

Numerical modelling of magma dynamics coupled to tectonic deformation of lithosphere and crust

Journal Article**Author(s):**

Keller, Tobias; May, Dave A.; Kaus, Boris J.P.

Publication date:

2013-12

Permanent link:

<https://doi.org/10.3929/ethz-b-000074808>

Rights / license:

[In Copyright - Non-Commercial Use Permitted](#)

Originally published in:

Geophysical Journal International 195(3), <https://doi.org/10.1093/gji/ggt306>

Numerical modelling of magma dynamics coupled to tectonic deformation of lithosphere and crust

Tobias Keller,¹ Dave A. May¹ and Boris J. P. Kaus^{2,3}

¹*Institute of Geophysics, ETH Zürich, Sonneggstr. 5, 8092 Zürich, Switzerland. E-mail: keller@erdw.ethz.ch*

²*Johannes Gutenberg University, Mainz, Germany*

³*University of Southern California, Los Angeles, CA, USA*

Accepted 2013 July 29. Received 2013 July 26; in original form 2012 November 06

SUMMARY

Many unresolved questions in geodynamics revolve around the physical behaviour of the two-phase system of a silicate melt percolating through and interacting with a tectonically deforming host rock. Well-accepted equations exist to describe the physics of such systems and several previous studies have successfully implemented various forms of these equations in numerical models. To date, most such models of magma dynamics have focused on mantle flow problems and therefore employed viscous creep rheologies suitable to describe the deformation properties of mantle rock under high temperatures and pressures. However, the use of such rheologies is not appropriate to model melt extraction above the lithosphere–asthenosphere boundary, where the mode of deformation of the host rock transitions from ductile viscous to brittle elasto-plastic. Here, we introduce a novel approach to numerically model magma dynamics, focusing on the conceptual study of melt extraction from an asthenospheric source of partial melt through the overlying lithosphere and crust. To this end, we introduce an adapted set of two-phase flow equations, coupled to a visco-elasto-plastic rheology for both shear and compaction deformation of the host rock in interaction with the melt phase. We describe in detail how to implement this physical model into a finite-element code, and then proceed to evaluate the functionality and potential of this methodology using a series of conceptual model setups, which demonstrate the modes of melt extraction occurring around the rheological transition from ductile to brittle host rocks. The models suggest that three principal regimes of melt extraction emerge: viscous diapirism, viscoplastic decompaction channels and elasto-plastic dyking. Thus, our model of magma dynamics interacting with active tectonics of the lithosphere and crust provides a novel framework to further investigate magmato-tectonic processes such as the formation and geometry of magma chambers and conduits, as well as the emplacement of plutonic rock complexes.

Key words: Fracture and flow; Dynamics of lithosphere and mantle; Mechanics, theory and modelling; Rheology: crust and lithosphere; Pluton emplacement.

1 INTRODUCTION

1.1 The problem of melt ascent through lithosphere and crust

The physics involved in the ascent of silicate melt from the upper mantle through the continental lithosphere and crust, as it occurs mainly along converging plate boundaries, poses a considerable number of challenges. As a pulse of partial melt formed in the asthenosphere rises, it reaches the thermal boundary layer beneath the lithosphere, where the competence of the host rock gradually increases, reaching a point where both pervasive melt percolation as well as viscous diapirism become an inefficient means of melt prop-

agation. As the host rock becomes more competent, the significance of viscous creep mechanisms is diminished. Thus, if treated with a purely viscous rheology, melt pulses would stagnate and crystallize at depth. Under these conditions, it becomes necessary to consider elasto-plastic modes of melt propagation. Additionally, the effect of tectonic deformation of the lithosphere and crust on potential modes of melt extraction is of much interest, as many magmatic systems on Earth coincide with zones of major tectonic activity. Moreover, many plutonic bodies as well as volcanic systems occur at plate boundaries or fault systems, the activity of which coincided or overlapped with the time of melt migration and magmatic rock formation (Pitcher 1979; Hollister & Crawford 1986; McCaffrey 1992; Petford & Atherton 1992; Hutton 2011).

The problem of melt ascent through the lithosphere and crust involves a great complexity of processes. Conditions from the asthenosphere up to the Earth's surface span far over a thousand degrees of temperature and up to several GPa of pressure variation. The continental lithosphere, in particular, may be highly heterogeneous as it accumulates compositional and structural inhomogeneities during its long history. Inclusions of more fusible compositions in the asthenosphere and lower lithosphere may influence the style and efficiency of melt production and the onset of melt extraction (Aharonov *et al.* 1995, 1997; Kelemen *et al.* 1997; Spiegelman *et al.* 2001; Weatherley & Katz 2012). Structurally weakened rock left behind by inherited fault systems may localize subsequent cycles of deformation (e.g. Barnes 1994; Corti *et al.* 2007) and thus form pathways for the melt to penetrate through more competent rock layers. Even though such complexities should eventually be considered when dealing with melt transport through lithosphere and crust, we will neglect most of these factors here in order to concentrate on the basic physics of a low-viscosity melt propagating through a visco-elasto-plastic host rock undergoing a variety of conditions as they occur on the way from the upper mantle towards the upper crust.

1.2 Geodynamic two-phase flow

The physics of two-phase flow in the context of geodynamics have been derived in various studies (e.g. Sleep 1974; McKenzie 1984; Fowler 1985; Scott & Stevenson 1986; Spiegelman 1993a,b; Bercovici *et al.* 2001a; Bercovici & Ricard 2003). Most of these formulations are equivalent to each other under certain assumptions (Bercovici *et al.* 2001a; Bercovici & Ricard 2003) and are typically referred to as two-phase flow equations.

Geodynamic two-phase flow is characterized by the interplay of various competing modes of melt propagation. Porous flow through a deforming host rock may be localized by channeling instabilities of both mechanical (Stevenson 1989; Richardson 1998; Katz *et al.* 2006; Connolly & Podladchikov 2007; Golabek *et al.* 2008; Kohlstedt & Holtzman 2009) and chemical nature (Aharonov *et al.* 1995, 1997; Kelemen *et al.* 1997; Spiegelman *et al.* 2001; Weatherley & Katz 2012). At low background viscosities, either pervasive melt transport in a compacting host rock, or advective melt transport in a convective host rock will govern the style of melt extraction. Thus, depending mainly on the ratio between shear and compaction viscosities, either compaction waves (symmetrical or channelized) or convective diapirism (upwellings driven by the buoyancy of locally accumulated melt) will be the main feature of melt extraction (Scott 1988).

Another significant regime transition takes place between ductile flow and brittle fracture. The study of this transition between viscously flowing and elasto-plastically fracturing modes of melt ascent and emplacement is an area of long-standing scientific debate, where no consensus has yet been reached as to which process governs melt extraction and emplacement (Paterson & Fowler 1993; Menand 2011). To put it in the words of Rubin (1993a), 'it seems that the interpretation of existing field observations is [...] hampered by an inadequate understanding of rock that can undergo both fracture and flow.' Therefore, the aim of the methodology proposed here is to create a physical modelling framework that is able to directly access all end members of melt extraction and emplacement mechanics, that is, compaction waves and channels, convective diapirism and brittle fractures, in a fully coupled, self-consistent continuum approach.

1.3 Previous work

The traditional approach to two-phase flow in partially molten rock describes viscous creep of the host rock only. Such an approach is sufficient for simulating processes in the mantle and asthenosphere, where the rock viscosity is weakened by high temperatures and therefore deformation is dominated by viscous creep. Previous studies that implemented various forms of viscous two-phase flow into numerical models of magma dynamics have dealt with (i) the study of the basic modes of melt propagation, such as solitary waves (Barcilon & Richter 1986; Barcilon & Lovera 1989; Spiegelman 1993a,b), (ii) the simulation of melt extraction at midoceanic ridges (e.g. Scott & Stevenson 1989; Katz 2010) and above subduction zones (Cagnioncle *et al.* 2007), (iii) the numerical study of various forms of channelized percolation, such as the shear-assisted channeling instability (Stevenson 1989; Richardson 1998; Katz *et al.* 2006), the reactive melt band instability (Aharonov *et al.* 1995, 1997; Kelemen *et al.* 1997; Spiegelman *et al.* 2001; Weatherley & Katz 2012) and the decompaction channeling instability (Connolly & Podladchikov 2007), or with (iv) the simulation of liquid iron segregation during core formation of terrestrial planets (Golabek *et al.* 2008; Šrámek *et al.* 2010).

Whereas most studies of computational magma dynamics have been concerned with the mechanical aspects of porous flow, some studies have coupled the dynamics of melt transport and host rock compaction with models of conservation of energy and composition (Richard *et al.* 2007; Šrámek *et al.* 2007; Katz 2008; Dufek & Bachmann 2010). Such treatments add considerable complexity to models of magma dynamics. As the main concern of this study is the expansion of existing two-phase flow models to the full visco-elasto-plastic deformation of the host rock, we will neglect these additional complexities for now and continue to address the purely mechanical aspects of melt propagation.

We note that within the engineering and geomechanics community, significant progress has been made in understanding the dynamics of viscoplastic, or elasto-plastic two-phase flow in polycrystalline metals and dry rocks (e.g. Loret & Prevost 1991; Ehlers & Volk 1998; Rubin *et al.* 2000; Lomov & Robin 2003; Khoei & Mohammadnejad 2011) and of porous rock containing hydrous fluids (e.g. Fournier 1996; Evans 2005; Gessner 2009; Yang 2002). These studies, however, are often based on different PDEs (e.g. poro-elastic equations) or use complex constitutive laws specific to engineering materials; thus, it is mostly not straightforward to apply this knowledge the problems of computational magma dynamics.

Within the computational magma dynamics community, only few studies have taken steps towards the introduction of a fully visco-elasto-plastic host rock rheology in models of magma dynamics. Some work has been done on solitary wave propagation in a viscoelastic host rock (Connolly & Podladchikov 1998). In a more recent follow-up study, Connolly & Podladchikov (2007) added a parametrized implementation of plastic decompaction failure, leading to the emergence of elongated melt tubes or channels, rather than spherical two-dimensional solitary waves. Morency *et al.* (2007) derived a system of equations that is able to deal with viscoelastic compaction flow in a host rock undergoing viscoplastic shear deformation. Their study, however, does not deal with magma dynamics but rather with the percolation of water in sedimentary rocks. Also, the formulation of plasticity in that study only considers shear fracture of the host rock and omits tensile modes of fracturing.

Melt-bearing tensile fractures, however, represent one of the most common features of melt propagation found in outcrops of

magmatic rock formations. Such dykes and sills therefore have been studied theoretically and experimentally for years (e.g. Murrell 1964a,b; Lister & Kerr 1991; Rubin 1993a,b, 1995; Taisne & Jaupart 2009). Rozhko *et al.* (2007) simulated the emergence of such fluid-filled brittle fractures in an elasto-plastic host rock using a poro-elastic model approach. Although their study involved drastic simplification (e.g. constant melt fraction), it is nevertheless highly relevant in our context, as it demonstrates the possibility of simulating the essentially discontinuous process of tensile fracturing in a continuum mechanics framework.

1.4 Proposed method

In an effort to extend numerical simulations of magma dynamics to the fully coupled visco-elasto-plastic two-phase physics, where both shear and compaction deformation may occur as viscous creep, elastic strain, plastic failure, or any combined mode of deformation, we will adapt the standard set of two-phase flow equations to a suitable form and introduce an appropriate rheology for both shear and compaction deformation. The resulting set of equations is then implemented into a finite-element code suitable for simulating melt extraction under asthenospheric, lithospheric and crustal conditions. The resulting numerical formulation, as we will show, represents a straightforward extension of the standard formulation of Stokes flow used in many geodynamic simulations of mantle or lithosphere deformation today (Moresi *et al.* 2003; Gerya & Yuen 2007; Kaus *et al.* 2008; May & Moresi 2008; Popov & Sobolev 2008; Kaus 2010).

To test the model and demonstrate its potential, we run a suite of simulations to explore the various regimes of melt transport that self-consistently emerge from the visco-elasto-plastic deformation of the host rock. In a first series of simulations, we impose a small volume of melt at the lower boundary of a homogenous model box. While we impose tectonic background deformation through kinematic boundary conditions on the sides of the box, a constant fluid pressure condition at the lower boundary serves to continuously supply melt from below. The dimensions of this problem are chosen so as to resolve processes on a scale of a few kilometres. Varying the intrinsic viscosity and the tensile strength of the host rock enables us to examine a range of conditions found in the asthenosphere, lithosphere and crust, and crossing the boundary from ductile to brittle modes of deformation.

In two additional simulations, we extend the model domain to include a full cross section through a compositionally and structurally homogeneous lithosphere and crust. An initial volume of melt is placed below the lithosphere. The model domain is again subject to tectonic deformation imposed by kinematic side boundaries. As the viscosity structure on a lithospheric scale is depth-dependent, the various end member models studied in the first series of simulations are likely to be found at various depths of the lithosphere and transitions between modes of deformation are thus expected to occur in a self-consistent manner.

2 PHYSICAL MODEL

2.1 Basic governing equations

2.1.1 Material fractions

Our formulation of the magma dynamics problem deals with two material phases: (i) The solid phase—in our context the host rock—deforming as an incompressible visco-elasto-plastic medium.

Properties of the solid phase are denoted with a subscript (\cdot)_s. (ii) The fluid phase—in our context the silicate melt—deforming as an incompressible Newtonian fluid. We denote fluid properties with a subscript (\cdot)_f.

Properties of the phase mixture are defined as averaged over the fractions of a unit volume occupied by the solid and fluid phase, expressed in terms of the melt fraction ϕ , which is defined as the volume fraction of the fluid phase (silicate melt) per unit volume of the two-phase mixture. At melt fractions below a certain threshold in the region of $0.25 \leq \phi \leq 0.35$, the melt fraction corresponds to the fully saturated porosity of the host rock. Above this disaggregation threshold, the interconnected, cohesive matrix of the host rock disaggregates to form a ‘mush’ of solid grains submerged in the fluid phase, a transition marked by a sharp decrease of mechanical strength in the system (Renner *et al.* 2000; Caricchi *et al.* 2007; Costa *et al.* 2009). As two-phase flow beyond the host rock disaggregation threshold is an important part of our model, the commonly used term ‘porosity’ would not be general enough to describe the fluid volume fraction in all potential states of the two-phase mixture, and thus we will consistently use the term ‘melt fraction’ instead.

Properties of the two-phase mixture are defined in general for any property a as

$$\bar{a} = (1 - \phi)a_s + \phi a_f, \quad (1)$$

denoted by an over bar. The phase difference of any property a is defined as

$$\Delta a = a_s - a_f. \quad (2)$$

2.1.2 Mass conservation

The mass conservation of both phases is well known (McKenzie 1984; Bercovici *et al.* 2001a). Both solid and fluid phases are considered to be intrinsically incompressible materials, and thus all compressibility in the model is accounted for by changes in melt fraction. Density changes are neglected in the mass balance and no melting or crystallization is allowed to occur. The mass conservation equations of the fluid and solid phase are

$$\frac{\partial \phi}{\partial t} + \nabla \cdot \phi \mathbf{v}_f = 0, \quad (3)$$

$$\frac{\partial (1 - \phi)}{\partial t} + \nabla \cdot (1 - \phi) \mathbf{v}_s = 0. \quad (4)$$

Summing the contribution of both phases results in the total mass conservation of the two-phase mixture:

$$\nabla \cdot \bar{\mathbf{v}} = 0. \quad (5)$$

Here, $\mathbf{v}_{s,f}$ are the solid and fluid velocity, and $\bar{\mathbf{v}} = (1 - \phi)\mathbf{v}_s + \phi\mathbf{v}_f$ is the bulk velocity. $\partial/\partial t$ denotes partial derivatives with respect to time. Note that the mass conservation of the mixture (eq. 5) becomes equal to the incompressibility condition of Stokes flow. It is evident therefore that the deformation of the phase mixture is divergence-free (incompressible). Conversely, deformation in the solid and fluid phase is not divergence-free. The divergence of solid and fluid velocity is, however, not due to intrinsic material compressibility of the solid and fluid phases, but due to changes in melt fraction resulting from relative movement of one phase with respect to the other. Such changes may be caused by compaction/decompaction of the host rock matrix at lower melt fractions, or by other phase separation processes like crystal settling at higher melt fractions.

2.1.3 Momentum conservation

The momentum conservation equations given here are those proposed by Bercovici *et al.* (2001a), except for the terms related to surface tensions, which we do not consider here, as they are likely negligible in the context of this study (Bercovici & Ricard 2003). As is well known, these equations are equivalent to the two-phase flow formulation proposed by McKenzie (1984), given a set of assumptions that are valid in the limit of magma dynamics (Bercovici & Ricard 2003). The momentum conservations of the fluid and solid phases are

$$\nabla \cdot \phi \boldsymbol{\tau}_f - \nabla \phi P_f - \phi \rho_f g \hat{\mathbf{z}} + h = 0, \quad (6)$$

$$\nabla \cdot (1 - \phi) \boldsymbol{\tau}_s - \nabla (1 - \phi) P_s - (1 - \phi) \rho_s g \hat{\mathbf{z}} - h = 0. \quad (7)$$

By summing the contributions of both phases, we find the total momentum conservation in the two-phase mixture to be

$$\nabla \cdot \bar{\boldsymbol{\tau}} - \nabla \bar{P} - \bar{\rho} g \hat{\mathbf{z}} = 0. \quad (8)$$

Here, $\boldsymbol{\tau}_{s,f}$ are the solid and fluid deviatoric stress tensors, $\rho_{s,f}$ the solid and fluid densities and $\hat{\mathbf{z}}$ the vector in direction of gravity g . The equal and opposite interface forces h acting between the phases are defined as $h = c \Delta \mathbf{v} + P^* \nabla \phi$ (Bercovici & Ricard (2003)), with the viscous drag coefficient c governing interface drag forces proportional to the velocity difference $\Delta \mathbf{v}$, and some interface pressure P^* acting on gradients of melt fraction ϕ . Note that the total momentum equation (eq. 8) is equal to the momentum equation of Stokes flow, with $\bar{\boldsymbol{\tau}}$, \bar{P} , $\bar{\rho}$ the total deviatoric stress, total pressure and total density, defined according to the relation given in eq. (1).

2.1.4 Magma dynamics limit

The eqs (3)–(8) represent a fully symmetrical approach to the continuum mechanics of a two-phase mixture. In the limit of magma dynamics, however, we are dealing with strongly asymmetric phase properties. The special case of a silicate melt interacting with a solid rock presents us with a situation where $\eta_f \ll \eta_s$, a scenario referred to as the magma dynamics limit or Darcy limit. Bercovici & Ricard (2003) give a detailed discussion of this issue, concluding that the fluid momentum equation (6) reduces to a form where the viscous drag force is balanced by the excess fluid pressure gradient (note that in the magma dynamics limit P^* reduces P_f):

$$c \Delta \mathbf{v} = \phi (\nabla P_f + \rho_f g \hat{\mathbf{z}}). \quad (9)$$

Under the same assumption that $\eta_f \ll \eta_s$, a simplified relation is found for the viscous drag coefficient in the form of $c = \eta_f \phi^2 / k_\phi$, depending on fluid viscosity η_f and permeability k_ϕ (see Bercovici *et al.* 2001a). Substituting this definition into eq. (9) and rearranging of terms recovers Darcy's law, defining the phase separation flux \mathbf{q} as

$$\mathbf{q} = -\phi \Delta \mathbf{v} = -K_D (\nabla P_f + \rho_f g \hat{\mathbf{z}}), \quad (10)$$

where the Darcy coefficient is defined as $K_D = k_\phi / \eta_f$. Details on the melt fraction-dependence of permeability are discussed below.

2.2 Rheology of the host rock

The rheology of the host rock is a key factor governing both style and efficiency of melt extraction and emplacement in magmatic systems. Therefore, employing a realistic rheology for the host rock is crucial. In contrast to incompressible Stokes flow, not only a rheol-

ogy for deviatoric, or shear deformation needs to be defined, but an additional flow rule for the volumetric, or compaction deformation governing changes in melt fraction is required. Both rheologies will be defined in terms of a constitutive law relating deviatoric and volumetric components of stress to their corresponding components of strain rate.

2.2.1 Strain rate and stress

The deformation of the solid phase is quantified by the solid strain rate tensor found from the gradient and divergence of the solid velocity field. The total strain rate tensor $\dot{\boldsymbol{\epsilon}}_s$ is defined as the sum of its deviatoric and volumetric components:

$$\dot{\boldsymbol{\epsilon}}_s = \dot{\boldsymbol{\epsilon}}'_s + \frac{1}{3} \dot{v}_s \mathbf{I} = \frac{1}{2} (\nabla \mathbf{v}_s + [\nabla \mathbf{v}_s]^T), \quad (11)$$

where $\dot{\boldsymbol{\epsilon}}'_s$ is the deviatoric strain rate tensor, \dot{v}_s is the scalar volumetric strain rate and \mathbf{I} is the identity matrix of appropriate dimensions. Volumetric or compaction strain rates are defined equal to the divergence of solid velocity:

$$\dot{v}_s = \nabla \cdot \mathbf{v}_s. \quad (12)$$

Deviatoric or shear strain rates are then derived from eqs (11) and (12) to be

$$\dot{\boldsymbol{\epsilon}}'_s = \frac{1}{2} (\nabla \mathbf{v}_s + [\nabla \mathbf{v}_s]^T) - \frac{1}{3} \nabla \cdot \mathbf{v}_s \mathbf{I}. \quad (13)$$

Shear and compaction deformation lead to the build-up of stress in the two-phase mixture. Total stress in the mixture is defined as the sum of the total shear stress $\bar{\boldsymbol{\tau}}$ and the total pressure \bar{P} (the volumetric stress), expressed in terms of the contributions of both material phases ($\boldsymbol{\tau}_{s,f}$, $P_{s,f}$) as

$$\begin{aligned} \bar{\boldsymbol{\sigma}} &= \bar{\boldsymbol{\tau}} + \bar{P} \mathbf{I} \\ &= (1 - \phi) \boldsymbol{\tau}_s + \phi \boldsymbol{\tau}_f + [(1 - \phi) P_s + \phi P_f] \mathbf{I}. \end{aligned} \quad (14)$$

The sign convention is such that compressive stresses are positive. In the magma dynamics limit, deviatoric stresses in the fluid phase $\boldsymbol{\tau}_f$ may be neglected (Sleep 1974; McKenzie 1984, see detailed discussion below), and therefore the total shear stress reduces to

$$\bar{\boldsymbol{\tau}} = (1 - \phi) \boldsymbol{\tau}_s. \quad (15)$$

2.2.2 The effective stress principle

The principle of effective stress, introduced by Terzaghi (1923, 1943) in soil mechanics, holds that the deformation of a fluid-filled porous rock is governed by an effective stress depending on the pressure of the pore fluid:

$$\bar{\boldsymbol{\sigma}}_e = \bar{\boldsymbol{\sigma}} - P_f \mathbf{I}. \quad (16)$$

According to the definition of total stress in the mixture (eq. 14), the effective stress may be rewritten in terms of shear stress and pressure as

$$\bar{\boldsymbol{\sigma}}_e = \bar{\boldsymbol{\tau}} + (\bar{P} - P_f) \mathbf{I}. \quad (17)$$

It becomes evident from this expression that fluid pressure has no influence on shear stresses in the phase mixture, and therefore the effective stress principle may be fully captured in a new pressure variable, the compaction pressure, governing the volumetric

deformation of a two-phase medium (Skempton 1960). This compaction pressure is then defined as

$$P_c = \bar{P} - P_f \\ = (1 - \phi) \Delta P. \quad (18)$$

Although no such compaction pressure variable is proposed in the original formulation of two-phase flow by McKenzie (1984), the concept of the pressure difference between the phases sustaining volumetric deformation features prominently in its derivation. Later, the more general analysis of two-phase physics proposed by Bercovici *et al.* (2001a) confirmed that the pressure difference ΔP takes the role of governing compaction deformation in the general case of any two-phase mixture of two incompressible fluids with general properties (*cf.* Katz *et al.* 2007; Morency *et al.* 2007 for similar use of compaction pressure).

According to Terzaghi's principle, the plastic yield strength of rock is reduced by the presence of a pressured fluid in the pore space of a rock. Therefore, we define one more pressure variable, the effective pressure (or effective mean stress) P_e governing the plastic yield strength of the host rock. The effective pressure is equal to the compaction pressure P_c , if a minimum critical amount of fluid phase is present (i.e. a large-scale interconnected pore space is achieved), but will assume the value of the total pressure \bar{P} in dry or unmolten rock (Skempton 1960):

$$P_e = \bar{P} - x_\phi P_f, \quad 0 \leq x_\phi \leq 1. \quad (19)$$

Here, x_ϕ is a parameter indicating when no significant amount of fluid phase is present ($x_\phi = 0$), and when sufficient amounts of fluid phase are present ($x_\phi = 1$) and thus two-phase physics and the reduction of plastic strength apply.

It follows from this discussion on effective stress and pressure that at least two independent pressure variables are needed to describe two-phase physics. Here, we choose fluid pressure P_f and compaction pressure P_c as independent variables featuring in the final governing equations. Fluid pressure, on the one hand, is a straightforward choice, as fluid velocities may be found from its gradient, thus eliminating the need to solve for fluid velocities with separate equations. The choice of compaction pressure, on the other hand, is convenient because of its central role in governing volumetric deformation. Note that all other relevant pressure components may be found from these two quantities (Table 1 summarizes all pressure components used in this study).

Summarizing the discussion on effective stress, we state that the stress state in the two-phase mixture is fully characterized in terms of the total shear stress $\bar{\tau}$ and compaction stress P_c . For these two components of stress, constitutive laws need to be found, relating them to shear and compaction strain rates to yield expressions of the form:

$$\bar{\tau} \propto f(\dot{\epsilon}'_s); \quad P_c \propto f(\dot{\nu}_s). \quad (20)$$

As both constitutive laws may be derived in close analogy, we will first rederive the well-known rheology for shear deformation before using the same concepts to derive the rheology for compaction deformation. To simplify notation, we will drop the subscript ($_s$) on all rheological parameters (viscosities, elastic moduli, etc.) relating to the solid phase, as there are no rheological parameters related to the fluid phase except for fluid viscosity η_f and, hence, no ambiguity.

2.2.3 Viscoelastic shear rheology

In order to find the constitutive relation for viscoelastic shear stresses in the solid matrix, we use a Maxwell body

$$\dot{\epsilon}'_{s,ve} = \frac{1}{2\eta} \boldsymbol{\tau}_s + \frac{1}{2G} \frac{\tilde{D}_s \boldsymbol{\tau}_s}{Dt}, \quad (21)$$

where $\dot{\epsilon}'_{s,ve}$ is the viscoelastic deviatoric strain rate tensor, $\boldsymbol{\tau}_s$ the deviatoric solid stress tensor, η the solid shear viscosity and G the elastic shear modulus. For the material time derivative of the solid deviatoric stress tensor, the Jaumann objective derivative is used in order to account for stress rotation (Bathe 1995):

$$\frac{\tilde{D}_s \boldsymbol{\tau}_s}{Dt} = \frac{D_s \boldsymbol{\tau}_s}{Dt} - \boldsymbol{W} \boldsymbol{\tau}_s + \boldsymbol{\tau}_s \boldsymbol{W}, \quad (22)$$

where $\boldsymbol{W} = \frac{1}{2}(\nabla \mathbf{v}_s - [\nabla \mathbf{v}_s]^T)$ is the vorticity tensor and $\frac{D_s}{Dt} = \frac{\partial}{\partial t} + \mathbf{v}_s \cdot \nabla$ is the general material time derivative of a property advected with solid velocity.

For the further construction of our method, it is of some advantage to discretize the time derivative in eq. (21) at this point (Schmalholz *et al.* 2001). Thus, we formulate the stress rate as an implicit backward finite difference in time, yielding

$$\dot{\epsilon}'_{s,ve} = \frac{1}{2\eta} \boldsymbol{\tau}_s + \frac{1}{2G} \frac{\boldsymbol{\tau}_s - \tilde{\boldsymbol{\tau}}_s^o}{\Delta t}, \quad (23)$$

with $\tilde{\boldsymbol{\tau}}_s^o$ denoting the rotated deviatoric stress tensor taken from the previous time step Δt :

$$\tilde{\boldsymbol{\tau}}_s^o = \boldsymbol{\tau}_s^o - (\boldsymbol{W} \boldsymbol{\tau}_s - \boldsymbol{\tau}_s \boldsymbol{W}) \Delta t. \quad (24)$$

Note that, as we will adopt an Arbitrary-Lagrangian-Eulerian marker-in-cell implementation (described below), no local advection terms ($\mathbf{v}_s \cdot \nabla \boldsymbol{\tau}_s$) are included in the statement above; instead, stress advection will be achieved by storing old stresses $\boldsymbol{\tau}_s^o$ on Lagrangian markers that are displaced by the solid velocity field. Rewriting eq. (23) in terms of deviatoric stress and grouping of terms yields

$$\boldsymbol{\tau}_s = \frac{2}{\frac{1}{\eta} + \frac{1}{G\Delta t}} \dot{\epsilon}'_{s,ve} + \frac{1}{1 + \frac{G\Delta t}{\eta}} \tilde{\boldsymbol{\tau}}_s^o, \quad (25)$$

which in simplified notation becomes the viscoelastic constitutive law for deviatoric stress:

$$\boldsymbol{\tau}_s = 2\eta_{ve} \dot{\epsilon}'_{s,ve} + \chi_\tau \tilde{\boldsymbol{\tau}}_s^o, \quad (26)$$

Table 1. Pressure components.

