Geophys. Astrophys. Fluid Dynamics, Vol. 80, pp. 57–74
Reprints available directly from the publisher
Photocopying permitted by license only

© 1995 OPA (Overseas Publishers Association)
Amsterdam B.V. Published under license by
Gordon and Breach Science Publishers SA
Printed in Singapore

DASPK: A NEW HIGH ORDER AND ADAPTIVE TIME-INTEGRATION TECHNIQUE WITH APPLICATIONS TO MANTLE CONVECTION WITH STRONGLY TEMPERATURE-AND PRESSURE-DEPENDENT RHEOLOGY

PETER E. VAN KEKEN

Department of Geological Sciences, University of Michigan, 1006 CC Little Building, Ann Arbor MI 48109

DAVID A. YUEN

Minnesota Supercomputer Institute, University of Minnesota, 1200 Washington Ave S, MN 55415 Minneapolis

LINDA R. PETZOLD

Computer Science Department & Army High Performance Computing Research Center, University of Minnesota, 1100 Washington Ave S, MN 55415 Minneapolis

(*Received 8 April 1994; in final form 26 January 1995*)

A new technique is presented for the efficient time-integration of the equations that describe the slow deformation in the Earth's mantle. This method is based on the adaptive, high order implicit solver for differential-algebraic equations (DASPK) and is independent of the choice of spatial discretization technique. Using a standard finite element package for the spatial discretization, it is shown that the solution of the 2-D convection-diffusion equation for temperature can be performed at much lower computational cost, but at the same or higher accuracy, compared to a traditional implicit second-order method. The solution to the full set of 2-D mantle convection equations is 3 to 4 times more efficient. Both in 2-D and 3-D, the memory and CPU-usage of this implementation depends linearly on the number of grid points and has good properties with respect to vectorization and parallelization. As an application of this technique, convection in the Earth's mantle with strongly temperature and pressure dependent rheology is studied in axisymmetric geometry. Models are developed that are consistent with current estimates of surface heat flow and radial viscosity distribution. General characteristics are: a dynamic upper mantle overlying a near-stationary lower mantle; strong plumes rising from the core-mantle boundary, even at high rates of internal heating; and an effective Rayleigh number of nearly two orders of magnitudes lower than commonly used values in the range of 10^7 to 10^8.

KEY WORDS: Mantle convection, mantle rheology, time-integration techniques.

1. INTRODUCTION

In computational geodynamics we are currently facing the task of finding efficient and accurate methods that allow for the solution of the full 3-D, time-dependent and

E-mail addresses: [1]keken @umich.edu; [2]davey@kissy.msi.umn.edu; [3]petzold@cs.umn.edu.

variable coefficient mantle dynamics equations using present-day computer capacity. Presently available time-dependent 3-D mantle convection codes use in general explicit and low order time-integration schemes for the heat equation and are optimized for special geometries, such as rectangular boxes or spherical shells (e.g., Gable, 1989; Travis *et al.*, 1990; Houseman, 1990; Christensen and Harder, 1991; Tackley, 1993; Parmentier *et al.*, 1994; Bunge and Baumgardner, 1995). In only a few applications third or higher order schemes are used to perform part of the computation (e.g., Malevsky and Yuen, 1991; Balachandar and Yuen, 1994). Ideally, we would like to use in the 3-D situation the implicit methods for general geometries that allow today for the accurate solution of 2-D problems. In comparison with these 2-D methods, it is then essential to reduce computational costs while maintaining (or improving) the accuracy and flexibility. In the most optimal sense, these methods should i) be implicit, ii) be able to handle general geometries, iii) not require major software development or recoding from the user and iv) be able to run both on general (e.g., workstation) and specialized (vector or parallel) machine architecture.

Here, we focus on the time-integration of the heat equation and describe a toolbox approach that combines the flexibility of finite element methods (using the general finite element package SEPRAN) and a solver for differential-algebraic equations (DASPK), that combines backward differentiation formula methods (up to order 5) and Krylov iterative methods. We will show that this method is more efficient than when using a traditional implicit time-integration method for 2-D geometries, can be efficiently extended to 3-D situations and has furthermore excellent properties with respect to vectorization and parallelization. The new aspect of this method is the high order time-integration technique that has wider application as it is independent of the spatial discretization method: it can be applied to finite difference and spectral methods, that can provide a more efficient spatial discretization than finite element methods.

First, a short description of the adaptive high order solver for differential algebraic equations (DASPK) and a general finite element package (SEPRAN) is given. Next, a comparison is given between DASPK and a commonly used second order predictor-corrector method for the solution of the 2-D and 3-D convection-diffusion equation. This is then applied to the solution of the 2-D free convection equations. As an application, models of mantle convection in axisymmetric geometry with strongly temperature- and depth-dependent rheology are presented.

2. DESCRIPTION OF DASPK AND SEPRAN

DASPK

DASPK (Maier and Petzold, 1993; Brown *et al.*, 1995) is a general solver for differential-algebraic equations (DAE's) of the form

$$\mathbf{F}(t, \mathbf{y}, \mathbf{y}') = \mathbf{0} \tag{1}$$

where $\mathbf{F}$, $\mathbf{y}$ and $\mathbf{y}'$ are N-dimensional vectors, and a consistent set of initial conditions $\mathbf{y}(t_0) = \mathbf{y}_0, \mathbf{y}'(t_0) = \mathbf{y}'_0$ is given, i.e., $\mathbf{F}(t_0, \mathbf{y}_0, \mathbf{y}'_0) = \mathbf{0}$. Of particular interest here is the form

of (1) that arises from partial differential equations after spatial discretization. Integration of (1) over a specified range of t is accomplished by replacing the derivatives with difference approximations and using a predictor-corrector method. An initial guess for the new solution is developed by evaluating a predictor polynomial, which interpolates solution values at previous time steps. The new solution is then computed in the corrector step by solving a nonlinear system of equations. The predictor and corrector polynomials are specified using backward difference formulae (BDF's) up to order 5. The method is adaptive: it selects step-size and integration order, depending on the temporal changes in the solution y and a user-specified error control (see below). A complete description of the integration method is given in Brenan *et al.* (1989). In each corrector step a sequence of nonlinear systems

$$F(t, \mathbf{y}, \alpha\mathbf{y} + \beta) \tag{2}$$

is solved by a Newton-type iteration, which requires the solution of a linear system $\mathbf{Cx} = \mathbf{d}$, where the iteration matrix $\mathbf{C}$ is given by

$$\mathbf{C} = \frac{\partial \mathbf{F}}{\partial \mathbf{y}} + \alpha \frac{\partial \mathbf{F}}{\partial \mathbf{y}'}. \tag{3}$$

DASPK is part of the DASSL family (Brenan *et al.*, 1989), that uses a direct method for the linear system solution. In DASPK there is the additional option to solve the linear systems approximately by a preconditioned generalized minimum residual (GMRES) method (Saad and Schultz, 1986). For further details, see Brown *et al.*, 1995; Maier and Petzold, 1993.

This algorithm has been implemented for massively parallel distributed memory architecture, using both data parallel (SIMD) and message passing (MIMD) paradigms (Maier *et al.*, 1995).

