Institute for Parallel Processing
Bulgarian Academy of Sciences
(IPP-BAS)

A comparison between two approaches to solve the equations of linear isostasy

Erik Bängtsson* and Bjorn Lund†

*Department of Information Technology, Uppsala University
Erik.Bangtsson@it.uu.se

†Department of Geophysics, Uppsala University SWEDEN
Bjorn.Lund@geo.uu.se

BIS 21++ report No. 2006-03 April 2006

Acad. G. Bonchev, Bl. 25A, 1113 Sofia, BULGARIA

http://bis-21pp.acad.bg

A comparison between two approaches to solve the equations of linear isostasy

Erik Bängtsson*and Björn Lund†

April 11, 2006

Abstract

We compare two approaches to compute the isostatic response of the Earths lithosphere to an external load. The lithosphere is modelled as an incompressible, linear elastic solid. The two approaches differ in the formulation of the problem and the capability to give accurate solutions in the incompressible limit, the choice of the finite elements used for discretization, and the solution strategy for the arising algebraic problem. Numerical experiments show that the when the two approaches are comparable, they give identical results.

1 Introduction

In many fields of science, due to practical, technical, and/or economical obstacles, it is not possible to perform classical experiments to obtain answers to our questions. In geophysics, for example, where the length and time scales are enormous, laboratory and field experiments are impossible to perform due to sheer size.

One particular problem from this field which has drawn lots of attention lately, is to simulate the response of the outer part of the Earth, the lithosphere, to glaciation and deglaciation. This in order to enhance the safety predictions for long term nuclear waste repositories in the bedrock of, for example, Scandinavia and Canada.

The scope of this paper is to compare the accuracy and efficiency of two approaches to discretize the equations of linear isostasy for glacial rebound, and to solve the arising linear system of equations. The first approach is to reformulate the equations of linear isostasy

*Department of Information Technology, Uppsala University
†Department of Geophysics, Uppsala University

onto a form that is suitable for standard finite element (FE) packages, see for example [8] and the references therein. The main problem with this approach is that the lithosphere is modelled as a purely incompressible solid, and the commercial packages often cannot handle the fully incompressible case.

The second approach is to pose the problem on mixed form which is well posed also for the fully incompressible solid, and solve the corresponding equations. For this we have developed our own code, which is not only capable of handling the purely incompressible solid, but also includes the first order terms in the equation of linear isostasy.

This paper is organized as follows. Section 2 contains a description of the problem we want to solve, and in Section 3 and Section 4 the two solution approaches are described. Section 5 contains a results from numerical experiments, and the paper concludes with Section 6 where some conclusions are drawn.

2 Problem description

The moment balance equation for a (visco)elastic, pre-stressed body in a constant gravity field reads,

$$-\nabla \cdot \sigma - \nabla(\mathbf{u} \cdot \nabla \mathbf{p_0}) + (\nabla \cdot \mathbf{u})\nabla \mathbf{p_0} = \mathbf{f}, \tag{1}$$

where σ is the Cauchy stress tensor, $\mathbf{u}$ are the displacements, p_0 is the so-called *pre-stress*, and $\mathbf{f}$ is a body force. The third term on the right hand side of Equation (1) describes the *bouyancy* of the compressed material, and it vanishes for purely incompressible material since $\nabla \cdot \mathbf{u} = \mathbf{0}$. The second term describes the advection of the pre-stress, which is a stress that is present in the solid before the application of the external load. For further details on the model and the origin of Equation (1) see, for example, [8].

Under the assumptions that the lithosphere is a homogeneous, isotropic, linear, and purely elastic solid, the Cauchy stress tensor is given by Hookes law,

$$\sigma(\mathbf{u}) = \mu\varepsilon(\mathbf{u}) + \lambda(\nabla \cdot \mathbf{u}), \tag{2}$$

where $\varepsilon(\mathbf{u}) = \mathbf{0.5}(\nabla\mathbf{u}+\nabla\mathbf{u}^{\mathbf{T}})$ is the strain tensor, and $\mu = E/2/(1+\nu)$, $\lambda = 2\mu\nu/(1-2\nu)$ are the Lamé coefficients. The material parameters E and ν are the Young modulus and the Poisson number, respectively.

