The fate of the slabs interacting with a density/viscosity hill in the mid-mantle

G. Morra a, b, *, D.A. Yuen c, d, L. Boschi b, P. Chatelain e, P. Koumoutsakos e, P.J. Tackley b

a Geology Department, University “Roma Tre”, Rome, Italy
b Geophysics Institute, ETH Zuerich, Switzerland
c, d Minnesota Supercomputing Institute, University of Minnesota, Minneapolis, MN 55455, USA
d Department of Geology and Geophysics, University of Minnesota, Minneapolis, MN 55455, USA
e Mechanical Engineering Department, ETH Zuerich, Switzerland

A R T I C L E   I N F O
Article history:
Received 16 June 2009
Received in revised form 24 March 2010
Accepted 6 April 2010

Guest Editors
Kei Hirose
Thorner Lay
David Yuen
Editor
G. Helffrich

Keywords:
Subduction
Slab–mantle interaction
Spin transition
Mid mantle viscosity

A B S T R A C T

In the last two decades it has been proposed several times that a non-monotonic profile might fit the average lower mantle radial viscosity. Most proposed profiles consist in a more or less broad viscosity hill in the middle of the mantle, at a depth roughly between 1200 km and 2000 km. Also many tomographic models display strong signals of the presence of “fast” material lying at mid mantle depths and a recent spectral analysis of seismic tomography shows a very clear transition for degree up to around 16 at a less than 1500 km depth. Finally latest works, both theoretical and experimental, on the high-to-low-spin transition for periclase, have suggested that the high-spin to low-spin transition of Fe++ might lie at the heart of all these observations. To verify the dynamical compatibility between possible mantle profile and observed tomographic images and compare them with possible mineral physics scenarios, such as the spin transition, we employ here a recently developed Fast Multipole-accelerated Boundary Element Method (FMM-BEM), a numerical approach for solving the viscous momentum equation in a global spherical setting, for simulating the interaction of an individual slab with a mid mantle smooth discontinuity in density and viscosity. We have focused on the complexities induced to the behaviour of average and very large plates (2000–10,000 km), characteristic of the Farallon, Tethys and Pacific plate subducting during the Cenozoic, demonstrating that the a mid mantle density and/or viscosity discontinuity produces a strong alteration of the sinking velocity and an intricate set of slab morphologies. We also employ the Kula–Farallon plate system subducting at 60 Ma as a paradigmatic case, which reveals the high resolution tomography models and clearly suggests an interaction with a strong and/or denser layer in the mantle. Our 38 models show that a plate might or might not penetrate into the lowest mantle and might stall in the mid lower mantle for long periods, depending on the radial profiles of density and viscosity, within a realistic range (viscosity 1, 10 or 100 times more viscous of the rest of the mantle, and a change of differential density in the range ~2% to 2%), of a transitional layer of 200 km or 500 km. We conclude that a layer with high viscosity or negative density would naturally trigger the observed geodynamic snapshot. We finally propose a scenario in which the long time accumulation of depleted slabs in the mid mantle would give rise to a partially chemically stratified mantle, starting from the less prominent high-spin to low-spin contribution on the basis of mantle density and rheology.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

The radial profile of the lower mantle viscosity is still largely unknown due to incomplete knowledge of composition, grain size, volatiles content, and to the lack of laboratory experimental data at such high pressure and temperatures and the difficulty of performing exhaustive ab-initio models of diffusion and dislocation creep for polycrystalline materials. However, geophysical inversion data based on geoid and post-glacial rebound (Forte and Mitrovica, 2001; Mitrovica and Forte, 1997) has suggested at the first order a peak of high viscosity in the middle of the lower mantle at around 2000 km and in a range between 10 and 100 times than the one in the upper mantle.

Few viscosity radial profiles for the lower mantle have been put forward. Among them (Ricard and Wuming, 1991) proposed that a peak of viscosity in the mid mantle might explain a set of geoid and topographical data. From geophysical considerations (Wen and Anderson, 1997) imposed a chemical barrier at 1200 km. (Kellogg et al., 1999) introduced the concept of a barrier at 1800 km depth based on geochemical and radiogenic heat production arguments, combined with tomographical evidences. Later a number of works...
came to similar conclusions, for example (Forte and Mitrovica, 2001) who combined satellite and ground data, (Ito and Toriumi, 2007) have found a peak of activation energy for vacancy diffusion at mid mantle pressures (both experimentally and numerically from large-scale molecular dynamics simulations) and Peltier (personal comm., 2008) looked at the GRACE dataset and found that the hypothesis of the viscosity hill in the mid mantle is still compatible with the most up-to-date gravity data, although this presents a controversy, as other studies, e.g., (Soldati et al., 2009) indicated that lower mantle viscosity structure can hardly be constrained from gravity data.

1.1. Tomographical evidences

A clearer picture of how anomalous the mid-lower mantle appears from mantle tomography. One of the better seismically resolved areas of the Earth’s mantle is under North America, where the Farallon slab is arguably the best-known subduction system characterized by likely penetration into the lower mantle, e.g., (Grand, 1994). The set of panels on the left of Fig. 1 illustrates the paleo plate reconstructions for the period from 80 Ma to 40 Ma (Müller et al., 2008) of the splitting of the wide oceanic plate with over 8000 km width subducting under North-America into two subplates called Farallon and Kula. A recent high resolution tomography (Sigloch et al., 2008), reproduced in Fig. 1, right panel, offers a very detailed image of the Farallon slab at mid mantle depth. The most prominent features are the thickening of the slab at depth between 1200 km and 1500 km and the segmentation into two plates, most probably the signature of the division of the paleo Farallon–Kula plate, also confirmed by the combination of palaeotectonic reconstruction and global tomography, as in Ren et al. (2007).

