Geophys. J. Int. (2005) **163,** 788–800 doi: 10.1111/j.1365-246X.2005.02742.x

Towards a self-consistent plate mantle model that includes elasticity: simple benchmarks and application to basic modes of convection

Hans-Bernd Muhlhaus[1,2] and Klaus Regenauer-Lieb[2]

[1]*ESSCC, The University of Queensland, St Lucia, QLD* 4072, *Australia. E-mail: muhlhaus@esscc.uq.edu.au*
[2]*CSIRO Division of Exploration and Mining,* 26 *Dick Perry Ave., Kensington WA* 6051, *Australia*

Accepted 2005 July 21. Received 2005 July 12; in original form 2005 January 25

SUMMARY

One of the difficulties with self-consistent plate-mantle models capturing multiple physical features, such as elasticity, non-Newtonian flow properties and temperature dependence, is that the individual behaviours cannot be considered in isolation. For instance, if a viscous mantle convection model is generalized naively to include hypo-elasticity, then problems based on Earth-like Rayleigh numbers exhibit almost insurmountable numerical stability issues due to spurious softening associated with the co-rotational stress terms. If a stress limiter is introduced in the form of a power-law rheology or yield criterion these difficulties can be avoided. In this paper, a novel Eulerian finite element formulation for viscoelastic convection is presented and the implementation of the co-rotational stress terms is addressed. The salient dimensionless numbers of viscoelastic–plastic flows such as Weissenberg, Deborah and Bingham numbers are discussed in a separate section in the context of Geodynamics. We present an Eulerian formulation for slow temperature-dependent, viscoelastic–plastic flows. A consistent tangent (incremental) formulation of the governing equations is derived. Numerical and analytical solutions demonstrating the effect of viscoelasticity, co-rotational terms are first discussed for simplified benchmark problems. For flow around cylinders we identify parameter ranges of predominantly viscous and viscoplastic and transient behaviour. The influence of locally high strain rates on the importance of elasticity and non-Newtonian effects is also discussed in this context. For the case of simple shear we investigate in detail the effect of different co-rotational stress rates and the effect of power-law creep. The results show that the effect of the co-rotational terms is insignificant if realistic stress levels are considered (e.g. deviatoric invariant smaller than 1/10 of the shear modulus say). We also consider the basic convection modes of stagnant lid, episodic resurfacing and mobile lid convection as applicable to a cooling planet. The simulations show that elasticity does not have a significant effect on global parameters such as the Nusselt number and the qualitative nature of the basic convection pattern. Our simple benchmarks show, however, also that elasticity plays a significant role for instabilities on the local scale of an individual subduction zone.

Key words: Visco-elasticity, plasticity, convection, Deborah number, co-rotational.

INTRODUCTION

Realistic simulations of earth processes such as faulting, shearing, magma flow, subduction and convection often require the consideration of non-Newtonian effects such as elasticity and power-law creep. Non-Newtonian effects combined in a single constitutive relationship allow, at least in principle, the modelling of key aspects of planetary behaviour, such as mantle convection, the emergence of an elastic or elasto-plastic boundary layer and even plate tectonics in a unified manner. Such models are unified in the sense that different modes of mechanical behaviour are represented as coincident features of a single physical model. Within a planet, the different convection modes are controlled by the usually strong temperature and pressure dependence of the coefficients of the viscous part of the model. The existing models of mantle convection with emergent plate tectonics, for example,], often concentrate on two key ingredients, namely the exponential dependence on the temperature of the viscosity (Arrhenius law) and a stress limiter in the form of a fracture, yield or damage criterion for example, Bercovici (1993), Solomatov (1995), Moresi & Solomatov (1998), Trompert & Hansen (1998), Ogawa (2003) and Tackley (1998, 2000a). There are but few papers in which the influence of elasticity is considered. In the past, strongly viscoelastic convection simulations with a lithospheric component have been limited to models with explicit layering in which a non-convecting viscoelastic layer is coupled to a viscous, convecting substratum (Poliakov *et al.* 1996). Viscoelastic

mantle convection has been limited to considering constant viscosity (Harder 1991). Harder's derivations are based on an upper convected Maxwell model that explains many of the difficulties he has struggled with: Maxwell models exhibit spurious strain softening and instabilities at elevated stresses. These 'side effects' recede significantly as soon as a yield criterion or a stress limiter such as power-law creep is considered. These issues will be discussed in the following sections. Models of subduction zones that incorporate viscoelasticity, faulting, and free surface behaviour have generally been limited to modest evolution times, after which further deformation produces severe remeshing problems (Melosh 1978; Gurnis *et al.* 1996), or separate the computational domains into a solid-mechanical domain where elasticity is considered (cold slab) and a fluid-dynamic domain (outer boundary of the slab) where elasticity is neglected (Regenauer-Lieb *et al.* 2001; Funiciello *et al.* 2003; Morra *et al.* 2004). However, these approaches cannot produce self-consistent plate tectonics emerging out of mantle convection. Although significant progress has been achieved by acknowledging non-Newtonian effects in fluid mechanical approaches, achieving basic plate-like features, spreading ridges and Earth-like toroidal/poloidal ratios (Tackley 2000a,b), it has been impossible to reproduce terrestrial subduction zones. Subduction zones appear as vertical two-sided downwellings. Comparatively, viscoelasto–plastic calculations produce asymmetric subduction zones (Regenauer-Lieb *et al.* 2001) suggesting that elasticity is required to produce consistent models. Recent results (Regenauer-Lieb & Yuen 2004) have shown that symmetry breaking in faulting arises as a result of elastic stored energy in a dissipative viscoelastic–plastic system.

In this context, an approach has been put forward by Moresi *et al.* (2002) where a novel finite element based on a moving integration point scheme is applied to the simulation of mantle convection based on temperature-dependent viscosity, Maxwell-type viscoelasticity and a pressure-dependant yield criterion. The latter acts as a stress limiter thus avoiding spurious softening effects otherwise associated with viscoelasticity. However the main purpose of Moresi *et al.*'s (2002) paper was the illustration of the applicability of the model and the Lagrangian integration point code in terms of a range of simple benchmark problems including a specific mantle convection study.

This paper presents an alternative and more general fully Eulerian model applicable to a wide range of existing fluid dynamic models. The approach considers combined Newtonian and power-law creep as well as elasticity and temperature dependence of the creep parameters. As the deformations involved in geological deformation are large the constitutive relationships must contain geometric terms ensuring that the tensor properties of the model are conserved. A model with such properties is described as being 'objective'. There exist a wide range of objective, viscoelastic–plastic models to choose from. The primary difference between these models consists in the choice of the objective stress rate, such as Jaumann, Oldroyd or Truesdell rates (see Kolymbas & Herle 2003, for a recent discussion). In the following we give an outline of the constitutive model, derive an incremental form for the constitutive operator and compare different models of viscoelasticity. The salient features of incompressible flows can be studied in homogeneous simple and pure shear flows, respectively (Appendix A1). For simple shear flow we compare the shear stress–shear strain curves for a constant applied shear strain rate, assuming an infinitesimal deformation model (i.e. no co-rotational stress terms), Jaumann and Naghdi models, respectively. Subsequently we consider the problem of incompressible flow around circular cylinders and highlight parameter ranges governed by viscous, viscoplastic and viscoelastic–plastic behaviour. Finally we explore the behaviour of the model in a study on 2-D (plane strain) natural convection in one by one and two by one domains for various ratios of the effective relaxation time to the thermal diffusion time (Deborah number), power-law exponents and Rayleigh numbers.

CONSTITUTIVE MODELS

We use Cartesian coordinates x_i with $i = 1, 2, 3$; v_i are the components of the velocity vector, partial differentiation with respect to x_j or the time t is written as a subscribed comma followed by the index of the coordinate with respect to which we differentiate; for example, the velocity gradient L_{ij} is defined as $L_{ij} = v_{i,j}$ and accordingly the partial time derivative of the velocity is written as $v_{i,t}$; σ_{ij} are the components of the stress tensor, δ_{ij} is the Kronecker unit tensor, T is the temperature. The stretching D_{ij} is defined as usual as the symmetric part of the velocity gradient L_{ij} and the spin $W_{ij} = -W_{ji}$ of an infinitesimal element of the continuum is given by the anti-symmetric part of L_{ij}.

In the formulation of the constitutive model we make the standard assumption that the stretching is the sum of an elastic, viscous and a plastic part, that is,

$$D_{ij} = D^e_{ij} + D^v_{ij} + D^P_{ij}. \tag{1}$$

For the elastic part we assume hypo-elasticity [see e.g. Prager 1961 for the definition of elastic, hypo-elastic (constitutive equations are in rate form; linear in the stretching) and hyper-elastic materials], for the viscous part we assume that the viscosity is made up of a Newtonian and a power-law contribution (to be specified below) and for the plastic deformation we assume von Mises plasticity combined with the usual Prandtl–Reuss flow rule (Hill 1998); thus

$$D_{ij} = \frac{1}{2\mu}\mathring{\tilde{\sigma}}'_{ij} + \frac{1}{2\eta}\sigma'_{ij} + \dot{\gamma}^P\frac{\sigma'_{ij}}{2\tau}, \qquad D_{kk} = div\mathbf{v} = 0, \tag{2}$$

where the prime designates the deviator of each tensor; for example:

$$\sigma'_{ij} = \sigma_{ij} - \frac{1}{3}\sigma_{kk}\delta_{ij}. \tag{3}$$

In the formulation of the equations we use index notation and adopt Einstein's summation convention, that is, summation over equal indexes. The pressure is defined as usual as $p = -1/3\ trace\ \sigma_{ij}$. Incompressibility is also assumed since we are mainly interested in large deformations. The significance of the objective or co-rotational stress rate $\mathring{\tilde{\sigma}}_{ij}$ will be discussed in a separate section below. In eq. (2) the quantity τ is the second deviatoric stress invariant defined as $\tau = \sqrt{1/2\sigma'_{ij}\sigma'_{ij}}$ and for later use we define the equivalent strain rate as $\dot{\gamma} = \sqrt{2D_{ij}D_{ij}}$; μ and η are the shear modulus and the shear viscosity, respectively. The shear viscosity is assumed in the form $1/\eta = 1/\eta_N + 1/\eta_P$, where $1/\eta_N$ is the Newtonian viscosity and η_P is the secant viscosity (instantaneous viscosity derived from the secant in the stress–strain rate plot) of the power-law creep component. The power-law viscosity depends on the second deviatoric stress invariant and in addition, both viscosities may depend on the pressure and the temperature. Specific forms will be assumed below. Here we adopt the usual practice in the earth science and assume that the shear modulus μ is a constant. The coefficient of the last term in eq. (2) is the plastic multiplier $\dot{\gamma}^P$. In the finite element simulations we shall assume a von Mises yield criterion with a constant yield stress τ_Y, that is, $\tau \leq \tau_Y$. In classical rate-independent plasticity, $\dot{\gamma}^P$ is determined from the so-called consistency condition. Here we adopt a simple

alternative to the rate-independent plasticity approach: we assume an evolution equation for $\dot{\gamma}^P$ in the form $\dot{\gamma}^P = \tau_Y/\eta_Y(\tau/\tau_Y)^{n_{pl}}$, where η_Y is a reference viscosity and $n_{pl} \gg 1$ is the power-law coefficient of the plastic deformation. This approach is computationally simpler than the conventional rate-independent plasticity approach. In fact for constant yield stress and $n_{pl} \to \infty$ the incremental constitutive equations of the power-law approach and the plasticity approach coincide up to one detail: in the plasticity approach the consistency condition is satisfied by the material stress derivative whereas in the power-law case it is the spatial stress derivative satisfies the consistency condition. See Appendix A2 for details.