Remark 2.1 *The description of the lithosphere as a purely elastic solid is a simplification, but nevertheless important to study. Efficient solution techniques for the purely elastic problem are of great* *importance for the simulation of a viscoelastic solid since the response of a viscoelastic solid at a given time t_k is a combination of the instantaneous, purely elastic response, and a memory term, an integral over previous responses. The memory term is numerically computed as a sum over weighted Hookean responses at all previous times t_i, $i = 0, \dots, k-1$, and these responses must be computed efficiently in order to achieve an efficient viscoelastic solver.*

Remark 2.2 *The lithosphere is assumed to be homogeneous with respect to the material parameters λ and μ. This is of coarse a simplification, but this assumption does not violate the generality of the obtained results.*

After combining Equation (1) and Equation (2), we arrive at the equations of linear isostasy, formulated in terms of the displacements $\mathbf{u}$,

$$-\nabla \cdot (\mu\varepsilon(\mathbf{u})) - \nabla(\mathbf{u} \cdot \nabla \mathbf{p_0}) + (\nabla \cdot \mathbf{u})\nabla \mathbf{p_0} - \lambda \nabla(\nabla \cdot \mathbf{u}) = \mathbf{f}, \qquad (3)$$

and a finite element discretization of Equation (3) gives rise to the algebraic problem

$$\mathcal{A}\mathbf{x} = \mathbf{b}, \qquad (4)$$

where $\mathcal{A} \in \mathbb{R}^{N_u \times N_u}$ is a large and sparse matrix, and $\mathbf{x} \in \mathbb{R}^{\mathbf{N_u}}$ and $\mathbf{b} \in \mathbb{R}^{\mathbf{N_u}}$ are vectors.

The two compared approaches differ in the formulation and FE discretization of Equation (3), and the solution strategy for Equation (4).

3 Approach I: ABAQUS

The first approach is based on the commercial finite element package ABAQUS. The code being commercial has the advantage that it is throughly debugged and optimized, but the disadvantage that it is not designed to solve problems of the type of Equation (1) or Equation (3), but rather the standard form of the moment balance equation for an elastic solid

$$-\nabla \cdot \sigma(\mathbf{u}) = \mathbf{f},$$

that is, a problem without the first order terms arising from the prestress advection and the bouyancy.

The remedy for this is to introduce the modified stress tensor

$$\mathbf{T}(\mathbf{u}) = \sigma(\mathbf{u}) - \mathbf{u} \cdot \nabla \mathbf{p_0} \mathbf{I}, \qquad (5)$$

where I is the identity tensor, and solve

$$-\nabla \cdot \mathbf{T}(\mathbf{u}) = \mathbf{f} \tag{6}$$

instead. The solution to Equation (3) is retrieved from the solution to Equation (6) in a post-processing step, see for example [8] for details.

The algebraic problem $\mathcal{A}\mathbf{x} = \mathbf{b}$ is solved with a direct solver, provided by ABAQUS.

Remark 3.1 *In Equation* (5)*, the bouyancy term* $(\nabla \cdot \mathbf{u})\nabla \mathbf{p_0}$ *is omitted because this formulation is tailored to the case of incompressible material (*$\nu = 0.5$*), where this term vanishes. In the problem formulation in Section 4, the bouyancy term is present for the sake of completeness of the analysis, but in the experiments reported in Section 5 the bouyancy is omitted also in Approach II.*

4 Approach II: Mixed u-p-formulation

In the incompressible limit, $\nu \to 0.5$, and the Lamé coefficient $\lambda \to \infty$. This makes the problem in Equation (2) ill-posed, and the corresponding stiffness matrix in Equation (4) extremely ill-conditioned. This is the mathematical formulation of the phenomenon known as *volumetric locking*, which leads to erroneous results when solving Equation (3) in the nearly incompressible limit. See, for example, [6], for further details on the locking effect.

