a fluid with a temperature-dependent viscosity beneath a stress-free upper boundary. *Geophysical Research Letters* 18: 2035–2038.

Weinstein SA and Olson PL (1992) Thermal-convection with non-Newtonian plates. *Geophysical Journal International* 111: 515–530.

Williams Q and Garnero EJ (1996) Seismic evidence for partial melt at the base of Earth's mantle. *Science* 273: 1528–1530.

Woods LC (1975) *The Thermodynamics of Fluid Systems*. Oxford: Clarendon Press.

Yuen D, Sabadini RCA, and Gasperini P (1986) On transient rheology and glacial isostasy. *Journal of Geophysical Research* 91: 1420–1438.

Zhang SX and Christensen U (1993) Some effects of lateral viscosity variations on geoid and surface velocities induced by density anomalies in the mantle. *Geophysical Journal International* 114: 531–547.

Zhang SX and Yuen DA (1996) Various influences on plumes and dynamics in time-dependent, compressible mantle convection in 3-D spherical shell. *Physics of the Earth and Planetary Interiors* 94: 241–267.

Zhong SJ, Gurnis M, and Moresi L (1996) Free surface formulation of mantle convection, part 1: Basic theory and implication to plumes. *Geophysical Journal International* 127: 708–718.

Zhong SJ, McNamara A, Tan E, Moresi L, and Gurnis M (2008) A benchmark study on mantle convection in a 3-D spherical shell using citcoms. *Geochemistry, Geophysics, Geosystems* 9: Q07005. 11063–11082. http://dx.doi.org/10.1029/2008GC002048.

Zhong SJ, Zuber MT, Moresi L, and Gurnis M (2000) Role of temperature-dependent viscosity and surface plates in spherical shell models of mantle convection. *Journal of Geophysical Research* 105: 11063–11082.