The evolution equation for the equivalent strain rate for the power-law part of the stretching we notice that a vast number of experimental data of the flow properties by dislocation glide have been fitted to power-law relationships of the type

$$\dot{\gamma}^{visP} = \dot{\gamma}_0 \exp(-AT_M/T)\left(\frac{\tau}{\tau_0}\right)^n, \tag{4}$$

where $\dot{\gamma}_0$ and τ_0 are a reference strain rate and stress, respectively; A is the activation energy and T_M is the melting temperature. In general, for dislocation creep or diffusion creep combined with stress dependence of the constitutive parameters such as grain size, n varies between 2 and 5. See Kirby (1983), and Karato & Wu (1993) for further details. A change in n of about unity is reported by some investigators at stresses less than $0.5\,10^7$–$1.0\ 10^7$ Pa. But overall the value $n = 3$ seems to be a reasonable approximation for most dislocation related geodynamic flows. The lower mantle seems to be governed by diffusion type creep with linear stress dependence of the stretching so that $n = 1$ in this case. We, therefore, adopt a combined rheology consisting of Newtonian and power-law creep contributions (see also discussion below eq. (7). Stresses associated with convecting flow ranges from 10^6 Pa in the asthenosphere to larger than 10^8 Pa in subducting plates. The high-stress regimes are governed largely by rate-independent or almost-rate-independent material behaviour, elasticity and brittle or plastic behaviour. In the present model the latter is described either by classic von Mises plasticity or by the power-law evolution for the equivalent plastic strain rate as discussed below eq. (3).

Some investigators prefer the approximation

$$\dot{\gamma}^{visP} = \dot{\gamma}_0 \exp(AT/T_M)\left(\frac{\tau}{\tau_0}\right)^n, \tag{5}$$

for small deviations of the temperature from the reference temperature (in this case T_M). However, as pointed out by Leroy & Molinari (1992), thermal runaway instabilities exist mathematically (Gruntfest 1963) for laws of type (5) with $n = 1$, but don't exist if the original Arrhenius relationship (4) is assumed.

We mentioned before that the Newtonian viscosity η_N depends on the temperature as well; for simplicity we assume the same style of temperature dependence as in (4), namely:

$$\eta_N = \eta_{N0} \exp(AT_M/T). \tag{6}$$

In the following applications we use a combined Newtonian and power-law viscosity. We assume that $\eta_{N0} = \tau_0/\dot{\gamma}_0$. In this case the effective viscosity is obtained as:

$$\frac{1}{\eta} = \frac{1}{\eta_N} + \frac{1}{\eta_N \xi_p} \quad \text{where} \quad \xi_p = (\tau/\tau_0)^{1-n}. \tag{7}$$

The stress parameter τ_0 has the significance of a transition stress: The flow is predominantly Newtonian for $\tau < \tau_0$ and predominantly power law for $\tau > 0$.

Computer simulations of problems involving non-Newtonian constitutive equations are often based on a scalar effective viscosity which, depends on the values of the state variables from the preceding time step. Such a procedure usually requires many iterations in each time step since the dependency should actually be on the values of the variables at the present time step. Geological problems are usually highly indefinite, that is, are not uniquely soluble and effective viscosity based approaches may favour continuations of the pre-bifurcation path. We, therefore, propose to base the computational model on a tangent form of the constitutive equations. This approach requires few if any iterations per step. For the tangent method the incremental form of our constitutive model is needed. Expansion of eq. (2) about the stress, temperature and pressure at time t and neglecting terms of order δt^3 yields in the rate-independent limit, $\eta_N \to \infty n\, n_{pl} \to \infty$ (the general incremental form as well as details of the derivation are represented in Appendix A2):

$$\partial\sigma'_{ij}/\partial t = \left(\mu(\delta_{ik}\delta_{jl} + \delta_{jk}\delta_{il}) - \mu\frac{\sigma'_{ij}}{\tau}\frac{\sigma'_{kl}}{\tau}\right)\left(D_{kl} - \frac{1}{2\mu}\sigma'_{ij,k}v_k\right) + (W_{ik}\sigma_{kj} - \sigma_{ik}W_{kj}). \tag{8}$$

In the above limit case the incremental form of the power-law plasticity model (8) and the corresponding conventional plasticity model are very similar. The only difference is that in conventional plasticity the consistency condition for continuous yielding is satisfied by the material stress rate whereas in the present power-law plasticity model the condition is satisfied by the spatial stress rate. Hence in the conventional plasticity model the spatial stress derivative on the left hand side of eq. (8) is replaced by the material rate (which includes stress advection) and the stress advection term on the right hand side of eq. (8) behind the stretching is not present. In the present model we assume that the yield stress is constant. A more general model will be presented in a forthcoming paper. We conclude this section with a comment on the iterative method used for example, by Moresi & Solomatov (1998) and Trompert & Hansen (1998) in connection with viscoplastic models. In this approach the viscosity at time $t + \delta t$ is defined as $\min(\eta, \tau_Y/\dot{\gamma}^{t+\delta t})$. Since the strain rate at $t + \delta t$ is not known at time t the problem has to be solved iteratively according to $\min(\eta, \tau_Y/\dot{\gamma}_\alpha^{t+\delta t})$, where $\alpha = 0, 1, 2, \ldots$, is an iteration counter and $\dot{\gamma}_0^{t+\delta t} = \dot{\gamma}^t$. A formulation, which is consistent with the incremental approach proposed here is obtained if $\dot{\gamma}^{t+\delta t}$ is replaced by $\dot{\gamma}^{t+\delta t} = \dot{\gamma}^t + \delta\dot{\gamma} = \dot{\gamma}^t + 2D^t_{kl}\delta D_{kl}/\dot{\gamma}^t$.

OBJECTIVE STRESS RATES

We consider a deforming continuum and assume that an infinitesimal neighbourhood of a spatial point, x_i say deforms momentarily like a rigid body. In this case the stretching vanishes and $L_{ij} = W_{ij}$. For momentarily rigid behaviour the stress rate is made up of an infinitesimal rigid rotation plus the usual contribution due to stress relaxation. Had we assumed $\mathring{\sigma}_{ij} = \dot{\sigma}_{ij}$ then the stress rotation would be neglected. If the relaxation time is infinite (purely elastic behaviour) then the stress rate is zero, which is obviously not correct.

What are physically meaningful choices for $\mathring{\sigma}_{ij}$? There are in fact infinitely many choices; the main requirement is that the definition contains the expression for an infinitesimal rotation of the stress tensor. Hence the simplest definition of the rate $\mathring{\sigma}_{ij}$, the so-called Jaumann stress, which we designate as, $\dot{\sigma}^J_{ij}$ reads:

$$\dot{\sigma}^J_{ij} = \dot{\sigma}_{ij} - W_{ik}\sigma_{kj} + \sigma_{ik}W_{kj}. \tag{9}$$

The Jaumann rate is the stress measured by an observer co-rotating with the spin W_{ij} of the infinitesimal volume element dV centred at x_i. Another popular choice in connection with the Maxwell model of viscoelasticity is the Oldroyd rate:

$$\dot{\sigma}_{ij}^{O} = \dot{\sigma}_{ij} - L_{ik}\sigma_{kj} - \sigma_{ik}L_{jk}. \tag{10}$$

We notice that the Jaumann rate is equal to the Oldroyd rate if $D_{ij} = 0$. The trace of the geometric terms $-W_{ik}\sigma_{kj} + \sigma_{ik}W_{kj}$ in the definition (9) of the Jaumann vanishes, that is, the geometric or co-rotational terms do not contribute to the pressure; there is another convenient property of the Jaumann rate in connection with the determination of the plastic multiplier from the von Mises yield criterion $\tau \leq \tau_Y$, where τ_Y is the yield stress. We have $\dot{\tau} = \sigma_{ij}\dot{\sigma}_{ij}/2\tau = \sigma_{ij}\dot{\sigma}_{ij}^{J}/2\tau$. Similar properties do not exist in general for the Oldroyd definition (10) of the stress rate. The definitions (9) and (10) differ by terms of the form $D_{ik}\sigma_{kj} + \sigma_{ik}D_{jk}$. Only terms of the type $-W_{ik}\sigma_{kj} + \sigma_{ik}W_{kj}$ as in (9) are necessary for the objectivity of the constitutive model. Depending on experimental evidence or micromechanical justification, terms like $D_{ik}\sigma_{kj} + \sigma_{ik}D_{jk}$ may or may not be present. The co-rotational terms in (9) depend only on the deviatoric stress (the pressure terms cancel because of the skew symmetry of W_{ij}) and the trace of these terms vanishes, that is, no contribution to the rate of volume change, which makes sense in the context of incompressible materials. In the definition (10) is neither pressure invariant nor does the trace of the co-rotational terms vanish. In connection with incompressible materials the following modification of the definition (10) seems natural: in (10) replace the stresses by the deviatoric stresses and of the resulting expression consider only the deviatoric part in the definition of the elastic strain rates. The modified version is invariant with respect to pressure, is deviatoric and coincides in important special cases such as general plane strain (not just simple shear) with the Jaumann definition (9). The co-rotational terms and in particular the convective stress terms complicate the implementation of the constitutive model considerably. How important are those terms?

The stress convection term is of the order stress times strain rate divided by shear modulus. Since the stress is limited by the yield stress, which is a small fraction of the shear modulus we expect—except if the stress gradients are high—the convection term to be of the order of a small fraction of the order of magnitude of the leading term, the stretching. High stress gradients occur at interfaces between hard and soft materials and geometric instabilities such as folding and buckling (see end of section on shear flows). The Jaumann terms (last term on the right hand side of eq. 8) usually are of a similar small order, however they are of crucial importance in special cases such as internal buckling (folding), surface instabilities and kinking of anisotropic materials. A wide variety of internal instability problems in anisotropic materials are discussed and solved in Biot's (1965) book on the mechanics of incremental deformations. (see also Muhlhaus 1985 for a study on buckling of layered materials with bending stiffness).

The Naghdi stress rate (Kolymbas & Herle 2003) is very similar to the Jaumann rate (9). The only difference is that the spin in the Naghdi definition is not equal to the non-symmetric part of the velocity gradient but equal to the spin of the rotation tensor of the polar decomposition of the deformation gradient. The gradient of the spatial coordinates of a material point with respect to the coordinates of the material point in the reference configuration is called the deformation gradient. The deformation gradient can always be decomposed in a multiplicative fashion into an orthogonal rotation tensor and a symmetric stretch tensor; see, for example, Malvern (1969) for details.

