Modeling of Viscoelastic Plume-Lithosphere Interaction Using Adaptive Multilevel Wavelet Collocation Method

Oleg V. Vasilyev\(^1\), Yuri Yu. Podladchikov\(^2\) and David A. Yuen\(^3\)

\(^1\) Department of Mechanical and Aerospace Engineering, University of Missouri-Columbia, Columbia, MO 65211. E-mail: VasilyevO@missouri.edu
\(^2\) Geologisches Institut, ETH-Zurich, Sonneggstr. 5, CH 8092, Zurich, Switzerland.
E-mail: yura@erdw.ethz.ch
\(^3\) Minnesota Supercomputer Institute and Department of Geology and Geophysics, University of Minnesota, Minneapolis, MN 55415. E-mail: davey@krisya.msi.umn.edu

Submitted to Geophysical Journal International, March 27, 2000

SUMMARY
Modeling of mantle flows with sharp viscosity contrasts in a viscoelastic medium is a challenging computational problem in geodynamics because of its multiple-scale nature in space and time. We have employed a recently developed adaptive multilevel wavelet collocation algorithm to study the dynamics of a small rising diapir interacting with a stiff lithosphere in a Maxwell viscoelastic mantle. In this kinematic model the upward velocity of the diapir is prescribed and only momentum equations involving the temporal evolution of the velocity components, pressure and the components of the stress tensor are integrated in time. The total number of collocation points did not exceed \(10^6\), which would require more than \(10^6\) grid points using conventional evenly spaced grid methods. Viscosity of the diapir is \(3 \times 10^{-4}\) times lower than the surrounding mantle, while the viscosity of the thin lithosphere, about 5 to 10\% of the entire layer depth, is \(10^6\) to \(10^8\) times stiffer than the ambient mantle. Our results show the efficacy of wavelets to catch the sharp gradients of the stress and pressure fields developed in the diapiric impingement process. Stress fields are focussed with time in the lithosphere. The magnitude of the stress fields can reach around 100 to 300 Mpa and are localized within the viscoelastic lithosphere from the interaction with the rising viscoelastic diapir. Viscoelasticity can play an important role in the dynamics of the lithosphere, especially concerning the potential severance of the lithosphere by mantle upwellings.

Key words: geodynamics, diapir, lithosphere, viscoelastic, wavelet, numerical method.

1 INTRODUCTION

There is a growing recognition of the important role played by viscoelasticity in mantle flow processes since the pioneering work by Harder (1991). Ricard et al. (1993) have attempted to model finite-amplitude subduction process with viscoelastic normal modes and recently Toth and Gurnis (1998) have studied the initiation of subduction with the finite-element method. Kameyama et al. (1999) have studied the thermal-mechanical evolution of viscoelastic shear zones and Schmalholz and Podladchikov (1999) have investigated the folding and buckling of viscoelastic medium in the large amplitude regime. Polakov et al. (1993) have investigated diapiric upwelling within a viscoelastic lithosphere using a finite-difference method. Modeling of viscoelastic flows in geodynamics remains still a very difficult problem because of both theoretical and numerical difficulties. The fundamental problem in modeling time-dependent viscoelastic flows is the mixed rheological properties, which result in a time-dependence of the stress on the history of the forcing function. The advection and rotation of stress fields are also important in problems with free-surface and faulting. Some finite-element models of viscoelastic behavior have been developed by Melosh and Raefsky (1980) for a fluid with a non-
Newtonian viscous rheology and by Chery et al. (1991) for coupled viscoelastic and plastic rheology. Both approaches are very powerful, but they can simulate relatively small deformations and are limited by the distortion of the Lagrangian grid. Therefore, it is important to consider new non-traditional methods for modeling complex rheologies over long timespans and which are adaptive in nature for capturing moving fronts involving sharp gradients in the viscosity and stress fields. In this paper we will introduce the adaptive wavelet collocation method to study the geophysical problem of a low-viscosity moving diapir impinging a highly viscous thin lithosphere in a viscoelastic mantle.

Traditional methods used for numerical solutions of geophysical problems mostly fall into three classes: finite-difference and finite-volume methods, finite-element methods, and spectral methods. The advantages of using either finite difference, finite volume, or finite element methods are the simplicity in adapting to complex geometries and sparseness of resulting matrices, while the basic advantage of spectral methods lies in the exponential rate of convergence.

If the solution of a geophysical problem has regular features, any of these numerical techniques can be applied. However, in many problems of geophysics there exists a multiplicity of very different spatial and temporal scales in the solution, as in strongly time-dependent non-Newtonian convection (Malesky & Yuen 1992), plume-lithosphere interaction (Larsen et al. 1993; Christensen & Ribe 1994), and diapir-lithosphere interaction (Poljakov et al. 1993). This particular attribute of multiple spatial scales, which possibly change over time, will put great strain on the aforementioned methods. Spectral methods would have some problems capturing large irregularities of the solutions. Accurate representation of the solution in regions where sharp physical transitions occur will require the implementation of dynamically adaptive finite-difference or finite-element methods (e.g. (Braun & Sambridge 1994)). Kameyama et al. (1999) employed an adaptive finite-difference method in a 1-D model but still needed 10,000 points to resolve the sharp shear zones being shrunk by viscous heating. Schmalholz and Podladchikov (1999) have developed a novel hybrid finite-difference spectral methods in 2-D and found that 1,000 points in the vertical are required in the large strain, O(1), regime. In these methods an automatic error estimation step should be employed to determine locally the accuracy of the solution. The main difficulties of existing adaptive methods are finding stable accurate spatial operators at the interface of computational molecules of very different sizes and developing computationally efficient robust adaptive procedure, which would dynamically adapt the computational grid to local structures of the solution.