SEPRAN

For the spatial discretization of the mantle convection equations we use the finite element package SEPRAN (Segal and Praagman, 1984; Cuvelier *et al.*, 1986), that has been developed at the Technical University Delft in the Netherlands. It is a toolbox package, from which general routines are available for mesh generation, matrix assembly, direct and iterative solution and postprocessing. Several standard differential equations can be solved for 2-D and 3-D geometries, of which the full Navier-Stokes and convection-diffusion equations are of specific geophysical interest. This package has been applied to a variety of geodynamical problems (Van den Berg *et al.*, 1991; Van Keken *et al.*, 1993; Vlaar *et al.*, 1994). Here SEPRAN is used to spatially discretize the temperature equation and to form the system of ordinary differential equations (ODE's), which are then solved using DASPK. This finite element package is commercially available (Sepra, Veursestraatweg 77, 2264 EB Leidschendam, The Netherlands).

3. DASPK FOR THE SOLUTION OF THE 2-D AND 3-D CONVECTION-DIFFUSION EQUATION

In this section the basic technique will be discussed with applications to the solution of the 2-D and 3-D time-dependent convection-diffusion equations.

Two-dimensional forced convection

Consider the 2-D forced advection-diffusion equation

$$\frac{\partial T}{\partial t} + (\mathbf{v}\cdot\nabla)\, T = \frac{1}{Pe}\nabla^2 T + q \qquad \text{on} \qquad \Omega = [0,1] \times [0,1], \tag{4a}$$

with initial condition

$$T(t=0) = 1 - y \tag{4b}$$

and boundary conditions

$$T = 1 \text{ on } y = 0; \qquad T = 0 \text{ on } y = 1; \qquad \partial T/\partial x = 0 \text{ on } x = 0 \text{ and } x = 1. \tag{4c}$$

The velocity is prescribed by

$$\mathbf{v} = [-\psi(t) \sin \pi x \cos \pi y, \psi(t) \cos \pi x \sin \pi y]^T \tag{5}$$

and the Peclet number is fixed at $Pe = 100$, q is the volumetric heat production rate.

The partial differential Equation (4a) is descretized on a equidistant mesh of linear triangles using SEPRAN. This gives a system of ordinary differential equations

$$\mathbf{M}\dot{\mathbf{T}} + \mathbf{S}\mathbf{T} = \mathbf{f} \tag{6}$$

where $\mathbf{M}$ is the mass matrix, $\mathbf{T}$ the vector containing the degrees of freedom in the nodal points, $\mathbf{S}$ the stiffness matrix (the dot denotes time derivative) and $\mathbf{f}$ is the discretized form of the source term q. The vector $\mathbf{T}$ is partitioned in unprescribed and prescribed degrees of freedom $\mathbf{T} = [\mathbf{T}_f, \mathbf{T}_p]^T$. The stiffness and mass matrices can be partitioned in a similar way, allowing for elimination of $\mathbf{T}_p$, which gives the smaller system of equations

$$\mathbf{M}_f\dot{\mathbf{T}}_f + \mathbf{S}_{ff}\mathbf{T}_f = \mathbf{f} - \mathbf{S}_{fp}\mathbf{T}_p. \tag{7}$$

Starting from given initial conditions $\mathbf{T}_0$ and $\dot{\mathbf{T}}_0$, DASPK integrates the discretized convection-diffusion Equation (6) over an interval that is specified by the user. During the integration, DASPK makes estimates of the time-derivative, $\dot{\mathbf{T}}_D$, for given t and $\mathbf{T}$ and requests the numerical value of the residual Δ to (1), which can be written in this

case as

$$\Delta = \mathbf{M}_f \dot{\mathbf{T}}_{f,D} + \mathbf{S}_{ff}\mathbf{T}_f + \mathbf{S}_{fp}\mathbf{T}_p - \mathbf{f}. \tag{8}$$

Implementation of the algorithm in the numerical code is straightforward. For the calculation of Δ(8) only matrix-vector and vector-vector multiplications are used. The matrix S can be stored in compressed format (that is, only the non-zero entries in the matrices are stored). For linear elements the mass matrix M can be made diagonal by lumping.

The algorithm and implementation given above have been compared with a popular second-order implicit predictor-corrector method to solve the heat equation in mantle convection modeling (method PC), that combines an implicit Euler predictor with a Crank-Nicolson corrector (e.g., Christensen, 1984; Hansen and Ebel, 1988; Van Keken, 1993). The accuracy of this method is governed by the time-step, which is limited to be a certain fraction of the Courant-Friedrichs-Levy criterion, Δt_{CFL}. This criterion gives the maximum time-step that can be used in explicit methods. The use of this criterion is essential to guarantee stability of explicit methods. In general, a certain fraction of the CFL-time-step is chosen even when using implicit methods to guarantee sufficient accuracy (e.g., Christensen, 1984). This predictor-corrector method generates a matrix-vector equation

$$\mathbf{AT} = \mathbf{b}, \tag{9}$$

which is then factored and solved. The matrices S and A are stored using a profile method (Cuvelier *et al.*, 1986). A renumbering algorithm by Sloan (1986) has been used to reduce the profile.

In the DASPK method (method D), the finite element stiffness matrix is stored in compressed sparse row storage format. The accuracy of the DASPK method is governed by an accuracy parameter *RTOL*, which is normally chosen to be much smaller than 1. DASPK can be used with both the direct solution technique and the Krylov subspace method (see Section 2). It turns out that for the solution of the convection-diffusion equation the direct solution technique provide sufficient accuracy and efficiency. Therefore the direct method has been used for the examples presented in this paper. A tri-diagonal approximation to the iteration matrix C, (3) gives sufficiently accurate results for both the 2-D and 3-D heat equation. The approximation is generated by lumping via finite-difference approximations using the methods of DASSL (Brenan *et al.*, 1989). The computational time of the direct solution method is similar to that of the iterative methods of DASPK, but there is a substantial reduction in memory requirements. In this particular application, the banded approximation is a sufficient approximation to the Jacobian to allow the Newton iteration to converge. Note that this may not work for all problems, in which case the more powerful iterative method can be employed.

In the predictor-corrector method PC, two copies of the stiffness matrix are used. Method D uses only one copy of a densely packed array. DASPK requires additional memory to store the solution at different time-levels (for the polynomial predictor), and the tri-diagonal approximation to the iteration matrix. In our implementation, the

maximum order of the BDF is limited to five. Under these conditions, DASPK needs storage for approximately $12\,N_{eq}$ double precision variables, where N_{eq} is the number of unknowns in the temperature equation. These storage requirements can be reduced by limiting the maximum order of the BDF. In the 2-D situation, the elements of the compressed stiffness matrix can be stored in $7\,N_{eq}$ unknowns. In addition to this, integer information on the location (i.e row and column number) of each matrix element has to be stored, which leads to extra memory usage of approximately $2 \times 7\,N_{eq}$ integers. Table I gives an indication of the minimum amount of memory (in megawords, one word = 8 bytes) that is required in the two methods.