The Naghdi rate has many appealing properties (see discussion in the next section). The main disadvantage is that in a numerical context the computation of the rotation tensor requires much more operations than are required for the Jaumann rate. In the section to come we explore which numerical implementation of objective, co-rotational stresses is ideally suited for the extremely large-scale deformation incurred by mantle convection. Such an application lies clearly beyond the traditional engineering benchmarks. In order to do this we investigate the essential features of non-Newtonian flow as described by (2) for the case of simple shear.

SIGNIFICANCE OF ELASTICITY AND PLASTICITY ON DIFFERENT TIME AND LENGTH SCALES

Non-Newtonian fluids have a characteristic time scale, the Maxwell time λ. In a flow with a characteristic shear rate $\dot{\gamma}_{ch}$, or a characteristic time T_{ch} (e.g. the service time of an engineering structure or a time of interest) two dimensionless groups can be formed: The Deborah number $De = \lambda/T_{ch}$ and the Weissenberg number $Wei = \lambda\dot{\gamma}_{ch}$. The Deborah number represents the transient nature of the flow relative to the Maxwell time whereas the Weissenberg number compares the elastic forces to the viscous effects; a more detailed outline including a discussion of the so-called Pipkin–Tanner diagram in which the horizontal axis scales with the Deborah number and the vertical axis with the Weissenberg number can be found in Phan-Thien (2002). In the Pipkin diagram different domains are described by different constitutive relationship, for example, $Wei = 0$, $De = 0$ corresponds to Newtonian flow; $De = 0$, $Wei > 0$ corresponds to linear viscometric flows; $Wei = 0$, $De > 0$ corresponds to linear viscoelasticity. We investigate the more general case where both Wei and De are nonzero. For such cases marked non-linear behaviour must be considered Phan-Thien (2002).

For mantle scale processes the mean relaxation time is of the order $10^{22}\,\mathrm{Pa\,s}/(4 \cdot 10^{10}\,\mathrm{Pa}) = 2.5 \cdot 10^{11}$ s and $\dot{\gamma} = 10^{-14}$ s so that $\lambda\dot{\gamma} \approx 2.5 \cdot 10^{-3}$. In the estimates we have assumed a typical lithospheric value for the shear modulus (Turcotte & Schubert 2002). If one is interested only in the global characteristics of the convective flow such as the Nusselt number then the effect of elasticity and Jaumann terms can safely be neglected. Larger values of $\lambda\dot{\gamma}$ may occur in subducting plates, where the relaxation time is of the order of 10^{14} s or more or in mechanically unstable geological structures such as folding and buckling of layered rock Schmalholz & Podladchikov (1999). Lithospheric deformations where elasticity is important such as elastic bending of the lithosphere under islands chains and at an ocean trench are discussed in Turcotte & Schubert (2002).

In a study on slab subduction Funiciello *et al.* (2003) define a characteristic time as the ratio between the deforming area in arc length of the bent arc (10–20 km) of the slab from the cold boundary into the mantle and the average plate speed (3×10^{-9} m s^{-1}). This leads to characteristic time scales between $10^5 - 2 \times 10^5$ yr ($(3.3–6.6) \times 10^{12}$ s). In the linear viscoelastic model realistic slab shapes were achieved with viscosities that are ranging between $\eta = 10^{23}$–10^{24} Pa s. Taking a shear modulus of $\mu = 4 \times 10^{10}$ Pa the *De* number thus ranges between 0.38 and 7.5. This is a range where in natural material non-linear behaviour may be prominent and must be considered as a possibility. A non-linear approach is characterized by a large *Wei* number. A non-linear thermally activated viscoelasto–plastic slab has also been modelled by Funiciello *et al.* (2003). Using material constants constrained by laboratory experiments an effective viscosity has been obtained that ranges between

$\eta = 10^{25} - 4.0 \times 10^{26}$ Pa s, while the corresponding characteristic strain rate varies between $\dot{\gamma}_{ch} = 3.0 \times 10^{-15}\,\mathrm{s}^{-1}$ to $\dot{\gamma}_{ch} = 1.0 \times 10^{-15}\,\mathrm{s}^{-1}$ (see viscosities and strain rates plots for Domain I in Fig. 12 of Funiciello *et al.* 2003). With a shear modulus of $\mu = 4 \cdot 10^{10}$ Pa the Weissenberg number associated with the plate bending is obtained between $Wei = 0.75$–10.00. The high *Wei* number for this particular subprocess is not reflected in the usual global parameters (e.g. the Nusselt number) of whole mantle convection where *Wei* is of the order of 10^{-3}. This is confirmed by our convection simulation results presented in the 'Natural Convection' section. Minor influences of elasticity are noticeable in Nusselt number histories only in connection with episodic lid behaviour (Fig. 11).

Elasticity is an instantaneous effect. The elastic constitutive relationships are invariant with respect to changes of the time scale. This property is sometimes called rate-independence in the literature. Another important rate - independent property of solids and fluids is plasticity. However unlike elasticity, the consideration of plasticity is of crucial importance for models producing plate-tectonics-like behaviour (Moresi & Solomatov 1998; Trompert & Hansen 1998). For instance a distinction into stagnant lid convection, mobile lid convection and episodic resurfacing if only possible is plasticity and yielding is considered. There are two dimensionless numbers associated with plasticity. The Bingham number $Bi = \tau_Y/(\eta\dot{\gamma}_{ch})$ where τ_Y is the v. Mises yield stress. The Bingham number is a measure for the ratio between plastic and viscous dissipation. Another dimensionless number associated with plasticity is the ratio between the yield stress and the shear modulus. This number is related to the Bingham and the Weissenberg number as $\tau_Y/\mu = Bi \cdot Wei$. Assuming a global Weissenberg number 2.5×10^{-3}, an average lithospheric strength of τ_Y/μ between 0.2×10^{-3} and 0.5×10^{-3} (e.g. Scholz 1990) we obtain values for the Bingham number *Bi* between 0.08 and 0.2; that is, plastic deformation will occur.

In the section following the computational formulation we demonstrate the salient features of viscoelastic flows in simple shear. We then include plasticity and consider viscoelastic–plastic flow around rigid cylinders for different values of the Weissenberg number. We shall use this case to highlight parameter ranges governed by viscous, viscoplastic and viscoelasto–plastic behaviour. The local Weissenberg number between the cylinders is much higher than the global Weissenberg number defined by the far field velocity and the cylinder radius and the Maxwell time. This strong variation between global and local flow characteristics is not unlike the situation in mantle convection where the stress in a subducting slab can be much higher than the average stress related to mantle size, average viscosity and average plate speed.

COMPUTATIONAL FORMULATION

In the next section we present the results of finite element simulations of plane strain, simple shear and natural convection problems in infinite strip in simple shear and in a rectangular L by H domain in convection, where H is the dimension in the direction of gravity. The governing equations consist of the constitutive relationships (A16), the stress equilibrium conditions

$$\sigma'_{ij,j} - p^{th}_{,i} + Ra^c T g_i = 0, \tag{11}$$

and the heat equation

$$T_{,t} + v_j T_{,j} = T_{,jj} + \frac{Di^c}{Ra^c}\sigma'_{ij} D_{ij}. \tag{12}$$

The comma followed by an index stands for partial differentiation with respect to the corresponding coordinate, that is, $a_{,i} = \partial a/\partial x_i$. In eq. (11) the unit vector g_i is parallel and opposite the direction of gravity; pth is the pressure due to convection. The parameters are explained ion the text further below. The stresses and the shear modulus μ are non-dimensionalized with respect to $\eta * /t_D$, where

$$t_D = \frac{\rho_0 c_p H^2}{\kappa} \tag{13}$$

is the characteristic thermal diffusion time, η^* is equal to the pre-exponential coefficient of the Newtonian viscosity, η_{N0} times a coefficient, which depends on the way the Arrhenius relationship is transformed for computational purposes (see eq. 19 for definition); ρ_0 is the surface density c_p is the heat capacity and κ is the thermal conductivity. Typical values are $H = 700$ km for upper mantle convection and $H = 3000$ km for whole mantle convection. With $\kappa/(\rho_0 c_p) = 10^{-6}\,\mathrm{m}^2\,\mathrm{s}^{-1}$ we obtain $10^{17} < t_D < 10^{19}$ s. In eqs (11) and (12) Ra^c and Di^c designate the computational Rayleigh number and the dissipation number respectively. The Ra^c and Di^c, consistent with the way the stresses are non-dimensionalized, are defined as

$$Ra^c = \frac{\rho_0^2 c_p g \alpha \Delta T H^3}{\kappa \eta^*} \text{ and } Di^c = \frac{\alpha \rho_0 g H}{\rho_0 c_p}. \tag{14}$$

In (14) α is the thermal expansion coefficient and ΔT is the temperature difference between the hot and the cold boundary of the domain under consideration. The temperature and the velocities are non-dimensionalized with respect to ΔT and Ht_D respectively. In all simulations we assume that the shear stresses and normal velocities vanish on all boundaries of the domain; the temperatures are fixed on the top and the bottom and the normal gradient of the temperature vanish on the sides. The governing equations have been implemented into the finite element based partial differential equation solver FASTFLO using the high level language FASTALK (see http://www.cmis.csiro.au/Fastflo/ for details). In the implementation we solve sequentially the stress equilibrium equation and the heat equation. The incompressibility constraint is satisfied iteratively by means of the algorithm

$$p^{\alpha+1} = p^{\alpha} - Pen\, v^{\alpha+1}_{j,j}, \alpha = 1, 2, 3, \ldots. \tag{15}$$

where α is an iteration counter, $Pen = Pen_0 \mu_{\mathrm{eff}} \delta t$ (see Appendix A2 for the definitions of the effective shear modulus μ_{eff} and the effective viscosity η_{eff}) are penalty functions. A typical value for the constant Pen_0 is 100; in connection with direct solvers convergence is of course fastest the larger Pen_0. However, there are usually limits to the value of Pen_0 in connection with iterative solvers.

After the pressure iteration the stresses are calculated by solution of matrix problems for each of the two columns (in 2-D) of the stress tensor (see A14). The solution of the stress problem is only necessary if stress advection is considered. If stress advection is neglected or not necessary (e.g. if elastic effects are not considered) then stress updates can be evaluated at the element level. For consistency the order of interpolation is one order less than the one for the velocities. For the velocities and the temperature we use six-noded triangular elements with bi-quadratic shape functions in connection with an unstructured mesh. For the stresses we use constant strain triangles and evaluate the values of the stresses at the mid-side nodes after the solution of the stress equations. Subsequently the heat equation is solved using backward Euler time differencing.

