ELSEVIER

Physics of the Earth and Planetary Interiors 97 (1996) 83–94

PHYSICS OF THE EARTH AND PLANETARY INTERIORS

The accuracy of finite element solutions of Stokes' flow with strongly varying viscosity

Louis Moresi *, Shijie Zhong, Michael Gurnis

Seismological Laboratory, California Institute of Technology, Pasadena, CA 91125, USA

Received 17 October 1995; accepted 21 February 1996

Abstract

We provide benchmark comparisons of two finite element (FE) mantle convection codes, CITCOM and CONMAN, against analytic solutions for Stokes' flow for strongly varying viscosity in the horizontal and vertical directions. The two codes use a similar FE formulation but different methods for solving the resulting equations. They both obtain accurate velocity, pressure and surface topography for viscosity structures which vary rapidly over a short distance, or discontinuously. The benchmarks help determine how many elements are needed to resolve a region of, for example, a convection simulation with high viscosity gradients. The overall accuracy does not depend on the global viscosity variation but on the gradients within individual elements. As a rule of thumb, accuracy can fall below 1% when there is a viscosity variation greater than a factor of two or three in an element. For iterative solution methods, necessary in three-dimensional modelling, these guidelines are required to determine the number of iterations to perform. We discuss a penalty based technique which improves the convergence of iterative solvers for general problems in which high viscosity gradients occurs spontaneously.

1. Introduction

The finite element method (FEM) has become a widespread computational tool because of its ability to provide reliable solutions to problems in irregular domains with highly variable boundary conditions while having discontinuities in driving forces and material properties. FEM is one of the methods of choice in codes developed to study convection in the Earth's interior (Christensen, 1984; Hansen and Ebel, 1984; Baumgardner, 1985; Schubert and Anderson, 1985; King et al., 1990; Van Keken et al., 1992; Zhong and Gurnis, 1995). Most studies of convection in the mantle use relatively simple domains, frequently with simple boundary conditions, but the material properties can vary rapidly in space and time because the rheology of the constituent minerals is strongly affected by temperature, pressure, stress, and composition (Karato et al., 1986; Karato and Wu, 1993)

Theoretical and experimental studies of thermal convection in fluids with strongly temperature-dependent viscosity (Fowler, 1985; Davaille and Jaupart, 1993b; Solomatov and Moresi, 1996) find that cold boundary layers naturally develop which localize the strongest viscosity variation in the fluid. These boundary layers are analogous to the mantle lithosphere and subducting slabs.

* Corresponding author present address: Research School of Earth Sciences, Australian National University, Canberra, A.C.T. 0200, Australia.

PII S0031-9201(96)03163-9

Thermal plumes within the convecting layer bring hot, low-viscosity material into contact with the cold, high-viscosity boundary layer. Where such plumes rise under the oceanic lithosphere, the bathymetry and geoid anomalies can be used to determine the structure of the upwelling and infer the viscosity of the mantle (Richards et al., 1988; Monnereau et al., 1993; Ribe and Christensen, 1994).

Although the highly viscous boundary layer provides a reasonable approximation to the interior of oceanic plates, these plates cannot move with respect to one another unless the boundaries between them are fairly weak (Kopitze, 1979; Schmeling and Jacoby, 1981). It is common to model plate boundaries by arbitrarily weakening the thermal lithosphere in narrow zones. Typically, these weak zones are given a viscosity equal to, or lower than, that of the mantle underlying the lithosphere and result in very large viscosity jumps across element boundaries (King and Hager, 1990, Gurnis, 1988).

Although creeping flow solutions with highly variable viscosity are primarily of interest to the mantle dynamics community, very similar conditions prevail in convection in lava lakes and magma chambers (Davaille and Jaupart, 1993a). In each of these examples, numerical methods must cope with very strong changes in the viscosity over a relatively small number of grid points. Knowing the behaviour of a specific algorithm in determination of a flow solution under these conditions is therefore critical, particularly when derived quantities such as topography or geoid are to be calculated for comparison with observations.

Previous mantle convection benchmarks with temperature-dependent viscosity have concentrated on comparisons between different numerical techniques and implementations (e.g. Blankenback et al., 1989; Travis et al., 1990) and laboratory experiments (Busse et al., 1994). Analytic solutions allow more exacting benchmarking of these numerical procedures, which are already in widespread use, for cases in which the viscosity varies rapidly or discontinuously within the fluid. One of the most important benefits from these benchmarks is that we can determine the number of elements required to resolve a given gradient in viscosity to a certain accuracy and construct meshes where high densities of elements appear only where they are needed.

We use four analytic solutions for Stokes' flow with viscosity which varies with depth or horizontal distance to benchmark two finite element convection codes. CONMAN uses the penalty function method and direct solution techniques (e.g. Hughes et al., 1979)—probably the most common finite element implementation for two-dimensional mantle convection problems. CITCOM uses a mixed formulation which is solved using a multigrid, iterative solver. The existence of accurate iterative methods is important for the solution of problems in three dimensions, where direct techniques are relatively slow and cumbersome (Cahouet and Chabard, 1988).

The iterative and direct solution schemes produce near-identical solutions for all of the benchmarks. Velocity can be determined with high accuracy given sufficient resolution. Accuracy is not influenced by the overall viscosity contrast in the system but by the local viscosity jump across an element. Pressure and surface topography can also be determined accurately, but are more strongly influenced by strong viscosity gradients and therefore some additional grid refinement may be needed.

2. Numerical methods and implementations

We solve Stokes' equation for creeping flow subject to incompressibility:

$$-\nabla(\eta\dot{\epsilon}) + \nabla p = \Delta\rho g\hat{z} \tag{1}$$

$$\nabla \cdot u = 0 \tag{2}$$

in which η is dynamic viscosity, $\dot{\epsilon} = [(\partial u_i/\partial x_j) + (\partial u_j/\partial x_i)]$ is the strain rate tensor, p is dynamic pressure, $\Delta\rho$ is the density anomaly, g is gravitational acceleration, $\hat{\mathbf{z}}$ is a unit vector pointing downwards, and u is velocity. We use two different finite element codes to solve Eq. (1) and Eq. (2): CITCOM (Moresi and Solomatov, 1995) and CONMAN (King et al., 1990). In some respects the codes are very similar—both use bilinear quadrilateral elements with constant pressure shape functions— but the equations are formulated and solved using different methods.

