

PHYSICS OF MANTLE CONVECTION

Yanick Ricard

Laboratoire de Science de la Terre

Université de Lyon; Université de Lyon 1; CNRS; Ecole Normale Supérieure de Lyon,
2 rue Raphaël Dubois, 69622, Villeurbanne Cedex, France.

abstract This chapter presents the fundamental equations necessary to understand the physics of fluids and some applications to mantle convection. The first section derives the equations of conservation for mass, momentum and energy, and the boundary and interface conditions for the various variables. The thermodynamic and rheological properties of solids are discussed in the second section. The mechanism of thermal convection and the classical Boussinesq and anelastic approximations are presented in the third section. As the subject is not often included in geophysical text books, an introduction to the physics of multicomponent and multiphase flows is given in a fourth section. At last the specifics of mantle convection are reviewed in the fifth section.

Contents

1	Introduction	4
2	Conservation equations	5
2.1	General expression of conservation equations	5
2.2	Mass conservation	7
2.3	Momentum conservation	8
2.3.1	General momentum conservation	8
2.3.2	Inertia and non-Galilean forces	9
2.3.3	Angular momentum conservation	11
2.4	Energy conservation	12
2.4.1	First law and internal energy	12
2.4.2	State variables	13
2.4.3	Temperature	14
2.4.4	Second law and entropy	15

2.5	Gravitational forces	17
2.5.1	Poisson equation	17
2.5.2	Self-gravitation	17
2.5.3	Conservative forms of momentum and energy equations	18
2.6	Boundary and interface conditions	20
2.6.1	General method	20
2.6.2	Interface conditions in the 1D case and for bounded variables	22
2.6.3	Phase change interfaces	22
2.6.4	Weakly deformable surface of a convective cell	24
3	Thermodynamic and rheological properties	25
3.1	Equation of State and solid properties	25
3.2	Rheology	28
3.2.1	Elasticity	28
3.2.2	Viscous Newtonian rheology	29
3.2.3	Maxwellian Visco-elasticity	30
3.2.4	Non-linear rheologies	32
4	Physics of convection	34
4.1	Convection experiments	34
4.2	Basic balance	35
4.3	Two simple solutions	36
4.3.1	The diffusive solution	36
4.3.2	The adiabatic solution	37
4.3.3	Stability of the adiabatic gradient	38
4.4	Approximate equations	39
4.4.1	Depth dependent reference profiles	39
4.4.2	Nondimensionalization	40
4.4.3	Anelastic approximation	42
4.4.4	Dimensionless numbers	43
4.4.5	Boussinesq approximation	44
4.4.6	Internal heating	45
4.4.7	Alternative forms	45
4.4.8	Change of nondimensionalization	47
4.5	Linear stability analysis for basally heated convection	47
4.6	Road to chaos	49
5	Introduction to physics of multicomponent and multiphase flows	51
5.1	Fluid dynamics of multicomponent flows in solution	53
5.1.1	Mass conservation in a multicomponent solution	53

5.1.2	Momentum and energy in a multicomponent solution	54
5.1.3	Entropy conservation in a multicomponent solution	56
5.1.4	Advection-diffusion equation and reaction rates	57
5.1.5	Conservation properties of the advection-diffusion equation	60
5.1.6	Laminar and turbulent stirring	62
5.1.7	Diffusion in Lagrangian coordinates	63
5.2	Fluid dynamics of two phase flows	65
5.2.1	Mass conservation for matrix and fluid	66
5.2.2	Momentum conservation of matrix and fluid	67
5.2.3	Energy conservation for two-phase flows	69
5.2.4	Phenomenological laws	70
5.2.5	Summary equations	73
6	Specifics of Earth's mantle convection	75
6.1	A mantle without inertia	75
6.1.1	Dynamic models	75
6.1.2	Mantle flow and post-glacial models	76
6.1.3	Time dependent models	77
6.2	A mantle with internal heating	79
6.3	A complex rheology	80
6.3.1	Temperature dependence of viscosity	82
6.3.2	Depth-dependence of viscosity	83
6.3.3	Stress-dependence of viscosity	84
6.3.4	Grain size dependence of viscosity	84
6.4	Importance of sphericity	85
6.5	Other depth-dependent parameters	85
6.5.1	Thermal expansivity variations	85
6.5.2	Increase in average density with depth	87
6.5.3	Thermal conductivity variations	88
6.6	Thermo-chemical convection	88
6.7	A complex lithosphere: plates and continents	90

1 Introduction

In many text books 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 ability to deform. Basically any material that appears as non-deformable in most common life experiments, with a crystalline structure (i.e., belonging to the solid state) or with a disordered structure (e.g., a glass, that 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 geological processes related to mantle convection, typically 10 Myrs (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...) 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 viscosity, μ , and elasticity (shear modulus), μ_r , called Maxwell time τ_M (Maxwell, 1831-1879),

$$\tau_M = \frac{\mu}{\mu_r}. \quad (1)$$

The rheological transition in some materials like silicon 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 years for the mantle (see 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 irreversibly deformable. Surface loading of the Earth by glaciation and deglaciation involves times of a few thousands 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. Convection can be kinematically forced by boundary conditions or induced by density variations. The density variations 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 petrological density anomalies or multi-phase physics. We will review some of these complexities.

The physics of fluid behavior, like the physics of elasticity is based on the general continuum hypothesis. This hypothesis requires that parameters like density, temperature, or momentum are defined everywhere, continuously. 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 grains.

2 Conservation equations

The basic equations of this section can be found in more detail in many classical text books (*Batchelor*, 1967; *Landau and Lifchitz*, 1980). We will only emphasize the aspects that are pertinent for the Earth.

2.1 General expression of conservation equations

Let us consider a fluid transported by the velocity field \mathbf{v} , 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, can be considered as a simple function of position and time, $A(\mathbf{X}, t)$, in a way very similar to the specification of the electromagnetic field. This is the Eulerian point of view (Euler, 1707-1783). The second point of view is traditionally attributed to Lagrange (Lagrange, 1736-1813). It considers the trajectory that a material element of the flow initially at \mathbf{X}_0 would follow $\mathbf{X}(\mathbf{X}_0, t)$. 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. For Euler the time derivative would simply be the rate of change seen by an observer at a fixed position, i.e., the partial derivative $\partial/\partial t$. For Lagrange, the time derivative noted with D would be the rate of change seen by an observer riding on a material particle

$$\frac{DA}{Dt} = \frac{dA(\mathbf{X}(\mathbf{X}_0, t), t)}{dt} = \sum_{i=1,3} \frac{\partial A}{\partial X_i} \frac{\partial X_i}{\partial t} + \frac{\partial A}{\partial t}, \quad (2)$$

where the X_i are the coordinates 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 material called the 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

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + (\mathbf{v} \cdot \nabla) \quad (3)$$

The operator $(\mathbf{v} \cdot \nabla)$ is the symbolic vector $(v_1\partial/\partial x_1, v_2\partial/\partial x_2, v_3\partial/\partial x_3)$ and a convenient mnemonic is to interpret it as the scalar product of the velocity field by the gradient operator $(\partial/\partial x_1, \partial/\partial x_2, \partial/\partial x_3)$. The operator $(\mathbf{v} \cdot \nabla)$ can be applied to a scalar or a vector. Notice that $(\mathbf{v} \cdot \nabla)\mathbf{A}$ is a vector that is neither parallel to \mathbf{A} nor to \mathbf{v} .

In a purely homogeneous fluid, the flow lines are not visible. In this case the point of view of Euler seems natural. On the other hand, a physicist describing elastic media can easily draw marks on the surface of deformable objects and flow lines become perceptible for him. He thus favors the Lagrangian point of view. 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 5.1.7).

A starting point for describing the physics of a continuum are the conservation equations. Consider a scalar or a vector extensive variable A (i.e., mass, momentum, energy, entropy, number of moles...) per unit volume and a virtual but fixed volume Ω enclosed in its boundary Σ . This virtual volume is freely crossed by the flow. The temporal change of A enclosed in Ω is

$$\frac{d}{dt} \int_{\Omega} A dV = \int_{\Omega} \frac{\partial A}{\partial t} dV. \quad (4)$$

Since Ω is fixed, the derivative of the integral is indeed 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 outflux) of A across the interface. This transport can either be a macroscopic advective transport by the flow or a more indirect transport, for example at a microscopic diffusive level. 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

$$\int_{\Omega} \frac{\partial A}{\partial t} dV = - \int_{\Sigma} \mathbf{J}_A \cdot d\mathbf{S} + \int_{\Omega} H_A dV, \quad (5)$$

where the infinitesimal surface is oriented with the outward unit normal. Equation (5) is the general form of any conservation equation. When the flux is regular

enough (when no surface of discontinuity crosses the volume Ω), we can make use of the divergence theorem

$$\int_{\Sigma} \mathbf{J}_A \cdot d\mathbf{S} = \int_{\Omega} \nabla \cdot \mathbf{J}_A dV, \quad (6)$$

to transform the surface integral into a volume integral. The divergence operator transforms the vector $\mathbf{J} = (J_1, J_1, J_2)$ into the scalar $\nabla \cdot \mathbf{J} = \sum_{i=1,3} \partial J_i / \partial x_i$, scalar product of the symbolic operator ∇ by the real vector \mathbf{J} (notice the difference between the scalar $\nabla \cdot \mathbf{J}$ and the operator $\mathbf{J} \cdot \nabla$). Since the integral equation (5) has been written for any volume Ω , we can deduce that the general differential form of the conservation equation is,

$$\frac{\partial A}{\partial t} + \nabla \cdot \mathbf{J}_A = H_A. \quad (7)$$

A similar expression can be used for a vector quantity \mathbf{A} with a tensor flux $\underline{\mathbf{J}}$ and a vector source term \mathbf{H}_A . In this case, the divergence operator converts the second-order tensor with components J_{ij} into a vector whose components are $\sum_{j=1,3} \partial J_{ij} / \partial x_j$.

We now apply this formalism to various physical quantities. Three quantities are strictly conserved: the mass, the momentum and the energy. 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). One very important physical quantity is not conserved - the entropy - but the second law of thermodynamics insures the positivity of the associated sources.

2.2 Mass conservation

The net rate at which mass is flowing is

$$J_\rho = \rho \mathbf{v}. \quad (8)$$

Using either the Eulerian or the Lagrangian time derivatives, the mass conservation can be written

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (9)$$

or

$$\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} = 0. \quad (10)$$

In an incompressible fluid, the Lagrangian observer does not see any density variation and $D\rho/Dt = 0$. In this case, mass conservation takes the simple form $\nabla \cdot \mathbf{v} = 0$. This equation is commonly called the continuity equation although this name is a little bit awkward.

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 scalar field A ,

$$\frac{\partial(\rho A)}{\partial t} + \nabla \cdot (\rho A \mathbf{v}) = \frac{D(\rho A)}{Dt} + \rho A \nabla \cdot \mathbf{v} = \rho \frac{DA}{Dt}, \quad (11)$$

and for any vector field \mathbf{A} ,

$$\frac{\partial(\rho \mathbf{A})}{\partial t} + \nabla \cdot (\rho \mathbf{A} \otimes \mathbf{v}) = \frac{D(\rho \mathbf{A})}{Dt} + \rho \mathbf{A} \nabla \cdot \mathbf{v} = \rho \frac{D\mathbf{A}}{Dt}, \quad (12)$$

where $\mathbf{A} \otimes \mathbf{v}$ is a tensor of components $A_i v_j$. When A is a constant the equalities (11) corresponds to the two expressions of mass conservation (9) and (10).

2.3 Momentum conservation

2.3.1 General momentum conservation

The changes of momentum can be easily deduced by balancing the changes of momentum with the body forces acting in the volume Ω and the surface force acting on its surface Σ . The total momentum is

$$\int_{\Omega} \rho \mathbf{v} dV, \quad (13)$$

and its variations are due to

- advective transport of momentum across the surface Σ ,
- forces acting on this surface and, last,
- internal body forces.

The momentum conservation can therefore be expressed in integral form as

$$\int_{\Omega} \frac{\partial(\rho \mathbf{v})}{\partial t} dV = - \int_{\Sigma} \rho \mathbf{v} (\mathbf{v} \cdot d\mathbf{S}) + \int_{\Sigma} \underline{\boldsymbol{\sigma}} \cdot d\mathbf{S} + \int_{\Omega} \mathbf{F} dV. \quad (14)$$

By definition, the tensor $\underline{\boldsymbol{\sigma}}$ corresponds to the total stresses applied on the surface Σ , (see Figure 1). The term \mathbf{F} represents the sum of all body forces, and in particular the gravitational forces $\rho \mathbf{g}$ (we will not consider the electromagnetic forces in this chapter).

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 $\rho (\mathbf{v} \otimes \mathbf{v}) \cdot d\mathbf{S}$) and the equality (12), the differential form of momentum conservation becomes

$$\rho \frac{D\mathbf{v}}{Dt} = \nabla \cdot \underline{\boldsymbol{\sigma}} + \mathbf{F}. \quad (15)$$

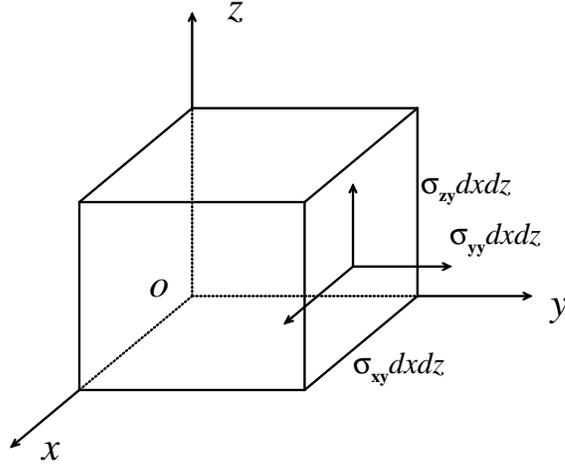


Figure 1: The force per unit area applied on a surface directed by the normal vector \mathbf{n}_i is by definition $\underline{\sigma} \cdot \mathbf{n}_i$. The component of this force along the unit vector \mathbf{e}_j therefore $\mathbf{e}_j \cdot \underline{\sigma} \cdot \mathbf{n}_i$.

It is usual to divide the total stress tensor into a thermodynamic pressure $-P\mathbf{I}$ where \mathbf{I} is the identity stress tensor, and a velocity-dependent stress $\underline{\tau}$. The relationship between the tensor $\underline{\tau}$ and the velocity field will be discussed later in section 3.2. Without motion, the total stress tensor is thus isotropic and equal to the usual pressure. In all the geophysical literature, it has been assumed that the velocity dependent tensor has no isotropic component, i.e., it is traceless $\text{tr}(\underline{\tau}) = 0$ or that $\text{tr}(\underline{\sigma}) = -3P$ (see section 3.2 for more details). The velocity-dependent stress tensor $\underline{\tau}$ is thus also the deviatoric stress tensor.

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

$$\rho \frac{D\mathbf{v}}{Dt} = -\nabla P + \nabla \cdot \underline{\tau} + \mathbf{F}. \quad (16)$$

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 (see 3.2).

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 and with respect to a Galilean frame

$$\frac{d\mathbf{e}_i}{dt} = \boldsymbol{\omega} \times \mathbf{e}_i, \quad (17)$$

where $\boldsymbol{\omega}$ is the angular velocity of Earth's rotation. A point $\mathbf{X} = \sum_i X_i \mathbf{e}_i$ of the Earth has therefore an acceleration $\boldsymbol{\gamma}$ in a Galilean frame (Galileo, 1564-1642)

$$\boldsymbol{\gamma} = \frac{\partial \mathbf{v}}{\partial t} + 2\boldsymbol{\omega} \times \mathbf{v} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{X}) + \frac{d\boldsymbol{\omega}}{dt} \times \mathbf{X}. \quad (18)$$

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

To quantify the importance of the three first acceleration terms (neglecting the Poincaré term), let us consider a characteristic scale for length (the Earth radius $a = 6371$ km), and mantle velocity ($U = 10$ cm yr⁻¹) and let us compare the various acceleration terms. One immediately gets

$$\frac{\text{inertia}}{\text{Coriolis}} = \frac{U}{2\omega a} = \frac{1}{2.9 \times 10^{11}}, \quad (19)$$

$$\frac{\text{Coriolis}}{\text{gravitational force}} = \frac{2\omega U}{g} = \frac{1}{2.1 \times 10^{13}}, \quad (20)$$

$$\frac{\text{centrifugal}}{\text{gravitational force}} = \frac{\omega^2 a}{g} = \frac{1}{291}. \quad (21)$$

A first remark is that the inertial term is much smaller than the Coriolis term (this ratio is also known as the Rossby number, (Rossby, 1898-1957)). The Coriolis term is itself perfectly negligible in front of the gravitational force. Even if we argue that a more meaningful comparison would be to compare the whole Coriolis force $2\rho\omega U$ to the lateral variations of the gravitational force $\delta\rho g$, (this ratio would be the inverse of the Eckman number, (Eckman, 1874-1954)), inertia and Coriolis accelerations play absolutely no role in mantle dynamics. Neglecting inertia means neglecting the changes in kinetic energy as inertia is the time-derivative of the kinetic energy. We can perform a simple numerical estimate to realize how ridiculously small the mantle kinetic energy is. The kinetic energy of a lithospheric plate (a square of size 2000 km, thickness 100 km, velocity 5 cm yr⁻¹ and density 3000 kg m⁻³) is 1.5 kJ, comparable to that of a middle size car (2000 kg) driven at only 8 km h⁻¹!

The centrifugal term is also quite small but not so small (1/291 of gravitational force). It controls two effects. The first is the Earth's flattening with an equatorial

bulge of 21 km (1/300 of Earth's radius). This is a static phenomenon that has no interactions with convection dynamics. The second is the possibility that the whole planet rotates along an equatorial axis in order to keep its main inertial axis coincident with its rotational axis (*Spada et al.*, 1992; *Ricard et al.*, 1993b). This rotational equilibrium of the Earth will not be discussed here (e.g., see *Munk and MacDonald* (1960) for the dynamics of a deformable rotating body).

We neglect all the acceleration terms in the following but we should remember that in addition to the convective motion of a non-rotating planet, a global rotation of the planet with respect to its rotation axis, is possible. This motion documented by paleomagnetism, is called True Polar Wander (*Besse and Courtillot*, 1991).

2.3.3 Angular momentum conservation

The angular momentum per unit mass $\mathbf{J} = \mathbf{r} \times \mathbf{v}$ is also a conserved quantity. Its conservation law can be obtained by 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 angular momentum sources, its variations are due to

- advective transport of angular momentum across the surface Σ ,
- torque of forces acting on this surface and, last,
- torque of internal body forces.

The resulting balance is therefore,

$$\int_{\Omega} \frac{\partial(\rho\mathbf{J})}{\partial t} dV = - \int_{\Sigma} \rho\mathbf{J}(\mathbf{v} \cdot d\mathbf{S}) + \int_{\Sigma} \mathbf{r} \times (\underline{\boldsymbol{\sigma}} \cdot d\mathbf{S}) + \int_{\Omega} \mathbf{r} \times \mathbf{F} dV. \quad (22)$$

The only difficulty to transform this integral form into a local equation is with the integral involving the stress tensor. After some algebra, the equation (22) becomes

$$\rho \frac{D\mathbf{J}}{Dt} = \mathbf{r} \times \nabla \cdot \underline{\boldsymbol{\sigma}} + \mathbf{r} \times \mathbf{F} + \mathcal{T}, \quad (23)$$

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

$$\rho \frac{D\mathbf{J}}{Dt} = \mathbf{r} \times \nabla \cdot \underline{\boldsymbol{\sigma}} + \mathbf{r} \times \mathbf{F}, \quad (24)$$

differs from (23) by the absence of the torque \mathcal{T} . This proves that in the absence of sources of angular momentum, the stress (either $\underline{\boldsymbol{\sigma}}$ or $\underline{\boldsymbol{\tau}}$) must be represented by a symmetrical tensor,

$$\underline{\boldsymbol{\sigma}} = \underline{\boldsymbol{\sigma}}^t, \quad \underline{\boldsymbol{\tau}} = \underline{\boldsymbol{\tau}}^t. \quad (25)$$

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

2.4 Energy conservation

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 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

$$\int_{\Omega} \rho \left(\mathcal{U} + \frac{v^2}{2} \right) dV. \quad (26)$$

A change of this energy content can be caused by

- advection of energy across the boundary Σ by the macroscopic flow,
- transfer of energy through the same surface, but at a microscopic level
- work of body forces,
- work of surface forces, and, last
- radioactive heat production

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

$$\frac{\partial}{\partial t} \left(\rho \left(\mathcal{U} + \frac{v^2}{2} \right) \right) = -\nabla \cdot \left(\rho \left(\mathcal{U} + \frac{v^2}{2} \right) \mathbf{v} + \mathbf{q} + P\mathbf{v} - \underline{\boldsymbol{\tau}} \cdot \mathbf{v} \right) + \mathbf{F} \cdot \mathbf{v} + \rho H, \quad (27)$$

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

This expression can be developed and simplified by using the equality (11) and the equations of mass and momentum conservation, (9) and (16) to reach the form

$$\rho \frac{D\mathcal{U}}{Dt} = -\nabla \cdot \mathbf{q} - P\nabla \cdot \mathbf{v} + \underline{\boldsymbol{\tau}} : \nabla \mathbf{v} + \rho H. \quad (28)$$

The viscous dissipation term $\underline{\boldsymbol{\tau}} : \nabla \mathbf{v}$ is the contraction of the two tensors $\underline{\boldsymbol{\tau}}$ and $\nabla \mathbf{v}$. Its expression is $\sum_{ij} \tau_{ij} \partial v_i / \partial x_j$ (be careful with the difference between the scalar $\nabla \cdot \mathbf{v}$ and the tensor $\nabla \mathbf{v}$ of components $\partial v_i / \partial x_j$).

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 thermodynamics presentation, but since we apply thermodynamics per unit mass, the volume V is in fact the volume per unit mass or $1/\rho$. We use the first Law 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 exchanges of heat and work are not exact differentials: the entire precise process has to be known to compute these exchanges, not only the initial and final stages. Using either $T - V$ or $T - P$ variables, we can write

$$\delta Q = C_V dT + l dV = C_P dT + h dP, \quad (29)$$

where C_P and C_V are the heat capacities at constant pressure and volume and h and l are two other calorimetric coefficients necessary to account for the heat exchanges at constant temperature. In the formalism of the physics of fluids the exchange of work is only due to the work of pressure forces

$$\delta W = -P dV. \quad (30)$$

This implies that only the pressure term corresponds to an energy capable to be stored and returned without loss when the volume change is reversed. On the contrary the stresses related to the velocity will ultimately appear in the dissipative, irrecoverable term of viscous dissipation. This point will be further commented in the section 3.2 about rheology.

Thermodynamics states that the total variations of energy, $d\mathcal{U} = \delta Q + \delta W$, enthalpy, $d\mathcal{H} = d\mathcal{U} + d(PV)$ or entropy, $d\mathcal{S} = \delta Q/T$, are exact differentials. This means that the difference in energy (enthalpy, entropy) between an initial and a final state only depends of the initial and final states themselves, not the intermediate stages. This implies mathematically that their second partial derivatives with respect to any pair of variables are independent of the order of differentiation. Using these rules a large number of equalities 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 and h ,

$$l = \alpha T K_T, \quad \text{and} \quad h = -\frac{\alpha T}{\rho}. \quad (31)$$

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

$$\begin{aligned}\alpha &= \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_P = -\frac{1}{\rho} \left(\frac{\partial \rho}{\partial T} \right)_P, \\ K_T &= -V \left(\frac{\partial P}{\partial V} \right)_T = \rho \left(\frac{\partial P}{\partial \rho} \right)_T.\end{aligned}\tag{32}$$

The thermodynamic laws and differentials apply to a closed deformable volume $\Omega(t)$. This corresponds to the perspective of Lagrange. We can therefore interpret the differential symbols "d" of the thermodynamic definitions (29) of (30) as Lagrangian derivatives "D".

To summarize, when the expressions for l and h are taken into account, (31), and when the differential symbols are interpreted as Lagrangian derivatives, the first law, $d\mathcal{U} = \delta Q + \delta W$ can be recast as

$$\frac{DU}{Dt} = C_V \frac{DT}{Dt} + (\alpha T K_T - P) \frac{\nabla \cdot \mathbf{v}}{\rho},\tag{33}$$

or as

$$\frac{DU}{Dt} = C_P \frac{DT}{Dt} - \frac{\alpha T}{\rho} \frac{DP}{Dt} - P \frac{\nabla \cdot \mathbf{v}}{\rho}.\tag{34}$$

In these equations we also have replaced the volume variation using mass conservation (9)

$$\frac{DV}{Dt} = \frac{D(1/\rho)}{Dt} = -\frac{1}{\rho^2} \frac{D\rho}{Dt} = \frac{\nabla \cdot \mathbf{v}}{\rho}.\tag{35}$$

2.4.3 Temperature

We can now identify these two thermodynamic equations (33) or (34), with our conservation equation deduced from fluid mechanics, (28), to express the conservation of energy in terms of temperature variations

$$\begin{aligned}\rho C_P \frac{DT}{Dt} &= -\nabla \cdot \mathbf{q} + \alpha T \frac{DP}{Dt} + \boldsymbol{\tau} : \nabla \mathbf{v} + \rho H, \\ \rho C_V \frac{DT}{Dt} &= -\nabla \cdot \mathbf{q} - \alpha T K_T \nabla \cdot \mathbf{v} + \boldsymbol{\tau} : \nabla \mathbf{v} + \rho H.\end{aligned}\tag{36}$$

In these two equivalent expressions, in addition to diffusion, three sources of temperature variations appear on the right side. 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 ^{235}U , ^{238}U , ^{236}Th and ^{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 6×10^9 inhabitants, the total natural radioactivity of the Earth is only ~ 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 drives the plates, raises mountains and produces a magnetic field. In addition to the present-day radioactivity, extinct radioactivities, like that of ^{36}Al (with a half life of 0.73 Myr), have played an important role in the initial stage of planet formation (*Lee et al.*, 1976).

The viscous dissipation term $\underline{\tau} : \nabla \mathbf{v}$ converts mechanical energy into a temperature increase. This term explains the classical Joule experiment (equivalence between work and heat, (Joule, 1818-1889)) 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 same dissipative term explains why clay (e.g., silly putty) heats when kneaded in our hands.

The remaining source term, containing the thermodynamic coefficients (α or αK_T in (36)), cancels 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 4.3.2.

2.4.4 Second law and entropy

We now apply the second law of thermodynamics and first express the entropy conservation law. Assuming local thermodynamic equilibrium, we have $d\mathcal{U} = Td\mathcal{S} - PdV$. Using the equation of conservation for the internal energy \mathcal{U} , (28), and expressing the volume change in term of velocity divergence (35), we obtain

$$\rho T \frac{D\mathcal{S}}{Dt} = -\nabla \cdot \mathbf{q} + \underline{\tau} : \nabla \mathbf{v} + \rho H. \quad (37)$$

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

$$\frac{\partial(\rho\mathcal{S})}{\partial t} = -\nabla \cdot \left(\rho\mathcal{S}\mathbf{v} + \frac{\mathbf{q}}{T} \right) - \frac{1}{T^2} \mathbf{q} \cdot \nabla T + \frac{1}{T} \underline{\tau} : \nabla \mathbf{v} + \frac{1}{T} \rho H. \quad (38)$$

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 , to radioactive production and to two other production terms.