Symbol	Name	Expression	$\phi \rightarrow 0$	$(1 - \phi) \rightarrow 0$
\bar{P}	Bulk pressure	$\bar{P} = (1 - \phi)P_s + \phi P_f$	P_s	P_f
P_f	Fluid pressure	$P_f = \bar{P} - P_c$	P_s	P_f
P_c	Compaction pressure	$P_c = \bar{P} - P_f = (1 - \phi)\Delta P$	0	0
P_e	Effective pressure	$P_e = \bar{P} - x_\phi P_f$	\bar{P}	0

Note: Definition of pressure components relevant to our formulation of two-phase physics. Note that Katz *et al.* (2007) and similar formulations of the McKenzie (1984) equations define compaction pressure as $P_c = -(\xi + 2/3\eta)\dot{\nu}$, a similar, but slightly different definition from the one in this study (eq. 44).

featuring two effective rheological parameters, that is, the viscoelasticity η_{ve} (with units of Pa s) and the non-dimensional elastic stress evolution parameter χ_τ :

$$\eta_{ve} = \frac{1}{\frac{1}{\eta} + \frac{1}{G\Delta t}}, \quad \chi_\tau = \frac{1}{1 + \frac{G\Delta t}{\eta}}. \quad (27)$$

As we neglect the contribution of shear stresses in the fluid phase, the total shear stress in the two-phase mixture may be described by multiplying eq. (26) by the volume fraction of the solid phase $(1 - \phi)$:

$$\bar{\tau} = (1 - \phi) (2\eta_{ve} \dot{\epsilon}'_{s,ve} + \chi_\tau \bar{\tau}'_s). \quad (28)$$

To simplify notation, we incorporate any occurrence of the solid volume fraction $(1 - \phi)$ in any further rheological statements into the relevant rheological parameters and denote this operation with the addition of an asterisk. In this case, we define the shear viscoelasticity and shear stress evolution parameter related to the two-phase mixture as

$$\eta_{ve}^* = \frac{1 - \phi}{\frac{1}{\eta} + \frac{1}{G\Delta t}}, \quad \chi_\tau^* = \frac{1 - \phi}{1 + \frac{G\Delta t}{\eta}}, \quad (29)$$

and thus the constitutive law for total shear stress (eq. 28) is finally expressed as

$$\bar{\tau} = 2\eta_{ve}^* \dot{\epsilon}'_{s,ve} + \chi_\tau^* \bar{\tau}'_s. \quad (30)$$

2.2.4 Viscoelastic compaction rheology

In the following, we proceed to define the viscoelastic constitutive law relating volumetric stresses to volumetric strain rates. Whereas McKenzie (1984) employed a viscous constitutive law for compaction stresses implied by his definition of the solid stress tensor, Bercovici *et al.* (2001a) explicitly state a constitutive relationship between the rate of change of melt fraction and the pressure difference between the phases of the form

$$\frac{D_s \phi}{Dt} = -K_0 \phi (1 - \phi) \frac{\Delta P}{\eta_0}, \quad (31)$$

with K_0 a constant of order $O(1)$ representing pore geometry, which we will take as unity here. We find that a constitutive law for viscoelastic compaction deformation consistent with eq. (31) may be formulated in close analogy to the one for deviatoric stress introduced above. Therefore, following the same strategy as in the previous section, we start to construct a viscoelastic constitutive law for ΔP by again assuming a Maxwell body analogue for the viscoelastic part of volumetric strain rate $\dot{v}_{s,ve}$:

$$\dot{v}_{s,ve} = -\frac{1}{\xi} \Delta P - \frac{1}{K_\phi} \frac{D_s \Delta P}{Dt}. \quad (32)$$

The pressure difference ΔP is negative where fluid is overpressured with respect to the pressure in the solid phase. ξ is the volumetric or compaction viscosity governing viscous compaction flow. That the viscous part of the proposed expression (first term on the right-hand side of eq. 32) is equivalent to eq. (31) is confirmed by dropping the elastic part (second term on right-hand side) from eq. (32) and substituting $\frac{1}{(1-\phi)} \frac{D_s \phi}{Dt}$ for $\dot{v}_{s,ve}$ (a relation that follows from eq. 4), and η_0/ϕ for ξ . Thus, the compaction flow law proposed by Bercovici *et al.* (2001a) is recovered in the purely viscous limit of the viscoelastic flow law proposed here.

For the elastic part of the constitutive law (second term on the right-hand side of eq. 32), we assume that all elastic volume changes are due to changes of melt fraction, with the elastic pore modulus

K_ϕ governing elastic compaction in response to changes in pressure difference over time. The assumption of a compressible pore space in intrinsically incompressible phase materials is commonly used in poro-elasto-plastic theory (Coussy 2010). D_s/Dt is the material derivative ($D_s/Dt = \partial/\partial t + \mathbf{v}_s \cdot \nabla$). Since ΔP is a scalar, the material derivative is objective and we are not required to consider objectivity with respect to rotation (as in eqs 21–22).

The viscoelastic part of the volumetric rheology is constructed in analogy to its deviatoric counterpart by implicitly discretizing the time derivative of pressure difference in eq. (32), rewriting the equation in terms of pressure difference, grouping of terms and introducing a simplified notation (*cf.* eqs 23–27 for step-by-step procedure). As a result, the viscoelastic constitutive law for compaction stress acting on the solid phase is

$$\Delta P = -\xi_{ve} \dot{v}_{s,ve} + \chi_p \Delta P^o, \quad (33)$$

where ΔP^o is the advected pressure difference of the last discrete time step. The compaction viscoelasticity ξ_{ve} and the compaction stress evolution parameter χ_p are defined as

$$\xi_{ve} = \frac{1}{\frac{1}{\xi} + \frac{1}{K_\phi \Delta t}}, \quad \chi_p = \frac{1}{1 + \frac{K_\phi \Delta t}{\xi}}. \quad (34)$$

As the pressure difference relates to compaction pressure in the same way as the solid shear stress relates to the total shear stress (*cf.* eqs 15 and 18), we multiply eq. (33) by the solid fraction $(1 - \phi)$ to find a constitutive law for compaction stress in the total mixture, that is, compaction pressure P_c :

$$P_c = -\xi_{ve}^* \dot{v}_{s,ve} + \chi_p^* \Delta P^o. \quad (35)$$

The compaction viscoelasticity ξ_{ve}^* and the compaction stress evolution parameter χ_p^* in the two-phase mixture then are defined as

$$\xi_{ve}^* = \frac{1 - \phi}{\frac{1}{\xi} + \frac{1}{K_\phi \Delta t}}, \quad \chi_p^* = \frac{1 - \phi}{1 + \frac{K_\phi \Delta t}{\xi}}. \quad (36)$$

2.2.5 Shear and tensile plastic failure

The theory of plasticity states that stresses induced by viscoelastic deformation of the host rock cannot exceed the plastic failure criterion. If viscoelastic stresses reach the maximum permissible stress state prescribed by the failure envelope, plastic failure occurs. A commonly used cohesive-frictional failure criterion for the mode-2 shear failure in geodynamic applications is that of Drucker–Prager plasticity (Paterson & Wong 2005), which is equivalent to the Mohr–Coulomb criterion in 2-D. In the case of two-phase flow, the dependence of plastic failure on effective pressure (eq. 19) according to Terzaghi's principle (Terzaghi 1923, 1943) needs to be considered, which states that increasing the fluid pressure in rock decreases the effective confining pressure, thus decreasing the yield strength for shear failure. Additionally, if effective pressure becomes negative (fluid is overpressured), tensile failure of the host rock may occur. The failure criterion for such mode-1 tensile fracturing is given by Griffith's criterion (Murrell 1964b). The combined failure criteria for both shear failure (mode-2 plasticity) and tensile failure (mode-1 plasticity) may be expressed in terms of the yield shear stress in the phase mixture $\bar{\tau}_y$ as

$$\bar{\tau}_{II} \leq \bar{\tau}_y = \begin{cases} C \cos \varphi + P_e \sin \varphi & \text{for } P_e > P_e^*, \\ \sigma_T - P_e & \text{for } P_e \leq P_e^*, \end{cases} \quad (37)$$

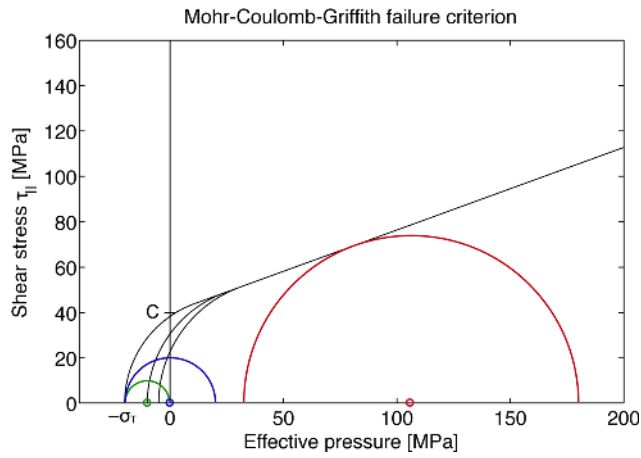


Figure 1. The combined shear and tensile plastic failure criteria expressed in terms of the Mohr–Coulomb and Griffith stress envelope, given for three values of tensile strength $\sigma_T = C/[2, 4, 8]$, displayed as the three black lines splitting of the frictional slope near the origin. Cohesion is $C = 40$ MPa, friction angle $\varphi = 30^\circ$. Also given are three characteristic stress states, plotted as Mohr–Coulomb circles for illustration. In green, a stress state leading to volumetric failure, in blue one leading to tensile fractures and in red, a stress state leading to shear fractures.

where $\bar{\tau}_{II} = (1 - \phi)\sqrt{\frac{1}{2}\tau_{ij}\tau_{ij}}$ (Einstein’s summation of repeated indices applies) is the second invariant of shear stress (radius of Mohr–Coulomb stress circle). P_e is the effective pressure (centre of stress circle) and $P_e^* = \frac{C \cos \varphi - \sigma_T}{1 - \sin \varphi}$ is the transition pressure where the two failure stress curves meet. C is the cohesion, φ the friction angle and $\sigma_T = C/R$ the tensile strength of the host rock with $2 \leq R \leq 8$ a parameter specifying the reduction of strength of the host rock under tensile stress (Cai 2010). Fig. 1 is a plot of the combined failure criteria. Note that the expression in eq. (37), in 2-D, states the radius of the maximum permissible stress circle as a function of its centre point. The yield envelope (black line in Fig. 1) is then defined as the line enveloping all permissible stress circles defined by the yield criteria in eq. (37).

The failure envelope not only sets a limit for the maximum permissible shear stress, but equally for the maximum permissible fluid overpressure (or minimum effective pressure P_e). This lower limit to permissible compaction stress may be expressed in terms of the yield pressure in the phase mixture P_y , which is found by rewriting the tensile yield criterion in eq. (37) in terms of P_e (bearing in mind that P_e is equal to compaction pressure P_c where a fluid phase is present):

$$P_e = P_c \geq P_y = \bar{\tau}_{II} - \sigma_T. \quad (38)$$

The implementation of the plastic component of shear and compaction deformation is achieved by an effective viscosity approach, where the viscoelastic constitutive law in locations of ongoing plastic failure is modulated by iteratively choosing both an effective shear η_{eff}^* and compaction viscosity ξ_{eff}^* such as to keep the local effective stress state $(\bar{\tau}, P_c)$ on the yield surface $(\bar{\tau}_y, P_y)$ (cf. Moresi et al. 2003; Kaus 2010).

The deviatoric part of the plastic failure rheology is first dealt with by restating eq. (30) in terms of second invariants of stress and strain rate tensors. During plastic failure, the yield shear stress $\bar{\tau}_y$ is substituted for $\bar{\tau}_{II}$ and the full visco-elasto-plastic shear strain rate $\dot{\mathbf{e}}'_s = \dot{\mathbf{e}}'_{s,\text{ve}} + \dot{\mathbf{e}}'_{s,\text{pla}}$ replaces the purely viscoelastic ones $\dot{\mathbf{e}}'_{s,\text{ve}}$. The resulting expression is a scalar (rather than tensorial) representation

of the constitutive law for shear stress, valid where plastic failure occurs:

$$\bar{\tau}_y = 2\eta_{\text{eff}}^* \dot{\mathbf{e}}'_{s,II} + \chi_\tau^* \tau_{s,II}^o \quad \text{for } \bar{\tau}_{II} = \bar{\tau}_y. \quad (39)$$

$\dot{\mathbf{e}}'_{s,II} = \sqrt{\frac{1}{2}\dot{\mathbf{e}}'_{s,ij}\dot{\mathbf{e}}'_{s,ij}}$ is the second invariant of deviatoric strain rates, and $\tau_{s,II}^o$ is the second invariant of advected deviatoric stresses of the previous time step. Solving this statement for the effective shear viscosity η_{eff}^* yields an expression by which the modified visco-elasto-plastic viscosity may be found. Note that the effective plastic viscosity remains equal to the viscoelasticity η_{ve}^* , where the local shear stress does not reach the failure criteria:

$$\eta_{\text{eff}}^* = \begin{cases} \frac{\bar{\tau}_y - \chi_\tau^* \tau_{s,II}^o}{2\dot{\mathbf{e}}'_{s,II}} & \text{for } \bar{\tau}_{II} = \bar{\tau}_y, \\ \eta_{\text{ve}}^* & \text{for } \bar{\tau}_{II} < \bar{\tau}_y. \end{cases} \quad (40)$$

The full visco-elasto-plastic constitutive law for shear stress in the host rock then becomes

$$\bar{\tau} = 2\eta_{\text{eff}}^* \dot{\mathbf{e}}'_s + \chi_\tau^* \bar{\tau}_s^o. \quad (41)$$

Although this effective viscosity approach to plasticity does not introduce an explicit statement for plastic shear strain rates $\dot{\mathbf{e}}'_{s,\text{pla}}$, these may be quantified in a post-processing step from the known values of total (eq. 13) and viscoelastic (eq. 21) shear strain rates as $\dot{\mathbf{e}}'_{s,\text{pla}} = \dot{\mathbf{e}}'_s - \dot{\mathbf{e}}'_{s,\text{ve}}$.

Next, the volumetric part of the plastic failure rheology is dealt with by substituting the compaction yield stress P_y for compaction pressure P_c and the total visco-elasto-plastic compaction strain rate $\dot{\nu}_s = \dot{\nu}_{s,\text{ve}} + \dot{\nu}_{s,\text{pla}}$ for the purely viscoelastic one $\dot{\nu}_{s,\text{ve}}$ in eq. (35). The resulting statement is again valid only where plastic failure occurs:

$$P_y = -\xi_{\text{eff}}^* \dot{\nu}_s + \chi_p^* \Delta P^o \quad \text{for } P_c = P_y. \quad (42)$$

Solving eq. (42) in terms of the effective compaction viscosity ξ_{eff}^* gives an expression, by which the modified visco-elasto-plastic viscosity may be found. Again, ξ_{eff}^* remains equal to the viscoelasticity ξ_{ve}^* where the local compaction pressure does not reach the failure criteria:

$$\xi_{\text{eff}}^* = \begin{cases} -\frac{P_y - \chi_p^* \Delta P^o}{\dot{\nu}_s} & \text{for } P_c = P_y, \\ \xi_{\text{ve}}^* & \text{for } P_c > P_y. \end{cases} \quad (43)$$

The full visco-elasto-plastic constitutive law for compaction stress in the host rock then becomes

$$P_c = -\xi_{\text{eff}}^* \dot{\nu}_s + \chi_p^* \Delta P^o. \quad (44)$$

The plastic compaction strain rates $\dot{\nu}_{s,\text{pla}}$ may be quantified in a post-processing step from the known values of total (eq. 12) and viscoelastic (eq. 32) compaction strain rates as $\dot{\nu}_{s,\text{pla}} = \dot{\nu}_s - \dot{\nu}_{s,\text{ve}}$.

The use of the effective viscosity approach to plasticity has the advantage of retaining the same form of the constitutive laws as in the viscoelastic case. The effect of plastic failure is captured in a continuum approach by limiting stress states in the model according to the combined shear and tensile failure criterion. As the numerical benchmarks and demonstration runs presented below indicate, the use of this rheology results in the localization of deformation in zones of shear or tensile failure, wherever stresses increase to the point of touching the failure envelope.

2.2.6 The viscous, elastic and plastic limits

The proposed constitutive laws for shear and compaction stress (eqs 41 and 44) describe visco-elasto-plastic deformation of the

solid phase. To arrive at these constitutive laws, we discretized the elastic time derivative of stresses with an implicit finite difference approximation. This yields constitutive laws for viscoelastic shear and compaction stress, on which a condition of maximum permissible stress, given by the combined failure criteria for shear and tensile plastic failure, is imposed. To confirm their validity, the properties of the two constitutive laws are examined in the three limiting cases of purely viscous, purely elastic and purely plastic deformation.

Both shear and compaction constitutive laws (eqs 41 and 44) consist of two terms. The first terms each take the form of a modified viscous rheology, where the shear and compaction viscosities are replaced by effective visco-elasto-plastic viscosities (η_{eff}^* , ξ_{eff}^*). Additionally, an elastic stress evolution term is introduced to both constitutive laws, which is governed by the dimensionless shear and compaction stress evolution parameters $\chi_{\tau,p}^*$.

If the effective stress state in the two-phase mixture ($\bar{\tau}$, P_c) does not reach the failure criteria ($\bar{\tau}_y$, P_y), the effective viscosities (η_{eff}^* , ξ_{eff}^*) remain equal to the shear and compaction viscoelasticities (η_{ve}^* , ξ_{ve}^*). The magnitudes of these viscoelastic parameters are governed by the magnitude of shear and compaction viscosities η and ξ and elastic moduli G and K_ϕ , and by the size of the time step under consideration. The timescale for which elastic deformation is dominant is given by the Maxwell time. The Maxwell time for shear deformation is $t_\tau = \eta/G$, and for compaction deformation $t_p = \xi/K_\phi$. On timescales larger than these Maxwell times, deformation is dominated by viscous deformation. Conversely, on timescales smaller than the Maxwell time, deformation is dominantly elastic. That the correct viscoelastic behaviour is indeed captured by such a combined viscoelastic rheology follows from eqs (29) and (36) (see Schmalholz *et al.* 2001; Kaus & Becker 2006), as

$$[\eta_{\text{ve}}^*, \xi_{\text{ve}}^*] \rightarrow \begin{cases} (1-\phi)[\eta, \xi] & \text{for } \Delta t \gg [t_\tau, t_p], \\ (1-\phi)[G\Delta t, K_\phi\Delta t] & \text{for } \Delta t \ll [t_\tau, t_p], \end{cases} \quad (45)$$

and

$$[\chi_\tau^*, \chi_p^*] \rightarrow \begin{cases} 0 & \text{for } \Delta t \gg [t_\tau, t_p], \\ 1 & \text{for } \Delta t \ll [t_\tau, t_p]. \end{cases} \quad (46)$$

Thus, the constitutive laws recover the purely viscous form, if the considered time step of deformation is large compared to the Maxwell time. For typical values of advective time step size and elastic modulus, this condition is satisfied where the intrinsic host rock viscosities are low:

$$\begin{aligned} \bar{\tau} &= (1-\phi)2\eta\dot{\epsilon}'_s, \\ P_c &= -(1-\phi)\xi\dot{\nu}_s. \end{aligned} \quad (47)$$

Furthermore, the constitutive laws recover the purely elastic form if the considered time step of deformation is small compared to the Maxwell time. For typical values of advective time step size and elastic modulus, this condition is satisfied where the intrinsic host rock viscosities are high:

$$\begin{aligned} \bar{\tau} &= (1-\phi)(2G\Delta t\dot{\epsilon}'_s + \bar{\tau}_s^o), \\ P_c &= (1-\phi)(-K_\phi\Delta t\dot{\nu}_s + \Delta P^o). \end{aligned} \quad (48)$$

Note that both the total deformation and stress evolution remain the same over time, independent of the considered time step.

If the viscoelastic stress evolution leads to effective stress levels in the two-phase mixture ($\bar{\tau}$, P_c) that locally reach the failure envelope ($\bar{\tau}_y$, P_y), the effective viscosities (η_{eff}^* , ξ_{eff}^*) are iteratively reduced according to eqs (40) and (43), in order to keep the local stress state

constant on the failure criterion. The resulting plastic deformation typically occurs in strongly localized failure zones and continues until local stresses are relaxed below the failure criterion again, at which point plastic failure ceases. The constitutive laws for the limiting case of pure plastic deformation then reduce to the simple statements of

$$\begin{aligned} \bar{\tau}_{II} &= \bar{\tau}_y, \\ P_c &= P_y. \end{aligned} \quad (49)$$

It may be argued that the elastic contribution to total deformation in most geodynamic contexts is insignificant compared to viscoplastic effects. Neglecting elasticity in computational models in geodynamics, however, leads to the numerically challenging situation where predicted viscous stresses are often far outside the failure criteria, and to iteratively bring them back to the yield surface poses problems (Kaus 2010). If elasticity is included, elastic stress evolution over time constrains the stresses in the model to approach the failure criteria step-by-step from below, a situation that is much less challenging for non-linear iterative solution strategies. The inclusion of elasticity thus not only adds to the physical description of rock deformation, but also helps to stabilize the numerical method, which was the primary motivation to include it here.

The proposed constitutive laws for shear and compaction stress are able to deal with all combinations of viscous, elastic and plastic deformation. A series of benchmark problems demonstrating the functionality and accuracy of the visco-elasto-plastic rheology and its implementation are presented in Appendix A2.

2.2.7 Melt-dependence of rheology

It has long been known that melt-bearing rock is weakened by the presence of even small amounts of melt. We therefore apply an exponential melt-weakening to the shear viscosity of the solid phase. This form of melt-weakening was proposed based on theory and experimental results (Kelemen *et al.* 1997; Mei *et al.* 2002; Simpson *et al.* 2010a,b) and has since been used by a number of studies on computational magma dynamics (e.g. Rabinowicz & Vigneresse 2004; Katz *et al.* 2006). It effectively captures the rapid decrease of shear viscosity at low to intermediate melt fractions and is expressed as

$$\eta = \eta_0 \exp(-\alpha_\phi \phi), \quad (50)$$

where η_0 is the intrinsic (or background) viscosity at zero melt fraction, which in this study is taken constant at typical values for Earth's mantle and crust. The exponential melt-weakening factor has been experimentally constrained to $25 \leq \alpha_\phi \leq 30$ (Mei *et al.* 2002). A fully consistent formulation of melt-weakened shear viscosity at all levels of melt fractions is discussed by Schmeling *et al.* (2012), however, as the shear viscosity for melt fractions larger than 30–50 per cent drops below the minimum cut-off value of 10^{16} Pa s necessary for numerical stability (see Section 2.5.3 below), the exponential weakening law is sufficient to capture melt-weakening of shear viscosity at melt fractions lower than that.

The compaction viscosity is defined relative to the intrinsic shear viscosity of the host rock. We use an inverse melt fraction law to relate the compaction viscosity to the shear viscosity of the host rock (Takei & Holtzman 2009; Simpson *et al.* 2010a; Schmeling *et al.* 2012):

$$\xi = \eta_0 \phi^{-p}. \quad (51)$$

The melt fraction exponent is taken as $p = 1$ in this study, consistent with theoretical work (Batchelor 1967; Schmeling 2000; Bercovici *et al.* 2001b; Simpson *et al.* 2010a).

The poro-elastic modulus of the host rock quantifies the change of melt fraction due to changes in compaction pressure over time; it depends on melt fraction in a way that lets compressible deformation cease when melt fraction goes to zero:

$$K_\phi = K_0 \phi^{-q}. \quad (52)$$

Here, K_0 is the reference pore modulus, which typically assumes values of the order of $1 \leq K_0 \leq 100$ GPa. The exponent of its melt fraction-dependence is taken as $q = 1/2$, a value found from experiments (Hall 1953; Jalal 2006a,b). Our formulation of elastic pore compressibility is equivalent to the form proposed in two previous numerical studies of viscoelastic two-phase flow in geodynamics (Connolly & Podladchikov 1998; Morency *et al.* 2007).

2.3 Final governing equations

We now return to the governing equations of two-phase flow in order to bring them into a more suitable form.

For the final form of the total momentum equation, we substitute the constitutive law for shear stress (eq. 41) and the sum of compaction pressure and fluid pressure for the total pressure (eq. 18) into the momentum equation of the two-phase mixture (eq. 8), yielding an expanded form of the momentum equation:

$$\nabla \cdot (2\eta_{\text{eff}}^* \dot{\boldsymbol{\epsilon}}'_s + \chi_r^* \boldsymbol{\tau}_s^o) - \nabla P_f - \nabla P_c - \bar{\rho} g \hat{\mathbf{z}} = 0. \quad (53)$$

By substituting the sum of solid velocity and phase separation flux $\mathbf{v}_s + \mathbf{q}$ for the averaged mixture velocity $\bar{\mathbf{v}}$ into the total mass conservation equation (eq. 5) and then replacing \mathbf{q} by Darcy's law according to eq. (10), we expand the total mass conservation to

$$\nabla \cdot \mathbf{v}_s - \nabla \cdot K_D [\nabla P_f + \rho_f g \hat{\mathbf{z}}] = 0. \quad (54)$$

The constitutive law for compaction deformation still needs to be incorporated into the final set of governing equations. Therefore, we choose eq. (44), rewritten in terms of the compaction strain rate as the third final equation.

$$\dot{\nu}_s + \frac{1}{\xi_{\text{eff}}^*} P_c - \frac{\chi_p^*}{\xi_{\text{eff}}^*} \Delta P^o = 0. \quad (55)$$

In the absence of melting and crystallization, changes in melt fraction are due to visco-elasto-plastic volumetric deformation only. In order to find a constitutive law for melt fraction evolution over time, we rearrange the solid mass conservation equation (with $\dot{\nu}_s = \nabla \cdot \mathbf{v}_s$) to yield

$$\frac{D_s \phi}{Dt} = (1 - \phi) \dot{\nu}_s, \quad (56)$$

thus relating the visco-elasto-plastic volumetric strain rate to the evolution of melt fraction over time. Adding eq. (56) as the fourth of the final governing equations, we now have a complete description of the mechanical aspects of two-phase flow in a visco-elasto-plastically deforming mantle, lithosphere and crust.

The final set of governing equations consists of four equations: the expanded total momentum and mass conservation equations for the two-phase mixture, the compaction equation and the melt evolution equation. These four equations contain the four unknowns: solid velocity \mathbf{v}_s , fluid pressure P_f , compaction pressure P_c and melt fraction ϕ . After moving terms independent of solution variables to the right-hand side and replacing strain rates with the respective

derivatives of velocity, the final governing equations become:

$$\begin{aligned} \nabla \cdot \left[\eta_{\text{eff}}^* (\nabla \mathbf{v}_s + [\nabla \mathbf{v}_s]^T) - \frac{2}{3} \eta_{\text{eff}}^* \nabla \cdot \mathbf{v}_s \mathbf{I} \right] - \nabla P_f - \nabla P_c \\ = \bar{\rho} g \hat{\mathbf{z}} - \nabla \cdot \chi_r^* \boldsymbol{\tau}_s^o; \end{aligned} \quad (57)$$

$$\nabla \cdot \mathbf{v}_s - \nabla \cdot K_D \nabla P_f = \nabla \cdot K_D \rho_f g \hat{\mathbf{z}}; \quad (58)$$

$$\nabla \cdot \mathbf{v}_s + \frac{P_c}{\xi_{\text{eff}}^*} = \frac{\chi_p^* \Delta P^o}{\xi_{\text{eff}}^*}; \quad (59)$$

$$\frac{D_s \phi}{Dt} - (1 - \phi) \nabla \cdot \mathbf{v}_s = 0. \quad (60)$$

As mentioned in Section 1, energy conservation, melting and crystallization and compositional evolution of the magmatic system are not dealt with here. In the future, such an addition to this physical model could be made without fundamental changes to the mechanical system of equations proposed above. The inclusion of such a treatment requires the addition of an equation for the conservation of energy in a two-phase system (see Bercovici & Ricard 2003; Šrámek *et al.* 2007) and one for the conservation of composition, along with a thermodynamically consistent closure condition determining melt fraction as a function of temperature, composition and pressure (Katz 2008).

In the limiting case of purely viscous deformation, these proposed governing equations are equivalent to previously proposed formulations of geodynamic two-phase flow, most of which are versions of the formulation introduced by McKenzie (1984). To confirm the validity of these equations for all possible melt fractions, their properties need to be considered in two significant limiting case scenarios: the zero melt limit and the full melt limit, which we proceed to do in the following two sections.

2.4 The zero melt limit

The particular form of writing the total momentum and mass conservation equations for the two-phase mixture adopted here (eqs 57 and 58) has a distinct advantage. In the zero melt limit ($\phi = 0$), the equations reduce to the standard form of the single-phase Stokes equations. If melt fraction goes to zero, the solid velocity \mathbf{v}_s becomes equal to the total velocity $\bar{\mathbf{v}}$ and solid shear stress $\boldsymbol{\tau}_s$ becomes equal to the total stress $\bar{\boldsymbol{\tau}}$. Furthermore, the compaction pressure P_c and its gradient vanish, leaving only one pressure component in the momentum equation, the fluid pressure P_f , which now becomes equal to the total pressure \bar{P} (see Table 1). The Darcy coefficient $K_D = k_\phi / \eta_f$ goes to zero as the permeability goes to zero, cancelling out the terms in eq. (58) relating to the divergence of phase separation flux \mathbf{q} .

Furthermore, the compaction and melt evolution equations (eqs 59 and 60) both become equal to the Stokes mass conservation or incompressibility condition and thus become redundant. Thus, the standard Stokes equations for a single-phase medium

$$\begin{aligned} \nabla \cdot \bar{\boldsymbol{\tau}} - \nabla \bar{P} &= \bar{\rho} g \hat{\mathbf{z}}, \\ \nabla \cdot \bar{\mathbf{v}} &= 0, \end{aligned} \quad (61)$$

are recovered in the zero melt limit, describing the deformation of the pure solid phase when no fluid phase is present.