The stress equations and the heat equation involve advective derivatives of the stresses and the temperature respectively. We use a basic upwinding scheme (Zienkiewicz & Taylor 2000, p. 30) to avoid spurious oscillations of the fields in advection dominated regimes. The common procedure to implement the standard streamline upwind Petrov–Galerkin (SUPG) formulation is to modify the test

or weight functions used in the formulation of the finite element method. However, in the present case it is more convenient to modify the differential equation: the function a is a scalar (e.g. temperature), a vector (normal vector on anisotropy surfaces) or a tensor (stress tensor). In our formulation we replace the material time derivative of a as follows:

$$a_{,t} + v_j a_{,j} \leftarrow a_{,t} + v_j a_{,j} - \left((a_{,t} + v_j a_{,j}) \frac{h}{2\sqrt{v_k v_k}} v_i \right)_{,i}. \qquad (16)$$

For pure advection problems, this approach is equivalent to the SUPG method. For unstructured meshes the discretization length scale h and the time step δt is determined by a Courant-like condition

$$h = \sqrt{\frac{\text{area}}{n_{\text{elem}}/2}} \text{ and } \delta t = C \frac{h}{v_{\max}}. \qquad (17)$$

The argument of the square root is a characteristic discretization length. Note that we are using unstructured meshes based on triangular elements; $v_{\max}$ is the maximum component of the magnitudes of the nodal point velocities. In eqs (16), and (17), C is a Courant-number-like quantity, which in explicit algorithms is put equal to 1/2 to ensure stability of the numerical solution in connection with regular grids. In the present calculations the Courant condition is not needed for numerical stability since we are using a fully implicit integration scheme but as a means to control the accuracy of the transient solution of the non-linear equations. C should be chosen in such a way that δt is always a fraction of the interest time (such as the Maxwell time for instance, when appropriate) of the problem.

For the definition of an effective Rayleigh number we require a measure for the effective viscosity. We use the following definition:

$$\bar{\eta} = \frac{\int_V \tau dV}{\int_V \frac{\tau}{\eta} dV}. \qquad (18)$$

Heat transport in convection simulations is usually advection dominated; hence the advection—and upwinding terms are crucial for a successful simulation. Since the elastic strain rate depends on the material stress rate, advective stress derivatives appear in the stress calculation as well. However, since the stresses are limited by the yield stress, which usually amounts only a small fraction of the shear modulus, the stress advection terms are not relevant in rock-like materials. If plastic deformations occur it is important that the stress is mapped exactly onto the yield surface which is in our case given by the second invariant of the deviatoric stress. Here we apply a simple but effective radial return scheme. The scheme works as follows: First the velocities are calculated using the tangent formulation described in the previous section (General case: eq. A16). Subsequently the stresses are calculated assuming viscoelastic behaviour everywhere.

In the convection study in the next section we ignore the pressure dependence of T_M in the Arrhenius relation. The main emphasis in the study will be on the role of elasticity, power-law creep and plastic yielding on the emergence of different convection styles. In the dimensionless formulation we write the Arrhenius relationship [eq. 6 as follows (see Tackley 2000a)]

$$\eta_{N0} \exp(A T_M / T) \rightarrow \eta_{N0} \exp\left(\frac{2\hat{A}}{3}\right) \times \exp\left(\hat{A}\left(\frac{1}{1+T} - \frac{2}{3}\right)\right) = \eta^* \exp\left(\hat{A}\left(\frac{1}{1+T} - \frac{2}{3}\right)\right) \qquad (19)$$

The exponent in (19) varies between 0 for $T = 0.5$ and $-\hat{A}/6$ and $\hat{A}/3$ for $T = 1$ and $T = 0$, respectively. For $\hat{A} = 23$ this corresponds to a Newtonian viscosity contrast of about 10^5 across the convection cell. In the absence of convection, the Newtonian viscosity varies slowly due to temperature change in the lower half of the cell, from 1 in the middle to 0.022 on the bottom and rapidly in the upper half from 1 to 2087 on the top. For temperature-dependent viscosities based on the approximation (5) the situation is slightly different: the viscosity variation is distributed more uniformly across the layer.

SHEAR FLOWS

Simple shear

In this section we illustrate the significance of elasticity when combined with viscous behaviour. The importance of certain geometric nonlinearities associated with elasticity is also discussed. We consider plane simple shear in the (x_1, x_2) plane. The shear is assumed parallel to x_1. The shear layer has width 1; the velocities at $x = 0$ are zero and the velocity at $x_2 = 1$ is $\dot{\gamma} = \text{const.}$, where γ is the shear strain (top displacement divided by layer thickness) of the layer. Details of the results presented are included in the Appendix (A1) for easy reference. The constitutive relations for D_{22} and D_{12} are obtained as

$$\frac{1}{2\mu}(\dot{\sigma}'_{22} + \dot{\gamma}\sigma_{12}) + \frac{1}{2\eta}\sigma'_{22} = 0,$$
$$\frac{1}{2\mu}(\dot{\sigma}_{12} - \dot{\gamma}\sigma'_{22}) + \frac{1}{2\eta}\sigma'_{12} = \dot{\gamma}/2. \qquad (20)$$

The terms $\dot{\gamma}\sigma'_{22}$ and $\dot{\gamma}\sigma_{12}$ are the co-rotational terms from the definition (9) of the Jaumann stress rate; the viscosity η is defined by (7). First we point out a number of properties of the steady state solution (i.e. zero stress rates). By multiplying the first eq. (20) by σ_{12}, the second equation by σ'_{22}, and subtracting the second from the first equation, we obtain:

$$\sigma'_{22} = -\sigma'_{11} = -\frac{\tau^2}{\mu}. \qquad (21)$$

Combining the eqs (20) for zero stress rates yields the shear stress at steady state:

$$\sigma_{12} = \frac{\eta\dot{\gamma}}{1 + (\lambda\dot{\gamma})^2}, \qquad (22)$$

where $\lambda = \eta/\mu$ is the local relaxation time. The quadratic rate term in the denominator of (22) is derives from the Jaumann terms in the constitutive model (2). For purely Newtonian flow the shear stress has a maximum at $\dot{\gamma} = \lambda^{-1} = \mu/\eta_N$ with strain-rate softening for $\dot{\gamma} > \lambda^{-1}$. We illustrate the various non-Newtonian effects in Figs 1 and 2 by means of results for the relative extreme case $Wei = (\eta_{N0}/\mu)\dot{\gamma} = 1$ (e.g. $(10^{25}\,\text{Pa s}/10^{11}\,\text{Pa})10^{-14}\,\text{s}^{-1}$. We compare the stress response for $\dot{\gamma} = \text{const}$ with and without Jaumann terms. Also shown for comparison is the response for a Maxwell model based on Naghdi's definition; see (Braun 1994; Kolymbas & Herle 2003) of the co-rotational rate.

The responses of the infinitesimal model (no co-rotational terms) and the Naghdi definition (—the spin is $\omega = -\dot{\gamma}/2/(1 + (\gamma/2)^2)$, whereas in the Jaumann model $\omega = -\dot{\gamma}/2$—) are qualitatively similar; the Jaumann model however displays the spurious softening behaviour as discussed below eq. (10). There is no experimental evidence for this kind of purely geometric softening in rocks and metals. Insofar the softening behaviour of the Jaumann model represents an unwanted side effect. However does this mean we have to abandon the Jaumann model, which is computationally much

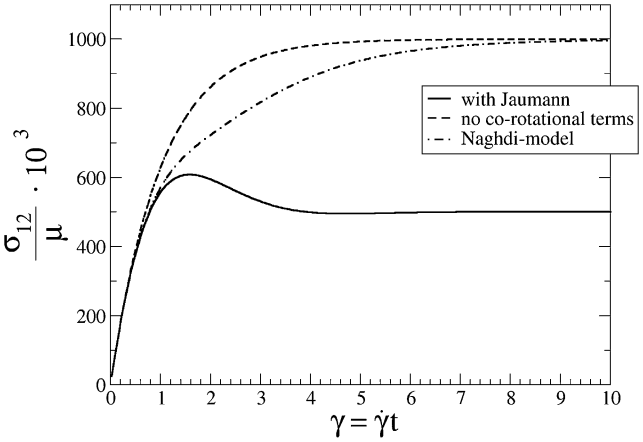

Figure 1. Simple shear of Maxwell model with Newtonian (linear) rheology . Weissenberg number $Wei = (\eta_{N0}/\mu)\dot{\gamma} = 1$.

more efficient and simpler to implement than the Naghdi model for instance?

Fig. 2 shows the results for a combined Newtonian and Power law rheology with $n = 3$ assuming the same dimensionless strain rate as in the examples displayed in Fig. 1. In this case the spurious softening behaviour disappears as do most of the differences between the three models. For combined Newtonian and power-law creep the argument of the term $(\lambda\dot{\gamma})^2$ in the denominator of (22) reads: $\lambda\dot{\gamma} = (\eta_N/\mu)\dot{\gamma}/(1 + (\tau/\tau_0)^2)$. To obtain an order of magnitude approximation for τ at steady state we assume pure power-law behaviour and derive the result $\tau/\tau_0 = (\tau_0/\mu)^{-1/3}$ for $\eta_N/\mu\dot{\gamma} = 1$ as was assumed in the shear examples. For $\tau_0/\mu = 10^{-3}$ as in Fig. 2 we find $(\lambda\dot{\gamma})^2 \approx 10^{-4}$ i.e. according to (22) Jaumann effects are not present in this case as confirmed by the results depicted in Fig. 2. However even for an unrealistically high transition stress of $\tau_0/\mu = 10^{-1}$ we find $(\lambda\dot{\gamma})^2 = 0.032$ meaning that co-rotational effects are barely noticeable even at unrealistically high transition stresses. We conclude that the spurious softening in simple shear cannot occur if stress limiters in the form of power-law creep or a yield criterion e.g. (Kolymbas & Herle 2003) are taken into account.

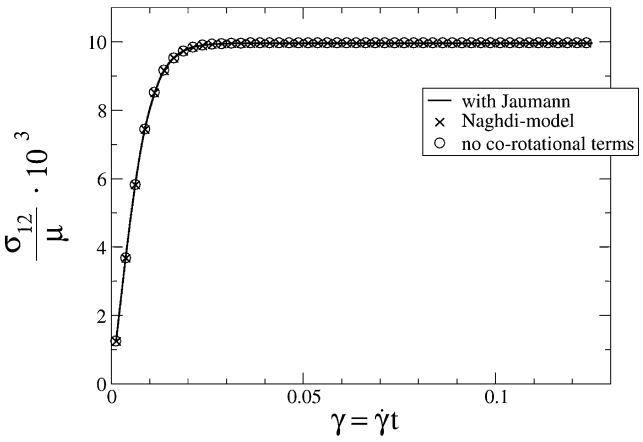

Figure 2. Simple shear of Maxwell model with combined Newtonian and power-law ($n = 3$) rheology. The dimensionless strain rate ($= Wei = (\eta_{N0}/\mu)\dot{\gamma}$) is $Wei = 1$ as in the previous case. The transition stress is $\tau_0 = 10^{-3}\ \mu$.