In CITCOM, Eq. (1) and Eq. (2) are recast into the following matrix form:

$$\boldsymbol{A}\boldsymbol{u} + \boldsymbol{B}\boldsymbol{p} = \boldsymbol{f} \tag{3}$$

$$\boldsymbol{B}^{\mathrm{T}}\boldsymbol{u} = 0 \tag{4}$$

where A is the stiffness matrix, u is a vector of velocities, B is the discrete gradient operator, p is a vector of pressures, and f is a vector of the body and boundary forces. The system of equations is solved using an Uzawa iteration scheme in which Eq. (3) is solved repeatedly with a Gauss–Seidel multigrid method so as to apply the constraint Eq. (4). This method has been described by Cahouet and Chabard (1988), and our specific implementation has been discussed by Moresi and Solomatov (1995). We also used the augmented Lagrangian formulation suggested by Fortin and Thomasset (1983) to improve the convergence of the pressure field for large viscosity variations. This method is related to the penalty method described below in reinforcing the constraint Eq. (4) upon Eq. (3) by a adding to the variational equation a large constant (Lagrange multiplier) times the square of the L^2 norm of the divergence of the velocity. It is usually necessary to use smaller values for the Lagrange multiplier with iterative solvers than with direct solvers. It should be noted that the penalization is added in conjunction with any preconditioning of the pressure iteration.

CONMAN uses a penalty function (Lagrange multiplier) approach coupled with direct solution methods to enforce incompressibility. Eq. (1) and Eq. (2) are solved as a single matrix equation for u,

$$A'u = f \tag{5}$$

in which A' is the stiffness matrix incorporating the additional incompressibility constraints. In this case, the velocity equation is solved directly in one step, and the pressure field found in post-processing.

In addition to velocity and pressure, we also benchmark the surface topography calculated from the vertical normal stress at the top boundary. Both codes make use of a consistent boundary flux method (Zhong et al., 1993) in which velocities and pressures are substituted back into Eq. (3) or Eq. (5) to calculate the surface normal stresses directly.

The accuracy of the direct solver is only dependent on the number of significant digits of the computer implementation, whereas the accuracy of an iterative method also depends on the number of iteration cycles performed. In general, we iterate until the residal r_1 defined by

$$r_1 = |f - Au - Bp|/|f| \tag{6}$$

is smaller than some target value. This target value needs to be small enough to minimize the solution error but not so small that many iterations are performed without improving the solution. In this study, obtaining the maximum accuracy is more important than solution time, therefore we chose a target value of 10^{-9}. We also monitor a second residual which indicates how well the solution obeys the discrete form of the incompressibility constraint:

$$r_2 = |B^{\mathrm{T}}u|/|u| \tag{7}$$

In the Uzawa scheme, r_1 is initially reduced to the target value by ignoring the pressure part of the equation. Then the pressure and velocity are adjusted iteratively to impose the incompressibility constraint. This pressure iteration is continued until the value of r_2 is smaller than the target value.

It is important to point out that these residuals are a measure not of the error from the true solution but of the error from the solution to a discrete form of the problem defined by the choice of shape-functions and element mesh. As we show below, continued reduction of the residual does not guarantee that the error relative to the true solution is decreasing as well. It is important to understand how the true errors relate to the calculated residual for different viscosity structures and grid resolutions so that we can determine the optimal target value for the residuals for other general problems in three dimensions where the solution takes much longer.

3. Benchmark cases

For incompressible Stokes' flow, there are analytic solutions available for four different viscosity structures: a step function in depth, z, measured downward from the surface, or distance, x, increasing from left to right, and exponential in depth or distance. All the benchmark comparisons were done in a 1×1 square box. Case (a), illustrated in Fig. 1(a), has an exponential increase or decrease in viscosity with depth and a harmonic temperature anomaly, δT, concentrated at a single depth—here we chose $\delta T = \delta(z - 0.5)\cos \pi x$ (Revenaugh and Parsons, 1987). Case (b), Fig. 1(b), has two horizontal layers separated by an interface at $z = 0.125$ with high viscosity in the upper layer and a value of unity

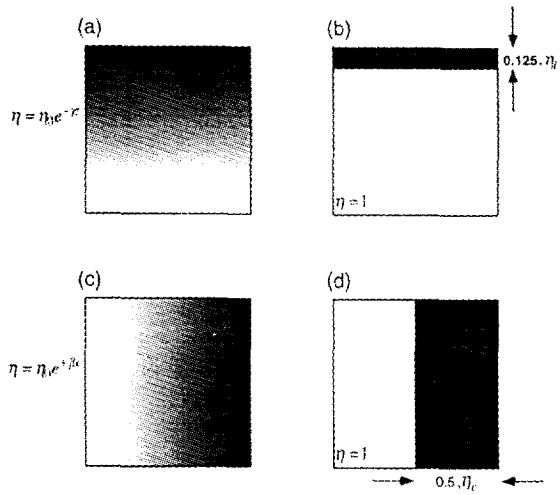

Fig. 1. Illustration of viscosity structures used for benchmark cases. (a) Exponentially increasing viscosity with depth. (b) High-viscosity layer of thickness 0.125 over constant viscosity. (c) Exponentially increasing viscosity with horizontal distance. (d) Columns of contrasting viscosity dividing the layer vertically in half.

in the lower layer, and uses the same temperature anomaly as Case (a) (Hager and O'Connell, 1981). Case (c), Fig. 1(c), has an exponential variation in viscosity with horizontal coordinate and uses a temperature anomaly which is harmonic in depth and horizontally, $\delta T = \sin \pi z \cos \pi x$ (Zhong, 1996). Case (d), Fig. 1(d), has two vertical columns dividing the box at $x = 0.5$, with the right-hand column having a high viscosity, and uses the same density structure as Case (c) (Zhong, 1996). Boundary conditions were free slip top and bottom, and reflecting sidewalls were used in the numerical computations to obtain harmonic solutions. The two solutions for discontinuous viscosities use a propagator matrix technique which is evaluated numerically.