The usual remedy to the locking problem is to introduce the *kinematic pressure* $p = \frac{\mu}{\lambda}\nabla \cdot \mathbf{u}$, and reformulate Equation (3) on mixed form, which yields

$$\begin{cases} -\nabla \cdot (\mu\varepsilon(\mathbf{u})) - \nabla(\mathbf{u} \cdot \nabla \mathbf{p_0}) + (\nabla \cdot \mathbf{u})\nabla \mathbf{p_0} - \mu\nabla \mathbf{p} = \mathbf{f} \\ \mu\nabla \cdot \mathbf{u} - \dfrac{\mu^{\mathbf{2}}}{\lambda}\mathbf{p} = 0. \end{cases} \tag{7}$$

4.1 Finite Element Discretization

In this subsection we look into the properties of the finite element problem corresponding to Equation (7) and derive a necessary boundary condition to ensure the solvability and stability of the FE formulation of it, which reads

$$\begin{array}{l} \text{Find } \mathbf{u} \in \mathbf{V} \subset \mathbf{H^1} \text{ and } p \in P = \{p \in L^2 : \int_\Omega p = 0\} \text{ such that} \\ \begin{array}{rcll} a(\mathbf{u}, \mathbf{v}) + \mathbf{b}(\mathbf{v}, \mathbf{p}) & = & f(v) + \langle \mathbf{l}, \mathbf{v_h} \rangle & \forall \mathbf{v} \in \mathbf{V}, \\ b(\mathbf{u}, \mathbf{q}) - \mathbf{c}(\mathbf{p}, \mathbf{q}) & = & 0, & \forall q \in P. \end{array} \end{array} \tag{8}$$

The bilinear forms in Equation (8) are

$$
\begin{aligned}
a(\mathbf{u}, \mathbf{v}) &= \int_{\Omega} \mu\varepsilon(\mathbf{u}) : \varepsilon(\mathbf{v}) - \nabla(\mathbf{u} \cdot \mathbf{b}) \cdot \mathbf{v} + (\nabla \cdot \mathbf{u})(\mathbf{c} \cdot \mathbf{v}) \mathbf{d\Omega} \\
b(\mathbf{u}, \mathbf{p}) &= \int_{\Omega} \mu(\nabla \cdot \mathbf{u}) \mathbf{p}\, \mathbf{d\Omega} \qquad \mathbf{c}(\mathbf{p}, \mathbf{q}) = \int_{\mathbf{\Omega}} \frac{\mu^{\mathbf{2}}}{\lambda} \mathbf{p}\, \mathbf{q}\, \mathbf{d\Omega} \\
\langle \mathbf{l}, \mathbf{v} \rangle &= \int_{\Gamma} \mathbf{v} \cdot \mathbf{l}\, \mathbf{d\Gamma}
\end{aligned} \tag{9}
$$

.

where the vector fields $\mathbf{b}$, $\mathbf{c}$, and $\mathbf{f}$ are introduced to generalize the problem. The boundary of the computational domain Ω is denoted with Γ.

Before we look into the properties of $a(\cdot, \cdot)$, $b(\cdot, \cdot)$, and $c(\cdot, \cdot)$, let us state some prerequisites.

A solution to the variational problem (8) exists and is unique if $a(\mathbf{u}, \mathbf{v})$, $c(p, p)$ and $b(\mathbf{u}, \mathbf{p})$ are bounded,

$$
\begin{aligned}
a(\mathbf{u}, \mathbf{v}) &\leq \overline{a} \|\mathbf{u}\|_{\mathbf{V}} \|\mathbf{v}\|_{\mathbf{V}} \quad \forall \mathbf{u}, \mathbf{v} \in \mathbf{V} && (10) \\
b(\mathbf{v}, \mathbf{p}) &\leq \overline{b} \|\mathbf{v}\|_{\mathbf{V}} \|\mathbf{p}\|_{\mathbf{P}} \quad \forall \mathbf{u} \in \mathbf{V},\ \mathbf{p} \in \mathbf{P} && (11) \\
c(p, q) &\leq \overline{c} \|p\|_P \|q\|_P \quad \forall p, q \in P, && (12)
\end{aligned}
$$

and if $a(\mathbf{u}, \mathbf{u})$ and $c(p, p)$ are coercive on $\mathbf{V}$ and P, respectively. That is, if

$$
\begin{aligned}
a(\mathbf{u}, \mathbf{u}) &\geq \underline{a} \|\mathbf{u}\|^{\mathbf{2}}_{\mathbf{V}}, \quad \underline{\mathbf{a}} > \mathbf{0} \quad \forall \mathbf{u} \in \mathbf{V} && (13) \\
c(p, p) &\geq \underline{c} \|p\|^2_P, \quad \underline{c} > 0 \quad \forall p \in P. && (14)
\end{aligned}
$$