Such anomalous mid-lower mantle behaviour is confirmed by the visual inspection of a global tomography as the the one shown in Fig. 2 where the isosurface corresponding 0.8% heterogeneity, associated with the shear-velocity model of (Grand, 1994) is originally displayed with the Paraview software (Henderson, 2007). While also here the intriguing pattern of the Farallon slab is visible, with its shape that largely broadens at 1500–2000 km depths, other similar patterns appear in other regions of the lower mantle. Finally the picture shown in Figs. 1 and 2 for the Farallon slab is shown to be robust and independent on the tomographic models assumption through the comparative analysis of five independent global shear velocity models in Fig. 3: tx2007 (Simmons et al., 2006), pri-s05 (Montelli et al., 2006), rmsl-s06 (Li et al., 2007), saw642an (Panning and Romanowicz, 2006) and smean (Becker and Boschi, 2002). All such models convey very clearly the picture of a Farallon slab that does not cross straightaway in the lower mantle, but instead flattens at depth of 1500–2000 km, as it would encounter an obstacle.
The 3D seismic structure of the lower mantle can be also analyzed globally, looking at the spectral signal of all the slabs together. Detailed analysis of spherical harmonics results as the one shown in Fig. 4 (modified from Fig. 1 of (Boschi et al., 2008)) illustrates an unexpected non-monotonic radial pattern. The logarithm of the ratio of positive-to-negative shear-velocity spectra from model SMEAN (Boschi et al., 2008) recently also observed in (Houser and Williams, 2009) elucidates, in fact, how the positive (blue) fast anomalies dominate until the depth of \( \sim 1500 \, \text{km} \) up to spherical harmonic degree 15, while the situation is reversed below where negative (red) values dominate at lower depths. This unquestionable tomographical observation is the premise for searching the causes and consequences of the phenomenon.

### 1.2. Mineral physics insights

The dichotomy between the deepest 1000 km and the rest of the lower mantle has been already proposed to be of compositional nature (Van der Hilst and Kárason, 1999). However this picture has been enriched by recent works that combine the interpretation of seismic data with the knowledge of the elastic properties of silicates whose spin varies at mid mantle conditions (da Silva et al., 2000). Such novel works depart from the most recent experiments on ferropericlase during its spin transition (Marquardt et al., 2009) and exploit the surprisingly transient and anomalous character change of the bulk modulus during the transition, that largely softens. Because the observed pattern of vertical gradient of bulk velocity \( (dV_p/dz) \) is instead smooth and apparently adiabatic (Cammarano et al., 2003; Matas et al., 2007), the only explanation of this phenomena is a mid-lower mantle thermo-chemical transition at the same depth of the spin transition in order to compensate for the bulk-modulus weakening (Cammarano et al., 2010).

If this hypothesis proves to be correct, this groundbreaking interpretation would then demand a thorough investigation of its geodynamical consequences and above all, of the conditions that might have produced it. Such a mid mantle transition is clearly not as sharp as the upper lower mantle one but it is believed to be smooth. The inversion of the laboratory data indicates a thickness...
of several hundred kms (Marquardt et al., 2009), which is less than the prior spin transition thickness i.e. up to 1000 km (da Silva et al., 2000). The combination of seismic and mineralogical investigations strongly suggests that the thermo-chemical should happen at depths around 1200–1600 km (Cammarano et al., 2010; Marquardt et al., 2009).

All these indications are converging toward a scenario in which the viscosity profiles in the lower mantle might be non-monotonic, which would be entirely different from the canonical picture one gets from an Arrhenius type of activation energy, while it would be better described by compositional differentiation. A fundamental point to clear is the origin of the hill shape of the mantle viscosity arising from some inversion. If the rising of viscosity can be associated with the differentiated composition, i.e. with more viscous rocks at higher depth, the following lowering viscosity would be simply due to the rising temperature, i.e. to the Arrhenius law again. In this sense a selection of stronger rocks at higher depth is the natural outcome of a very long-term dynamics in which the deep Earth is naturally less mobile.

With the exception of the work by (Wen and Anderson, 1997), who postulated a chemical layer at around 1200 km depth based on geoid inversion, this is probably the first attempt to fit the observed mantle observable with an explicit compositional well defined radial transition. The details of the generation of such a three-layered mantle require careful study and are still speculative, however the strength of the tomographic indications listed above are so neat that require models able to explain them. Yet there is only one conclusive evidence of the existence of such dramatic alteration, i.e. (Cammarano et al., 2010) and more independent studies are required to confirm this observation. Finally definitive estimates of this new radial profile have not been presented.

1.3. Geodynamic scenario and our parameterization

If the above mineral-physics and seismological interpretations are both correct, it remains to be explained the geodynamic origin of the lower mantle transition and what is its impact for rising plumes and sinking slabs. This work aims at understanding just the fate of a slab in the lower mantle. The novel geodynamic scenario here suggested is the one in which a sinking slab must cross three distinct mantle layers in its way to the core: the upper mantle, the external-lower mantle (i.e. above 1500–2000 km) and the interior-lower mantle (i.e. below 1500–2000 km). In particular, we propose that the viscosity hill often suggested in the literature was indicating such a mid-lower mantle viscosity transition.