The velocity field is generated using the Lorenz equations (Lorenz, 1963) in a parameter range that guarantees a non-periodic solution. The non-dimensional parameter r represents the Rayleigh number and indicates the vigor of the flow. A typical solution using $r = 50$ is shown in Figure 1. These type of solutions are used to describe the velocity (5). The differential Equation (4a) is now a linear equation with time varying coefficients. Figure 2 shows the temporal evolution between $t = 0$ to $t = 1$ of the temperature field for $r = 50$. Mesh size is 40×40. Note the strong spin-up effect in the beginning of the calculation, and the later change between importance of diffusion and advection, which depends on the magnitude of velocity (Fig. 1). To test the accuracy

Table I Memory (in megawords, 1 MW = 8 Mbyte) required for the storage of the matrix arising from the discretization of the sample 2-D advection diffusion equation. PC is the Sepran method that uses a profile method for storage. D is the DASPK method, using compressed row storage format. N_{eq} is the number of equations.

Mesh	N_{eq}	*PC*	*D*
20 × 20	361	0.034	0.013
40 × 40	1521	0.265	0.055
60 × 60	3481	0.885	0.125
80 × 80	6241	2.086	0.225

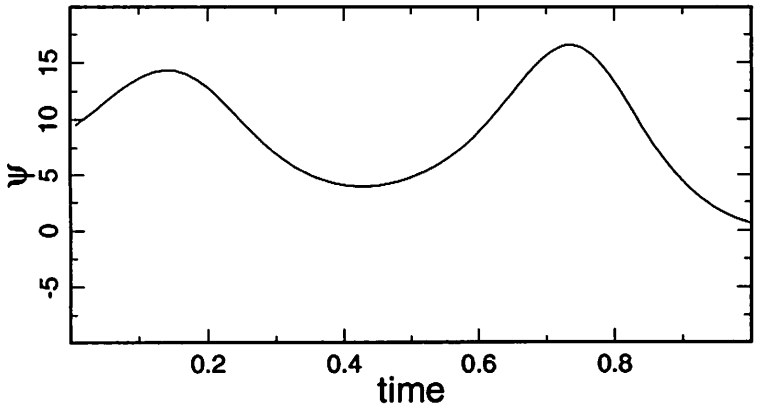

Figure 1 Solution Ψ of the Lorenz Equations (10a–c) $Pr = 10$, $b = 8/3$, $r = 50$.

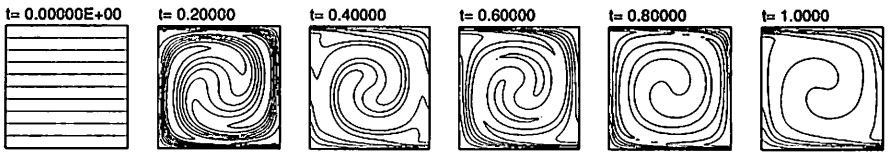

Figure 2 Evolution of temperature in forced convection problem with $r = 50$, $Pe = 100$.

and efficiency of the different methods, we have compared the values of the Nusselt number (normalized heat flux out of the upper surface) at $t = 0.1$ for different values of r and mesh discretizations. Higher values of r indicate stronger convective motions. The accuracy parameter RTOL is varied between 10^{-3} and 10^{-6}. In the fixed time-step method PC the accuracy is governed by the fraction $a = \Delta t/\Delta t_{CFL}$, which is varied between 1 and 0.001. Figure 3 shows the Nusselt number (Nu) as a function of CPU-time. The results indicate that the DASPK method is particularly efficient for this kind of spin-up problem. In general, a convergent solution is obtained nearly 100 times faster compared with the other methods. DASPK achieves very high accuracy by using high order difference schemes in the calculation. Figure 4 shows the order at each time-step, as chosen by DASPK for the calculation with $r = 50$ and mesh 40×40. It may be expected that this aspect of the DASPK algorithm can lead to a large reduction in CPU time for convection calculations with non-Newtonian rheology, where this kind of spin-up frequently occurs (Van Keken *et al.*, 1992; Van den Berg *et al.*, 1993).

Three-dimensional convection-diffusion equation

In the formulation given above, the solution of the heat equation requires only matrix-vector multiplications, which allows for a compact storage of the matrix. This creates the possibility of using this method in 3-D models, where use of the profile

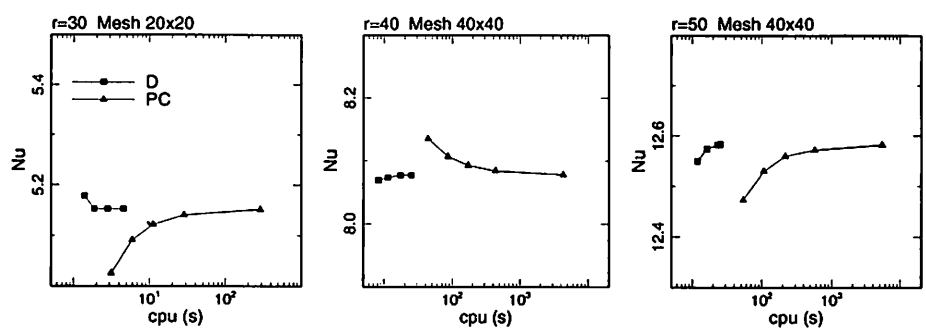

Figure 3 Nusselt number at $t = 0.1$ as function of CPU-time for varying r and mesh discretization. The accuracy parameter for the DASPK methods D is varied between 10^{-3} and 10^{-6}. The accuracy of method PC is governed by the time-step, which is given as a fraction a of the Courant time-step ($\Delta t = a\Delta t_{cfl}$), where $a = 1, 0.5, 0.25, 0.1$ and 0.01. These calculations are performed using 64-bit arithmetic on a Silicon Graphics workstation with a MIPS R4400 @ 150 Mhz processor, specified by the vendor to have a maximum speed of 25 Mflops. Using a standard Linpack double precision benchmark we find a speed of 18 Mflops for the inversion of a 300×300 matrix, when using the vendor supplied BLAS-library.

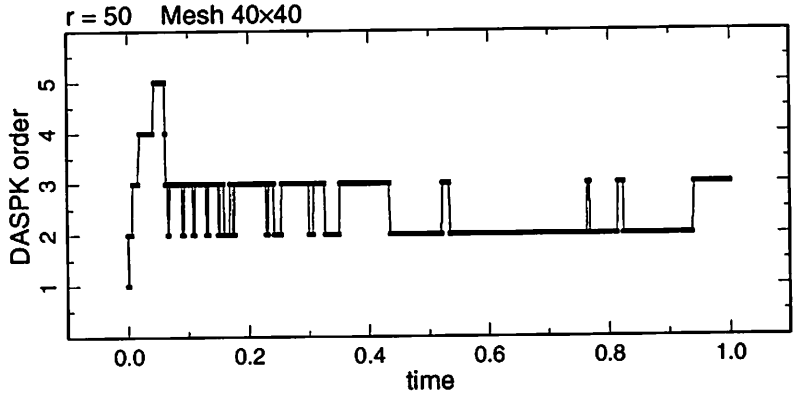

Figure 4 DASPK order of finite difference scheme shown at 687 time-steps for $r = 50$. Mesh is 40 × 40, $RTOL = 10^{-4}$.

method becomes prohibitively expensive. To illustrate the extension to 3-D, we have considered a simple forced convection problem in a rectangular box in which a constant velocity is prescribed.