2.5 The full melt limit

Most treatments of two-phase flow in geodynamics since Sleep (1974) have neglected shear stresses in the fluid phase. As a consequence, the fluid momentum equation may be stated in the form of Darcy's law (eq. 10) and the total shear stress tensor is reduced to the contribution of solid shear stresses (eq. 15). This simplification is well accepted for two-phase flow at low melt fractions, where the fluid phase percolates through the pore space of a structurally coherent solid phase. In the low melt fraction context, the argument for the validity of this assumption is twofold: (i) due to the inherent tortuosity of fluid percolation in the intergranular pore space, the contributions of the fluid shear stress divergence in the momentum conservation of the fluid phase likely cancel out over a typical volume considered in the continuum approach; (ii) due to the large contrast in viscosities, the magnitude of the fluid shear stresses in the total shear stress tensor is very small compared to the magnitude of the solid shear stresses, and therefore fluid shear stresses may be neglected in the total shear stress tensor of the two-phase mixture.

2.5.1 Host rock disaggregation

However, if melt accumulation leads to the disaggregation of the rock matrix and potentially the emergence of fully molten magma bodies, these assumptions need to be revisited to ensure the validity of the proposed governing equations at high melt fractions. As mentioned previously, at melt fractions above a threshold located at around 25 per cent, the solid phase starts to disaggregate into a mush of grains partly or fully submerged in the abundant melt phase. In this configuration, the grains may still interlock to a certain degree, and thus the solid phase may continue to sustain some level of shear stress up to melt fractions of around 60 per cent, depending on grain size and geometry. Phase separation beyond the disaggregation threshold occurs by gravitational grain settling, which is hampered by entrainment of grains in the increasingly unhindered flow of the accumulating melt phase. The effective solid shear viscosity decreases drastically at intermediate melt fractions, until it eventually drops to zero (Schmeling *et al.* 2012). Thus, the behaviour of the two-phase mixture at very high melt fractions is increasingly dominated by the flow of the melt phase, until the few remaining solid grains become fully entrained in the melt flow.

2.5.2 Permeability

The question is often posed of whether Darcy's law (eq. 10) is a valid representation of the fluid momentum conservation at melt fraction beyond host rock disaggregation and up to the full melt limit. The answer to this question, among other factors discussed below, hinges upon the physical understanding of permeability. As the name suggests, permeability is understood as a property of the host rock matrix quantifying how efficiently fluids may permeate through its pore space. The term permeability therefore seems to imply the presence of a coherent rock matrix. This particular understanding of permeability evidently breaks down when the host rock matrix disaggregates. However, it has been previously argued that Darcy's law (eq. 10) is general enough to capture the relative motion of the components of disaggregated host rock (grains, blocks, etc.) against the fluid phase flowing around them (Batchelor 1967). In this case, permeability, interpreted as the obstruction to phase separation flux exerted by the solid phase on the fluid phase and vice versa, remains a valid concept. Abe (1995) suggests a modified permeability law that captures this situation more adequately

by replacing permeability at intermediate and high melt fraction with a parametrized law representing the obstruction to flow characteristic of ideal Stokes spheres sinking through the melt phase. Additionally, the effects of surface tension on grains submerged in melt might have to be considered, as it may reduce the efficiency of complete separation of grains from melt in a high melt fraction magma body (Hier-Majumder *et al.* 2006). For the sake of simplicity, we will here assume a straightforward modification of a standard melt fraction law for permeability given by the relation

$$k_\phi = k_0 \phi^n (1 - \phi)^m, \quad (62)$$

with k_0 denoting the reference permeability. The melt fraction exponent is set to $n = 3$, a value representing interconnected disc-shaped melt pockets along grain boundaries (Faul 1997). The exponent assigned to the solid fraction we set to $m = 2$. The choice to include a dependence on solid fraction into the permeability law is based on the concept that a permeability law should reflect the most important aspects of two-phase flow at high melt fractions as well as the more well-established behaviour of porous flow. First, physical consistency demands that drag forces in the zero and full melt limits decrease to zero. With a permeability law depending on both melt and solid fraction, this condition is satisfied. Second, a dependence on solid fraction should in some way reflect grain size (see discussion of drag forces in Bercovici *et al.* 2001a), thus giving a valid reason to include the square of the solid fraction. Here, we use the solid fraction as a non-dimensional representation of grain size, capturing the reduction of drag forces with increasing melt fraction to a good first order. The resulting material fraction-dependent behaviour of phase separation flux is given in Fig. 2. The curves are found by measuring the relative magnitude of phase separation flux \mathbf{q} taken from a series of single time step, purely viscous, homogeneous box simulations of phase separation performed at various level of melt fraction. The blue curve gives phase separation flux values resulting from the permeability law used here, with $n = 3$ and $m = 2$, and the red curve gives phase separation flux values resulting from a permeability law with $n = 3$ and $m = 1$ for comparison (*cf.* phase separation flux curve given in Abe 1995). Note that the inclusion of the solid fraction only significantly affects the permeability law at melt fraction larger than 70 per cent. In Fig. 2, we additionally mark the approximate domain boundaries, where the two-phase mixtures take on the different styles of phase separation discussed above.

2.5.3 Total stress approximation

Moving on to the examination of the system of equations for melt fraction going to unity, we find that fluid velocity \mathbf{v}_f becomes equal to the total velocity $\bar{\mathbf{v}}$ and the fluid shear stress $\boldsymbol{\tau}_f$ becomes equal to the total stress $\bar{\boldsymbol{\tau}}$. The compaction pressure P_c and its gradient vanish as in the zero melt limit, leaving only one pressure component in the momentum equation, the fluid pressure P_f , which again becomes equal to the total pressure \bar{P} (see Table 1). Thus, the total momentum equation becomes equal to the Stokes momentum equation. With no fluid phase present, the total mass conservation, the compaction equation and the melt fraction evolution equation all take the form of the incompressibility condition, and thus, the standard Stokes equations for a single-phase medium are recovered again in the full melt limit (see eq. 61), now describing the deformation of the pure fluid phase when no solid phase is present.

Returning to the assumption of negligible fluid shear stresses, we argue that in the definition of the Darcy flux (eq. 10), this assumption remains valid. The convective flow field of the fluid

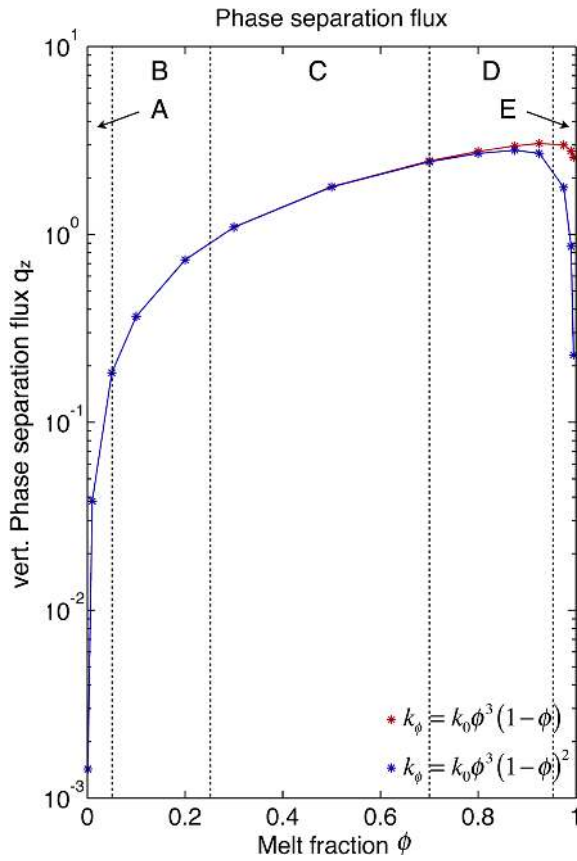


Figure 2. Phase separation flux \mathbf{q} as a function of melt fraction. Data points are relative magnitudes of \mathbf{q} measured in the middle of a series of single time step, homogeneous compaction models run at various melt fractions. High values of \mathbf{q} stand for high-velocity difference between the phases, values of \mathbf{q} going to zero signify ceasing of two-phase flow in favour of single-phase Stokes flow of the remaining phase. The red curve gives values resulting from a permeability function with a melt fraction exponent of $n = 3$ and a solid fraction exponent of $m = 1$; the blue curve gives values with $n = 3$ and $m = 2$, the values used in simulations here. Domains A–E mark melt fractions at which the two-phase mixture takes on a different style of phase separation. In domains A and E, single-phase Stokes flow of the solid and melt phase is dominant, respectively. In domain B, melt percolates through an increasingly weakened, but coherent rock matrix. In domain C, the two-phase mixture becomes a mush of disaggregated grains settling with respect to melt flow. In domain D, grain settling becomes hampered as remaining grains are entrained by the melt flow.

phase in partial melt at higher melt fractions may be of a somewhat lower tortuosity than that of pervasive flow through a low-porosity matrix rock, yet the contributions of shear stress divergence to the momentum conservation of the fluid phase likely still average out over a typical volume considered in the continuum approach. Due to the low viscosity of silicate melt, the vigorous convective motion in the melt phase occurs on a time and length scale that may not be adequately resolved by current numerical methods on a crustal or lithospheric scale. As, however, a melt-filled dyke, channel or other magma body larger than a metre in size could be resolved even by one-phase Stokes flow, two things need to be considered in addition to the argument of tortuous convective flow. First, the solid stresses in the walls of such fully molten structures are orders of magnitude larger than the fluid stresses in the molten body, and thus, the overall deformation of the system will be governed by the deformation of the solid. The deformation of the melt itself mainly determines the efficiency of internal mixing in a magma body, or the rate of flux

through a dyke or channel. Second, the use of real melt viscosities of the order of 1–1000 Pa s in fully molten bodies of crustal-scale numerical simulations invariably leads to numerical instabilities in the velocity solution. Mainly for this numerical reason, a minimum cut-off viscosity $\eta_{\text{cut-off}}$ is applied in the total shear stress tensor, resulting in a high melt fraction approximation of shear stresses in the two-phase mixture (bearing in mind that towards the full melt limit $\dot{\epsilon}'_s \approx \dot{\epsilon}'_f$):

$$\bar{\tau} \approx 2\eta_{\text{cut-off}}\dot{\epsilon}'_s \quad \text{for } \eta_{\text{eff}}^* \leq \eta_{\text{cut-off}}. \quad (63)$$

By imposing this lower cut-off viscosity, small-scale flow features are effectively filtered out in regions of high melt fraction. A fully resolved treatment of the fluid phase at high melt fractions was recently presented by Dufek & Bachmann (2010). Their numerical models of magma chamber evolution, however, are performed with a spatial resolution of 0.5 m, on a model scale of tens of metres. For the problem of melt ascent through lithosphere and crust, spatial resolution is typically limited to an element size upwards of 100 m, and therefore, a lower viscosity cut-off of the order of $1e15$ to $1e17$ Pa s remains necessary.

2.6 Numerical implementation

Here, we give a short overview over the numerical implementation of the methodology derived above. For the details of the implementation, including the discretized form of the governing equations, along with a set of benchmark problems to demonstrate the functionality and accuracy of the methodology, the readers are referred to the Appendix.

The simulation code employs an Arbitrary–Lagrangian–Eulerian primitive variable approach on a deformable, structured quadrilateral finite-element mesh with a marker-based advection scheme. Linear and elementwise constant shape and test functions are used to formulate the equations in the weak form and discretize them on a finite-element mesh.

The system of equations derived above is nonlinear in terms of melt fraction and solid velocity, where nonlinearities are introduced in the form of melt fraction-dependent material properties, and through the two plastically weakened effective viscosities. Therefore, the solution strategy for each time step involves an iterative solver method to treat the nonlinearities, where each iteration consists of a direct solve of eqs (57)–(59) as a linear system of equations for the solution variables of solid velocity, fluid pressure and compaction pressure, followed by an update of the current melt fraction and all nonlinear material properties (i.e. permeability, elastic pore modulus, effective shear and compaction viscosities) according to the current velocity–pressure solution and melt fraction. More details on the nonlinear solver are given in Appendix A1.

As the melt evolution over time (eq. 56) is treated as an explicit update, the time derivative is discretized using a second-order Crank–Nicolson scheme, to yield:

$$\phi^{i+1} = \phi^o + \frac{1}{2} [(1 - \phi^o) \dot{v}_s^o + (1 - \phi^i) \dot{v}^i] \Delta t, \quad (64)$$

where superscript $()^o$ refers to properties retrieved from the previous time step and superscript $()^i$ refers to properties retrieved from the i th non-linear iteration.

Material properties are advected with either solid or fluid velocity, depending on which material phase they are derived from. Note that although melt fraction is related to the melt phase, we choose a form of the solid mass conservation equation (eq. 60) to track the time evolution of the melt fraction, and therefore, the melt

fraction is advected with solid velocity. Advection is implemented into the code by using a marker-in-cell approach, where all advected properties are interpolated onto one of two sets of marker particles (one set of markers for each material phase), which are then moved with their respective velocity field (Gerya & Yuen 2007). The main advantage of utilizing markers for advection is that with appropriate interpolation functions from numerical grid to marker, it may reduce the numerical diffusion typically associated with grid-based advection schemes. However, the large number of marker particles needed for accurate results renders this method fairly expensive in terms of memory, and interpolation methods need to be formulated efficiently. Details about the advection scheme and the interpolation method used in our code are provided in the Appendix.

The overall workflow of the code may be outlined by the following sequence:

1. Pre-processing and initialization of all variables and material properties.
2. Start main time-stepping loop.
 - a. Solve system of equations for velocity–pressure solution and update melt fraction.
 - i. Update nonlinear material properties according to current solution guess.
 - ii. Assemble coefficient matrix and right-hand side of linear system of equations.
 - iii. Check non-linear residual of current solution guess.
 - iv. Proceed with steps iv to vi if convergence criterion is not achieved.
 - v. Solve linear system of equations (using Matlab’s ‘backslash’ direct solver).
 - vi. Update melt fraction according to current solid velocity divergence.
 - vii. Check time step according to courant criterion and adapt if necessary.
 - viii. Go to step i for next non-linear iteration.
 - b. Advect free surface if specified and adapt finite-element mesh accordingly.
 - c. Advect properties of host rock and melt phase on marker fields.
 - d. Check time and loop back to step a for next time step if desired.
3. Post-processing and visualization.

In order to obtain an efficient finite-element implementation, able to deal with several hundred thousand degrees of freedom, a MILAMIN-style vectorization is used to speed up the assembly of the global stiffness matrix (Dabrowski *et al.* 2008).

3 NUMERICAL RESULTS

3.1 Model setup

In this section, we demonstrate the effects of the proposed visco-elasto-plastic rheology on the mode and efficiency of melt extraction by running a suite of numerical simulations on a simple model setup. See Table 2 for all relevant model parameter values chosen to represent a silicate melt ascending through a homogeneous host rock. Fig. 3 displays the initial melt distribution used in all runs, a Gaussian melt pulse located in the middle of the lower boundary with an amplitude of 20 per cent melt fraction, surrounded by un-molten host rock.

To simulate a progression from asthenospheric conditions (weak, ductile host rock) towards crustal conditions (competent, brittle

Table 2. Model parameters.

Symbol	Name	Unit	Value
D	Box depth	km	4
W	Box width	km	6
ρ_s	Host rock density	kg m^{-3}	3000
ρ_f	Melt density	kg m^{-3}	2500
k_0	Reference permeability	m^2	$1e-8$
η_f	Fluid viscosity	Pa s	10
η_0	Reference solid viscosity	Pa s	$1e+18 - 1e+23$
G	Shear modulus	GPa	50
K_0	Ref. pore modulus	GPa	5
C	Cohesion	MPa	40
φ	Friction angle	$^\circ$	30
α_ϕ	Melt weakening factor	—	27
$\dot{\epsilon}_{\text{BG}}$	background strain rate	s^{-1}	$1e-15$
ϕ_{crit}	critical melt fraction	per cent	$1e-3$

Note: Model parameter values used for all simulation runs, if not specified otherwise in the text. Values are chosen to represent generic host rock and silicate melt properties in the context of mantle melt percolating through a continental lithosphere and crust.

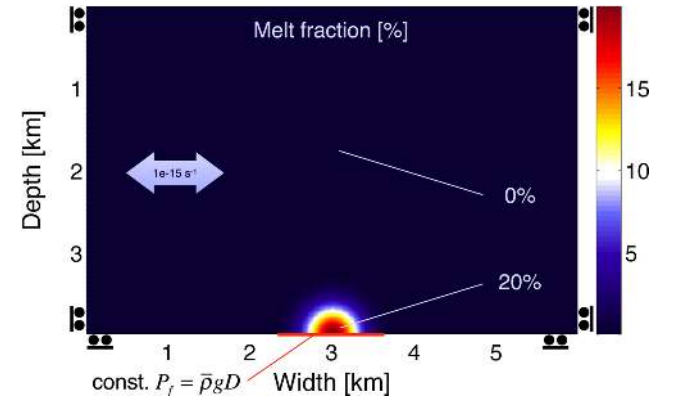


Figure 3. Initial melt distribution for suite of rheological demonstration runs, consisting of a 2-D Gaussian pulse with amplitude of 20 per cent against a homogeneous background of 0 per cent melt fraction. The red bar underneath the lower boundary indicates the area where a constant lithostatic boundary condition is applied on the fluid pressure, mimicking a reservoir of melt in pressure equilibrium with the host rock beneath the model. An extensional normal velocity, corresponding to a constant strain rate of $\dot{\epsilon}_{\text{BG}} = 1e-15 \text{ s}^{-1}$, is imposed on the side boundaries, along with a zero tangential stress boundary condition. The top boundary is held stress-free.

host rock), a series of runs with increasing reference viscosity of the lithospheric host rock η_0 are employed. In addition, we also vary the tensile strength σ_T of the rock phase (see Fig. 1 for the meaning of σ_T in relation to plastic failure) in order to study different onset levels of volumetric plastic yielding, a feature relevant to the mode and efficiency of melt transport.

Field evidence suggests that major plutonic rock complexes are often closely associated with zones of tectonic activity (Pitcher 1979; Pelletier *et al.* 1989; McCaffrey 1992; Petford & Atherton 1992). In order to include the effects of regional tectonic stresses, kinematic boundary conditions are applied on the side boundaries (constant strain rate). Apart from this uniform background strain rate, the boundaries are shear stress free on the side and the bottom boundaries and stress free on the top. For the moment, we limit our focus on extensional tectonics. The direction of least compressive stress for a pure shear extensional stress field is horizontal, and

Table 3. Simulation runs.

Run ID	η_0 [Pa s]	σ_T
d118r[2,4,8]	$1e + 18$	$C/[2, 4, 8]$
d119r[2,4,8]	$1e + 19$	$C/[2, 4, 8]$
d120r[2,4,8]	$1e + 20$	$C/[2, 4, 8]$
d121r[2,4,8]	$1e + 21$	$C/[2, 4, 8]$
d122r[2,4,8]	$1e + 22$	$C/[2, 4, 8]$
d123r[2,4,8]	$1e + 23$	$C/[2, 4, 8]$

Note: List of model runs presented in Section 3 (Results). Given are the choice of intrinsic host rock viscosity η_0 and tensile rock strength σ_T in relation to cohesion $C = 40$ MPa for each run.

thus, extensional boundary conditions likely produce subvertical melt extraction features propagating upwards from the initial melt pulse, along the direction normal to the least compressive stress.

On the lower boundary, underlying the initial melt pulse, a region of constant lithostatic fluid pressure is imposed, thus simulating ongoing melt supply at the lower boundary (displayed as a red bar in Fig. 3). On the top boundary, a surface pressure of 50 MPa is added, equivalent to around 1.5–2 km of rock overburden, thereby increasing the confining pressure governing the frictional yield strength, without otherwise influencing the dynamics of the simulation. This condition serves to adjust the onset of shear plasticity for brittle modes of deformation, representing conditions within the upper crust. Runs at lower background viscosity, representing lower crustal or asthenospheric conditions, are not affected by the increased confining pressure.

All simulations are run in a rectangular 2-D box of 360×240 elements (347 400 degrees of freedom, 2.76 million markers), resulting in a spatial resolution of 16.6 m over a model depth of 4 km. Each simulation typically runs to a total of 500–1500 time steps, taking up to 128 hr to complete on a single AMD Opteron node of the Brutus cluster in Zurich. All model parameters employed in the simulations presented in this section are listed in Table 3.

We find that at least three distinct regimes of melt extraction emerge, involving both fracture and flow of the host rock. These three regimes are best characterized by their primary feature of melt transport, which are melt diapirs, decompaction channels and tensile fractures, forming in this sequence with increasing viscosity of the host rock. We will now discuss the results in terms of three criteria, which we will assess with the following questions: (i) Geometry: is there a characteristic shape, size and orientation of melt transport features and are they of a distributed or localized nature? (ii) Deformation: What is the dominant style of deformation (viscous, elastic and plastic)? Is the flow field characterized by convection, compaction or fracturing? (iii) Efficiency: Are there melt transport features that are more efficient in extracting melt from the source than others, and on which property of the system does this difference depend?

3.2 Melt diapirism

The first regime of melt transport emerges at low host rock viscosities of $\eta_0 \leq 1e + 20$ Pa s, values that are typical for the asthenospheric mantle. The characteristic melt transport feature in this regime is a melt diapir: a local upwelling characterized by convective flow driven by the buoyancy of accumulated melt, by which melt is transported away from the source region. Fig. 4 displays three snapshots of the run d118r2, the results of which are representative of this regime ($\eta_0 = 1e + 18$ Pa s, $\sigma_T = 2 = 20$ MPa,

see Movies S1 and S2 for animated time evolution), giving the melt fraction (left), along with the volumetric strain rate field (right column of panels) at three points of time during the simulation. Overlaying the melt fraction and volumetric strain rate plots are arrows of melt velocity and solid velocity, respectively. For improved clarity of observation, velocity arrows are plotted after subtracting the extensional pure shear component of the velocity fields imposed by the kinematic boundaries.

3.2.1 Geometry of melt diapirs

The dominant feature of melt transport in this regime approximates the shape of a Stokes sphere, rising vertically through the ductile host rock. In Fig. 4, we identify three stages of diapir evolution, which are incipient compaction wave formation, followed by diapir growth, and finally, diapir ascent. In the initial stage (top panels), melt collects in an approximately spherical, distributed peak characteristic of two-dimensional compaction waves (Scott & Stevenson 1984; Scott 1988).

In a second stage (middle panels), with more melt supplied from the source region imposed on the lower boundary, the incipient compaction wave takes on the characteristics of a more sharply bounded magma body with up to 100 per cent melt content. At this stage, however, the diapir remains stationary at the lower boundary, as it grows in size and melt content by percolation from below. In a third stage, the diapir reaches a stable radius, detaches from the boundary and rises up in an approximately constant spherical shape with constant melt content.

The initial volume of melt is given by a 2-D Gaussian imposed on an otherwise zero melt fraction domain. The Gaussian remains non-zero across the domain, yet the melt fraction at some point drops below 10^{-3} which we choose as the cut-off level for two-phase physics to apply (see detailed discussion in Appendix A4). Permeability is therefore non-zero everywhere as well, although it is cut-off at 10^{-19} m² in our simulations to ensure stability of the Q1Q1 elements. Therefore, the injected melt may permeate into the parts of the domain where the initial melt fraction is below 10^{-3} , although that permeable flow is very slow and thus easily overtaken by convective motion at low rock viscosities, or other modes of deformation at later stages (see subsequent sections). In this context, the mode of two-phase flow with the originally imposed melt volume can thus be classified within the one observed in 2-D compaction waves in contrast to convective flow observed in the diapirism mode. As such, length scales of the upwelling below depend on the compaction length scale.

In a partially molten region of the asthenosphere, melt is likely to collect into spherical solitary waves, if the melt source region is larger than the characteristic length scale of melt percolation, given by the compaction length δ_c (McKenzie 1985)

$$\delta_c = \sqrt{K_D \xi^*}. \quad (65)$$

Melt extraction by percolation is driven mainly by the buoyancy contrast between solid and fluid phases, characterized by a characteristic fluid overpressure p_c

$$p_c = \Delta \rho g \delta_c, \quad (66)$$

and a characteristic percolation velocity w_c

$$w_c = K_D \Delta \rho g. \quad (67)$$

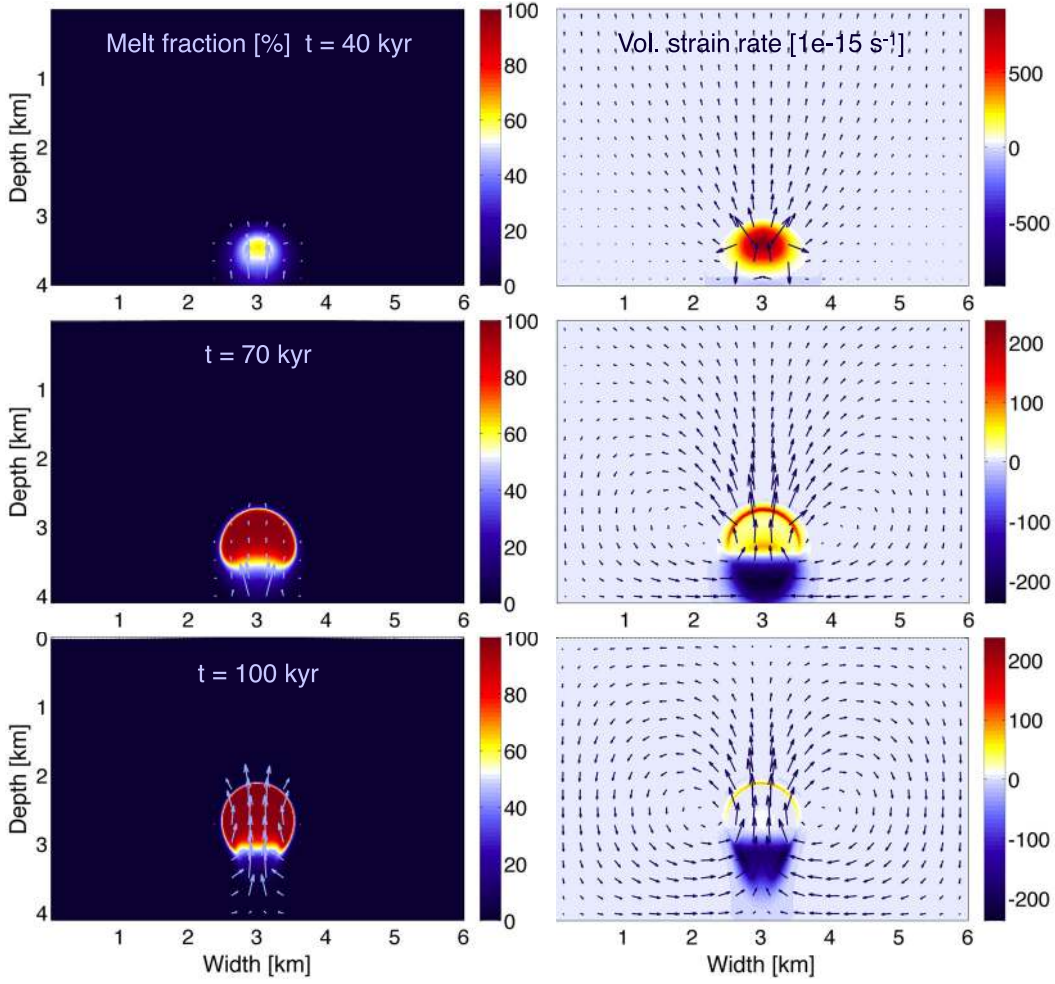


Figure 4. Snapshots of run d118r2 ($\eta_0 = 1e+18$ Pa s, $\sigma_T = 20$ MPa) resulting in melt diapirism, displaying melt fraction in per cent (left) overlaid with melt velocity arrows, and volumetric strain rate scaled by background strain rate (right) overlaid with solid velocity arrows. Both velocity fields are reduced by the pure shear extensional component. Velocity arrows are scaled with relative magnitude.

The related compaction flow of the rock matrix is characterized by a compaction velocity scale v_c , given by

$$v_c = \frac{\Delta\rho g \delta_c^2}{\xi^*}. \quad (68)$$

The rise of a fully molten diapir of radius R , however, is characterized by the velocity of a Stokes sphere u_c , which, apart from some geometric constant, is given by the relation

$$u_c \sim \frac{\Delta\rho g R^2}{\eta^*}. \quad (69)$$

It follows from eqs (65) and (69) that both the compaction length δ_c and the diapir radius R are related to the host rock viscosity η . Therefore, with an increase in background viscosity, we expect to find larger melt diapirs forming from larger incipient compaction waves, and building up higher fluid overpressure along the way. Comparing the outcome of the three runs d118r2, d119r2 and d120r2 with increasing background viscosities from $1e+18$ to $1e+20$ Pa s in Fig. 5 demonstrates that the diapir radius and the maximum fluid overpressures indeed both increase with increasing host rock viscosity. Here, the theory suggests that both diapir radius and magnitude of overpressure should scale with the square root of the viscosity. In the runs presented in Fig. 5, however, this theoretical prediction

cannot be directly verified, probably because factors such as locally variable compaction length and interaction with model boundaries obscure the situation.