For each viscosity structure, we ran cases with maximum viscosity contrasts of 1, 10^6, 10^{12}, 10^{18}, 10^{24}, and 10^{30} using meshes consisting of 32×32, 64×64, or 128×128 uniformly distributed square elements. We were also able to use 256×256 element meshes with CITCOM, which needs less mem-

Table 1
Root mean square errors in velocity field as a percentage of the magnitude of the velocity for all calculations presented in this paper

Viscosity model	$\Delta\eta$	CITCOM				CONMAN		
		32×32	64×64	128×128	256×256	32×32	64×64	128×128
a	10^0	0.046	0.012	0.0029	0.00073	0.20	0.054	0.0013
a	10^6	0.38	0.091	0.024	0.011	0.46	0.079	0.032
a	10^{12}	1.65	0.37	0.090	0.021	1.58	0.37	0.10
a	10^{18}	3.35	0.89	0.21	0.046	3.37	0.88	0.21
a	10^{24}	4.94	1.65	0.38	0.086	4.98	1.64	0.38
a	10^{30}	6.54	2.57	0.81	0.14	6.58	2.57	0.64
b	10^0	0.046	0.011	0.0029	0.0073	0.20	0.054	0.0013
b	10^6	0.17	0.042	0.011	0.0026	0.18	0.044	0.011
b	10^{12}	0.17	0.042	0.011	0.0026	0.18	0.044	0.011
b	10^{18}	0.17	0.042	0.011	0.0026	0.18	0.044	0.011
b	10^{24}	0.17	0.042	0.011	0.0026	0.18	0.044	0.011
b	10^{30}	0.17	0.042	0.011	0.0026	0.18	0.044	0.011
c	10^0	0.12	0.030	0.0075	0.0019	0.12	0.030	0.0019
c	10^6	0.82	0.20	0.050	0.013	0.82	0.20	0.049
c	10^{12}	3.01	0.74	0.18	0.043	3.01	0.73	0.171
c	10^{18}	5.81	1.68	0.40	0.096	5.81	1.68	0.40
c	10^{24}	7.57	2.96	0.718	0.17	7.44	2.99	0.85
c	10^{30}	9.75	4.41	1.14	0.27	9.75	4.41	1.67
d	10^0	0.12	0.030	0.0075	0.019	0.12	0.030	0.0019
d	10^6	0.90	0.22	0.056	0.014	0.90	0.22	0.056
d	10^{12}	0.90	0.22	0.056	0.014	0.90	0.22	0.056
d	10^{18}	0.90	0.22	0.056	0.014	0.90	0.22	0.056
d	10^{24}	0.90	0.22	0.056	0.014	0.90	0.22	0.056
d	10^{30}	0.90	0.22	0.056	0.013	0.90	0.22	0.056

Viscosity models are taken from Figs. 1 and 1. $\Delta\eta$ is the maximum viscosity contrast in the system.

ory. The maximum viscosity contrasts are higher than we expect to be important in mantle convection problems, but are required for the cases with exponentially varying viscosity to obtain viscosity gradients comparable with those expected in the lithosphere. The analytic solutions require very accurate arithmetical computation, particularly for large viscosity variations. We found that 64 bit precision was only adequate for viscosity contrasts of about 10^{14} or smaller, so we used the multiple precision arithmetic package MPFUN (Bailey, 1993) to calculate every solution to a precision of 100 digits.

4. Results

4.1. Comparisons with analytic solutions

We use the following method to define a single measure of the error, E, between numerical and analytic solutions:

$$E = \left(\frac{|x_{\text{num}} - x_{\text{an}}|^2}{|x_{\text{an}}|^2} \right)^{\frac{1}{2}} \tag{8}$$

where x_{num} is the numerical solution and x_{an} is the analytic solution for velocity, pressure or surface topography. Although this measure is not strongly biased by errors near the stagnation points in the flow, it does not tell us whether the errors are localized or are uniformly distributed across the domain.

The results are listed in Tables 2 and 3, for both CONMAN and CITCOM. The errors are essentially the same for both codes, although on the largest mesh only the solutions from CITCOM are available.

4.1.1. Exponentially varying viscosity

The tabulated results for viscosity varying exponentially in the horizontal and vertical directions are

Table 2
Root mean square errors in surface topography (%)

Viscosity model	$\Delta\eta$	CITCOM				CONMAN		
		32 × 32	64 × 64	128 × 128	256 × 256	32 × 32	64 × 64	128 × 128
a	10^0	0.0099	0.0025	0.00063	0.00021	0.25	0.062	0.016
a	10^6	0.19	0.047	0.011	0.0030	0.052	0.012	0.0032
a	10^{12}	1.20	0.30	0.075	0.019	0.96	0.24	0.065
a	10^{18}	3.23	0.80	0.20	0.050	2.98	0.76	0.19
a	10^{24}	6.31	1.56	0.39	0.097	6.05	1.52	0.38
a	10^{30}	10.45	2.57	0.64	0.16	10.18	2.55	0.63
b	10^0	0.059	0.025	0.0063	0.0021	0.25	0.062	0.016
b	10^6	0.13	0.035	0.0090	0.0020	0.23	0.056	0.014
b	10^{12}	0.13	0.035	0.0088	0.0024	0.23	0.056	0.014
b	10^{18}	0.13	0.034	0.0084	0.0021	0.23	0.056	0.014
b	10^{24}	0.13	0.035	0.0085	0.0020	0.23	0.056	0.014
b	10^{30}	0.14	0.036	0.0090	0.0027	0.23	0.056	0.014
c	10^0	0.080	0.020	0.0050	0.0013	0.038	0.0095	0.0016
c	10^6	10.81	4.58	1.66	0.64	10.78	4.36	1.65
c	10^{12}	23.15	10.76	4.32	1.73	23.71	10.83	4.39
c	10^{18}	34.37	18.56	7.25	3.30	35.11	17.83	7.27
c	10^{24}	44.67	22.45	10.81	5.06	44.24	24.52	11.30
c	10^{30}	50.21	31.07	15.83	6.56	51.4	30.60	14.91
d	10^0	0.040	0.012	0.0050	0.0013	0.038	0.0095	0.0016
d	10^6	0.82	0.31	0.10	0.049	0.80	0.29	0.11
d	10^{12}	0.80	0.29	0.11	0.051	0.80	0.29	0.11
d	10^{18}	0.80	0.28	0.11	0.049	0.80	0.29	0.11
d	10^{24}	0.87	0.29	0.12	0.047	0.80	0.29	0.11
d	10^{30}	0.81	0.32	0.14	0.054	0.80	0.29	0.11