A brief introduction to the general principles of non-equilibrium thermodynamics will be given in 5.1.4. Here, we simply state that the second law requires that in all situations, the total entropy production is positive. When the entropy terms are of different tensor orders (tensors, vectors or scalars), the fact that they are affected differently by a change of coordinates, implies for isotropic systems that the production terms of different tensor orders 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

$$-\mathbf{q} \cdot \nabla T \geq 0, \quad \text{and} \quad \boldsymbol{\tau} : \nabla \mathbf{v} \geq 0. \quad (39)$$

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

$$\mathbf{q} = -k \nabla T, \quad (40)$$

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

$$\kappa = \frac{k}{\rho C_P}, \quad (41)$$

(in principle isobaric and isochoric thermal diffusivities should be defined). In situations with uniform conductivity, without motion and radioactivity sources, the energy equation (36) becomes the standard diffusion equation

$$\frac{\partial T}{\partial t} = \kappa \nabla^2 T. \quad (42)$$

The relation between stress and velocity will be discussed in detail in section 3.2. We will show that the relationship

$$\boldsymbol{\tau} = 2\mu \left(\dot{\boldsymbol{\epsilon}} - \frac{1}{3} \nabla \cdot \mathbf{v} \right) \quad (43)$$

is appropriate for the mantle, where the strain rate tensor $\dot{\boldsymbol{\epsilon}}$ is defined by

$$\dot{\boldsymbol{\epsilon}} = \frac{1}{2} \left([\nabla \mathbf{v}] + [\nabla \mathbf{v}]^t \right). \quad (44)$$

Using this relation and assuming μ uniform, the divergence of the stress tensor that appears in the momentum conservation equation has the simple form

$$\nabla \cdot \boldsymbol{\tau} = \mu \nabla^2 \mathbf{v} + \frac{\mu}{3} \nabla (\nabla \cdot \mathbf{v}), \quad (45)$$

where the vectorial Laplacien ∇^2 stands for $\nabla(\nabla \cdot) - \nabla \times (\nabla \times)$. This relationship suggests a meaningful interpretation of the viscosity. The momentum equation (16), can be written

$$\frac{\partial \mathbf{v}}{\partial t} = \frac{\mu}{\rho} \nabla^2 \mathbf{v} + \text{other terms} \dots \quad (46)$$

forgetting the other terms, a comparison with the thermal diffusion equation (42) shows that the kinematic viscosity ν , defined by

$$\nu = \frac{\mu}{\rho}, \quad (47)$$

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.

2.5 Gravitational forces

2.5.1 Poisson equation

In this chapter, the only force is the gravitational body force. The gravity is the sum of this gravitational body force and the centrifugal force already discussed (see 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)

$$\mathbf{g} = -\nabla\psi, \quad \text{and} \quad \nabla^2\psi = 4\pi G\rho, \quad (48)$$

where G is the gravitational constant (some textbooks define $-\psi$ as the gravitational potential). 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.

2.5.2 Self-gravitation

When dealing with fluid dynamics at the laboratory scale, the gravitational force can be considered as constant. The gravitational force is related to the entire distribution of mass in the Earth (in the Universe) and is practically independent of the local changes in density in the experimental room. Therefore, at the laboratory scale, it is reasonable to forget Poisson's equation and to assume that $\mathbf{g} = \mathbf{g}_0$, 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 $\delta\rho$, origin of a gravitational perturbation $\delta\mathbf{g}$. The force term, \mathbf{F} is therefore $\rho_0\mathbf{g}_0 + \delta\rho\mathbf{g}_0 + \rho_0\delta\mathbf{g}$ at first order. It is tempting to assume that each term in the previous 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 feed back between density perturbations and flow). Practically, this 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 above idea and show that unfortunately the third term, namely the force due to the perturbation of gravitation acting on the average density, may be of the same order as the second one (*Ricard et al.*, 1984; *Richards and Hager*, 1984; *Panasyuk et al.*, 1996). 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 just like 2D 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 degrees) 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 which allows an easy 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 ($\delta(r - a)$ is the Dirac delta function (*Dirac*, 1902-1984), σ has unit of kg m^{-2}). This mass distribution generates the radial gravitational perturbation field inside the planet

$$\delta g = 4\pi G\sigma \frac{l}{2l+1} \left(\frac{r}{a}\right)^{l-1} Y_{lm}(\theta, \phi), \quad (49)$$

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 $g_0 = 4/3\pi G\rho_0 a$. Since $\langle\delta\rho\rangle = \sigma Y_{lm}(\theta, \phi)$, we get

$$\frac{\rho_0\langle\delta g\rangle}{g_0\langle\delta\rho\rangle} = \frac{3}{2l+1}. \quad (50)$$

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 $\rho_0\delta g$ is about 50% of the direct effect $\delta\rho g_0$ and reaches 10% of it only near $l \sim 15$. Self-gravitation has been taken into account in various models intended to explain the Earth's gravity field from mantle density anomalies (see Chapter 7 and Chapter 8 of this Treatise). Some spherical convection codes seem to neglect this effect although it is important at the longest wavelengths.

2.5.3 Conservative forms of momentum and energy equations

In the general remarks on conservation laws in section 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 (27) two terms, ρg and $\rho g \cdot \mathbf{v}$, appear

as sources (we also said that the radioactive term ρH , appears because the classical physics does not identify mass and energy. A negligible term $-\rho H/c^2$, 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 such a form is appropriate to deal with global balance, interfaces and boundaries (see section 2.6). We can obtain conservative equations by using Poisson's relation and performing some algebra

$$\rho \mathbf{g} = -\frac{1}{4\pi G} \nabla \cdot \left(\mathbf{g} \otimes \mathbf{g} - \frac{1}{2} g^2 \mathbf{I} \right) \quad (51)$$

$$\rho \mathbf{g} \cdot \mathbf{v} = -\rho \frac{D\psi}{Dt} - \frac{1}{4\pi G} \nabla \cdot \left(\mathbf{g} \frac{\partial \psi}{\partial t} \right) - \frac{1}{8\pi G} \frac{\partial g^2}{\partial t} \quad (52)$$

If we substitute these two expressions in the momentum and the energy conservation equations, (16) and (27), we obtain the conservative forms

$$\frac{\partial(\rho \mathbf{v})}{\partial t} = -\nabla \cdot \left(\rho \mathbf{v} \otimes \mathbf{v} + P \mathbf{I} - \boldsymbol{\tau} + \frac{1}{4\pi G} \mathbf{g} \otimes \mathbf{g} - \frac{1}{8\pi G} g^2 \mathbf{I} \right). \quad (53)$$

$$\begin{aligned} \frac{\partial}{\partial t} \left(\rho \left(\mathcal{U} + \psi + \frac{v^2}{2} \right) + \frac{g^2}{8\pi G} \right) = \\ -\nabla \cdot \left(\rho \mathbf{v} \left(\mathcal{U} + \psi + \frac{v^2}{2} \right) + \mathbf{q} + P \mathbf{v} - \boldsymbol{\tau} \cdot \mathbf{v} + \frac{\mathbf{g}}{4\pi G} \frac{\partial \psi}{\partial t} \right) + \rho H, \end{aligned} \quad (54)$$

When the gravitational force is time independent, a potential energy 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...), 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 magneto-hydrodynamics (Poynting, 1852-1914)).

In a permanent or in a statistically steady regime, the time-dependent terms of energy equation (54) can be neglected. The equation can then be integrated over the volume of the Earth. The natural assumption is that at the Earth's surface the velocities are perpendicular to the Earth's normal vector and that the surface is stress free. Using the divergence theorem to transform the volume integral of the divergence back to a surface integral of flux, most terms cancel and only remains

$$\int_{\Sigma} \mathbf{q} \cdot d\mathbf{S} = \int_{\Omega} \rho H dV. \quad (55)$$

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

It is surprising at first, that heat dissipation does not appear in this balance. To understand this point, we can directly integrate the energy equation written in terms of temperature (36),

$$\int_{\Omega} \rho C_P \frac{DT}{Dt} dV = - \int_{\Sigma} \mathbf{q} \cdot d\mathbf{S} + \int_{\Omega} \left(\alpha T \frac{DP}{Dt} + \boldsymbol{\tau} : \nabla \mathbf{v} \right) dV + \int_{\Omega} \rho H dV. \quad (56)$$

On the right side, the first and last terms cancel each other out. On the left side, we can use (11) to replace $\rho DT/Dt$ by $\partial(\rho T)/\partial t + \nabla \cdot (\rho \mathbf{v} T)$. The modest assumptions that C_P is a constant and that the temperature is statistically constant lead to

$$\int_{\Omega} \left(\alpha T \frac{DP}{Dt} + \boldsymbol{\tau} : \nabla \mathbf{v} \right) dV = 0. \quad (57)$$

The total heat production due to dissipation is balanced by the work due to compression and expansion over the convective cycle (*Hewitt et al., 1975*). 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.

2.6 Boundary and interface conditions

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. To obtain the interface conditions for a quantity A , the general method is to start from the conservation equation of A in its integral form (see (5)). We choose a cylindrical volume Ω (a pill-box) 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). 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 for the mantle). Since the surface of the pill-box $\Sigma(\epsilon)$ remains finite we must have

$$\lim_{\epsilon \rightarrow 0} \int_{\Sigma(\epsilon)} \mathbf{J}_A \cdot d\mathbf{S} = 0, \quad (58)$$

(the demonstration is here written for a vector flux, but is easily extended to tensor flux). Let us assume that the quantity \mathbf{J} varies very rapidly but continuously across

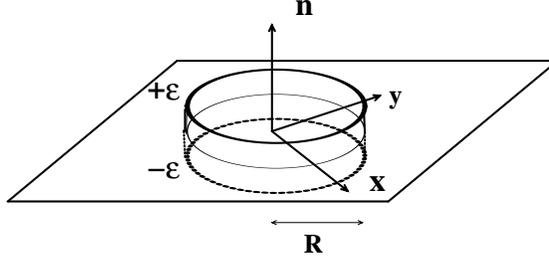


Figure 2: The pillbox volume used to derive the interface and boundary conditions.

$\pm\epsilon$ so that the divergence theorem can be applied

$$\lim_{\epsilon \rightarrow 0} \int_{\Omega(\epsilon)} \nabla \cdot \mathbf{J}_A dV = 0. \quad (59)$$

We now consider a reference frame where z is along the normal to the interface, \mathbf{n} , and x and y tangent to the interface. In the previous integral, the divergence can be splitted into a the 2D divergence using $\nabla_H = (\partial/\partial x, \partial/\partial y, 0)$ and a vertical derivative,

$$\int dx dy \lim_{\epsilon \rightarrow 0} \left(\int_{-\epsilon}^{\epsilon} \frac{\partial \mathbf{J}_A \cdot \mathbf{n}}{\partial z} dz + \nabla_H \cdot \int_{-\epsilon}^{\epsilon} \mathbf{J}_A dz \right) = 0. \quad (60)$$

The boundary condition is therefore

$$[\mathbf{J}_A] \cdot \mathbf{n} + \nabla_H \cdot \lim_{\epsilon \rightarrow 0} \int_{-\epsilon}^{\epsilon} \mathbf{J}_A dz = 0, \quad (61)$$

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} are bounded, or is exactly zero when the flux is only function of z . In this case, the boundary condition for A becomes

$$[\mathbf{J}_A] \cdot \mathbf{n} = 0. \quad (62)$$

At an interface, the normal flux of A must therefore be continuous. However in some cases, e.g., when \mathbf{J} varies with x and y but contains a z -derivative, the second term may not cancel and this happens in the case of boundaries associated with phase changes.

2.6.2 Interface conditions in the 1D case and for bounded variables

Using the mass, momentum, energy and entropy conservations in their conservative forms in (9), (53), (54) and (38) and assuming for now that no variable becomes infinite at an interface, the interface conditions in the reference frame where the interface is motionless are

$$\begin{aligned}
 [\rho \mathbf{v}] \cdot \mathbf{n} &= 0, \\
 [\boldsymbol{\tau}] \cdot \mathbf{n} - [P] \mathbf{n} &= 0, \\
 [\rho \mathbf{v} \mathcal{U} + \mathbf{q} - \boldsymbol{\tau} \cdot \mathbf{v} + P \mathbf{v}] \cdot \mathbf{n} &= 0, \\
 [\rho \mathbf{v} \mathcal{S} + \frac{\mathbf{q}}{T}] \cdot \mathbf{n} &= 0,
 \end{aligned} \tag{63}$$

(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 (63) 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 $[\rho \mathbf{v} v^2 / 2] \cdot \mathbf{n}$ to the third), these equations are known as Hugoniot-Rankine conditions (Hugoniot, 1851-1887; Rankine, 1820-1872).

At the surface of a fluid, and on any impermeable interfaces where $\mathbf{v} \cdot \mathbf{n} = 0$, the general jump conditions (63) without inertia, imply that the heat flux, $[\mathbf{q}] \cdot \mathbf{n}$, the entropy flux $[\mathbf{q}/T] \cdot \mathbf{n}$ (and therefore the temperature T) and the stress components $[\boldsymbol{\tau}] \cdot \mathbf{n} - [P] \mathbf{n}$ are continuous.

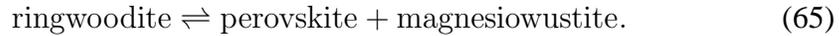
In 3-D, four boundary conditions are necessary on a surface to solve for the three components of velocity and for the temperature. The temperature (or the heat flux) can be imposed and, for the velocity, either free slip ($\mathbf{v} \cdot \mathbf{n} = 0$ and $\boldsymbol{\tau} \cdot \mathbf{n} - (\mathbf{n} \cdot \boldsymbol{\tau} \cdot \mathbf{n}) \mathbf{n} = 0$), or no slip ($\mathbf{v} = 0$), boundary conditions are generally used.

2.6.3 Phase change interfaces

The mantle minerals undergo several phase transitions at depth and at least two of them, the olivine-wadsleyite and the ringwoodite-perovskite-magnesiowustite transitions around 410 and 660 km depth



and



are sharp enough to be modeled by discontinuities. Conditions (63) suggest that $[\rho \mathbf{v}] \cdot \mathbf{n} = 0$ and $[\boldsymbol{\tau} \cdot \mathbf{n} - P \mathbf{n}] = 0$ and these seem to be the conditions used in convection codes. However as pointed by (Corrieu *et al.*, 1995), the first condition is

correct, not the second one. The problem arises from the term in $\partial v_z/\partial z$ present in the rheological law (98) that becomes infinite when the flow is asked 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 pill-box does not cancel when the pill-box height is decreased.

The only terms may be unbounded on the interface are σ_{xx} , σ_{yy} and σ_{zz} . Because of the presence of ∇_H , only the first two normal stress components appear in the integral of (61). The continuity of stress is thus

$$[\boldsymbol{\tau}] \cdot \mathbf{n} - [P]\mathbf{n} + \mathbf{e}_x \frac{\partial}{\partial x} \lim_{\epsilon \rightarrow 0} \int_{-\epsilon}^{\epsilon} \sigma_{xx} dz + \mathbf{e}_y \frac{\partial}{\partial y} \lim_{\epsilon \rightarrow 0} \int_{-\epsilon}^{\epsilon} \sigma_{yy} dz = 0. \quad (66)$$

Using $\sigma_{xx} = \sigma_{zz} + 2\mu \partial v_x/\partial x - 2\mu \partial v_z/\partial z$, and assuming that the viscosity remains uniform, we see that

$$\lim_{\epsilon \rightarrow 0} \int_{-\epsilon}^{\epsilon} \sigma_{xx} dz = -2\mu \lim_{\epsilon \rightarrow 0} \int_{-\epsilon}^{\epsilon} \frac{\partial v_z}{\partial z} dz = -2\mu[v_z]. \quad (67)$$

The same result holds for the σ_{yy} term. Since v_z is discontinuous, forcing a sudden change in volume implies a discontinuity of the tangential stresses. The boundary conditions are thus

$$[\tau_{xz}] - 2\mu \frac{\partial}{\partial x}[v_z] = [\tau_{yz}] - 2\mu \frac{\partial}{\partial y}[v_z] = [\tau_{zz} - P] = [v_x] = [v_y] = [\rho v_z] = 0. \quad (68)$$

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,

$$\begin{aligned} [\rho \mathbf{v}(\mathcal{U} + \frac{P}{\rho})] \cdot \mathbf{n} + [\mathbf{q}] \cdot \mathbf{n} &= 0, \\ [\rho \mathbf{v} \mathcal{S}] \cdot \mathbf{n} + \frac{1}{T}[\mathbf{q}] \cdot \mathbf{n} &= 0. \end{aligned} \quad (69)$$

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

$$\Delta \mathcal{H} = T \Delta \mathcal{S}, \quad (70)$$

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

$$\Delta \mathcal{H} \rho \mathbf{v} \cdot \mathbf{n} + [\mathbf{q}] \cdot \mathbf{n} = 0, \quad (71)$$

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

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 resulting. These stresses, according to the second interface condition (63), 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, rigid enough to resist the 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 discussing this point in more detail.

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

$$(\underline{\boldsymbol{\tau}} \cdot \mathbf{n} - P\mathbf{n})_{\text{on } z=h} = 0, \quad (72)$$

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

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

$$\frac{\partial h}{\partial t} + \mathbf{v}_H \cdot \nabla_H h - v_z^0 = 0. \quad (73)$$

This equation expresses the fact that a material particle on the surface remains always on it. In this expression \mathbf{v}_H and v_z^0 are the horizontal and vertical velocity components at the surface of the planet. We will see in section 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, $|\nabla_H h| \ll 1$. Boundary condition (72) and topography advection (73) can therefore be expanded at first order to give

$$(\underline{\boldsymbol{\tau}} \cdot \mathbf{e}_z - P\mathbf{e}_z)_{\text{on } z=0} \sim -\rho_0 g_0 h \mathbf{e}_z, \quad (74)$$

$$\frac{dh}{dt} = v_z^0, \quad (75)$$

where we again make use of the fact that the total stress remains close to hydrostatic (ρ_0 and g_0 are the surface values of density and gravity). At first order, the stress boundary condition on a weakly deformable top surface is therefore a zero shear stress and a time dependent topography related to the surface velocity. This topography applies an equivalent normal stress at the reference level $z = 0$.

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 $\bar{\mu}/\rho_0 g_0 \lambda$ where $\bar{\mu}$ 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.

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 short wavelength structures and for rapid events (e.g., for a localized thermal anomaly impinging the Earth's surface), the time for topographic equilibration becomes comparable to the time scale of the internal process. 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*).

3 Thermodynamic and rheological properties

The section 2 on conservation equations is valid for all fluids. 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 this section 3 and will be more specific in the last section 6.

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, $PV/T = \text{constant}$, is well known, but irrelevant for solids. Unfortunately there is no equation for solids based on a simple and efficient theoretical model. In the Earth mineralogical community, the third order finite strain Birch-Murnaghan EoS seems highly favored (*Birch, 1952*). This equation is cumbersome and does not rest on any sound physical basis. 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.

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

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

where the thermal expansivity α and incompressibility K_T have been defined in (32). This expression is a first order expansion of any EoS.

Equation (76) can be misleading if one forgets that the parameters α and K_T cannot be constant but must be related through Maxwell relations (for example, their definitions (32) imply that $\partial(\alpha\rho)/\partial P = -\partial(\rho/K_T)/\partial T$). Some models can be found in the geophysical literature in which assumptions made inconsistently about thermodynamic parameters (either constant or depth-dependent) violate the Maxwell rules.

Equation (76) can be used for a very simple numerical estimate that illustrates an important characteristics of solid Earth geophysics. Typically for silicates $\alpha \sim 10^{-5} \text{ K}^{-1}$, $K_T \sim 10^{11} \text{ Pa}$, 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 ($\alpha\Delta T \ll 1$ but $\Delta P/K_T \sim 1$). In planets, at first order, the radial density is only a function of pressure, not of temperature. This is opposite to most physical common sense gained from 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)

$$\Gamma = \frac{\alpha K_T}{\rho C_V} = \frac{1}{\rho C_V} \left(\frac{\partial P}{\partial T} \right)_V. \quad (77)$$

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,

$$\Gamma = \Gamma_0 \left(\frac{\rho_0}{\rho} \right)^q, \quad (78)$$

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 equipartioned in the three directions 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 7 atoms for a molar mass of 140 g, its heat capacity at constant volume is therefore close to $C_{Vm} = 21R = 174.56 \text{ J K}^{-1}\text{mol}^{-1}$ or $C_V = 1247 \text{ J K}^{-1}\text{kg}^{-1}$. The approximate constancy of C_V and the fact that Γ is only a function of ρ allow us to integrate (77)

$$P = F(\rho) + \alpha_0 K_T^0 (T - T_0) \left(\frac{\rho}{\rho_0} \right)^{1-q}, \quad (79)$$

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 dynamists, is the Murnaghan EoS (Murnaghan, 1951) at constant T that allows us to write an EoS for solids of the form

$$P = \frac{K_T^0}{n} \left[\left(\frac{\rho}{\rho_0} \right)^n - 1 \right] + \alpha_0 K_T^0 (T - T_0) \left(\frac{\rho}{\rho_0} \right)^{1-q}, \quad (80)$$

with an exponent n of order of 3. This equation could easily be used to derive any thermodynamic property like $\alpha(P, T)$ or $K_T(P, T)$. This equation has been used implicitly in various models of mantle convection (e.g., Glatzmaier, 1988; Bercovici *et al.*, 1989a,). An important consequence of this EoS assuming $q \sim 1$ is that αK_T is more or less constant and that

$$K_T \sim K_T^0 \left(\frac{\rho}{\rho_0} \right)^n, \quad \alpha \sim \alpha_0 \left(\frac{\rho}{\rho_0} \right)^{-n}. \quad (81)$$

In the mantle, the incompressibility increases and the thermal expansion decreases significantly with depth. The geophysical consequences are further discussed in 6.5.1.

Two other thermodynamic equalities that can also be straightforwardly deduced by chain rules of derivatives will be also used in the following. They are

$$\frac{C_P}{C_V} = \frac{K_S}{K_T} = 1 + \Gamma \alpha T, \quad (82)$$

These equalities relate the two heat capacities C_P and C_V of the energy equations (36). Since $\Gamma \sim 1$ and since $\alpha T \ll 1$, the two heat capacities are basically equal. It seems safer to assume that C_V is constant (the Dulong and Petit rule) and infer C_P from it, but most convection papers have done the reverse assumption of constant C_P .

The incompressibility K_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

$$K_S = \rho \left(v_p^2 - \frac{4}{3} v_s^2 \right), \quad (83)$$

where v_p and v_s are the p and s wave velocities. Through this important parameter a bridge can be built between geodynamics and seismology.

3.2 Rheology

In the section 2.3 on momentum conservation, no assumption is made on the rheology of the fluid, i.e., on the relation between the stress tensor and the flow itself. In contrast, the discussion of energy conservation, 2.4, is based on a strong assumption. It assumes that the pressure-related work is entirely recoverable, (30), and, as a consequence, the work of the deviatoric stresses ends up entirely as a dissipative term, 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, in the second case, bulk viscosity.

3.2.1 Elasticity

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

$$\underline{\epsilon}^e = \frac{1}{2}(\nabla \mathbf{u} + [\nabla \mathbf{u}]^t), \quad (84)$$

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

$$\sigma_{ij}^e = A_{ijkl}^e \epsilon_{kl}^e. \quad (85)$$

where $\underline{\mathbf{A}}^e$ is the rank four stiffness tensor. Since both the stress and the strain tensors are symmetric and because of the Maxwell thermodynamic relationships, $\partial^2 \mathcal{U} / \partial \epsilon_{ij} \partial \epsilon_{kl} = \partial^2 \mathcal{U} / \partial \epsilon_{ij} \partial \epsilon_{kl}$, the elastic tensor is invariant by permutations of i and j , k and l , ij and kl . This leaves in the most general case of anisotropy, 21 independent stiffness coefficients. 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_T (we assume here, that the deformation is isothermal) and the rigidity μ_R and the elastic behavior satisfies

$$\underline{\sigma}^e = K_T \text{tr}(\underline{\epsilon}^e) \underline{\mathbf{I}} + 2\mu_R \left(\underline{\epsilon}^e - \frac{1}{3} \text{tr}(\underline{\epsilon}^e) \underline{\mathbf{I}} \right), \quad (86)$$

where $\text{tr}(\underline{\underline{\epsilon}}^e) = \nabla \cdot \mathbf{u}$. (This expression assumes that the displacement vector is computed from an initial situation where the solid is perfectly stress-free, i.e., $\underline{\underline{\sigma}}^e = 0$ when $\underline{\underline{\epsilon}}^e = 0$. In practical problems, only incremental displacements with respect to an initial situation where internal pre-stress existed are known and $\underline{\underline{\sigma}}^e$ has to be understood as a variation of the stress tensor). The other elastic parameters that are often introduced, Poisson's ratio, Young's modulus (Young, 1773-1829), Lamé's parameters (Lamé, 1795-1870), are simple functions of incompressibility and rigidity. Since the term proportional to μ_R is traceless, equation (86) leads to, $\text{tr}(\underline{\underline{\sigma}}^e) = 3K_T \text{tr}(\underline{\underline{\epsilon}}^e)$, the rheology law can also be written in terms of compliance (i.e., getting $\underline{\underline{\epsilon}}^e$ as a function of $\underline{\underline{\sigma}}^e$),

$$\underline{\underline{\epsilon}}^e = \frac{1}{9K_T} \text{tr}(\underline{\underline{\sigma}}^e) \mathbf{I} + \frac{1}{2\mu_R} \left(\underline{\underline{\sigma}}^e - \frac{1}{3} \text{tr}(\underline{\underline{\sigma}}^e) \mathbf{I} \right). \quad (87)$$

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

$$\text{tr}(\underline{\underline{\sigma}}^e) = -3P. \quad (88)$$

The momentum equation (15) remains valid in a purely elastic solid (except that the advective transport is generally neglected, $D/Dt \sim \partial/\partial t$), but the discussion of energy conservation and thermodynamics is different for elastic and viscous bodies. The work of the elastic stresses is entirely recoverable: a deformed elastic body returns to its undeformed shape when the external forces are turned off. The work of the elastic stresses is $\delta W = V \underline{\underline{\sigma}}_e : d\underline{\underline{\epsilon}}^e$ instead of $\delta W = -PdV$ and, as a consequence, no dissipative entropy source related to deformation remains in the final temperature or entropy equation. For an elastic body, the temperature equations (36) and the entropy equation (37) hold when the $\underline{\underline{\tau}} : \nabla \mathbf{v}$ source term is removed.

3.2.2 Viscous Newtonian rheology

On a very long time scale, 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 $\underline{\underline{\tau}} : \nabla \mathbf{v} = 2 \underline{\underline{\tau}} : \underline{\underline{\dot{\epsilon}}}^v$ 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

$$\sigma_{ij}^v = -P\delta_{ij} + A_{ijkl}^v \dot{\epsilon}_{kl}^v, \quad (89)$$

δ_{ij} being the Kronecker symbol (Kronecker, 1823-1891). Except for the time derivative, the only formal difference between this expression and (85), if that

pressure exists in a motionless fluid but disappears in an undeformed elastic solid (with no pre-stress).