3.2.2 Deformation in melt diapirism regime

Examining the relative magnitude of the viscous, elastic and plastic components of deformation reveals that low viscosity runs, in which diapirs are formed, are entirely dominated by viscous deformation. A convective flow pattern forming around the diapir in both phase velocities (see velocity arrows in Fig. 4) is characteristic of the melt diapirism regime. With ongoing melt supply from the lower boundary, the melt fraction inside the diapir increases beyond the host disaggregation threshold, creating a magma body of melt fractions of up to 100 per cent, constituting the head of the diapir. As the diapirs increase in size with increasing background viscosity, some small-scale downwellings, driven by differences in solid fraction, start dropping from the roof of the magma body (Fig. 5). This process is triggered when small-scale internal convection is faster than the overall rise of the diapir. We assume that a more inhomogeneous model setup would favour the onset of such small-scale convection, as it is driven by perturbations in the melt fraction distribution,

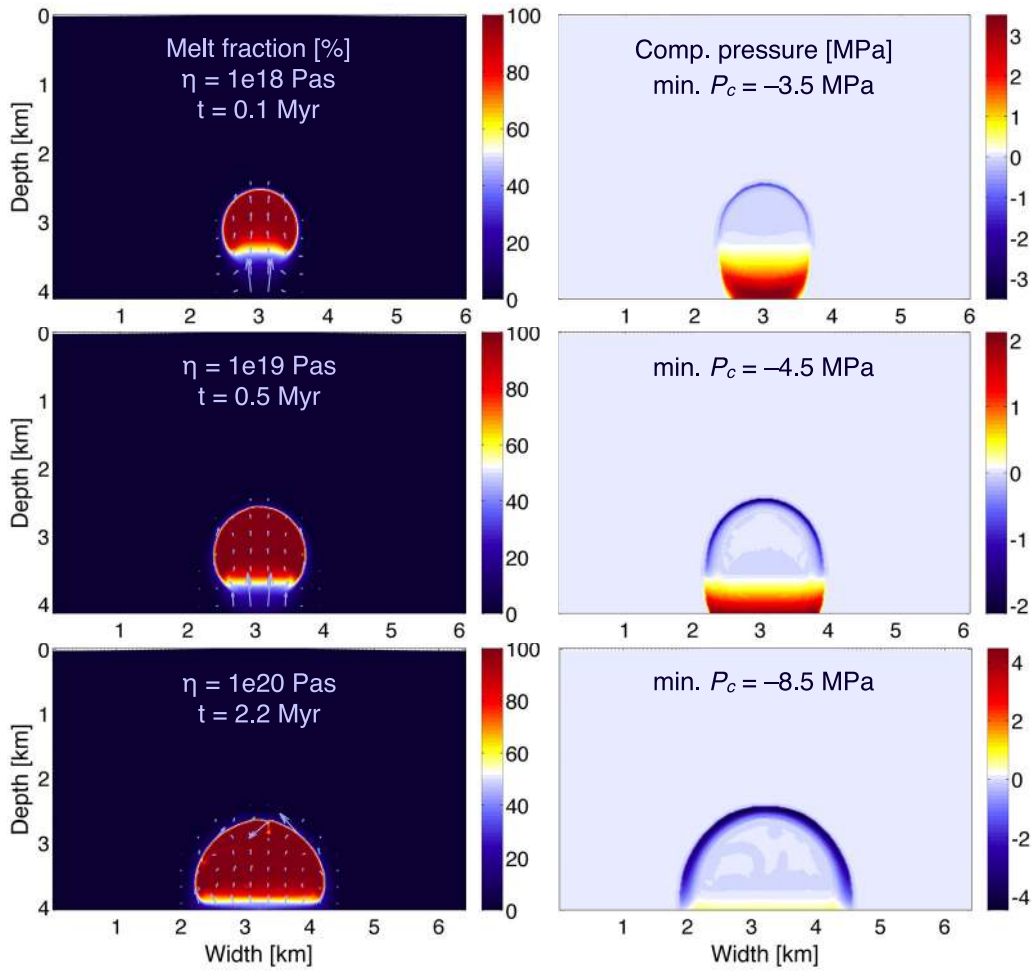


Figure 5. Snapshots of simulations d118r2 (top), d119r2 (middle) and d120r2 (bottom), taken at the time when the melt diapir has reached a stable radius. Displayed are melt fractions in per cent (left), overlaid with arrows of melt velocity as in Fig. 4, and compaction pressure P_c in MPa (right), negative values correspond to melt overpressure, minimum value of P_c measured over the model time is given for each simulation.

which here only derive from fluctuations due to limited numerical accuracy along the relatively sharp upper boundary of the diapir.

Neither shear nor tensile plastic failure are relevant in runs d118r2, d119r2 and d120r2 (Fig. 5), as stress states in these runs remain below the failure criteria. A minimum magnitude of compaction pressure of around -3.5 MPa is reached in the run d118r2, which is only about 20 per cent the magnitude of the tensile strength of the rock ($\sigma_T = 20$ MPa). A decrease of tensile rock strength to values of 10 and 5 MPa (runs d118r4 and d118r8) still does not lead to any plastic failure. Shear failure of the host rock is not observed either, as the magnitude of deviatoric stresses here remains below 1 MPa. For runs with rock viscosities of $\eta_0 \leq 1e20$ Pa s and lower tensile yield strength (d119r8, d120r4 and d120r8, see summary Fig. 13 for the outcome of these runs), the increased magnitude of fluid overpressure reaches the limit set by the tensile yield strength, and thus volumetric plastic failure sets in. As a consequence, the style of melt transport transitions into a next regime, as discussed in Section 3.3.

3.2.3 Efficiency of melt extraction by diapirism

In these runs, a competition between melt transport by compaction waves and by convective diapirism is observed. Although early melt

accumulation occurs in a mode characteristic of compaction wave (i.e. dominated by percolation), once the diapir has grown to a certain radius, the upwards speed of the diapir exceeds the speed of melt percolation. Consequently, the diapir detaches from the boundary and rises through the model domain, with the surrounding host rock undergoing convective flow to accommodate the rise of the melt diapir. The stable radius of a diapir is reached when the characteristic speed of convective rise of the diapir, u_c (eq. 69), exceeds the characteristic speed of percolation, w_c (eq. 67), feeding the diapir from below. Thus, setting the two characteristic velocity scales as equal and solving for the diapir radius yields an expression for the stable radius of melt diapirism similar to the compaction length:

$$R = \sqrt{K_D \eta^*}. \quad (70)$$

We do not observe any runs, where two-dimensional compaction waves prevail over melt diapirism. First, this is due to the fact that melt fraction drops to zero away from the initial melt volume and thus the compaction length drops to zero, and compaction of the host rock ceases. Second, the melt-dependent shear viscosity law causes convective flow to be more efficient than compaction flow. Scott (1988) shows that the ratio r_η between shear viscosity η^* and compaction viscosity ξ^* determines the competition between

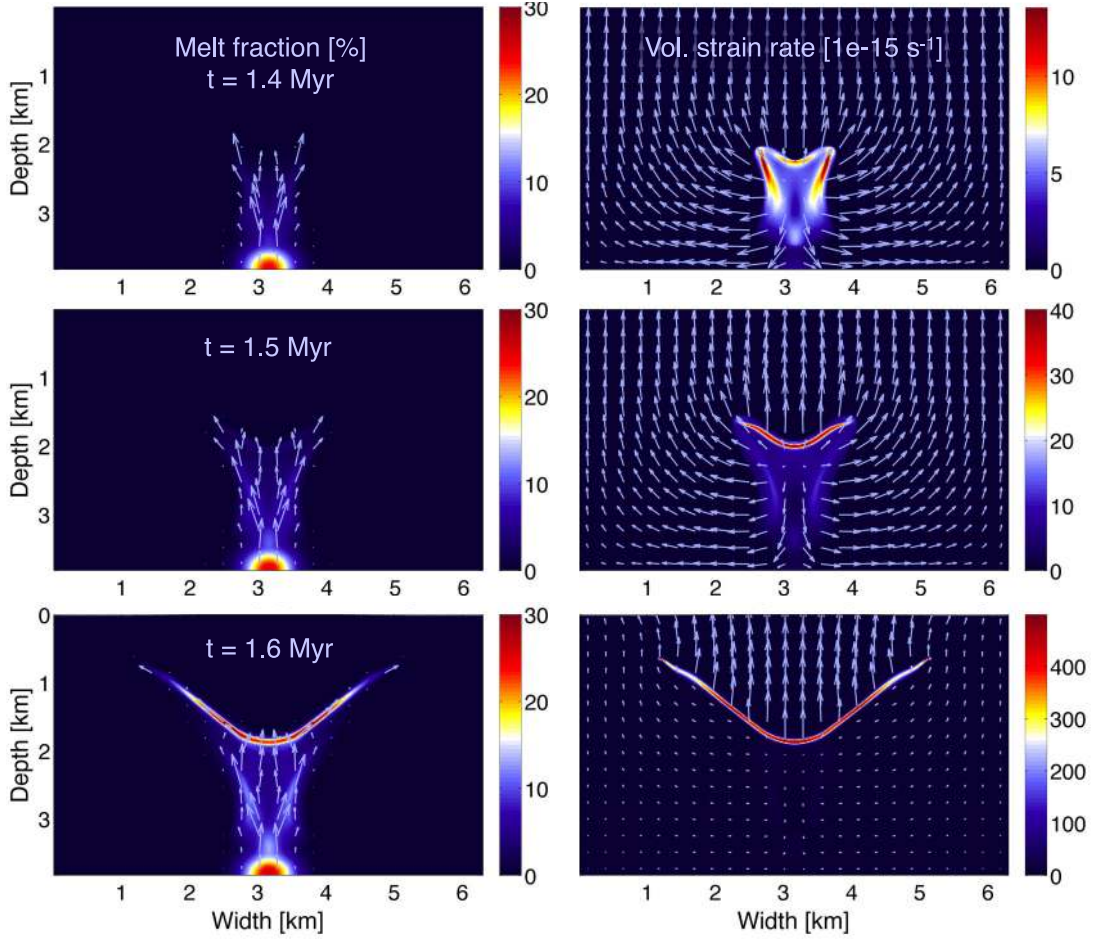


Figure 6. Snapshots of run d121r4 ($\eta_0 = 1e + 21$ Pa s, $\sigma_T = 10$ MPa) resulting in decompaction channeling. Visualization as in Fig. 4.

melt transport by compaction waves and convective diapirism. This ratio is also implied by comparing the characteristic velocity for compaction v_c (eq. 68) and convection u_c (eq. 69), if the diapir radius is of the order of the compaction length. Here, in the dominantly viscous region of the parameter space, this ratio takes the form of

$$r_\eta = \eta^* / \xi^* = \phi \exp(-\alpha_\phi \phi), \quad (71)$$

resulting in values two orders of magnitude below unity, indicating that convective transport will prevail at all levels of melt fraction. The highest values of r_η (0.001–0.013) are attained at melt fraction of 1–20 per cent, which may explain why the initial stages of diapir formation show some characteristic of compaction wave formation, even though no long-term propagation of a compaction wave is observed later on during the model runs.

3.3 Decompaction channeling

The second regime of tectonically coupled melt extraction emerges at intermediate host rock viscosities of $1e + 20$ Pa s $\leq \eta_0 \leq 1e + 22$ Pa s, values found around the lithosphere–asthenosphere boundary (LAB) or in ductile regions of the lower crust. The regime is observed to set in as the compaction pressure P_c reaches the magnitude of the tensile strength of rock σ_T at low shear stresses. The characteristic melt transport features in this regime are elongated bands of localized volumetric deformation, referred to as decompaction channels. Along these features, melt is accumulated and transported

away from the source region. Fig. 6 displays three snapshots of a run representative of this regime (d121r4, $\eta_0 = 1e + 21$ Pa s, $\sigma_T = C/4 = 10$ MPa, see Movies S3 and S4 for animated time evolution).

3.3.1 Geometry of decompaction channels

In Fig. 6, we observe melt transport features that are clearly different from the melt diapirs discussed above. An area of partial melt now extends upwards from the initial melt pulse. This area is initially bounded on both sides by two subvertical zones of localized volumetric deformation and melt accumulation, angled at around 20° towards the vertical (Fig. 6, top panels). Further along the simulation, a more strongly localized band of melt accumulation forms along the upper boundary of the area of partial melt (Fig. 6, middle panels). Both ends of this decompaction channel propagate diagonally upwards into the melt-free host rock. At first, this upper decompaction band is not stationary, but travels upwards as the zone of partial melt extends further into the host rock. Later in the simulation run, however, this bowl-shaped melt channel begins to stagnate in a final position, while both ends of the channel continue to extend outwards into the host rock (Fig. 6, bottom panels). At this stage, melt fraction is increased to more than 30 per cent along the channel, compared to values of around 15 per cent in the zone of partial melt below. Note that decompaction channels are only observed inside, or at the boundary of partially molten areas.

These channelized features of melt transport form as a consequence of volumetric failure of the host rock matrix, caused by high fluid overpressure at low shear stress ($\tau_{II} < |P_c|$), as illustrated by a representative stress circle in the Fig. 1 (green circle). In this simulation (d121r4), the magnitude of the extensional background stress field caused by the kinematic boundary conditions has a magnitude of 2 MPa, while fluid overpressures reach values of up to 10 MPa. The two subvertical melt bands are observed to form at steep angles of about 70° to the direction of least compressive stress above the initial melt pulse. The bowl-shaped orientation of the final channel dominating the later stages of the simulation is formed as the direction of least compressive stress is deflected outwards as the flow field interacts with the stress-free condition imposed on the surface. More systematic tests need to be performed to determine the exact nature of how decompaction channels align with principal stress directions. However, some first tests reveal that for compressive kinematic boundary conditions, the angles of initially emerging decompaction channels that are not yet interacting with the free surface are subhorizontal rather than subvertical.

The localization of compaction deformation emerging here as a consequence of volumetric failure may be quantified in terms of the volumetric plastic yield ratio r_{pla} , which is the ratio of the minimum compaction pressure P_y given by the failure criterion (eq. 38) over the predicted viscous compaction pressure,

$$r_{\text{pla}} = \frac{P_y}{-\xi^* \dot{\epsilon}_s}, \quad (72)$$

neglecting elastic effects, as the elastic contribution to total deformation is small. The effective compaction viscosity may then be expressed depending on the plastic yield ratio as $\xi_{\text{eff}}^* = \xi^* r_{\text{pla}}$. Replacing ξ^* by ξ_{eff}^* in the definition of the compaction length (eq. 65) yields the characteristic length scale associated with decompaction failure $\delta_{c,\text{pla}}$ to be (Connolly & Podladchikov 2007)

$$\delta_{c,\text{pla}} = \delta_c \sqrt{r_{\text{pla}}}. \quad (73)$$

The ratio r_{pla} typically takes values of $1e-1$ to $1e-3$, depending on the background viscosity and local strain rates, and thus the characteristic width of visco-elasto-plastic decompaction channels may be orders of magnitude below the viscous compaction length.

In their study of decompaction channeling, Connolly & Podladchikov (2007) used a constant value of r_{pla} to weaken compaction viscosity where fluid overpressure occurred. This simplification, along with a two-phase model neglecting shear stresses in the host rock, lead to the emergence of vertical decompaction tubes of constant width (due to constant r_{pla}), rather than angled channels, as they are observed here. This comparison additionally confirms our interpretation that decompaction channels align with the stress field.

Other runs, where decompaction channels emerge are d119r8, where a low tensile rock strength enable volumetric failure at a viscosity as low as $1e19$ Pa s, d120r4, where a set of diagonal decompaction channels form on top an already well-developed melt diapir, d120r8, d121r8 and d122r2, where two subvertical decompaction bands propagate upwards until a new regime emerges from top of the decompaction channels, leading to a melt-bearing tensile fracture propagating ahead of the channels (see summary Fig. 13 for a summary).

3.3.2 Deformation in decompaction channeling regime

From the solid and fluid velocity fields (arrows in Fig. 6), it is evident that melt is generally extracted upwards, percolating through the area of partial melt above the initial melt pulse. The pattern

of percolation, however, is not distributed homogeneously, but the melt flow is rather concentrated inside the decompaction channels, as seen by the melt velocity vectors aligning with the direction of channel propagation. Inside the areas of melt percolation, the deformation of the host rock consists of a downwards compaction flow. Outside the partially molten zone, host rock deformation accommodates the additional melt volume by outwards and upwards flow directions. As the upper, bowl-shaped decompaction channel becomes stationary towards the end of the simulation, we observe a strong increase of vertical uplift of the host rock above the forming melt band, caused by a strongly extensional volumetric deformation perpendicular to the melt band. The reasons for uplift in an otherwise extensional environment are that the rates of uplift are much larger than the extension velocity. Additionally, some shear deformation is observed along each decompaction channel as solid velocity arrows change both magnitude and direction across each channel.

To further characterize the style of deformation in the decompaction channeling regime, snapshots of the viscous, elastic and plastic components of both shear and compaction deformation in the run d121r4 are displayed in Fig. 7. Viscous deformation (top panels) is observed to be active in the area of the initial melt pulse, where viscous diapirism is ongoing, although relatively slowly. Other viscous deformation, especially in the shear domain, is concentrated within the decompaction bands, facilitated by the melt-weakening of shear viscosity cause by melt accumulating in the channels.

As expected, amplitudes of the elastic component of deformation (middle panels) are small compared to viscous and plastic ones. Most elastic deformation is concentrated in the propagating tips of decompaction channels, where shear stress and fluid overpressure accumulate until the effective stress state reaches the failure criterion. In both volumetric and deviatoric strain rates, a strongly localized band of plastic deformation is observed along the top of the partial melt zone, indicative of the strong opening component facilitated by volumetric plastic failure.

The combined pattern of deformation is one of bands or channels of localized viscoplastic decompaction deformation, the propagation of which into undeformed host rock is aided by elasticity. Decompaction channels are observed to be mobile, both propagating outwards from their tips and moving through the host rock along their whole length in the style of strongly asymmetrical compaction waves. Once enough melt is accumulated and stress directions remain favourable, a travelling decompaction band may become stationary. We note again that decompaction channels generally emerge inside or along the boundary of a partially molten area and may be fed by low melt fraction percolation along much of their length, thus potentially channeling low degree partial melt away from a larger melt region.

3.3.3 Efficiency of melt extraction by decompaction channels

The enhanced efficiency of melt transport relative to diapirism is manifest from the fact that the decompaction channels form and propagate ahead of the incipient diapir forming at the bottom of the box. Viscous percolation occurs on the characteristic timescale set by the characteristic percolation velocity (eq. 67) and the compaction length (eq. 65):

$$t_c = \frac{\delta_c}{w_c}. \quad (74)$$

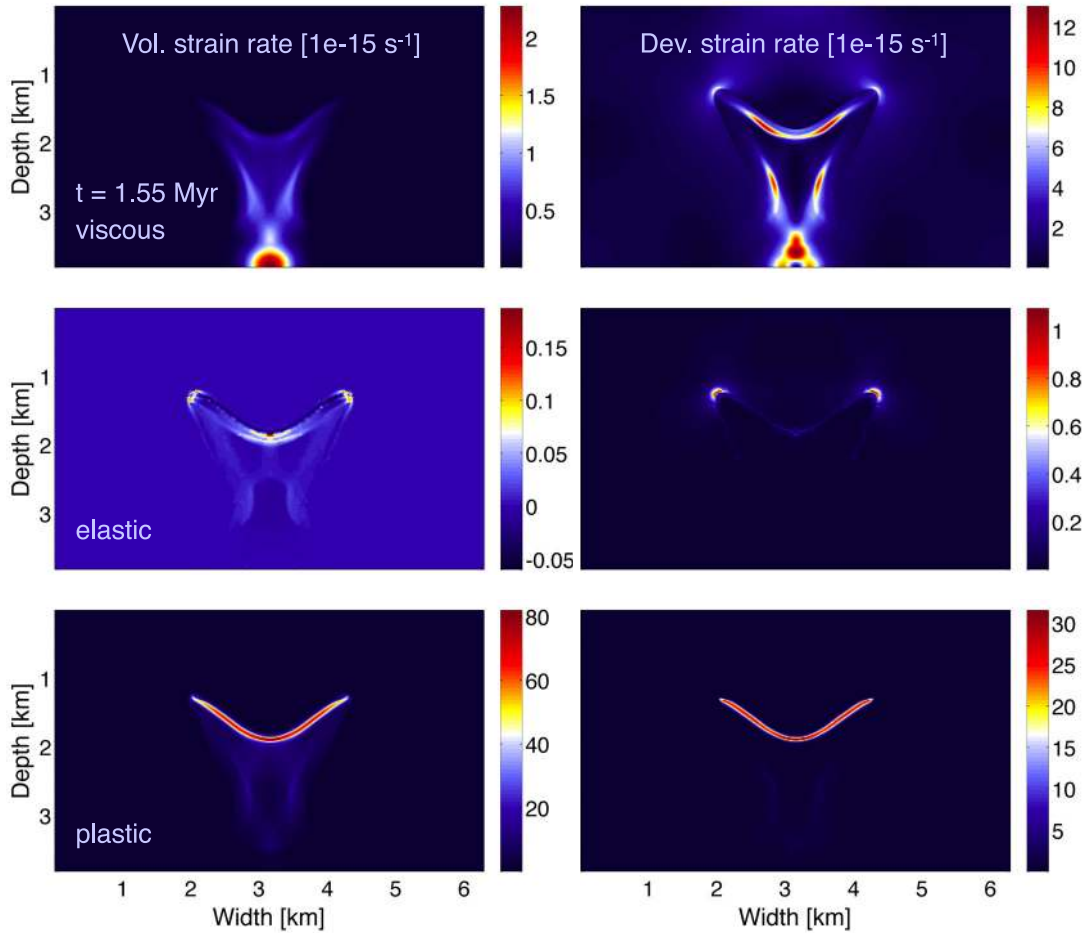


Figure 7. Viscous, elastic and plastic components of deformation from a snapshot of run d121r4 ($\eta_0 = 1e + 21$ Pa s, $\sigma_T = 10$ MPa), displaying compaction strain rate components (left) and shear strain rate components (right).

In the case of decompaction failure, the timescale of percolation is reduced proportional to the plastic reduction in compaction length given in eq. (73) (Connolly & Podladchikov 2007)

$$t_{c,pla} = t_c \sqrt{r_{pla}}, \quad (75)$$

and thus the efficiency of melt extraction again depends on the square root of the plastic yield ratio r_{pla} .

As it is evident that melt transport in the decompaction channeling regime is facilitated by localized melt percolation in a compacting host rock, the competition between compaction and convection needs to be re-examined in the case of volumetric failure. As introduced above (eq. 71), the ratio r_η between shear and compaction viscosity gives a measure of the relative efficiency of melt propagation by advective transport along a convecting host rock against pervasive transport in a compacting host rock. Once decompaction failure sets in the plastically modified viscosity ratio $r_{\eta,pla}$ becomes

$$r_{\eta,pla} = r_\eta / r_{pla}. \quad (76)$$

The meaning of this relation is that an increase in volumetric failure (smaller r_{pla}), the viscosity ratio $r_{\eta,pla}$ assumes higher values, indicating that the relative efficiency of compaction over convection is increased. This simple analysis explains how the increased efficiency of pervasive melt transport comes about under conditions where purely viscous compaction without weakening by decompaction plasticity would be highly inefficient.

3.4 Tensile failure

The third regime of tectonically coupled melt extraction emerges at high host rock viscosities of $\eta_0 \geq 1e + 22$ Pa s, values that are found in relatively cool regions of the lithosphere and crust, especially the upper crust. This regime is observed to set in as the shear stress $\bar{\tau}_{II}$ reaches the magnitude of the tensile yield strength σ_T at compaction pressures P_c close to zero. The dominant features of melt extraction in this regime are subvertical, sharply localized zones of deformation, in which melt is transported away from the source region. These features are a continuum representation of tensile fractures; in this case melt-bearing dykes. Fig. 8 displays two snapshots of the run d122r4, the results of which are representative of this regime ($\eta_0 = 1e + 22$ Pa s, $\sigma_T = C/4 = 10$ MPa, see Movies S5 and S6 for animated time evolution).

3.4.1 Geometry of tensile fractures

The features of melt transport found in this regime are a set of narrow, subvertical deformation zones propagating upwards from the initial melt pulse. If these features are indeed to be identified as tensile fractures or dykes, they should align perpendicular to the direction of least compressive stress. As the pure shear extensional stress field caused by the kinematic side boundaries is slightly perturbed by the presence of the melt source in the middle of the lower boundary, the orientation of the two fractures in Fig. 8, with

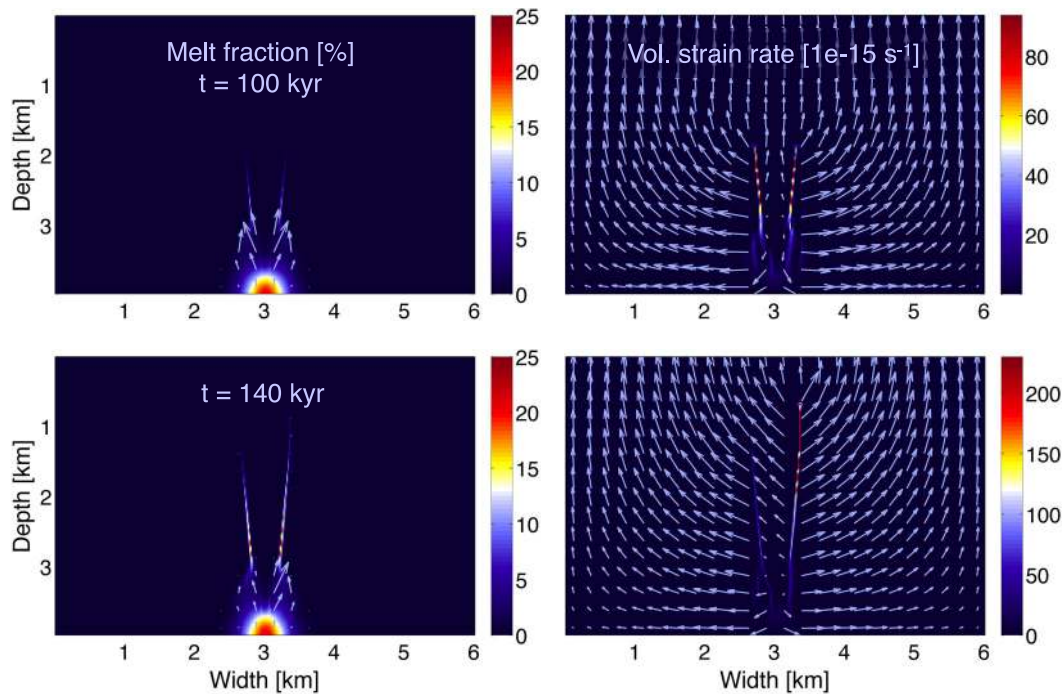


Figure 8. Snapshots of run d122r4 ($\eta_0 = 1e + 22$ Pa s, $\sigma_T = 10$ MPa) resulting in tensile fracturing. Visualization as in Fig. 4.

a slight deviation from the vertical, pointing radially away from the initial melt pulse, supports the hypothesis that these are, in fact, tensile fractures. Additionally, preliminary tests revealed that if compressive kinematic boundary conditions are applied instead, tensile fractures extend horizontally away from an initial melt pulse, which is again the direction perpendicular to the least compressive stress.

Furthermore, the width of these fractures typically extends over no more than one or two elements of the finite-element mesh, a characteristic shared by brittle plastic shear fractures in computational continuum mechanics. If numerical resolution is increased, the orientation of brittle fractures remains the same, but the number of parallel fractures is increased, while each individual fracture again localizes down to the grid level (Buitter *et al.* 2006). Preliminary high-resolution tests reveal that the same is true for the tensile fractures observed here. This property of brittle plasticity in computational continuum mechanics implies that brittle fractures are not resolved numerically on their characteristic length scale, which from field geology and rock physics is known to be as small as a few millimetres.

Tensile fractures in Fig. 8 are observed to propagate into melt-free rock, while continually being infiltrated by melt from below. While dykes in nature are discrete fractures completely filled by melt, the simulated tensile fractures here only accumulate a rather low melt fraction of 10–25 per cent. We interpret this result in the light of numerical resolution. As the characteristic width of a tensile fracture is below the spatial resolution of the numerical method, the volumetric strain rate obtained on the available numerical resolution is underestimated by the ratio of the grid spacing towards the real width of a fracture. The following example should serve to illustrate this point: If a real fracture is narrower than the numerical element size by a factor of 100, the volumetric strain rate measured on the numerical grid would be smaller than the actual value by a factor of 100 as well, and thus, the evolution of melt content in the simulated dyke is underestimated by the same factor. While there are means

by which this underestimation of melt transport by dykes simulated on a regular continuum grid could be corrected (e.g. increasing permeability as a function of plastic yielding, parametrized increase of volumetric strain rates inside dykes, etc.), we have made no such attempts here, but simply observed the features that self-consistently arise from the visco-elasto-plastic rheology.

3.4.2 Deformation in tensile failure regime

Closer examination of the melt velocity (arrows in Fig. 8) reveals that melt is extracted upwards inside the dykes. All melt extraction is limited to the narrow zones of volumetric deformation visible in Fig. 8, meaning that no pervasive melt transport occurs outside these fractures. The flow field of the host rock conversely displays a 180° change of horizontal velocity across the width of a single finite element, while the vertical velocity component remains mostly unchanged across the fracture. This flow signature constitutes the closest possible continuum approximation of the discontinuous flow field caused by a brittle fracture.

Fig. 9 displays the viscous, elastic and plastic components of shear and compaction deformation during the propagation of tensile fractures in the simulation d122r2. Viscous compaction deformation (top left panel) is active mainly in the area of the initial melt pulse, albeit only at a small relative amplitude. Viscous shear deformation (top right panel), however, is mainly active at the trailing ends of both dykes, where the most melt has been accumulated. Once again, viscous shear deformation is concentrated by the melt-weakening of the solid shear viscosity, as observed in the decompression channeling regime above.