Wavelet analysis is a new numerical concept which allows one to represent a function as a linear combination of building blocks (basis functions), called wavelets, which are localized in both location and scale (Daubechies 1992; Louis et al. 1997; Meyer 1992; Strang & Nguyen 1996). Good wavelet localization properties in physical and wavenumber spaces are to be contrasted with the spectral approach, which employs infinitely differentiable functions but with global support and small discrete changes in the resolution. On the other hand, finite-difference, finite-volume and finite-element methods have small compact support but poor continuity properties. Consequently, spectral methods have good spectral localization (which results in exponential convergence rates), but poor spatial localization (which results in Gibbs phenomena in regions of fast transitions), while finite-difference, finite-volume, and finite-element methods have good spatial localization but poor spectral localization (which results in algebraic convergence rates). Wavelets appear to combine the advantages of both spectral and finite-difference bases. One of the principal purpose of this paper is to present both the essence of the wavelet method and to demonstrate the prowess in solving geophysical problems with a localization of physical properties and the computational economy in doing so.

The paper is organized as follows. In Section 2 we present in detail the mathematical formulation of the problem. Numerical implementation of the adaptive wavelet collocation algorithm is described in Section 3. In Section 4 we discuss the time-dependent results of the viscoelastic flow in a model mantle driven by a rising diapir. The geophysical implications and conclusions of viscoelastic flows in the mantle are given there as well.

2 MATHEMATICAL MODEL

Viscoelastic processes play an important role for short time-scale transient phenomena. In this paper we will consider a plane-strain viscoelastic flow in the mantle with a strongly variable viscosity driven by density inhomogeneities in a vertical rectangular domain [0, L*] × [0, L*]. L* can be considered to have the dimension of the upper mantle. Super-script “*” denotes dimension quantities. The viscoelastic flow processes involves a thin highly viscous upper boundary layer (lithosphere) which interacts with highly variable viscous interior (the mantle) associated with a rising diapir, which is modeled kinematically by a rising vertically small lower density sphere with a viscosity considerably lower than the ambient mantle.

This viscoelastic model problem involves six unknowns (two velocity components V1* and V2*, pressure p*, (the isotropic part of the three dimensional stress tensor with minus sign), and three in-plane deviatoric components of stress tensor τ11*, τ22*, and τ12*). The dimensional scales are the size of the domain L*, the dynamic viscosity μ*, the shear elastic modulus G*, and gravity force per unit volume ρ*g*, where ρ* is the characteristic scale for the density deviations and g* is the gravity acceleration. We emphasize that both pressure p* and density ρ* are the deviations from the reference hydrostatic state. There is also an extra parameter “inertial density” ρ0*, which may not equal to the p* (Poljakov et al. 1993). The three independent characteristic scales used in dimensional analyses are L*, μ* and ρ* g*, which makes the stress, time and velocity scales to be ρ* g* L*, μ* / (ρ* g* L*)
Table 1. The choice of parameters for five different cases studied.

<table>
<thead>
<tr>
<th>Case</th>
<th>Case II</th>
<th>Case III</th>
<th>Case IV</th>
<th>Case V</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_1$</td>
<td>$10^4$</td>
<td>$10^6$</td>
<td>$10^6$</td>
<td>$10^6$</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>$10^{-2}$</td>
<td>$10^{-2}$</td>
<td>$10^{-2}$</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>$2 \times 10^{-2}$</td>
<td>$2 \times 10^{-2}$</td>
<td>$5 \times 10^{-2}$</td>
<td>$2 \times 10^{-2}$</td>
</tr>
<tr>
<td>$v_{20}$</td>
<td>$4 \times 10^{-2}/Dc$</td>
<td>$10^{-1}/Dc$</td>
<td>$10^{-2}/Dc$</td>
<td>$10^{-3}/Dc$</td>
</tr>
</tbody>
</table>

and $(\rho^* g^* (L^*)^2)/\mu^*$ respectively. Finite amplitude effects of stress advection have not been included.