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

$$\underline{\boldsymbol{\sigma}}^v = (-P + \zeta \text{tr}(\underline{\dot{\boldsymbol{\epsilon}}})\mathbf{I} + 2\mu \left(\underline{\dot{\boldsymbol{\epsilon}}} - \frac{1}{3} \text{tr}(\underline{\dot{\boldsymbol{\epsilon}}})\mathbf{I} \right), \quad (90)$$

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

$$\underline{\dot{\boldsymbol{\epsilon}}} = \frac{1}{9\zeta}(3P + \text{tr}(\underline{\boldsymbol{\sigma}}^v))\mathbf{I} + \frac{1}{2\mu} \left(\underline{\boldsymbol{\sigma}}^v - \frac{1}{3} \text{tr}(\underline{\boldsymbol{\sigma}}^v)\mathbf{I} \right). \quad (91)$$

The two parameters μ 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 (Newton, 1642-1727). 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*).

Since $\text{tr}(\underline{\boldsymbol{\sigma}}^v)/3 = -P + \zeta \nabla \cdot \mathbf{v}$, 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, $2 \underline{\boldsymbol{\tau}} : \underline{\dot{\boldsymbol{\epsilon}}}$, during isotropic compaction could be dissipated in the form of the heat source $2\zeta(\nabla \cdot \mathbf{v})^2$. A density-independent bulk viscosity allows an infinite compression under a finite isotropic stress. The elusive bulk viscosity parameter ζ is generally only introduced to be immediately omitted and we will do the same. However, using (90) with $\zeta = 0$ but $\nabla \cdot \mathbf{v} \neq 0$ does not seem valid since it would remove all resistance to compaction. We will see that considering $\zeta = 0$ in (90) is formally correct although the real physical explanation is more complex: elastic stresses must be present to provide a resistance to isotropic viscous compaction. The bulk viscosity, or some equivalent concept, is however necessary to handle two phase compaction problems (*McKenzie, 1984; Bercovici et al., 2001a*)(see section 5.2).

3.2.3 Maxwellian Visco-elasticity

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 (91) with the time derivative of the elastic compliance equation, (87) and in the case of an infinite bulk viscosity ζ , we get

$$\underline{\dot{\boldsymbol{\epsilon}}} = \frac{1}{9K_T} \text{tr}(\underline{\dot{\boldsymbol{\sigma}}})\mathbf{I} + \frac{1}{2\mu} \left(\underline{\boldsymbol{\sigma}} - \frac{1}{3} \text{tr}(\underline{\boldsymbol{\sigma}})\mathbf{I} \right) + \frac{1}{2\mu_R} \left(\underline{\dot{\boldsymbol{\sigma}}} - \frac{1}{3} \text{tr}(\underline{\dot{\boldsymbol{\sigma}}})\mathbf{I} \right), \quad (92)$$

where $\underline{\sigma} = \underline{\sigma}^v = \underline{\sigma}^e$ and $\underline{\epsilon} = \underline{\epsilon}^v + \underline{\epsilon}^e$. 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 (92). First, we can consider the case where stress and strain are simple time-dependent sinusoidal functions with frequency ω (i.e., $\underline{\sigma} = \underline{\sigma}_0 \exp(i\omega t)$ and $\underline{\epsilon} = \underline{\epsilon}_0 \exp(i\omega t)$). The solution to this problem can then be used to solve other time-dependent problems by Fourier or Laplace transforms. The equation (92) becomes

$$\underline{\epsilon}_0 = \frac{1}{9K_T} \text{tr}(\underline{\sigma}_0) \mathbf{I} + \frac{1}{2\mu_R} \left(1 - \frac{i}{\omega\tau}\right) \left(\underline{\sigma}_0 - \frac{1}{3} \text{tr}(\underline{\sigma}_0) \mathbf{I}\right), \quad (93)$$

where $\tau = \mu/\mu_R$ is the Maxwell time, (1). This equation can be compared to (87), and shows that the solution of a visco-elastic 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 1D Maxwellian body (only σ_{zz} and ϵ_{zz} are non zero), 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 $\epsilon_{zz} = \epsilon_0 H(t)$. The solutions are, for $t \geq 0$,

$$\epsilon_0 = \frac{1}{3k\mu_R} \sigma_0 + \frac{1}{3\mu} \sigma_0 t, \quad (94)$$

and

$$\tau_0 = 3k\mu_R \exp\left(-k\frac{t}{\tau}\right) \epsilon_0, \quad (95)$$

respectively, with $k = 3K_T/(3K_T + \mu_R)$. In the first case, a finite elastic deformation occurs instantaneously, followed by a perfectly Newtonian flow. In the second case, the initial deformation is immediately resisted by the elastic stresses, that 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 non-deviatoric 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 mid-mantle, $K_T \sim 2\mu_R \sim 200$ GPa).

From equation (92), 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 4 and 6), the appropriate Maxwell times are in the range 30 yr-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

$$\dot{\underline{\epsilon}} = \frac{1}{9K_T} \text{tr}(\dot{\underline{\sigma}}) \underline{\mathbf{I}} + \frac{1}{2\mu} \left(\underline{\sigma} - \frac{1}{3} \text{tr}(\underline{\sigma}) \underline{\mathbf{I}} \right). \quad (96)$$

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

$$\underline{\sigma} = -P \underline{\mathbf{I}} + 2\mu \left(\dot{\underline{\epsilon}} - \frac{\dot{P}}{3K_T} \underline{\mathbf{I}} \right). \quad (97)$$

This equation is intrinsically a visco-elastic equation, that can be replaced by a purely viscous equation plus an EoS

$$\underline{\sigma} = -P \underline{\mathbf{I}} + 2\mu \left(\dot{\underline{\epsilon}} - \frac{1}{3} \text{tr}(\dot{\underline{\epsilon}}) \underline{\mathbf{I}} \right), \quad (98)$$

$$\text{tr}(\underline{\epsilon}) = \frac{P}{K_T}. \quad (99)$$

The equation (98) is therefore the appropriate limit of the equation (90) for slow deformation, when $\zeta = +\infty$ and when isotropic compaction is resisted by the elastic stresses.

The use of a Maxwell visco-elastic body to represent the mantle rheology on short time scale 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 be used for the Earth but the data that could support or dismiss them is scarce (*Yuen et al.*, 1986).

3.2.4 Non-linear 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 the shear viscosity could be any function of the invariants of the strain rate tensor. There are three invariants of the strain rate tensor; its trace (but $\text{tr}(\underline{\tau}) = 0$, in the absence of bulk viscosity), its determinant

and the second invariant $I_2 = \sqrt{\dot{\underline{\epsilon}} : \dot{\underline{\epsilon}}}$ (where as in (28), the double dots denote tensor contraction). The viscosity could therefore be a function of $\det(\dot{\underline{\epsilon}})$ and I_2 .

The main mechanisms of solid state deformation pertinent for mantle conditions (the lithosphere with brittle and plastic deformations is excluded) are 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 a Boltzmann statistics (Boltzmann, 1844-1906),

$$C \propto \exp\left(-\frac{PV}{RT}\right), \quad (100)$$

(V is the atomic volume, 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 ($\tau V \ll RT$). This induces the flux of atoms (number of atoms per unit surface and unit time)

$$J \propto D \frac{C_0}{d} \frac{\tau}{RT}, \quad (101)$$

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_h (with grain boundary diffusion D_h) along the grains interfaces through a surface hd (h being the thickness of the grain boundary), according to

$$\frac{d^2 \delta d}{V} \sim J_D d^2 \delta t, \quad \text{or} \quad \frac{d^2 \delta d}{V} \sim J_h h d \delta t. \quad (102)$$

As $\dot{\epsilon} = (\delta d/\delta t)/d$, the previous equations lead to the stress-strain rate relationship

$$\dot{\underline{\epsilon}} \propto \frac{V}{d^2 RT} \left(D_v + D_h \frac{h}{d} \right) \underline{\tau}. \quad (103)$$

This 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) and $\eta \propto d^3$ for Coble grain-boundary creep (Coble, 1928-1992). The viscosity is also very strongly

T -dependent not so much because of the explicit factor T in (103), but because diffusion is a thermally activated process, $D \propto \exp(E_{dif}/RT)$, where E_{dif} is an activation enthalpy of diffusion.

In the case of dislocation creep, lines or planar imperfections are present in the crystalline lattice and macroscopic deformation occurs by 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 often found to vary as $1/I_2$. Therefore, instead of a diffusion creep with a viscosity in d^n the resulting rheology is rather in I_2^{-n} and is also thermally activated with an activation energy E_{dis} . Dislocation creep leads to a non-linear regime where the equivalent viscosity varies with the second invariant with a power $-n$, where n is typically of order 2,

$$\dot{\underline{\epsilon}} \sim I_2^n \exp(E_{dis}/RT) \underline{\tau}, \quad (104)$$

(this relationship is often written, in short, $\dot{\underline{\epsilon}} \sim \underline{\tau}^m$ with a stress exponent m of order 3 but $\underline{\tau}^m$ really means $I_2^{m-1} \underline{\tau}$).

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 non-linear (stress dependent), viscosities are also strongly dependent upon temperature, pressure, melt content, water content, mineralogical phase and oxygen fugacity (e.g., *Hirth and Kohlstedt (1996)*). In section 6.3 we will further discuss the rheological mechanisms appropriate for the Earth.

4 Physics of convection

4.1 Convection experiments

The complex and very general system of equations that we have discussed in sections 2 and 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 Chapter 5 and Chapter 8 of this Treatise). It can also be induced by compositional density anomalies (some examples are discussed in Chapter 5 and Chapter 12 of this Treatise). However, a fundamental cause of motion is due to the interplay between density 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 also Chapter 4). 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)*).

Using the shadowgraph method, it is easy to perform a few experiments. When a liquid is heated or cooled from its side (as in a glass of cold water with a hand holding one side of the glass), motion always occurs. The body forces in the fluid ρg are different on the sides of the fluid because of the density-temperature relationship. Motion therefore starts in the direction opposite to g on the heated side of the tank with a downwelling along the coldest side.

If a homogeneous fluid is heated from the top or cooled from the bottom, no motion is induced. Since temperature variations are along gravity and with no horizontal components of the gravitational forces, the symmetry of this situation seems to justify that the fluid remains motionless. What is, on the contrary, unexpected (or should be unexpected from symmetry arguments), is that, when the fluid is heated from the bottom or cooled from the top, a spontaneous motion can occur if a strong enough temperature gradient is imposed. This is what physicists call a spontaneous symmetry breaking: the fluid properties and the temperature forcing are invariant with respect to horizontal translations, but this invariance (that physicists call symmetry) is surprisingly lost when convection starts.

4.2 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} , it has an average density ρ and thermal expansivity α . The volume Ω , because of its anomalous temperature, experiences an Archimedian force, or buoyancy (Archimede around 287-212 BC) given by

$$\mathbf{F} = -c_1 a^3 \rho \alpha \Delta T \mathbf{g}, \quad (105)$$

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

$$\mathbf{v}_s = -c_1 c_2 \frac{a^2 \rho \alpha \Delta T \mathbf{g}}{\mu}, \quad (106)$$

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

During its motion, the volume Ω exchanges heat by diffusion with the rest of the fluid and the diffusion equation (42) tells us that a time of order

$$t_d = c_3 \frac{\rho C_p a^2}{k}. \quad (107)$$

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 $l \gg a$, i.e., when the fluid volume can be displaced by several times its size, motion will be possible. On the contrary when $l \ll a$, 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 above expressions, depends on only one quantity, the Rayleigh number Ra

$$Ra = \frac{\rho^2 \alpha \Delta T g a^3 C_P}{\mu k} = \frac{\alpha \Delta T g a^3}{\kappa \nu}, \quad (108)$$

in terms of which motion occurs when $Ra \gg 1$ (assuming that $c_1 c_2 c_3 \sim 1$). The Rayleigh number compares the driving mechanism (e.g., the Archimedian buoyancy) to the two resistive mechanisms, the diffusion of heat, represented by κ (see (41)), and the diffusion of momentum, represented by ν (see (47)).

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 4.5. 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.

4.3 Two simple solutions

4.3.1 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,

$$0 = -\nabla P + \rho \mathbf{g}, \quad (109)$$

and the energy equation (36) when the temperature is diffusive (using the Fourier law (40)),

$$\nabla \cdot (k \nabla T) + \rho H = 0. \quad (110)$$

Solving analytically for the hydrostatic pressure and the diffusive temperature is trivial when H , k and ρ are uniform. For example, choosing z positive downward, we get

$$P = \rho g z, \quad T = T_0 + \Delta T \frac{z}{h} + \frac{1}{2} \rho H z (h - z), \quad (111)$$

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 remains feasible, but could be quite cumbersome if one introduces a realistic EoS and computes gravity in agreement with the density distribution using the Poisson equation (48). In section 4.5 we will understand why the system does not necessarily choose this solution.

4.3.2 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 (36) can also be written

$$\begin{aligned} \rho C_V T \left(\frac{D \ln T}{Dt} - \Gamma \frac{D \ln \rho}{Dt} \right) &= -\nabla \cdot \mathbf{q} + \boldsymbol{\tau} : \nabla \mathbf{v} + \rho H, \\ \rho C_P T \left(\frac{D \ln T}{Dt} - \frac{\alpha}{\rho C_P} \frac{DP}{Dt} \right) &= -\nabla \cdot \mathbf{q} + \boldsymbol{\tau} : \nabla \mathbf{v} + \rho H. \end{aligned} \quad (112)$$

The right sides of these equations were previously shown to be equal to $\rho T D S / Dt$ in (37). If we decrease the viscosity in a fluid in a thought experiment, the convective velocity will increase. With a large velocity, the advection terms, $\mathbf{v} \cdot \nabla T$, $\mathbf{v} \cdot \nabla \ln \rho$ and $\mathbf{v} \cdot \nabla P$, will become much larger than the time dependent, diffusion and radioactive production terms. A large velocity also implies a well-mixed fluid in which the lateral temperature variations and thus the lateral density variations, will be small. The viscous stresses, related to these density differences, will therefore decrease. As a consequence, when convection is vigorous enough, the fluid

should evolve toward a situation where

$$\begin{aligned} (\nabla \ln T)_S - \Gamma (\nabla \ln \rho)_S &= 0, \\ (\nabla \ln T)_S - \frac{\alpha}{\rho C_P} (\nabla P)_S &= 0. \end{aligned} \quad (113)$$

Since such equations imply that the entropy is exactly conserved, $DS/Dt = 0$, this equilibrium is called the adiabatic equilibrium. We added a subscript $[]_S$ to remember this conservation.

Notice that, since the Grüneisen parameter is only density-dependent, (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 (78)), the first of the equations (113) implies

$$T = T_0 \left(\frac{\rho}{\rho_0} \right)^{\Gamma_0}, \quad (114)$$

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 K to 2230 K) from the asthenosphere ($\rho_0 \sim 3200 \text{ kg m}^{-3}$) to the CMB ($\rho \sim 5500 \text{ kg m}^{-3}$), if we assume $\Gamma_0 = 1$.

4.3.3 Stability of the adiabatic gradient

In section 4.2 we discussed a laboratory convection experiment but we totally neglected the adiabatic temperature variations. When a fluid is compressed, it heats up and it cools down when decompressed. This is the very same physics than explains why the atmospheric temperature decreases with altitude. Of course this adiabatic effects are vanishingly small in a tank experiment.

In fact we can redo the same estimate as in section 4.2 but account for the adiabatic gradient with only a little change. 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 $a(dT/dz)_S$. However the surrounding fluid will be at the temperature $a(dT/dz)$ where dT/dz is just the temperature gradient, not necessarily adiabatic, of the fluid at rest. We can define ΔT_{na} as the non-adiabatic temperature $\Delta T_{na} = a(dT/dz - (dT/dz)_S)$. The parcel being suddenly warmer or colder than the surroundings will rise or sink with a Stokes velocity that, rather than (106) will be of order

$$\mathbf{v}_s = -c_1 c_2 \frac{a^2 \rho \alpha \Delta T_{na} \mathbf{g}}{\mu}. \quad (115)$$

Since we chose $dz \geq 0$ along \mathbf{g} , the adiabatic gradient is positive and the fluid parcel will be locally unstable if the gradient in the surrounding fluid is larger (superadiabatic) than the adiabatic gradient. On the contrary a sub-adiabatic gradient

will be stable with respect to convection. It is therefore not the total temperature difference between the top and bottom of the fluid that drives motion, but only its non-adiabatic part.

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

$$Ra = \frac{\alpha \Delta T_{na} g_0 a^3}{\kappa \nu}. \quad (116)$$

This number is based on the non-adiabatic 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. This superadiabaticity being the driving source of convection implies that too large a superadiabaticity would provide negative feedback and rapidly reduce its excess temperature returning to adiabaticity. This mechanism suggests that an adiabatic reference background should not be such a bad assumption for 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 K) 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.

4.4 Approximate equations

4.4.1 Depth dependent reference profiles

Since we guessed that the thermodynamic state should not be too far from an hydrostatic adiabat, we are going to choose this state as a reference and rewrite the equations of fluid dynamics in term of perturbations. We denote all the reference variables with an overbar. We choose a reference hydrostatic pressure

$$\nabla \bar{P} = \bar{\rho} \bar{\mathbf{g}}, \quad (117)$$

and adiabatic temperature and densities according to (113)

$$\begin{aligned} \nabla \bar{T} &= \frac{\bar{\alpha} \bar{\mathbf{g}}}{\bar{C}_p} \bar{T} \\ \nabla \bar{\rho} &= \frac{\bar{\alpha} \bar{\mathbf{g}}}{\bar{C}_p \Gamma} \bar{\rho}, \end{aligned} \quad (118)$$

where all the parameters are computed along the reference geotherm and where $\bar{\mathbf{g}}$ has been solved using Poisson's equation (48) with the reference density $\bar{\rho}$.

The reference parameters are depth-dependent and usually, even for a simple EoS cannot be analytically obtained. They can however be computed numerically from (113) 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 profile can be obtained by quadratures.

4.4.2 Nondimensionalization

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 appears in the equations. A necessary starting point is therefore to rephrase any fluid dynamics problem involving N dimensional parameters in term of M dimensionless quantities ($N - M \leq 0$ is the number of independent physical dimensions of the problem).

It is not very convenient to perform the nondimensionalization of the convection equations using the variable reference profiles. We must therefore introduce constant parameters, with indices $[\]_0$, corresponding to some typical or mantle-averaged values of the depth dependent reference values. We introduce for example α_0 , ρ_0 , g_0 or C_p^0 .

The adiabatic equations (118) impose the natural scale for temperature variation, $\bar{C}_p/(\bar{\alpha}||\bar{\mathbf{g}}||)$. This scale varies with depth but should not be too different from the value computed with the constant parameters (with indices 0). We therefore introduce a nondimensional number, the dissipation number D_0 ,

$$D_0 = \frac{\alpha_0 g_0 a}{C_p^0}, \quad (119)$$

that compares the natural scale of temperature variations 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 hydrostatic and adiabatic reference profiles satisfy approximately

$$\frac{d\bar{T}}{dz} \sim D_0 \frac{\bar{T}}{a}, \quad \text{and} \quad \frac{d\bar{\rho}}{dz} \sim \frac{D_0 \bar{\rho}}{\Gamma_0 a}. \quad (120)$$

A zero dissipation number leads to uniform reference temperature and pressure, as the adiabatic compression effects are not large enough to affect these quantities.

From top to bottom, the reference temperature increases adiabatically by ΔT_S while a total temperature jump $\Delta T_S + \Delta T_{na}$ is imposed. Only the excess non adiabatic temperature ΔT_{na} is really useful to drive convection. The term driving

convective instability is thus the dimensionless quantity ϵ ,

$$\epsilon = \alpha_0 \Delta T_{na}, \quad (121)$$

which is always a small number (all the necessary numerical values are listed in Table 1).

All these preliminaries have been somewhat lengthy but we are now ready to nondimensionalize the various equations. Then, we will get the approximate equations by simply taking into account that $\epsilon \ll 1$ (anelastic equations) or more crudely that both $\epsilon \ll 1$ and $D_0 \ll 1$ (Boussinesq equations), (*Jarvis and Mckenzie*, 1980). The mathematical formulation is heavy because of the number of symbols with or without overbars, tildes or indices, but is straightforward. The reader may jump directly to the results 4.4.3.

To nondimensionalize the equations we can use the quantities a and ΔT_{na} ; we also need a characteristic velocity and pressure. Following our discussion of the basic force balance in section 4.2, we use a typical Stokes velocity, $V_0 = a^2 g_0 \rho_0 \alpha_0 \Delta T_{na} / \mu_0$, time a/V_0 , and pressure $P_0 = \mu_0 V_0 / a$. Using the definitions of D_0 , (119), Γ_0 , (77) and ϵ , (121), we perform the following change of variables

$$\begin{aligned} \mathbf{v} &= \epsilon \frac{a K_T^0 D_0 C_P^0}{\mu_0 \Gamma_0 C_V^0} \tilde{\mathbf{v}}, \\ P &= \bar{P} + \epsilon K_T^0 \frac{D_0 C_P^0}{\Gamma_0 C_V^0} \tilde{P}, \\ \alpha_0 T &= \alpha_0 \bar{T} + \epsilon \tilde{T}, \\ \nabla &= \frac{1}{a} \tilde{\nabla}, \\ \frac{\partial}{\partial t} &= \epsilon \frac{K_T^0 D_0 C_P^0}{\mu_0 \Gamma_0 C_V^0} \frac{\partial}{\partial \tilde{t}} \end{aligned} \quad (122)$$

The EoS (76) can then be expanded at first order for a thermodynamic state close to the hydrostatic adiabatic reference

$$\rho = \bar{\rho} + \frac{\bar{\rho}}{\bar{K}_T} (P - \bar{P}) - \bar{\alpha} \bar{\rho} (T - \bar{T}). \quad (123)$$

After nondimensionalization, the EoS becomes

$$\rho = \bar{\rho} \left(1 + \epsilon \left(\frac{K_T^0 D_0 C_P^0}{\bar{K}_T \Gamma_0 C_V^0} \tilde{P} - \frac{\bar{\alpha}}{\alpha_0} \tilde{T} \right) \right). \quad (124)$$

As $C_P^0 \sim C_V^0$, $K_T \sim K_T^0$, $\alpha \sim \alpha_0$, $\tilde{P} \sim \tilde{T} \sim 1$ and $\epsilon \ll 1$ it shows that the density remains close to the reference profile within terms of order ϵ , $\rho = \bar{\rho}(1 + O(\epsilon))$.

4.4.3 Anelastic approximation

To approximate the equations of convection we perform the change of variables, (122), and use the fact that ϵ is small. The choice of the reference state leads to the cancelation of the terms $O(1)$ of the mass, momentum and energy equations, (9), (16) and (36).

At order $O(\epsilon)$, the first terms that remain in each equation are

$$\begin{aligned} \tilde{\nabla} \cdot \left(\frac{\bar{\rho}}{\rho_0} \tilde{\mathbf{v}} \right) &= 0. \\ \frac{Ra}{Pr} \frac{D\tilde{\mathbf{v}}}{D\tilde{t}} &= -\tilde{\nabla} \tilde{P} + \tilde{\nabla} \cdot \tilde{\boldsymbol{\tau}} + \frac{\bar{\rho}}{\rho_0} \frac{\bar{\mathbf{g}}}{g_0} \frac{K_T^0}{\bar{K}_T} \frac{D_0}{\Gamma_0} \frac{C_P^0}{C_V^0} \tilde{P} - \frac{\bar{\rho}}{\rho_0} \frac{\bar{\mathbf{g}}}{g_0} \frac{\bar{\alpha}}{\alpha_0} \tilde{T}. \\ \frac{\bar{\rho}}{\rho_0} \frac{\bar{C}_P}{C_P^0} \frac{D\tilde{T}}{D\tilde{t}} &= \frac{1}{Ra} \tilde{\nabla} \cdot \left[\frac{\bar{k}}{k_0} \tilde{\nabla} \left(\frac{\tilde{T}}{\Delta T_{na}} + \tilde{T} \right) \right] + \frac{\bar{\alpha}}{\alpha_0} \frac{\bar{\rho}}{\rho_0} \frac{\bar{g}}{g_0} D_0 \tilde{T} \tilde{v}_g \\ &\quad + \frac{\bar{\rho}}{\rho_0} \frac{1}{Ra} \frac{\rho_0 Ha^2}{k_0 \Delta T} + D_0 \tilde{\boldsymbol{\tau}} : \nabla \tilde{\mathbf{v}} \end{aligned} \quad (125)$$

In the last equation, \tilde{v}_g is the component of the velocity field is along the radial reference gravity $\bar{\mathbf{g}} \cdot \tilde{\mathbf{v}} = \bar{g} \tilde{v}_g$. This new set of equations constitutes the equation of fluid dynamics in the anelastic approximation. The term ‘‘anelastic’’ comes from the fact that the propagation of sound waves is impossible since the term in $\partial\rho/\partial t$ is neglected.

In equations (125), we introduced the Rayleigh, Ra , and Prandtl, Pr , numbers

$$Ra = \frac{\alpha_0 \Delta T_{na} \rho_0^2 g_0 a^3 C_P^0}{\mu_0 k_0} = \frac{\alpha_0 \Delta T_{na} g_0 a^3}{\nu_0 \kappa_0}, \quad (126)$$

$$Pr = \frac{\mu_0 C_P^0}{k_0} = \frac{\nu_0}{\kappa_0}. \quad (127)$$

The physical meaning of the Rayleigh number as a measure of convective vigor has already been discussed. The Prandtl number (Prandtl, 1975-1953) compares the two diffusive processes: namely the diffusion of momentum and heat.

In the momentum equation, the nondimensionalized stress tensor is $\tilde{\boldsymbol{\tau}} = \boldsymbol{\tau} a / (\mu_0 V_0)$ which, in the Newtonian case (without bulk viscosity), becomes

$$\tilde{\boldsymbol{\tau}} = \frac{\mu}{\mu_0} (\tilde{\nabla} \tilde{\mathbf{v}} + [\tilde{\nabla} \tilde{\mathbf{v}}]^t) - \frac{2}{3} \frac{\mu}{\mu_0} \nabla \cdot \tilde{\mathbf{v}}. \quad (128)$$

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 2.5.2). To account for this term we should have added on the right side of second equation

of (125) a term $-\bar{\rho}/\rho_0 \tilde{\nabla} \tilde{\psi}$, where the perturbed gravitational potential due to the departure of the density from the reference profile, satisfies Poisson's equation

$$\tilde{\nabla}^2 \tilde{\psi} = 3 \frac{\bar{\rho}}{\langle \rho \rangle} \left(\frac{K_T^0 D_0 C_P^0}{\bar{K}_T \Gamma_0 C_V^0} \tilde{P} - \frac{\bar{\alpha}}{\alpha_0} \tilde{T} \right), \quad (129)$$

where $\langle \rho \rangle$ is the average density of the Earth.

4.4.4 Dimensionless numbers

The ratio Ra/Pr is also called the Grashof number $Gr = \alpha g \Delta T_{na} a^3 / \nu^2$ (Grashof, 1826-1893). This number can also be written as $V_0 a / \nu$ 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 is a parameter imposed by a boundary condition, using it to perform the nondimensionalization and speaking in terms of Reynolds number would be more natural than using the Grashof number. If a thermal structure is imposed by a velocity boundary condition (for example by the thickening of the oceanic lithosphere with age), it would seem natural to introduce a Péclet number $V a / \kappa$ (Péclet, 1973-1857), which is nothing else than the Rayleigh number of the flow if the velocity is imposed by the internal dynamics.