We consider 4 fundamental parameters influencing the fate of the slab. Other factors taken fixed in all models are the thickness of the plate, set to 200 km, which should account for the broadening

Fig. 5. Viscosity and density radial profiles for the lower mantle. Above are shown two models proposed in the past by (Ricard and Wuming, 1991) and (Forte and Mitrovica, 2001), both displaying a non-monotonous viscosity profile in the lower mantle, with a maximum at mid mantle depth. The lower panel illustrates instead the viscosity and density profiles that we test in this work. Out lower mantle model exhibit a less prominent viscosity peak respect to the one proposed above, as we observe in our model that 100× has already a strong influence on the geodynamic evolution. The density profile shown in bottom right figure illustrates both the hill-mantle and litho-mantle density differential. Those values are end-member, corresponding to typical seismic fluctations. We model all the nine combinations.
that the plate undergoes due to diffusion mechanisms and maybe thickening in the early stages of subduction. The plate starts each model from an initial straight shape, inclined of 20 degrees from the vertical, just above the hill.

The first examined parameter is the differential density between slabs and mantle. Because the slab composition cannot change in its way to the core (except some devolatilization phenomena, mostly in the upper mantle), and the spin transition happens in the slab at shifted depths compared to the rest of the mantle, due to lithosphere’s cooler temperature, we consider the entire range of possible differential densities, 2%, 0% and −2%, from the lower mantle during the transition. Before and after the mid mantle transition we assume instead unvaried 2% negative buoyancy for the slab (Fig. 5). Real values can be different for each real case, but our end-member scenarios should encompass all possible events.

The second parameter is mantle viscosity. We investigate a non-homogeneous mantle with a mid mantle peak viscosity 1, 10 and 100 times of the far field mantle, as they are end members and are of lower or equal order of magnitude to the proposed mantle viscosity hills (Fig. 5). We do not vary the mantle-lithosphere viscosity ratio. Such ratio has been amply studied in most geodynamic simulations of subduction that give a value of about 100 (Capitanio et al., 2007) in the upper mantle. In the lower mantle such a value is typically estimated as being much lower. We assume therefore a slab with a viscosity ratio of 10 between slab and mantle. Interesting tests of how a homogenously viscous slab might evolve in the lower mantle have been proposed by (Kaˇırason, PhD Thesis) showing a complexity of shapes, also seismically detected and interpreted (Ren et al., 2007).

In the technical implementation of the viscosity hill, as explained in Appendix A, only the high shear stresses due to the highly viscous mantle have been implemented. Several reasons justify this assumption: (a) the spin transition should take place at different depths inside and outside the slab, due to the differential temperature between slab and mantle; (b) we follow the conceptual model that the viscosity rises in the mantle due to a compositional variation, not directly due to the spin change (however, this is not clear yet); (c) we are interested in the fate of the slab in the particular context of their ability to cross a mid mantle viscosity peak and reach the core-mantle boundary contributing to the D” composition or to a more or less thick differentiated mid mantle layer, and this is due more to the mantle shear stresses applied to the slab than to the slab strength itself. In other words, we are interested only in the case in which the lithosphere remains a coherent entity, while we can only detect, and not directly model, the case in which it becomes unstable, thus producing fragments that drop in the underlying mantle.

The last two parameters are geometrical, i.e. the thickness of the transition, in which two end-members cases are tested: a thin 200 km and a thick 500 km and two different slab width cases: very wide, 10,000 km and relatively narrow, 2000 km, corresponding respectively to the largest plates on Earth and to a typical subduction zone.

The total of 36 possible combinations gives origin to a very large set of possible outcomes that we summarize in three categories: (1) Stokes flow, for slabs that directly penetrate through the transition and do not feel the presence of mantle peak effects, (2) transient, for slab that sensibly slow down at the mid mantle transition and then go through it arriving at the lowest mantle and (3) stalling for slabs that remain entrapped in the mid mantle and never cross the transition.

2. Numerical simulations

The simulation of the flow of creeping systems with large and sharp viscosity and density variations presents formidable challenges, due to the extremely complex evolving heterogeneities. We employ an indirect boundary integral formulation representing the boundary between a strong slab a less viscous mantle. The resulting formulation, based on bounded Stokes flow laws, is a set of Fredholm integral equations of the second kind. While the methods based on boundary integral equations (Brebbia, 1978) reduce the effort necessary for representing complex geometries, their standard formulation is computationally very expensive for large problem due to the construction of dense matrices. In order to accelerate the solution, we replace the calculation of the dense matrix with the implementation of an algorithmic “matrix mul-
Fig. 7. Evolution of a sinking slab through a viscosity peak with two different lithospheric-mantle differential densities $\Delta \rho$ for a not very large plate (2000 km) and very broad peak zone (500 km). The viscosity hill is chosen 10 times the background lower mantle viscosity. When the differential density between slab and hill density is null, the transition is crossed but the slab morphology changes dramatically, generating a Rayleigh Taylor instability from the center of the stalling plate, with a ring shape. The time evolution shown in the lower panels illustrates how negative buoyancies (i.e., hill density higher than lithosphere density) inhibits the crossing inducing slab stalling. In the long term, if the differential buoyancy remains respect to the lowest mantle, how small scale or even a subduction-like instability triggered, as illustrated in Fig. 8.