The viscoelastic equations described in (Poliakov et al., 1993) are given by

$$\frac{\partial V_i}{\partial t} = \frac{1}{Re} \left[ -\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j} \left( \frac{\partial V_j}{\partial x_i} \right) + \rho \frac{\partial \phi}{\partial x_i} \right] ,$$

$$\frac{\partial p}{\partial t} = -K \frac{\partial V_i}{\partial x_i} ,$$

$$\frac{\partial \tau_{ij}}{\partial t} = \frac{1}{Dc} \left[ \frac{\tau_{ij}}{2\mu} + \frac{1}{2} \left( \frac{\partial V_j}{\partial x_i} + \frac{\partial V_i}{\partial x_j} \right) - \frac{1}{3} \frac{\partial V_k}{\partial x_k} \delta_{ij} \right] ,$$

where $i, j, k = 1, 2$, repeated indices imply the summation, $\delta_{ij}$ is the Kronecker delta, and $\epsilon_{xy} = (0, -1)$ is the unity vector along gravity. The independent dimensionless parameters appearing in the equations are

$$Re = \frac{\rho^* \phi^* (L^*)^2}{(\mu^*)^2} , \quad Dc = \frac{\rho^* g^* L^*}{G^*} , \quad K = \frac{K^*}{G^*} .$$

These parameters represent the Reynolds and Deborah numbers, and the measure of the ratio of bulk and shear elastic moduli. The Deborah number is the ratio of the Maxwell relaxation time $\mu^*/G^*$ to the characteristic time scale defined above. The Deborah number $De$ (Reiner 1964) is a measure of the relative influence of the viscous versus elastic deformation modes, such that when $Dc = O(1)$ both modes are comparable, while $Dc \rightarrow 0$ and $Dc \rightarrow \infty$ respectively represent the viscous and elastic limits. For realistic Earth parameters, $Dc$ may lie between $10^{-2}$ to $10^2$ with corresponding time scales between $10^{-2}$ and $10^2$ Myrs and spatial scales between 5 and 500 km (Vaslyev et al., 1998). Note that although Eqs. (1-3) can be rescaled in such a way that some of the parameters are omitted, we have chosen this form for geophysical relevance.

The density perturbation serves the purpose of driving the flow. The non-dimensional density perturbation $\rho$ in Eq. 1 is chosen to be

$$\rho(x_1, x_2, t) = -\rho_0 f_0(x_1, x_2, t) ,$$

where

$$f_0(x_1, x_2, t) = e^{-((x_1-x_{20})^2+(x_2-x_{20}-v_{20}t-x_{40})^2)/\sigma^2} ,$$

$v_{20}$ is the speed with which the diapir rises, $(x_{10}, x_{20})$ is the diapir location at the initial time $t_0$. In order for results to be of the order of unity we choose $\rho_0 = 10^5$.

The non-dimensional viscosity $\mu$ is taken to be $\mu(x_1, x_2) = (\mu_0 + (\mu_1 - \mu_0) f_1(x_2))(1 - f_0(x_1, x_2, t)) + \mu_2 f_2(x_1, x_2, t) , (7)$

where

$$f_1(x_2) = e^{-(x_2-x_{10})^2}/\lambda^2 .$$

Constants $\mu_0, \mu_1, \mu_2, \lambda, r_0, v_{20}$ and functions $f_1$ and $f_2$ are chosen in such a way that high viscosity region is centered near the top. The problem is solved for five different cases. The following parameters are common for all the cases: $Re = 10^{-3}$, $Dc = 10^{-2}$, $K = 1$, $\mu_0 = 1$, $r_0 = 0.05$, $(x_{10}, x_{20}) = (0.0, 0.0)$, while parameters $\mu_1$, $\mu_2$, $\lambda$, $v_{20}$ are varied. Five sets of parameters corresponding to different cases are given in Table 1.

The initial and boundary conditions are respectively given by

$$V_i(x_1, x_2, 0) = p(x_1, x_2, 0) = \tau_{ij}(x_1, x_2, 0) = 0, \quad i, j = 1, 2 ,$$

and

$$V_i(0, x_2, t) = V_i(1, x_2, t) = V_2(x_1, 0, t) = V_2(1, 1, t) = 0 ,$$

$$\tau_{12}(x_1, 0, t) = \tau_{12}(x_1, 1, t) = \tau_{12}(0, x_2, t) = \tau_{12}(1, x_2, t) = 0 .$$

The initial condition for the Case V is taken to be the solution of the Case I at time $t = 3.0 Dc$.

3 NUMERICAL METHOD

In this section we will describe basic elements of dynamically adaptive wavelet collocation method used for this viscoelastic problem. For the simplicity of presentation we will formulate the algorithm for two-dimensional problem. The generalization of the algorithm into three dimensions is straightforward. In describing the method we will try to present it in a way that is easy to understand and implement without profound knowledge of wavelet theory. For theorem proofs and other mathematical aspects of the algorithm we refer to the articles of Vaslyev et al. (1995; 1996; 1997; 1997a; 1997b).