Consider 3-D forced convection in a unit cube, using the same boundary conditions as in the 2-D case: symmetry conditions along the vertical sides, and constant temperature at top ($T = 0$) and bottom ($T = 1$). The initial condition is given by the conductive solution. The velocity is now prescribed by

$$\mathbf{v} = (-100 \sin \pi x \cos \pi y \cos \pi z,\ 100 \cos \pi x \sin \pi y \cos \pi z,\ 100 \sin \pi x \cos \pi y \cos \pi z)^T. \tag{10}$$

This is compared with the 2-D model [equations (4) and (5)], with $\psi = 100$ and $Pe = 1$. The heat equation is integrated from $t = 0$ to $t = 0.01$, where a relative tolerance of $RTOL = 10^{-4}$ for the DASPK-method.

The advection-diffusion equation is discretized using the finite elements method. For the 3-D problem linear tetrahedral elements are used. In 3-D, the connectivity of each nodal-point is larger, therefore leading to a larger number of non-zeroes in each row of the stiffness matrix. The storage requirements are doubled with respect to the 2-D case, and now $14\,N_{eq}$ double precision variables are needed to store the real matrix coefficients in compressed format. In the case of the free convection, the matrix would have to be assembled at each time-step. Table II lists some performance statistics for both 2-D and 3-D models at different discretizations. Shown are the total CPU-time minus the time spent in the matrix assembly process (CPU_{solve}), number of equations (N_{eq}), number of time-steps (N_t), and CPU-time per time-step per equation. The CPU-time spent in the actual solution of the heat equation after spatial discretization is an approximately linear function of number of equations (= nodal points) and number of time-steps, both in 2-D and 3-D. If the computational costs per nodal point are compared, the 3-D method is more expensive than the 2-D equivalent. The matrix assembly is twice as expensive (because of the higher connectivity) and the actual

Table II Comparison of performance of DASPK solution to 2-D and 3-D convection-diffusion equation. Shown are as a function of discretization, the total CPU-time minus the time used for matrix assembly (CPU_{solve}), the number of equations (N_{eq}), the number of time-steps (N_t) needed to integrate the solution from $t = 0$ to $t = 0.1$, and the CPU-time per equation per time-step. See the caption of Figure 3 for details on the CPU speed.

Discretization	CPU_{solve}(*seconds*)	N_{eq}	N_t	$CPU_{solve}/(N_{eq} * N_t)$ (*seconds*)
10^2	0.17	99	35	5.5×10^{-5}
20^2	0.56	399	45	3.1×10^{-5}
25^2	1.13	624	62	2.9×10^{-5}
30^2	2.05	899	88	2.6×10^{-5}
40^2	5.74	1599	155	2.3×10^{-5}
50^2	13.6	2499	247	2.2×10^{-5}
60^2	29.2	3599	365	2.2×10^{-5}
10^3	1.7	1089	30	5.2×10^{-5}
20^3	23.3	8379	67	4.1×10^{-5}
25^3	58.6	16224	98	3.6×10^{-5}
30^3	132.1	27869	138	3.7×10^{-5}
40^3	556.8	65559	247	3.4×10^{-5}

DASPK solution is 1.5 times more expensive (in large part because of the larger number of time-steps: compare for example N_t for 30^2 and 30^3 in Table II).

4. FREE CONVECTION APPLIED TO THE EARTH'S INTERIOR

The slow deformation of the Earth's mantle can be described by ductile creep of solid mantle silicates under high pressure and temperature conditions. Mantle convection can approximately be described by the non-dimensional equations of conservation of mass, momentum and thermal energy for a Boussinesq fluid

$$\nabla \cdot v = 0, \tag{11a}$$

$$-\nabla P + \nabla \cdot (\eta \dot{\varepsilon}) = Ra\, T \hat{\mathbf{z}}, \tag{11b}$$

$$\frac{\partial T}{\partial t} + (\mathbf{v} \cdot \nabla)\, T = \nabla^2 T + q, \tag{11c}$$

where P is dynamic pressure, η is viscosity, $\dot{\varepsilon}$ is the strain rate tensor ($\dot{\varepsilon} = \nabla \mathbf{v} + \nabla \mathbf{v}^T$), Ra the Rayleigh number and q the internal heat production rate. The equations of motion contain no inertial terms as a consequence of the very high viscosity of the Earth's interior. The viscosity η is given here as a generalized Newtonian viscosity, which for the Earth's mantle is a strong function of temperature, pressure and stress.

The set of Equations (11a–c) is normally solved by integrating (11c) in time and finding the instantaneous velocity field $\mathbf{v}$ at each time-step from the elliptic system (11a–b). Here, the equations of motion and the incompressibility constraint are solved

using a penalty function method (Cuvelier *et al.*, 1986). Without loss of generality, this solution method can be replaced by other approaches, such as spectral or multigrid methods. The penalty function solution method uses quadratic finite elements, which are defined using the same nodal point distribution as used for the temperature equation. This leads to a nested mesh, where each quadratic triangular element for the Stokes equations contains four linear quadratic elements for the heat equation (Van den Berg *et al.*, 1993). Because of the larger number of degrees of freedom in the Stokes problem, the size of the Stokes stiffness matrix dominates the memory usage. For the heat-equation the memory used in the DASPK method is strongly reduced with respect to the PC-method as described above and this leads to an overall memory reduction of 30%. The set of equations is integrated by the second-order predictor-corrector (PC) method (see Section 3):

Specify initial temperature distribution.
Find velocity at this time.
Integrate until end,
  determine time-step, taken to be 50% of the CFL time-step,
  predict temperature at new time by an implicit Euler method,
  find velocity at this time given the predicted temperature,
  correct temperature field by the Crank-Nicolson scheme,
  correct velocity field.

We tested the performance of DASPK in solving (11a–c) in comparison with the fixed order PC method. The integration was started by providing DASPK with the system of Equations (8), where the matrices were built using the instantaneous velocity at time $t = 0$. Each time that DASPK changes to a new time-level, the velocity field was recalculated and the Equation (8) rebuilt. We show examples of convection with strongly temperature- and pressure-dependent rheology. The Stokes equation is computationally the most difficult, especially with variable viscosity. Then, the matrix for the Stokes equation varies in time and it has to be assembled and decomposed each time the Stokes equation has to be solved.