Fig. 8. The final crossing morphology of 6 different models, 3 characterized by a small 2000 km wide plate and 3 by a wide one (10,000 km), all crossing a 200 km thick transition zone and viscosity hill of 10 times the background mantle. The models from left to right change through differential lithosphere-mantle density. The positive to null lithosphere-hill differential density comparison shows little effect, while the negative differential density (hill density > slab viscosity) instead triggers in the long term long residency times in the mid mantle and an unstable behaviour, coherently found for small or very wide plates, at different scales.
Fig. 9. Summary of all models tested, through the radial trajectory of center of mass of the sinking slab. Small plates (2000 km wide) are displayed in the panel above, while very wide plate (10,000 km) in the lower panel. The three families of models also recalled in the tables are (a) Stokes-like, which sink down to the core almost not interacting with the viscosity hill; (b) stalling, with plates entrapped close to the viscosity–density hill for a period of the order of lower mantle Stokes crossing time or longer; (c) transient, spend a moderate time around the viscosity hill, typically no longer of the lower mantle Stokes crossing time. Meaning of the legend: the line style indicates the viscosity ratio between inner and outer slab requires a careful implementation. The approximation here employed is explained in detail in Appendix A, formula (9), and has been tested with excellent results, giving only local errors in proximity to the mantle–hill zone. In practice, it corresponds to a mean-field approximation, neglecting the effects of the far field viscosity transition; (b) the non-homogeneous differential density requires the calculation of a one-dimensional radial finite difference implementation of the pressure, applied to the boundary element formulation, replacing the component of the stress in the Stokeslet component, as explained in recalculating the formula (10). This is done at each step, with minimum computational time consumption, by integrating over the entire mantle density profile.

3. Model results

In this section we describe in detail the results of our experiments in numerical modelling; we provide in Section 4 a geophysical explanation of our theoretical insights. Let us consider first the interaction of a small plate (2000 km × 2000 km) that encounters a purely 500 km thick viscosity hill, without considering any density variation, i.e. only the case in which \( \Delta \rho_{\text{hill}} \) is 2% higher in the slab compared to the surrounding mantle. The transition from the 10 × to the 100 × viscosity hill model is paradigmatic of the dramatic change due to the pure mantle shear stresses to the slab sinking. Three steps whose time evolution for the two models are illustrated in Fig. 6. The slab in both cases crosses the viscosity hill, but the time required for going through it is one order of magnitude higher for the 100 × model. Furthermore the impact with the high viscosity layer changes dramatically the morphology of the down welling, which converts from a slab like structure into a cylindrical drop, growing as Rayleigh–Taylor (RT) instability. We call the case shown above Fig. 6 as being Stokes sinking, while the one on the lower mantle is defined as being transient.

In order to show the effect of a density variation superposed to a viscosity hill, the following Fig. 7 illustrates the time sequences of two 10 × viscosity models with a small 2000 km × 2000 km plate testing 2 different smooth density variations \( \Delta \rho_{\text{hill}} = \rho_{\text{litho}} - \rho_{\text{mantle}} \) at the peak either 0% or −2% compared to surrounding mantle. The 0% model, whose time evolution is illustrated in Fig. 7 (top panels), displays an intermediate outcome between the two models shown in Fig. 6, with the final morphology (after the hill is crossed) resembling partly such as slab and partly such as a RT instability, falling in a ring shape. The lower panel of Fig. 7 shows that in this simple configuration, the opposing buoyancy model (\( \Delta \rho_{\text{ill}} = -2\% \)) is able to halt the plate at the viscosity transition. The two models here displayed are paradigmatic for a transition from a transient to a stalling case.

Fig. 8 illustrates the final stage of 6 different modeled slabs, focusing on the role of differential density, summarizing their morphology. Only the mild viscosity hill (10 ×) is considered as they show a larger variation compared to the 100 × cases. Here upper and lower panels represent the comparison between a small (2000 km × 2000 km, top panels) and a very wide plate (10,000 km × 2000 km, bottom panels). Although the timing of the deformation is very different for small and large plates (see Fig. 9), the final configurations are similar. The most surprising
ters involved in the system. In order to detect the equivalent
of possible sinking rates that depend on the physical parame-
trajectories of the centre of mass for the small and the 5 ×
panels and a 10,000 km
the radial position of the center of mass of each plate versus a
laterally in form of drops.

instability might grow from segments inducing several mesoscale
is therefore foreseeable that for a morphological complex hill this
sides of the plate itself. As it is shown in Fig. 9, such a down welling
rounded long-term shape of the plate, with down welling at the

Table 1
Summary of the results of the 36 experiments.