3.1 Multilevel Wavelet Approximation

Wavelets are constructed by the discrete (typically dyadic) dilation $a_j = a_0 2^{-j}$ and translation $b_j = a_0 k$ of a single function, called “mother” wavelet, $\psi(x)$, which has good localization properties in physical as well as wave number spaces. Examples of “mother” wavelets are shown in Fig. 1.
The discrete dilations and translations generate a doubly-indexed set of wavelet functions $\psi^j_k(x)$ given by

$$\psi^j_k(x) = a_j^{-1/2}\psi\left(\frac{x - b^j_k}{a_j}\right),$$

where $\psi^j_k(x)$ is a wavelet of scale $a_j$ located at position $b^j_k$. For notational convenience we use the superscript to denote the level of resolution and the subscript to denote the location in physical space (with the exception of $a_j$). The wavelet decomposition of a one-dimensional function $u(x)$ may be written as:

$$u(x) = \sum_{j=-\infty}^{\infty} \sum_{k \neq 0} c^j_k \psi^j_k(x).$$

In other words wavelet decomposition can be viewed as multi-level or multi-resolution representation of a function, where each level of wavelet consists of waves having the same scale $a_j$ but different locations $b^j_k$. For details of wavelet theory we refer to the following books (Daubechies 1992; Louis et al. 1997; Meyer 1992; Strang & Nguyen 1996).

The more familiar trigonometric decompositions of a periodic function can be expressed using above notation as:

$$u(x) = \sum_{j=0}^{\infty} \sum_{k \neq 0} c^j_k \sin(jx + \pi k/2),$$

where $j$ corresponds to the frequency (the scale) of the basis function and $k$ describes its location. In this particular example there are only two locations of the basis function, which is due to the global support of the basis function.

Wavelet decompositions can be introduced the same way in two dimensions. In this paper we construct two-dimensional wavelet basis as a tensor product of two one-dimensional bases:

$$\psi^j_{k_1,k_2}(x, y) = \psi^j_{k_1}(x)\psi^j_{k_2}(y),$$

where $\psi^j_{k_1}(x)$ and $\psi^j_{k_2}(y)$ are two one-dimensional wavelets. Therefore, a function $u(x, y)$, defined in rectangular domain $\Omega = [X_1, X_2] \times [Y_1, Y_2]$, is approximated as a linear combination of basis functions (12) and is given by

$$u^j(x, y) = \sum_{j=0}^{J_a} \sum_{(k, l) \neq 0} c^j_{k_1,k_2} \psi^j_{k_1,k_2}(x, y).$$

(13)

After discretization of this rectangular domain the scales and locations are defined by:

$$a_{j_k} = 2^{-(L_x + j_x)} (X_2 - X_1), \quad b^j_{k_1} = \frac{X_1 + X_2}{2} + a_{j_k}k$$

for $j_k = 0, 1, \ldots, J_x$ and

$$a_{j_l} = 2^{-(L_y + j_y)} (Y_2 - Y_1), \quad b^j_{k_2} = \frac{Y_1 + Y_2}{2} + a_{j_l}l$$

for $j_l = 0, 1, \ldots, J_y$. For clarity of presentation we use subscripts $x$ and $y$ to denote all the parameters associated with directions $x$ and $y$ correspondingly. Integer parameters $L_x$ and $L_y$ determine the largest scales of the approximation in $x$ and $y$ directions, while integers $J_x$ and $J_y$ define the finest scales of the approximation (13).

For a clarity of discussion, all wavelets whose centers are located within the rectangular domain will be called internal wavelets; the other wavelets will be called external wavelets. Due to the local or approximately local (the exponential decay) support of a wavelet, only those external wavelets which located close to the boundary contribute to the approximation $u^j(x, y)$. Thus wavelets which are located far away from the domain can be omitted from the approximation (13). On each side of the computational domain we will keep only $N_x$ and $N_y$ external wavelets in the directions $x$ and $y$ respectively. Then the approximation (13) can be rewritten as

$$u^j(x, y) = \sum_{j=0}^{J_a} \sum_{(k, l) \neq 0} c^j_{k_1,k_2} \psi^j_{k_1,k_2}(x, y).$$

(14)

where $Z^j = [-K_2, K_2] \times [-K_2, K_2]$, $K_2^x = 2^{L_x + j_x - 1} + N_x$, and $K_2^y = 2^{L_y + j_y - 1} + N_y$.  

3.2 Wavelet Interpolation

The approximation (14) can be considered as a wavelet interpolation provided there is a way to obtain values of wavelet coefficients from the values of the function at a discrete set of grid (collocation) points. Due to the multilevel character of approximation (14) we choose the set of collocation points in such a way that any finer level contains the collocation points of the coarser level functions. This relation between collocation points of different levels of resolution enables us to have the same values of the function at different levels of resolution at the same collocation points.

We associate each wavelet with one particular collocation point. Note that wavelets at different levels of resolution may have the same collocation point. For internal wavelets we chose wavelet locations $(b^j_{k_1}, b^j_{k_2})$ to be
the collocation points \((x_k^j, y_l^j)\). The collocation points of external wavelets of a finer level of resolution. For example let \(b_1^j < X_1\) and \(Y_1 < b_2^j < Y_2\), then collocation point is taken to be \((x_k^j, y_l^j) = (b_2^j - b_1^j, b_2^j)\), where \(j = j + \log_2(N_z) + 1\) (\([\cdot]\) integer part) and \((b_2^j - b_1^j) / a_{zz^j} = (X_1 - b_1^j) / a_{zz^j}.