summarized in Fig. 2. The accuracy of the velocity, pressure and surface topography solutions all depend on the viscosity contrast and grid dimension in exactly the same manner. Errors increase with total viscosity contrast, and decrease as more grid points are added. A closer inspection of the tables shows that the controlling factor is the relative viscosity jump across each element. When all elements are the same size, and with an exponential variation in viscosity, this value is the same for every element; for example, in a 32×32 box with an overall viscosity contrast of 10^{12}, the viscosity jump across every element is 2.37. Thus, if the exponent of the viscosity variation, and the number of grid points per unit length, are both doubled, then the error remains approximately constant. In Fig. 2 this is illustrated by the grey shading, which joins calculations sharing the same viscosity contrast per element. In all but two of the 36 comparisons (shown in lighter grey in Fig. 2(c)), the shaded areas do not overlap any calculations with different viscosity contrast per element. In realistic simulations of the Earth's mantle where analytic solutions are not known, this property should prove extraordinarily useful in determining the appropriate numbers of elements to use in regions with strong viscosity gradients.

Using the definition of Eq. (8) above, the velocity errors are considerably smaller than those of the pressure by one or two orders of magnitude. The large pressure error is not surprising, because, for bilinear-velocity elements, the discretization for pressure is constant within each element, and pressure values must be averaged at the nodes (Hughes et al., 1979). Surface topography errors are comparable with velocity errors for the cases with depth-dependent viscosity (compare Fig. 2(a) and Fig. 2(c)), but comparable with the pressure errors when the viscosity varies horizontally (compare Fig. 2(e) and Fig. 2(f)). As the surface topography is determined by substituting nodal velocity and element pressure fields into the equation of motion, the accuracy is dependent on obtaining reliable pressure as well as

Table 3
Root mean square errors in pressure field (%)

Viscosity model	$\Delta\eta$	CITCOM				CONMAN		
		32×32	64×64	128×128	256×256	32×32	64×64	128×128
a	10^0	0.075	0.019	0.0046	0.0012	0.17	0.036	0.0084
a	10^6	12.18	4.72	1.74	0.63	3.48	0.80	0.18
a	10^{12}	27.87	12.43	4.80	1.85	12.53	3.49	1.09
a	10^{18}	41.11	20.73	8.88	3.46	23.25	7.70	2.01
a	10^{24}	51.25	28.57	13.10	5.31	33.26	12.77	3.65
a	10^{30}	58.91	35.60	17.36	7.30	41.85	18.20	5.68
b	10^0	0.075	0.019	0.0046	0.0012	0.17	0.036	0.0084
b	10^6	0.34	0.12	0.040	0.014	0.35	0.12	0.040
b	10^{12}	0.35	0.12	0.041	0.014	0.35	0.12	0.040
b	10^{18}	0.35	0.12	0.040	0.014	0.35	0.12	0.040
b	10^{24}	0.35	0.11	0.041	0.014	0.35	0.12	0.040
b	10^{30}	0.34	0.12	0.041	0.014	0.35	0.12	0.040
c	10^0	0.40	0.15	0.025	0.0063	0.40	0.15	0.025
c	10^6	10.24	4.04	1.51	0.55	10.24	4.04	1.51
c	10^{12}	24.14	10.75	4.29	1.61	24.08	10.75	4.29
c	10^{18}	37.04	18.55	7.92	3.08	37.08	18.54	7.92
c	10^{24}	47.44	26.70	11.97	4.52	47.37	26.19	11.94
c	10^{30}	55.39	33.22	16.09	6.74	55.32	33.22	16.09
d	10^0	0.40	0.15	0.025	0.0063	0.40	0.15	0.025
d	10^6	0.82	0.29	0.10	0.037	0.82	0.29	0.10
d	10^{12}	0.98	0.29	0.11	0.036	0.82	0.29	0.10
d	10^{18}	0.98	0.29	0.11	0.036	0.82	0.29	0.10
d	10^{24}	0.88	0.29	0.11	0.036	0.82	0.29	0.10
d	10^{30}	0.98	0.29	0.11	0.036	0.82	0.29	0.10