The Rayleigh and Prandtl numbers can be estimated in different ways. In fact the only difficult parameter 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 geographical variability. The mantle viscosity and its depth dependence can be constrained by post-glacial 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, between 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 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 e.g., *Peltier* (1989); *Sabadini and Yuen* (1989); *Lambeck and Johnston* (1998) or *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 (125) can safely be set to zero (see numerical values in Table 1).

We can also introduce the Rayleigh and Prandtl number in the EoS (124) and define a new number, M ,

$$\rho = \bar{\rho} \left(1 + \frac{K_T^0}{\bar{K}_T} Ra Pr M^2 \tilde{P} - \frac{\bar{\alpha}}{\alpha_0} \tilde{T} \right). \quad (130)$$

This number is the Mach number $M = v_D/v_M$, ratio of the velocity of thermal diffusion, κ_0/a , to the velocity $v_M = \sqrt{K_T/\rho}$ (Mach, 1838-1916). This last velocity is very close to the bulk velocity $v_\phi = \sqrt{K_S/\rho}$ (see (83)). The anelastic approximation is sometimes called small Mach number approximation. The Earth's Mach number is indeed of order 10^{-15} since thermal diffusion is much slower than the sound speed. However the anelastic approximation really requires a small $RaPrM^2$ and this quantity is just $\epsilon D_0 C_P^0/\Gamma_0 C_V^0$, i.e., of order $\epsilon = 10^{-2}$. 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).

4.4.5 Boussinesq approximation

As was expected from section 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.

In the approximation $D_0 = 0$, the so-called Boussinesq approximation (Boussinesq, 1842-1929), the reference density and temperature become constants according to (120) (when the reference temperature is uniform we can also choose $C_P = C_V = \bar{C}_P = \bar{C}_V$). The non adiabatic 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 (124) indicates that the density is only a function of temperature and the fluid dynamics equations become

$$\begin{aligned} \tilde{\nabla} \cdot \tilde{\mathbf{v}} &= 0, \\ \frac{Ra}{Pr} \frac{D\tilde{\mathbf{v}}}{D\tilde{t}} &= -\tilde{\nabla}\tilde{P} + \tilde{\nabla} \cdot \left(\frac{\mu}{\mu_0} (\tilde{\nabla}\tilde{\mathbf{v}} + [\tilde{\nabla}\tilde{\mathbf{v}}]^t) \right) - \frac{\bar{\mathbf{g}}}{g_0} \frac{\bar{\alpha}}{\alpha_0} \tilde{T}, \\ \frac{D\tilde{T}}{D\tilde{t}} &= \frac{1}{Ra} \tilde{\nabla} \cdot \left[\frac{\bar{k}}{k_0} \tilde{\nabla}\tilde{T} \right] + \frac{1}{Ra} \frac{\rho_0 Ha^2}{k_0 \Delta T}. \end{aligned} \quad (131)$$

Here again as in the mantle, $Gr = Ra/Pr \ll 1$, the inertia in the momentum equation (131) can be neglected. The self-gravitational term $-\bar{\rho}/\rho_0 \tilde{\nabla}\tilde{\psi}$ should be added to the momentum equation (second equation of (131)) for large scale simulations, the gravitational potential being solution of (129) where only the thermal part of the density variation needs to be taken into account.

The physical behavior of a large Rayleigh number is obvious in (131). When

$Ra \rightarrow \infty$, the temperature becomes a purely advected and conserved quantity, $D\tilde{T}/D\tilde{t} = 0$.

4.4.6 Internal heating

In the nondimensionalization, we assumed that the non-adiabatic 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 anymore a free parameter but must result from the properties of the flow itself. In the nondimensionalization, we can replace ΔT_{na} by $\rho_0 H a^2 / k_0$ in such a way that the radioactive heat source of the anelastic or Boussinesq energy equations, (125) or (131), are simply $1/Ra$. This choice requires the introduction of a somewhat different Rayleigh number (internally heated Rayleigh number)

$$Ra_H = \frac{\alpha_0 H \rho_0^3 g_0 a^5 C_P^0}{\mu_0 k_0^2}. \quad (132)$$

4.4.7 Alternative forms

Using the formulation of internal energy in terms of C_V , we would have reached the equivalent anelastic energy equation

$$\begin{aligned} \frac{\bar{\rho}}{\rho_0} \frac{\bar{C}_V}{C_P^0} \frac{D\tilde{T}}{D\tilde{t}} &= \frac{1}{Ra} \tilde{\nabla} \cdot \left[\frac{\bar{k}}{k_0} \tilde{\nabla} \left(\frac{\bar{T}}{\Delta T} + \tilde{T} \right) \right] + \\ &+ \frac{\bar{\alpha}}{\alpha_0} \frac{\bar{\rho}}{\rho_0} \frac{\bar{g}}{g_0} \frac{\bar{K}_T}{\bar{K}_S} D_0 \tilde{T} \tilde{v}_z + \frac{\bar{\rho}}{\rho_0} \frac{1}{Ra} \frac{\rho_0 H a^2}{k_0 \Delta T} + D_0 \tilde{\boldsymbol{\epsilon}} : \tilde{\nabla} \tilde{\mathbf{v}}, \end{aligned} \quad (133)$$

where the reference incompressibility \bar{K}_S is the incompressibility measured along the reference adiabatic profile

$$\bar{K}_S = \bar{\rho} \left(\frac{\partial P}{\partial \rho} \right)_S = \bar{\rho} \frac{\|\nabla \bar{P}\|}{\|\nabla \bar{\rho}\|} = \frac{\bar{\rho}^2 \bar{g}}{\|\nabla \bar{\rho}\|}. \quad (134)$$

This relationship is given here as a definition of \bar{K}_S , and this incompressibility is built from a theoretical hydrostatic and adiabatic model. However if the real Earth is indeed hydrostatic and adiabatic, then this relationship (134) connects a seismological observation $K_S(r)/\rho(r)$ to the density gradient of the real Earth $\rho(r)g(r)/\|\nabla \rho(r)\|$. This is the important Bullen hypothesis (*Bullen, 1940*) used to build the reference density of the Earth (e.g., *Dziewonski and Anderson, 1981*).

Table 1: Typical parameter values for numerical models of mantle convection. 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 non-adiabatic 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} !

		Mantle	Core	
size	a	$3 \cdot 10^6$	$3 \cdot 10^6$	m
dyn. viscosity	μ_0	10^{21}	10^{-3}	Pa s
heat capacity	C_P^0 ou C_V^0	1000	700	$\text{J K}^{-1} \text{kg}^{-1}$
density	ρ_0	4000	11000	kg m^{-3}
heat cond.	k_0	3	50	$\text{W m}^{-1} \text{K}^{-1}$
expansivity	α_0	$2 \cdot 10^{-5}$	10^{-5}	K^{-1}
temperature excess	ΔT_{na}	1500	1?	K
radioactivity prod.	H	$7 \cdot 10^{-11}$	0?	W kg^{-1}
gravity	g_0	9.8	5	m s^{-2}
incompressibility	K_T^0	10^{11}	10^{12}	Pa
kin. viscosity	$\nu = \mu_0/\rho_0$	$2.5 \cdot 10^{17}$	$9.1 \cdot 10^{-8}$	$\text{m}^2 \text{s}^{-1}$
thermal diff.	$\kappa = k_0/(\rho_0 C_P^0)$	$7.5 \cdot 10^{-7}$	$6.5 \cdot 10^{-6}$	$\text{m}^2 \text{S}^{-1}$
	ϵ	$3.0 \cdot 10^{-2}$	1.0^{-5}	
dissip. number	D_0	0.59	0.21	
Grüneisen par.	Γ	0.50	1.3	
Rayleigh	Ra	$4.2 \cdot 10^7$	$2.2 \cdot 10^{27}?$	
Intern. Rayleigh	Ra_H	$2.4 \cdot 10^{10}$	0?	
Prandlt	Pr	$3.3 \cdot 10^{23}$	$1.4 \cdot 10^{-2}$	
Reynolds	$Re = Ra/Pr$	$1.3 \cdot 10^{-15}$	$1.6 \cdot 10^{29}$	
Mach	M	$5.0 \cdot 10^{-15}$	$2.3 \cdot 10^{-16}$	
	$RaPrM^2$	$3.5 \cdot 10^{-2}$	$1.6 \cdot 10^{-6}$	

4.4.8 Change of nondimensionalization

We thought after the discussion of the basic balance of convection in section 4.2, that it was logical to use a Stokes velocity V_0 , to normalize the equations. We have therefore introduced a velocity and a time of order of 300 m yr^{-1} and of $a/V_0=10000 \text{ yr}$ (see Table 1). This is certainly very fast or short compared with geological scales. Most physical and geophysical textbooks (e.g., *Schubert et al.*, 2001) use instead a diffusive time $t_D = \rho C_P a^2 / k_0$ and velocity a/t_D . This is perfectly valid but Table 1, shows that the diffusive time and velocity amount to $t_D = 400 \text{ byr}$ and $V_D = 7 \cdot 10^{-6} \text{ m yr}^{-1}$. This is even less Earth-like which means that the nondimensionalized values using a diffusive time may have rather unequal orders of magnitude. Using this approach, we would have obtained the anelastic equations

$$\begin{aligned}
\tilde{\nabla} \cdot \left(\frac{\bar{\rho}}{\rho_0} \tilde{\mathbf{v}} \right) &= 0, \\
\frac{1}{Pr} \frac{D\tilde{\mathbf{v}}}{D\tilde{t}} &= -\tilde{\nabla} \tilde{P} + \tilde{\nabla} \cdot \tilde{\boldsymbol{\tau}} + \frac{\bar{\rho}}{\rho_0} \frac{\bar{\mathbf{g}}}{g_0} \frac{K_T^0}{\bar{K}_T} \frac{D_0}{\Gamma_0} \frac{C_P^0}{C_V^0} \tilde{P} - \frac{\bar{\rho}}{\rho_0} \frac{\bar{\mathbf{g}}}{g_0} \frac{\bar{\alpha}}{\alpha_0} Ra \tilde{T} \\
\frac{\bar{\rho}}{\rho_0} \frac{\bar{C}_P}{C_P^0} \frac{D\tilde{T}}{D\tilde{t}} &= \tilde{\nabla} \cdot \left[\frac{\bar{k}}{k_0} \tilde{\nabla} \left(\frac{\bar{T}}{\Delta T} + \tilde{T} \right) \right] + \\
&\quad \frac{\bar{\alpha}}{\alpha_0} \frac{\bar{\rho}}{\rho_0} \frac{\bar{g}}{g_0} D_0 \tilde{T} \tilde{v}_g + \frac{\bar{\rho}}{\rho_0} \frac{\rho_0 H a^2}{k_0 \Delta T} + \frac{D_0}{Ra} \tilde{\boldsymbol{\tau}} : \tilde{\nabla} \tilde{\mathbf{v}}
\end{aligned} \tag{135}$$

and the Boussinesq equations,

$$\begin{aligned}
\tilde{\nabla} \cdot \tilde{\mathbf{v}} &= 0, \\
\frac{1}{Pr} \frac{D\tilde{\mathbf{v}}}{D\tilde{t}} &= -\tilde{\nabla} \tilde{P} + \tilde{\nabla} \cdot \left(\frac{\mu}{\mu_0} (\tilde{\nabla} \tilde{\mathbf{v}} + [\tilde{\nabla} \tilde{\mathbf{v}}]^t) \right) - \frac{\bar{\mathbf{g}}}{g_0} \frac{\bar{\alpha}}{\alpha_0} Ra \tilde{T}, \\
\frac{D\tilde{T}}{D\tilde{t}} &= \tilde{\nabla} \cdot \left[\frac{\bar{k}}{k_0} \tilde{\nabla} \tilde{T} \right] + \frac{1}{Ra} \frac{\rho_0 H a^2}{k_0 \Delta T}.
\end{aligned} \tag{136}$$

Notice that the Ra number appears in different places than in (125) or (131). Of course after their appropriate changes of variables the solutions back with real dimensions are the same.

4.5 Linear stability analysis for basally heated convection

Using the Boussinesq approximation, it is easy to understand why the diffusive solution is not necessarily the solution chosen by the fluid. The standard way to test the stability of a solution is what physicists call a study of marginal stability (see also Chapter 5). It consists of substituting into the basic equations a known

solution plus an infinitely small perturbation and checking whether or not this perturbation amplifies, decreases or propagates. Its is only if the perturbation decreases in amplitude, that the tested solution is stable.

Since we benefit from the assumption that the perturbation has initially an infinitely small amplitude, computing its time evolution is simpler than solving the general equations since the nonlinear products can be neglected. The marginal stability study is therefore powerful for mapping the stability domain of a solution and describing its destabilization. At the same time it is also somewhat frustrating as the real unstable solution cannot be obtained.

This approach can be employed to understand the destabilisation of the diffusive solution. We use the Boussinesq approximation, with constant viscosity and conductivity, neglecting inertia and without internal heating. The nondimensionalized equations (131) (the tilde sign has been omitted for simplicity) write

$$\begin{aligned}\nabla \cdot \mathbf{v} &= 0, \\ -\nabla P + \nabla^2 \mathbf{v} - T \mathbf{e}_z &= 0, \\ \frac{\partial T}{\partial t} + \mathbf{v} \cdot \nabla T &= \frac{1}{Ra} \nabla^2 T,\end{aligned}\tag{137}$$

(\mathbf{e}_z is a normal vector directed along \mathbf{g}). The steady diffusive nondimensional temperature solution is $T = z$ with z directed along \mathbf{g} , 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 $z = 1$. As in the diffusive case the velocity is zero, the velocity induced by δT will be infinitely small $\delta \mathbf{v}$. In the nonlinear term, we can do the approximation $\mathbf{v} \cdot \nabla T = \delta \mathbf{v} \cdot \nabla (z + \delta T) \sim \delta v_z = v_z$. 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 $z = 1$ 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 2D, 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 Chapter 5). It is then straightforward to deduce that for such a thermal anomaly, the energy equation imposes a vertical velocity

$$v_z = - \left(\dot{\theta} + \frac{(k^2 + \pi^2)}{Ra} \theta \right) \sin(\pi z) \sin(kx).\tag{138}$$

From mass conservation the x -component of the velocity must be

$$v_x = - \frac{\pi}{k} \left(\dot{\theta} + \frac{(k^2 + \pi^2)}{Ra} \theta \right) \cos(\pi z) \cos(kx).\tag{139}$$

This flow has a vertical component that vanishes on the top and bottom surfaces where the horizontal component is maximum. The choice of the temperature structure corresponds to free-slip velocity conditions. When the velocity and the temperature are introduced in the momentum equation, the time evolution of the temperature perturbation is found

$$\dot{\theta} = \theta \left(\frac{k^2}{(\pi^2 + k^2)^2} - \frac{(\pi^2 + k^2)}{Ra} \right). \quad (140)$$

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

$$Ra = \frac{(\pi^2 + k^2)^3}{k^2}. \quad (141)$$

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

$$k = \frac{\pi}{\sqrt{2}}, \quad Ra_c = \frac{27}{4}\pi^4 \sim 657. \quad (142)$$

What can be interpreted as the size of one convective cell is π/k since one wavelength corresponds to two contrarotating 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 convection heated from below with free-slip boundary conditions. As soon as $Ra > Ra_c$ there is a wavenumber 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 $\delta v \cdot \nabla \delta T \ll \delta v \cdot \nabla z$ becomes invalid.

4.6 Road to chaos

Following the same approach, but with some additional complexities, the critical Rayleigh number for convection between no-slip surfaces, mixed free-slip/no-slip, with internal heating in Cartesian and spherical geometries could be obtained (e.g., see *Schubert et al. (2001)*). In all cases, critical Rayleigh numbers of order of 10^3 are found.

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. Using the values of Table 1, the critical Rayleigh number of the mantle would be attained for a non-adiabatic temperature difference between the surface and the CMB of only 0.025 K! The mantle Rayleigh number

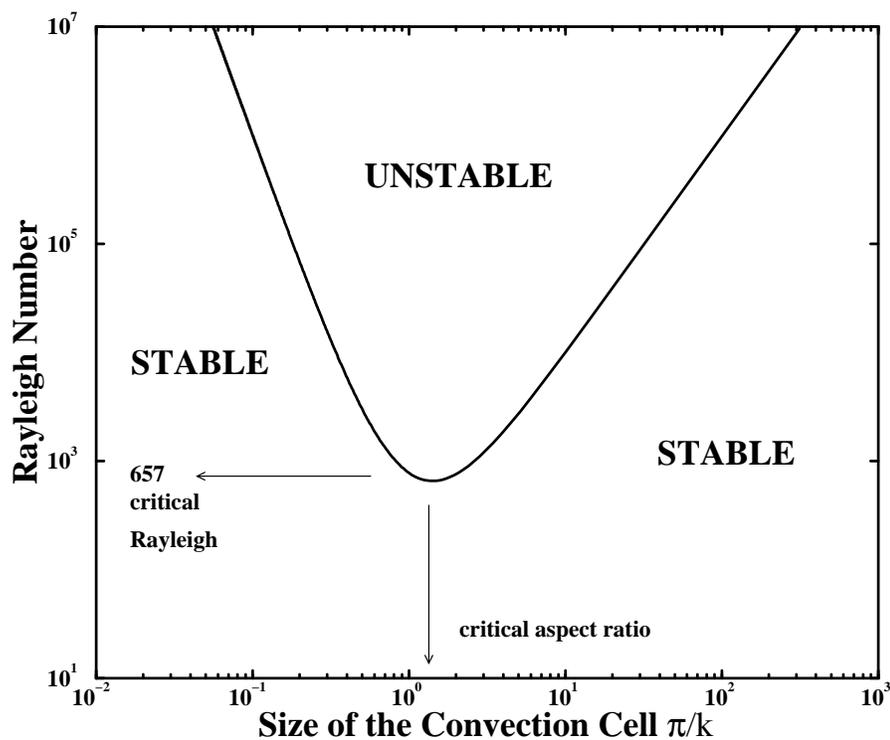


Figure 3: Critical Rayleigh number as a function of the half wavelength π/k (the size of the convection cells). Above this curve, convection occurs with a whole range of unstable wavelengths. Below this curve, the conductive temperature is stable since temperature perturbations of any wavelength, decrease. 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.

is 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 e.g., Chapter 4).

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 gases. 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. First, 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 the applications and illustrations to other chapters of this Treatise (e.g., Chapter 11 and Chapter 12). 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., *Haase*, 1990; *de Groot and Mazur*, 1984). 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 sepa-

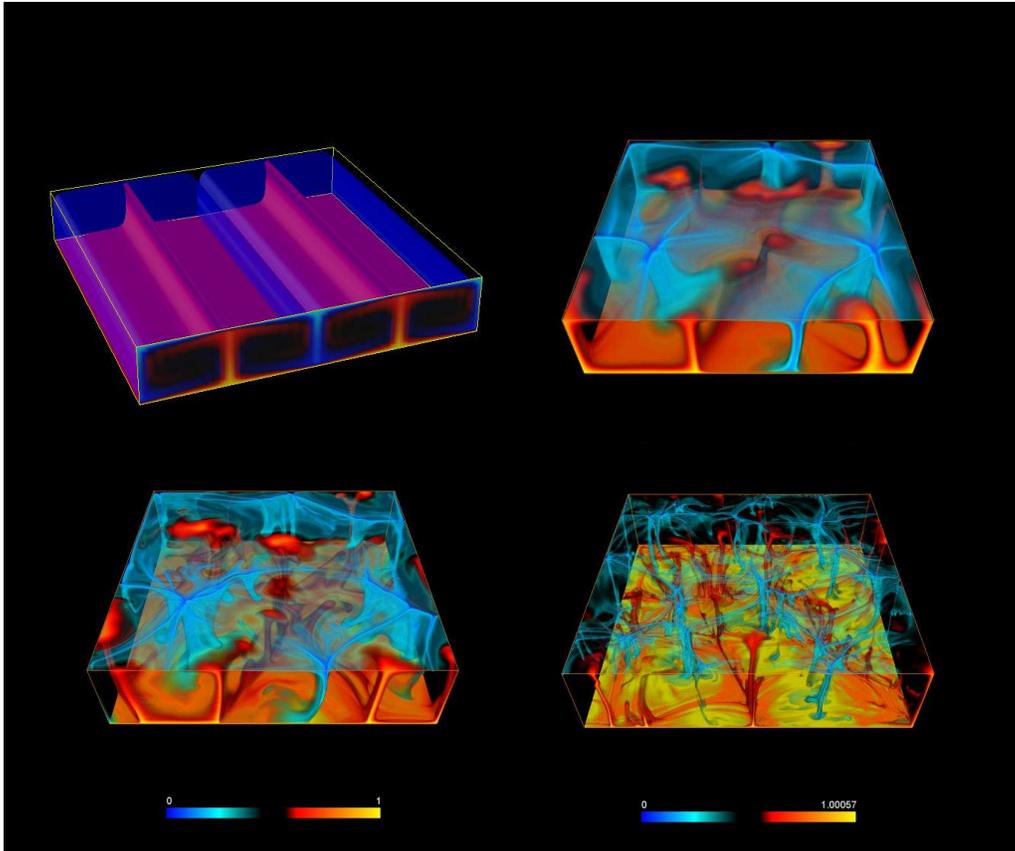


Figure 4: Convection patterns of a fluid heated from below at Rayleigh number 10^5 , 10^6 , 10^7 , 10^8 . The temperature color bars range from 0 (top boundary) to 1 (bottom boundary). The Boussinesq approximation was used (numerical simulations by F. Dubuffet). The increase in Rayleigh number corresponds 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 $\sim \sqrt{2}$ as predicted by marginal stability. For higher Rayleigh number, the patterns are highly time-dependent.

rated by physical interfaces which 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. 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 (e.g., in studies of dike propagation through hydraulic fracturing), or where fusion is parameterized in some way. We do not discuss these approaches.

5.1 Fluid dynamics of multicomponent flows in solution

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 , heat capacities C_p^i ...).

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

$$\frac{\partial(\phi_i \rho_i)}{\partial t} + \nabla \cdot (\phi_i \rho_i \mathbf{v}_i) = \Gamma_i, \quad (143)$$

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

$$\bar{\rho} = \sum \phi_i \rho_i, \quad (144)$$

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

$$\mathbf{v}_b = \frac{\sum \phi_i \rho_i \mathbf{v}_i}{\bar{\rho}}. \quad (145)$$

The average mass conservation can be obtained by summing the equations of component conservation (143),

$$\frac{\partial \bar{\rho}}{\partial t} + \nabla \cdot (\bar{\rho} \mathbf{v}_b) = 0, \quad (146)$$

since the sum of the rates of mass production is zero

$$\sum_i \Gamma_i = 0. \quad (147)$$

In equation (143), instead of the various component velocities \mathbf{v}_i , we can introduce the barycentric velocity \mathbf{v}_b and the diffusive flux of the component i with respect to this average flow,

$$\frac{\partial(\phi_i \rho_i)}{\partial t} + \nabla \cdot (\phi_i \rho_i \mathbf{v}_b) = -\nabla \mathbf{J}_i + \Gamma_i, \quad (148)$$

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

$$\mathbf{J}_i = \phi_i \rho_i (\mathbf{v}_i - \mathbf{v}_b). \quad (149)$$

By definition of the barycentric velocity (145), the sum of the diffusive flows just cancels out,

$$\sum_i \mathbf{J}_i = 0. \quad (150)$$

A diffusive transport is nothing else than an advective transfer with respect to the average barycentric velocity. We will show later in simple cases, that the diffusive transports are driven by concentration gradients.

If we introduce the mass fraction $C_i = \phi_i \rho_i / \bar{\rho}$ (in kg of i per kg of mixture), we can easily show from (146) and (148) that

$$\bar{\rho} \frac{DC_i}{Dt} = -\nabla \mathbf{J}_i + \Gamma_i, \quad (151)$$

where the Lagrangian derivative is defined with the barycentric velocity,

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v}_b \cdot \nabla. \quad (152)$$

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}_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

$$\nabla \cdot \boldsymbol{\tau} - \nabla P + \bar{\rho} \mathbf{g} = 0, \quad (153)$$

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 (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 Σ . Instead of the one component equation (28), we have to sum up various contributions and we get

$$\sum_i \frac{\partial(\phi_i \rho_i \mathcal{U}_i)}{\partial t} = -\nabla \cdot \left(\sum_i \phi_i \rho_i \mathcal{U}_i \mathbf{v}_i + P \sum_i \phi_i \mathbf{v}_i + \mathbf{q} - \boldsymbol{\tau} \cdot \mathbf{v}_b \right) + \mathbf{g} \cdot \sum_i \phi_i \rho_i \mathbf{v}_i + \bar{\rho} \bar{H}. \quad (154)$$

In this expression, we recognize the temporal changes in energy (\mathcal{U}_i is the component internal energy per unit mass, the kinetic energies are neglected), the energy advection, the pressure work, the thermal diffusion, the viscous stress work, the gravity work and the radioactivity production ($\bar{\rho} \bar{H} = \sum \phi_i \rho_i H_i$). The various ϕ_i come from the fact that each component i is present in proportion ϕ_i , in the volume Ω and on its surface Σ . 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 (145), of the diffusive fluxes, (149), of the momentum conservation, (153), and using $\sum_i \phi_i = 1$, the energy expression can be simplified to

$$\sum_i \phi_i \rho_i \frac{D\mathcal{H}_i}{Dt} = -\nabla \cdot \mathbf{q} - \sum_i \mathbf{J}_i \cdot \nabla \mathcal{H}_i + \frac{DP}{Dt} - \sum_i \Gamma_i \mathcal{H}_i + \boldsymbol{\tau} : \nabla \mathbf{v}_b + \bar{\rho} \bar{H}, \quad (155)$$

where \mathcal{H}_i are the component enthalpies

$$\mathcal{H}_i = \mathcal{U}_i + \frac{P}{\rho_i}. \quad (156)$$

The enthalpy variation for each component i can be expressed as a function of the state variables P and T . From $d\mathcal{H}_i = \delta Q_i + V_i dP$ and the expression of the exchanged heat (29), we can write

$$\rho_i \frac{D\mathcal{H}_i}{Dt} = \rho_i C_p^i \frac{DT}{Dt} + (1 - \alpha_i T) \frac{DP}{Dt}, \quad (157)$$

Finally, the expression for the temperature evolution is

$$\bar{\rho} \bar{C}_p \frac{DT}{Dt} = -\nabla \cdot \mathbf{q} - \sum_i \mathbf{J}_i \cdot \nabla \mathcal{H}_i + \bar{\alpha} T \frac{DP}{Dt} - \sum_i \Gamma_i \mathcal{H}_i + \boldsymbol{\tau} : \nabla \mathbf{v}_b + \bar{\rho} \bar{H}, \quad (158)$$

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

Compared to the homogeneous case (36), 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, $\sum_i \mathbf{J}_i \cdot \nabla \mathcal{H}_i$.