Wavelet coefficient are found in two steps:

(i) **Initialization Step.**
At the coarsest level of resolution \((j = 0)\) we set the values of the approximation at collocation points \((x_k^j, y_l^j)\) equal to values of the function at those locations. Then in order to obtain wavelet coefficients \(c_{k,l}^j\) we solve the following linear system of equations:

\[
u^j(x_k^0, y_l^0) = \sum_{(m,n) \in Z_x^j \times Z_y^j} c_{m,n}^j \psi_{m,n}^j(x_k^0, y_l^0). (15)
\]

(ii) **Recursion Step.**
For each level of resolution \(0 < j \leq J\) we fix the values of wavelet coefficients at the lower levels of resolution. Then we form the residual \(\Delta^j(x_k^j, y_l^j)\) between values of a function at collocation points of finest level of resolution and wavelet approximation at the previous level of resolution:

\[
\Delta^j(x_k^j, y_l^j) = u^j(x_k^j, y_l^j) = \sum_{k=0}^{j-1} \sum_{(m,n) \in Z_x^j \times Z_y^j} c_{m,n}^{j-1} \psi_{m,n}^{j-1}(x_k^0, y_l^0).
\]

Then we set the values of the individual contribution of \(j\) level of resolution at collocation points \((x_k^j, y_l^j)\) equal to values of the residual \(\Delta^j(x_k^j, y_l^j)\) at these locations. Thus in order to obtain wavelet coefficients \(c_{k,l}^j\) we solve the following linear system of equations:

\[
\Delta^j(x_k^j, y_l^j) = \sum_{(m,n) \in Z_x^j \times Z_y^j} c_{m,n}^j \psi_{m,n}^j(x_k^j, y_l^j).
\]

In other words, wavelet coefficient are found in recursive manner. We start from the coarsest level of resolution and progressively move to the finest level. On each level of resolution the coefficients of the lower levels are fixed, so we only obtain the coefficients corresponding to that level.

In actual implementation of the algorithm the values of the residual \(\Delta^j(x_k^j, y_l^j)\) are found recursively by using the following relation:

\[
\Delta^j(x_k^j, y_l^j) = \Delta^{j-1}(x_k^j, y_l^j) = \sum_{(m,n) \in Z_x^j \times Z_y^j} c_{m,n}^{j-1} \psi_{m,n}^{j-1}(x_k^0, y_l^0)
\]

for \(1 \leq j \leq J\), \((k, l) \in Z_x^j \times Z_y^j\), while \(\Delta^0(x_k^0, y_l^0) = u^0(x_k^0, y_l^0)\).

The solution of equations (15) and (17) for general wavelet might be expensive. In order to make algorithm computationally efficient we utilize cardinal interpolating wavelets (Donoho 1992; Sweldens & Schröder 1996). By cardinal wavelets we mean wavelet which satisfy the following relation

\[
\psi(k) = \delta_{k,g},
\]

where \(\delta_{k,g}\) is the Kronecker delta symbol. Example of such cardinal scaling function is shown on Fig. 1(d).

The procedure of finding wavelet coefficients is the following. First we find coefficients for internal wavelets, then we find coefficients for all external wavelets except corner wavelets, finally we find coefficients for corner wavelets. This three step approach enables us to find wavelet coefficients without matrix inversion. The cost of finding wavelet coefficient at \(j\) level of resolution is \(N^j\), where \(N^j\) is the number of wavelets at \(j\) level of resolution.

Knowing the values of wavelet coefficients the function \(u(x, y)\) can be approximated by (14). The derivatives \(u^{(m,n)}(x, y)\) of a function are approximated by differentiating (14), where \(m, n = 0, 1\). Thus the derivative of a function \(u(x, y)\) is found using the following relation:

\[
u^{(m,n)}(x, y) = \sum_{j=0}^{J} \sum_{k \in Z_x^j} \sum_{l \in Z_y^j} c_{k,l}^{j} \psi_{k,l}^{j}(m,n)(x, y).
\]

Derivatives higher than first order can be found by consecutive applications of the above algorithm.

### 3.3 Wavelet Compression

The absolute value of the wavelet coefficient \(c_{k,l}^j\) appearing in the approximation (14) depends upon the local regularity of \(u(x, y)\) in the neighborhood of location \((b_1^j, b_2^j)\). A good approximation of a function is retained even when wavelets whose coefficients are below a prescribed threshold parameter \(c\) are omitted and only those wavelets are kept whose coefficients are above the threshold. In order words \(u(x, y)\) can be approximated by

\[
\hat{u}(x, y) = \sum_{j=0}^{J} \sum_{k \in Z_x^j} \sum_{l \in Z_y^j} c_{k,l}^{j} \psi_{k,l}^{j}(x, y).
\]

where two-dimensional integer set \(Z_x^j \times Z_y^j\) consists of wavelets whose amplitudes satisfy the following inequality:

\[
|c_{k,l}^j| \geq a_1^{1/2} a_2^{1/2} c.
\]

We also define the two-dimensional integer set \(Z^j \times Z^j\) is such a way that collocation points \((x_k^j, y_l^j)\), \((k, l) \in Z^j \times Z^j\) at finest level of resolution are constructed as a union of collocation points at all levels of resolution. It is easy to show that

\[
\|u(x, y) - \hat{u}(x, y)\|_{L^2(\Omega^j)} \leq C \|u^{(j)}(x, y)\|_{L^2(\Omega^j)},
\]

where \(C\) is a constant of order unity. In Fig. 2 we sketch the locations of wavelets used in a four-level approximation of a one-dimensional function. Figure 2 also shows schematically the locations of wavelets whose coefficients are above a given threshold parameter for a function that has a sharp
transition in the middle of the domain. This schematic example illustrates the considerable savings one can achieve by keeping in the approximation only significant wavelets and omitting those, which play an insignificant role in the approximation.