In the examples shown below, we are particularly interested in the rheology of the lower mantle and its influence on the dynamics of the mantle. The viscosity η is described by the Weertman formulation of the Arrhenius rheology

$$\eta = \eta_0 \exp\left(\gamma \frac{T_m}{T}\right) \tag{12}$$

where T_m is the melting temperature, and η_0 and γ parameters describing the visocity distribution (Weertman, 1970). The activation temperature T_m represents here both the temperature- and pressure-dependence of viscosity. The composition of the lower mantle is most likely an assemblage of perovskite and magnesiowüstite. Recent static high pressure melting experiments to 65 GPa indicate a very high melting temperature for perovskite of nearly 8000 K near the coremantle boundary (CMB) (Zerr and Boehler, 1993; Shen and Lazor, 1995). Similar type of experiments on magnesiowüstite gives an extrapolated melting temperature of 5000 K at the CMB (Zerr and Boehler, 1994). Recent shock wave experiments on olivine give a slightly lower melting

temperature of 4500 K at the CMB (Gallagher and Ahrens, 1994), which possibly indicates that magnesiowüstite is the mineral controlling the melting temperature of the lower mantle. All these experiments indicate a strong increase of melting temperature and consequently of activation temperature with depth. For thermally controlled creep processes in the lower mantle this leads to both a high viscosity and a stronger temperature dependence. Van Keken *et al.* (1994a) gave a preliminary investigation into the implications of this strong increase of melting temperature on the dynamics of the lower mantle using a simple steady-state, Cartesian model. Their main conclusion was that this type of viscosity can lead to a radial viscosity profile that has a distinct maximum of 10^{23} Pa·s at approximately 2000 km depth. At larger depths the viscosity drops to a value of 10^{22} Pa·s at the CMB. This viscosity profile is similar to that found from the simultaneous inversion of geoid, topography, moment of inertia and plate velocities (Ricard and Bai Wuming, 1991; Ivins, *et al.*, 1993; Corrieu *et al.*, 1994).

Here, we investigate this further using axisymmetric geometry to simulate thermal convection in a spherical Earth. Figure 5 shows the model geometry. The mesh of

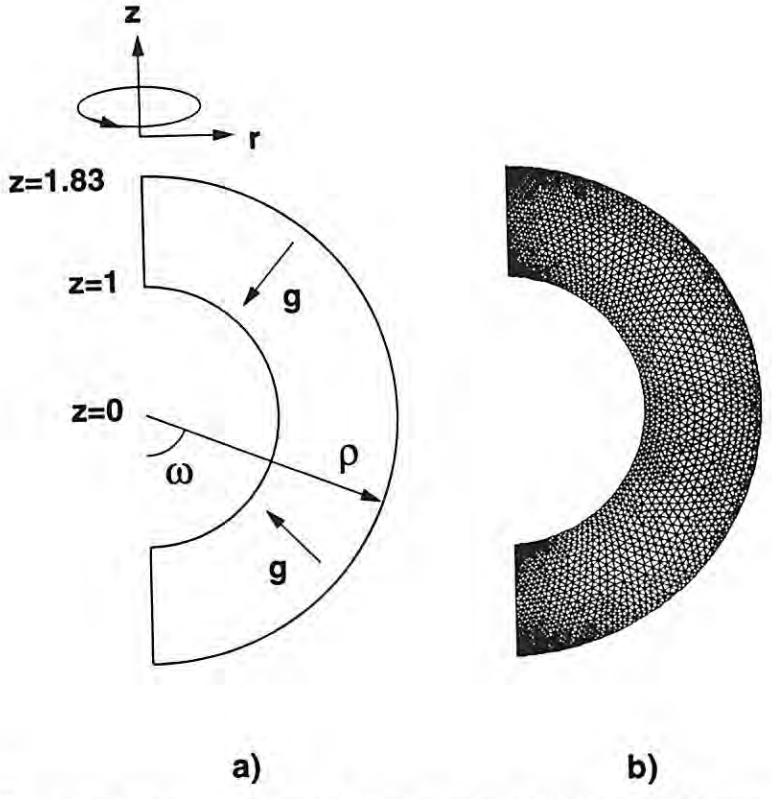

Figure 5 Geometry and computational mesh for convection models in axisymmetric coordinates. The spherical shell model has rotational symmetry around an axis through the poles (left frame). A typical mesh of quadratic triangles is shown in the right frame. The mesh of linear triangles for the heat equation is four times as dense.

quadratic triangles that has been employed for low Rayleigh number calculations is shown in Figure 5b. The use of finite elements now allows for refinement near the boundary layers and the poles of the model.

The viscosity parameters η_0 and γ (12) are chosen such that the average radial viscosity profile is similar to that obtained from the geodetic inversions, with an upper mantle viscosity of 10^{21} Pa·s and maximum viscosity in the lower mantle of 10^{23} Pa·s. The heat production rate q is based on approximately 1.5 times the value given by the chondritic abundance for radiogenic elements, to simulate a primitive lower mantle that is enriched in U and Th. The Rayleigh number Ra is chosen in a range that gives reasonable values of heat flow at the Earth's surface. Present day average heat flow out of oceans is approximately 100 mW/m^2, which corresponds to a Nusselt number of 15–20, based on a CMB-temperature of 4000 K.

Figure 6 shows part of the time evolution of an axisymmetric model with this strongly temperature- and pressure-dependent rheology. The top row shows the stream function ψ, the bottom row temperature T, both as non-dimensional quantities.

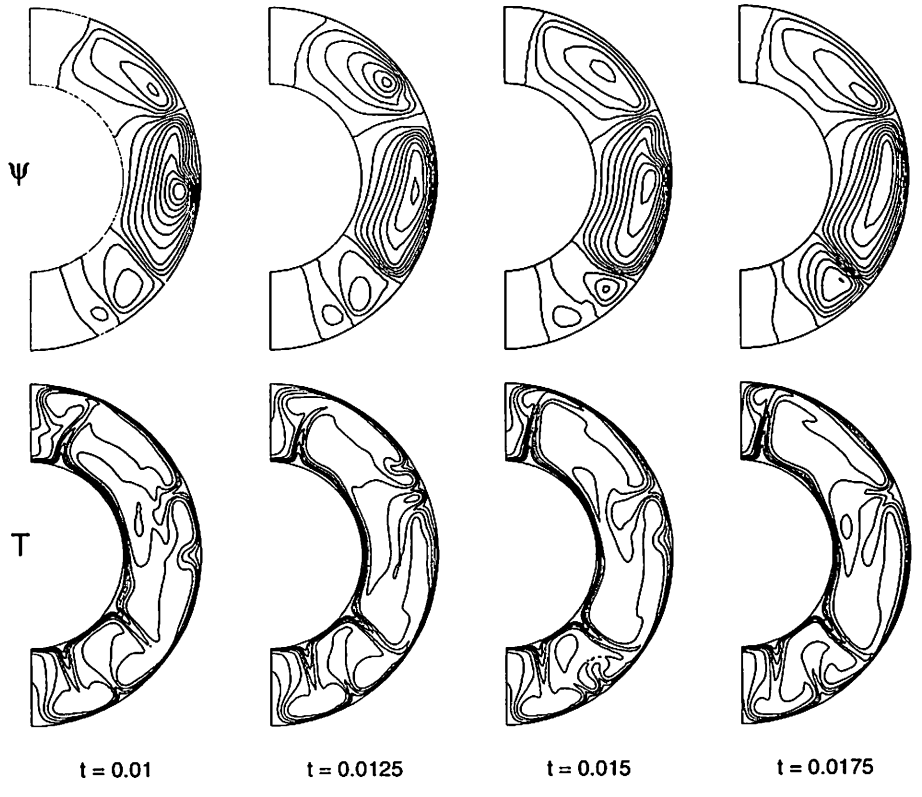

Figure 6 Evolution of stream function Ψ and temperature T in time for an axisymmetric model with strongly temperature and pressure dependent rheology. Spacing between the contours of stream function is 100. Temperature is contoured at an interval of 0.1. The time period $\Delta t = 0.0025$ between the frames represents approximately 200 Myrs. The average Nusselt number for this case is $Nu = 16$, the effective Rayleigh number is $\langle Ra \rangle = 5 \times 10^5$ and 60% of the heat at the Earth's surface is generated by internal heating.