<table>
<thead>
<tr>
<th>Hill thickness (km)</th>
<th>Viscosity ratio (η_l = η_lith/η_mantle)</th>
<th>Differential density (Δρ = ρ_lith−ρ_lith) (%)</th>
<th>Slab width (km)</th>
<th>Fate of the slab (Stokes, transient or stalling)</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>1</td>
<td>-2</td>
<td>2000</td>
<td>Transient</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>-2</td>
<td>10,000</td>
<td>Transient</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>0</td>
<td>2000</td>
<td>Stokes</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>0</td>
<td>10,000</td>
<td>Stokes</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>2</td>
<td>2000</td>
<td>Stokes</td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td>2</td>
<td>10,000</td>
<td>Stokes</td>
</tr>
<tr>
<td>200</td>
<td>10</td>
<td>-2</td>
<td>2000</td>
<td>Transient</td>
</tr>
<tr>
<td>200</td>
<td>10</td>
<td>-2</td>
<td>10,000</td>
<td>Transient</td>
</tr>
<tr>
<td>200</td>
<td>10</td>
<td>0</td>
<td>2000</td>
<td>Stokes</td>
</tr>
<tr>
<td>200</td>
<td>10</td>
<td>0</td>
<td>10,000</td>
<td>Stokes</td>
</tr>
<tr>
<td>200</td>
<td>10</td>
<td>2</td>
<td>2000</td>
<td>Stokes</td>
</tr>
<tr>
<td>200</td>
<td>10</td>
<td>2</td>
<td>10,000</td>
<td>Stokes</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>-2</td>
<td>2000</td>
<td>Stalling</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>-2</td>
<td>10,000</td>
<td>Stokes</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>0</td>
<td>2000</td>
<td>Stokes</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>0</td>
<td>10,000</td>
<td>Stokes</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>2</td>
<td>2000</td>
<td>Stokes</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>2</td>
<td>10,000</td>
<td>Stokes</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>2</td>
<td>2000</td>
<td>Stalling</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>2</td>
<td>10,000</td>
<td>Stokes</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>2</td>
<td>2000</td>
<td>Stalling</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>2</td>
<td>10,000</td>
<td>Stokes</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>2</td>
<td>2000</td>
<td>Stalling</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>2</td>
<td>10,000</td>
<td>Stokes</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>2</td>
<td>2000</td>
<td>Stalling</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>2</td>
<td>10,000</td>
<td>Stokes</td>
</tr>
</tbody>
</table>

result is that, while a variation from 2% to 0% triggers little effect,
the negative-buoyancy (Δρ_lith = -2%) model is characterized by a
rounded long-term shape of the plate, with down welling at the
sides of the plate itself. As it is shown in Fig. 9, such a down welling
arises a long time after the plate has met the mid mantle hill. It
is therefore foreseeable that for a morphological complex hill this
instability might grow from segments inducing several mesoscale
down welling (Cizkova and Matyiska, 2004), but possibly also being
an origin for upwelling plumes. In general, the models with inverse
density hill (able to invert the sign of the slab buoyancy) display an
unstable upside-down effect for which the slab splits and then fall
laterally in form of drops.

All models are finally summarized in Fig. 9 where it is displayed
the radial position of the center of mass of each plate versus a
common non-dimensional time (a 2000 km × 2000 km plate in top
panels and a 10,000 km × 2000 km plate in the bottom panels). The
trajectories of the centre of mass for the small and the 5 × wider
plate display a similar pattern but the mean-velocity is about dou-
brable for the wider plate (see x-axis of the upper and lower panel).
This can be explained by the prediction of (Capitanio et al., 2007)
based on the classical work of (Happel and Brenner, 1983), that
the drag of a plate in the mantle is proportional to [1 + ln(L/T)],
where L is the length of the plate and T is the short-
est. In our case DL = 1 + ln(10,000 km/2000 km) = 1 + ln(5) = 2.6 while
DS = 1 + ln(2000 km/2000 km) = 1 + ln(1) = 1. The sinking velocity is
given by the ratio between plate total buoyancy (volume × Δρ) and
drag, υ = (5 × 1)/2.6 ~ 1.9 and υ ~ (1 × 1)/1 = 1, therefore the ratio
υ/υ = 1.9 explains the observed sinking velocity.

The non-dimensional time employed in Fig. 9 generates a range
of possible sinking rates that depend on the physical param-
eters involved in the system. In order to detect the equivalent
order of magnitude for the modelled plates, we can convert
the non-dimensional time using the velocity renormalization factor
2/9 × Δρ_ga^2 |η and the time renormalization 9/2 × |η/Δρ_ga, where
η is the mantle viscosity, Δρ is the differential density, g is the
gravity and a is the effective length of the slab a = 5 × [1 + ln(L/T)],
where S and L are as above, 2000 km and 2000–10,000 km. Assuming
a lower mantle background viscosity in the range of 10^22 to
3 × 10^22 Pa s and an average differential density in the range 1–2% (~50 kg/m^3) for the sinking lithosphere, one finds that a time unit
is equivalent to 2–6 Myrs, which can be linearly renormalized for
a different lower mantle viscosity. This implies that we predict a
crossing time of the lower mantle of the order of 20–60 Myrs for
very wide plates, as Farallon, Tethys or Pacific, and of 40–120 Myrs
for small plates as single pacific slabs, both for the models that
directly cross the mid mantle transition. Comparing this time with
the diffusive timescale of a plate in the mantle, this could explain
why only the largest plates have been detected in the deepest
mantle.

For both, wide and narrow plates, the trajectories can be
grouped in three families (Fig. 9 and Table 1), defined roughly by the
same parameter subset. The first family is specified by the cluster
of trajectories that cross the hill in an almost unperturbed form and
is called Stokes. For a hill of only 200 km, such models are defined
by the ones that cross a small or no viscosity transition hill and
have a lithosphere–mantle density difference positive or null, while
for a hill thickness of 300 km already when the density difference
is null the trajectory is perturbed. The second family is composed
by the plates that remain permanently or for extremely long time
stalled at the hill. We call this dynamics stalling. This involves all
the models with a hill-thickness of 500 km and negative Δρ and
high viscosity and negative Δρ (see Table 2 for better visualizing

Table 2
Synthetic summary of the results.