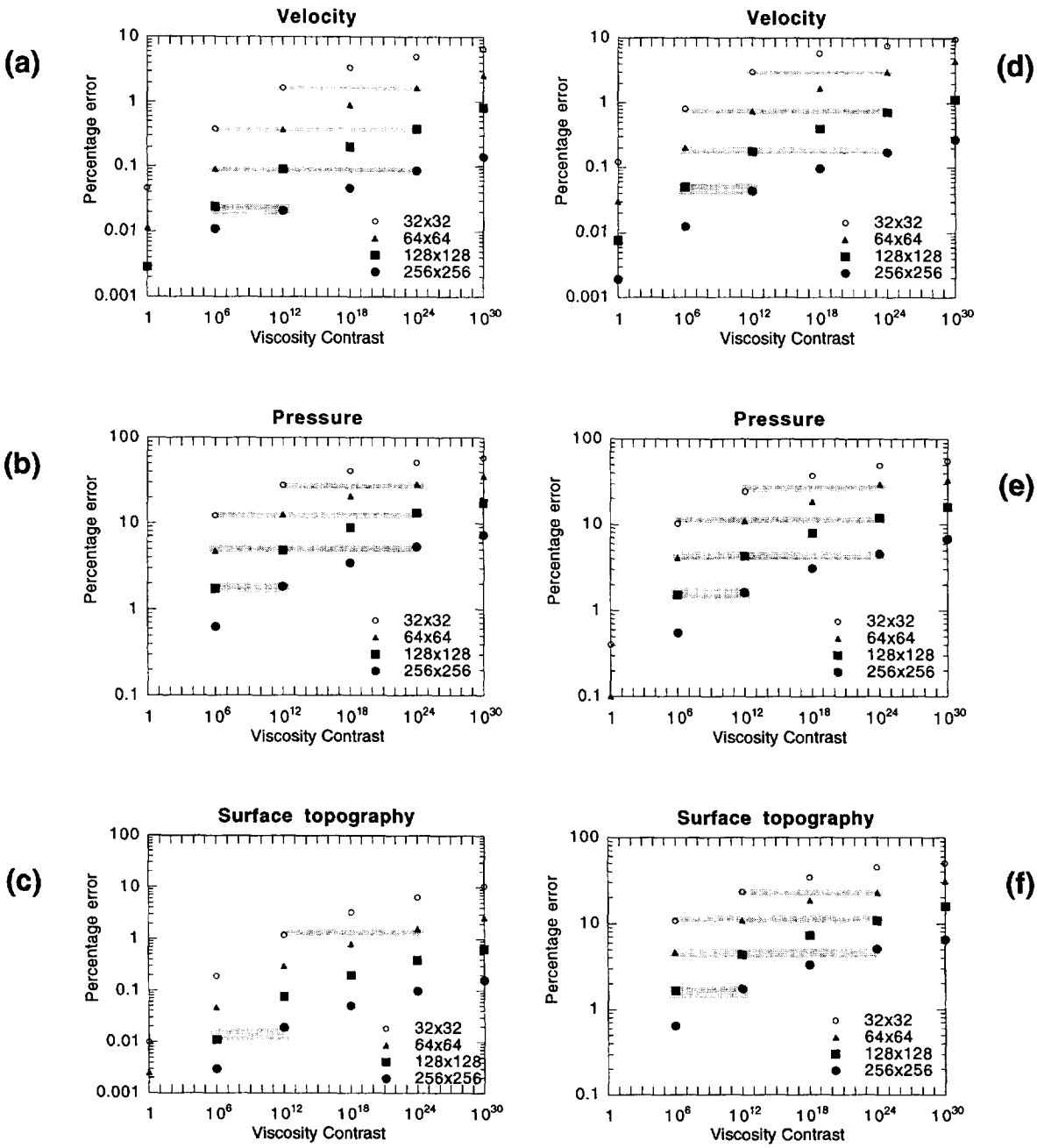

Fig. 2. Errors in velocity, surface topography and dynamic pressure as a function of viscosity contrast and mesh size for viscosity varying exponentially with depth (a–c) and with horizontal distance (d–f). Grey shading underlies calculations with the same relative viscosity jump across individual elements. In (c) the scatter is larger and lighter shading is used to indicate that the shading also overlaps points where the element viscosity jump is different.

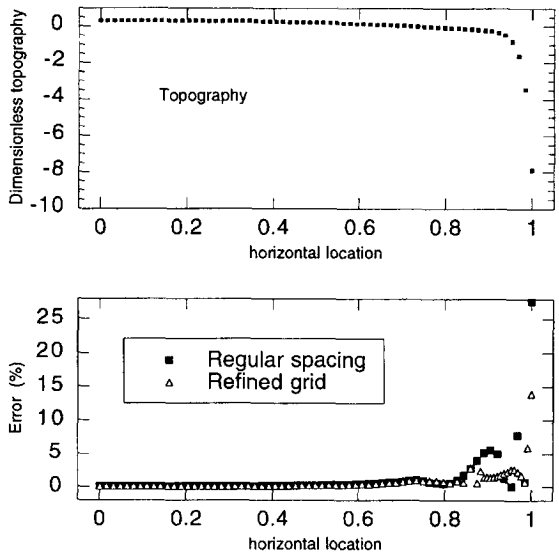

Fig. 3. Surface topography and percentage error in topography as function of position. The viscosity varies exponentially as a function of horizontal position and the total viscosity contrast is 10^{24}. The grids are 64×64 square elements of equal size and 73×64 elements obtained from the regular grid by adding and refining elements in the region between $x = 0.875$ and $x = 1.0$.

velocity fields. In the case with depth-dependent viscosity, the error in the pressure field close to the surface, away from the discontinuous driving buoyancy forces, happens to be comparable with the error in the velocity.

The distribution of errors across the mesh is highly uniform for the velocity solutions. However, the errors in the pressure and surface topography are largest in regions where the viscosity gradients are highest. For a total horizontal viscosity contrast of 10^{24}, Fig. 3 shows analytic topography plotted together with the percentage error in topography at each point on the top of a 64×64 element grid (the relative errors are not calculated near the zero-crossing). The reported error of 22% (Table 2 and Fig. 2(f)) is dominated by the single value at the right-hand end—everywhere else the error is below 5%, which is closer to the velocity error for this solution (3%). The lateral variation in the topography error follows exactly the pattern of errors in the pressure field in the upper part of the layer. A second calculation includes eight additional element so that the spacing of node points is halved between $x = 0.875$ and $x = 1.0$. The relative error in this region is reduced, approximately, by a factor of two, and the errors remain flat much closer to the edge of the box. In the region where the node spacing is unchanged, the topography is indistinguishable from the original solution. Clearly, there are significant benefits, when determining pressure and surface topography, in refining the finite element mesh in regions with high viscosity gradients.

4.1.2. Discontinuous viscosity

The most obvious property of our solutions to problems in which viscosity is discontinuous across element boundaries is that the error in velocity, pressure and surface topography is independent of the viscosity jump. Standard texts such as Becker et al. (1981) suggest that accuracy will be maintained if changes in material moduli coincide with element boundaries, but it has not been clear whether the resulting matrix equations are so ill conditioned that practical solutions could not be obtained without special treatment (Wissmann, 1977). The r.m.s. error listed in the tables is representative of the error in both the high- and low-viscosity regions of the system. Although the ultimate errors remain the same, the convergence rate of the unpenalized pressure iteration degrades rapidly as the viscosity contrast is increased. In this case, the use of the augmented Lagrangian approach is invaluable in obtaining the most accurate pressure and surface topography solutions in a reasonable number of iterations.