Time is non-dimensionalized by the diffusion time scale. The initial condition of this run (not shown) was taken from a previous model calculation at constant viscosity and higher effective Rayleigh number. The time interval between the snapshots shown from left to right is approximately 200 million years. General characteristics of these type of models are the low average temperature (approx. 2000 K just above the thermal boundary layer) and the strong lower thermal boundary layer. Although the rate of internal heating is high (approximately 60% of the surface heat flow is generated by internal heating), strong plumes develop from the CMB. The upper mantle is much more dynamic, as indicated by the mobility of the upper boundary layer instabilities and the close spacing of the stream function contours in the upper mantle. As a consequence of the higher viscosity in the lower mantle, the plumes are near-stationary features with very low mobility. This difference in upper mantle dynamics compared to that of the lower mantle is evident in the plots of stream function as well. The effective Rayleigh number $\langle Ra \rangle$, based on the average viscosity is 5×10^5 for this case and the average Nusselt number is 16.

The evolution of a model at higher Ra is shown in Figure 7. Here, the average Nusselt number is 22 (which is at the high end of the range realistic for the Earth) and the rate of

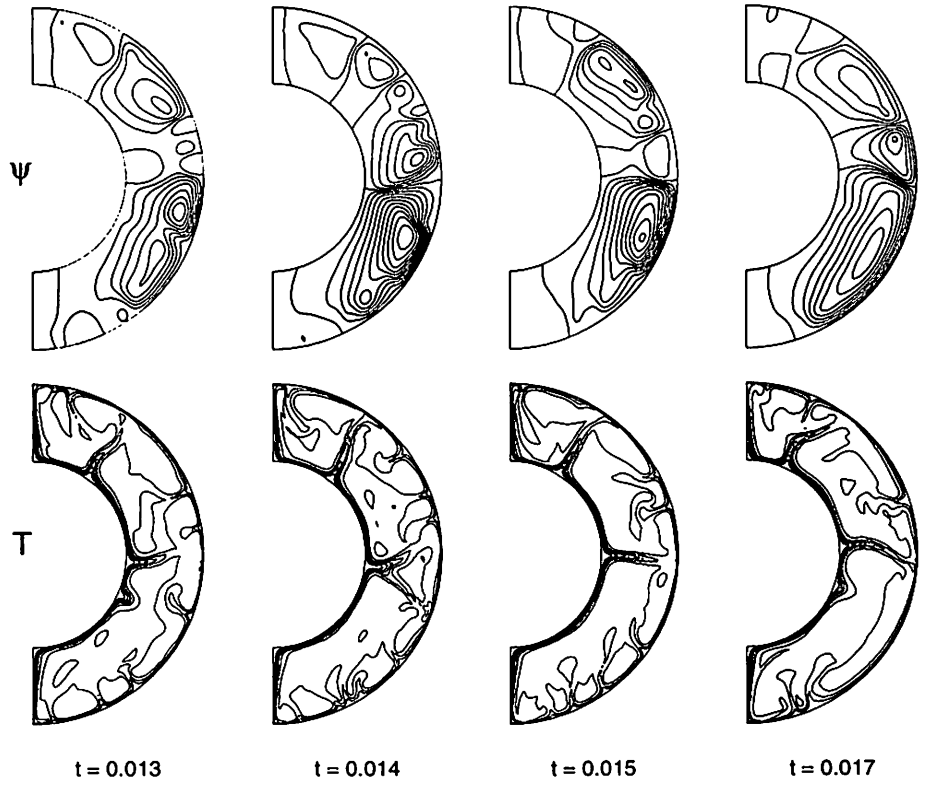

Figure 7 As Figure 5, but now for higher effective Rayleigh number $\langle Ra \rangle = 2 \times 10^6$. The spacing between the stream lines is 250. The time interval between the frames is $\Delta t = 0.001$ or approximately 100 Myr. The average Nusselt number is $Nu = 22$ and the ratio of internal to total heating is 56%.

internal heating is 56%. The effective Rayleigh number $\langle Ra \rangle$ is now 2×10^6. Snapshots are now shown at intervals of approximately 100 million years. The plumes in the lower mantle are substantially thinner, compared to the lower $\langle Ra \rangle$ simulation (compare Figure 6). The upper mantle shows strong time-dependent behavior on top of sluggishly moving lower mantle. Thus, it is shown that this type of rheology can explain the existence of stationary plumes underneath fast moving plates. It is essential to incorporate both a high viscosity in the lower mantle and a strong temperature dependence, as without the temperature dependence, the plumes would have difficulties penetrating the high viscosity in the lower mantle. Figure 8 shows the horizontally averaged viscosity at several points in time for the two models, indicating the close resemblance of these models to the inversion results of Ricard and Bai Wuming (1991).

Note that the effective Rayleigh numbers are substantially lower than the often quoted range of 10^7–10^8 (e.g., Peltier, 1989). This range is based on upper mantle values of the thermodynamical properties and does not take into account the likely depth variations, such as the decrease in thermal expansion (Chopelas and Boehler, 1989), increase in thermal diffusivity (Anderson, 1987) and increase in viscosity, that all work together to lower the effective Rayleigh. Using the high estimates of *Ra* leads to models with much higher Nusselt numbers, and consequently to unrealistically high surface heat flow values.

Returning to the main theme of our paper, these calculations are performed at substantially higher computational efficiency than with the traditional second order,

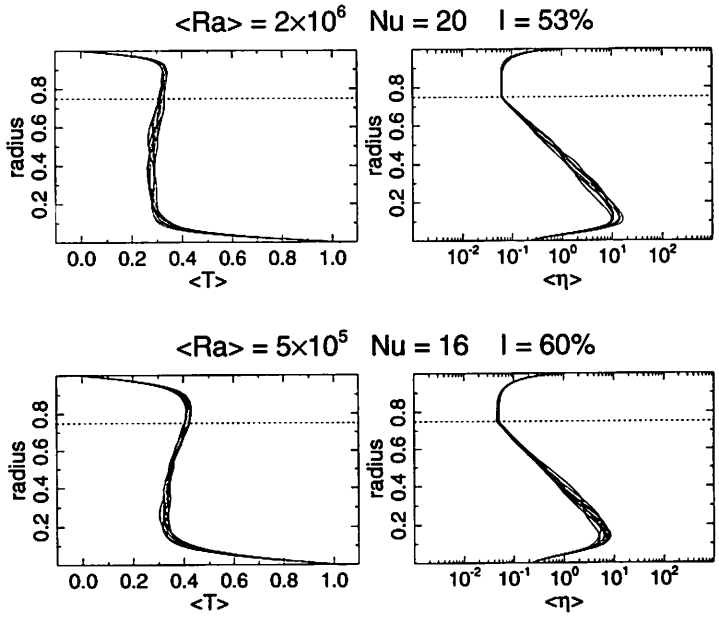

Figure 8 Radial (horizontally averaged) viscosity and temperature profiles for several time steps taken from the two cases shown in Figures 6 and 7. The temperature profiles clearly indicate the low average temperature in the lower mantle above the thermal boundary layer and the near-adiabatic conditions for the higher $\langle Ra \rangle$ case. The viscosity profile shows the marked viscosity increase as a consequence of the increase in melting temperature with depth.