<table>
<thead>
<tr>
<th>$\Delta \rho = \rho_{mantle} - \rho_{slab}$</th>
<th>$\eta = \eta_{mantle}/\eta_{slab}$</th>
<th>2%</th>
<th>0%</th>
<th>−2%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Stokes</td>
<td>Stokes</td>
<td>Stalling (T = 200 km)</td>
<td>Transient (T = 200 km)</td>
</tr>
<tr>
<td>10</td>
<td>Stokes (T = 200 km)</td>
<td>Transient (T = 500 km)</td>
<td>Transient (T = 200 km)</td>
<td>Transient (T = 500 km)</td>
</tr>
<tr>
<td>100</td>
<td>Transient (T = 200 km)</td>
<td>Stalling (T = 500 km)</td>
<td>Stalling</td>
<td></td>
</tr>
</tbody>
</table>

the transition). Collected into a third family there is then a large set of models that are characterized by moderate anomalous viscosity and moderate density anomaly, displaying a transient behaviour. They sensibly slow around the viscosity/density hill for periods of the same order of magnitude of the ones that penetrate down to the bottom of the lower mantle and then sink down, often displaying a new morphology. Those models are called transient in the tables. Although their definition is here slightly arbitrary, the direct analysis of the models shows their actual existence. Their importance, as discussed in the next section, resides in the possibility that the long transient time spent around the mid mantle transition might act as a reservoir that in the life of the Earth evolution would originally cause the compositional anomaly observed in the lower mantle, as better discussed in the next session.

4. Discussion and new scenarios for the Earth’s mantle

Convergent results taken from gravity data, and more recently mineral physics and seismology, indicate that a non-monotonic profile of viscosity/density might better describe the mid-lower mantle dynamics. In particular, it has been proposed that a high viscosity layer several hundred km thick might exist at depths 1200–1600 km. More recently a study that compares mineral physics and seismological data introduces evidences for a compositionally distinct layering below ~1500 km depth (Cammarano et al., 2010).

We find little difference between the dynamics of a very wide plate and a small plate, except for the sinking-velocity, which respects a logarithmic law (Capitanio et al., 2007; Morra et al., 2009). This terminal velocity diversity explains why only very large-scale plates (Farallon, Tethys and Pacific), while smaller slabs would tend cross, although being slowed down, by the non-monotonic radial pattern of the lower mantle viscosity. This filtering effect could be enhanced by a mid mantle layering that induces a transient dynamics (Fig. 9), which induces small slabs to stall in the mid mantle after a period of several tens Myrs, diffusing heat and thereby dissipating away their differential buoyancy.

We modelled numerically the effects of a wide but realistic range of thicknesses, densities and viscosities for a non-homogeneous layer in the lower mantle and on the lower-mantle subduction of a relatively small and a very wide plate. Depending on such parameters, a variety of different behaviours is possible. The change in thickness induces alone stalling if a sufficiently strong negative buoyancy (to the mantle) is applied, while the transient behaviour is common to both 200 km and 500 km models (Table 1), changing only the timing of the transition. We conclude that even a mild viscosity and density transition, if broad enough, can dramatically change the fate of the lower mantle slabs.

An unexpected outcome of the model is the fate of slabs that during their sinking encounter a combination of a high viscosity zone with a negative density anomaly. We predict that they might stall a long time in the mid mantle transition zone, furthermore they might tear and/or fragment because of the opposite buoyancy forces encountered and then amplified by the local high viscosity of the mantle. This would reduce the wavelength of the seismic anomalies observed below in the deep mantle, which can explain the observation of Fig. 4.

Plotting the hill–mantle viscosity ratio versus slab–hill differential density in Table 2 shows the most distinct transition between the three dynamic domains that we show in this work, with a banded configuration that illustrates how both the negative buoyancy and the high viscosity ratio are able to create a transient and a stalling dynamics. This somehow predictable result is here directly quantified, illustrating how (a) if the viscosity peak proposed before for the mid mantle hill is realistic (100 x) then the slab must either stall or have a long residential time in the mid mantle (see bottom line of Table 2); (b) if the compositional heterogeneity in the deep lower mantle is more dense then the above one (as it is predictable) the dynamics of the lower portion of the lower mantle will be strongly inhibited and the slabs will always display a transient or stalling behaviour (see right column of Table 2); (c) if the two effects act in a combined way, with close to neutral buoyancy and a mild hill–mantle viscosity ratio, the thickness of the transitional layer will play an important role controlling the slab behaviour and also the slab width will be relevant because inducing a larger mantle flow.

If a density and viscosity non-homogenous layer of thickness of 200 km or 500 km exists in the lower mantle, this must be necessary due to a compositional heterogeneity, either within locally in the layer, or between a sandwich consisting of a top and a bottom layer in the lower-mantle, because the spin transition in the ferropericlase alone cannot induce it (Wentzcovitch et al., 2009). Furthermore, the bulk modulus softening in ferropericlase during the spin transition is not observed in the dVp/dz data, which also implies that a compositional transition might compensate for the lack of dVp/dz jump. Such a configuration, if confirmed, requires a geodynamic system that at steady state is able to create such compensation. We envisage here this new mechanism, proposing that many sinking slabs would have transient time or completely stall in the mid-lower mantle. Such slabs represent a compositional differentiated reservoir, as it has been for example already proposed for an hazzburgite component in the D” layer (Hirose et al., 1999). In the long term, a differentiation in the mid-lower mantle would dynamically develop by a feedback due to the already existing compositional lowest mantle segregation, in a self-sustained form. This mechanism could induce over a long term the viscosity–density hill itself, causing a radial separation in two layers of the lower mantle, therefore preventing global mantle convection and favouring a more layered form of mantle dynamics.