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

$$\sum_i \frac{\partial(\phi_i \rho_i \mathcal{S}_i)}{\partial t} = -\nabla \cdot \mathbf{J}_S + H_S, \quad (159)$$

where \mathbf{J}_S and H_S are the yet unknown entropy flux and source. The entropy of the various components take into account their specific entropies but also 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

$$\sum_i \phi_i \rho_i \frac{D\mathcal{S}_i}{Dt} = \nabla \cdot \left(\sum_i \phi_i \rho_i \mathcal{S}_i \mathbf{v}_b + \sum_i \mathcal{S}_i \mathbf{J}_i - \mathbf{J}_S \right) + H_S - \mathcal{S}_i \Gamma_i - \mathbf{J}_i \cdot \nabla \mathcal{S}_i. \quad (160)$$

However a second expression of the entropy conservation can be obtained from the enthalpy conservation, (155), using $d\mathcal{H}_i = Td\mathcal{S}_i + V_i dP$ which, in our case can be expressed as

$$\rho_i \frac{D\mathcal{H}_i}{Dt} = \rho_i T \frac{D\mathcal{S}_i}{Dt} + \frac{DP}{Dt}, \quad (161)$$

we derive

$$\sum_i \phi_i \rho_i T \frac{D\mathcal{S}_i}{Dt} = -\nabla \cdot \mathbf{q} - \sum_i \mathbf{J}_i \cdot \nabla \mathcal{H}_i - \sum_i \Gamma_i \mathcal{H}_i + \boldsymbol{\tau} : \nabla \mathbf{v}_b + \bar{\rho} \bar{H}. \quad (162)$$

A comparison of the two expressions for the entropy conservation, (160) and (162) allows us to identify the total entropy flux

$$\mathbf{J}_S = \frac{\mathbf{q}}{T} + \mathbf{v}_b \sum_i \phi_i \rho_i \mathcal{S}_i + \sum_i \mathcal{S}_i \mathbf{J}_i, \quad (163)$$

and the entropy sources

$$TH_S = -\left(\frac{\mathbf{q}}{T} + \sum_i \mathcal{S}_i \mathbf{J}_i\right) \cdot \nabla T - \sum_i \mathbf{J}_i \cdot \nabla \mu_i - \sum_i \Gamma_i \mu_i + \boldsymbol{\tau} : \nabla \mathbf{v}_b + \bar{\rho} \bar{H}. \quad (164)$$

where we introduced the chemical potentials $\mu_i = \mathcal{H}_i - T\mathcal{S}_i$. The total entropy flux, (163), is related to thermal diffusion and to advection and chemical diffusion of component entropies.

In (164), the two 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, using

$$\nabla \mu_i = T \nabla \frac{\mu_i}{T} + \mu_i \frac{\nabla T}{T}, \quad (165)$$

and $\mu_i = \mathcal{H}_i - T\mathcal{S}_i$, the entropy source (164) can be written as

$$TH_S = -(\mathbf{q} + \sum_i \mathcal{H}_i \mathbf{J}_i) \frac{\nabla T}{T} - T \sum_i \mathbf{J}_i \cdot \nabla \frac{\mu_i}{T} - \sum_i \Gamma_i \mu_i + \boldsymbol{\tau} : \nabla \mathbf{v}_b + \bar{\rho} \bar{H}. \quad (166)$$

We can also introduce the gradient of μ at constant temperature $\nabla_T \mu$ as

$$\nabla_T \mu_i = \nabla \mu_i + \mathcal{S}_i \nabla T, \quad (167)$$

which leads to

$$TH_S = -\frac{1}{T} \mathbf{q} \cdot \nabla T - \sum_i \mathbf{J}_i \cdot \nabla_T \mu_i - \sum_i \Gamma_i \mu_i + \boldsymbol{\tau} : \nabla \mathbf{v}_b + \bar{\rho} \bar{H}. \quad (168)$$

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).

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 Curie's principle. The positivity of the entropy production imposes three conditions, coupling tensors, vectors, and scalars.

$$\boldsymbol{\tau} : \nabla \mathbf{v}_b \geq 0, \quad -\mathbf{q} \cdot \frac{\nabla T}{T} - \sum_i \mathbf{J}_i \cdot \nabla_T \mu_i \geq 0 \quad - \sum_i \Gamma_i \mu_i \geq 0. \quad (169)$$

The first term relates tensors and we have already discussed its implications for the rheology in section 3.2.

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

$$\begin{pmatrix} \mathbf{J}_1 \\ \mathbf{J}_2 \\ \dots \\ \mathbf{q} \end{pmatrix} = - \begin{pmatrix} m_{11} & m_{12} & \dots & m_{1q} \\ m_{21} & m_{22} & \dots & m_{2q} \\ \dots & \dots & \dots & \dots \\ m_{q1} & m_{q2} & \dots & m_{qq} \end{pmatrix} \begin{pmatrix} \nabla_T \mu_1 \\ \nabla_T \mu_2 \\ \dots \\ \nabla T/T \end{pmatrix} \quad (170)$$

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

$$TH_S^{(v)} = \mathbf{X}^t M \mathbf{X} = \mathbf{X}^t \frac{M + M^t}{2} \mathbf{X}. \quad (171)$$

According to the second law of thermodynamics, the symmetric part of the matrix M , $(M + M^t)/2$ must be positive definite, i.e., the right hand side of equation (171) must be positive for any vectors \mathbf{X} .

Thermodynamic equilibrium is always dynamic on a microscopic scale. To maintain equilibrium, a process and its reverse must occur at the same microscopic rate. A consequence, known as the Onsager reciprocal relations, is the existence of symmetry or antisymmetry between m_{ij} and m_{ji} (Onsager, 1903-1976). A general discussion can be found in e.g., *de Groot and Mazur* (1984) or *Woods* (1975). When the forces are even functions of the velocities and in the absence of magnetic field, the matrix M must be symmetric. As $\nabla T/T$, and $\nabla_T \mu_i$ are even functions, as independent of the velocities, $m_{ij} = m_{ji}$.

In the general case, the transport of heat by concentration gradients (Dufour effect, (Dufour, 1832-1892)) or the transport of concentration by temperature gradients (Soret effect, (Soret, 1827-1890)) are 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 definite positivity of a symmetric i by i matrix is not a very strong constraint). We therefore restrict our study to a simple two-component system where

$$\begin{pmatrix} \mathbf{J}_1 \\ \mathbf{J}_2 \end{pmatrix} = - \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} \nabla_T \mu_1 \\ \nabla_T \mu_2 \end{pmatrix} \quad (172)$$

For such a simple case, the sum of the flux must cancel, see (150), and since the Onsager relations impose the symmetry of the matrix, the coefficients m_{ij} must verify

$$m_{11} + m_{21} = m_{12} + m_{22} = m_{12} - m_{21} = 0. \quad (173)$$

Only one coefficient, for example m_{11} can be freely chosen, the fluxes can be written

$$\mathbf{J}_1 = m_{11} \nabla_T (\mu_2 - \mu_1), \quad (174)$$

$$\mathbf{J}_2 = m_{11} \nabla_T (\mu_1 - \mu_2), \quad (175)$$

and the second law requires $m_{11} > 0$. If the component 1 is in small quantity (the solute) and the component 2 is in large quantity (the solvent, with $\mu_2 = 0$), we can easily track the evolution of solute concentration C_1 . Its chemical diffusion flux is $\mathbf{J}_1 = -m_{11} \nabla_T \mu_1$ and according to (151), its concentration satisfies

$$\bar{\rho} \left(\frac{\partial C}{\partial t} + \mathbf{v}_b \cdot \nabla C \right) = \nabla \cdot (m \nabla_T \mu) + \Gamma, \quad (176)$$

where the subscripts 1 have been omitted.

For a solute the chemical potential is a standard chemical potential μ_0 plus a mixing term expressing the entropy gain (configuration 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 remains uniform, the advection-diffusion equation is obtained

$$\frac{\partial C}{\partial t} + \mathbf{v}_b \cdot \nabla C = \nabla \cdot (D \nabla C) + \Gamma \quad (177)$$

with a diffusion coefficient $D = m/\bar{\rho}(\partial\mu/\partial C)$, most likely T -dependent. The negative linear relationship between chemical diffusion and concentration gradient 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 $[\mu] = [\mu_0 + RT \log a(C)] = 0$. The concentration ratio of C (or partition coefficient of C), must therefore verify

$$\frac{a(C)^+}{a(C)^-} = \exp\left(-\frac{\mu_0^+ - \mu_0^-}{RT}\right), \quad (178)$$

where $[\]^+$ and $[\]^-$ denote the values on the two sides of the discontinuity. This equation corresponds to the general rule of chemical equilibria.

The last entropy source in (169) 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 $j^{th} = (1\dots r)$ chemical reaction with reaction rate, Γ_j . We can express Γ_i as

$$\Gamma_i = \sum_{j=1\dots r} n_i^j \Gamma_j, \quad (179)$$

and the second law imposes

$$-\sum_{j=1\dots r} \Gamma_j \sum_i n_i^j \mu_i \geq 0, \quad (180)$$

The positivity of the entropy source is satisfied if the kinetic rates of the $j^{th} = 1\dots r$ 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

$$\Gamma_j = -\mathcal{R}_j \Delta G_j. \quad (181)$$

Chemical reaction rates are 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 . In the case of exact thermodynamic equilibrium, $\sum_i n_i^j \mu_i = 0$, the second law is of course satisfied.

In the same way as we defined the affinity ΔG_j of the reaction j , we can define its enthalpy $\Delta \mathcal{H}_j = \sum_i n_i^j \mathcal{H}_i$. The enthalpy exchange term of the energy equation (158), $\sum_i \Gamma_i \mathcal{H}_i$ can also be written $\sum_j \Gamma_j \Delta \mathcal{H}_j$, products of the reaction rates by 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 6.6.

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 (177), and that this solute is not involved in any chemical reaction, $\Gamma = 0$. For simplicity, we assume that the barycentric flow is incompressible (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, i.e., $\mathbf{v} \cdot \mathbf{n} = 0$ and $\nabla C \cdot \mathbf{n} = 0$ on the surface Σ 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 (177) implies

$$\frac{d}{dt} \int_{\Omega} C dV = - \int_{\Sigma} (C \mathbf{v} - D \nabla C) \cdot d\mathbf{S} = 0. \quad (182)$$

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. Multiplying (177) by $2C$, we get after some algebra

$$\frac{\partial C^2}{\partial t} + \nabla \cdot (\mathbf{v}_b C^2) = D \nabla^2 C^2 - 2D |\nabla C|^2. \quad (183)$$

This expression when integrated over the closed volume Ω implies that

$$\frac{d}{dt} \int_{\Omega} C^2 dV = -2D \int_{\Omega} |\nabla C|^2 dV. \quad (184)$$

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

The concept of mixing is associated with the idea that the concentration variance decreases with time. Since the average mixing rate is proportional to the diffusion D , (184), we note however, that a non-diffusive flow does not mix at

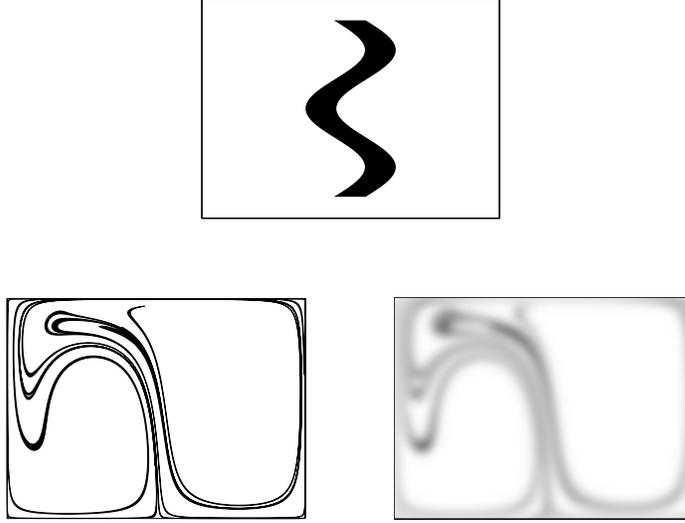


Figure 5: An initial heterogeneity (top) is introduced at $t = 0$ into a time-dependent convection cell. Without diffusion, $D = 0$, (bottom left), the heterogeneity is stirred by convection and then stretched on the form of thin ribbons. However the variance of the heterogeneity concentration remains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

all. A diffusive flow just stirs the heterogeneities. In other terms, if the initial concentration is either $C = 1$ or $C = 0$, a perfect mixing is achieved after a time t if the concentration is everywhere $C = \bar{C}$, the average concentration. 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 ($D = 10^{-19} \text{ m}^2 \text{ s}^{-1}$ for uranium, $D = 10^{-13} \text{ m}^2 \text{ s}^{-1}$ for helium (see Table 2)) and many studies have totally neglected chemical diffusion. We see, at face value, that these models are not really mixing, only stirring the heterogeneities. Without diffusion a chemical heterogeneity (e.g., a piece of subducted oceanic crust) will forever remain the same petrological heterogeneity, only its shape will change.

Since the mixing rate is related to the compositional gradient (184) we should discuss the evolution of this gradient. We multiply (177) by the operator $2\nabla C \cdot \nabla$ to obtain

$$\frac{\partial |\nabla C|^2}{\partial t} + \nabla \cdot (\mathbf{v}_b |\nabla C|^2) + 2\nabla C \cdot \dot{\underline{\epsilon}} \cdot \nabla C = 2D \left(\nabla \cdot (\nabla^2 C \nabla C) - (\nabla^2 C)^2 \right), \quad (185)$$

which can be integrated as

$$\frac{d}{dt} \int_{\Omega} |\nabla C|^2 dV = -2 \int_{\Omega} \nabla C \cdot \underline{\dot{\epsilon}} \cdot \nabla C dV - 2D \int_{\Omega} (\nabla^2 C)^2 dV. \quad (186)$$

The rate of gradient production is related to the flow properties through the strain rate tensor $\underline{\dot{\epsilon}}$ 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 (186)), 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 becomes 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 (186) is thus of order of $-\dot{\epsilon}_{min} |\nabla C|^2$ ($\dot{\epsilon}_{min}$ is the local, negative eigenvalue of the strain rate tensor). Stirring is thus the source of production of concentration gradient.

We can now understand the interplay between advection and diffusion. Even when the diffusion coefficient D is vanishingly small in (186), the stirring of the flow by convection will enhanced the concentration gradients until the average diffusion term, proportional to the concentration gradients, will become large enough (see (184)), for a rapid decrease of the concentration variance. We illustrate this behavior in the next two paragraphs by choosing a simple expression for the strain rate and computing the evolution of concentration through time.

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 a little bit more the stirring properties of a flow. Let us consider a vertical piece of heterogeneity of width $2d_0$, height $2L$ ($L \ll d_0$) in a simple shear flow (or Couette flow, (Couette, 1858-1943)) $u_x = \dot{\epsilon}z$. Its top and bottom ends are at $(0, L_0)$, $(0, -L_0)$ and they will be advected to $(\dot{\epsilon}tL_0, L_0)$, $(-\dot{\epsilon}tL_0, -L_0)$ after a time t . As the heterogeneity length increases as $2L_0(1 + \dot{\epsilon}^2t^2)$, mass conservation implies that its half width $d(t)$ decreases as

$$d(t) = \frac{d_0}{\sqrt{1 + \dot{\epsilon}^2t^2}}. \quad (187)$$

Such flows, in which heterogeneities are stretched at rate $\sim 1/t$, are called flows with laminar stirring. They are not very efficient in enhancing the diffusion be-

cause they do not increase the concentration gradients, typically of order $1/d(t)$, fast enough.

On the contrary, in a pure shear flow, $u_z = z\dot{\epsilon}$, the length of the heterogeneity would increase as $L = L_0 \exp(\dot{\epsilon}t)$ and its width would shrink as

$$d(t) = d_0 \exp(-\dot{\epsilon}t). \quad (188)$$

Such a flow is called flow with turbulent stirring. This usual vocabulary is very poorly chosen 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 Ra number mixes more efficiently than low Ra number convection is often true (*Schmalzl et al.*, 1996). However steady 3D 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).

5.1.7 Diffusion in Lagrangian coordinates

In section 5.1.5 we discussed the mixing properties from an 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), i.e., by solving the advection-diffusion equations in a coordinate frame that follows the deformation.

Let us consider a strip of thickness $2l_0$ with an initial concentration C_0 embedded in a infinite matrix of concentration C_∞ . In the absence of motion, the solution of the advection-diffusion equation (177) can be expressed using the error function and the time dependent concentration $C(x, t)$ is given by

$$\frac{C(x, t) - C_\infty}{C_0 - C_\infty} = \frac{1}{2} \left[\operatorname{erf} \left(\frac{l_0 - x}{2\sqrt{Dt}} \right) + \operatorname{erf} \left(\frac{l_0 + x}{2\sqrt{Dt}} \right) \right], \quad (189)$$

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

$$\frac{C(0, t) - C_\infty}{C_0 - C_\infty} = \operatorname{erf} \left(\frac{l_0}{2\sqrt{Dt}} \right), \quad (190)$$

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

$$t_0 \sim \frac{l_0^2}{D}. \quad (191)$$

($\text{erf}(1/2)$ 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). In a motionless mantle, even the relatively mobile helium would be frozen since the Earth formed since it would only have migrated around 50 cm.

However, this idea of a \sqrt{t} diffusion is totally wrong since the flow stirs the heterogeneity and increases the quantity of compositional gradients (186) which in turn accelerates the mixing process (184). Assuming that the problem remains two-dimensional enough so that diffusion only occurs perpendicular to the deforming heterogeneity, let $l(t)$ be the thickness of the strip containing a chemical heterogeneity. The velocity perpendicular to the strip would locally be at first order

$$v_x = \frac{x}{l(t)} \frac{d}{dt} l(t), \quad (192)$$

(each side of the stripe, at $x = \pm l(t)$ moves at $\pm dl(t)/dt$).

We can choose as a new space variable $\tilde{x} = xl_0/l(t)$, in such a way that the Lagrangian coordinate \tilde{x} will vary between the fixed values $-l_0$ and l_0 . The diffusion equation becomes

$$\left(\frac{\partial C}{\partial t} \right)_{\tilde{x}} = D \left(\frac{l_0}{l(t)} \right)^2 \frac{\partial^2 C}{\partial \tilde{x}^2}, \quad (193)$$

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 $D (l_0/l(t))^2$. This equivalent diffusivity is larger than D and increases with time as $l(t)$ decreases.

To solve analytically equation (193) it is appropriate to rescale the time variable by defining $\tilde{t} = F(t)$ with

$$F(t) = \int_0^t (l_0/l(u))^2 du, \quad (194)$$

and the resulting advection-diffusion equation in Lagrangian coordinates becomes the simple diffusion equation with constant diffusivity. Its solution is given by (189) where t and x are replaced by \tilde{t} and \tilde{x} . For example, the concentration at the center of the deformable strip varies like

$$\frac{C(0, t) - C_\infty}{C_0 - C_\infty} = \text{erf} \left(\frac{l_0}{2\sqrt{DF(t)}} \right), \quad (195)$$

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

$$F(t) \sim \frac{l_0^2}{D}. \quad (196)$$

Table 2: Homogenization times for helium and uranium assuming an heterogeneity of initial thickness $2l_0 = 7$ km and a strain rate of $5 \times 10^{-16} \text{ s}^{-1}$.

	Uranium	Helium
$D \text{ m}^2 \text{ s}^{-1}$	10^{-19}	10^{-13}
$t_0 \text{ ma}$	3.88×10^{13}	3.88×10^7
$t_L \text{ ma}$	3.60×10^5	3.60×10^3
$t_T \text{ ma}$	1920	1490

To perform a numerical application let us consider that the flow is either a simple shear (187), or a pure shear deformation, (188). Computing $F(t)$ from equation (194) is easy and, assuming $\dot{\epsilon}t \gg 1$, we get from (196), the homogenization times

$$t_L \sim \frac{3^{1/3}l_0^{2/3}}{\dot{\epsilon}^{2/3}D^{1/3}}, \quad (197)$$

and

$$t_T \sim \frac{1}{2\dot{\epsilon}} \log \frac{2l_0^2\dot{\epsilon}}{D}, \quad (198)$$

respectively. For the same oceanic crust of initial thickness 7 km, we get homogenization times of about 1.49 byr for He and 1.92 byr for U if we use the pure shear mechanism and assume rather arbitrary that $\dot{\epsilon} = 5 \times 10^{-16} \text{ s}^{-1}$ (this corresponds to a typical plate velocity of 7 cm yr⁻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 may be comparable.

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., *van Keken et al.*, 2002; *Tackley and Xie*, 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 diffused totally into the background. In other words, even if diffusion seems negligible, diffusion will erase all heterogeneities after a finite time mostly controlled by the stirring properties of the flow.

5.2 Fluid dynamics of two phase flows

Up to now, in all of section 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 hotspots, but they may also be present in the middle and deep mantle (*Williams and Garnero, 1996; Bercovici and Karato, 2003*) and they were certainly more frequent when the Earth was younger. We discuss the situation where two phases, fluid and matrix, each having the same composition, can interact. In contrast to section 5.1, where the proportion and velocity of each component in solution was 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, $\tilde{\mathbf{v}}_m$, or the interstitial velocity of the melt, $\tilde{\mathbf{v}}_f$.

We assume that the two phases are individually homogeneous, incompressible and with densities ρ_f and ρ_m . They have Newtonian rheologies with viscosities μ_f and μ_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 $\text{kg m}^{-3} \text{s}^{-1}$). 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 $1 - \phi$, $\Delta\Gamma$ and $-\Delta\Gamma$, must leave the equations unchanged. This rule is both a physical requirement and a strong guidance in establishing the general equations.

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).

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 convoluted topology. From it, we can define first, the porosity (volume fraction of fluid) ϕ , then, the fluid and matrix averaged velocities, \mathbf{v}_f and \mathbf{v}_m by

$$\phi = \frac{1}{\delta V} \int_{\delta V} \theta dV, \quad (199)$$

$$\phi \mathbf{v}_f = \frac{1}{\delta V} \int_{\delta V} \theta \tilde{\mathbf{v}}_f dV, \quad (1 - \phi) \mathbf{v}_m = \frac{1}{\delta V} \int_{\delta V} (1 - \theta) \tilde{\mathbf{v}}_m dV. \quad (200)$$

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 (*McKenzie, 1984; Bercovici et al., 2001a*). They are

$$\frac{\partial \phi}{\partial t} + \nabla \cdot [\phi \mathbf{v}_f] = \frac{\Delta\Gamma}{\rho_f}, \quad (201)$$

$$-\frac{\partial \phi}{\partial t} + \nabla \cdot [(1 - \phi)\mathbf{v}_m] = -\frac{\Delta\Gamma}{\rho_m}. \quad (202)$$

We get the same equations as in (143) except that we refer to ϕ , $1 - \phi$, $\Delta\Gamma$ instead of ϕ_1 , ϕ_2 and Γ_1 . When averaged, the mass conservation equations of two separated phase takes the same form as the mass conservation equations of two components in a solution.

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

$$\bar{q} = \phi q_f + (1 - \phi)q_m, \quad \Delta q = q_m - q_f. \quad (203)$$

The velocity $\bar{\mathbf{v}}$ is volume averaged and is different from the barycentric velocity (145), $\mathbf{v}_b = (\phi\rho_f\mathbf{v}_f + (1 - \phi)\rho_m\mathbf{v}_m)/\bar{\rho}$. By combining the fluid and matrix mass conservation equations we get the total mass conservation equation

$$\frac{\partial \bar{\rho}}{\partial t} + \nabla \cdot (\bar{\rho}\mathbf{v}_b) = 0, \quad (204)$$

(as before, (146)), and the time rate of change in volume during melting

$$\nabla \cdot \bar{\mathbf{v}} = \Delta\Gamma \frac{\Delta\rho}{\rho_f\rho_m}. \quad (205)$$

5.2.2 Momentum conservation of matrix and fluid

Total momentum conservation, i.e., the balance of the forces applied to the mixture is

$$\nabla \cdot \bar{\boldsymbol{\tau}} - \nabla \bar{P} + \bar{\rho}\mathbf{g} = 0. \quad (206)$$

We have considered that the only force is due to gravity, although surface tension between the two phases could also be introduced (*Bercovici et al.*, 2001a). In this equation, \bar{P} , $\bar{\boldsymbol{\tau}}$ and $\bar{\rho}$ are the average pressure, stress and density. The equation is not surprising and looks identical to its counterpart for a multicomponent solution (153). However the average pressure and stresses, $\bar{P} = \phi P_f + (1 - \phi)P_m$ and $\bar{\boldsymbol{\tau}} = \phi\boldsymbol{\tau}_f + (1 - \phi)\boldsymbol{\tau}_m$ 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 may feel the same pressure does not rest on any physical justification and certainly cannot hold if surface tension is present. We will show later that their pressure difference controls the rate of porosity change.

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

$$-\nabla[\phi P_f] + \phi\rho_f\mathbf{g} + \nabla \cdot [\phi\boldsymbol{\tau}_f] + \mathbf{h}_f = 0, \quad (207)$$

$$-\nabla[(1-\phi)P_m] + (1-\phi)\rho_m\mathbf{g} + \nabla \cdot [(1-\phi)\boldsymbol{\tau}_m] + \mathbf{h}_m = 0. \quad (208)$$

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 parametrized 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*) (Darcy, 1803-1858). In the absence of gravity and when the pressures are uniform and equal, no motion should occur even in the presence of non-uniform porosity. In this situation where $\Delta\mathbf{v} = \boldsymbol{\tau}_f = \boldsymbol{\tau}_m = 0$ 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 h_f and 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

$$\begin{aligned} \mathbf{h}_f &= c\Delta\mathbf{v} + P_f\nabla\phi + \omega\Delta P\nabla\phi, \\ \mathbf{h}_m &= -c\Delta\mathbf{v} + P_m\nabla(1-\phi) + (1-\omega)\Delta P\nabla\phi. \end{aligned} \quad (209)$$

These expressions verify $\mathbf{h}_f + \mathbf{h}_m = 0$, are Galilean and material invariant and allow equilibrium of a mixture with non uniform porosity but uniform and equal pressures.

The interaction coefficient c is related to permeability which is itself a function of porosity. A symmetrical form compatible with the usual Darcy term is (see *Bercovici et al., 2001a*)

$$c = \frac{\mu_f\mu_m}{k_0[\mu_f(1-\phi)^{n-2} + \mu_m\phi^{n-2}]}, \quad (210)$$

where the permeability of the form $k_0\phi^n$ was used (usually $n \sim 2-3$). Assuming $n = 2$ and $\mu_m \gg \mu_f$, the interaction coefficient becomes a constant, $c = \mu_f/k_0$.

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

$\mu_f = 0$). A symmetrical form like

$$\omega = \frac{\phi\mu_f}{\phi\mu_f + (1 - \phi)\mu_m} \quad (211)$$

satisfies these conditions.

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

$$-\phi[\nabla P_f - \rho_f \mathbf{g}] + \nabla \cdot [\phi \boldsymbol{\tau}_f] + c\Delta \mathbf{v} + \omega \Delta P \nabla \phi = 0, \quad (212)$$

$$-(1 - \phi)[\nabla P_m - \rho_m \mathbf{g}] + \nabla \cdot [(1 - \phi) \boldsymbol{\tau}_m] - c\Delta \mathbf{v} + (1 - \omega) \Delta P \nabla \phi = 0. \quad (213)$$

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

$$\boldsymbol{\tau}_j = \mu_j \left(\nabla \mathbf{v}_j + [\nabla \mathbf{v}_j]^t - \frac{2}{3} \nabla \cdot \mathbf{v}_j \mathbf{I} \right), \quad (214)$$

where j stands for f or m . There is no difference between the expression of the rheology for the isolated component and for the component in the mixture.