The derivatives $u^{(m,n)}(x,y)$ of a function are approximated by differentiating (21) and are given by

$$u^{(m,n)}_2(x,y) = \sum_{j=0}^{J} \sum_{(k,l) \in Z_{2^j \geq 2}} c_{j,k,l}^{m,n}(x,y).$$

For convenience of the discussion all wavelets $\psi_{j,k,l}^{(m,n)}$ whose coefficients satisfy inequality (22) will be called active wavelets. Thus the set $Z_{2^j \geq 2}$ consists of indexes of active wavelets of $j$ level of resolution.

Note that if function $u(x,y)$ has localized structures, then approximation (21) has considerably less wavelets then approximation (14). We also note that the accuracy of the approximation is controlled by $c$: the smaller the epsilon the closer is the approximation to $u^2(x,y)$.

### 3.4 Fast Wavelet Collocation Transform

In the application of partial differential equations it essential to find values of derivative of a function at collocation points by providing values of the function at the same collocation points. This can be achieved by the means of Fast Wavelet Collocation Transform (FWCT). The essential ingredients of FWCT are described in Section 3.2. Assuming that locations of active wavelets are prescribed a priori. Then FWCT consists of two steps:

(i) Given values of function at collocation points $(x_{k,j}, y_{l,j})$ $(k,l) \in Z_{2^j \geq 2}$ we find values of wavelet coefficients at all levels of resolution. The procedure is the same as described in Section 3.2. We start from the coarsest level of resolution and progressively move to the finest level. On each level of resolution $j$ the coefficients of the lower levels are fixed, so we only obtain the coefficients $c_{j,k,l}^{m,n}$, $(k,l) \in Z_{2^j \geq 2}$ corresponding to that level of resolution.

(ii) Values of derivative of a function are found by evaluating (21) at $(x_{k,j}, y_{l,j})$ $(k,l) \in Z_{2^j \geq 2}$ collocation points.

Due to the fact that wavelet function has compact or effectively compact support the overall computational cost of finding wavelet coefficients and values of the derivatives of a function at collocation points is $O\left((J+1)M^2N\right)$, where $N$ is the total number of collocation points and $M$ is wavelet support. Thus the smaller wavelet support is the less expansive is to evaluate the derivatives of the function given its values at collocation points.

### 3.5 Application to Partial Differential Equations

Let us discuss the application of wavelet interpolation and wavelet compression algorithm into the numerical solution of partial differential equation. We will focus on evolutionary equations which can be applied in geophysics. The most general form of the system of partial differential equations can be written in the following form:

$$\frac{\partial u}{\partial t} = F(x, y, t, u, u_x, u_y),$$

$$\Phi(x, y, t, u, u_x, u_y) = 0,$$

where $u = (u_1, \ldots, u_3)$ is a vector of unknown functions, equations (25) describes time evolution and equations (26) represent either algebraic differential constrains, boundary conditions, or elliptic partial differential equations. The numerical method is formally derived by evaluating the governing partial differential equations (25), (26) at collocation points, which results in a system of nonlinear ordinary differential-algebraic equations describing the evolution of the solution at these collocation points. Spatial derivatives of a function $u_k(x, y, t)$ at collocation points are found using Fast Wavelet Collocation Transform (FWCT).

In order for the algorithm to resolve all the structures appearing in the solution, the basis of active wavelets and, consequently, the computational grid should adapt dynamically in time to reflect local changes in the solution. The adaptation of the computational grid is based on the analysis of wavelet coefficients. The contribution of a wavelet into the approximation is considerable if and only if the nearby structures of the solution have comparable size with the wavelet scale. Thus, we may drop the large number of fine scale wavelets with small coefficients in the regions where the solution is smooth. In wavelet collocation method every wavelet is uniquely associated with a collocation point. Consequently the collocation point should be omitted from the computational grid if the associated wavelet is omitted from the approximation. This property of the multilevel wavelet approximation allows local grid refinement up to an arbitrary small scale without a drastic increase of the number of grid points.