As viscosity values are defined on the interiors of the element (at the Gaussian integration points), the viscosity for any element in this calculation is constant, i.e. the relative viscosity jump for every element is unity. If the discontinuity does not lie along an element boundary, at least one row or column of nodes will have the global viscosity contrast associated with individual elements. We performed a calculation with 32×32 elements using a columnar viscosity structure (Case (d)) with a viscosity contrast of 10^6, and moved the interface to $x = 0.5125$, bisecting one column of elements. Compared with the case where the interface was at $x = 0.5$, errors increased from 0.17% in velocity to 15%, and 1% in pressure to 80%. The lesson is clear: when arbitrary zones of strength or weakness are imposed in the domain, they should be assigned to individual ele-

ments. In an Eulerian formulation regridding may be required.

4.2. Error reduction during progress of iteration

We next examine the relation between the residuals of the discrete problem (Eq. (6) and Eq. (7)), and the actual error (Eq. (8)). For iterative schemes, it is important to know the point at which further iteration to reduce the residual no longer reduces the actual error. Fig. 4 shows the relative error in the numerical solution as a function of CPU time in seconds (all calculations were performed on the same machine) for two calculations on a 64×64 mesh with different viscosity contrasts. As the number of Gauss–Seidel relaxations in each multigrid cycle often needs to be increased when the augmented Lagrangian is introduced, the overall CPU time is the best measure of whether there is a benefit obtained by penalizing the problem.

To illustrate the use of the augmented Lagrangian, we first run two calculations without any penalization for incompressibility. Each calculation has an exponential variation of viscosity horizontally with overall viscosity contrast of 10^6 (Fig. 4(a)) and 10^{24} (Fig. 4(b)). At the lower viscosity contrast, the pressure and velocity errors reduce at comparable rates until 20 s have elapsed. At this point, the errors flatten out, and further iteration does not improve the solution despite the continued reduction of the discrete divergence residual r_2 (Eq. (7)). The higher viscosity contrast is more problematic. In this case, the pressure error initially increases dramatically. The velocity solution achieves maximum accuracy after 80 s but the iteration must be continued to 200 s before the pressure error reaches the value obtained by the penalty method. Errors in surface topography are similar to the pressure errors: initially becoming worse before converging to the eventual solution. This situation is wasteful of computer time, for cases where pressure or surface topography are not required, but it also may lead to numerical problems if the pressure residual becomes so large that numerical noise is introduced, in which case the velocity field can be degraded irreparably during the pressure iteration. Unfortunately, this problem becomes most pronounced on the finer grids.

Fig. 5 illustrates the behaviour of the pressure

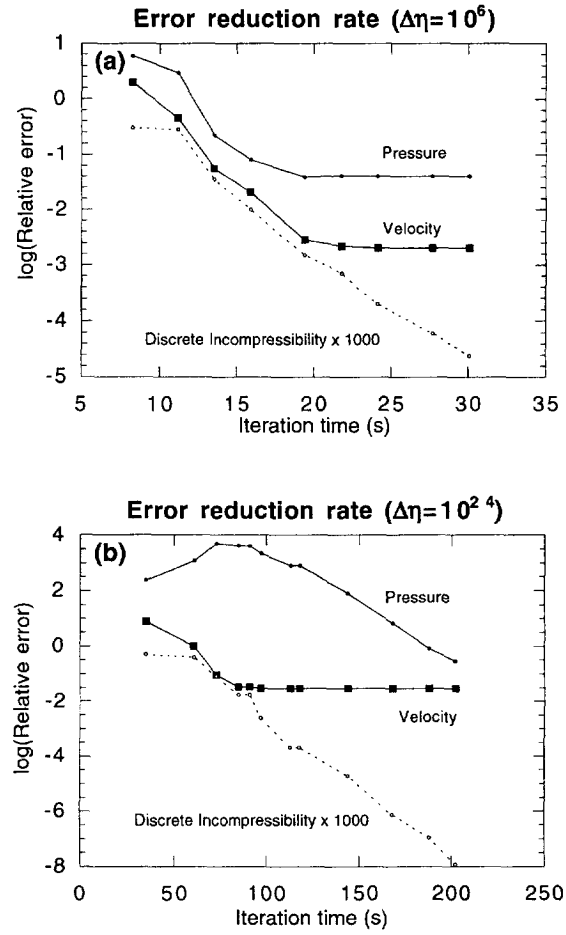

Fig. 4. Relative error reduction for pressure and velocity during progress of iteration for pressure for (a) moderate viscosity contrast, 10^6, and (b) very large viscosity contrast, 10^{24}. The discrete divergence is given by Eq. (7) and is the residual for the discretized matrix equation. After a number of iterations the residual continues to decrease but the true error remains approximately constant.

residual when the problem is penalized using the augmented Lagrangian approach for both calculations shown in Fig. 4—a penalty value of 10^5 was used. In this case, the rate of reduction of the pressure error is almost independent of the viscosity contrast and both calculations flatten out after 60 s. The velocity and topography errors also take the same number of iterations to reach their final values. For low viscosity contrast cases, the augmented Lagrangian method does not offer many clear benefits, and the solution time is likely to increase owing to

the additional iterations required to solve a strongly penalized matrix problem. However, in practical cases such as convection problems where the temperature and viscosity fields are not known in advance, the augmented Lagrangian approach may be the safest approach to take.