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 (Sramek et al., 2006). Otherwise the global conservation is straightforward and can be expressed by the following equation where the left side represents the temporal change of 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,

$$\begin{aligned} & \frac{\partial}{\partial t} [\phi \rho_f \mathcal{U}_f + (1 - \phi) \rho_m \mathcal{U}_m] \\ &= \bar{\rho} \bar{H} - \nabla \cdot \mathbf{q} - \nabla \cdot [\phi \rho_f \mathcal{U}_f \mathbf{v}_f + (1 - \phi) \rho_m \mathcal{U}_m \mathbf{v}_m] \\ &+ \nabla \cdot [-\phi P_f \mathbf{v}_f - (1 - \phi) P_m \mathbf{v}_m + \phi \mathbf{v}_f \cdot \boldsymbol{\tau}_f + (1 - \phi) \mathbf{v}_m \cdot \boldsymbol{\tau}_m] \\ &+ \phi \mathbf{v}_f \cdot \rho_f \mathbf{g} + (1 - \phi) \mathbf{v}_m \cdot \rho_m \mathbf{g}. \end{aligned} \quad (215)$$

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

$$\begin{aligned} & \phi \rho_f C_f \frac{D_f T}{Dt} + (1 - \phi) \rho_m C_m \frac{D_m T}{Dt} \\ &= -\nabla \cdot \mathbf{q} - \Delta P \frac{D_\omega \phi}{Dt} + \Delta \mathcal{H} \Delta \Gamma + \Psi + \bar{\rho} \bar{H}, \end{aligned} \quad (216)$$

where Ψ is the rate of deformational work

$$\Psi = \phi \nabla \mathbf{v}_f : \boldsymbol{\tau}_f + (1 - \phi) \nabla \mathbf{v}_m : \boldsymbol{\tau}_m + c(\Delta v)^2. \quad (217)$$

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

$$\frac{D_j}{Dt} = \frac{\partial}{\partial t} + \mathbf{v}_j \cdot \nabla, \quad (218)$$

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

$$\mathbf{v}_\omega = \omega \mathbf{v}_f + (1 - \omega) \mathbf{v}_m, \quad (219)$$

In contrast to section 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 diffusion terms. Here the two components may have very different velocities and we have to define various substantial derivatives.

Since ω represents a partitioning of pressure jump, it is not surprising to find the velocity \mathbf{v}_ω (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 $(q_m - q^\omega) - (q_f - q^\omega)$. 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 to define interface values by

$$q^\omega = (1 - \omega)q_f + \omega q_m. \quad (220)$$

Notice by comparing the expressions of the interface velocity \mathbf{v}_ω , (219), and interface value, q^ω (220), that ω swaps with $1 - \omega$.

The right side of (216) 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 (155).

5.2.4 Phenomenological laws

Entropy conservation is needed to constrain the the pressure jump between phases and the melting rate. Starting from entropy conservation

$$\frac{\partial}{\partial t} [\phi \rho_f \mathcal{S}_f + (1 - \phi) \rho_m \mathcal{S}_m] = -\nabla \cdot \mathbf{J}_S + H_S, \quad (221)$$

where \mathbf{J}_S is the total entropy flux and H_S is the internal entropy production, we compare the energy and the entropy equations (216) and (221) taking into account that, for each incompressible phase, $ds_j = C_j dT/T = du_j/T$. After some algebra, one gets

$$\mathbf{J}_S = \phi \rho_f \mathcal{S}_f \mathbf{v}_f + (1 - \phi) \rho_m \mathcal{S}_m \mathbf{v}_m + \frac{\mathbf{q}}{T}, \quad (222)$$

$$T H_S = -\frac{1}{T} \mathbf{q} \cdot \nabla T - \Delta P \frac{D_\omega \phi}{Dt} + \Delta \mu \Delta \Gamma + \Psi + \bar{\rho} \bar{H}, \quad (223)$$

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

$$\Delta \mu = \Delta \mathcal{H} - T \Delta S \quad (224)$$

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

Following the standard procedure of non-equilibrium 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

$$\begin{pmatrix} D_\omega \phi / Dt \\ \Delta \Gamma \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} -\Delta P \\ \Delta \mu \end{pmatrix} \quad (225)$$

The matrix of phenomenological coefficients m_{ij} is positive definite and symmetrical by Onsager's theorem, $m_{12} = m_{21}$, as the two thermodynamic forces ($-\Delta P$ and $\Delta \mu$) are even in velocity. For a 2x2 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 (201) and (202) and the definitions of \mathbf{v}_ω and ρ^ω , (219) and (220), we can combine equations (225) to get

$$\Delta P = -\frac{m_{22}}{m_{11} m_{22} - m_{12}^2} \left[(1 - \omega)(1 - \phi) \nabla \cdot \mathbf{v}_m - \omega \phi \nabla \cdot \mathbf{v}_f + \left(\frac{\rho^\omega}{\rho_f \rho_m} - \frac{m_{12}}{m_{22}} \right) \Delta \Gamma \right]. \quad (226)$$

In the limiting case where the two phases have the same density $\rho_f = \rho_m = \rho^\omega$, melting can occur with no motion, $\mathbf{v}_m = \mathbf{v}_f = 0$, (226) 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, $\nabla \cdot \mathbf{v}_m = 0$, and thus the two pressures should also be the same, $\Delta P = 0$. In this situation all the terms in equation (226) are 0 except for the term proportional to $\Delta\Gamma$. This suggests in the general case, that

$$\frac{m_{12}}{m_{22}} = \frac{\rho^\omega}{\rho_f \rho_m}. \quad (227)$$

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 equations (225) as

$$\Delta P = -\zeta \left[\frac{D_\omega \phi}{Dt} - \frac{\rho^\omega}{\rho_f \rho_m} \Delta\Gamma \right], \quad (228)$$

$$\Delta\Gamma = \mathcal{R} \left[\Delta\mu - \frac{\rho^\omega}{\rho_f \rho_m} \Delta P \right] \quad (229)$$

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 3.2 (see also the summary section 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 $\lim_{\phi \rightarrow 0} \zeta(\phi) = \lim_{\phi \rightarrow 1} \zeta(\phi) = +\infty$. Simple micromechanical models (e.g., *Nye, 1953; Bercovici et al., 2001a*) allow us to evaluate the bulk viscosity as

$$\zeta = K_0 \frac{(\mu_f + \mu_m)}{\phi(1 - \phi)}. \quad (230)$$

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 (228). This hypothesis led to a damage theory developed in (*Bercovici et al., 2001b; Ricard and Bercovici, 2003; Bercovici and Ricard, 2005*). Here we assume that the system remains close enough to mechanical equilibrium that damage does not occur.

The second equation (229) 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,

$$\mu_i^* = \mathcal{U}_i + \frac{P^\omega}{\rho_i} - T\mathcal{S}_i, \quad (231)$$

where i stands for f or m and write the kinetic equation (229),

$$\Delta\Gamma = \mathcal{R}\Delta\mu^*. \quad (232)$$

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 (228) and (229), we can show that the entropy production is indeed positive and given by

$$TH_S = k\frac{1}{T}|\nabla T|^2 + \frac{\Delta P^2}{\zeta} + \frac{(\Delta\mu^*)^2}{\mathcal{R}} + \Psi + \bar{\rho}\bar{H}. \quad (233)$$

Chemical relaxation and bulk compression are associated with dissipative terms.

5.2.5 Summary equations

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

The mass conservations equations are (201) and (202). The equation of conservation of momentum for the fluid phase is

$$-\phi\nabla[P_f + \rho_f gz] + c\Delta\mathbf{v} = 0, \quad (234)$$

(assuming z positive upward). This is a typical Darcy equilibrium with $c = \mu_f\phi^2/k(\phi)$, where k is the permeability (often varying as $k_0\phi^n$ with $n = 2$ or 3). The second momentum equation could be the matrix momentum (208) or a combined force-difference (or action-reaction) equation

$$-\nabla[(1-\phi)\Delta P] + (1-\phi)\Delta\rho\mathbf{g} + \nabla \cdot [(1-\phi)\boldsymbol{\tau}_m] - \frac{c\Delta\mathbf{v}}{\phi} = 0, \quad (235)$$

where the deviatoric stress in the matrix is given by

$$\boldsymbol{\tau}_m = \mu_m \left(\nabla\mathbf{v}_m + [\nabla\mathbf{v}_m]^T - \frac{2}{3}\nabla \cdot \mathbf{v}_m\mathbf{I} \right). \quad (236)$$

and the pressure jump between phases, (228), becomes

$$\Delta P = -\zeta(1-\phi)\nabla \cdot \mathbf{v}_m, \quad (237)$$

if an equivalent bulk viscosity is used, or

$$\Delta P = -K_0 \mu_m \frac{\nabla \cdot \mathbf{v}_m}{\phi}, \quad (238)$$

from the micromechanical model, (230), of *Bercovici et al.* (2001a).

The action-reaction equation (235) can also be written in a different way for example by elimination of $\Delta \mathbf{v}$ taken from the Darcy equilibrium (234),

$$-\nabla P_f + \nabla \cdot [(1 - \phi) \boldsymbol{\tau}_m^*] + \bar{\rho} \mathbf{g} = 0, \quad (239)$$

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

$$\boldsymbol{\tau}_m^* = \mu_m \left(\nabla \mathbf{v}_m + [\nabla \mathbf{v}_m]^T - \frac{2}{3} \nabla \cdot \mathbf{v}_m \mathbf{I} \right) + \zeta (1 - \phi) \nabla \cdot \mathbf{v}_m \mathbf{I}. \quad (240)$$

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 (91)), with a bulk viscosity $(1 - \phi)\zeta \sim \zeta$ (*McKenzie*, 1984). This analogy only holds without surface tension between phases.

The rate of melting is controlled by

$$\Delta \Gamma = \mathcal{R} \left(\Delta u + P_f \left(\frac{1}{\rho_m} - \frac{1}{\rho_f} \right) - T \Delta s \right), \quad (241)$$

and the energy equation is

$$\begin{aligned} \rho_f \phi C_f \frac{D_f T}{Dt} + \rho_m (1 - \phi) C_m \frac{D_m T}{Dt} - \Delta \mathcal{H} \Delta \Gamma = \\ \bar{\rho} \bar{H} - \nabla \cdot \mathbf{q} + \frac{\Delta \Gamma^2}{\mathcal{R}} + K_0 \mu_m \frac{1 - \phi}{\phi} (\nabla \cdot \mathbf{v}_m)^2 + \Psi, \end{aligned} \quad (242)$$

where we have assumed the relation (238).

These equations have been used by many authors with various levels of approximation. 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., *McKenzie*, 1984; *Scott and Stevenson*, 1984; *Richter and Mckenzie*, 1984; *Ribe*, 1985a,b; *Scott and Stevenson*, 1986, 1989; *Stevenson*, 1989; *Kelemen et al.*, 1997; *Choblet and Parmentier*, 2001; *Spiegelman et al.*, 2001; *Spiegelman and Kelemen*, 2003; *Katz et al.*, 2004). This overestimates the possibilities of matrix compaction at low porosity. Porosity dependent parameters have been explicitly accounted for in other papers (e.g., *Fowler*, 1985; *Connolly and Podladchikov*, 1998; *Schmeling*, 2000; *Bercovici et al.*, 2001a; *Ricard et al.*, 2001; *Rabinowicz et al.*, 2002). The melting rates are sometimes imposed instead of solving the energy equation (e.g.,

Turcotte and Morgan, 1992) or solved for assuming univariant transition (*Fowler*, 1989; *Sramek et al.*, 2006). Surface tension is added in *Riley et al.* (1990) and (*Hier-Majumder et al.*, 2006). Similar equations have also been used to describe the interaction between iron and silicates near the CMB (*Buffett et al.*, 2000), the formation of dendrites (*Poirier*, 1991) or the compaction of lava flows (*Massol et al.*, 2001).

6 Specifics of Earth’s mantle convection

In this last section we discuss various aspects of large-scale mantle convection. We leave the problems of partial melting to Chapter 9 and Chapter 11 of this Treatise. 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 (see also the books by *Schubert et al.* (2001) and *Davies* (1999)).

6.1 A mantle without inertia

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

$$\begin{aligned}
 \nabla \cdot (\bar{\rho}\mathbf{v}) &= 0, \\
 -\nabla P + \nabla \cdot \boldsymbol{\tau} + \delta\rho\mathbf{g} + \bar{\rho}\delta\mathbf{g} &= 0, \\
 \nabla^2\psi &= 4\pi G\delta\rho, \\
 \delta\mathbf{g} &= -\nabla\psi,
 \end{aligned}
 \tag{243}$$

in agreement with (125). In this set of equations we kept the self-gravitation term as appropriate at long wavelengths. If the internal loads $\bar{\rho}$ are known, the flow can be computed independently of the temperature equation. This time-independent system has been used by many authors to try to infer the mantle properties.

6.1.1 Dynamic models

The system of equations (243) can be solved analytically for a depth-dependent viscosity, when variables are expressed on the basis of spherical harmonics (see *Hager and Clayton* (1989) and also Chapter 5 and Chapter 8). 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...) can be expressed on the basis of spherical harmonics with

components O_{lm} . Through (243), they are related to the spherical harmonics components of the internal density variations $\delta\rho_{lm}(r)$ by various degree-dependent Green's functions (Green, 1793-1841)

$$O_{lm} = \int G_l^O(r) \delta\rho_{lm}(r) dr \quad (244)$$

(see Chapter 5 for analytical details). The Green's functions $G_l^O(r)$ can be computed from the averaged density and viscosity profiles.

Before seismic imaging gave us a proxy of the 3D 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 the Earth gravity field (or to the geoid, proportional to ψ), to the lithospheric stress regime or to the topography (*Runcorn, 1964; Parsons and Daly, 1983; Lago and Rabinowicz, 1984; Richards and Hager, 1984; Ricard et al., 1984*).

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 equals at first order the mass 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 ≥ 30 viscosity increase around the upper-lower mantle interface (*Lago and Rabinowicz, 1984; Hager et al., 1985*). 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.

6.1.2 Mantle flow and post-glacial 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 et al.*, 1985; *Hager and Clayton*, 1989; *King and Hager*, 1994). 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).

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 post glacial rebound (*Peltier*, 1996). The various approaches (time-dependent for the post-glacial 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 the presence of a kilometeric surface topography, induced by mantle convection and called dynamic topography (by opposition with the isostatic topography related to crustal and lithospheric density variations). Its direct observation, from the Earth's topography corrected from isostatic crustal contributions, is difficult and remains controversial (e.g., *Colin and Fleitout*, 1990; *Kido and Seno*, 1994; *Lestunff and Ricard*, 1995; *Lithgow-Bertelloni and Silver*, 1998).

6.1.3 Time dependent models

The thermal diffusion in the mantle is so slow that even over 100-200 myrs it can be neglected in some long wavelength global models. The equations (243) can thus be solved by imposing the known paleo-plate velocities at the surface and advect the mass anomalies with the flow without solving explicitly the energy equation. This forced-convection approach has shown that the deep mantle structure is mostly inherited from the Cenozoic and Mesozoic plate motion (*Richards and Engebretson*, 1992; *Lithgow-Bertelloni and Richards*, 1998). From plate paleo-slab reconstructions only, a density models 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 hotspot fixity (*Richards*, 1991; *Steinberger and O'Connell*, 1998), the sea level changes (*Lithgow-Bertelloni and Gurnis*, 1997) or the polar wander of the Earth (*Spada*

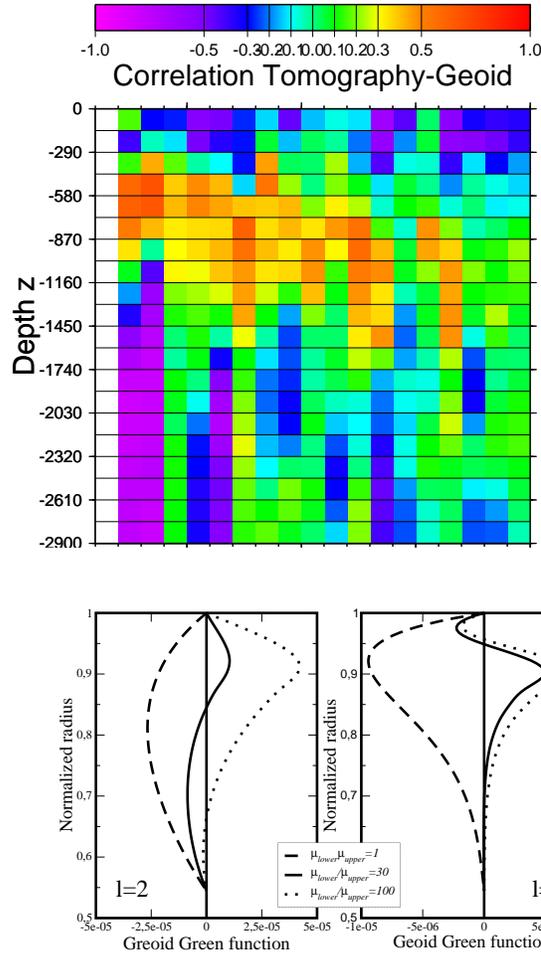


Figure 6: Correlations between gravity and the synthetic tomographic Smean model (*Becker and Boschi, 2002*) as a function of degree l and normalized radius (top). The seismic velocities, proxy of the density variations, are positively correlated with gravity around the upper-lower mantle interface (warm 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 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, 10 times more viscous than the upper mantle.

et al., 1992; *Richards et al.*, 1997).

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 symmetric 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 mid-depth 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. 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 is rising passively, i.e., 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 during 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 (hotspots). 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 primitive mantle composition (*McDonough and Sun*, 1995; *Rudnick and Fountain*, 1995) imply that about 50% of the surface heat flux is due to radioactive decay and 50% to mantle and core cooling. Generally

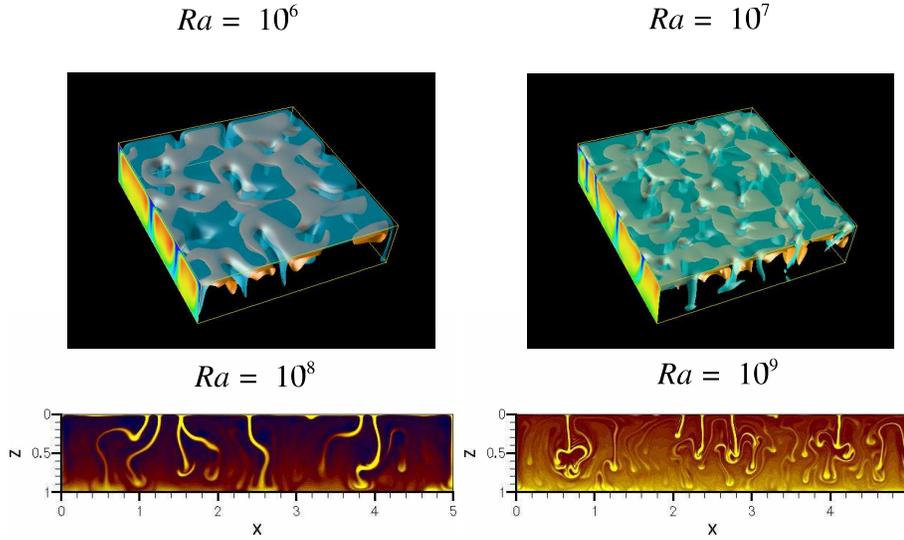


Figure 7: Convection patterns of a fluid entirely heated from inside at Rayleigh number 10^6 , 10^7 , 10^8 , 10^9 (simulations ran by Fabien Dubuffet). Cold finger-like instabilities are sinking from the top boundary layer, and spread on the bottom boundary layer. No active upwellings are present (compare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

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 order of 80% radioactive and 20% cooling (see Chapter 2). The basal heat flux at the CMB represents the core cooling component, part of the total cooling rate of the Earth. The secular cooling and the presence of internal sources tends 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 hotspots).

6.3 A complex rheology

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

$$\mu = AI_2^n d^{-m} \exp \frac{E^* + PV^*}{RT}, \quad (245)$$

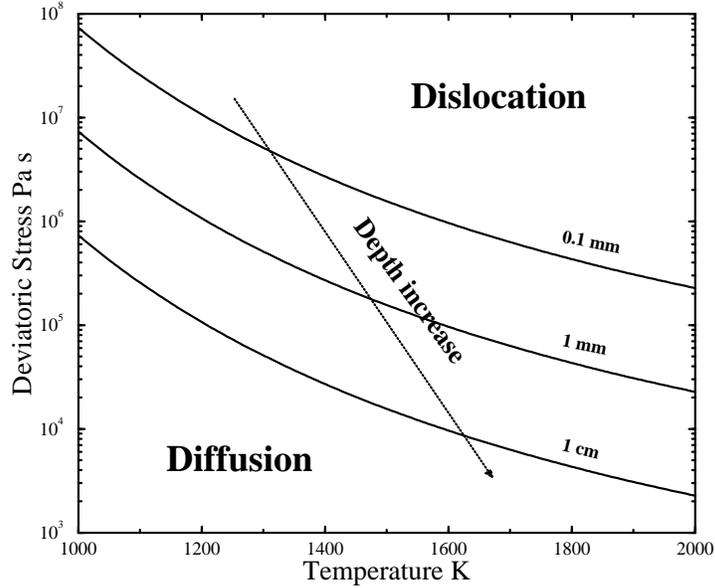


Figure 8: Creep regime map for dry olivine. High deviatoric stress or temperature favor a dislocation mechanism. A decrease in the grain size favors diffusion. In the upper mantle the stress and temperature conditions tend to bring the creep regime from dislocation to diffusion at depth ($n = 2.5$, $m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

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 (Weertman and Weertman, 1975; Ranalli, 1995). 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, 1 cm) the stress rate at which the strain rate predicted for the dislocation and diffusion mechanisms are the same (see (104) and (103)). The data corresponds to dry upper mantle (Karato and Wu, 1993). Low stress and temperature favor diffusion creep while high stress and high temperature favor dislocation creep. Below the lithosphere, in the upper mantle or at least in its shallowest part, non-linear creep is likely to occur. As depth increase, the decrease in the average deviatoric stress favors a diffusive regime with a Newtonian viscosity. The observation of rheological parameters at lower mantle conditions are more difficult but the lower mantle should mostly be in diffusive linear 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 Christensen, 1993*). Near the surface, viscosity variations are present, at least between continental and oceanic lithosphere (*Ricard et al., 1991; Cadek and Fleitout, 2003*). The gain in fitting the Earth's gravity or post glacial rebound data with 3D 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*).

6.3.1 Temperature dependence of viscosity

The 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 to 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 $T \sim 1600$ K). The effect of temperature dependence of viscosity on the planform of convection was recognized experimentally using oils or syrups (*Booker, 1976; Richter, 1978; Nataf and Richter, 1982*). 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, μ_{max}/μ_{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 (*Parmentier et al., 1976; Christensen, 1984b; Tackley et al., 1993; Trompert and Hansen, 1998b; Kameyama and Ogawa, 2000*). 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, eases 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.

Various modelers have thus tried to use T -dependent rheologies but have imposed a plate-like 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 (*Kincaid and Olson, 1987; Guilloufrottier et al., 1995*).

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. On the contrary, a too slowly convecting planet would not extract its radioactive energy, and would heat up until the viscosity is reduced sufficiently (see also Chapter 2). 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.

6.3.2 Depth-dependence of viscosity

The activation volume of the viscosity is typically around $10^{-5} \text{ m}^3 \text{ mol}^{-1}$. 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. The viscosity increase by a factor 30 to 100 suggested by geodynamics (see paragraph 6.1 and Chapter 7) is probably a constrain as robust as what could be deduced from mineralogic experiments.

The effect of a depth dependent viscosity on the planform of convection has been studied by e.g., *Gurnis and Davies (1986); Cserepes (1993) or 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 hotspots. 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 6.5.3) favors the relative hotspot 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 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)). This thickening is observed in tomographic models (*van der Hilst et al.*, 1997) and can be inferred from geoid modeling (*Ricard et al.*, 1993a).

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

At large Rayleigh number, however, non-linear convection becomes more unstable (*Malevsky and Yuen*, 1992) and the combination of non-linear 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).

6.3.4 Grain size dependence of viscosity

The factor A of the viscosity law (245), can also be strongly variable and it is for example a function of the grain size d with $A \propto d^m$ and m of order 3 in the diffusion regime (*Karato et al.*, 1986) (when the rheology is linear in terms of stress, it becomes non-linear in terms of grain size). There is a clear potential feed back interaction between deformation, grain size reduction by dynamic recrystallization, viscosity reduction and further localization (*Jaroslow et al.*, 1996). Grain size reduction is balanced by Ostwald ripening (e.g., the fact that surface energy drives diffusion from small grains to larger grains) which provides a healing mechanism to increase on the long term the average size of grains (*Hillert*, 1965). This increase of the average size of grains is in fact obtained by the decrease in size of small grains by dissolution into larger ones.

A grain size-dependent viscosity has been introduced into geodynamic models (e.g., *Braun et al.*, 1999; *Kameyama et al.*, 1997). The effect is potentially important in the mantle and even more important in the lithosphere. However, up to now, the basic physics remains problematic since the general conservation equations and the empirical laws of grain size evolution are not self consistent. The changes in surface energy cannot be easily accounted for in the total energy conservation and the experimental laws of grain reduction (in the dislocation regime) and viscosity reduction and healing (in the diffusion regime) have been obtained in mutually exclusive conditions.

6.4 Importance of sphericity

An obvious difference between the convection planform in a planet and in an experimental tank is due to the sphericity of the former. In the case of purely heated from below 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 4 times smaller than the top surface, this implies a 4 times larger thermal gradient through the bottom boundary layer than across the lithosphere. A bottom boundary layer thicker than the top 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 6.2) or T -dependent viscosity (see 6.3). Although numerically more difficult to handle, spherical convection models are more and more common (*Bercovici et al.*, 1989b; *Tackley et al.*, 1993; *Bunge et al.*, 1997; *Zhong et al.*, 2000). Figure 9 is an example of the complex patterns that can be obtained with convection in a sphere.

6.5 Other depth-dependent parameters

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

$$\alpha = \frac{\alpha_0}{(\rho/\rho_0)^{n-1+q} + \alpha_0(T - T_0)}. \quad (246)$$

It decreases with both temperature and density, and thus with depth. The expansivity varies from $\sim 4 \times 10^{-5} \text{ K}^{-1}$ near the surface to $\sim 8 \times 10^{-6} \text{ K}^{-1}$ near the CMB (*Chopelas and Boehler*, 1992). This diminishes the buoyancy forces

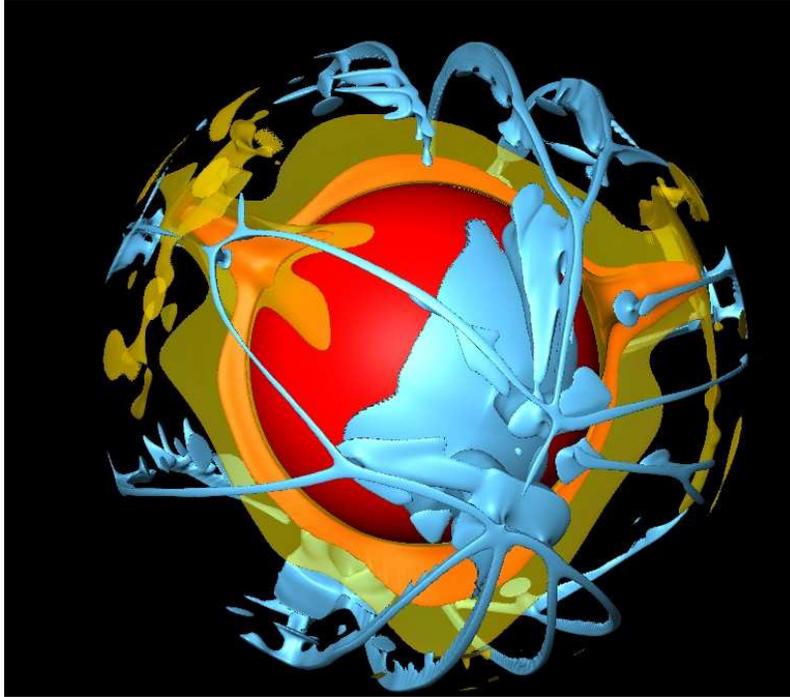


Figure 9: Isothermal surfaces in a spherical convection model (simulations by Marc Monnereau, cold=blue, yellow and orange=hot). Cold sheet like downwellings are sinking from the surface. Large-scale upwellings are reaching the upper-lower mantle interface. The presence of a phase change (see below 6.6), impedes the large hotspot penetration but generates secondary smaller plumes in the upper mantle.

and slows down the 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 (*Schmelting et al.*, 2003).