To ensure accuracy of the approximation, the basis
should consist of wavelets whose coefficients are or can possibly become significant during the period of time when the basis remains unchanged. Thus, at any instant of time, the basis should include wavelets belonging to an adjacent zone of wavelets for which the magnitude of coefficients satisfy criteria (22). We say that the two-dimensional wavelet \( \psi_{m,n}(x, y) \) belongs to the adjacent zone of wavelet \( \psi_{k,l}(x, y) \) if the following relations are satisfied:

\[
|\omega| \leq L_u, \quad |\psi_{m,n} - \psi_{k,l}| \leq M_x a_{x x}, \quad |\psi_{m,n} - \psi_{k,l}| \leq M_y a_{y t}, \quad (27)
\]

where \( L_u \) determines the extent of which coarser and finer scales are included into the adjacent zone and \( M_x \) and \( M_y \) define the width of the adjacent zone in physical space. The concept of adjacent zone is illustrated in Fig. 3 by showing the adjacent zone of one-dimensional wavelet \( \psi_k \) for \( L_u = 1 \) and \( M_x = 1 \).

The values of \( L_u, M_x, \) and \( M_y \) affect the total number of collocation points present in the irregular grid at any instant of space and the time interval during which the calculations can be carried out without modifying the computational grid. For efficiency we want to keep the number of collocation points as small as possible, while at the same time we would like to minimize changes in the collocation grid.

Let us denote by \( G_{x,y} \) the irregular grid of wavelet collocation points that are retained to approximate the solution at time \( t \). Following the classical collocation approach and evaluating (25), (26) at collocation points \( (x_k, y_l) \) \( (k, l) \in Z_{x,y} \geq \) we obtain a system of ordinary differential equations.

The present numerical algorithm consists of three steps:

(i) Knowing the values of the solution \( u_{k,l}^t(t) \) we compute the values of wavelet coefficients at all levels of resolution.

For a given threshold \( c \) we adjust \( G_{x,y}^{t+\Delta t} \) based on the magnitude of wavelet coefficients.

(ii) If there is no changes between computational grids \( G_{x,y}^t \) and \( G_{x,y}^{t+\Delta t} \) at time \( t \) and \( t + \Delta t \), we go directly to step 3. Otherwise, we compute the values of the solution at the collocation points \( G_{x,y}^{t+\Delta t} \), which are not included in \( G_{x,y}^t \).

(iii) Integrate the resulting system of ordinary differential equations to obtain new values \( u_{k,l}^t(t + \Delta t) \) at positions on the irregular grid \( G_{x,y}^{t+\Delta t} \) and go back to step 1.

The basic hypothesis motivating the algorithm is that during a time interval \( \Delta t \), the domain of wavelets with significant coefficients does not move in phase space beyond its border. With such an algorithm the irregular grid of wavelet collocation points is dynamically adapted in time and follows the local structures that appear in the solution. Note that by omitting wavelets with coefficients below a threshold parameter \( c \) we automatically control the error of approximation. Thus wavelet collocation method has another important feature such as active control of the accuracy of the solution. The smaller \( c \) is chosen the smaller the error of the solution is. In typical application the value of \( c \) varies between \( 10^{-2} \) and \( 10^{-4} \). We also note that the larger the value of \( c \) is set the fewer number of grid points are used to obtain the solution.

4 RESULTS AND DISCUSSION

The results presented in this section have been obtained by using the dynamically adaptive multilevel collocation method. The cardinal interpolating wavelet of order ten (Donoho 1992; Sweldens & Schröder 1996), which represents a very accurate scheme, was employed with the threshold pa-
Figure 6. Absolute value of the $\tau_{12}$ component of the stress tensor for the case I at four different times.

Figure 7. Absolute value of the pressure $p$ for the case I at four different times.

Figure 8. Absolute value of $\tau_{11} + \tau_{22}$ for the case I at four different times.

Figure 9. Computational grid for the case I at four different times.

rameter $c = 5 \times 10^{-7}$, which means that the local relative error is everywhere less than $5 \times 10^{-7}$.

The adaptation of the computational grid is based on the analysis of coefficients associated with all six dependent variables of Eqs. (1-3). The irregular grid $\mathcal{G}_h$ of wavelet collocation points is constructed as a union of irregular grids corresponding to each dependent variable.

The solution for the velocity field $\vec{V}$, $\tau_{11}$, $\tau_{12}$ components of the stress tensor, pressure $p$, $\tau_{11} + \tau_{22}$ and the corresponding computational grid for case I are shown in Figs. 4-9 for four different times. The location of the diapir can be easily seen on Fig. 4. Due to the low viscosity, the velocity inside of the diapir is substantially higher than in the surrounding mantle. The build up of the high stress region in the lithosphere just above the diapir is illustrated in Fig. 5. We call this phenomenon a stress-focusing effect. The absolute value of the shear stress component $\tau_{12}$ is shown in Fig. 6. In contrast to the $\tau_{11}$ stress tensor component the shear stress component $\tau_{12}$ is practically zero in the lithosphere, which is due to the high viscosity of the lithosphere, which results in substantial decrease of the velocity and the absence of the shear. The distribution of the absolute value of the pressure is shown in Fig. 7. High pressure region is concentrated just under the lithosphere and above the di-
Figure 10. Absolute value of the $\tau_{11}$ component of the stress tensor for cases I-IV of Table 1 for the same diapir location.