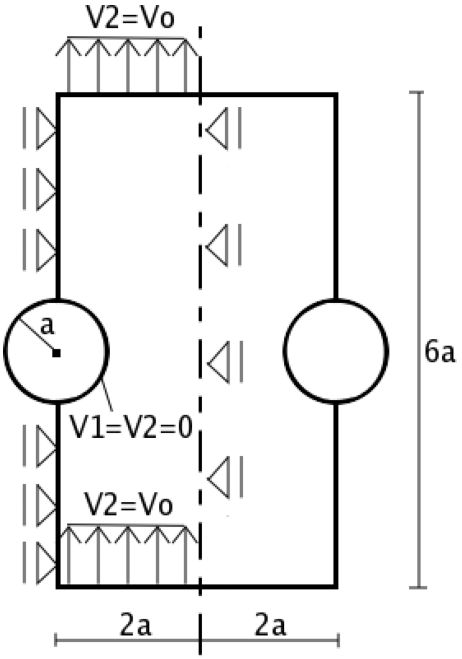

Figure 3. Flow around circular cylinders; the spacing between the centres of the cylinders is $2a$ and the radius of the cylinders is a. the spacing between the upper and lower boundaries is 6a and the vertical velocity at the upper and lower boundary is assumed as constant equal to v_0. In the present case $Wei = \eta_{N0}v_0/(\mu a)$. The computational domain is $2a$, $6a$ considering symmetries. The velocity components on the cylinder surfaces are equal to zero.

Flow around rigid circular cylinders

We consider viscoelastic–plastic flow between circular cylinders (Fig. 3). The problem is symmetric so that only one half of the domain needs to be discretised. The assumed geometry and the boundary conditions are shown in Fig. 3.

The cylinder radius is a, the spacing between cylinders is $2a$ and the model height is $6a$. The cylinder walls are assumed as rough so that the velocity components vanish on the cylinder boundaries. The horizontal velocities (v_1) as well as the shear stress vanish on the sides of the model. On the top and bottom surfaces we also assume that the shear stress vanishes and we also assume $v_2(x_1, x_2 = \pm 3a) = v_0 = \text{const}$. The mesh consists of 718 six-noded triangular elements. The material is described by a constant shear modulus μ, a constant viscosity η and yielding is described by the power-law model introduced in the Constitutive Relations section repeated here for easy reference: $\dot{\gamma}^P = \tau_Y/\eta_Y(\tau/\tau_Y)^{n_{pl}}$; In this application we use a power law coefficient of $n_{pl} = 25$. The Weissenberg number is defined as $Wei = \frac{\eta}{\mu}\frac{v_0}{a}$. In the non-dimensional form of the governing equation (the time is non-dimensionalized with respect to the Maxwell time), *Wei* is the coefficient of the deviatoric stress rate and since we wish to model close to rate-independent plasticity behaviour we also define the coefficient in the power law as $Wei = \frac{\eta_Y}{\mu}\frac{v_0}{a}$. We assume a weak rock with $\tau_Y/\mu = 1/1000$.

To get an idea of the order of magnitude of the maximum strain rate on the cylinder wall we assume a parabolic distribution of v_2 between the cylinders. It follows from continuity that the maximum velocity on the symmetry line between the cylinders is $3v_0$ (the difference to the corresponding finite element result proved to be less than 20 per cent for purely viscous flow) and the strain rate on the cylinder wall is $6v_0/a$. Hence we expect viscoelastic or elastic–plastic behaviour to dominate for $Wei > 1/6$. In fact if yielding takes

place then the strain rate localizes on the cylinder walls, the strain rate can be significantly higher than our simple estimate and rate-independent behaviour (elastic–plastic) takes over at much smaller values of *Wei* than 1/6. Our conjecture is indeed confirmed by the numerical simulations.

In Fig. 4(a) we show the history on the force acting on the cylinder over time for various values of *Wei*. On start up of the flow the system behaves elastically. If v_0 is large we observe a rapid build up of strain and corresponding elastic stress on the sides of the cylinder. The stress level is ultimately limited by the yield stress. If v_0 is small that is, $Wei \ll 1$ then the stress relaxes before yielding can occur.

In Fig. 4(b) we represent the quasi steady state value of the force on the cylinder at $v_0 t/a = 0.004$ versus the Weissenberg number. Because of the strain concentration on the sides of the cylinder the transition to plastic, quasi-rate-independent flow occurs already for $Wei > 0.02$. The flow is predominantly viscous for $Wei < 0.006$ and viscoplastic in between. A higher or lower yield stress will not affect the non-dimensional value on the force on the cylinder in Fig. 4(b) as long as the yield stress is smaller than the shear modulus. If the yield stress is of the same order of magnitude as the shear modulus then side effects as described in the previous subsection which are associated with the co-rotational stress terms may occur.

For each loading curve in Fig. 4(a) the duration of transient behaviour may be defined by the dimensionless time interval Δt_{rans} v_0/a from 0 to the point where the dimensionless force approaches a constant value. From Fig. 4(a) we find that for predominantly viscous behaviour ($Wei < 0.006$) the interval Δt_{trans} is of the order of one-half of the Maxwell time; Δt_{trans} decreases for increasing *Wei* in the plastic range (Fig. 4a) and approaches 0 for $Wei \to \infty$.

Such short time intervals are insignificant on most geological time scales. In the present example we have assumed ideal plasticity, that is, $\tau_Y = \text{const}$. If strain softening is considered due to some form of damage accumulation then localization and rupture altering the nature of the system may take place during the transient phase. In such cases the evolution of the system has to be modelled as accurately as possible which includes elasticity. In the plastic range (Fig. 4b) we found the inclusion of elasticity convenient from a numerical point of view even if we are interested in steady states only. The time is then functioning as a natural underrelaxation factor.

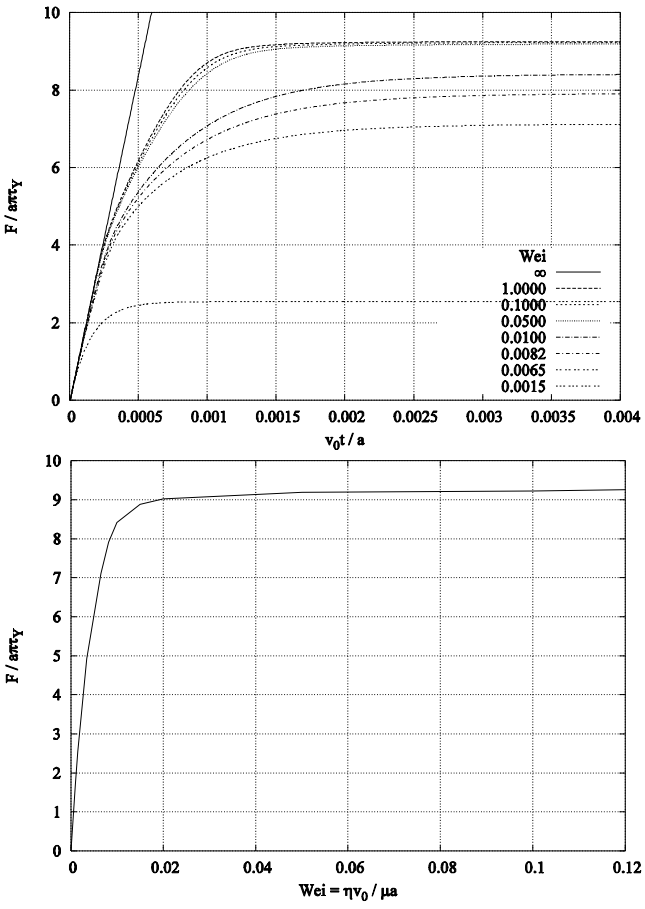

Figure 4. (a) Resulting force on cylinder versus time for different values of $Wei = \eta v_0/\mu a$. See Fig. 3 for geometry and boundary conditions. We assumed $\tau_Y/\mu = 10^{-3}$ for $Wei < \infty$ and $\tau_Y/\mu = \infty$ for $Wei = \infty$ (elastic limit). (b) Force on cylinders at steady states vs. Weissenberg number Wei. The system is ideal plastic for $Wei > 0.02$, predominantly viscous for $Wei < 0.006$ and viscoplastic in between.

NATURAL CONVECTION

Mesh dependence and transition stress

In the following study on mesh dependence of the results and subsequently on the influence of the transition stress τ_0 we ignore the temperature dependence of the viscosity. The results displayed in Fig. 5 are obtained assuming $Ra^c = 10^6$ (see eq. 14), $(\mu/\eta^*)t_D = \infty$, $n = 3$, $\tau_0 = 10^4$ and $\tau_Y/\tau_0 = \infty$. In all simulations we shall assume that $Di = 0$ and the pressure dependence of the melting temperature is ignored. More sophisticated models including pressure and state-variable-dependent yield stress will be considered in a forthcoming paper. The different graphs of the Nusselt number correspond to different discretizations of the problem. The coarsest discretization (154 six-noded triangles) produces the usual result: a steady state is reached after a few decaying oscillations. An intermediate convection structure consisting of three convection cells (one major cell in the middle flanked by to smaller cells on the sides) gets temporarily locked in, in the finer discretizations, until the intermediate structure becomes unstable and the system settles into a single cell steady state. The Nusselt numbers at steady state are remarkably similar for all three discretizations.

The results for the medium and the fine discretization are qualitatively the same: both predict the existence of an intermediate state, which eventually becomes unstable. Qualitative changes in the convection pattern are associated with oscillation of $Nu(t)$. In the finer

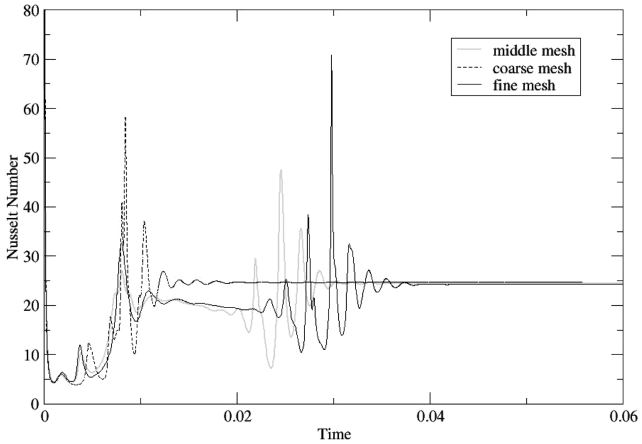

Figure 5. Mesh dependence of solution of finite element models. $Ra^c = 10^6$. The average dimensionless viscosity at steady state is 0.7. The mesh consists of 154, 734 and 1488 six-noded triangular elements respectively.

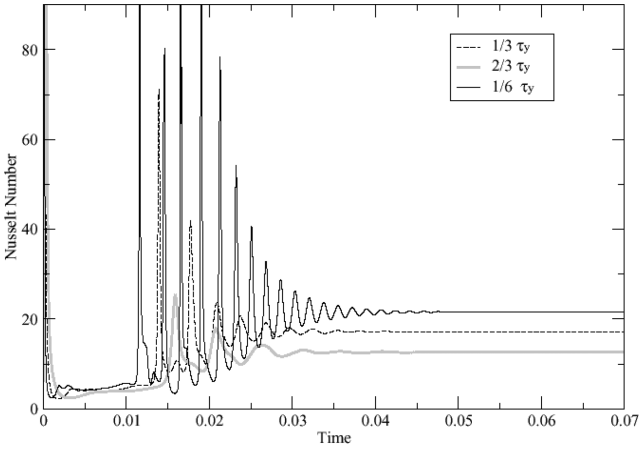

Figure 6. Nusselt number versus time for transition stresses $\tau_0 = 10^{2.5}, 10^{2.5}$ and $0.5 \times 10^{2.5}$ The yield stress $\tau_Y = 3 \times 10^{2.5}$. Computational Rayleigh number: $Ra^c = 10^5$. The average viscosities at steady state are 0.44, 0.21 and 0.10.