Elastic volumetric deformation (middle left panel) has a small relative amplitude and is somewhat erratically distributed along the fracture zones, probably an artefact of the numerical grid resolution. A more consistent pattern, however, is observed in the partially molten area of the initial melt pulse, where a set of subvertical bands of elastic compaction strain rate is found. Observing these

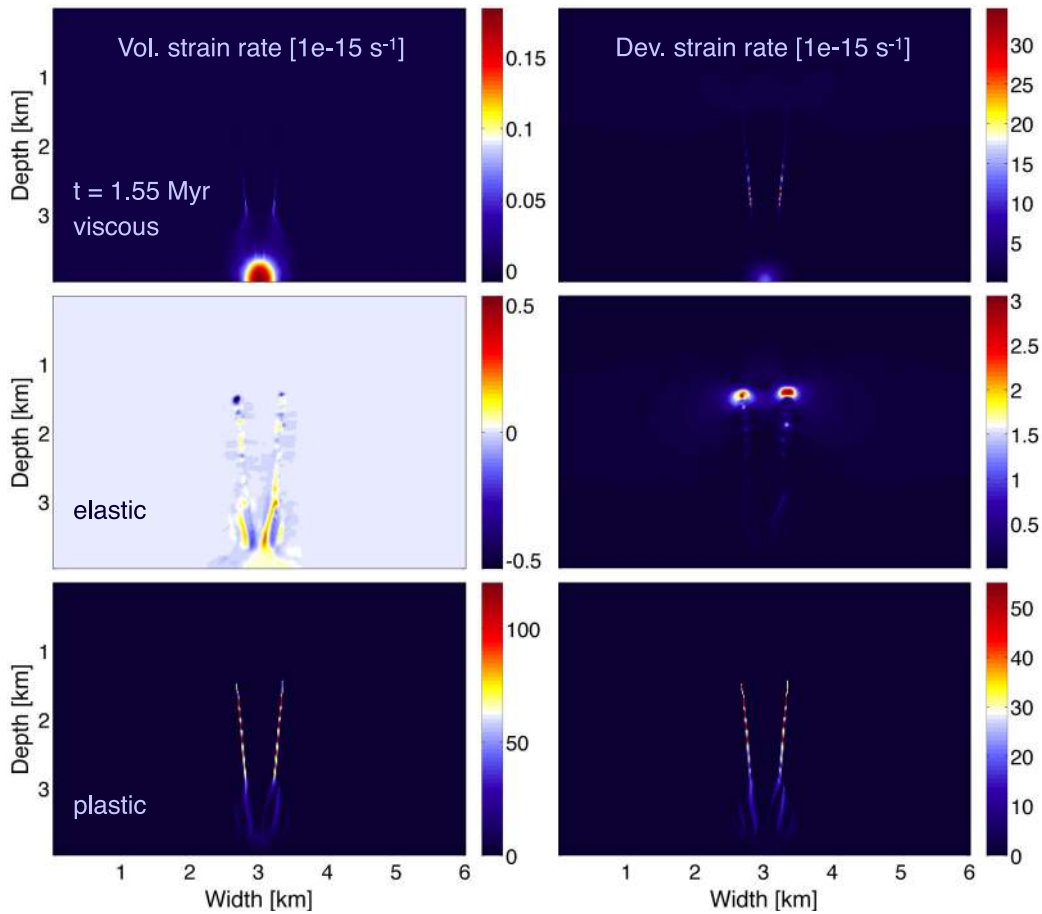


Figure 9. Viscous, elastic and plastic components of deformation from a snapshot of run d122r4 ($\eta_0 = 1e+22$ Pa s, $\sigma_T = 10$ MPa), displaying compaction strain rate components (left) and shear strain rate components (right).

bands through their time evolution (Movies S5 and S6) reveals that they are mobile, travelling through the partially molten area at a characteristic wave length and speed. Thus, these features are identified with the mobile decompaction bands observed in the previous regime. Elastic shear deformation (middle right panel) is concentrated to the propagating tip of each fracture, where elastic stress concentration facilitates the propagation of the crack tip.

Not surprisingly, plastic deformation accounts for the major part of deformation associated with tensile fractures (bottom panels). Note that the plastic shear strain rates are found at approximately half the amplitude of the plastic volumetric component, a ratio that is predicted by plasticity theory for an opening fracture (effective dilatancy angle of 90°) (Vermeer & De Borst 1984). The decompaction bands found in the partially molten area at the root of the dykes are clearly visible in the plastic strain rate fields, as expected from observation above.

Other runs at solid viscosities of $1.e+22$ Pa s or higher that are similarly dominated by tensile fractures as the preferred mode of melt transport include d122r2/8 (see summary Fig. 13 for the final outcome of these runs) and d123r2/4/8 (see Fig. 10 for results of d123r4, representative of the three runs at $1e+23$ Pa s). The only additional features observed at viscosities greater than $1.e+22$ are normal faults that emerge where shear stresses outside the partially molten regions of the model box reach stresses high enough to cause shear failure. Fig. 10 gives two snapshots of the run d123r4,

showing melt fraction with melt velocity arrows (left column of panels) and solid shear strain rates with solid velocity arrows (right column of panels). Initially, three vertical dykes are observed to emerge from the melt source (best visible in the strain rate field, upper right panel). In the first snapshot, the two outer dykes are connected at their tip to a set of conjugated normal faults, whereas the middle is not connected to any shear fractures. Later, in the second snapshot, the two outer dykes have ceased to propagate, whereas the middle dyke is still active and is now, in turn, connected to a pair of normal faults. These observations are relevant as they demonstrate the vital connection between magmatism and tectonics. Magmatic structures always constitute weak zones in the fabric of the lithosphere and crust, thus providing stress inhomogeneities on which brittle fractures preferentially nucleate. On the other hand, this close connection between magmatism and brittle tectonics serves to explain why many areas of magmatic activity are found in close proximity of major fault systems.

3.4.3 Efficiency of melt extraction by tensile fractures

Crack propagation in nature may be recorded on the timescale of earthquakes, and thus constitutes the fastest known process of melt extraction. However, the transported volumes of melt in dykes are relatively small, as dykes found in the geological record are typically reported at width to length ratios of $1e-2$ to $1e-4$ (Rubin 1995).

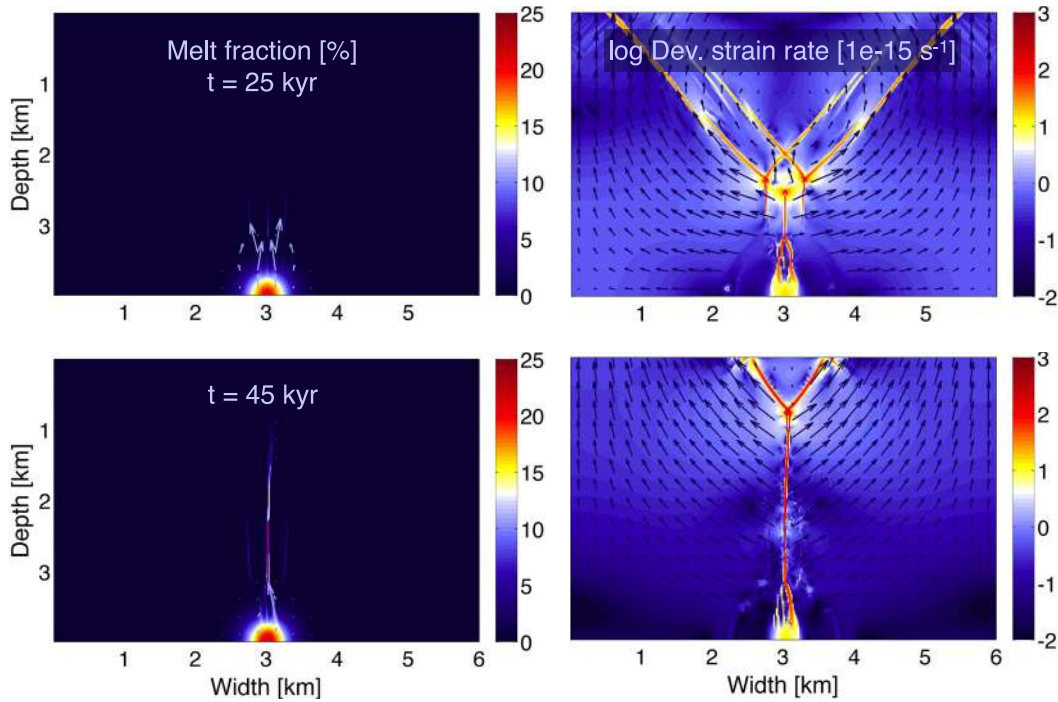


Figure 10. Snapshots of run d123r4 ($\eta_0 = 1e + 23$ Pa s, $\sigma_T = 10$ MPa) resulting in tensile and shear fracturing. Visualization as in Fig. 4.

Still, melt flow in a tensile fracture is very efficient, as the pure melt phase has a very low viscosity and may thus move at high speed through a tensile fracture once it is opened.

As we assess the efficiency of melt extraction by dyking, the question arises how properties like the width, speed of propagation and the accumulated melt content of the dykes simulated here are to be interpreted. As discussed above, the dyke width is limited to the element size of the numerical grid. Simulated dykes propagate at speeds of the order of centimetres per year, depending on model properties such as the solid shear viscosity and shear modulus, the permeability and fluid viscosity and the amplitude of imposed extension. Melt fractions accumulated inside the dykes reach values of around 25 per cent in this suite of simulations. Comparing these findings to the theoretical and observational characteristics of melt-bearing tensile fractures, we note that the simulated dykes are much wider, slower in propagation and lower in melt content. All of these limitations are bound up with the spatial resolution available to numerical models on a geodynamically relevant scale. However, despite these limitations, dyking is the most efficient feature of melt transport in all simulation runs at high viscosities. Where tensile fractures emerge, they consistently propagate with significantly higher speed than all other modes of melt transport and thus create conduits for melt extraction even under the mechanically most competent conditions in the host rock.

3.5 Numerical models of melt ascent from asthenosphere to crust

Having sampled host rock conditions from ductile to brittle crustal levels in a simple setup, we now proceed to two additional simulations on a lithospheric scale in order to see how the observed modes of melt transport discussed above self-consistently emerge and interact as melt is extracted from the LAB to the upper crust. To this end, we construct a model setup representing a continental

lithosphere and crust with background viscosity profiles established from two different thermal profiles. Note that temperature advection or diffusion is not treated here, but a realistic geotherm obtained from a standard half-space cooling model is used to constrain the initial viscosity profile of the model runs.

The model box has dimensions of 160×240 km and consists of three layers: 20 km of upper crust with a density of $\rho_{UC} = 2500$ kg m⁻³, followed by 20 km of lower crust with a density of $\rho_{LC} = 2700$ kg m⁻³ and a mantle lithosphere with a density of $\rho_{ML} = 3000$ kg m⁻³. With a melt density of $\rho_f = 2600$ kg m⁻³, a reversal of density contrast at the lower to upper crust boundary is achieved. The depth profiles of intrinsic rock viscosity are calculated from the geotherm using an Arrhenius law, with slightly varying pre-exponential factor and activation energy to reflect lower strength of crustal rocks with respect to the mantle. The resulting strength profiles of a thermally younger and a thermally older lithosphere are given in Figs 11 and 12, respectively. As initial melt fraction distribution, a circular region with a diameter of 12 km and a melt content of 80 per cent is imposed at the lower boundary. The reference permeability is set to $k_0 = 1e - 7$ m² and the extensional kinematic boundary condition is given by a background strain rate of $\dot{\epsilon}_{BG} = -5e - 15$ s⁻¹. The constant fluid pressure boundary condition is switched off and thus pressure boundaries are zero flux on all sides. All other model parameters are as listed in Table 2.

To provide a seed for plastic shear and tensile failure to nucleate, random noise is added to both friction angle (maximum amplitude of 1°) and cohesion (maximum amplitude of 0.2 MPa). Plastic shear strain rates are integrated over time and a linear weakening of the friction angle proportional to accumulated plastic strain is added to the method in order to provide a damage memory, by which plastic faults constitute tectonic weak zones, even after they are no longer active.

Snapshots of melt fraction and shear strain rate of runs with a thermally younger, and thus mechanically weaker, and a thermally

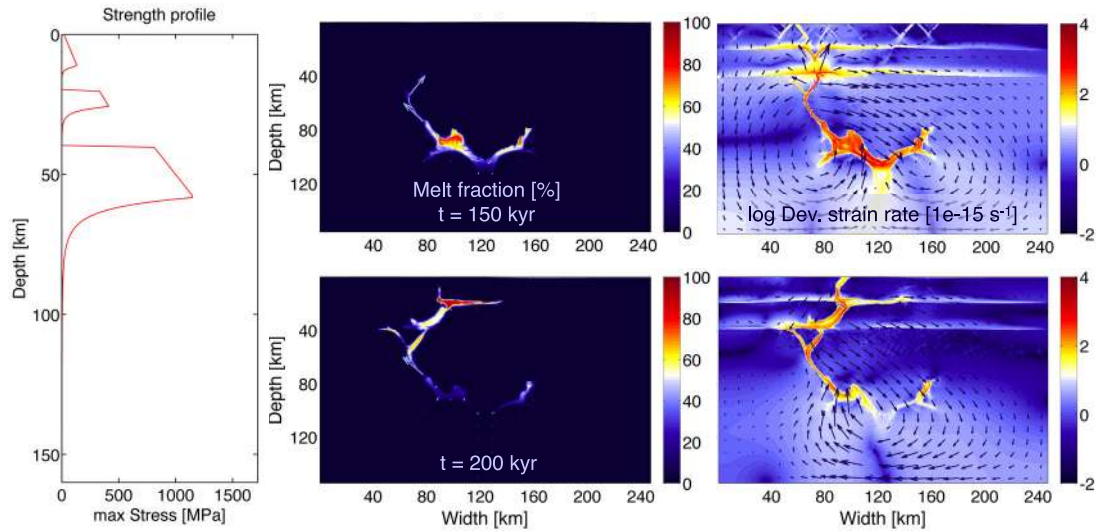


Figure 11. Snapshots of whole-lithosphere run with a thermally younger lithosphere, displaying the strength profile of lithosphere and crust with depth (left), melt fractions (per cent) with melt velocity arrows (middle) and shear strain rates relative to background strain rate (right). The applied density structure is such that the depth of neutral buoyancy is at 20 km.

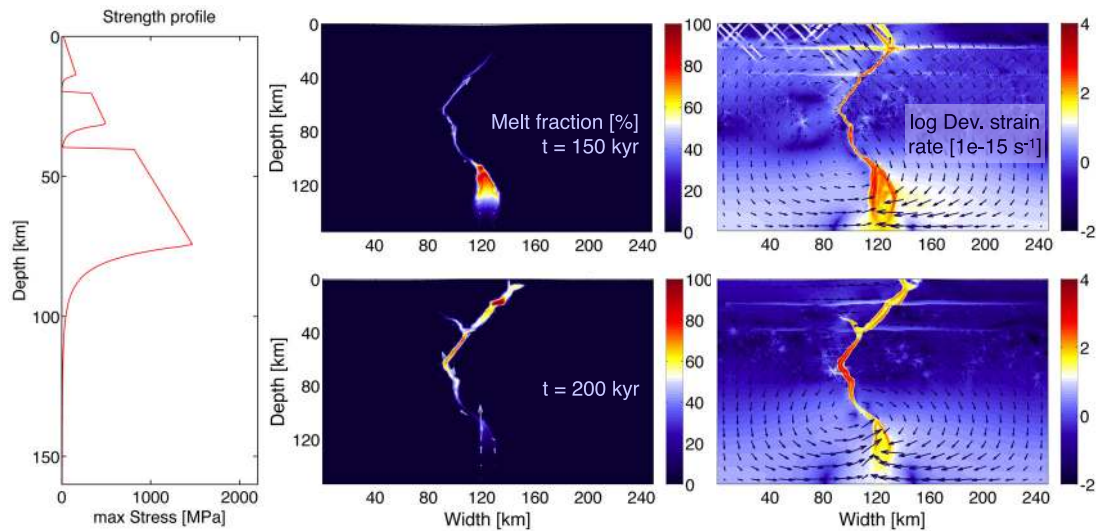


Figure 12. Snapshots of whole-lithosphere run with a thermally older lithosphere. Visualization as in Fig. 11.

older, and mechanically stronger, lithosphere are given in Figs 11 and 12 (see Movies S7–S10 for animated time evolutions of these results). In both runs, most of the melt transport features described above emerge and interact self-consistently over the range of host rock conditions found from the LAB to the surface. We do not observe any melt diapirism here, as the initial melt volumes are placed directly on the LAB, where viscosities are already higher than values for which diapirism would be the preferred mode of melt transport.

In both runs, the circular melt volume first penetrates the LAB by means of a process not present in the runs presented above. Here, a layer of overlying host rock along the top boundary of the melt pulse is weakened by the infiltration of overpressured melt, leading to volumetric failure and causing the weakened layer to delaminate into the magma body below. This interaction of an advancing percolation front with volumetric failure by melt overpressure and subsequent delamination of the resulting layer of crystal mush was

not observed earlier. This particular mode of deformation requires melt overpressures of amplitudes reaching the tensile strength of rock (here $\sigma_T = 10$ MPa) to occur in a low-viscosity host rock. Yet, because a magma body of sufficiently large dimension is present, delamination of the weakened rock into the magma body occurs on a timescale similar to the advancement of the percolation front. The resulting magma body takes the shape of a funnel (most likely tube-like in 3-D), narrowing towards the top as a function of the increasing background viscosity (best visible in Fig. 12, top panels).

As host rock viscosities rise to $1e+21$ Pas, the first major regime transition takes place, as channels of localized volumetric deformation propagate diagonally away from the main magma body. Although spatial resolution renders the distinction more difficult, we identify these features as decompaction channels, recognized by their orientation and deformation pattern. In the simulation with a weaker lithosphere, two decompaction channels form above the initial magma body. As they provide a weak zone in the host

rock, a return flow pattern sets in, causing the lower lithosphere to flow down and in towards the initial magma body, while large volumes of melt are moved up and outwards into the widening channels, up to the point where two new magma bodies are established further up and on either side of the initial melt volume (Fig. 11, top panels). However, the further model evolution shows that only one of these magma bodies is able to propagate further up towards the lower crust. Interestingly, a large portion of the melt trapped in the inactive secondary channel is eventually forced back by regional tectonic deformation and diverted into the primary channel.

The second major regime transition occurs, where the imposed viscosity profile causes a steep increase in mechanical strength up to the point of brittle strength levels (see strength profiles in Figs 11 and 12). As expected, melt-bearing tensile fractures emerge. Unlike in other runs presented above (d122r4, d123r4), the orientation of these fractures is often not vertical, as expected from the regional extension applied at the side boundaries. Instead, tensile fractures preferentially form along the weak zones created by pre-existing shear faults, a behaviour known from geological observation (Rubin 1995). As overpressured melt propagates along pre-existing shear faults, the walls of the fault are forced apart and the previous pattern of shear deformation along the fault is changed into an extensional flow perpendicular to the fault plane. This process, best visible in Fig. 12 (top panels), illustrates that even though dykes in these models are limited to the available spatial resolution, the geometry of tensile fractures is self-consistently determined by material properties and local stress conditions and, in the long-term evolution, such fractures evolve into wider conduits for melt extraction through the mechanically most competent parts of the lithosphere.

As melt eventually reaches the upper crust in both simulations, the buoyancy contrast between solid and melt phase is inverted and thus further melt ascent is mostly prohibited. As a consequence, melt accumulates in the weak lower layer of the upper crust, forming a magma body of up to 100 per cent melt content. In the simulation with the weaker strength profile, the shape of this final magma chamber is that of a laccolith fed from a feeder zone at one end (Fig. 11, bottom panels), a geometry that incidentally bears some resemblance to the Torres del Paine plutonic rock complex, although it occurs at a larger depth in this simulation (e.g. Leuthold *et al.* 2012).

The geometry of the upper crustal magma body in the second simulation consists of two roughly elliptical magma bodies linked by a diagonal conduit of similar dimensions (Fig. 12, lower panels). In both runs, some melt transport above the depth of neutral buoyancy is observed. In fact, even though the geometry of the upper crustal magma bodies is quite different, both runs feature a single, vertical dyke propagating from the uppermost magma body upwards to the surface, where eruption would occur. These results demonstrate that the depth of neutral buoyancy is not the primary control of the emplacement level of a magma body. Instead, it is the integrated buoyancy contrast of the interconnected column of melt through the lithosphere and crust, combined with the forcing imposed by regional tectonic stress that governs the depth of emplacement or the ultimate eruption of a magmatic system.

These numerical results of melt ascent from the asthenosphere to the upper crust again need to be understood with the caveat that no thermal and chemical coupling of magmatism is included. Nevertheless, they demonstrate the potential of the proposed method to investigate open questions of melt extraction and magma emplacement tectonics under a wide variety of conditions found in a continental lithosphere and crust.

4 DISCUSSION

4.1 Melt transport in a visco-elasto-plastic host rock

These numerical results demonstrate that the implementation of a visco-elasto-plastic rheology in two-phase flow simulations may result in at least three distinct modes of melt transport under geodynamically relevant host rock conditions. Whereas melt diapirism generally occurs at low viscosities, decompaction channels are formed at intermediate viscosities, and tensile fractures emerge at high viscosities. The results also indicate that the proposed methodology provides an opportunity to address some of the open issues related to melt extraction through lithosphere and crust, as it is able to simulate melt extraction through host rock undergoing both fracture and flow, a feature previously inaccessible to computational models of magma dynamics.

The aim of this study was to introduce a visco-elasto-plastic two-phase rheology and demonstrate its potential for the computational study of melt transport through the lithosphere and crust. The three observed regimes of tectonically coupled melt transport that self-consistently emerged from a suite of simple numerical problems have been characterized in a preliminary way, as a full discussion of particularly the plastic features of melt transport goes beyond the scope of this study and would require a much more comprehensive coverage of the relevant parameter space. Still, these observations allow some preliminary discussion of implications these findings might have for the understanding of how magmatic systems evolve coupled to the rock conditions and tectonic deformation of a continental lithosphere. In the following, we will summarize our analysis of the three regimes of melt transport described above, discuss some of the implications with respect to magmatic systems and review the limitations of the numerical method.

4.1.1 Melt transport by diapirism

Summarizing the findings on melt diapirism, we note that melt extraction under mechanically weak asthenospheric conditions (lower than $1.e+20$ Pa s) is dominated by viscous diapirism constituted by local upwellings driven by the buoyancy of accumulated melt. As no crystallization is considered here, quantitative aspects of these results need to be interpreted with caution. Only at very low host rock viscosities will enough heat be advected for a diapir to keep a stable melt content over the timescale of vertical transport. Nevertheless, the combination of melt accumulation by incipient compaction wave formation and subsequent melt ascent by diapirism is in our simulations the dominant process by which melt from a wider region of partial melting in the asthenosphere is collected and brought upwards towards the LAB. Melt thus arrives at the LAB in discrete pulses of high melt fraction rather than in a continuous percolative flux, providing an important constraint for possible initial conditions for further models investigating melt extraction through the lithosphere.

Most likely, the composition of melt collected in a diapir closely reflects the degree of melting of the asthenospheric source of partial melt the diapir is derived from. During the rise of a melt diapir, little equilibration with surrounding rock material is expected to occur and convection dominates over compaction. As any inherited or newly crystallized grain assemblage is entrained along the convective flow, rather than being removed from the remaining melt, not much magmatic differentiation may be expected to occur during melt transport by diapirism.

4.1.2 Melt transport by decompaction channeling

From our findings on decompaction channeling, we conclude that melt extraction in rock of intermediate mechanical strength, as it is found in the relatively high-temperature parts of the lithosphere and lower crust ($1.e+20$ to $1.e+22$ Pa s), is dominantly facilitated by decompaction channeling. From the observations presented above, we arrive at a definition of decompaction channels as elongated bands of channelized melt percolation in a compacting host rock, characterized by volumetric failure due to fluid overpressure, and aligned at steep angles with respect to the least compressive stress direction.

So far, we have not addressed the question of the physical process that underlies volumetric, or decompaction failure. As this process occurs at low levels of shear stress and in areas of melt percolation, we propose that the volumetric deformation is accommodated mainly by the opening of grain boundaries, forced apart by the presence of an overpressured melt. This interpretation is in line with previous theoretical considerations (Connolly & Podladchikov 1998, 2007; Ricard & Bercovici 2003). Conceptually, the nature of these channels is thought of not as a brittle, macroscopic mode of plastic failure of the host rock, but rather as a distributed microscopic failure mode manifesting along grain boundaries in an area of partially molten rock under fluid overpressure. Connolly & Podladchikov (2007) argue that this type of failure should lead to an overall viscoplastic compaction flow characterized by a weakening of the effective compaction viscosity under fluid overpressure, which coincides with our observations.

Furthermore, our numerical results demonstrate that decompaction bands may travel through the host rock as a kind of elongated visco-elasto-plastic compaction wave, driven by shear stresses in the matrix rather than buoyancy alone. Although the physical origin of this process cannot be inferred from our models, these decompaction bands may be conceptually related to the so-called vug waves, a concept proposed by Phipps Morgan & Holtzmann (2005) in order to explain melt band formation in a deforming host rock. The connection between decompaction channeling and melt band formation certainly deserves more attention. The volumetric failure described here effectively behaves as a power-law viscous rheology with a very high stress exponent, a condition that, according to the analysis of Katz *et al.* (2006), favours increased growth rates of melt bands in a partially molten rock under shear deformation. Also, the melt bands observed in analogue simple shear experiments form at angles of 55° – 70° to the least compressive stress direction (Holtzman 2003). We find that decompaction channels grow at similar angles in our simulations.

For now, with the conclusive physical interpretation of these features left to be determined, localized melt percolation by volumetric failure of the host rock and the subsequent formation of decompaction channels is identified as the likely process, by which melt may penetrate the thermal boundary layer of the LAB and ascend into the mechanically stronger lithosphere.

The flow patterns observed in the decompaction channeling regime lead to an efficient phase separation and, by implication, to strong magmatic differentiation. Pervasive flow of melt through the compacting solid phase likely leads to a higher degree of chemical equilibration with the host rock than convective transport of whole magma bodies by melt diapirism. Additionally, if crystallization occurs, forming crystals would follow the downwards compaction flow of the solid, whereas the remaining melt would continually be removed, causing fractional crystallization of magmas as they penetrate into the lower lithosphere.

4.1.3 Melt transport by tensile fracturing

The numerical results demonstrate that melt extraction in rock of high mechanical strength, as it found in the relatively low-temperature parts of the lithosphere and crust (viscosity greater than $1.e+22$ Pa s), is dominantly facilitated by the opening of melt-bearing tensile fractures. Such dykes in our simulations are clearly distinct from decompaction channels by the following four phenomenological criteria. (i) Width: Tensile fractures have a natural length scale far below the numerically available spatial resolution, and therefore dykes in our simulations always localize down to the width of a single element. Conversely, even narrow decompaction channels are more diffuse features than tensile cracks, assuming a width of at least several elements, even at low grid resolutions. (ii) Position: Simulated tensile fractures are always found to be stationary features, whereas decompaction bands mostly move through the rock matrix until they finally assume a fixed location, where they develop into stationary melt channels. Even stationary decompaction channels are easily distinguished from tensile fractures, as the former generally occur inside or along the boundary of a partially molten area, whereas tensile fractures protrude from a melt source into melt-free rock. (iii) Orientation: We find that tensile fractures form perpendicular to the least compressive stress direction, whereas decompaction channels form at angles of around 70° . Even when dykes propagate along pre-existing shear faults, the pattern of deformation indicates that the local stress field has been locally rotated in such a way that the dyke again is positioned normal to the least compressive stress. (iv) Stress conditions: Tensile fracturing is only observed as shear stresses in the mixture approach the magnitude of tensile rock strength, whereas volumetric failure leading to decompaction channeling occurs at lower levels of shear stress, but higher levels of fluid overpressure.

In the first suite of simulations above, dykes are observed to propagate relatively quickly into the host rock, but to transport rather low volumes of melt through it. If, on the other hand, a dyke is established, local extension is ongoing and enough melt is available from below, a tensile fracture may be used by larger volumes of melt and even crystal-bearing magma as a conduit, thus allowing large volumes of magma to ascend rapidly through the widening conduit. The two simulations of melt extraction from the LAB to the upper crust demonstrate that our method is able to capture this process of how initially narrow dykes widen to provide conduits for large-scale melt extraction.

4.1.4 Summary of melt transport regimes

Fig. 13 summarizes the findings on the three regimes of tectonically coupled melt transport as a function of tensile strength and intrinsic viscosity of the host rock. For clarity of visualization, a background colour is assigned to each of the three regimes of melt transport. The red background colour in Fig. 13 indicates the stability field of melt diapirism (A), the stability field of decompaction channeling (B) is coloured in yellow, and the one of tensile fracturing (C) in blue. The characteristic flow pattern of each mode of melt transport is depicted in the three panels at the bottom of Fig. 13.