Figure 11. Absolute value of the $\tau_{12}$ component of the stress tensor for cases I-IV of Table 1 for the same diapir location.

Figure 12. Absolute value of the pressure $p$ for cases I-IV of Table 1 for the same diapir location.

Figure 13. Absolute value of $\tau_{11} + \tau_{22}$ for cases I-IV of Table 1 for the same diapir location.

Viscoelastic Plume-Lithosphere Interaction

...
results in lower values of the \( \tau_{11} \) component of the stress tensor, appearance of the stronger wake region and lower gradients for all the components of the stress tensor and pressure. The absence of very large gradient can be easily seen by comparing computation grids for these two cases. We see that computational grid has less grid points, which is the signature of lower gradients.

Comparison of the results of Figs. 10-14 for cases II and III illustrate the stress focusing phenomenon is the signature of the lithosphere, the high viscosity region. The thickness of the lithosphere for the case III is four times less then the one for the case II. To emphasize the stress focusing effects, the zoomed-in view of the high viscosity region for \( \tau_{11} \) component of the stress tensor, pressure, and \( \tau_{11} + \tau_{22} \) are shown in Figs. 15-17 respectively. As it is expected, the stress for the case III is focussed in four times thinner lithosphere when compared to the results of case II. We also note that the numerical algorithm adapt to the thinning of the lithosphere by increasing the resolution only in the region of the lithosphere. It can be easily seen comparing computational grid for cases II and III shown in Fig. 14.

Comparison of the results of Figs. 10-14 for cases II and
Viscoelastic Plume-Lithosphere Interaction

Figure 18. Absolute value of the velocity $|\vec{V}|$ and velocity vectors for the case V at four different times.

Figure 19. Absolute value of the $\tau_{11}$ component of the stress tensor for the case V at four different times.

Figure 20. Absolute value of the $\tau_{12}$ component of the stress tensor for the case I at four different times.

Figure 21. Absolute value of the pressure $p$ for the case V at four different times.

Figure 22. Absolute value of $\tau_{11} + \tau_{22}$ for the case V at four different times.

IV illustrate a very important point of the stress focusing phenomenon. If one looks at the boundary of the high stress, then it is located at approximately the same viscosity values. This is because the velocity is decreases substantially because of high viscosity and it is not that important whether the viscosity is only ten or hundred times higher, what is important is the viscosity contrast. If one base the lithosphere thickness on the value of $\lambda$, then one should expect to have the same lithosphere thickness. However it is not the case, as it easily can be seen on the figures. If one base the definition of thickness based on the absolute value of the viscosity, then the lithosphere thickness should be approximately 1.4 times larger for the case IV when compared with the case II. This is confirmed in Figs. 10-14 and especially in zoomed-in view shown in Figs. 15-17.

The long-term viscoelastic memory effects on the stress distribution are shown in Figs. 18-22, where the solution for the velocity field $\vec{V}$, $\tau_{11}$, $\tau_{12}$ components of the stress tensor, pressure $p$, and $\tau_{11} + \tau_{22}$ for case V are shown. This solution has been obtained by setting the rising speed of the diapir velocity to zero and then monitoring the subsequent relaxation. This condition effectively locks the diapir at a fixed position. But the stress distribution then decreases with time from the viscoelastic mechanism. This process takes a long time on the order of ten Maxwell relaxation times.

5 CONCLUSIONS

In this paper we have demonstrated the prowess of the dynamically adaptive wavelet collocation method to approach the regime of realistic mantle-lithospheric interaction by going to thin layers, on the order of $O(10)$ km and viscosity contrasts up to $10^5$. The multilevel wavelet approximation has allowed for local grid refinement up to an arbitrary small scale, on the order of 1 km, without a drastic increase in the number of collocation points, about $10^5$. Conversely, this approach can be used in 3-D with $10^7$ points for obtaining a resolution in 3-D comparable to the one obtained here for 2-D.

We have developed a new wavelet-based method for solving partial differential equations in a geophysical situation involving viscoelastic diapiric flows with strong localized variations in the viscosity. We have verified the early findings of stress amplification in a thin very viscous layer (Polakow et al. 1993; Podlachikov et al. 1993) in that several hundreds of MPa were found to develop in the lithosphere. This raises the possibility for the potential breaking of the
viscoelastic lithosphere by the interaction with a rising diapir. With the use of the highly adaptive method, we can reach much thinner lithosphere, on the order of a few km, as compared to more conventional finite-difference method used by Poliakov et al. (1993). Future work would entail going to 3-D adopting this method to treat the mechanism of stress advection in the finite-amplitude regime (Schmalholz & Podladchikov 1999), to incorporate thermal-mechanical effects into localization situations (Kameyama et al. 1999) and to include a geophysically realistic yield stress criterion (Kohlstedt et al. 1995) allowing for the potential seavage of the lithosphere. Moresi and Solomatov (1998) have already shown the possibilities of lithospheric breaking by the vigor of thermal convection.

ACKNOWLEDGMENTS

We acknowledge fruitful discussions with Arkady Ten and A. Bobby Poliakov. This research has been supported by the geophysics program of National Science Foundation.

REFERENCES