Acknowledgements

Comments by two anonymous reviewers greatly helped to improve the manuscript. We thank discussions with Fabio Cammarano, Anne Hofmeister, Matt Knepley, Marc Monnereau, Fabio Capitanio. This work, as part of the Eurohors/ESF European Young Investigators Awards Scheme, was supported by funds from the National Research Council of Italy and other National Funding Agencies participating in the 3rd Memorandum of Understanding, as well as from the EC Sixth Framework Programme.
Appendix A. Fast multipole boundary element method for Stokes flow

In order to capture the most fundamental modes that characterize how a slab can or not cross a density and/or a viscosity transition we simplified the system to its fundamentals. The effective viscosity \( \eta(T,C) \) (Kellogg et al., 1999) that characterizes rheological behaviour of subducted plates has been replaced by two mean values: \( \eta_{\text{int}} \) inside the plate and \( \eta_{\text{ext}} \) for the external mantle. Using such rheological definition the gradient of velocity and \( \sigma \) is indicated by the viscous stress tensor:

\[
\sigma = -pI + \eta (\nabla \mathbf{u} + \nabla^T \mathbf{u}) = -pI + \eta \hat{\mathbf{e}}
\]

We solve the generalized Stokes equations that comprise the momentum conservation and incompressibility condition:

\[
\nabla p + \rho \mathbf{b} = 0 \quad \nabla \mathbf{u} = 0
\]

For our problem, this amounts to subdividing the mantle into several closed regions, each characterized by a homogenous density and viscosity. When such an approximation is acceptable, it is possible rewrite (2) as a boundary integral equation for each sub-domain. This constitutes a decrease in the problem dimensionality and therefore potential computational gains. Moreover, this approach offers inherent multiscale capabilities as the surface mesh resolution can vary dynamically and track the physics of interest.

We show later how a system characterized by a perturbed viscosity \( \eta \) and/or density \( \rho \) in function of space:

\[
\begin{align*}
\eta &= \eta(z) \\
\rho &= \rho(z)
\end{align*}
\]

as is our case, can be solved using a perturbative approach to our boundary equations.

A.1. Boundary equations

We consider Stokes flow as described by the Eqs. (1) and (2). Assuming a constant \( \eta \) in each domain, the velocity of each point inside the domain can be written as a sum of two surface integrals (Pozrikidis, 1992), called single and double layer integrals respectively, representing the effects of the forcing \( \mathbf{f}(\mathbf{x}) \) and velocity \( \mathbf{u}(\mathbf{x}) \)

\[
-\frac{1}{8\pi \eta} \int_{\partial D} \mathbf{f}(\mathbf{x}) \cdot \mathbf{n} G_{ij}(\mathbf{r}, \mathbf{x}, \mathbf{x}_0) dS(\mathbf{x}) + \frac{1}{8\pi} \int_{\partial D} \mathbf{u}(\mathbf{x}) \cdot \mathbf{n} T_{ijk}(\mathbf{r}, \mathbf{x}, \mathbf{x}_0) dS(\mathbf{x})
\]

\[
= \begin{cases} 
\mathbf{u}(\mathbf{x}_0) & \text{if } \mathbf{x}_0 \in D, \\
0 & \text{otherwise}
\end{cases}
\]

where \( G_{ij}, T_{ijk} \) are the steady Green’s functions for velocity and stress respectively, also known as the Stokeslet and the Stresslet:

\[
G_{ij}(\mathbf{x} - \mathbf{x}_0) = \frac{\delta_{ij}}{r} + \frac{\hat{x}_i \hat{x}_j}{r^3}, \quad \hat{x} = \mathbf{x} - \mathbf{x}_0 \quad \text{and} \quad r = |\hat{x}|
\]

\[
T_{ijk}(\mathbf{x} - \mathbf{x}_0) = -\frac{6}{r^5} (\hat{x}_i \hat{x}_j \hat{x}_k).
\]

In turn, the above equation is implemented both for the inner and the outer fluid, in this way the boundary equations are cast into a form more appropriate for a quasi-steady multiphase flows in the presence of a gravity field. Hence for a point \( \mathbf{x} \) on the surface \( \Sigma \) that separates different fluids, we obtain:

\[
\frac{1 + \lambda}{2} \mathbf{u}^{\text{PV}}(\mathbf{x}) - \frac{1 - \lambda}{2\pi} \int_{\Sigma} \mathbf{n} \cdot \mathbf{T} \cdot \mathbf{u} dS = -\frac{1}{8\pi \eta_0} \int_{\Sigma} G \Delta f dS,
\]

where \( \mathbf{PV} \) denotes the principal value of the integral, \( \eta_0 \) is the viscosity of the external fluid, taken as a reference and \( \lambda = \eta_{\text{int}}/\eta_0 \) is the viscosity ratio between inner and outer fluid and \( \Delta f \) is a normal stress jump that in our case accounts for gravity.