4.3. Choice of iteration parameters

To minimize the solution time for a given grid resolution and viscosity variation, the target value of the solution residual must be chosen so that iteration does not continue after the error curves flatten towards the inherent discretization error. Here we can only suggest a conservative rule of thumb. We observed that the target value of the final residual which gave the optimal number of iterations did not change very strongly with viscosity contrast, despite the fact that the solution was not as accurate at high values. We suggest setting the target residual to be an order of magnitude smaller than the best accuracy achieved for velocity or pressure in our tests for each resolution (generally the cases with constant viscosity). The recommended values are listed in Table 4,

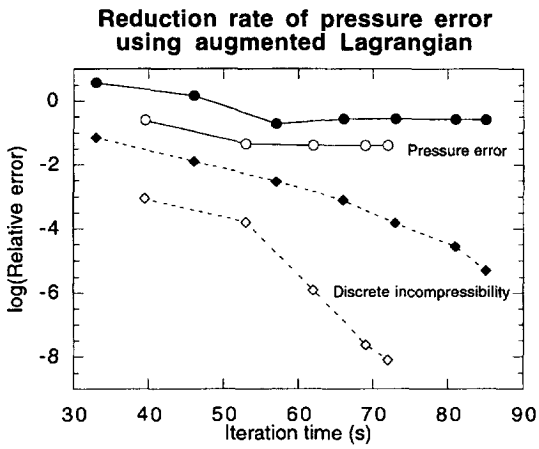

Fig. 5. The same cases as in Fig. 4 but using the augmented Lagrangian method to improve the convergence of the pressure iteration. Open symbols are used for the case with viscosity contrast of 10^6, and filled symbols are used for the case with viscosity contrast of 10^{24}. Errors in the velocity field reduce to their minimum values at the same rate as those for the pressure field. The solution time using the augmented Lagrangian is longer for the calculation with lower viscosity variation but significantly shorter for the high viscosity contrast case.

Table 4
Recommended iteration parameters

Elements	Grid spacing	Target residual	Penalization
32×32	0.03125	3×10^{-5}	3×10^{4}
64×64	0.015625	10^{-5}	7×10^{4}
128×128	0.0078125	3×10^{-6}	10^{5} [a]
256×256	0.0039063	10^{-6}	2×10^{5} [a]

Iteration continues until r_1 and r_2 are both smaller than the target value. The penalization which is recommended is the smallest value which we have found gives a good improvement in convergence of the pressure iteration for large viscosity variations.
[a] A coarse mesh solution or other good estimate of the velocity and pressure fields is required if the velocity iteration is to converge.

and, in each case tested, these values allowed the velocity error to reach its minimum level.

The choice of penalty parameter in the augmented Lagrangian method does not influence the value of the target residual but it must be sufficiently large to allow the pressure error to be reduced to a minimum at the same rate as the velocity error, if pressures are required. We found that values larger than 5×10^4 were usually sufficient, and a value of 10^5 was almost always effective. We expect that the most effective value of the penalization would be the reciprocal of the solution accuracy. Unfortunately, use of these large penalty numbers can affect convergence of the iterative scheme adversely, necessitating prohibitively large numbers of velocity iterations; so values much higher than 10^5 usually lead to much longer solution times. However, we found that the use of interpolated information from a coarse mesh solution allowed the use of significantly larger penalizations, and that the number of velocity iterations was also reduced. This suggests the possibility of developing a very effective full multigrid approach to this problem using augmented Lagrangian methods to achieve the robustness of the standard penalty method for large, three-dimensional calculations.

5. Discussion and conclusions

We have demonstrated that we can use both penalty and mixed iterative finite element formulations to obtain accurate velocity fields for incompressible Stokes' flow problems where there are very

strong variations in viscosity. The error is dictated by the ratio of viscosities associated with individual elements, not with the viscosity contrast across the mesh as a whole. Thus to obtain errors of no more than approximately 1% in the velocity solution, we recommend having no element in the mesh with a viscosity jump greater than approximately two (or, roughly, three elements to span one order of magnitude viscosity variation).

One consequence of this property is that problems which have discontinuities of viscosity along element boundaries can be solved to essentially the same accuracy as problems with constant viscosity. The fact that the errors scale so predictably with element viscosity contrast is useful because it suggests that we can employ grid refinement to minimize this quantity to try to construct an optimally resolved mesh for a given problem.

The surface topography is more difficult to analyse, because the errors are concentrated in regions where the horizontal viscosity gradients at the surface are largest. Clearly, a grid which is optimal for the velocity field may not have sufficient resolution to obtain extremely accurate topography in regions where horizontal viscosity gradients are very high. For mantle convection modelling, such regions would correspond to plate boundaries—the very regions where we most want accurate topography. However, we have demonstrated that the mesh can be refined to obtain improved topography in regions such as plate boundaries where viscosity gradients are high, and errors in the remainder of the domain are unaffected.

Errors in the dynamic pressure are always larger than those in the velocity field, in part because the interpolation is coarser for the pressure field than velocity. Fortunately, we seldom require accurate pressure fields provided that the topography is calculated using the consistent boundary flux method, and provided that non-Newtonian viscosity is formulated in terms of strain rates rather than stresses.

The solution time for the penalty method is independent of the viscosity structure whereas the iterative method can take considerably longer to obtain accurate velocity and especially pressure when viscosity gradients are high. The target value of the finite element residual in the iterative method is a weak function of the viscosity structure, but the individual iteration steps are more time consuming for cases with high viscosity gradients.

We examined the possibility of using a penalty approach with the iterative scheme (the augmented Lagrangian formulation). When viscosity contrasts are relatively small, the pressure and topography errors can be minimized without using any penalization to impose incompressibility. For higher viscosity contrasts, the iterative scheme becomes significantly less efficient at reducing the pressure and topography error, and the application of the augmented Lagrangian formulation is beneficial. The inclusion of this penalty does also affect the solution time for the Gauss–Seidel relaxation so that, despite requiring fewer pressure iterations, the overall solution time may be several times greater than that required to obtain just the velocity field. The optimal choice of penalty number for the augmented Lagrangian method certainly depends on the convergence rate of the iterative method chosen. In our case, we found that values greater than 10^5 gave significantly longer solution times. However, our preliminary results also suggested that a multigrid approach to the pressure problem, in which coarse solutions are interpolated to provide the starting point for higher-resolution calculations, could provide a way to incorporate the augmented Lagrangian method into a robust iterative method.

Acknowledgements

This research was supported by the David and Lucille Packard Foundation and NSF Grants EAR-9496185 and EAR-9417645. This paper is Contribution 5606 of the Division of Geological and Planetary Sciences, California Institute of Technology.

References

Bailey, D.H., 1993. Multiprecision translation and execution of FORTRAN programs. ACM Trans. Math. Software, 19: 288–319.

Baumgardner, J., 1985. Three-dimensional treatment of convective flow in the Earth's mantle. J. Stat. Phys., 39: 501–511.