4.1.5 Regime transitions

We find from the numerical results that the first regime boundary from diapirism to decompaction channeling occurs where fluid overpressure reaches the magnitude of tensile rock strength at shear stresses close to zero. Setting the characteristic pressure scale

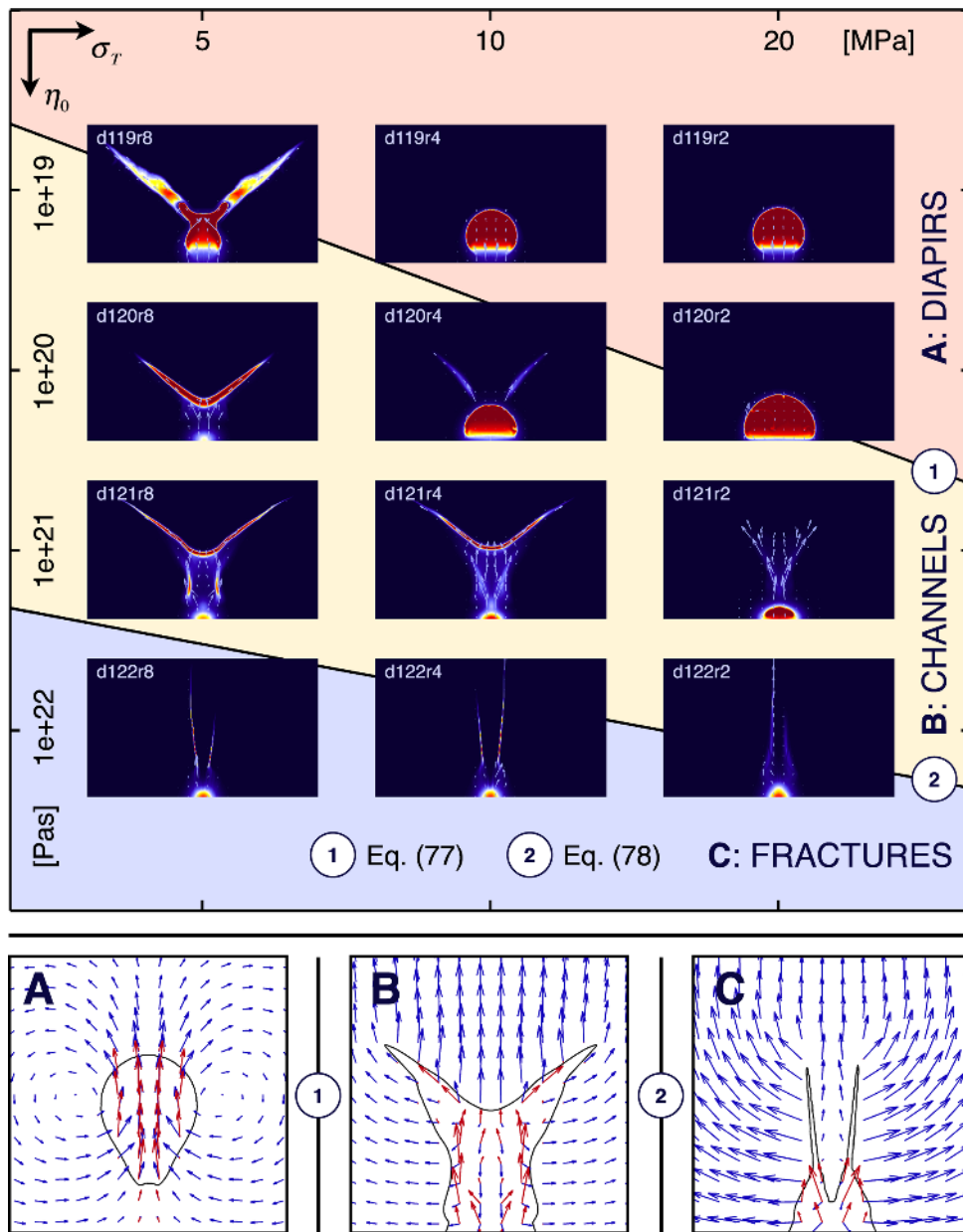


Figure 13. Summary of model results as a function of host rock viscosity and tensile rock strength. Regions of the parameter space are coloured in red where (A) melt diapirism is dominant, yellow where (B) decompaction channeling is observed and blue where (C) tensile fracturing emerges. Transitions from one regime to the next occur (1) where fluid overpressure reaches the magnitude of tensile rock strength at low shear stresses (eq. 77), and (2) where shear stresses reach the magnitude of tensile rock strength at high fluid pressure (eq. 78). Three panels at the bottom, marked (A), (B) and (C), give characteristic flow patterns taken from representative runs of each regime, with dark blue arrows for the host rock, and red arrows for melt. A contour line encircles regions of melt content over 5 per cent.

$p_c = \Delta\rho g\delta_c$ equal to the tensile rock strength and solving for the intrinsic rock viscosity η_0 (contained in the definition of compaction length) reveals that the level of viscosity at which volumetric failure may emerge, with given values for other material properties, is proportional to the square of the tensile rock strength σ_T :

$$\eta_0 = \frac{\phi}{(1-\phi)} \left(\frac{\sigma_T}{\Delta\rho g} \right)^2 K_D^{-1}. \quad (77)$$

The second regime boundary between decompaction channels and tensile fractures occurs as shear stresses increase to the magnitude of tensile rock strength at fluid pressures equal to the total pressure. Setting the shear stress induced by the kinematic boundary condition ($2\eta^*\dot{\epsilon}_{BG}$) equal to σ_T and again solving for the intrinsic

solid viscosity η_0 reveals that the level of viscosity necessary to excite tensile failure at given values for other material properties is a linear function of the tensile rock strength σ_T :

$$\eta_0 = \frac{\exp(\alpha_\phi\phi)}{(1-\phi)} \frac{\sigma_T}{2\dot{\epsilon}_{BG}}. \quad (78)$$

The two lines forming the boundaries between regimes in Fig. 13 are found by evaluating eqs (77) and (78) at a melt fraction of 1 per cent (other property values as in Table 2) and plotting the resulting lines in log-log space. The emergence of the three melt transport regimes in the numerical results stands in excellent agreement with the predicted regime boundaries.

4.3 Model limitations

The main limitations of the proposed methodology are related to numerical resolution. The physics of tectonically coupled melt extraction involves processes spanning length scales of millimetres to hundreds of kilometres and timescales spanning seconds to millions of years. Any modelling approach needs to be limited to a certain bandwidth of processes, as no simulation framework is able to consider all relevant scales simultaneously. The method adopted here consists of an approach to simulate magma dynamics as a continuum process at spatial scales of hundreds of metres to hundreds of kilometres, over timescales of hundreds to millions of years. Typical rates of deformation therefore should remain below values of a few metres per year. This focus leads to a choice of model resolutions that render it necessary to filter out smaller scale processes, as, for example, small-scale mixing in a magma chamber.

There are two aspects of our method where such limitations apply: First, a lower cut-off viscosity is needed to ensure that length scales of simulated flow features remain larger than the grid spacing, and thus, numerical instabilities are avoided.

Second, fractures cannot localize below the scale of one grid spacing. Both limitations are common to numerical models of geodynamic processes and cannot be circumvented in any straightforward manner. In general, however, the characteristics of the features of melt transport described in this study should not be greatly influenced by our choice of model resolution, as long as it is kept in mind that both speed and complexity of flow features in regions of very high melt fraction upwards of 60 per cent are likely underestimated, and that tensile and shear fractures may only localize to grid level and thus efficiency of melt transport by dyking is likely underestimated. However, even with these limitations, the spontaneous and self-consistent emergence of melt-bearing tensile fracturing presents an interesting opportunity for the computational study of magma dynamics in a host rock that may undergo both fracture and flow.

In this study, we focussed on two-dimensional simulations with moderate numerical resolutions. We employed direct solvers in combination with a fixed point iteration method to deal with nonlinearities. If one wishes to solve the same set of equations in three dimensions at sufficiently large resolutions, efficient (multigrid) iterative methods need to be tested, which is likely to be challenging. Furthermore, adaptive mesh refinement methods could prove helpful to better resolve the tensile dykes in our models.

Another model limitation relates to the fact that the proposed methodology concentrates on the mechanical side of magma dynamics in lithosphere and crust, thereby neglecting the thermal and chemical evolution that a magmatic system would undergo. The fundamental regimes of melt transport described here are expected to equally apply to fully coupled thermochemical simulations of magma dynamics, even though a greater wealth of possible features is expected to be uncovered, as many feedbacks between composition, temperature, pressure and stress will modify the style of melt transport.

5 CONCLUSIONS

In this study, we described how to include visco-elasto-plastic constitutive laws for both shear and compaction deformation in an otherwise standard numerical model of two-phase magma dynamics and discussed how to solve it numerically in two dimensions. The proposed formulation is found able to capture the most important modes of melt extraction that are to be expected in a tectonically

active lithosphere and crust. Three regimes of melt transport emerge in sequence and sometimes overlap, as the intrinsic viscosity of the host rock increases with decreasing temperatures from the asthenosphere towards the surface. These regimes are (i) melt diapirism, (ii) decompaction channeling and (iii) tensile failure. However, to physically quantify the detailed properties of each of these deformation regimes, a more thorough investigation of the model parameter space and choice of boundary conditions is required. In particular, the regime of decompaction channeling deserves further investigation, as the underlying physics are not yet fully understood.

The main strength of the proposed approach is the ability to simulate magma dynamics in a host rock undergoing both fracture and flow within a self-consistent physical framework. This study mainly aimed at introducing the visco-elasto-plastic two-phase rheology and demonstrating its potential. The proposed formulation of the two-phase flow equations and their numerical implementation has been developed using operators and techniques utilized in standard finite-element Stokes flow discretizations. The evident similarity between these two types of computational problems suggests that much of the experience gained from Stokes flow modelling of lithosphere deformation may straightforwardly be applied to tackle tectonically coupled magma dynamics problems in the future. Moreover, coupling such models with energy conservation and magmatic evolution will allow to address various open problems in the field of igneous rock formation.

ACKNOWLEDGEMENTS

The authors would like to thank Richard Katz and Guillaume Richard for their helpful and thorough reviews, which greatly helped to improve this manuscript. This work was supported by the Swiss National Science Foundation's research grant PDFMP2_123087 as part of the ProDoc 4D-Adamello doctoral school. Author DAM was supported by the ETH Zurich Postdoctoral Fellowship Program, additionally partial support was provided to authors DAM and BJKP by the European Research Council under the European Community's Seventh Framework Program (FP7/2007–2013) through ERC Starting Grant agreement 258830.

REFERENCES

- Abe, Y., 1995. Basic equations for evolution of partially molten mantle and core, in *The Earth's Central Part: Its Structure and Dynamics*, pp. 215–235, ed. Yukutabe, T., Terrapub.
- Aharonov, E., Spiegelman, M. & Kelemen, P., 1997. Three-dimensional flow and reaction in porous media: implications for the Earth's mantle and sedimentary basins, *J. geophys. Res.*, **102**, 14 821–14 833.
- Aharonov, E., Whitehead, J.A., Kelemen, P.B. & Spiegelman, M., 1995. Channeling instability of upwelling melt in the mantle, *J. geophys. Res.*, **100**, 20 433–20 450.
- Barcilon, V. & Lovera, O.M., 1989. Solitary waves in magma dynamics, *J. Fluid Mech.*, **204**, 121–133.
- Barcilon, V. & Richter, F.M., 1986. Nonlinear waves in compacting media, *J. Fluid Mech.*, **164**, 429–448.
- Barnes, P.M., 1994. Inherited structural control from repeated cretaceous to recent extension in the North Mernoo Fault Zone, western Chatham Rise, New Zealand, *Tectonophysics*, **237**, 27–46.
- Batchelor, G.K., 1967. *An Introduction to Fluid Dynamics*, 61 pp., Cambridge University Press.
- Bathe, K.-J., 1995. *Finite Element Procedure*, 1037 pp., Prentice Hall.
- Bercovici, D. & Ricard, Y., 2003. Energetics of a two-phase model of lithospheric damage, shear localization and plate-boundary formation, *Geophys. J. Int.*, **152**, 581–596.

- Bercovici, D., Ricard, Y. & Schubert, G., 2001a. A two-phase model for compaction and damage 1. General theory, *J. geophys. Res.*, **106**, 8887–8906.
- Bercovici, D., Ricard, Y. & Schubert, G., 2001b. A two-phase model for compaction and damage 3. Applications to shear localization and plate boundary formation, *J. geophys. Res.*, **106**, 8925–8939.
- Buiter, S.J.H., Babeyko, A.Y., Ellis, S., Gerya, T.V., Kaus, B.J.P., Kellner, A., Schreurs, G. & Yamada, Y., 2006. The numerical sandbox: comparison of model results for a shortening and an extension experiment, *Geol. Soc. London Spec. Publ.*, **253**, 29–64.
- Cagnioncle, A.-M., Parmentier, E.M. & Elkins-Tanton, L.T., 2007. Effect of solid flow above a subducting slab on water distribution and melting at convergent plate boundaries, *J. geophys. Res.*, **112**, doi:10.1029/2007JB004934.
- Cai, M., 2010. Practical estimates of tensile strength and Hoek-Brown strength parameter $m(i)$ of brittle rocks, *Rock Mech. Rock Eng.*, **43**, 167–184.
- Caricchi, L., Burlini, L., Ulmer, P., Gerya, T.V., Vassalli, M. & Papale, P., 2007. Non-Newtonian rheology of crystal-bearing magmas and implications for magma ascent dynamics, *Earth planet. Sci. Lett.*, **264**, 402–419.
- Connolly, J.A.D. & Podladchikov, Y.Y., 1998. Compaction-driven fluid flow in viscoelastic rock, *Geodin. Acta*, **11**, 55–84.
- Connolly, J.A.D. & Podladchikov, Y.Y., 2007. Decompaction weakening and channeling instability in ductile porous media: implications for asthenospheric melt segregation, *J. geophys. Res.*, **112**, B10205, doi:10.1029/2005JB004213.
- Corti, G., van Wijk, J., Cloetingh, S. & Morley, C.K., 2007. Tectonic inheritance and continental rift architecture: numerical and analogue models of the East African Rift system, *Tectonics*, **26**, TC6006, doi:10.1029/2006TC002086.
- Costa, A., Caricchi, L. & Bagdassarov, N., 2009. A model for the rheology of particle-bearing suspensions and partially molten rocks, *Geochem. Geophys. Geosyst.*, **10**, Q03010, doi:10.1029/2008GC002138.
- Coussy, O., 2010. *Mechanics and Physics of Porous Solids*, 297 pp., Wiley.
- Dabrowski, M., Krotkiewski, M. & Schmid, D.W., 2008. MILAMIN: MATLAB-based finite element method solver for large problems, *Geochem. Geophys. Geosyst.*, **9**, Q04030, doi:10.1029/2007GC001719.
- Deubelbeiss, Y. & Kaus, B.J.P., 2008. Comparison of Eulerian and Lagrangian numerical techniques for the Stokes equations in the presence of strongly varying viscosity, *Phys. Earth planet. Inter.*, **171**, 92–111.
- Dufek, J. & Bachmann, O., 2010. Quantum magmatism: magmatic compositional gaps generated by melt-crystal dynamics, *Geology*, **38**, 687–690.
- Ehlers, W. & Volk, W., 1998. On shear band localization phenomena of liquid-saturated granular elastoplastic porous solid materials accounting for fluid viscosity and micropolar solid rotations, *Mech. Cohesive-Frictional Mater.*, **2**, 301–320.
- Elman, H.C., Silvester, D.J. & Wathen, A.J., 2005. *Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics*, 400 pp., Oxford University Press.
- Evans, K.F., 2005. Permeability creation and damage due to massive fluid injections into granite at 3.5 km at Soultz: 2. Critical stress and fracture strength, *J. geophys. Res.*, **110**, B04204, doi:10.1029/2004JB003169.
- Faul, U.H., 1997. Permeability of partially molten upper mantle rocks from experiments and percolation theory, *J. geophys. Res.*, **102**, 10 299–10 311.
- Fournier, R.O., 1996. Compressive and tensile failure at high fluid pressure where preexisting fractures have cohesive strength, with application to the San Andreas Fault, *J. geophys. Res.*, **101**, 25 499–25 509.
- Fowler, A.C., 1985. A mathematical model of magma transport in the asthenosphere, *Geophys. Astrophys. Fluid. Dyn.*, **33**, 63–96.
- Gerya, T.V. & Yuen, D.A., 2003. Characteristics-based marker-in-cell method with conservative finite-differences schemes for modeling geological flows with strongly variable transport properties, *Phys. Earth planet. Inter.*, **140**, 293–318.
- Gerya, T.V. & Yuen, D.A., 2007. Robust characteristics method for modelling multiphase visco-elasto-plastic thermo-mechanical problems, *Phys. Earth planet. Inter.*, **163**, 83–105.
- Gessner, K., 2009. Coupled models of brittle-plastic deformation and fluid flow: approaches, methods, and application to mesoproterozoic mineralisation at Mount Isa, Australia, *Surv. Geophys.*, **30**, 211–232.
- Golabek, G.J., Tackley, P.J. & Schmeling, H., 2008. Earth's core formation aided by flow channelling instabilities induced by iron diapirs, *Earth planet. Sci. Lett.*, **271**, 24–33.
- Hall, H., 1953. Compressibility of reservoir rocks, *J. Petrol. Tech.*, **5**, 17–19.
- Hier-Majumder, S., Ricard, Y. & Bercovici, D., 2006. Role of grain boundaries in magma migration and storage, *Earth planet. Sci. Lett.*, **248**, 735–749.
- Hollister, L.S. & Crawford, M.L., 1986. Melt-enhanced deformation: a major tectonic process, *Geology*, **14**, 558–561.
- Holtzman, B.K., 2003. Stress-driven melt segregation in partially molten rocks, *Geochem. Geophys. Geosyst.*, **4**, 8607, doi:10.1029/2001GC000258.
- Hutton, D.H.W., 2011. Granite emplacement mechanisms and tectonic controls: inferences from deformation studies, *Trans. R. Soc. Edinburgh Earth Sci.*, **79**, 245–255.
- Jalalh, A.A., 2006a. Compressibility of porous rocks: Part I. Measurements of Hungarian reservoir rock samples, *Acta Geophys.*, **54**, 319–332.
- Jalalh, A.A., 2006b. Compressibility of porous rocks: Part II. New relationships, *Acta Geophys.*, **54**, 399–412.
- Katz, R.F., 2008. Magma dynamics with the enthalpy method: benchmark solutions and magmatic focusing at mid-ocean ridges, *J. Petrol.*, **49**, 2099–2121.
- Katz, R.F., 2010. Porosity-driven convection and asymmetry beneath mid-ocean ridges, *Geochem. Geophys. Geosyst.*, **11**, Q0AC07, doi:10.1029/2010GC003282.
- Katz, R.F., Knepley, M.G., Smith, B., Spiegelman, M. & Coon, E.T., 2007. Numerical simulation of geodynamic processes with the Portable Extensible Toolkit for Scientific Computation, *Phys. Earth planet. Inter.*, **163**, 52–68.
- Katz, R.F., Spiegelman, M. & Holtzman, B.K., 2006. The dynamics of melt and shear localization in partially molten aggregates, *Nature*, **442**, 676–679.
- Kaus, B.J.P., 2010. Factors that control the angle of shear bands in geodynamic numerical models of brittle deformation, *Tectonophysics*, **484**, 36–47.
- Kaus, B.J.P. & Becker, T.W., 2006. Effects of elasticity on the Rayleigh–Taylor instability: implications for large-scale geodynamics, *Geophys. J. Int.*, **168**, 843–862.
- Kaus, B.J.P., Gerya, T.V. & Schmid, D.W., 2008. Recent advances in computational geodynamics: theory, numerics and applications, *Phys. Earth planet. Inter.*, **171**, 2–6.
- Kelemen, P.B., Hirth, G., Shimizu, N., Spiegelman, M. & Dick, H.J., 1997. A review of melt migration processes in the adiabatically upwelling mantle beneath oceanic spreading ridges, *Phil. Trans. R. Soc. A Math. Phys. Eng. Sci.*, **355**, 283–318.
- Khoei, A.R. & Mohammadnejad, T., 2011. Numerical modeling of multiphase fluid flow in deforming porous media: a comparison between two- and three-phase models for seismic analysis of earth and rockfill dams, *Comput. Geotech.*, **38**, 142–166.
- Kohlstedt, D.L. & Holtzman, B.K., 2009. Shearing melt out of the Earth: an experimentalist's perspective on the influence of deformation on melt extraction, *Ann. Rev. Earth planet. Sci.*, **37**, 561–593.
- Leuthold, J., Müntener, O., Baumgartner, L.P., Putlitz, B., Ovtcharova, M. & Schaltegger, U., 2012. Time resolved construction of a bimodal laccolith (Torres del Paine, Patagonia), *Earth planet. Sci. Lett.*, **325**, 85–92.
- Lister, J.R. & Kerr, R.C., 1991. Fluid-mechanical models of crack propagation and their application to magma transport in dykes, *J. geophys. Res.*, **96**, 10 049–10 077.
- Lomov, I. & Robin, M.B., 2003. Numerical simulation of damage using an elastic-viscoplastic model with directional tensile failure, *J. Phys.*, **110**, 281–286.
- Loret, B. & Prevost, J.H., 1991. Dynamic strain localization in fluid-saturated porous media, *J. Eng. Mech.*, **117**, 907–922.
- May, D.A. & Moresi, L., 2008. Preconditioned iterative methods for Stokes flow problems arising in computational geodynamics, *Phys. Earth planet. Inter.*, **171**, 33–47.

- McCaffrey, K., 1992. Igneous emplacement in a transpressive shear zone: Ox Mountains igneous complex, *J. geol. Soc.*, **149**, 221–235.
- McKenzie, D., 1984. The generation and compaction of partially molten rock, *J. Petrol.*, **25**, 713–765.
- McKenzie, D., 1985. The extraction of magma from the crust and mantle, *Earth planet. Sci. Lett.*, **74**, 81–91.
- Mei, S., Bai, W., Hiraga, T. & Kohlstedt, D.L., 2002. Influence of melt on the creep behavior of olivine–basalt aggregates under hydrous conditions, *Earth planet. Sci. Lett.*, **201**, 491–507.
- Menand, T., 2011. Physical controls and depth of emplacement of igneous bodies: a review, *Tectonophysics*, **500**, 11–19.
- Morency, C., Huismans, R.S., Beaumont, C. & Fullsack, P., 2007. A numerical model for coupled fluid flow and matrix deformation with applications to disequilibrium compaction and delta stability, *J. geophys. Res.*, **112**, B10407, doi:10.1029/2006JB004701.
- Moresi, L., Dufour, F. & Mühlhaus, H.B., 2003. A Lagrangian integration point finite element method for large deformation modeling of viscoelastic geomaterials, *J. Comput. Phys.*, **184**, 476–497.
- Murrell, S.A.F., 1964a. Theory of the propagation of elliptical Griffith cracks under various conditions of plane strain or plane stress: part 1, *Br. J. appl. Phys.*, **15**, 1195–1210.
- Murrell, S.A.F., 1964b. The theory of the propagation of elliptical Griffith cracks under various conditions of plane strain or plane stress: parts II and III, *Br. J. Appl. Phys.*, **15**, 1211.
- Paterson, M.S. & Wong, T., 2005. *Experimental Rock Deformation - The Brittle Field*, 351 pp., Springer.
- Paterson, S.R. & Fowler, T.K. Jr., 1993. Re-examining pluton emplacement processes, *J. Struct. Geol.*, **15**, 191–206.
- Pelletier, D., Fortin, A. & Camarero, R., 1989. Are FEM solutions of incompressible flows really incompressible? (or how simple flows can cause headaches!), *Int. J. Num. Methods Fluids*, **9**, 99–112.
- Petford, N. & Atherton, M.P., 1992. Granitoid emplacement and deformation along a major crustal lineament: the Cordillera Blanca, Peru, *Tectonophysics*, **205**, 171–185.
- Phipps Morgan, J. & Holtzman, B.K., 2005. Vug waves: a mechanism for coupled rock deformation and fluid migration, *Geochem. Geophys. Geosyst.*, **6**, Q08002, doi:10.1029/2004GC000818.
- Pitcher, W.S., 1979. The nature, ascent and emplacement of granitic magmas, *J. geol. Soc.*, **136**, 627–662.
- Popov, A.A. & Sobolev, S.V., 2008. SLIM3D: a tool for three-dimensional thermomechanical modeling of lithospheric deformation with elasto-visco-plastic rheology, *Phys. Earth planet. Inter.*, **171**, 55–75.
- Rabinowicz, M. & Vigneresse, J.L., 2004. Melt segregation under compaction and shear channeling: application to granitic magma segregation in a continental crust, *J. geophys. Res.*, **109**, B04407, doi:10.1029/2002JB002372.
- Ramberg, H., 1981. *Gravity, Deformation, and the Earth's Crust: In Theory, Experiments, and Geological Application*, 452 pp., Academic Press.
- Renner, J., Evans, B. & Hirth, G., 2000. On the rheologically critical melt fraction, *Earth planet. Sci. Lett.*, **181**, 585–594.
- Ricard, Y. & Bercovici, D., 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.
- Richard, G.C., Kanjilal, S. & Schmeling, H., 2012. Solitary-waves in geophysical two-phase viscous media: a semi-analytical solution, *Phys. Earth planet. Inter.*, **198–199**, 61–66.
- Richard, G., Monnereau, M. & Rabinowicz, M., 2007. Slab dehydration and fluid migration at the base of the upper mantle: implications for deep earthquake mechanisms, *Geophys. J. Int.*, **168**, 1291–1304.
- Richardson, C.N., 1998. Melt flow in a variable viscosity matrix, *Geophys. Res. Lett.*, **25**, 1099–1102.
- Rozhko, A.Y., Podladchikov, Y.Y. & Renard, F., 2007. Failure patterns caused by localized rise in pore-fluid overpressure and effective strength of rocks, *Geophys. Res. Lett.*, **34**, L22304, doi:10.1029/2007GL031696.
- Rubin, A.M., 1993a. Dikes vs. diapirs in viscoelastic rock, *Earth planet. Sci. Lett.*, **119**, 641–659.
- Rubin, A.M., 1993b. Tensile fracture of rock at high confining pressure: implications for dike propagation, *J. geophys. Res.*, **98**, 15 919–15 935.
- Rubin, A.M., 1995. Propagation of magma-filled cracks, *Ann. Rev. Earth planet. Sci.*, **23**, 287–336.
- Rubin, M.B., Vorobiev, O.Y. & Glenn, L.A., 2000. Mechanical and numerical modeling of a porous elastic–viscoplastic material with tensile failure, *Int. J. Solids Struct.*, **37**, 1841–1871.
- Schmalholz, S.M., Podladchikov, Y.Y. & Schmid, D.W., 2001. A spectral/finite difference method for simulating large deformations of heterogeneous, viscoelastic materials, *Geophys. J. Int.*, **145**, 199–208.
- Schmeling, H., 2000. Partial melting and melt segregation in a convecting mantle, *Phys. Chem. Molten Rocks*, **11**, 141–178.
- Schmeling, H., Kruse, J.P. & Richard, G., 2012. Effective shear and bulk viscosity of partially molten rock based on elastic moduli theory of a fluid filled poroelastic medium, *Geophys. J. Int.*, **190**, 1571–1578.
- Scott, D.R., 1988. The competition between percolation and circulation in a deformable porous medium, *J. geophys. Res.*, **93**, 6451–6462.
- Scott, D.R. & Stevenson, D.J., 1984. Magma solitons, *Geophys. Res. Lett.*, **11**, 1161–1164.
- Scott, D.R. & Stevenson, D.J., 1986. Magma ascent by porous flow, *J. geophys. Res.*, **91**, 9283–9296.
- Scott, D.R. & Stevenson, D.J., 1989. A self-consistent model of melting, magma migration and buoyancy-driven circulation beneath mid-ocean ridges, *J. geophys. Res.*, **94**, 2973–2988.
- Simpson, G. & Spiegelman, M., 2011. Solitary wave benchmarks in magma dynamics, *J. Sci. Comput.*, **49**, 268–290.
- Simpson, G., Spiegelman, M. & Weinstein, M.I., 2010a. A multiscale model of partial melts: 1. Effective equations, *J. geophys. Res.*, **115**, B04410, doi:10.1029/2009JB006375.
- Simpson, G., Spiegelman, M. & Weinstein, M.I., 2010b. A multiscale model of partial melts: 2. Numerical results, *J. geophys. Res.*, **115**, B04411, doi:10.1029/2009JB006376.
- Skempton, A.W., 1960. Effective stress in soils, concrete and rocks, in *Proceedings of the Conference on Pore Pressure and Suction in Soils*, Butterworths, London, pp. 4–16.
- Sleep, N.H., 1974. Segregation of magma from a mostly crystalline mush, *Geol. Soc. Am. Bull.*, **85**, 1225–1232.
- Spiegelman, M., 1993a. Flow in deformable porous media. Part 1. Simple analysis, *J. Fluid. Mech.*, **247**, 17–38.
- Spiegelman, M., 1993b. Flow in deformable porous media. Part 2. Numerical analysis—the relationship between shock waves and solitary waves, *J. Fluid. Mech.*, **247**, 39–63.
- Spiegelman, M., Kelemen, P.B. & Aharonov, E., 2001. Causes and consequences of flow organization during melt transport: the reaction infiltration instability in compactible media, *J. geophys. Res.*, **106**, 2061–2077.
- Stevenson, D.J., 1989. Spontaneous small-scale melt segregation in partial melts undergoing deformation, *Geophys. Res. Lett.*, **16**, 1067–1070.
- Šrámek, O., Ricard, Y. & Bercovici, D., 2007. Simultaneous melting and compaction in deformable two-phase media, *Geophys. J. Int.*, **168**, 964–982.
- Šrámek, O., Ricard, Y. & Dubuffet, F., 2010. A multiphase model of core formation, *Geophys. J. Int.*, **181**, 198–220.
- Tagliasacchi, A., 2008. *Matlab kdtree*, Available at: <http://www.mathworks.com/matlabcentral/fileexchange/21512> (accessed 16 October, 2012).
- Taisne, B. & Jaupart, C., 2009. Dike propagation through layered rocks, *J. geophys. Res.*, **114**, B09203, doi:10.1029/2008JB006228.
- Takei, Y. & Holtzman, B.K., 2009. Viscous constitutive relations of solid-liquid composites in terms of grain boundary contiguity: 1. Grain boundary diffusion control model, *J. geophys. Res.*, **114**, B06205, doi:10.1029/2008JB005850.
- Terzaghi, K., 1923. *Die Berechnung der Durchlässigkeitsziffer des Tonnes aus dem Verlauf der Hydrodynamischen Spannungserscheinungen*, 14 pp., Sitzungsber Akad Wis Wien, Math Nat Klasse, Abt Ila.
- Terzaghi, K., 1943. *Theoretical Soil Mechanics*, 12 pp., Wiley.
- Van Keken, P.E., King, S.D., Schmeling, H., Christensen, U.R., Neumeister, D. & Doin, M.P., 1997. A comparison of methods for the modeling of thermochemical convection, *J. geophys. Res.*, **102**, 22 477–22 495.
- Vermeer, P.A. & De Borst, R., 1984. Non-associated plasticity for soils, concrete and rock, *Heron*, **29**, 1–63.