6.5.2 Increase in average density with depth

To take into account the depth dependence of density, the Boussinesq approximation has to be replaced by the anelastic approximation. The investigations have been carried out by *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 4.3.3), 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 the order of magnitude equations (120) as real equalities) 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 α with depth insufficiently balanced by the density increase, see (118)). 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 hotspots (*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 mid-mantle. Locally, viscous dissipation and adiabatic heatings can be larger than radiogenic heat production although integrated over the whole mantle and averaged over time, the adiabatic and dissipative sources cancel out (see (57)).

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 $\sim 2-3$. 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. It also broadens the thermal anomalies, increases the average mantle temperature and thins the bottom boundary layer (*Dubuffet et al.*, 1999).

6.6 Thermo-chemical convection

Except in section 5, a simple negative relationship was assumed between density variations and temperature variations, through the thermal expansivity, $\Delta\rho = -\alpha\Delta T$. However, in the mantle several sources of density anomalies are present. 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, 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

$$\Delta\rho = \left(\frac{\partial\rho}{\partial T}\right)_{\phi} \Delta T + \left(\frac{\partial\rho}{\partial\phi}\right)_T \left(\frac{\partial\phi}{\partial T}\right)_{\chi} \Delta T + \left(\frac{\partial\rho}{\partial\phi}\right)_T \left(\frac{\partial\phi}{\partial\chi}\right)_T \Delta\chi \quad (247)$$

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 term is responsible for a rise of the 410 km depth interface and deepens the 660 km depth interface in the presence of cold downwellings (*Irfune 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).

The effect of the second term has been rather well studied (*Schubert et al.*, 1975; *Christensen and Yuen*, 1984; *Machetel and Weber*, 1991; *Peltier and Solheim*, 1992; *Tackley et al.*, 1993; *Tackley*, 1995; *Christensen*, 1996). 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. The existence of latent heat release during phase change (see (158)) is a secondary and minor effect. The recent discovery of a phase transformation in the deep lower mantle (*Murakami et al.*, 2004) (the post-perovskite phases) suggests that part of the complexities of the D'' layer are related to the interaction between a phase change and the hot boundary layer of the mantle (*Nakagawa and Tackley*, 2006).

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 depth 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 nineties. 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 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.

The last contribution to the density anomalies are related to variations in chemical composition. There are indications of large-scale depth and lateral variations of Fe or Si contents in the mantle (*Bolton and Masters*, 2001; *Saltzer et al.*, 2004). A large well-documented elemental differentiation is between the oceanic crust (poor in Mg, rich in Al and Si) and the mantle. The density differences between a normal pyrolitic mantle and an oceanic basalt, both at lower mantle pressure (and appropriate mineralogical phases) is larger than the thermal density variations that can be expected in the mantle. The oceanic crust in its high pressure eclogitic facies, is ~ 5% denser than the average mantle density in most of the mantle except in the shallowest 100 kilometers 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). Since the thermal expansivity decreases with depth the compositional density variations, though small, remain important in the lower mantle. 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 by e.g., *Tackley* (2000) and *Davies* (2002). These compositional pyramids may anchor the hotspots (*Jellinek and Manga*, 2002; *Davaille et al.*, 2002). The presence of a petrologically dense component of the source of hotspots also seems necessary to explain their excess temperature (*Farnetani*, 1997).

Not only present-day subductions can generate compositional anomalies in the mantle. Geochemists have often argued for a deep layer of primitive material. This layer should be intrinsically denser to resist entrainment by convection. The stability of such a layer has been discussed by various authors (*Davaille*, 1999; *Kellogg et al.*, 1999; *Tackley and Xie*, 2002; *LeBars and Davaille*, 2002; *Samuel and Farnetani*, 2003). Depending of the ratio between the intrinsic density increase to the thermal density variations, a deep layer may be stable, slowly eroded, may undergo large oscillating motions (the so-called lava-lamp mode of thermochemical convection), or be rapidly mixed. Numerical simulations of thermo-chemical convection are certainly going to replace the thermal convection models in the next years. They will help to bridge the gap between geochemical observations and convection modeling (*Coltice and Ricard*, 1999; *van Keken et al.*, 2002).

6.7 A complex lithosphere: plates and continents

The lithosphere is part of the convection cell and plate tectonics and mantle convection cannot be separated. The role and specifics of the lithosphere are complex and are more deeply discussed in Chapter 8.

The fact that the cold lithosphere is much more viscous and concentrates most of the mass heterogeneities of the mantle, makes it behaving by some aspect as 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.

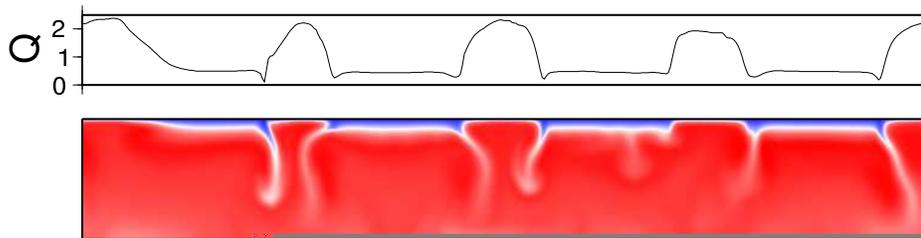
The Earth's mantle is certainly not controlled by surface tension. Marangoni convection, strictly speaking has nothing to do with mantle convection. However the equations of thermal convection with cooling from the top and with a highly viscous lithosphere can be shown to be mathematically close (through a change of variables) to those of Marangoni convection (*Lemery et al.*, 2000). This analogy has been advocated as a "top-down" view of the mantle dynamics (*Anderson*, 2001). More 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*) belong to the same approach.

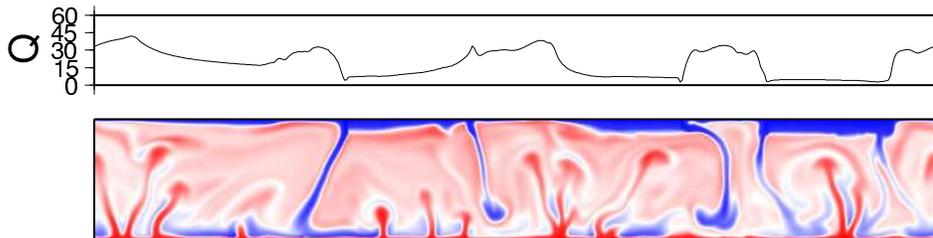
Due to the complexities of the lithosphere properties, the boundary condition at the surface of the Earth is far from being a uniform free-slip condition. Both continents and tectonic plates impose their own wavelengths and specific boundary conditions on the underlying convecting asthenosphere. Of course 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 piece-wise 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 (*Gurnis and Hager, 1988; Grigné and Labrosse, 2001*) 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 10). 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 non-linear rheologies have failed. Although in 2D some successes can be obtained in localizing deformation in plate-like domains, (*Schmeling and Jacoby, 1981; Weinstein and Olson, 1992; Weinstein, 1996*), they are obtained with stress exponents (e.g., $n \geq 7$) that are larger than what can be expected from laboratory experiments ($n \sim 2$). The problems are however worst in 3D. 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 plate-like surface boundary conditions. These studies have been performed in 2D and 3D (*Ricard and Vigny, 1989; Gable et al., 1991; King et al., 1992; Monnereau and Quéré, 2001*). 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 11 illustrates the organizing effect of plates in spherical, internally heated compressible convection with depth dependent viscosity (*Bunge and Richards, 1996*).



Internally heated



Bottom heated

Figure 10: Convection patterns in the presence of 4 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, hotspots tend to be preferentially anchored below continents where they bring an excess heat. This tends to reduce the surface heat flux variations.

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., *Moresi and Solomatov, 1998; Auth et al., 2003*). The existence of plates seems also to require the existence of a weak sub-lithospheric asthenosphere (*Richards et al., 2001*). In the last years, the first successes in computing 3D models that spontaneously organize their top boundary layer into plates have been reached (*Tackley, 1998; Trompert and Hansen, 1998a; Stein et al., 2004*). Although the topological characteristics of the predicted plates and their time evolution may be still far from the observed characteristics of plate tectonics, and often too episodic (stagnant-lid convection punctuated by plate-like events), a very important breakthrough has been made by modelers (see Figure 12).

The Earth's plate boundaries keep the memory of their weakness over geological times (*Gurnis et al., 2000*). This implies that the rheological properties cannot be a simple time-independent function of stress or temperature but has 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*).

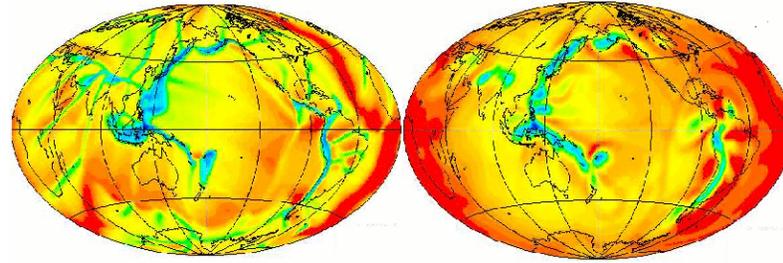
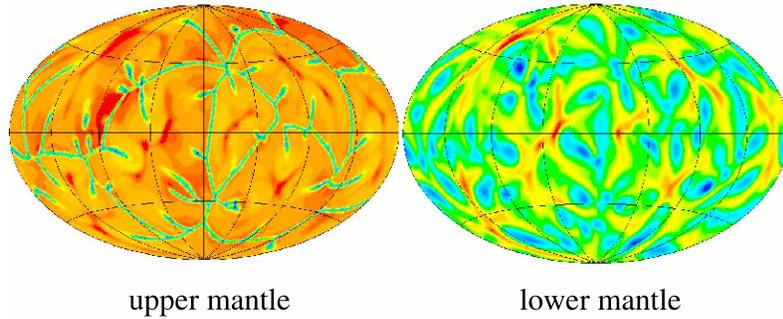
Acknowledgements:

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

References

- Anderson, D. L. (2001), Geophysics - Top-down tectonics?, *Science*, 293, 2016–2018.
- Anderson, O. L. (1979), Evidence supporting the approximation $\gamma\rho = \text{const}$ for the gruneisen parameter of the earth lower mantle, *J. Geophys. Res.*, 84, 3537–3542.
- Auth, C., D. Bercovici, and U. R. Christensen (2003), Two-dimensional convection with a self-lubricating, simple-damage rheology, *Geophys. J. Int.*, 154, 783–800.

Free slip surface



Imposed plate motion

Figure 11: 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, 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, 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).

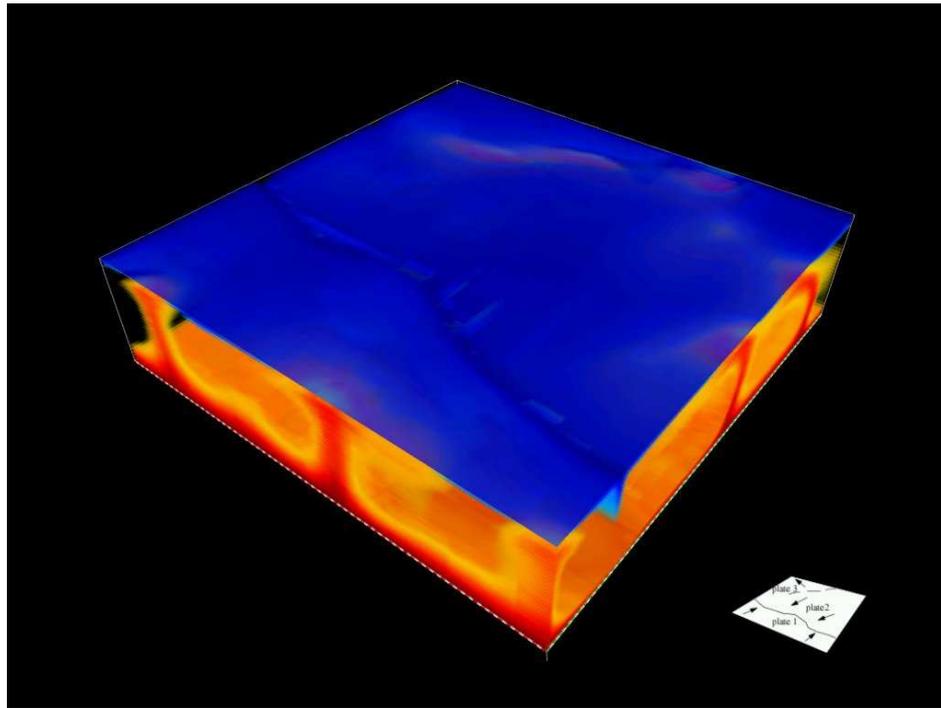


Figure 12: Convection models with self-coherent plate generation (*Stein et al., 2004*). Snapshot of the temperature field for a model calculation in a box of aspect ratio 4. The viscosity is temperature, pressure and stress dependent. The flow pattern reveals cylindrical upwellings and sheet-like downflow. Three plates have formed sketched in the small plot.

- Badro, J., J. P. Rueff, G. Vanko, G. Monaco, G. Fiquet, and F. Guyot (2004), Electronic transitions in perovskite: Possible nonconvecting layers in the lower mantle, *Science*, *305*, 383–386.
- Bai, W. M., C. Vigny, Y. Ricard, and C. Froidevaux (1992), On the origin of deviatoric stresses in the lithosphere, *J. Geophys. Res.*, *97*, 11,729–11,737.
- Balachandar, S., D. A. Yuen, and D. Reuteler (1992), Time-dependent 3-dimensional compressible convection with depth-dependent properties, *Geophys. Res. Lett.*, *19*, 2247–2250.
- Balachandar, S., D. A. Yuen, and D. Reuteler (1993), Viscous and adiabatic heating effects in 3-dimensional compressible convection at infinite prandtl number, *Phys. Fluids A-fluid Dynamics*, *5*, 2938–2945.
- Barenblatt, G. I. (1996), *Scaling*, Cambridge University Press.
- Batchelor (1967), *An Introduction to Fluid Dynamics*, Cambridge University Press.
- Beaumont, C. (1978), Evolution of sedimentary basins on a viscoelastic lithosphere - theory and examples, *Geophys. J. Royal Astron. Soc.*, *55*, 471–497.
- Becker, T. W., and L. Boschi (2002), A comparison of tomographic and geodynamic mantle models, *Geochem. Geophys. Geosyst.*, *3*, 2001GC000,168.
- Bercovici, D., and S. Karato (2003), Whole-mantle convection and the transition-zone water filter, *Nature*, *425*, 39–44.
- Bercovici, D., and Y. Ricard (2003), Energetics of a two-phase model of lithospheric damage, shear localization and plate-boundary formation, *Geophys. J. Int.*, *152*, 581–596.
- Bercovici, D., and Y. Ricard (2005), Tectonic plate generation and two-phase damage: Void growth versus grain size reduction, *J. Geophys. Res.*, *110*.
- Bercovici, D., G. Schubert, and G. A. Glatzmaier (1989a), Three dimensional spherical models of convection in the earth's mantle, *Science*, *244*, 950–955.
- Bercovici, D., G. Schubert, and G. A. Glatzmaier (1989b), 3-dimensional spherical-models of convection in the earths mantle, *Science*, *244*, 950–955.
- Bercovici, D., G. Schubert, and G. A. Glatzmaier (1992), 3-dimensional convection of an infinite-prandtl-number compressible fluid in a basally heated spherical-shell, *J. Fluid Dyn.*, *239*, 683–719.

- Bercovici, D., Y. Ricard, and M. Richards (2000), *The relation between mantle dynamics and plate tectonics: A primer*, in: *History and Dynamics of Global Plate Motions*, *Geophys. Monogr. Ser.*, vol. 121, 5–46 pp., M.A. Richards, R. Gordon, and R. van der Hilst, AGU, Washington D.C.
- Bercovici, D., Y. Ricard, and G. Schubert (2001a), A two-phase model of compaction and damage, 1. general theory, *J. Geophys. Res.*, *106*, 8887–8906.
- Bercovici, D., Y. Ricard, and G. Schubert (2001b), A two-phase model for compaction and damage 3. Applications to shear localization and plate boundary formation, *J. Geophys. Res.*, *106*, 8925–8939.
- Besse, J., and V. Courtillot (1991), Revised and synthetic apparent polar wander paths of the african, eurasian, north-america and indian plates, and true polar wander since 200 ma, *J. Geophys. Res.*, *96*, 4029–4050.
- Birch, F. (1952), Elasticity and constitution of the earth's interior, *J. Geophys. Res.*, *57*, 227–286.
- Bolton, H., and G. Masters (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, *J. Geophys. Res.*, *106*, 13,527–13,540.
- Booker, J. R. (1976), Thermal-convection with strongly temperature-dependent viscosity, *J. Fluid Dyn.*, *76*, 741–754.
- Braun, J., J. Chery, A. Poliakov, D. Mainprice, A. Vauchez, A. Tomassi, and M. Daignieres (1999), A simple parameterization of strain localization in the ductile regime due to grain size reduction: A case study for olivine, *J. Geophys. Res.*, *104*, 25,167–25,181.
- Buffett, B. A., E. J. Garnero, and R. Jeanloz (2000), Sediments at the top of Earth's core, *Science*, *290*, 1338–1342.
- Bullen, K. E. (1940), The problem of the earth's density variation, *Seismol. Soc. Am. Bull.*, *30*, 235–250.
- Bunge, H. P., and M. A. Richards (1996), The origin of large scale structure in mantle convection: Effects of plate motions and viscosity stratification, *Geophys. Res. Lett.*, *23*, 2987–2990.
- Bunge, H. P., M. A. Richards, and J. R. Baumgardner (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, *J. Geophys. Res.*, *102*, 11,991–12,007.

- Bunge, H. P., Y. Ricard, and J. Matas (2001), Non-adiabaticity in mantle convection, *Geophys. Res. Lett.*, *28*, 879–882.
- Cadek, O., and L. Fleitout (2003), Effect of lateral viscosity variations in the top 300 km on the geoid and dynamic topography, *Geophys. J. Int.*, *152*, 566–580.
- Choblet, G., and E. M. Parmentier (2001), Mantle upwelling and melting beneath slow spreading centers: effects of variable rheology and melt productivity, *Earth Planet. Sci. Lett.*, *184*, 589–604.
- Chopelas, A., and R. Boehler (1992), Thermal expansivity in the lower mantle, *Geophys. Res. Lett.*, *19*, 1983–1986.
- Christensen, U. (1984a), Convection with pressure-dependent and temperature-dependent non-newtonian rheology, *Geophys. J. Royal Astron. Soc.*, *77*, 343–384.
- Christensen, U. R. (1984b), Heat-transport by variable viscosity convection and implications for the earths thermal evolution, *Phys. Earth Planet. Inter.*, *35*, 264–282.
- Christensen, U. R. (1996), The influence of trench migration on slab penetration into the lower mantle, *Earth Planetary Science Lett.*, *140*, 27–39.
- Christensen, U. R. (1997a), Some geodynamical effects of anisotropic viscosity, *Geophys. J. Royal Astron. Soc.*, *91*, 711–736.
- Christensen, U. R. (1997b), Influence of chemical buoyancy on the dynamics of slabs in the transition zone, *J. Geophys. Res.*, *102*, 22,435–22,443.
- Christensen, U. R., and H. Harder (1991), 3-d convection with variable viscosity, *Geophys. J. Int.*, *104*, 213–226.
- Christensen, U. R., and A. Hofmann (1994), Segregation of subducted oceanic crust in the convecting mantle, *J. Geophys. Res.*, *99*, 19,867–19,884.
- Christensen, U. R., and D. A. Yuen (1984), The interaction of a subducting lithospheric slab with a chemical or phase-boundary, *J. Geophys. Res.*, *89*, 4389–4402.
- Colin, P., and L. Fleitout (1990), Topography of the ocean-floor - thermal evolution of the lithosphere and interaction of deep mantle heterogeneities with the lithosphere, *Geophys. Res. Lett.*, *17*, 1961–1964.
- Coltice, N., and Y. Ricard (1999), Geochemical observations and one layer mantle convection, *Earth Planetary Science Lett.*, *174*, 125–137.
- Connolly, J. A. D., and Y. Y. Podladchikov (1998), Compaction-driven fluid flow in viscoelastic rock, *Geodinamica Acta*, *11*, 55–84.

- Corrieu, V., C. Thoraval, and Y. Ricard (1995), Mantle dynamics and geoid green-functions, *Geophys. J. Int.*, *120*, 516–523.
- Cserepes, L. (1993), Effect of depth-dependent viscosity on the pattern of mantle convection, *Geophys. Res. Lett.*, *20*, 2091–2094.
- Davaille, A. (1999), Two-layer thermal convection in viscous fluids, *J. Fluid Dyn.*, *379*, 223–253.
- Davaille, A., and C. Jaupart (1993), Transient high-rayleigh-number thermal-convection with large viscosity variations, *J. Fluid Dyn.*, *253*, 141–166.
- Davaille, A., F. Girard, and M. LeBars (2002), How to anchor hotspots in a convecting mantle?, *Earth Planet. Sci. Lett.*, *203*, 621–634.
- Davies, G. F. (1980), Thermal histories of convective earth models and constraints on radiogenic heat-production in the earth, *J. Geophys. Res.*, *85*, 2517–2530.
- Davies, G. F. (1999), *Davies, G. F., 1999, Dynamic Earth Plates, Plumes and Mantle Convection*, Cambridge University Press, 458 p.
- Davies, G. F. (2002), Stirring geochemistry in mantle convection models with stiff plates and slabs, *Geochimica Et Cosmochimica Acta*, *66*, 3125–3142.
- de Groot, S. R., and P. Mazur (1984), *Non-Equilibrium Thermodynamics*, Dover Publications, New York.
- Doin, M. P., L. Fleitout, and U. Christensen (1997), Mantle convection and stability of depleted and undepleted continental lithosphere, *J. Geophys. Res.*, *102*, 2771–2787.
- Dombre, T., U. Frisch, J. M. Greene, M. Henon, A. Mehr, and A. M. Soward (1986), Chaotic streamlines in the abc flows, *J. Fluid Dyn.*, *167*, 353–391.
- Drew, D. A., and L. A. Segel (1971), Averaged equations for 2-phase flows, *Studies In Appl. Mathematics*, *50*, 205–220.
- Drury, M. R., and J. D. FitzGerald (1998), *The Earth's mantle: composition, structure, and evolution*, chap. Mantle rheology: insights from laboratory studies of deformation and phase transition, Cambridge University Press, Cambridge, England.
- Dubuffet, F., D. A. Yuen, and M. Rabinowicz (1999), Effects of a realistic mantle thermal conductivity on the patterns of 3-D convection, *Earth Planet. Sci. Lett.*, *171*, 401–409.
- Dubuffet, F., M. Rabinowicz, and M. Monnereau (2000), Multiple scales in mantle convection, *Earth Planet. Sci. Lett.*, *178*, 351–366.
- Dziewonski, A. M., and D. Anderson (1981), Preliminary reference earth model, *Phys. Earth Planet. Inter.*, *25*, 297–356.

- England, P., and D. McKenzie (1982), A thin viscous sheet model for continental deformation, *Geophys. J. Royal Astron. Soc.*, *70*, 295–321.
- Farnetani, C. G. (1997), Excess temperature of mantle plumes: The role of chemical stratification across d", *Geophys. Res. Lett.*, *24*, 1583–1586.
- Ferrachat, S., and Y. Ricard (1998), Regular vs, chaotic mantle mixing, *Earth Planet. Sci. Lett.*, *155*, 75–86.
- Forte, A. M., and W. R. Peltier (1987), Plate tectonics and aspherical earth structure – the importance of poloidal-toroidal coupling, *J. Geophys. Res.*, *92*, 3645–3679.
- Forte, A. M., and W. R. Peltier (1991), Viscous flow models of global geophysical observables .1. forward problems, *J. Geophys. Res.*, *96*, 20,131–20,159.
- Fowler, A. C. (1985), A mathematical model of magma transport in the asthenosphere, *Geophys. Astrophys. Fluid Dyn.*, *33*, 63–96.
- Fowler, A. C. (1989), Generation and creep of magma in the earth, *Siam J. On Appl. Mathematics*, *49*, 231–245.
- Gable, C. W., R. J. Oconnell, and B. J. Travis (1991), Convection in 3 dimensions with surface plates - generation of toroidal flow, *J. Geophys. Res.*, *96*, 8391–8405.
- Gasperini, P., G. D. Forno, and E. Boschi (2004), Linear or non-linear rheology in the Earth's mantle: the prevalence of power-law creep in the postglacial isostatic readjustment of laurentia, *Geophys. J. Int.*, *157*, 1297–1302.
- Glatzmaier, G. A. (1988), Numerical simulation of mantle convection: Time-dependent, three dimensional, compressible, spherical shell, *Geophys. Astrophys. Fluid Dyn.*, *43*, 223–264.
- Grigné, C., and S. Labrosse (2001), Effects of continents on Earth cooling: thermal blanketing and depletion in radioactive elements, *Geophys. Res. Lett.*, *28*, 2707–2710.
- Guillou, L., and C. Jaupart (1995), On the effect of continents on mantle convection, *J. Geophys. Res.*, *100*, 24,217–24,238.
- Guilloufrottier, L., J. Buttles, and P. Olson (1995), Laboratory experiments on the structure of subducted lithosphere, *Earth Planet. Sci. Lett.*, *133*, 19–34.
- Gurnis, M., and G. F. Davies (1986), Numerical study of high rayleigh number convection in a medium with depth-dependent viscosity, *Geophysical J. Royal Astronomical Soc.*, *85*, 523–541.
- Gurnis, M., and B. H. Hager (1988), Controls of the structure of subducted slabs, *Nature*, *335*, 317–321.