Becker, E.B., Carey, G.F. and Oden, J.T., 1981. Finite Elements, Vol. 1, an Introduction. Prentice–Hall, Englewood Cliffs, NJ.

Blankenback, B. et al., 1989. A benchmark comparison for mantle convection codes. Geophys. J. Int., 98: 23–38.

Busse, F.H., Christensen, U., Clever, R., Cserepes, L., Gable, C., Giannandrea, E., Guillou, L., Houseman, G., Nataf, H.-C., Ogawa, M., Parmentier, M., Sotin, C. and Travis, B., 1994. 3D convection at infinite Prandtl number in Cartesian geometry—a benchmark comparison. Geophys. Astrophys. Fluid Dyn., 75: 39–59.

Cahouet, J. and Chabard, J.-P., 1988. Some fast 3D finite element solvers for the generalized Stokes problem. Int. J. Numer. Methods Fluids, 8: 869–895.

Christensen, U., 1984. Convection with pressure- and temperature-dependent non-Newtonian rheology. Geophys. J. R. Astron. Soc., 77: 343–384.

Davaille, A. and Jaupart, C., 1993a. Thermal convection in lava lakes. Geophys. Res. Lett., 20: 1827–1830.

Davaille, A. and Jaupart, C., 1993b. Transient high Rayleigh number convection with large viscosity variations. J. Fluid Mech., 253: 141–166.

Fortin, M. and Thomasset, F., 1983. Augmented Lagrangian methods: application to the Stokes and Navier–Stokes equations. In: M. Fortin and R. Glowinski (Editors), Augmented Lagrangian Methods. North-Holland, Amsterdam, pp. 47–95.

Fowler, A.C., 1985. Fast thermoviscous convection. Stud. Appl. Math., 72: 1–34.

Gurnis, M., 1988. Large-scale mantle convection and the aggregation and dispersal of supercontinents. Nature, 332: 695–699.

Hager, B.H. and O'Connell, R.J., 1981. A simple global model of plate dynamics and mantle convection. J. Geophys. Res., 86: 4843–4867.

Hansen, U. and Ebel, A., 1984. Experiments with a numerical model related to mantle convection: boundary layer behaviour of small and large scale flows. Phys. Earth Planet. Inter., 36: 374–390.

Hughes, T.J.R., Liu, W.K. and Brooks, A., 1979. Finite element analysis of incompressible viscous flows by the penalty function formulation. J. Comput. Phys., 30: 1–60.

Karato, S.-I. and Wu, P., 1993. Rheology of the upper mantle: a synthesis. Science, 260: 771–778.

Karato, S.-I., Paterson, M.S. and Fitzgerald, J.D., 1986. Rheology of synthetic olivine aggregates: influence of grain size and water. J. Geophys. Res., 91: 8151–8176.

King, S.D. and Hager, B.H., 1990. The relationship between plate velocity and trench viscosity in Newtonian and power-law subduction calculations. Geophys. Res. Lett., 17: 2409–2412.

King, S.D., Raefsky, A. and Hager, B.H., 1990. CONMAN—vectorizing a finite-element code for incompressible 2-dimensional convection in the Earth's mantle. Phys. Earth Planet. Inter., 59: 195–207.

Kopitze, U., 1979. Finite element convection models: comparison of shallow and deep mantle convection, and temperatures in the mantle. J. Geophys., 46.

Monnereau, M., Rabinowicz, M. and Arquis, E., 1993. Mechanical erosion and reheating of the lithosphere: a numerical model for hotspot swells. J. Geophys. Res., 98: 809–823.

Moresi, L. and Solomatov, V.S., 1995. Numerical investigations of 2D convection with extremely large viscosity contrasts. Phys. Fluids, 7: 2154–2162.

Revenaugh, J. and Parsons, B., 1987. Dynamic topography and gravity anomalies for fluid layers whose viscosity varies exponentially with depth. Geophys. J. R. Astron. Soc., 90: 349–368.

Ribe, N.M. and Christensen, U.R., 1994. Three-dimensional modeling of plume–lithosphere interaction. J. Geophys. Res., 99: 669–682.

Richards, M.A., Hager, B.H. and Sleep, N.H., 1988. Dynamically supported geoid highs over hotspots: observation and theory. J. Geophys. Res., 93: 7690–7708.

Schmeling, H. and Jacoby, W.R., 1981. On modelling the lithosphere in mantle convection with non-linear rheology. J. Geophys., 50: 89–100.

Schubert, G. and Anderson, C.A., 1985. Finite element calculations of very high Rayleigh number thermal convection. Geophys. J. R. Astron. Soc., 80: 575–601.

Solomatov, V.S. and Moresi, L., 1996. Stagnant lid convection on Venus. J. Geophys. Res., 101: 4737–4753.

Travis, B., Anderson, C., Baumgardner, J., Gable, C.W., Hager, B.H., O'Connell, J., Olson, P., Raefsky, A. and Schubert, G., 1990. A benchmark comparison of numerical methods for infinite Prandtl number thermal convection in two-dimensional Cartesian geometry. Geophys. Astrophys. Fluid Dyn., 55: 137–160.

Van Keken, P., Yuen, D.A. and van den Berg, A., 1992. Pulsating diapiric flows: consequences of vertical variations in mantle creep laws. Earth Planet. Sci. Lett., 112: 179–194.

Wissmann, J.W., 1977. A rigid body transformation algorithm for structures with large stiffness variations, with an application to incompressible elements. In: K.-J. Bathe, J.T. Oden and W. Wunderlich (Editors), Formulations and Computational Algorithms in Finite Element Analysis: US–Germany Symposium. Massachusetts Institute of Technology, Cambridge, MA, pp. 109–130.

Zhong, S., 1996. Analytic solutions for Stokes' flow with lateral variations in viscosity. Geophys. J. Int., 124: 18–28.

Zhong, S. and Gurnis, M., 1995. Towards a realistic simulation of plate margins in mantle convection. Geophys. Res. Lett., 22: 981–984.

Zhong, S., Gurnis, M. and Hulbert, G., 1993. Accurate determination of surface normal stress in viscous flow from a consistent boundary flux method. Phys. Earth Planet. Inter., 78: 1–8.