fixed time-step predictor-corrector method. The DASPK method is also more stable as a consequence of the use of the backward difference formulae. In addition to the higher efficiency and stability of the solution of the temperature equation, the higher order that is employed allows for larger time-steps and therefore the momentum equations have to be solved less often. In the examples shown above, this leads to an overall factor of three to four improvement in computational speed in addition to the reduction in memory. This improvement is similar to that for isoviscous cases and Cartesian geometries (Van Keken *et al.*, 1994b) and has been noted in a wide variety of cases. In all situations, the comparison was made between solutions that were converged in time over the time interval of interest. Note that this is a substantial improvement with respect to the explicit first order schemes. It can be easily shown that use of a time-step that is close to the CFL-time-step leads to inaccurate results. To reach comparable accuracy as the results presented in this paper, a time-step of 10% of the CFL-criterion or less has to be used. Then, the number of solutions to the momentum equation increases strongly as well, leading to a strong decrease in overall efficiency.

The fact that only matrix-vector multiplications are necessary in the solution of the spatially discretized equations makes this formulation very attractive for use on vector and parallel machine architecture. DASPK itself has shown excellent vectorization and parallelization properties on specialized hardware. For the details see Maier *et al.* (1995). For the combination of DASPK and the finite element package SEPRAN as discussed in this paper, see Van Keken *et al.* (1994b).

5. DISCUSSION

We have presented here a new combination of techniques that allows for a more efficient, implicit solution to advection-diffusion equations and the full set of mantle convection equations, both for isoviscous and temperature-dependent rheology, compared with traditional implicit methods. Most important aspects of this combination are i) ease of implementation, ii) implicit scheme for the time-integration, iii) flexibility of the finite element method, iv) general applicability of DASPK to 2-D and 3-D problems and v) good parallelization and vectorization properties.

We have chosen to use a finite element package for the spatial discretization, which allows us to use this method for general and unstructured grids. In many cases, however, the geometries used in studies of mantle convection can very well be described by regular and structured grids, for which highly efficient discretization techniques exists, such as finite-difference and spectral methods.

We concentrated on a high order time-integration scheme. We expect that the combination of this with high-order or pseudo-spectral finite difference (Fornberg and Sloan, 1995; Moser *et al.*, 1995) or high-order finite element methods (Fischer and Ronquist, 1992) will lead to a more accurate and efficient solution scheme.

DASPK has been used here to solve the system of ordinary differential equations arising after spatial discretization of the heat equation. An interesting possiblity would be to use the Stokes equations as algebraic constraints, which would then give a system of differential-algebraic equations, for which DASPK is well suited (Brenan *et al.*,

1989). This method has been applied to the mean-field convection equations by Quareni and Yuen (1984).

Another application of DASPK can be found in the calculation of thermal-kinetic phase boundaries in subducting slabs (Dässler and Yuen, 1993) where the temperature equation must be solved simultaneously with a set of highly nonlinear ordinary differential equations at each slab grid point. Furthermore, DASPK can be applied to ground-water flow problems in a similar formulation as presented above for the mantle convection equations through Darcy's law, the porous flow analogue of the Navier-Stokes equations (e.g., Nield and Bejan, 1992).

Note that the spatially discretized system of Equations (6) is an expression of general time-dependent partial differential equations after spatial discretization. The algorithm shown here can thus be applied to other equations, as e.g. the Navier-Stokes or wave equation, as well.

6. CONCLUSIONS

The use of DASPK gives a large reduction of memory size for the heat equation in comparison with traditional implicit time-integration methods. DASPK is 3–4 times more efficient in the solution of the full set of convection equations than the second-order predictor-corrector method at the same or higher accuracy. The method is well suited for use on specialized vector and parallel machine architecture. In both 2-D and 3-D applications, the memory and CPU requirements of this implementation of the DASPK method grow linearly with number of nodal points. For a comparable 2-D simulation, approximately 1.5 times the amount of memory and CPU-time is needed per nodal point in 3-D. The formulation is well suited for other time-dependent problems (Navier-Stokes, wave equation) and other methods of spatial discretization, such as finite difference or spectral methods that are better suited for regular , structured grids. The method is efficiently applied to study the development of axisymmetric models of mantle convection, where the strongly temperature- and pressure-dependent viscosity is governed by an increase of melting temperature of lower mantle silicates with depth. The combined increase in both viscosity and temperature dependence may provide an explanation for the occurrence of near-stationary plumes in the lower mantle under a rapidly evolving upper mantle.

Acknowledgments

This research was supported by the Army High Performance Computer Research Center under ARO contract number DAAL03-89-C-0038 during the visiting fellowship of PvK. Discussions with Carl Gable, Hans-Peter Bunge, Paul Fischer, Gary Glatzmaier and John Baumgardner are greatly appreciated. We thank two anonymous referees for their constructive comments.

References

Anderson, D. L., "A seismic equation of state II. Shear properties and thermodynamics of the lower mantle," *Phys. Earth Planet. Int.* **45**, 307–323 (1987).

Balachandar, S. and Yuen, D. A., "Three-dimensional fully spectral numerical method for mantle convection with depth-dependent properties," *J. Comput. Phys.* **113**, 62–74 (1994).

Brenan, K., Campbell, S. and Petzold, L. R., *Numerical solution of initial-value problems in differential-algebraic equations*, Elsevier (1989).

Brown, N. P., Hindmarsh, A. C. and Petzold, L. R., "Using Krylov methods in the solution of large-scale differential-algebraic systems," preprint 93-033, Army High Performance Computing Research Center, 1100 Washington Ave S, Minneapolis, MN 55415, to appear in *SIAM J. Sci. Comp.* (1995).

Bunge, H.-P. and Baumgardner, J. R., "Mantle convection modeling on the parallel virtual machine," *Computers in Physics*, in press (1995).

Chopelas, A. and Boehler, R., "Thermal expansion measurements at very high pressures, systematics and a case for a chemically homogeneous mantle," *Geophys. Res. Lett.* **16**, 1347–1350 (1989).

Christensen, U., "Convection with pressure- and temperature-dependent non-Newtonian rheology," *Geophys. J. R. astr. Soc.* **77**, 343–34 (1984).

Christensen, U. and Harder, H., "3-D convection with variable viscosity," *Geophys. J. Int.* **104**, 213–226 (1991).

Corrieu, V., Ricard, Y. and Froidevaux, C., "Converting mantle tomography into mass anomalies to predict the Earth's radial viscosity" *Phys. Earth Planet Int.* **84**, 3–13 (1994).