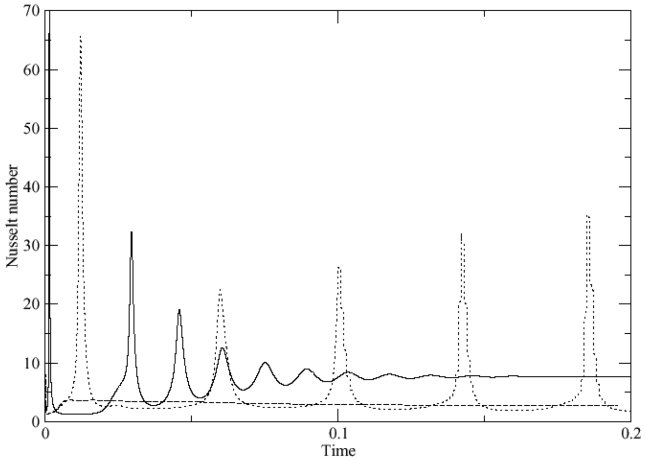

Figure 7. Nusselt number vs. time. Parameters: yield stress $\tau_Y = (1, 3, >6)\tau_0$—corresponding to mobile, episodic and stagnant lid behaviour; transition stress $\tau_0 = 0.866\,10^{2.5}$; $Ra^c = 10^4$; $(\mu/\eta^*)\ t_D = 10^4$; discretization unstructured, 734 six-noded triangles; Courant number $C = 0.1$. The average viscosity at the end of the stagnant lid calculation was 0.27, 0.33 for the steady state of the mobile lid, and between 0.20 and 0.40 of 0.40 for the episodic behaviour. Mobile and stagnant lid cases have steady states. The higher Nusselt number at steady state corresponds to the mobile lid case.

discretization the transition from the intermediate to the steady state appears delayed somewhat, however studies in a different context indicate that the delay diminishes for smaller time steps. The coarseness of the 154 element discretization acts as a filter suppressing the emergence of the intermediate structure. The dimensionless average steady state viscosity is about 0.7 in all three cases.

Next we consider the influence of the transition stress. Fig. 6 shows histories of the Nusselt number for $\tau_0 = 2 \times 10^{2.5}, 1 \times 10^{2.5}$, $0.5 \times 10^{2.5}$. In all cases we assume again $n = 3$ and an infinite shear modulus. The steady state Nusselt numbers at steady state are 12.61, 17.21 and 21.70; the average dimensionless viscosity at steady state is 0.44, 0.21 and 0.10 respectively and $Ra^c = 10^5$. In all cases we have assumed a dimensionless yield stress of $(\tau_Y/\eta^*)t_D =$ 10^2.5 and a power-law representation of ideal plasticity; that is, $\tau_Y = \text{const}$; $\eta_Y = \eta_{N0}$, $n_{pl} = 15$ (see Appendix A2 for details).

The approach to steady state becomes increasingly more oscillatory with decreasing transition stress, as expected. The effective Rayleigh numbers, based on the average viscosities are $Ra^{\text{eff}} = 2.27 \times 10^5$, 4.76×10^5 and 10^6. The Nusselt numbers obtained from our non-Newtonian analyses match very closely the ones from a Newtonian analysis based on the average viscosity at steady state. There are differences in the transient phase, however. The consideration of elasticity does not have a major effect. We have conducted simulations assuming $(\mu/\eta^*)t_D = 10^5$. The result was that the peaks during the transient phase were somewhat higher however the steady states turned out almost identical to the infinite elasticity simulations. The influence of elasticity will be investigated further in the context of simulations with temperature-dependent viscosity.

Temperature-dependent viscosity

The temperature dependence of the viscosity is considered as defined by (19) with $\hat{A} = 23$. As we mentioned above, the viscosity ratio from the cold to the hot boundary due to temperature alone is 10^5. More extreme viscosity contrasts are not a problem in the present formulation because the upper limit for the effective dimensionless viscosity is set by the dimensionless elastic shear modulus and the time increment $\mu t_D/\eta^* \delta t$. We assume $Ra^c = 10^4, \tau_0 = 0.866 \times 10^{2.5}, \tau_Y = 3\tau_0$, ideal plasticity i.e. $\tau_Y = \text{const}$ and $(\mu/\eta^*)t_D = 10^4$. In the simulations we use the power-law plasticity model with $n_Y = 15$ and $\eta_Y = \eta_{N0}$ (see A2 for details).

The corresponding Nusselt number *vs*. time plot is shown in Fig. 7. After initial, rapidly decaying primarily elastic oscillations, the system settles temporarily into stagnant lid type convection (Fig. 8). During this initial convection stresses build up until the yield stress is locally reached. The locally increased mobility is accompanied by thermal advection, a narrow plume is forming, hot material is advected underneath a narrow cold boundary layer until finally the cold layer plunges into the model mantle along the boundary opposite to the plume (Fig. 9).

This process repeats itself in apparently regular intervals. The time intervals between the first peaks of the episodic case in Fig. 7 are 0.071, 0.044, 0.045, 0.044, const. spacing from then on. Also shown in Fig. 7 are the Nusselt numbers for mobile lid convection (obtained for $\tau_Y = \tau_0$) and stagnant lid convection ($\tau_Y \geq 6\tau_0$). The isotherms and velocity arrows displayed in Figs 8 and 9 are representative for the stagnant lid phases (minima of Nusselt plots in Fig. 7) and the subduction events (maxima of Nusselt plots), for episodic convection. Episodic behaviour based on 3-D rigid visco plastic model is also observed by Trompert & Hansen (1998) and Stein *et al.* (submitted 2004). Based on numerical experiments, Stein *et al.* have identified sub-domains in the parameter space of their model of stagnant lid convection, episodic behaviour and mobile lid convection.

The vertical lines plotted along the cold boundary of the velocity plots are the difference between the magnitude of the largest horizontal velocity anywhere in the domain and the magnitude of the local horizontal velocity on the cold boundary, divided by the largest horizontal velocity; i.e. if the local horizontal velocity on the cold boundary happens to be the largest one, then the line has locally the length zero. If the horizontal velocity on the cold boundary is zero (stagnant lid behaviour) the lines have uniformly the length one (Fig. 8). The fact that the horizontal velocity distribution on the cold boundary layer is flat on large parts of the body is indicative of

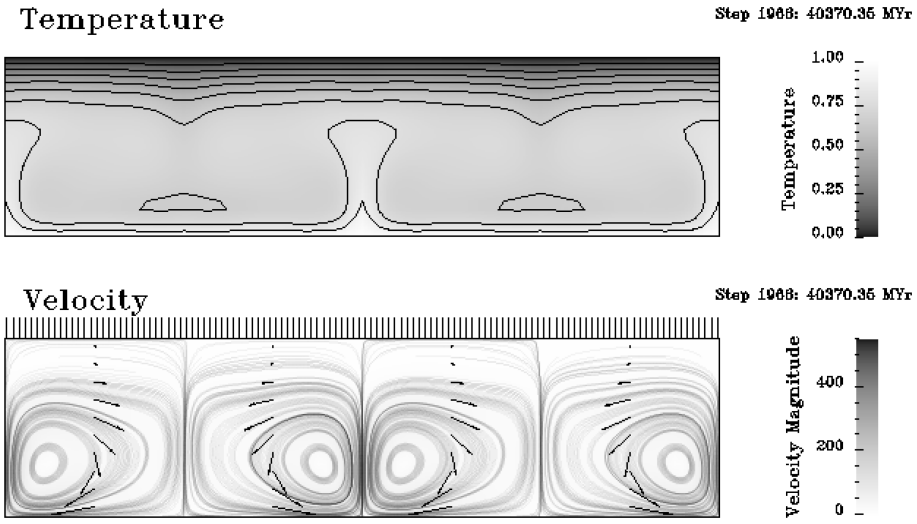

Figure 8. Typical temperature and velocity distributions for episodic convection at a minimum of the Nusselt number (refer to Fig. 7).

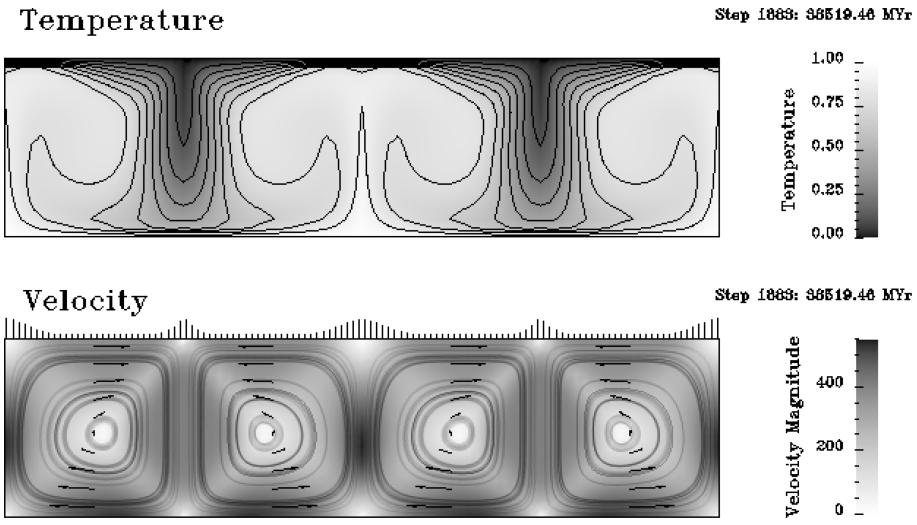

Figure 9. Typical temperature and velocity distributions for episodic convection at a maximum of the Nusselt number (refer to Fig. 7).

plate-tectonics like behaviour; that is, parts of the boundary move not unlike rigid bodies.

The question is what controls this kind of behaviour. A key model property, beside the elastic shear modulus is obviously the existence of a finite yield stress. The influence of the yield stress is represented in Fig. 10.

Simulations were conducted for $\tau_Y = (3, 4, 5)\tau_0$. The spacing between the five peaks of the $\tau_Y = 3\tau_0$ simulation is 0.071, 0.044, 0.045 and 0.044; the spacing between the four peaks for the $\tau_Y = 4\tau_0$ is 0.055, 0.051 and 0.054. The effect of the increasing yield stress is to somewhat delay the onset of the first episodic event and extend the period between the events. No episodic behaviour is observed for $\tau_Y > 6\tau_0$. Finally we consider the sensitivity of the episodic modes to values of the elastic shear modulus (Fig. 11).

The influence of the calculation to the magnitude of the elastic shear modulus becomes important if $(\mu/\eta)^* t_D < 10^5$ for the parameters assumed. Virtually no difference can be observed between the $(\mu/\eta)^* t_D = 10^5$ and the $(\mu/\eta)^* t_D = 10^{20}$ cases. Smaller shear moduli cause a shortening of the period between the subduction events. The area under a spike in the Nusselt plot is directly related to the mechanical work associated with a subduction event.[1] The area underneath the peaks—and thus the work associated with a subduction event—decreases with decreasing value of the elastic shear modulus. It is remarkable that this effect is noticeable even in global energy measure such as the Nusselt number.