- Gurnis, M., L. Moresi, and R. D. Mueller (2000), *The History and Dynamics of Global Plate Motions*, Richards, M. A., Gordon, R., and van der Hilst, R. editors, chap. Models of mantle convection incorporating plate tectonics: The Australian region since the Cretaceous, pp. 211–238, American Geophysical Union, Washington D.C.
- Haase, R. (1990), *Thermodynamics of irreversible processes*, Dover Publications, New York.
- Hager, B. H., and R. W. Clayton (1989), *Constraints on the structure of mantle convection using seismic observation, flow models and the geoid*, in: *Mantle Convection, Plate Tectonics and Global Dynamics*, W.R. Peltier, Gordon and Breach Science Publishers, New York.
- Hager, B. H., and R. J. Oconnell (1979), Kinematic models of large-scale flow in the earths mantle, *J. Geophys. Res.*, *84*, 1031–1048.
- Hager, B. H., and R. J. O’Connell (1981), A simple global model of plate tectonics and mantle convection, *J. Geophys. Res.*, *86*, 4843–4867.
- Hager, B. H., R. W. Clayton, M. A. Richards, R. P. Comer, and A. M. Dziewonski (1985), Lower mantle heterogeneity, dynamic topography and the geoid, *Nature*, *313*, 541–546.
- Hansen, U., D. A. Yuen, S. E. Kroening, and T. B. Larsen (1993), Dynamic consequences of depth-dependent thermal expansivity and viscosity on mantle circulations and thermal structure, *Phys. Earth Planet. Inter.*, *77*, 205–223.
- Haskell, N. A. (1937), The viscosity of the asthenosphere, *Am. J. Sci.*
- Hewitt, J. M., D. P. McKenzie, and N. O. Weiss (1975), Dissipative heating in convective flows, *J. Fluid Dyn.*, *68*, 721–738.
- Hier-Majumder, S., Y. Ricard, and D. Bercovici (2006), Role of grain boundaries in magma migration and storage, *Earth Planet. Sci. Lett.*, *248*, 735–749.
- Hillert, M. (1965), On theory of normal and abnormal grain growth, *Acta Metallurgica*, *13*, 227–235.
- Hirth, G., and D. L. Kohlstedt (1996), Water in the oceanic upper mantle: implications for rheology, melt extraction and the evolution of the lithosphere, *Earth Planet. Sci. Lett.*, *144*, 93–108.
- Hofmeister, A. M. (1999), Mantle values of thermal conductivity and the geotherm from phonon lifetimes, *Science*, *283*, 1699–1706.
- Irifune, T., and A. Ringwood (1993), Phase transformations in subducted oceanic crust and buoyancy relationships at depths of 600-800 km in the mantle, *Earth Planet. Sci. Lett.*, *117*, 101–110.

- Irifune, T., and A. E. Ringwood (1987), Phase transformations in a harzburgite composition to 26 GPa: implications for dynamical behaviour of the subducting slab, *Earth Planet. Sci. Lett.*, *86*, 365–376.
- Jaroslow, G. E., G. Hirth, and H. J. B. Dick (1996), Abyssal peridotite mylonites: Implications for grain-size sensitive flow and strain localization in the oceanic lithosphere, *Tectonophysics*, *256*, 17–37.
- Jarvis, G. T., and D. P. Mckenzie (1980), Convection in a compressible fluid with infinite prandtl number, *J. Fluid Dyn.*, *96*, 515–583.
- Jaupart, C., and J. C. Mareschal (1999), The thermal structure and thickness of continental roots, *Lithos*, *48*, 93–114.
- Jeanloz, R., and S. Morris (1987), Is the mantle geotherm subadiabatic, *Geophys. Res. Lett.*, *14*, 335–338.
- Jellinek, A. M., and M. Manga (2002), The influence of a chemical boundary layer on the fixity, spacing and lifetime of mantle plumes, *Nature*, *418*, 760–763.
- Kameyama, M., and M. Ogawa (2000), Transitions in thermal convection with strongly temperature-dependent viscosity in a wide box, *Earth Planet. Sci. Lett.*, *180*, 355–367.
- Kameyama, M., D. A. Yuen, and H. Fujimoto (1997), The interaction of viscous heating with grain-size dependent rheology in the formation of localized slip zones, *Geophys. Res. Lett.*, *24*, 2523–2526.
- Karato, S. I. (1998), Seismic anisotropy in the deep mantle, boundary layers and the geometry of mantle convection, *Pure Appl. Geophys.*, *151*, 565–587.
- Karato, S. I., and P. Wu (1993), Rheology of the upper mantle: A synthesis, *Science*, *260*, 771–778.
- Karato, S. I., M. S. Paterson, and J. D. F. gerald (1986), Rheology of synthetic olivine aggregates - influence of grain-size and water, *J. Geophys. Res.*, *91*, 8151–8176.
- Katz, R. F., M. Spiegelman, and S. M. Carbotte (2004), Ridge migration, asthenospheric flow and the origin of magmatic segmentation in the global mid-ocean ridge system, *Geophys. Res. Lett.*, *31*.
- Kelemen, P. B., G. Hirth, N. Shimizu, M. Spiegelman, and H. J. B. Dick (1997), A review of melt migration processes in the adiabatically upwelling mantle beneath oceanic spreading ridges, *Philosophical Transactions Royal Soc. London Series A-mathematical Phys. Engineering Sciences*, *355*, 283–318.
- Kellogg, L. H., and D. L. Turcotte (1987), homogeneization of the mantle by convective mixing and diffusion, *Earth Planet. Sci. Lett.*, *81*, 371–378.

- Kellogg, L. H., B. H. Hager, and R. D. van der Hilst (1999), Compositional stratification in the deep mantle, *Science*, 283, 1881–1884.
- Kennett, B. L. N. (2001), *The seismic wavefield*, Cambridge University Press.
- Kido, M., and T. Seno (1994), Dynamic topography compared with residual depth anomalies in oceans and implications for age-depth curves, *Geophys. Res. Lett.*, 21, 717–720.
- Kincaid, C., and P. Olson (1987), An experimental-study of subduction and slab migration, *J. Geophys. Res.*, 92, 13,832–13,840.
- King, S. D., and B. H. Hager (1994), Subducted slabs and the geoid .1. numerical experiments with temperature-dependent viscosity, *J. Geophys. Res.*, 99, 19,843–19,852.
- King, S. D., C. W. Gable, and S. A. Weinstein (1992), Models of convection-driven tectonic plates - a comparison of methods and results, *Geophys. J. Int.*, 109, 481–487.
- Lago, B., and M. Rabinowicz (1984), Admittance for convection in a layered spherical shell, *Geophys. J. Royal Astron. Soc.*.
- Lambeck, K., and P. Johnston (1998), *The Earth's mantle: composition, structure, and evolution*, chap. The viscosity of the mantle, pp. 461–502, Cambridge University Press, Cambridge, England.
- Landau, L., and E. Lifchitz (1980), *An Introduction to Fluid Dynamics*, Pergamon Press, Oxford.
- Landau, L., and E. Lifchitz (2000), *Theory of elasticity*, Butterworth- Heinemann.
- Larsen, T. B., and D. A. Yuen (1997), Fast plumeheads: Temperature-dependent versus non-Newtonian rheology, *Geophys. Res. Lett.*, 24, 1995–1998.
- Larsen, T. B., A. V. Malevsky, D. A. Yuen, and J. L. Smedsmo (1993), Temperature-dependent newtonian and non-newtonian convection - implications for lithospheric processes, *Geophys. Res. Lett.*, 20, 2595–2598.
- LeBars, M., and A. Davaille (2002), Stability of thermal convection in two superimposed miscible viscous fluids, *J. Fluid Dyn.*, 471, 339–363.
- Lee, T., D. A. Papanastassiou, and G. L. Wasserburg (1976), Demonstration of ^{26}Mg excess in allende and evidence for al^{26} , *Geophys. Res. Lett.*, pp. 41–44.
- Lemery, C., Y. Ricard, and J. Sommeria (2000), A model for the emergence of thermal plumes in Rayleigh-Benard convection at infinite Prandtl number, *J. 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, *Geophys. J. Int.*, 134, 706–720.

- Lestunff, Y., and Y. Ricard (1995), Topography and geoid due to lithospheric mass anomalies, *Geophys. J. Int.*, *122*, 982–990.
- Lithgow-Bertelloni, C., and M. Gurnis (1997), Cenozoic subsidence and uplift of continents from time-varying dynamic topography, *Geology*, *25*, 735–738.
- Lithgow-Bertelloni, C., and J. H. Guynn (2004), Origin of the lithospheric stress field, *J. Geophys. Res.*, *109*, doi: 10.1029/2003JB002467.
- Lithgow-Bertelloni, C., and M. A. Richards (1998), Dynamics of cenozoic and mesozoic plate motions, *Reviews of Geophys.*, *36*, 27–78.
- Lithgow-Bertelloni, C., and P. G. Silver (1998), Dynamic topography, plate driving forces and the African superswell, *Nature*, *395*, 269–272.
- Machetel, P., and P. Weber (1991), Intermittent layered convection in a model mantle with an endothermic phase change at 670 km, *Nature*, *350*, 55–57.
- Malevsky, A. V., and D. A. Yuen (1992), Strongly chaotic non-newtonian mantle convection, *Geophysical Astrophysical Fluid Dynamics*, *65*, 149–171.
- Massol, H., C. Jaupart, and D. W. Pepper (2001), Ascent and decompression of viscous vesicular magma in a volcanic conduit, *J. Geophys. Res.*, *106*, 16,223–16,240.
- McDonough, W. F., and S. S. Sun (1995), The composition of the earth, *Chem. Geology*, *120*, 223–253.
- McKenzie, D. (1984), The generation and compaction of partially molten rock, *J. Petrol.*
- McNamara, A. K., S. I. Karato, and P. E. van Keken (2001), Localisation of dislocation creep in the lower mantle: Implications for the origin of seismic anisotropy, *Earth Planet. Sci. Lett.*, *191*, 85–99.
- Mitrovica, J. X., and A. M. Forte (1997), Radial profile of mantle viscosity: Results from the joint inversion of convection and postglacial rebound observables, *J. Geophys. Res.*, *102*, 2751–2769.
- Monnereau, M., and S. Quéré (2001), Spherical shell models of mantle convection with tectonic plates, *Earth Planet. Sci. Lett.*, *184*, 575 – 588).
- Moresi, L., and V. Solomatov (1998), Mantle convection with a brittle lithosphere: thoughts on the global tectonic styles of the Earth and venus, *Geophys. J. Int.*, *133*, 669–682.
- Muhlhaus, H. B., L. Moresi, and M. Cada (2004), Emergent anisotropy and flow alignment in viscous rocks, *Pure Appl. Geoph.*, *161*, 2451–2463.

- Munk, W. H., and G. J. F. MacDonald (1960), *The rotation of the Earth: A geophysical discussion*, Cambridge.
- Murakami, T., K. Hirose, K. Kawamura, N. Sata, and Y. Ohishi (2004), Post-perovskite phase transition in MgSiO₃, *Science*, *304*, 855–858.
- Murnaghan, F. D. (1951), *Finite deformation of an elastic solid*, John Wiley and Sons, New York.
- Nakagawa, T., and P. J. Tackley (2006), Three-dimensional structures and dynamics in the deep mantle: Effects of post-perovskite phase change and deep mantle layering, *Geophys. Res. Lett.*, *33*(L12S11, doi:1029/2006GL025719).
- Nataf, H. C., and F. M. Richter (1982), Convection experiments in fluids with highly temperature-dependent viscosity and the thermal evolution of the planets, *Phys. Earth Planet. Inter.*, *29*, 320–329.
- Nield, D. A. (1964), Surface tension and buoyancy effects in cellular convection, *J. Fluid Dyn.*, *19*, 341–352.
- Nye, J. F. (1953), The flow law of ice from measurements in glacier tunnels, laboratory experiments and the jungfrau firn borehole experiment, *Proc. Royal Soc. London Series A-mathematical Phys. Sciences*, *219*, 477–489.
- Ogawa, M., G. Schubert, and A. Zebib (1991), Numerical simulations of 3-dimensional thermal-convection in a fluid with strongly temperature-dependent viscosity, *J. Fluid Dyn.*, *233*, 299–328.
- Olson, P., D. A. Yuen, and D. Balsiger (1984), Mixing of passive heterogeneties by mantle convection, *J. Geophys. Res.*, *89*, 425–436.
- Ottino, J. M. (1989), *The kinematics of mixing: stretching, chaos and transport*, Cambridge University Press, New-York.
- Panasyuk, S., B. Hager, and A. Forte (1996), Understanding the effects of mantle compressibility on geoid kernels, *Geophys. J. Int.*, *124*, 121–133.
- Parmentier, E. M., D. L. Turcotte, and K. E. Torrance (1976), Studies of finite-amplitude non-newtonian thermal convection with application to convection in earths mantle, *J. Geophys. Res.*, *81*, 1839–1846.
- Parmentier, E. M., C. Sotin, and B. J. Travis (1994), Turbulent 3-d thermal-convection in an infinite prandtl number, volumetrically heated fluid - implications for mantle dynamics, *Geophys. J. Int.*, *116*, 241–251.
- Parsons, B., and S. Daly (1983), The relationship between surface topography, gravity anomaly and the temperature structure of convection, *J. Geophys. Res.*, *88*, 1129–1144.

- Peltier, W. R. (1989), *Mantle convection: plate tectonics and Global geodynamics*, chap. Mantle viscosity, pp. 389–478, Gordon and Breach Science Publishers, New York.
- Peltier, W. R. (1996), Mantle viscosity and ice-age ice sheet topography, *Science*, *273*, 1359–1364.
- Peltier, W. R., and L. P. Solheim (1992), Mantle phase transition and layered chaotic convection, *Geophys. Res. Lett.*, *19*, 321–324.
- Poirier, J. P. (1991), *Introduction to the physics of the Earth's interior*, Cambridge University Press.
- Poirier, J. P., and A. Tarantola (1998), A logarithmic equation of state, *Phys. Earth Planet. Inter.*, *109*, 1–8.
- Rabinowicz, M., Y. Ricard, and M. Gregoire (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 Planet. Sci. Lett.*, *203*, 205–220.
- Ranalli, G. (1995), *Rheology of the Earth*, Kluwer academic publishers.
- Ribe, N. M. (1985a), The deformation and compaction of partial molten zones, *Geophys. J. Royal Astron. Soc.*, *83*, 487–501.
- Ribe, N. M. (1985b), The generation and composition of partial melts in the earths mantle, *Earth Planet. Sci. Lett.*, *73*, 361–376.
- Ricard, Y., and D. Bercovici (2003), Two-phase damage theory and crustal rock failure: the theoretical 'void' limit, and the prediction of experimental data, *Geophys. J. Int.*, *155*, 1057–1064.
- Ricard, Y., and N. Coltice (2004), *Geophysical and geochemical models of mantle convection: Successes and future challenges*, *The state of the planet: Frontiers and challenges in geophysics*, vol. 150, 59-68 pp., Geophys. Monogr. Ser.
- Ricard, Y., and C. Vigny (1989), Mantle dynamics with induced plate-tectonics, *J. Geophys. Res.*, *94*, 17,543–17,559.
- Ricard, Y., L. Fleitout, and C. Froidevaux (1984), Geoid heights and lithospheric stresses for a dynamic Earth, *Ann. Geophysica*, *2*, 267–286.
- Ricard, Y., C. Vigny, and C. Froidevaux (1989), Mantle heterogeneities, geoid and plate motion: a monte carlo inversion, *J. Geophys. Res.*, *94*, 13,739–13,754.
- Ricard, Y., C. Doglioni, and R. Sabadini (1991), Differential rotation between lithosphere and mantle - a consequence of lateral mantle viscosity variations, *J. Geophys. Res.*, *96*, 8407–8415.

- Ricard, Y., M. Richards, C. Lithgow-Bertelloni, and Y. L. Stunff (1993a), A geodynamic model of mantle density heterogeneity, *J. Geophys. Res.*, *98*, 21,895–21,909.
- Ricard, Y., G. Spada, and R. Sabadini (1993b), Polar wander of a dynamic earth, *Geophys. J. Int.*, *113*, 284–298.
- Ricard, Y., D. Bercovici, and G. Schubert (2001), A two-phase model for compaction and damage 2. Applications to compaction, deformation, and the role of interfacial surface tension, *J. Geophys. Res.*, *106*, 8907–8924.
- Ricard, Y., E. Mattern, and J. Matas (2005), *Earth's Deep Mantle: Structure, Composition, and Evolution*, chap. Synthetic tomographic images of slabs from mineral physics, pp. 285–302, AGU monograph, van der Hilst, Bass, Matas and Trampert eds.
- Richards, M. (1991), *Hotspots and the case for a high viscosity lower mantle*, in *Glacial Isostasy, Sea-Level and Mantle Rheology*, R. Sabadini et al., eds, 571–587 pp., Kluwer Academic Publishers, Dordrecht.
- Richards, M. A., and D. C. Engebretson (1992), Large-scale mantle convection and the history of subduction, *Nature*, *355*, 437–440.
- Richards, M. A., and B. H. Hager (1984), Geoid anomalies in a dynamic earth, *J. Geophys. Res.*, *89*, 5987–6002.
- Richards, M. A., R. Y. C. Lithgow-Bertelloni, G. Spada, and R. Sabadini (1997), An explanation for earth's long-term rotational stability, *Science*, *275*, 372–375.
- Richards, M. A., W. S. Yang, J. R. Baumgardner, and H. P. Bunge (2001), Role of a low-viscosity zone in stabilizing plate tectonics: Implications for comparative terrestrial planetology, *Geochem. Geophys. Geosys.*, *2*.
- Richter, F. M. (1978), Experiments on the stability of convection rolls in fluids whose viscosity depends on temperature, *J. Fluid Dyn.*, *89*, 553–.
- Richter, F. M., and D. Mckenzie (1984), Dynamical models for melt segregation from a deformable matrix, *J. Geology*, *92*, 729–740.
- Ricolleau, A., G. Fiquet, J. Perillat, I. Daniel, N. Menguy, H. Cardon, A. Addad, C. Vanni, and N. Guignot (2004), The fate of subducted basaltic crust in the earth's lower mantle: an experimental petrological study, in *Eos Trans. AGU, Fall Meet. Suppl.*, vol. 85(47), abstract U33B-02.
- Riley, G. N., D. L. Kohlstedt, and F. M. Richter (1990), Melt migration in a silicate liquid-olivine system - an experimental test of compaction theory, *Geophys. Res. Lett.*, *17*, 2101–2104.

- Ringwood, A. E., and T. Irifune (1988), Nature of the 650-km seismic discontinuity - implications for mantle dynamics and differentiation, *Nature*, *331*, 131–136.
- Rudnick, R. L., and D. M. Fountain (1995), Nature and composition of the continental-crust - a lower crustal perspective, *Rev. Geophys.*, *33*, 267–309.
- Runcorn, S. K. (1964), Satellite gravity measurements and the laminar viscous flow model of the earth's mantle, *J. Geophys. Res.*.
- Sabadini, R., and D. A. Yuen (1989), Mantle stratification and long-term polar wander, *Nature*, *339*, 373–375.
- Saltzer, R. L., E. Stutzmann, and R. D. van der Hilst (2004), Poisson's ratio in the lower mantle beneath Alaska: Evidence for compositional heterogeneity, *J. Geophys. Res.*, *109*.
- Samuel, H., and C. G. Farnetani (2003), Thermochemical convection and helium concentrations in mantle plumes, *Earth Planet. Sci. Lett.*, *207*, 39–56.
- Schmalzl, J., G. A. Houseman, and U. Hansen (1996), Mixing in vigorous, time-dependent three-dimensional convection and application to Earth's mantle, *J. Geophys. Res.*, *101*, 21,847–21,858.
- Schmeling, H. (2000), *Physics and chemistry of partially molten rocks*, chap. Partial Melting and melt migration in a convecting mantle, Kluwer Academic Publisher, Dordrecht.
- Schmeling, H., and W. R. Jacoby (1981), On modeling the lithosphere in mantle convection with non-linear rheology, *J. Geophysics-zeitschrift Fur Geophysik*, *50*, 89–100.
- Schmeling, H., G. Marquart, and T. Ruedas (2003), Pressure- and temperature-dependent thermal expansivity and the effect on mantle convection and surface observables, *Geophys. J. Int.*, *154*, 224–229.
- Schubert, G., D. A. Yuen, and D. L. Turcotte (1975), Role of phase-transitions in a dynamic mantle, *Geophys. J. Royal Astron. Soc.*, *42*, 705–735.
- Schubert, G., D. Stevenson, and P. Cassen (1980), Whole planet cooling and the radiogenic heat-source contents of the earth and moon, *J. Geophys. Res.*, *85*, 2531–2538.
- Schubert, G., D. Turcotte, and P. Olson (2001), *Mantle convection in the Earth and planets*, Cambridge University Press.
- Scott, D. R., and D. J. Stevenson (1984), Magma solitons, *Geophys. Res. Lett.*, *11*, 9283–9296.
- Scott, D. R., and D. J. Stevenson (1986), Magma ascent by porous flow, *J. Geophys. Res.*, *91*, 9283–9296.

- Scott, D. R., and D. J. Stevenson (1989), A self-consistent model of melting, magma migration and buoyancy-driven circulation beneath mid-ocean ridges, *J. Geophys. Res.*, *94*, 2973–2988.
- Solomatov, V. S. (1995), Scaling of temperature-dependent and stress-dependent viscosity convection, *Phys. Fluids*, *7*, 266–274.
- Spada, G., Y. Ricard, and R. Sabadini (1992), Excitation of true polar wander by subduction, *Nature*, *360*, 452–454.
- Spear, F. S. (1993), *Metamorphic Phase Equilibria and Pressure-Temperature-Time Paths*, 799p, Mineralogical Society of America, Washington, D. C.
- Spiegelman, M., and P. B. Kelemen (2003), Extreme chemical variability as a consequence of channelized melt transport, *Geochem. Geophys. Geosys.*, *4*.
- Spiegelman, M., P. B. Kelemen, and E. Aharonov (2001), Causes and consequences of flow organization during melt transport: The reaction infiltration instability in compactible media, *J. Geophys. Res.*, *106*, 2061–2077.
- Sramek, O., Y. Ricard, and D. Bercovici (2006), Simultaneous melting and compaction in deformable two phase media, *Geophys. J. Int.*.
- Stacey, F. D. (1977), Application of thermodynamics to fundamental earth physics, *Geophys. Surveys*, *3*, 175–204.
- Stacey, F. D., and P. M. Davis (2004), High pressure equations of state with applications to the lower mantle and core, *Phys. Earth Planet. Inter.*, *142*, 137–184.
- Stein, C., J. Schmalzl, and U. Hansen (2004), The effect of rheological parameters on plate behaviour in a self-consistent model of mantle convection, *Phys. Earth Planet. Inter.*, *142*, 225–255.
- Steinberger, B., and R. J. O’Connell (1998), Advection of plumes in mantle flow: implications for hotspot motion, mantle viscosity and plume distribution, *Geophys. J. Int.*, *132*, 412–434.
- Stevenson, D. J. (1989), Spontaneous small-scale melt segregation in partial melts undergoing deformation, *Geophys. Res. Lett.*, *16*, 1067–1070.
- Tackley, P. (1995), On the penetration of an endothermic phase transition by upwellings and downwellings, *J. Geophys. Res.*, *100*, 15,477–15,488.
- Tackley, P. J. (1998), Self-consistent generation of tectonic plates in three-dimensional mantle convection, *Earth Planetary Science Lett.*, *157*, 9–22.
- Tackley, P. J. (2000), Mantle convection and plate tectonics: Toward an integrated physical and chemical theory, *Science*, *288*, 2002–2007.

- Tackley, P. J., and S. X. Xie (2002), The thermochemical structure and evolution of Earth's mantle: constraints and numerical models, *Philosophical Transactions Royal Soc. London Series A-mathematical Phys. Engineering Sciences*, *360*, 2593–2609.
- Tackley, P. J., D. J. Stevenson, G. A. Glatzmaier, and G. Schubert (1993), Effects of an endothermic phase-transition at 670 km depth in a spherical model of convection in the earths mantle, *Nature*, *361*, 699–704.
- Tao, W. C., and R. J. Oconnell (1993), Deformation of a weak subducted slab and variation of seismicity with depth, *Nature*, *361*, 626–628.
- Toth, J., and M. Gurnis (1998), Dynamics of subduction initiation at preexisting fault zones, *J. Geophys. Res.*, *103*, 18,053–18,067.
- Toussaint, V., P. Carriere, J. Scott, and J. N. Gence (2000), Spectral decay of a passive scalar in chaotic mixing, *Phys. Fluids*, *12*, 2834–2844.
- Tozer, D. C. (1972), The present thermal state of the terrestrial planets, *Phys. Earth Planet. Inter.*, *6*, 182–197.
- Tritton, D. J. (1988), *Physical fluid dynamics*, Oxford University Press.
- Trompert, R., and U. Hansen (1998a), Mantle convection simulations with rheologies that generate plate-like behaviour, *Nature*, *395*, 686–689.
- Trompert, R. A., and U. Hansen (1998b), On the rayleigh number dependence of convection with a strongly temperature-dependent viscosity, *Phys. Fluids*, *10*, 351–360.
- Turcotte, D. L., and J. P. Morgan (1992), *Mantle flow and melt generation at mid-ocean ridges*, chap. The physics of magma migration and mantle flow beneath a mid-ocean ridge, pp. 155–182, eds J. Phipps Morgan, D. K. Blackmann and J. M. Simpson, AGU New York.
- Turcotte, D. L., and G. Schubert (1982), *Geodynamics: applications of continuum physics to geological problems*, 1-450 pp., Wiley, New York.
- Urey, H. C. (1951), The origin and development of the earth and other terrestrial planets, *Geochim. Cosmochim. Acta*, *1*, 209–277.
- van der Hilst, R. D., S. Widiyantoro, and E. R. Engdahl (1997), Evidence for deep mantle circulation from global tomography, *Nature*, *386*, 578–584.
- van Keken, P., D. A. Yuen, and A. Vandenberg (1992), Pulsating diapiric flows - consequences of vertical variations in mantle creep laws, *Earth Planetary Science Lett.*, *112*, 179–194.

- van Keken, P. E., E. H. Hauri, and C. J. Ballentine (2002), Mantle mixing: The generation, preservation, and destruction of chemical heterogeneity, *Ann. Rev. Earth Planetary Sciences*, *30*, 493–525.
- Vinet, P., J. Ferrante, J. H. Rose, and J. R. Smith (1987), Compressibility of solids, *J. Geophys. Res.*, *92*, 9319–9325.
- Weertman, J., and J. R. Weertman (1975), High-temperature creep of rock and mantle viscosity, *Ann. Rev. Earth Planet. Sci.*, *3*, 293–315.
- Weinstein, S. A. (1996), Thermal convection in a cylindrical annulus with a non-Newtonian outer surface, *Pure Appl. Geophys.*, *146*, 551–572.
- Weinstein, S. A., and P. L. Olson (1992), Thermal-convection with non-newtonian plates, *Geophys. J. Int.*, *111*, 515–530.
- Williams, Q., and E. J. Garnero (1996), Seismic evidence for partial melt at the base of Earth's mantle, *Science*, *273*, 1528–1530.
- Woods, L. C. (1975), *The thermodynamics of fluid systems*, Clarendon Press, Oxford.
- Yuen, D., R. C. A. Sabadini, and P. Gasperini (1986), On transient rheology and glacial isostasy, *J. Geophys. Res.*, *91*, 1420–1438.
- Zhang, S. X., and U. Christensen (1993), Some effects of lateral viscosity variations on geoid and surface velocities induced by density anomalies in the mantle, *Geophys. J. Int.*, *114*, 531–547.
- Zhang, S. X., and D. A. Yuen (1996), Various influences on plumes and dynamics in time-dependent, compressible mantle convection in 3-D spherical shell, *Phys. Earth Planetary Interiors*, *94*, 241–267.
- Zhong, S. J., M. Gurnis, and L. Moresi (1996), Free surface formulation of mantle convection, part 1: basic theory and implication to plumes, *Geophys. J. Int.*, *127*, 708–718.
- Zhong, S. J., M. T. Zuber, L. Moresi, and M. Gurnis (2000), Role of temperature-dependent viscosity and surface plates in spherical shell models of mantle convection, *J. Geophys. Res.*, *105*, 11,063–11,082.