Weatherley, S.M. & Katz, R.F., 2012. Melting and channelized magmatic flow in chemically heterogeneous, upwelling mantle, *Geochem. Geophys. Geosyst.*, **13**, Q0AC18, doi:10.1029/2011GC003989.

Yang, X.-S., 2002. A mathematical model for Voigt poro-visco-plastic deformation, *Geophys. Res. Lett.*, **29**, 1066, doi:10.1029/2001GL014014.

Zienkiewicz, O.C., Taylor, R.L. & Nithiarasu, P., 2005. *The Finite Element Method for Fluid Dynamics*, 6th edn, 441 pp., Elsevier Butterworth-Heinemann.

APPENDIX A: NUMERICAL METHOD

A.1 Implementation of governing equations

To find a suitable implementation of the conservation eqs (57)–(59), we follow a similar strategy as is used for implementing variable viscosity Stokes flow in finite-element modelling (Zienkiewicz *et al.* 2005). First, eqs (57)–(59) are rewritten in a simplified matrix–vector form in terms of the solution variables to give

$$\begin{bmatrix} \mathbf{B}^T \mathbf{D} \mathbf{B} & -\mathbf{G} & -\mathbf{G} \\ -\mathbf{G}^T & -\mathbf{G}^T \mathbf{K}_D \mathbf{G} & 0 \\ -\mathbf{G}^T & 0 & -C \end{bmatrix} \begin{bmatrix} \mathbf{v}_s \\ P_f \\ P_c \end{bmatrix} = \begin{bmatrix} \bar{\rho} \mathbf{g} \hat{\mathbf{z}} - \mathbf{B}^T \chi_\tau^* \tilde{\boldsymbol{\tau}}_s^o \\ \mathbf{G}^T \mathbf{K}_D \rho_f \mathbf{g} \hat{\mathbf{z}} \\ -C \chi_p^* \Delta P^o \end{bmatrix}, \quad (\text{A1})$$

with the volumetric rheological constant defined as $C = 1/\xi_{\text{eff}}^*$, and using the following relation for the deviatoric stress tensor

$$\bar{\boldsymbol{\tau}} = \mathbf{D} \mathbf{B} \mathbf{v}_s + \chi_\tau^* \tilde{\boldsymbol{\tau}}_s^o, \quad (\text{A2})$$

where the operator \mathbf{B} is used to obtain strain rates from the velocity field

$$\mathbf{B} = \begin{bmatrix} \partial_x & 0 \\ 0 & \partial_y \\ \partial_y & \partial_x \end{bmatrix}. \quad (\text{A3})$$

For shorter notation, we use ∂_x to denote a partial derivative with respect to coordinate x . \mathbf{G} is the gradient operator defined as

$$\mathbf{G} = \begin{bmatrix} \partial_x \\ \partial_y \end{bmatrix}. \quad (\text{A4})$$

The constitutive tensor for deviatoric stresses \mathbf{D} contains the effective shear viscosity of the solid phase:

$$\mathbf{D} = \eta_{\text{eff}}^* \begin{bmatrix} 4/3 & -2/3 & 0 \\ -2/3 & 4/3 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (\text{A5})$$

The Darcy tensor \mathbf{K}_D contains the Darcy coefficients $K_{D,ij}(x, y) = k_\phi/\eta_f$ in direction of coordinates, which in this study are kept isotropic, but could theoretically describe an anisotropic permeability structure:

$$\mathbf{K}_D = \begin{bmatrix} K_{D,xx} & K_{D,xy} \\ K_{D,yx} & K_{D,yy} \end{bmatrix}. \quad (\text{A6})$$

In a second step, the problem is discretized on a rectangular finite-element mesh. Solid velocity and fluid pressure are approximated on Q1 elements with their respective linear shape functions $\mathbf{N}_v = [N_1 N_1 N_2 N_2 N_3 N_3 N_4 N_4]$ and $\mathbf{N}_f = [N_1 N_2 N_3 N_4]$, whereas compaction pressure is defined as piecewise constant on P0 elements with shape functions $\mathbf{N}_p = 1$. After reformulating the equations in

the weak form, we arrive at the following expression (see below for the detailed finite-element formulation of each submatrix)

$$\begin{bmatrix} \mathbf{V} \mathbf{V} & \mathbf{V} \mathbf{F} & \mathbf{V} \mathbf{C} \\ \mathbf{V} \mathbf{F}^T & \mathbf{F} \mathbf{F} & \mathbf{0} \\ \mathbf{V} \mathbf{C}^T & \mathbf{0} & \mathbf{C} \mathbf{C} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{v}}_s \\ \tilde{\mathbf{p}}_f \\ \tilde{\mathbf{p}}_c \end{bmatrix} = \begin{bmatrix} \mathbf{r}_v \\ \mathbf{r}_f \\ \mathbf{r}_c \end{bmatrix}, \quad (\text{A7})$$

which are compactly expressed as

$$\mathbf{L} \mathbf{x} = \mathbf{r}. \quad (\text{A8})$$

Here, \mathbf{L} is the global stiffness matrix, \mathbf{x} is the solution vector containing the discretized solution variables $\tilde{\mathbf{v}}_s$, $\tilde{\mathbf{p}}_f$ and $\tilde{\mathbf{p}}_c$, and \mathbf{r} the right-hand-side vector containing all forcing terms. The coefficient matrices $\mathbf{V} \mathbf{V}$, $\mathbf{V} \mathbf{F}$ and $\mathbf{V} \mathbf{C}$ contain coefficients related to solid deviatoric stress divergence and gradients of fluid and compaction pressure, respectively. The matrix $\mathbf{F} \mathbf{F}$ contains coefficients related to the divergence of the Darcy flux and $\mathbf{C} \mathbf{C}$ contains coefficients related to compaction/decompaction rheology of the host rock. The coefficient matrices arising from the finite-element discretization take the following form:

$$\begin{aligned} \mathbf{V} \mathbf{V} &= \int_V \tilde{\mathbf{B}}^T \mathbf{D} \tilde{\mathbf{B}} dV, & \mathbf{F} \mathbf{F} &= - \int_V \tilde{\mathbf{G}}_f^T \mathbf{K}_D \tilde{\mathbf{G}}_f dV, \\ \mathbf{V} \mathbf{F} &= - \int_V \tilde{\mathbf{G}}_v^T \mathbf{N}_f dV, & \mathbf{C} \mathbf{C} &= - \int_V \mathbf{N}_p^T \mathbf{N}_p C dV, \\ \mathbf{V} \mathbf{C} &= - \int_V \tilde{\mathbf{G}}_v^T \mathbf{N}_c dV, \end{aligned} \quad (\text{A9})$$

where $\tilde{\mathbf{B}}$, $\tilde{\mathbf{G}}_v$ and $\tilde{\mathbf{G}}_f$ are the discretized counterparts of the strain rate and gradient operators introduced above (eqs A1–A4), containing suitably arranged derivatives of the linear shape functions:

$$\begin{aligned} \tilde{\mathbf{B}} &= \begin{bmatrix} \partial_x N_1 & 0 & \partial_x N_2 & 0 & \partial_x N_3 & 0 & \partial_x N_4 & 0 \\ 0 & \partial_y N_1 & 0 & \partial_y N_2 & 0 & \partial_y N_3 & 0 & \partial_y N_4 \\ \partial_y N_1 & \partial_x N_1 & \partial_y N_2 & \partial_x N_2 & \partial_y N_3 & \partial_x N_3 & \partial_y N_4 & \partial_x N_4 \end{bmatrix}, \\ \tilde{\mathbf{G}}_v &= [\partial_x N_1 \quad \partial_y N_1 \quad \partial_x N_2 \quad \partial_y N_2 \quad \partial_x N_3 \quad \partial_y N_3 \quad \partial_x N_4 \quad \partial_y N_4], \\ \tilde{\mathbf{G}}_f &= \begin{bmatrix} \partial_x N_1 & \partial_x N_2 & \partial_x N_3 & \partial_x N_4 \\ \partial_y N_1 & \partial_y N_2 & \partial_y N_3 & \partial_y N_4 \end{bmatrix}. \end{aligned} \quad (\text{A10})$$

And finally, the components of the right-hand-side vector \mathbf{r} are defined as

$$\begin{aligned} \mathbf{r}_v &= \int_V \mathbf{N}_v^T [\bar{\rho} \mathbf{g} \hat{\mathbf{z}} - \mathbf{B}^T \chi_\tau^* \tilde{\boldsymbol{\tau}}_s^o] dV, \\ \mathbf{r}_f &= \int_V \tilde{\mathbf{G}}_f^T \mathbf{K}_D \rho_f \mathbf{g} \hat{\mathbf{z}} dV, \\ \mathbf{r}_c &= - \int_V \mathbf{N}_c^T C \chi_p^* \Delta \bar{p}^o dV, \end{aligned} \quad (\text{A11})$$

where $\tilde{\boldsymbol{\tau}}_s^o$ and $\Delta \bar{p}^o$ denote the discrete stresses and pressure differences obtained from the previous time step, resulting from the elastic stress rates that are discretized in order to derive the rheological constitutive laws (eqs 23–27 and 33). The coefficients related to solid velocity divergence in the mass conservation and compaction equation can, due to the symmetry of the finite element formulation,

be expressed as the transpose matrices of the two pressure gradients, \mathbf{VF}^T and \mathbf{VC}^T . Therefore, since \mathbf{VV} , \mathbf{FF} and \mathbf{CC} each are symmetric (provided that the boundary conditions are set properly), the matrix \mathbf{L} is symmetric. Note that the formulation of the governing equations enables the discrete form of these equations to be symmetric. Combining this symmetric property, together with the fact that the bulk mass conservation and compaction equations (unlike the incompressible Stokes mass conservation) have non-zero diagonals, we expect to be able to exploit solution strategies which have been developed to solve variable viscosity Stokes problems.

In this implementation, we solve the linear system of governing equations using Matlab's 'backslash' direct solver. Some of the coefficients in these equations, however, are nonlinear. The nonlinearity is caused by plasticity-related and melt-fraction-dependent material properties, those being both effective viscosities, permeability, pore modulus and bulk density. Therefore, we employ a defect correction, fixed point iteration scheme (Picard) to solve the system of nonlinear equations. During each nonlinear iteration, nonlinear material properties in the submatrices \mathbf{VV} , \mathbf{FF} and \mathbf{CC} are updated to the current solution and the system is solved again to update the solution. These nonlinear iterations are repeated until the nonlinear residuals of the solution are decreased below a given limit. The iteration procedure for solving the governing equations is executed in the following sequence of steps (i and k denote the current iterative step and time step, respectively).

1. Assemble the global stiffness matrix and right-hand-side vector using the values of nonlinear material properties from previous time step (or initial condition at first time step): $\mathbf{L}(\mathbf{x}^{k-1})$, $\mathbf{r}(\mathbf{x}^{k-1})$.

2. Solve the full system as an initial solution guess for the nonlinear solve,

$$\mathbf{x}^i = \mathbf{L}(\mathbf{x}^{k-1}) \setminus \mathbf{r}(\mathbf{x}^{k-1}). \quad (\text{A12})$$

3. Enter nonlinear iteration loop:

a. Update the nonlinear material properties according to the current solution and re-assemble global stiffness matrix and right-hand-side vector $\mathbf{L}(\mathbf{x}^i)$, $\mathbf{r}(\mathbf{x}^i)$.

b. Compute the nonlinear residual vector $\mathbf{f}(\mathbf{x}^i)$ of the current guess using the updated global stiffness matrix and right-hand-side vector. Use the norm of residuals normalized by the norm of the right-hand side (f^i) to monitor convergence of nonlinear iterations:

$$\mathbf{f}(\mathbf{x}^i) = \mathbf{L}(\mathbf{x}^i) \mathbf{x}^i - \mathbf{r}(\mathbf{x}^i), \quad f^i = \frac{\|\mathbf{f}(\mathbf{x}^i)\|}{\|\mathbf{r}(\mathbf{x}^i)\|}. \quad (\text{A13})$$

c. Compute iterative update for current solution:

$$\mathbf{dx}(\mathbf{x}^i) = \mathbf{L}(\mathbf{x}^i) \setminus \mathbf{f}(\mathbf{x}^i). \quad (\text{A14})$$

d. Update the current guess of solution using computed iterative update:

$$\mathbf{x}^{i+1} = \mathbf{L}(\mathbf{x}^i) + \mathbf{dx}(\mathbf{x}^i). \quad (\text{A15})$$

e. Repeat steps a to d until converged and $f^i \leq f_{\text{crit}}$.

Generally, the convergence of these iterations is very well-behaved with up to one order of magnitude reduction of the nonlinear residuals per iterative update. There are, however, a number of issues concerning convergence of the nonlinear iterations especially relating to plasticity, and some other measures are required to optimize the performance of the code; all of these are outlined in more detail below.

A.2 Matrix scaling

Rather than using non-dimensionalized equations, a strategy otherwise common in numerical modelling of magma dynamics (Barcilon & Lovera 1989; Spiegelman 1993b), we use the dimensional form of the two-phase problem in our simulation code. The main advantage of a non-dimensional approach would be to extract non-dimensional numbers from the set of equations that help to understand how physical processes are balanced against each other. However, we prefer to separate this useful step of additional analysis from the simulation of the actual physics, as it renders it more straightforward to relate input or output parameter values to physical quantities of interest (May & Moresi 2008).

Due to large variations in magnitude of some material properties (i.e. viscosity and permeability), the coefficients in the global stiffness matrix \mathbf{L} may vary by up to 10 orders of magnitude. Such a matrix configuration is not optimal for use in a direct solver like Matlab's 'backslash'. To improve solution stability and solver speed, a scaling of the linear problem is applied with the use of a diagonal scaling matrix $\mathbf{\Lambda}$ (Pelletier *et al.* 1989). The scaled problem is expressed as

$$\hat{\mathbf{x}} = \mathbf{\Lambda} \mathbf{L} \mathbf{\Lambda} \setminus \mathbf{\Lambda} \mathbf{r}, \quad \mathbf{x} = \mathbf{\Lambda} \hat{\mathbf{x}}, \quad (\text{A16})$$

where $\hat{\mathbf{x}}$ is the scaled solution vector resulting from the scaled linear solve. The scaling matrix $\mathbf{\Lambda}$ is a diagonal matrix containing elementwise scaling factors $\lambda_{v,f,c}^{\text{el}}$ on its diagonal. One possible choice of finding these scaling factors is related to the diagonal coefficients of the unscaled global stiffness matrix itself:

$$\mathbf{\Lambda} = \frac{1}{\sqrt{|\mathbf{L}|}} \mathbf{I}. \quad (\text{A17})$$

This choice leads to the diagonal entries of the scaling matrix to be of order $O(1)$. Scaling the linear problem may result in a solver speed-up of up to 400 per cent with Matlab's 'backslash' function, which is due to the fact that Matlab chooses a more efficient factorization based on the properties of the coefficient matrix.

A.3 Solver stabilization

As mentioned above, we use quadrilateral brick elements to discretize the two-phase problems. We chose a first-order element type, primarily because it is the least expensive in terms of memory requirements and solver time. We use linear Q1 shape functions for solid velocity and fluid pressure. The linear shape functions facilitate the retrieval of strain rates and stresses from the solid velocity gradients, as well as Darcy flux and fluid velocities from the fluid pressure gradient. Compaction pressure mainly relates to rheology and is therefore defined as piecewise constant on P0 elements, as are all rheological parameters.

However, in the limit of zero melt fraction, the system of equations reduces to a Stokes flow problem (see Section 2.4.2) discretized on Q1Q1 elements. This element type is not LBB-stable for a Stokes problem, and thus it will be susceptible to instabilities in form of oscillating pressure solutions. Such instabilities mostly occur in regions of the model, where the compaction length drops below the resolution of the finite-element mesh due to a very small or zero melt fraction.

One of the standard strategies used to stabilize the Q1Q1 Stokes problem is the use of a Laplacian term, introducing a small amount of diffusivity on the pressure field discretized on Q1 elements (Elman *et al.* 2005). Such a term is naturally included in our set of equations arising from the divergence of Darcy flux in the bulk

mass conservation equation (eq. 58). In regions, where melt fraction goes to zero, we therefore introduce a non-zero, lower cut-off to the permeability field to ensure that the stabilizing effect of this term on the fluid pressure field stays effective. A cut-off value of $1.e - 20 \leq k_{\text{cut-off}} \leq 1.e - 18$ is used in this study, depending on the grid resolution.

A second instability linked to the use of linear elements is the occurrence of checkerboard oscillations in the compaction pressure field, especially around sharply localized zones of deformation due to brittle plasticity. This instability may be avoided by applying a small amount of smoothing to the quantities related to plastic yielding, such as differential strain rates and compaction pressure. The smoothing is accomplished by an operation representing an averaging over adjacent elements. For a general property a on the i th element of the j th row of the regular element mesh (with i, j denoting the element indices of the regular mesh), we obtain the smoothed field \hat{a} with the following relation:

$$\hat{a}_{i,j} = (1 - \gamma) a_{i,j} + \gamma (a_{i-1,j} + a_{i+1,j} + a_{i,j-1} + a_{i,j+1}). \quad (\text{A18})$$

The smoothing parameter is chosen as $0.01 \leq \gamma \leq 0.1$. Doing so does not qualitatively alter the solution, but renders it slightly smoother by limiting solution variability from one element to the next. We apply this smoothing to compaction pressure as well as deviatoric and volumetric strain rate fields, and use these smoothed values to update nonlinear material properties (see end of Section A1.1).

As mentioned above, convergence of the nonlinear iterations is typically very well-behaved. However, testing the code for various geodynamically relevant setups, we find two major issues leading to poor convergence of the nonlinear solver. The first occurs, when few elements, usually located at sharp edges of brittle fracture zones, go in and out of plastic yielding from one iteration to the next. This oscillatory behaviour manifests itself in alternatively weakening and strengthening the effective viscosity of those elements concerned. To avoid this instability, we apply a rheological damping parameter α_{rheo} used to update deviatoric and volumetric viscosities from one iterative step to the next as

$$[\eta_{\text{eff}}^*, \xi_{\text{eff}}^*]^{i+1} = \left([\eta_{\text{eff}}^*, \xi_{\text{eff}}^*]^{i+1} \right)^{\alpha_{\text{rheo}}} \left([\eta_{\text{eff}}^*, \xi_{\text{eff}}^*]^{i+1} \right)^{1-\alpha_{\text{rheo}}}. \quad (\text{A19})$$

The second issue is related to plastic yielding representing a horizontal cut-off in a stress–strain rate diagram. Around the cut-off corner, where the yield criterion intersects the viscoelastic stress curve, the nonlinear iterations are most susceptible to such oscillations. To avoid the issues related to a sharp cut-off corner, we may calculate each visco-elasto-plastic effective viscosity ($\eta_{\text{eff}}^*, \xi_{\text{eff}}^*$) in a way that it fits the stress–strain rate curve asymptotically, thus avoiding the corner. These alternative plastic yield viscosities introduce a small amount of additional weakening of material under stress conditions just prior to reaching the yield stress. We use the parameter $0 < \beta_{\text{rheo}} < 1$ to weight between the viscosities representing a cut-off or asymptotical stress curve, so as to avoid the corner effect with as little additional pre-plastic yielding as possible. We calculate the effective visco-elasto-plastic viscosities in the following way:

$$\begin{aligned} [\eta_{\text{eff}}^*, \xi_{\text{eff}}^*]^{\text{cut-off}} &= \min([\eta_{\text{ve}}^*, \xi_{\text{ve}}^*], [\eta_{\text{eff}}^*, \xi_{\text{eff}}^*]), \\ [\eta_{\text{eff}}^*, \xi_{\text{eff}}^*]^{\text{asympt}} &= \frac{1}{\frac{1}{[\eta_{\text{ve}}^*, \xi_{\text{ve}}^*]} + \frac{1}{[\eta_{\text{eff}}^*, \xi_{\text{eff}}^*]}}, \\ [\eta_{\text{eff}}^*, \xi_{\text{eff}}^*] &= \left([\eta_{\text{eff}}^*, \xi_{\text{eff}}^*]^{\text{cut-off}} \right)^{\beta_{\text{rheo}}} \left([\eta_{\text{eff}}^*, \xi_{\text{eff}}^*]^{\text{asympt}} \right)^{1-\beta_{\text{rheo}}}. \quad (\text{A20}) \end{aligned}$$

A.4 Limits of two-phase flow

In the presence of deep melt (pore space not interconnected to the surface), the effective pressure generally varies around zero, thus reducing plastic strength of the host rock to around cohesion. In regions with very small fluid fraction, the assumption of an interconnected pore space sustaining any significant influence on plastic rock strength becomes invalid. Therefore, we choose to fix compaction pressure to zero and set the effective pressure to equal the total pressure, where $\phi < \phi_{\text{crit}}$ and introduce a parameter x_ϕ indicating the presence of fluid phase beyond the critical melt fraction necessary for two-phase flow physics to become effective:

$$x_\phi = \begin{cases} 0 & \text{for } \phi < \phi_{\text{crit}}, \\ 1 & \text{for } \phi \geq \phi_{\text{crit}}. \end{cases} \quad (\text{A21})$$

The parameter x_ϕ indicates where two-phase flow is active (i.e. effective pressure is equal to compaction pressure, shear plasticity is weakened, tensile plasticity is enabled) or inactive (i.e. effective pressure is equal to total pressure, compaction pressure is held at zero, shear plasticity is frictional, tensile plasticity is switched off, $k_\phi = k_{\text{cut-off}}$). The appropriate value to use for critical melt fraction probably depends on the grain size and geometry of the host rock as well as the viscosity of the melt ($\phi_{\text{crit}} \sim K_D$). For reasons of simplicity, however, we use a constant critical melt fraction value of $\phi_{\text{crit}} = 0.1$ per cent.

Additionally, in order to prevent extreme values for melt-fraction-dependent properties where either fluid or solid fractions go to zero, we cut off melt and solid fraction at $1.e-6$ and $1.e-2$, respectively, wherever they go into the computation of any material-fraction-dependent properties. For all other situations, we still let material fractions go from 0 to 1.

A.5 Marker-in-cell advection

For the marker-in-cell advection scheme employed in this study, material properties are generally defined on two sets of Lagrangian marker particles, one each for properties of the host rock and of the silicate melt phase. Each set of marker particles is then advected with the respective velocity field. To update the material properties according to the current solution variables, these are interpolated to the marker particles after each solve. Material properties are only interpolated to elements or nodes when required to evaluate the weak form associated with the finite-element discretization.

Solid velocities and fluid pressures are defined on linear shape functions and are thus consistently interpolated from nodes to markers using these shape functions. The nodal coordinates of each element are used to perform an inverse coordinate transformation determining the local (or natural) coordinates ξ of each marker within a given element. The four linear shape functions may now be evaluated at the marker's local coordinates $N_i(\xi_m)$. Now, a simple operation will yield the interpolated property value a_m at each marker's location as

$$a_m = \sum_i N_i(\xi_m) a_i, \quad (\text{A22})$$

with a_i the nodal values, and N_i are the corresponding linear shape functions evaluated at local coordinates of marker m .

For the inverse interpolation, summing marker contributions a_m found around each node means that the sum contains shape functions that are evaluated in different local coordinate frames of the four elements surrounding each node. To take into account that elements might be slightly deformed in case of a free surface or kinematic

boundary conditions, each contribution needs to be weighted by the volume of each element, ΔV_{el} , leading to the following expression for marker to node interpolation:

$$a_i = \frac{\sum_m a_m N_i(\xi_m) \Delta V_{el}}{\sum_m N_i(\xi_m) \Delta V_{el}}. \quad (\text{A23})$$

Element properties (i.e. compaction pressure, melt fraction and all material properties like density, permeability, viscosity, etc.) are defined piecewise constant over an element, which leads to a slightly different interpolation procedure. The shape functions for element properties are defined as unity; therefore, the interpolation from elements to marker particles is a simple matter of copying element properties (a_{el}) to all markers located within an element:

$$a_m = a_{el}. \quad (\text{A24})$$

The inverse interpolation from marker to elements is accomplished similarly to eq. (A23), only that the element volumes cancel out, as the sum of marker properties is assembled in one element only at one time:

$$a_{el} = \frac{\sum_m a_m}{n_{el}}, \quad (\text{A25})$$

with n_{el} denoting the total number of marker particles found in an element.

It has been found that using a harmonic averaging scheme to interpolate viscosity from markers to elements leads to more accurate results in the presence of large-viscosity contrasts (Deubelbeiss & Kaus 2008). We applied the same logic to the permeability field as well, as it may also vary over orders of magnitude, and thus use the following relation for interpolations of viscosity and permeability from markers to elements:

$$a_{el} = \frac{n_{el}}{\sum_m 1/a_m}. \quad (\text{A26})$$

In order to ensure an accurate time evolution, we update time-dependent properties on markers incrementally as

$$a_m(t + \Delta t) = a_m(t) + [a_{el}(t + \Delta t) - a_{el}(t)], \quad (\text{A27})$$

where each marker property is updated with the incremental change the property a_{el} has experienced during the current time step. Such incremental updates coupled with strongly variable velocity fields may lead to artificial heterogeneities in the marker values. These may be avoided by applying a small amount of subgrid diffusion to such marker fields (Gerya & Yuen 2003), accomplished by the following operation:

$$a_m = (1 - \kappa_{sub}) a_m + \kappa_{sub} a(\mathbf{x}_m), \quad (\text{A28})$$

with $a(\mathbf{x}_m)$ being the property value fully interpolated from nodes to marker position. The subgrid diffusion coefficient is chosen at small values of $0.001 \leq \kappa_{sub} \leq 0.1$. Incremental time updates combined with subgrid diffusion are applied to compaction pressure P_c and solid deviatoric stresses τ_s (both time-dependent due to elastic stress build-up), and melt fraction ϕ .

A final issue to consider when employing a marker-in-cell advection scheme in two-phase flow is that both solid and fluid velocity fields are divergent, thus accumulating marker particles in regions of $\dot{v}_s < 0$, and creating gaps in the marker distribution where $\dot{v}_s > 0$. We have introduced a particle reseeding algorithm in our code to ensure that the marker distribution remains approximately uniform. If the number of markers per element differs from the initial value

by more than 25 per cent, a new set of evenly distributed markers is inserted into the element under consideration. Marker properties of newly inserted particles are copied from the nearest particle presently located in that element. After this reseeding operation, the original markers in the element under treatment are discarded.

The complete sequence of operations performed to accomplish marker-in-cell advection after a completed solver cycle during a given time step is outlined here:

1. Update material properties and solution variable fields on markers according to the shape functions on which they are discretized (eqs A22 and A24). Time-dependent fields are updated incrementally according to eq. (A27).
2. Move two sets of marker particles according to the solid and fluid velocity field, respectively, to obtain new location of each marker in global coordinates \mathbf{x}_m . We employ a 4th-order Runge–Kutta scheme to ensure accuracy of advective transport.
3. Identify which markers are located in each element after advection. This is accomplished efficiently through a closest point algorithm provided by Matlab Central (Tagliasacchi 2008).
4. Calculate local coordinates of each marker particle ξ_m inside an element with nodal coordinates \mathbf{x}_i by solving the inverse of the nonlinear equation describing the element geometry

$$\mathbf{x}_m = \sum_i N_i(\xi_m) \mathbf{x}_i. \quad (\text{A29})$$

Eq. (A29) is solved by employing a Picard iterative procedure.

5. Interpolate material properties and solution variable fields back to elements and nodes.
6. Check number of markers per element after advection. Add/delete markers where necessary.
7. Apply subgrid diffusion according to eq. (A28) to time-dependent properties.
8. Correct time-dependent marker fields for potential inaccuracies introduced by either adding/deleting markers or applying subgrid diffusion, using a correction Δa_{el} found by comparing interpolated values on elements *before* to interpolated values on elements *after* adding/deleting and/or subgrid diffusion:

$$\Delta a_{el} = \left(\frac{\sum_m a_m}{n_{el}} \right)^{before} - \left(\frac{\sum_m a_m}{n_{el}} \right)^{after}, \quad (\text{A30})$$

$$\hat{a}_m = a_m + \Delta a_{el}.$$

The resulting advection scheme is both robust and accurate and avoids numerical diffusion that is inevitable with most grid-based advection schemes. The only drawback here is the large number of markers (16–25 particles per element) needed to achieve adequate resolution. However, if appropriate infrastructure is available, the marker treatment may nevertheless be done efficiently. In the case of this study, the factor limiting computational speed is the direct solver, using up to five times more CPU-time per time step than the marker-in-cell treatment.