The equations have been widely used and tested for homogenous media. In this work we considered two types of inhomogeneities parameterized into a spatial variation of the coefficients (1) viscosity ratio \( \lambda = \eta_{\text{int}}/\eta_0 \); (2) density differential \( \Delta \rho = \Delta \rho_{\text{int}} - \Delta \rho_{\text{ext}} \). We assumed that a zone of thickness between 200 km and 500 km, within a domain of several 1000s km, could have viscosities of 1–2 orders of magnitude higher then the surrounding and \( \Delta \rho_0 = 0 \) or even negative. In practice for each panel \( \mathbf{x}_0 \) at the center of the panel(\( \mathbf{x}_0 \)), \( \lambda \) has been rescaled as:

\[
\lambda(\mathbf{x}_0) = \begin{cases} 
\eta_{\text{int}}/\eta_0 & \text{if } \mathbf{x}_0 \in \text{viscosity hill}, \\
\eta_{\text{int}}/\eta_{\text{ext}}(\mathbf{x}_0) & \text{otherwise}
\end{cases}
\]

The implementation of non-homogeneous density is slightly more complex. Exploiting a first-order hydrostatic approximation the non-uniform density is implemented as a space dependent shift of the pressure term in the single layer integral (in our case, the space variable is the radius and the pressure is defined by a profile of the pressure of the mantle to the lithosphere). We recalculated the entire radial pressure profile at each time step and the external forces applied in the single layer integral to the system. Hence the pressure \( P \) at the radius \( r = \left( \mathbf{x}_0 - \mathbf{x} \right)^2 / r^2 \) becomes:

\[
P(\mathbf{x}_0) = \int_{r(\mathbf{x}_0)}^{\infty} g \Delta \rho(\mathbf{r}(\mathbf{x})) d\mathbf{r}
\]

that is applied in the single layer of the boundary equation, replacing \( \Delta f \) with \( \mathbf{P} \).

A.2. Multipole approach

The sinking slab surface \( S \) in Fig. 5 and the supported quantities velocity \( \mathbf{u} \) and stress tensor at the boundary \( \mathbf{\sigma} \) are discretized with boundary elements (also called panels). The boundary integral equation thus becomes a linear system:

\[
\left( \frac{1 + \lambda}{2} + T \right) \mathbf{U} = \mathbf{F}.
\]

Many approaches carry out the construction of the matrix which scales as \( N^2 \) both memory- and computation time-wise though, making it impractical for large systems.

We use a Fast Multipole Method (FMM) (Barnes and Hut, 1986; Greengard and Rokhlin, 1987) in Eq. (8). The FMM algorithm, illustrated in Fig. 10, dramatically reduces the complexity of matrix-vector multiplication involving a certain type of dense matrix, which can arise out of many physical systems.

The FMM scales as \( N \log(N) \), which is far more tractable and still allows the use of a Generalized Minimized Residual method GMRES or any Krylov space based method that does not rely on the storage of the full matrix. By treating the interactions between asymptotic basis functions using the FMM, the corresponding matrix elements do not need to be explicitly stored, resulting in a significant reduction in required memory.

A multipole method exploits the decay of the kernel to convolve and makes a controlled approximation. It does this by expanding the system Green’s function using a multipole expansion, which allows one to group sources that lie close together and treat them as if they are a single source.

More explicitly, let us compute:

\[
\mathbf{u}(\mathbf{x}_0) = \int_{D} G(\mathbf{x}_0 - \mathbf{x}) \mathbf{P}(\mathbf{x}) dV(\mathbf{x})
\]
Fig. 10. A general sketch of the multipole method employed. The summation on each surface is done using local (point) or multipole terms, depending on the relative position of different centers. Right panel: The components of the summation are the integrals on each panel (=boundary element). Summing the integral contribution within a certain domain, the contribution of a set of panels is compressed into the one of one pole. Left panel: There are three possible summation methods. local (=point)–multipole, local–local, multipole–multipole. They depend on the relative contribution of the green function.

We consider the contribution from $D_i$, a part of $D$ that is far enough from our evaluation point $x_0$ and proceed with a Taylor expansion of the kernel $G$ about $x_0$, $D_i$:

$$u(x_0) = \int\! G(x_0 - x)\rho(x)dV(x)$$
$$\approx \int\! (G(x_0 - x_c) - \nabla G(x_0 - x_c)(x_0 - x_c) + \ldots)\rho(x)dV(x)$$
$$\approx G(x_0 - x_c) \int\! \rho(x)dV(x) - \nabla G(x_0 - x_c) \int\! (x_0 - x_c)\rho(x)dV(x) \ldots$$

We note that the equation involves successive moments of the $\rho$ distribution in $D_i$.

The present FMM code can handle convolutions with the Green's functions for the Poisson equation, the Stokeslet and the Stresslet. It employs up to the second order moments of the source distribution (quadrupoles). The reader is referred to (Barnes and Hut, 1986; Greengard and Rokhlin, 1987; Warren and Salmon, 1993) for general information on multipole methods and the work of (Tornberg and Greengard, 2008) for the transposition of harmonic multipoles to the evaluation of the Stokeslet and Stresslet.

The FMM algorithm thus sorts the sources in a tree structure whose cells contain the moment integrals – or multipoles – and carries out a field evaluation through a tree traversal (Fig. 11). The refinement of the interactions is determined by a tree traversal stopping criterion based on a prescribed tolerance.

The FMM-BEM drastically improves the computational cost of the method. For the coarse resolutions, the method displays the nominal $N^2$ scaling of a direct interaction code. This is attributed to the relatively few elements and tree cells. The scaling then quickly approaches a nearly linear one $N \log(N)$ for the finer resolutions.

The FMM-BEM has been parallelized using MPI. The parallel efficiency has been tested on a Opterion cluster with Quadrics connections. The scaling is very good up to 64 CPUs, still keeping 90% of efficiency. In its current implementation the FMM-BEM uses a shared (not distributed) tree, thus reducing the communication load at the expense of memory requirements.

References