CONCLUSIONS AND FUTURE DIRECTIONS

We have discussed the salient features of viscoelastic–plastic flows in a section of shear flows in which we consider the cases of simple shear and flow around circular cylinders. We derive an analytical solution for transient viscoelastic flow and discuss possible unphysical side effects associated with the choice of particular co-rotational stress rates. We conclude that these unphysical side effects do not

[1] By multiplying (11) with v_i, integration of the resulting expression over the domain V and application of Gauss theorem we obtain the total mechanical power $\dot{W}$ as $\dot{W} = RaV(Nu - 1)$. See also Parmentier *et al.* (1976).

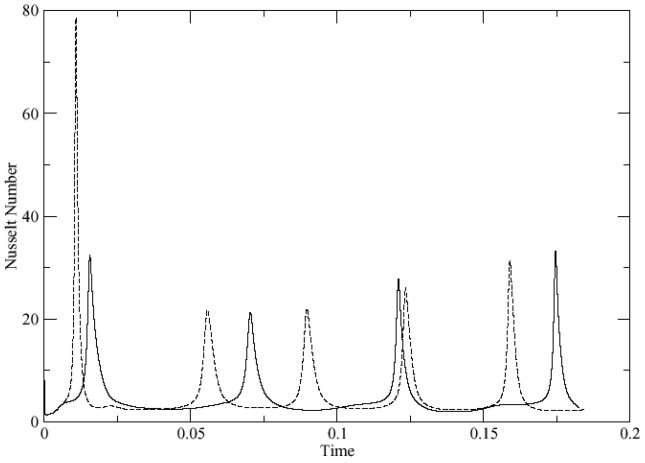

Figure 10. The influence of the yield stress: Simulation for $\tau_Y = (3, 4)\tau_0$. No episodic behaviour is observed for $\tau_Y > 6\,\tau_0$.

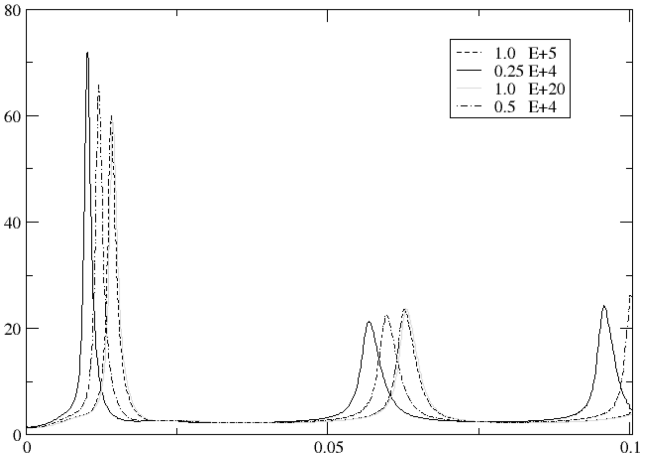

Figure 11. Influence of elastic shear modulus: $(\mu/\eta)^* t_D = 0.25\,10^4$, $0.5\,10^4$, 10^5, 10^{20}; $n = 3$, $n_{pl} = 15$, $Ra^c = 10^4$, $\tau_0 = 0.866\,10^{2.5}$, $\tau_Y = 3\tau_0$.

occur for stress levels possible in rocks. The latter is guaranteed if a suitable yield criterion is included in the constitutive description as proposed here and elsewhere (e.g. Moresi *et al.* 2002).

In the subsequent section we used the case of flow around cylinders to highlight parameter ranges governed by viscous, viscoplastic and viscoelasto–plastic behaviour. The local Weissenberg number between the cylinders is much higher than the global Weissenberg number defined by the far field velocity and the cylinder radius and the Maxwell time. This strong variation between global and local flow characteristics is not unlike the situation in mantle convection where the stress in a subducting slab can be much higher than the average stress related to mantle size, average viscosity and average plate speed. In the plot of the steady state values of the resulting force on the cylinders versus the Weissenberg number we observe a sharp transition between predominantly viscous to rigid plastic behaviour with a relatively narrow interval of viscoplastic behaviour in between.

The formulation described for viscoelastic–plastic geological flows is based on a combined Newtonian and power-law rheology; the effect of plastic yielding is considered by an additional power-law term with a high ($n \geq 15$) power-law coefficient (eq. A13). The model is valid for studying the geodynamics of mantle convection among other problems. The non-linear equations of motion are solved incrementally based on a consistent tangent formulation producing second-order accurate results so that iterations within each time step are not necessary in most cases. In Moresi & Solomatov (1998) and Tackley (1998) plastic yielding is considered by introducing an upper limit to the viscosity given by the ratio of the yield stress and the equivalent viscous strain rate. Since the strain rate at the current time is unknown, an initial estimate has to be based on the strain rate from the last time step producing first-order accurate results; hence a time consuming, iterative approach is necessary. The iterative approach is usually more time consuming than the present incremental approach with occasional iterative reduction of residuals. In the iterative approach the constitutive operator is more sparse than in the consistent incremental approach, which sometimes can be used to advantage. The convection problem with strongly temperature-dependent viscosity has some unique characteristics: the strain in much of the system is very large, necessitating a fluid dynamics formulation, yet the relaxation time in the cool thermal boundary layer is significant compared to the characteristic time associated with fluid flow.

In the bulk of the fluid the relaxation time is small compared to the time taken for convective features to evolve due to the much lower viscosity of the warm fluid. Because elastic stresses in the strongly convecting part of the system relax rapidly, the introduction of elasticity does not produce a qualitative change to the stagnant lid convection regime (see Solomatov 1995). In episodic and mobile lid regimes, there is a competition between the build-up of stresses in the cool lid, and the stress-limiting effect of the yield criterion. The introduction of elastic deformation does not influence this balance either, although we do expect a difference in the distribution of stresses in the lid, which explains the variation in the onset of overturns and their increasing frequency which we observed as the elastic shear modulus was reduced. We expect also that the presence of an elastic deformation mechanism allows deformation of the highly viscous lid with lower viscous energy-dissipation rates. This is reflected in the lower energy dissipation during episodic overturns which we observed by integrating the system Nusselt number. In the Earth this effect may be important in subduction zones where prediction of dissipation rates due to slab bending is unphysically large. In a next step towards self-consistent plate mantle instabilities at the local scale of slab bending or subduction initiation are important. We have shown that elasticity plays a crucial role at this local scale.

The present power-law representation of plastic flow has the advantage that strain localization due to strain softening or the fact that pressure sensitivity is not matched by a corresponding volumetric dilatancy and is not accompanied by a change of type of the governing equations (Needleman 1988). The notorious mesh sensitivity of finite element solutions (beyond the usual discretization error) does not occur in this case. We plan to expand our model to include pressure sensitivity and history dependence in the form of suitable self-consistent hardening/ softening relationships and apply the extended model to more detailed studies of subduction and large-scale shear banding.

ACKNOWLEDGMENTS

The authors would like to acknowledge valuable comments by A/Prof L Moresi and the financial support by the Australian Computational Earth System Simulator (ACcESS) MNRF (H M) and the Predictive Minerals Discovery CRC (PMD CRC).

REFERENCES

Bercovici, D., 1993. A simple model of plate generation from mantle flow, *Geophys. J. Int.,* **114,** 635–650.

Biot, M., 1965. *The Mechanics of Incremental Deformations,* John Wiley, New York.

Braun, J., 1994. 3-dimensional numerical simulations of crustal-scale wrenching using a nonlinear failure criterion, *Journal of Structural Geology,* **16**(8), 1173–1186.

Funiciello, F., Morra, G., Regenauer-Lieb, K. & Giardini, D., 2003. Dynamics of retreating slabs: 1. Insights from two-dimensional numerical experiments, *Journal of Geophysical Research-Solid Earth,* **108**(B4), art. no.-2206.

Gruntfest, I.J., 1963. Thermal feedback in liquid flow; plane shear at constant stress, *Transactions of the Society of Rheology*, **8,** 195–207.

Gurnis, M., Eloy, C. & Zhong, S.J., 1996. Free-surface formulation of mantle convection.2. Implication for subduction-zone observables, *Geophys. J. Int.,* **127**(3), 719–727.

Harder, H., 1991. Numerical-simulation of thermal-convection with maxwellian viscoelasticity, *Journal of Non-Newtonian Fluid Mechanics,* **39**(1), 67–88.

Hill, R., 1998. *The Mathematical Theory of Plasticity,* Oxford University Press, Oxford.

Karato, S. & Wu, P., 1993. Rheology of the Upper Mantle: A Synthesis, *Science* **260,** 771–778.

Kirby, S.H., 1983. Rheology of the Lithosphere, *Reviews of Geophysics and Space Physics*, **21,** 1458–1487.

Kolymbas, D. & Herle, I., 2003. Shear and objective stress rates in hypoplasticity, *International Journal for Numerical and Analytical Methods in Geomechanics,* **27**(8), 733–744.

Leroy, Y.M. & Molinari, A., 1992. Stability of steady states in shear zones, *J. Mech. Phys. Solids*, **40,** 181–212.

Malvern, L.E., 1969. *Introduction to the Mechanics of a Continuous Medium,* Prentice-Hall, New Jersey.

Melosh, H.J., 1978. Dynamic support of the outer rise, *Geophys. Res. Lett.,* **5,** 321–324.

Moresi, L. & Solomatov, V., 1998. Mantle convection with a brittle lithosphere: thoughts on the global tectonic styles of the Earth and Venus, *Geophys J. Int.,* **133**(3), 669–682.

Moresi, L., Dufour, F. & Muhlhaus, H., 2002. Mantle convection models with viscoelastic/brittle lithosphere: numerical methodology and plate tectonic modeling, *Pageoph,* **159**(9), 2335–2356.

Morra, G. & Regenauer-Lieb, K., 2005. A coupled solid-fluid method for modelling subduction, *Philosophical magazine,* London, in press.

Muhlhaus, H.B., 1985. Surface instability of a half space with bending stiffness (in German), *Ing. Archive,* **56,** 383–388.

Needleman, A., 1988. Material rate dependence and mesh sensitivity in localization problems. *Comput. Methods Appl. Mech. Engrg.,* **67,** 69–85.

Ogawa, M., 2003. Plate-like regime of a numerically modeled thermal convection in a fluid with temperature-, pressure-, and stress-history-dependent viscosity, *Journal of Geophysical Research-Solid Earth,* **108**(B2), 2067–2072.

Parmentier, E.M., Turcotte, D.L. & Torrance, K.E., 1976. Studies of finite amplitude non-Newtonian thermal convection with application to convection in the Earth mantle, *J. geophys. Res.,* **81,** 1839–1846.

Phan-Thien, N., 2002. *Understanding Viscoelasticity-Basics of Rheology,* Springer Verlag, Berlin Heidelberg New York.

Poliakov, A., Podladchikov, Y., Dawson, E. & Talbot, C.J., 1996. Salt diapirism with simultaneous brittle faulting and viscous flow, *Geological Society,* **100,** 291–302.

Prager, W., 1961. *Introduction to Mechanics of Continua,* Boston, MA: Ginn.