Cuvelier, C., Segal, A. and Van Steenhoven, A. A. *Finite element methods and Navier-Stokes equations*, Reidel, Dordrecht (1986).

Dässler, R. and Yuen, D. A., "The effects of phase transition kinetics of subducting slabs, "*Geophys. Res. Lett.* **20**, 2603–2606 (1993).

Fischer, P. F. and Ronquist, E. M., "Spectral element methods for large scale parallel Navier-Stokes calculations," Proc. 2nd Internation conference on spectral and high-order methods, Montpellier, France (1992).

Fornberg, B. and Sloan, D. M., *A review of pseudospectral methods for solving partial differential equations*, Acta Numerica, Cambridge Univ. Press, in press (1995).

Gable, C. W., "Numerical models of plate tectonics and mantle convection in three dimensions," PhD Thesis, Harvard University, Cambridge MA (1989).

Gallagher, K. G. and Ahrens, T. J., "Shock-induced melting of olivine (high-pressure phase(s)) and its geophysical implications" *Eos Trans. Amer. Geophys. Un.* **75**, 653 (1994).

Hansen, U. and Ebel, A., "Time-dependent thermal convection, a possible explanation for multiscale flow in the Earth's mantle," *Geophys. J.* **94**, 181–191 (1988).

Houseman, G., "Boundary conditions and efficient solution algorithms for the potential function formulation of the three-dimensional viscous flow equation," *Geophys. J. Int.* **100**, 33–38 (1990).

Ivins, E. R., Sammis, C. G. and Yoder, C. F., "Deep mantle viscosity structure with prior estimate and satellite constraint" *J. Geophys. Res.* **98**, 4579–4609 (1993).

Lorenz, E., "Deterministic non-periodic flow," *J. Atmos. Sci.* **20**, 130–141 (1963).

Maier, R. S. and Petzold, L. R., *User's guide to DASPKMP and DASPKF90*, preprint 93-034, Army High Performance Computing Research Center preprint, 1100 Washington Ave S, Minneapolis, MN 55415 (1993).

Maier, R. S., Petzold, L. R. and Rath, W., "Parallel solution of large-scale differential-algebraic systems," preprint 94-014, Army High Performance Computing Research Center, 1100 Washington Ave S, Minneapolis, MN 55415. Also in: *Concurrency: Practice and Experience*, in press (1995).

Malevsky, A. V. and Yuen, D. A., "Characteristics-based methods applied to infinite Prandtl number thermal convection in the hard turbulent regime," *Phys. Fluids A* **3**, 2105–2115 (1991).

Moser, J., Yuen, D. A., Larsen, T. and Fornberg. B., "A new finite-difference method with variable mesh and order-application to mantle convection modeling," submitted to *Geophys. Astrophys. Fluid Dynam.* (1995).

Nield, D. A. and Bejan, A., *Convection in porous media*, Springer Verlag (1992).

Parmentier, E. M., Sotin, C. and Travis, B. J., "Turbulent 3-D thermal convection in an infinite Prandtl number, volumetrically heated fluid: Implications for mantle dynamics," *Geophys. J. Int.* **116**, 241–251 (1994).

Peltier, W. R., "Mantle viscosity," in: *Mantle convection: plate tectonics and global dynamics* (ed. W. R. Peltier) Gordon and Breach Publ. New York (1989).

Ricard, Y. and Bai Wuming, "Inferring the viscosity and 3-D density structure of the mantle from geoid, topography and plate velocities," *Geophys. J. Int.* **105**, 561–571 (1991).

Quareni, F. and Yuen, D. A., "Time-dependent solutions of mean-field equations with applications for mantle convection," *Phys. Earth Planet Int.* **36**, 337–353 (1984).

Saad, Y. and Schultz, M. H., "GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems," *SIAM J. Sci. Stat. Comp.* **7**, 856–869 (1986).

Segal, A. and Praagman, N., *Sepran User's guide*, Sepra, Leidschendam, The Netherlands (1984).

Shen, G. and Lazor, P., "Melting of minerals under lower mantle conditions: experimental results," *J. Geophys. Res.* in press, (1995).

Sloan, S. W., "An algorithm for profile and wavefront reduction of sparse matrices." *Int. J. Num. Meth. Engng.* **23**, 239–251 (1986).

Tackley, P., "Effects of strongly temperature-dependent viscosity on time-dependent, three-dimensional models of mantle convection," *Geophys. Res. Lett.* **20**, 2187–2190 (1993).

Travis, B. J., Weinstein, S. and Olson, P., "Three-dimensional convection planforms with internal heat generation," *Geophys. Res. Lett.* **17**, 243–246 (1990).

Van den Berg, A. P., Yuen, D. A. and Van Keken, P. E., "Effects of depth-variations in creep laws on the formation of plates in mantle dynamics," *Geophys. Res. Lett.* **18**, 2197–2200 (1991).

Van den Berg, A. P., Van Keken, P. E. and Yuen, D. A., "The effects of a composite non-Newtonian and Newtonian rheology on mantle convection," *Geophys. J. Int.* **115**, 62–78 (1993).

Van Keken, P. E., "Numerical modelling of thermochemically driven fluid flow with non-Newtonian rheology, applied to the Earth's lithosphere and mantle," *Geologica Ultraiectina*, **107**. PhD thesis, University of Utrecht, The Netherlands (1993).

Van Keken, P. E., Yuen, D. A. and Van den Berg, A. P., "Pulsating diapiric flows: consequences of vertical variations in mantle creep laws," *Earth Plan. Sci. Lett.* **112**, 179–194 (1992).

Van Keken, P. E., Yuen, D. A. and Van den Berg, A. P., "Implications for mantle dynamics from the high melting temperature of perovskite," *Science* **264**, 1437–1440, (1994a).

Van Keken, P. E., Yuen, D. A. and Petzold, L. R., "Time-integration techniques for the combined Stokes and heat equations using an implicit high order solver for differential-algebraic equations equations (DASPK), with applications to mantle convection," Report 94-019, Army High Performance Computing Research Center, University of Minnesota, 1100 Washington Ave S, Minneapolis MN 55415, (1994b)

Van Keken, P. E., Spiers, C. J., Van den Berg, A. P. and Muyzert, E. J., "The effective viscosity of rocksalt: implementation of steady-state creep laws in numerical models of salt diapirism," *Tectonophysics*, **225**, 457–476 (1993).

Vlaar, N. J., Van Keken, P. E. and Van den Berg, A. P., "Cooling of the Earth in the Archaean: consequences of pressure-release melting in a hotter mantle," *Earth Plan. Sci. Lett.* **121**, 1–18 (1994).

Weertman, J., "The creep strength of the Earth's mantle," *Rev. Geophys. Space Phys.* **8**, 145–168 (1970).

Zerr, A. and Boehler, R., "Melting of $MgSiO_3$-perovskite to 625 kilobars: indication of a high melting temperature in the lower mantle," *Science*, **262**, 553–555 (1993).

Zerr, A. and Boehler, R., "Constraints on the melting temperature of the lower mantle from high-pressure experiments on MgO and magnesiowüstite," *Nature* (1994).