APPENDIX B: CODE BENCHMARKS

B.1 Rayleigh–Taylor benchmarks

The methodology presented above reduces to the standard Stokes flow problem in the absence of any melt phase. In typical applications of melt extraction from the upper mantle through lithosphere and crust, up to 90 per cent of the model domain may

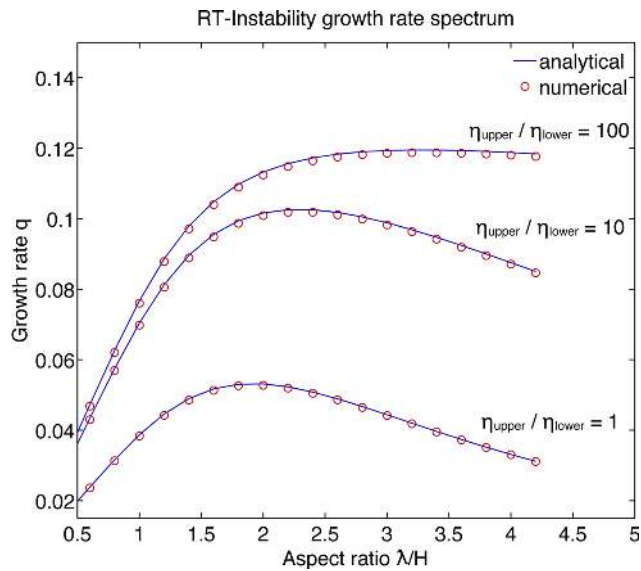


Figure B1. Relative growth rate q plotted against aspect ratio λ/H for a two-layer compositional RT-instability. Blue lines indicate analytical solutions, red circles mark numerical results of individual runs at viscosity contrasts between the two layers of $\eta_{\text{upper}}/\eta_{\text{lower}} = 100$ (uppermost curve), $\eta_{\text{upper}}/\eta_{\text{lower}} = 10$ (middle curve) and $\eta_{\text{upper}}/\eta_{\text{lower}} = 1$ (lowermost curve).

remain unmolten for the duration of the simulation. Therefore, we need to ensure that the numerical results for unmolten material reproduce some common features of Stokes flow, such as compositional Rayleigh–Taylor (RT) instabilities. Here, we present two benchmark problems employing a two-layer compositional model (layer 2 on top of layer 1) with an inverse density difference of $\Delta\rho_c = \rho_{\text{upper}} - \rho_{\text{lower}}$ (lower density layer below) and a small sinusoidal initial amplitude perturbation that will give rise to a RT instability.

B.1.1 RT growth rate spectrum benchmark

For a two-layer instability with small-amplitude perturbation, the initial growth rate is known analytically (Ramberg 1981). Plotting this growth rate at various wave length λ of initial perturbation relative to the model box thickness H results in a RT growth rate spectrum, revealing a maximum growth rate corresponding to a certain wave length. Such a growth rate depends on a number of parameters, notably the viscosity contrast between the two layers. For a model box of depth H , a layer thickness of $H/2$, a constant density difference $\Delta\rho_c$, a viscosity contrast of $\eta_{\text{upper}}/\eta_{\text{lower}} = [1, 10, 100]$ and an initial amplitude perturbation of $a = a_0 \cos 2\pi/\lambda$ with $a_0 \ll H$ a very small amplitude, the growth rate spectrum given in Fig. B1 is obtained, run at a numerical resolution of 160×80 elements. The numerical data fit the analytical spectrum very well for all three viscosity contrasts we considered.

B.1.2 Van Keken benchmark

This next benchmark problem (Van Keken *et al.* 1997) again employs a two-layer compositional RT-instability. Rather than focusing on initial growth rates, this benchmark deals with the long-term evolution of an isoviscous compositional overturn problem and thus will above all test the accuracy of the marker-in-cell advection scheme over long-time evolutions. The initial setup features a model box of depth H , a buoyant bottom layer of thickness $H/5$, an aspect ratio

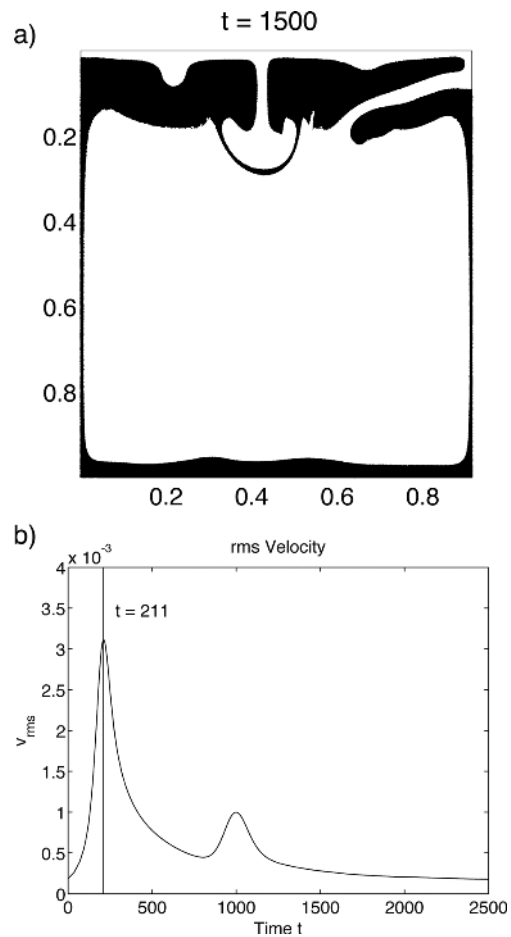


Figure B2. (a) Compositional field of Van Keken convection benchmark after non-dimensional model time $t = 1500$; (b) root-mean-square velocity evolution with time of the same benchmark run, with peak velocities reached at $t = 211$, in agreement with the benchmark results.

of $r_a = 0.9142$ (set to this value as it corresponds to the maximum growth rate of a RT-instability of given geometry), a density contrast $\Delta\rho_c$, an isoviscous rheology ($\eta_2/\eta_1 = 1$) and an initial amplitude of $a = 0.02H \sin \pi/r_a$. This simulation was performed using a mesh of 200×200 elements containing 25 markers per element. Markers are advected with a 4th-order Runge–Kutta scheme along the solid velocity field.

The results at non-dimensional model time of 1500 are given along with the evolution of rms velocity over time in Fig. B2. The peak velocity occurs at a non-dimensional time of $t = 211$. Please refer to the original paper of Van Keken *et al.* (1997) for a comparison of these results. Both the timing of the layer overturn as well as the shape of the resulting flow patterns compares well with results presented in the original paper, leading us to conclude that the advection scheme in our code is able to accurately advect material properties in problems with large total deformation.

B.2 Visco-elasto-plastic benchmarks

An important feature of visco-elasto-plastic Stokes flow of lithosphere and crust is the interplay of viscous, elastic and plastic modes of deformation. To demonstrate the functionality of the visco-elasto-plastic rheology, we present three benchmark problems testing viscoelastic deviatoric stress build-up and plastic shear failure.

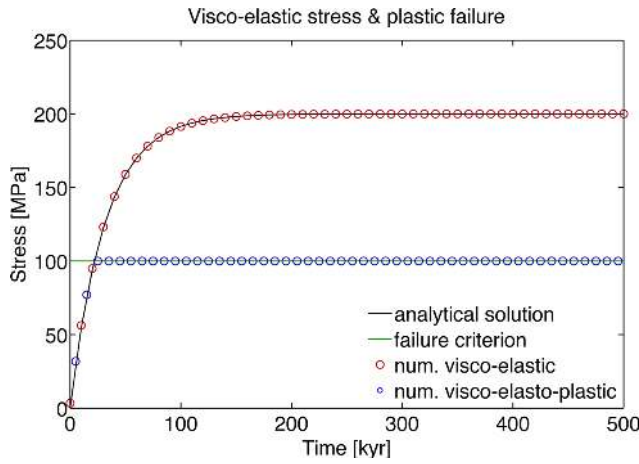


Figure B3. Viscoelastic stress evolution with time under pure shear deformation. The black line gives analytical stress solution, the dark green horizontal line indicates the plastic failure criterion, red circles mark numerical results for a run without plastic failure rheology and blue circles mark results of a run with plastic failure switched on.

B.2.1 Visco-elasto-plastic evolution of shear stress

In a homogeneous 0-D numerical experiment under pure shear imposed by constant strain rate side boundaries ($\dot{\epsilon}'_{BG}$), the stress due to viscoelastic deviatoric deformation in the model box should evolve according to the relation $\bar{\tau}_{II} = 2\eta\dot{\epsilon}'_{BG}[1 - \exp(-Gt/\eta)]$. If plastic shear failure is considered as well, the deviatoric stress will be limited by the given failure criterion $\bar{\tau}_{II} \leq \sigma_y$. We have performed two such 0-D tests, the first one allowing only the viscous and elastic modes of deformation, whereas in the second test plastic shear failure is considered as well. The numerical results depicted in Fig. B3 show exact agreement with the analytically predicted stress curve; therefore, we are confident that the visco-elasto-plastic stress evolution is reproduced accurately by our rheological implementation.

B.2.2 Visco-elasto-plastic evolution of compaction stress

Similarly, in a homogeneous 0-D numerical experiment under volumetric expansion imposed by constant strain rate boundaries (\dot{v}_{BG}), the compaction due to viscoelastic volumetric deformation in the model box should evolve according to the relation $P_c = -(1 - \phi)\xi\dot{v}_{BG}[1 - \exp(-K_\phi t/\xi)]$. If plastic shear failure is considered as well, the compaction pressure will be limited by the given failure criterion $P_c \geq P_y$. We have performed two such 0-D tests, the first one allowing only the viscous and elastic modes of deformation, whereas in the second test plastic shear failure is considered as well. The numerical results depicted in Fig. B4 again show exact agreement with the analytically predicted pressure curve, which illustrates that the visco-elasto-plastic pressure evolution is implemented accurately as well.

B.2.3 Viscoelastic bending beam

For this next benchmark problem, we employ an initial compositional field representing a dominantly elastic beam, fixed to, and protruding horizontally from the left wall of the model box. Surrounding the elastic beam is a viscous, but inelastic fluid. All boundaries are free slip, except for the left wall, which is set to no slip in order to keep the bending beam fixed to the wall. The beam has a higher density than the surrounding fluid and

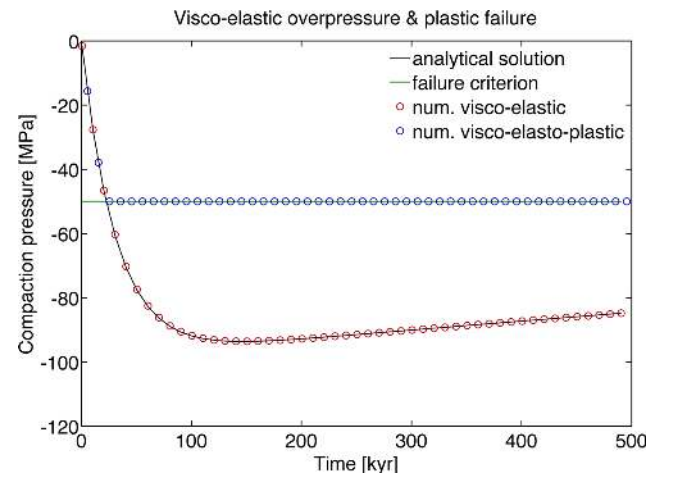


Figure B4. Viscoelastic pressure evolution with time under constant volumetric expansion. The black line gives analytical pressure solution, the dark green horizontal line indicates the plastic failure criterion, red circles mark numerical results for a run without plastic failure rheology and blue circles mark results of a run with plastic failure included. Note that the positive slope of the viscoelastic pressure curve is due to the steadily increasing melt fraction caused by the imposed constant matrix expansion.

thus will bend down elastically driven by gravity. After the beam has accumulated some elastic strain through bending down, we switch off gravity. If the stress evolution is implemented accurately, the elastic beam should now, free from the pull of gravity, move upwards again and restore its initial position. We run this setup twice, with constant density difference and constant properties of the fluid ($\Delta\rho_c = 500 \text{ kg m}^{-3}$, $\eta_{\text{fluid}} = 1e + 18 \text{ Pa s}$ and $G_{\text{fluid}} = 1e + 11 \text{ Pa}$), while varying the viscosity of the bending beam (first run $\eta_{\text{beam}} = 1e + 24 \text{ Pa s}$, $G_{\text{beam}} = 1e + 10 \text{ Pa}$; second run $\eta_{\text{beam}} = 5e + 22 \text{ Pa s}$, same shear modulus). This choice of parameters leads to a Maxwell time $t_m = 0.32 \text{ yr}$ for the background fluid and Maxwell times of $t_m = 3.2 \text{ Myr}$ (first run) and $t_m = 0.16 \text{ Myr}$ (second run), meaning that the deformation in this benchmark problem, which occurs on a timescale of thousands to a million years, will lead to dominantly viscous deformation in the fluid, and dominantly elastic behaviour of the beam in the first run, and mixed viscoelastic deformation of the beam in the second run. The numerical resolution used was 300×200 elements, with 16 markers per elements for stress advection.

Two time frames of each run, given in Fig. B5, demonstrate how, after gravity is switched off at 50 kyr model time, the dominantly elastic beam of the first run recovers all elastic deformation and thus returns to its initial position. The mixed viscoelastic beam of the second run, however, only recovers part of the total deformation. These results demonstrate how the viscoelastic rheology may represent both purely elastic deformation, as well as intermediate combinations of restorable elastic strain and non-restorable viscous flow.

B.2.4 Visco-elasto-plastic shear bands

Plastic shear failure under pure shear extension or compression leads to the formation shear bands at characteristic angles. A homogeneous 2-D test setup with a small, circular weak inclusion in the middle should produce shear bands with an angle of around 60° relative to the horizontal direction under extension, whereas shear bands at an angle of around 30° are to be expected under compression.

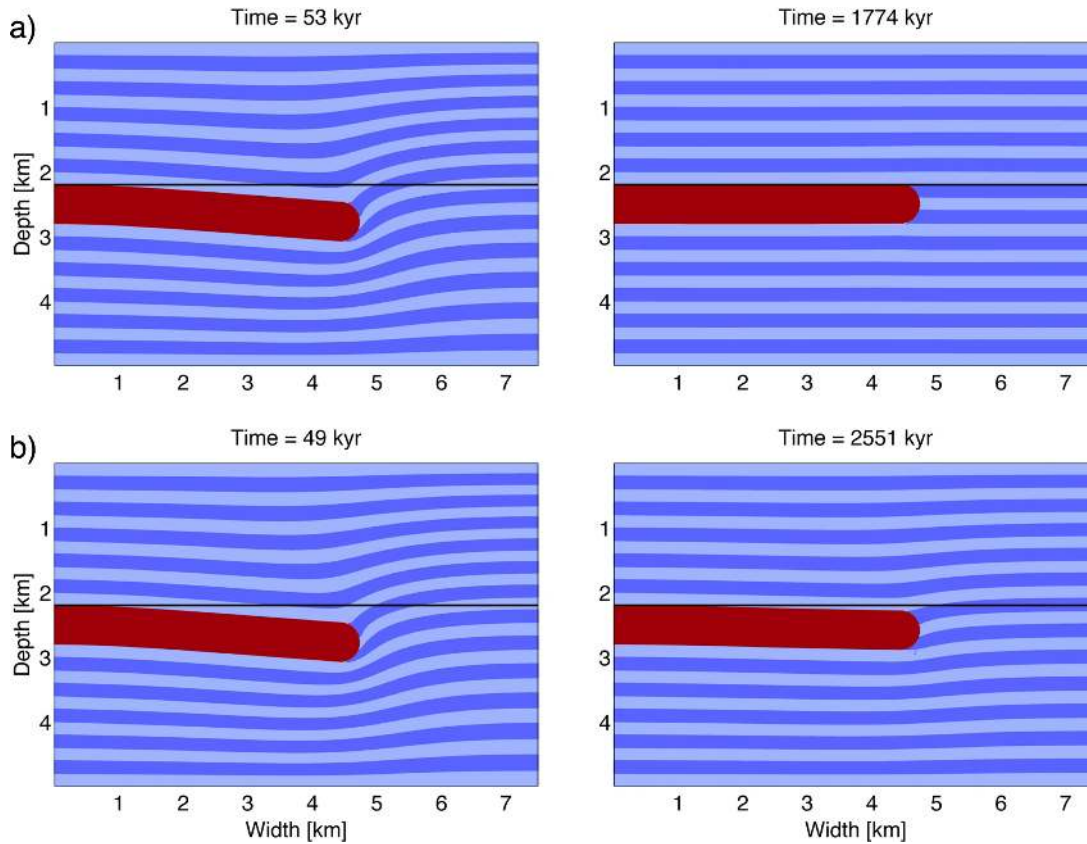


Figure B5. Two time frames of two benchmark runs featuring (a) a dominantly elastic beam (red colour, $\eta = 1e + 24$ Pa s, $G = 1e + 10$ Pa, $t_m = 3.2$ Myr), and (b) a mixed viscoelastic beam (red colour, $\eta = 5e + 22$ Pa s, $G = 1e + 10$ Pa, $t_m = 0.16$ Myr), each bending inside a box of viscous fluid (blue colours), given at time of maximum deformation (panels to the left) and after full relaxation, after gravity was switched off (panels to the right). Black line indicates the initial level of the beam for reference. As expected, the elastic beam fully recovers, whereas the viscoelastic beam only partly recovers.

Fig. B6 displays the results of this test, run on a mesh of 240×120 elements, at a reference viscosity of $\eta_0 = 1e + 23$ Pa s, a shear modulus of $G = 10$ GPa, a cohesion of $C = 100$ MPa, a friction angle of $\phi = 30^\circ$ and a background strain rate of $\dot{\epsilon}' = \pm 1e - 14$ s $^{-1}$. The results demonstrate that our implementation of viscoelastic stress build-up to plastic failure leads to the formation of shear bands at the correct characteristic angles under both extension and compression.

B.3 Solitary wave benchmark

A well-known feature of geodynamic two-phase flow equations such as the ones used in our code is the emergence of melt fraction instabilities in the form of solitary waves. These have been studied both analytically (Scott & Stevenson 1984; Barcilon & Richter 1986) and numerically (Barcilon & Lovera 1989; Spiegelman 1993b; Connolly & Podladchikov 1998) in the context of magma dynamics. Solitary waves thus serve as an apt benchmark problem for numerical codes of magma migration (Simpson & Spiegelman 2011; Richard *et al.* 2012). We perform runs on a pseudo-1-D profile of uniform background melt fraction with depth, perturbed by a discrete representation of a 1-D solitary wave solution. If our numerical implementation is accurate, this initial wave with relative amplitude A_ϕ should propagate upwards at constant non-dimensional speed of $c = 2A_\phi + 1$ without changing shape. Note, however, that this analytical solution is valid for the small melt fraction limit only.

We employ a series of runs at various resolutions in space and time. The 2-D finite-element mesh is used as a pseudo-1-D section

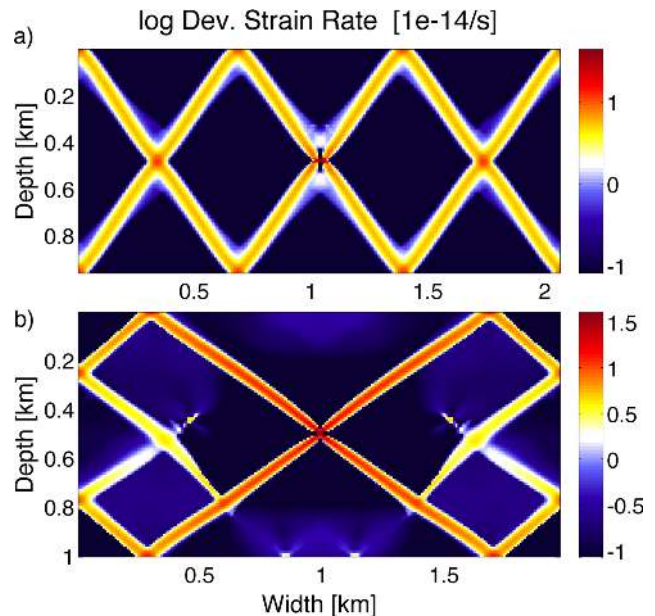


Figure B6. Results of two benchmark runs testing the emergence of characteristic angled shear bands under (a) pure shear extensional and (b) pure shear compressional boundary conditions. Displayed are deviatoric differential strain rates relative to background strain rate of $\dot{\epsilon}_{BG} = 1e - 14$ s $^{-1}$ in logarithmic scaled.

by keeping the number of elements in horizontal direction constant at $n_x = 4$. Vertical resolution is varied as multiples by two of the coarsest grid with $n_z = 100$. The depth of the modelled profile is $D = 200\delta_0$, with δ_0 being the compaction length at background melt fraction of 0.1 per cent. Other material parameters are chosen to represent conditions in a partially molten upper mantle at very low melt fractions. Deviatoric and volumetric solid viscosities are held equal and constant at $1.e20$ Pa s, fluid viscosity is 100 Pa s and permeability at background melt fraction is $5.e-18$ m², with a melt fraction exponent in the permeability law of $n = 3$.

To quantify the accuracy of the solution, three measures arise naturally from the problem set: (i) the shape of the solitary wave should be conserved; (ii) the shape of the compaction pressure anomaly linked to the solitary wave should be conserved; (iii) the speed of wave propagation should be constant. To determine all three quantities, we need to find the phase shift Δ , defined as the difference between the analytically predicted and the numerically computed z -coordinate of the wave peak with time (Simpson & Spiegelman 2011). In the following, we will denote numerically computed quantities with a curled over bar. The phase shift is given by

$$\Delta = z_p - \tilde{z}_p. \quad (\text{B1})$$

The numerical phase speed may subsequently be found from the relationship

$$\tilde{c} = c - \frac{\Delta}{t}. \quad (\text{B2})$$

We measure the relative error of the phase speed as

$$e_c = \left| \frac{\tilde{c}}{c} - 1 \right|. \quad (\text{B3})$$

The preservation of the shape of melt fraction and compaction pressure fields is quantified by computing the root-mean-square error of the melt fraction and compaction pressure solution normalized by their respective peak amplitude as

$$e_f = \frac{\|[\tilde{\phi}(z) - \phi^0(z - \Delta)]/A_\phi\|_2}{\sqrt{n_z}},$$

$$e_p = \frac{\|[\tilde{P}_c(z) - P_c^0(z - \Delta)]/A_p\|_2}{\sqrt{n_z}}. \quad (\text{B4})$$

The numerical resolution is chosen relative to the compaction length at background melt fraction so that $h_z = [1/2, 1, 2, 4]\delta_0$. The simulations of each spatial resolution are run with different time step lengths. Time steps are chosen as multiples of the courant step for the appropriate phase speed c so that $\Delta t = [1/4, 1/2, 1, 2]\delta_0/c$ with resulting courant numbers $CLF = c/h_z = [1/8, 1/4, 1/2, 1]$. The models were run long enough for the solitary wave to travel a distance at least four times greater than its wave length.

Fig. B7 displays the shape of the numerically computed solitary wave along with the compaction pressure anomaly after it had travelled approximately three times its own wave length. The analytical solution is shown for comparison (black line). Two numerical solutions are presented in this figure, one computed with a first-order accurate time evolution for melt fraction (backward finite difference in time, blue line), and another one computed with the second-order accurate time evolution used in all other computations in this study (Crank–Nicholson scheme in time, see eq. 64, red line). It is evident that the first-order accurate melt fraction evolution leads to a strongly diffused wave peak. The second-order melt fraction evolu-

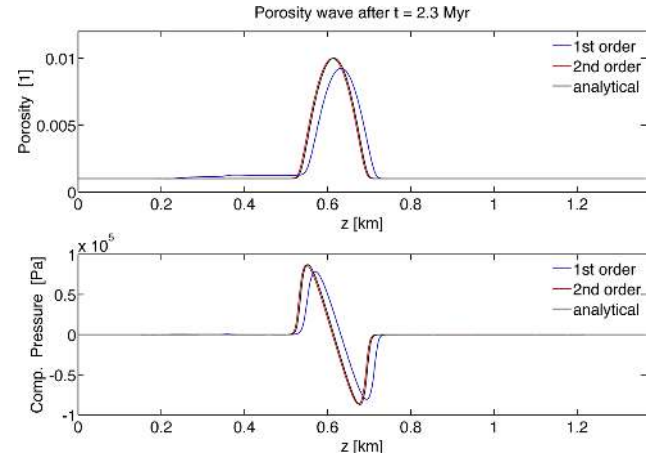


Figure B7. Melt fraction (top panel) and compaction pressure (lower panel) plotted against vertical coordinate for the solitary wave benchmark. Black lines give analytical solution for 1-D solitary wave, blue curves indicate numerical results obtained with a first-order accurate (backward finite difference) time discretization for melt fraction evolution, red curves indicate numerical results obtained with a second-order accurate (Crank–Nicholson scheme) time discretization for melt fraction evolution.

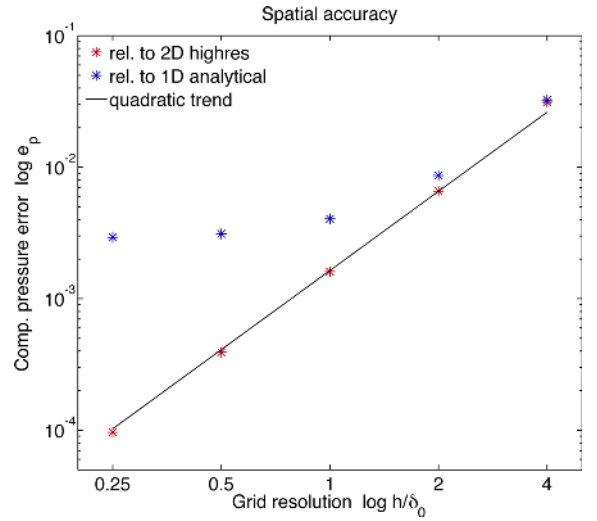


Figure B8. Results of spatial convergence test, plotting errors of compaction pressure solution relative to 1-D analytical solution (blue asterisks) and relative to 2-D high-resolution numerical solution (red asterisks), plotted against grid resolution relative to background compaction length h/δ_0 . Black line gives a quadratic convergence trend as reference.

tion, however, retains the shape of the solitary wave very well, but propagates slightly slower than the analytical solution.

Comparing the error in the initial pressure solution in runs with increasing spatial resolution reveals that the accuracy at first converges almost quadratically with increasing resolution, but then saturates at a level of around $2.e-3$ (Fig. B8). Further convergence of the numerical solution cannot be expected, as the solitary wave solution used as a benchmark is in reality the solution of slightly simplified system of equations, whereas our code solves the full equations of two-phase flow. If we compare the same error measurement to a very high-resolution numerical solution, the quadratic convergence is perfect, demonstrating the coherent functionality of the solver.

Comparing the error in the shape of the solitary wave and the compaction pressure field at a given time towards the end of each

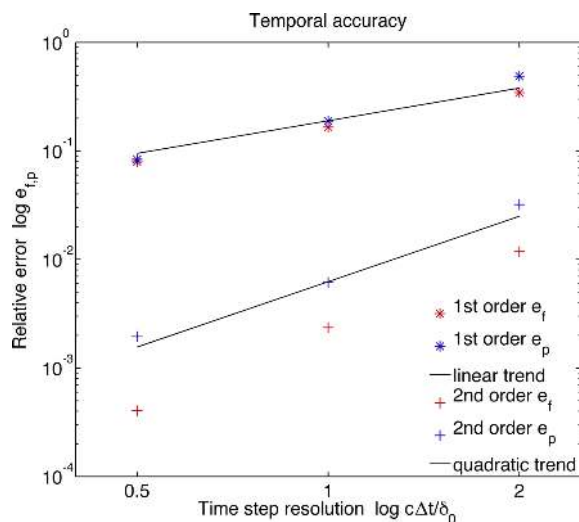


Figure B9. Results of temporal convergence test, giving errors of melt fraction and compaction pressure solution obtained with a first-order accurate time discretization (red and blue asterisks), and melt fraction and compaction pressure errors obtained with a second-order accurate time discretization, plotting against distance a solitary wave travels per time step relative to background compaction length $c\Delta t/\delta_0$. Black lines give a linear and a quadratic convergence trends as reference.

run, we find that the time-dependent solution converges quadratically with smaller time step size runs employing the second-order melt fraction evolution scheme, and linearly for runs employing the first-order melt fraction evolution scheme (Fig. B9).

The results of this suite of benchmarks sufficiently demonstrate the functionality and accuracy of the code developed in this study. Moreover, it is evidently advisable to only use second-order melt fraction evolution (eq. 64), as the time-dependent error of melt fraction propagation would be considerable otherwise.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Movie S1. Animation of melt fraction and melt velocity arrows of run d118r2 (melt diapirism). Colours and scaling the same as in Fig. 4.

Movie S2. Animation of volumetric strain rate and solid velocity arrows of run d118r2 (melt diapirism). Colours and scaling the same as in Fig. 4.

Movie S3. Animation of melt fraction and melt velocity arrows of run d121r4 (decompaction channeling). Colours and scaling the same as in Fig. 6.

Movie S4. Animation of volumetric strain rate and solid velocity arrows of run d121r4 (decompaction channeling). Colours and scaling the same as in Fig. 6.

Movie S5. Animation of melt fraction and melt velocity arrows of run d122r4 (tensile fracturing). Colours and scaling the same as in Fig. 8.

Movie S6. Animation of volumetric strain rate and solid velocity arrows of run d122r4 (tensile fracturing). Colours and scaling the same as in Fig. 8.

Movie S7. Animation of melt fraction and melt velocity arrows of melt extraction through a thermally younger lithosphere. Colours and scaling the same as in Fig. 11.

Movie S8. Animation of shear strain rate and solid velocity arrows of melt extraction through a thermally younger lithosphere. Colours and scaling the same as in Fig. 11.

Movie S9. Animation of melt fraction and melt velocity arrows of melt extraction through a thermally older lithosphere. Colours and scaling the same as in Fig. 12.

Movie S10. Animation of shear strain rate and solid velocity arrows of melt extraction through a thermally older lithosphere. Colours and scaling the same as in Fig. 12 (<http://gji.oxfordjournals.org/lookup/suppl/doi:10.1093/gji/ggt306/-/DC1>).

Please note: Oxford University Press are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.