Regenauer-Lieb, K. & Yuen, D.A., 2004. Positive feedback of interacting ductile faults from coupling of equation of state, rheology and thermal-mechanics, *Phys. Earth planet. Int.,* **142**(1–2), 113–135.

Regenauer-Lieb, K., Yuen, D. & Branlund, J., 2001. The initation of subduction: criticality by addition of water?, *Science,* **294,** 578–580.

Schmalholz, S.M. & Podladchikov, Y., 1999. Buckling versus folding: importance of viscoelasticity. *Geophys. Res. Lett.,* **26**(17), 2641–2644.

Scholz, C.H., 1990. *The Mechanics of Earthquakes and Faulting,* Cambridge University Press, Cambridge.

Solomatov, V., 1995. Scaling of temperature-dependent and stress-dependent viscosity convection. *Physics of Fluids,* **7**(2), 266–274.

Stein, C.A., Schmalzl, J. & Hansen, U., 2004. The effect of rheological parameters on plate behaviour in a self-consitent model of mantle convection, *Phys. Earth planet. Int.,* **142(3–4)**, 225–255.

Tackley, P., 1998. Self-consistent generation of tectonic plates in three-dimensional mantle convection, *Earth planet. Sci. Lett.,* **157,** 9–22.

Tackley, P., 2000a. Self-consistent generation of tectonic plates in time-dependent, three-dimensional mantle convection simulations, 1. Pseudo-plastic yielding, *G3*, **01**(23), 1525.

Tackley, P., 2000b. Self-consistent generation of tectonic plates in time-dependent, three-dimensional mantle convection simulations. 2. Strain weakening and asthenosphere, *G3*, **01**(25), 2027.

Trompert, R. & Hansen, U., 1998. Mantle convection simulations with rheologies that generate plate-like behavior, *Nature,* **395,** 686–689.

Turcotte, D.L. & Schubert, G., 2002. *Geodynamics,* 2nd edn, Cambridge University Press, Cambridge.

Zienkiewicz, O.C. & Taylor, R.L., 2000. *The Finite Element Method,* 5th edn, Butterworth/Heinemann, Oxford, ISBN 0 7506 5050 8, 3.

APPENDIX

A1 Simple shear

The following derivations are based on the assumption of the Jaumann rate in the constitutive relationships (2). It should be mentioned that for simple shear the Jaumann rate based model and the Oldroyd based model are identical. For isothermal, plane simple shear parallel to x_1, the equilibrium conditions read:

$$\sigma_{12,2} = 0,$$
$$\sigma_{22,2} = 0. \tag{A1}$$

From the general definition of the Jaumann stress rate

$$\dot{\sigma}_{ij}^{J} = \dot{\sigma}_{ij} - W_{ik}\sigma_{kj} + \sigma_{ik}W_{kj}, \tag{A2}$$

we obtain for simple shear:

$$\begin{aligned}
\dot{\sigma}_{22}^{\prime J} &= -\dot{\sigma}_{11}^{\prime J} = \dot{\sigma}_{22}^{\prime} - \frac{1}{2}(v_{2,1} - v_{1,2})\sigma_{12} - \sigma_{21}\frac{1}{2}(v_{2,1} - v_{1,2}) \\
&= \dot{\sigma}_{22}^{\prime} + v_{1,2}\sigma_{12}, \\
\dot{\sigma}_{12}^{J} &= \dot{\sigma}_{12} - \frac{1}{2}v_{1,2}\sigma_{22}^{\prime} + \sigma_{11}^{\prime}\frac{1}{2}v_{1,2} \\
&= \dot{\sigma}_{12} - v_{1,2}\sigma_{22}^{\prime}.
\end{aligned} \tag{A3}$$

Insertion into the constitutive relationship (2) yields:

$$\begin{aligned}
&\lambda\dot{\sigma}_{12} - \lambda\dot{\gamma}\sigma_{22}^{\prime} + \sigma_{12} = \eta\dot{\gamma}, \\
&\lambda\dot{\sigma}_{22}^{\prime} + \lambda\dot{\gamma}\sigma_{12} + \sigma_{22}^{\prime} = 0,
\end{aligned} \tag{A4}$$

where

$$\dot{\gamma} = v_{1,2}, \quad \lambda = \frac{\eta}{\mu} \text{ and } \eta = \left(\frac{1}{\eta_N} + \frac{1}{\eta_N(\tau/\tau_0)^{1-n}}\right)^{-1}. \tag{A5}$$

In steady state the stress rate in A4 disappears; elimination of σ_{22} yields eq. (22) in the simple shear section.

A2 Tangent operator

Incremental expansion of the constitutive relation

$$D_{ij} = \frac{1}{2\mu}\dot{\sigma}_{ij}^{\prime J} + \frac{1}{2\eta}\sigma_{ij}^{\prime} + \dot{\gamma}^{p}\frac{\sigma_{ij}^{\prime}}{2\tau}, \tag{A6}$$

with a power-law evolution equation for $\dot{\gamma}^p$

$$\dot{\gamma}^p = \frac{\tau_Y}{\eta_Y}\left(\frac{\tau}{\tau_Y}\right)^{n_{pl}}, \tag{A7}$$

and η is defined by eq. (7), yields:

$$D_{ij} - \frac{\delta t}{2\eta_{\rm eff}}(W_{ik}\sigma_{kj} - \sigma_{ik}W_{kj}) - \frac{1}{2\mu}\sigma'^{\,}_{ij,k}v_k = \frac{1}{2\mu_{\rm eff}\delta t}\delta\sigma'^{J}_{ij} + \underbrace{\left(\frac{1}{h_{\rm vis}} + \frac{1}{h_Y}\right)}_{1/h}\frac{\sigma'_{ij}}{2\tau}\frac{\sigma'_{kl}}{2\tau}\delta\sigma'^{J}_{kl} + \frac{1}{2\eta_{\rm eff}}\sigma'_{ij} + \frac{1}{2\eta_{\rm eff}}\sigma'_{ij}\left(\frac{\delta T}{h_T} + \frac{\delta p}{h_p}\right), \tag{A8}$$

where

$$\dot{\sigma}'_{ij} = \sigma'_{ij,t} + \sigma'_{ij,k}v_k, \tag{A9}$$

$$\delta\sigma^J_{ij} = (\sigma_{ij,t} - W_{ik}\sigma_{kj} + \sigma_{ik}W_{kj})\delta t, \quad \delta T = T_{,t}\delta t, \quad \delta p = p_{,t}\delta t \tag{A10}$$

$$h_{\rm vis} = \frac{1}{(n-1)}\eta_N\left(\frac{\tau}{\tau_0}\right)^{1-n}, \quad \text{and } h_Y = \frac{1}{(n_{pl}-1)}\eta_Y\left(\frac{\tau}{\tau_Y}\right)^{1-n_{pl}} \tag{A11}$$

$$h_T = \frac{T^2}{AT_M}, \quad h_p = -\frac{T}{A\frac{\partial T_M}{\partial p}}, \tag{A12}$$

where η_N is defined by (6) and

$$\eta_{\rm eff} = \left(\frac{1}{\eta_N} + \frac{1}{\eta_N\left(\frac{\tau}{\tau_0}\right)^{1-n}} + \frac{1}{\eta_Y\left(\frac{\tau}{\tau_Y}\right)^{1-n_{pl}}}\right)^{-1},$$

$$\mu_{\rm eff} = \left(\frac{1}{\mu} + \frac{\delta t}{\eta_{\rm eff}}\right)^{-1}. \tag{A13}$$

Inversion of A8 yields:

$$\delta\sigma'^{J}_{ij} = \left(\mu_{\rm eff}\delta t(\delta_{ik}\delta_{jl} + \delta_{jk}\delta_{il}) - \frac{(\mu_{\rm eff}\delta t)^2}{h + \mu_{\rm eff}\delta t}\frac{\sigma'_{ij}}{\tau}\frac{\sigma'_{kl}}{\tau}\right) \times \left(D_{kl} - \frac{\sigma'_{kl,m}v_m}{2\mu}\right) - \frac{\mu_{\rm eff}\delta t}{\eta_{\rm eff}}\left(1 - \frac{\mu_{\rm eff}\delta t}{h + \mu_{\rm eff}\delta t}\right) \times \left(1 + \frac{\delta T}{h_T} + \frac{\delta p}{h_p}\right)\sigma'_{ij} - \frac{\mu_{\rm eff}\delta t^2}{\eta_{\rm eff}}(W_{ik}\sigma_{kj} - \sigma_{ik}W_{kj}), \tag{A14}$$

where

$$h = \frac{h_{\rm vis}h_Y}{h_{\rm vis} + h_Y}. \tag{A15}$$

The function h_p in eqs (A8) and (A14) considers the pressure dependence of T_M. The slope of the T_M–p curve is usually assumed as constant; a typical order of magnitude value for the slope is 10^{-7} K Pa^{-1}. For $n = 1$ and $\mu \to \infty$ we obtain Newtonian flow with the viscosity $1/2\eta_N$.

In many practical applications the velocity problem, the pressure/incompressibility problem and the heat equation are solved sequentially. In this case the terms associated with δT and δp in eq. (A14) are not needed. In linear instability analyses however the full incremental form (A14) is required.

In the viscous limit, $\mu \to \infty$, (A14) reduces to:

$$\delta\sigma'_{ij} = \left(\eta_{\rm eff}(\delta_{ik}\delta_{jl} + \delta_{jk}\delta_{il}) - \frac{\eta^2_{\rm eff}}{h + \eta_{\rm eff}}\frac{\sigma'_{ij}}{\tau}\frac{\sigma'_{kl}}{\tau}\right)D_{kl} - \left(1 - \frac{\eta_{\rm eff}}{h + \eta_{\rm eff}}\right)\left(1 + \frac{\delta T}{h_T} + \frac{\delta p}{h_p}\right)\sigma'_{ij}. \tag{A16}$$

In steady states, the stress, temperature and pressure increments vanish so that the remaining terms in (A16) have to cancel. The latter is indeed the case as can be shown by insertion of $D_{ij} = 1/2\eta_{\rm eff}\,\sigma'_{ij}$.

The limit $\eta_N \to \infty$ does not yield a simpler expression for the incremental relationship A14 since the effective moduli still depend on the viscosity $\eta_Y(\tau_Y/\tau)^{n_{pl}-1}$. In the rate-independent limit, which is obtained for $\eta_N \to \infty$, $n_{pl} \to \infty$ and $\delta t \to 0$, the expression A14 reduces to

$$\partial\sigma'_{ij}/\partial t = \left(\mu(\delta_{ik}\delta_{jl} + \delta_{jk}\delta_{il}) - \mu\frac{\sigma'_{ij}}{\tau}\frac{\sigma'_{kl}}{\tau}\right)\left(D_{kl} - \frac{1}{2\mu}\sigma'_{ij,k}v_k\right) + (W_{ik}\sigma_{kj} - \sigma_{ik}W_{kj}). \tag{A17}$$

In the above derivations it was always assumed that the yield stress is constant. A large deformation model with power-law plasticity with state variable dependence of the yield stress will be presented in a forthcoming paper.