As is clear from Equation (9) $c(p, q) = 0$, $\forall p, q \in P$ corresponds to $\nu = 0.5$. In this case, , Equation (8) is solvable if

- the conditions in Equation (10) - (12) hold,
- $a(\mathbf{u}, \mathbf{u})$ is coercive on the null-space of $b(\mathbf{u}, \mathbf{q})$,
- $b(\mathbf{u}, \mathbf{q}) = \mathbf{0} \quad \Rightarrow \mathbf{q} = \mathbf{0} \quad \forall \mathbf{u} \in \mathbf{V}$.

Furthermore, Equation (8) is stable if the following inf-sup (or Ladyzhenskaya-Babuška-Brezzi or LBB) conditions are fulfilled,

$$
\inf_{\mathbf{u} \in \mathbf{V}} \sup_{\mathbf{v} \in \mathbf{V}} \frac{a(\mathbf{u}, \mathbf{v})}{\|\mathbf{u}\|_{\mathbf{V}} \|\mathbf{v}\|_{\mathbf{V}}} \geq \underline{a}' > 0, \tag{15}
$$

and

$$
\inf_{q \in P} \sup_{\mathbf{v} \in \mathbf{V}} \frac{b(\mathbf{u}, \mathbf{q})}{\|\mathbf{v}\|_{\mathbf{V}} \|\mathbf{q}\|_{\mathbf{P}}} \geq \underline{b} > 0. \tag{16}
$$

Note that when $a(\mathbf{u}, \mathbf{v})$ is coercive, Equation (15) is automatically satisfied. See, for example [7] for details.

The coercivity of $c(\cdot, \cdot)$ is straightforwardly seen and Equation 16 is guaranteed by the theory for the Stokes problem [6], but for $a(\cdot, \cdot)$, the situation is more complicated. In [3] we show that the bilinear forms in Equation (8) are bounded, but that $a(\mathbf{u}, \mathbf{v})$, in general is not coercive due to the first order terms, and the rest of this section will be devoted to a discussion on the coercivity of $a(\mathbf{u}, \mathbf{v})$, in general and for some special choices of boundary conditions and finite element spaces.

In order to clarify the coercivity of $a(\mathbf{u}, \mathbf{u})$, we split it into two parts,

$$a(\mathbf{u}, \mathbf{v}) = \widehat{\mathbf{a}}(\mathbf{u}, \mathbf{v}) + \widetilde{\mathbf{a}}(\mathbf{u}, \mathbf{v}), \tag{17}$$

where

$$\widehat{a}(\mathbf{u}, \mathbf{v}) = \int_{\mathbf{\Omega}} \mu \varepsilon(\mathbf{u}) : \varepsilon(\mathbf{v}), \tag{18}$$

and

$$\widetilde{a}(\mathbf{u}, \mathbf{u}) = -\int_{\mathbf{\Omega}} \nabla(\mathbf{u} \cdot \mathbf{b}) \cdot \mathbf{v} + \int_{\mathbf{\Omega}} (\nabla \cdot \mathbf{u})(\mathbf{c} \cdot \mathbf{v}). \tag{19}$$

When homogeneous Dirichlet conditions are imposed on $\Gamma_0 \subset \Gamma$, Korns inequality gives

$$K_2 \mu \|\mathbf{u}\|_{\mathbf{1}}^{\mathbf{2}} \leq \widehat{\mathbf{a}}(\mathbf{u}, \mathbf{u}) \leq \mathbf{K_1} \mu \|\mathbf{u}\|_{\mathbf{1}}^{\mathbf{2}} \qquad \forall \mathbf{u} \in \mathbf{V}, \tag{20}$$

for $K_1, K_2 > 0$, and the coercivity of $\widehat{a}(\mathbf{u}, \mathbf{u})$, the elastic part of $a(\mathbf{u}, \mathbf{u})$ is guaranteed.

From Equation (17) it follows that

$$a(\mathbf{u}, \mathbf{u}) \geq \widehat{\mathbf{a}}(\mathbf{u}, \mathbf{u}) - |\widetilde{\mathbf{a}}(\mathbf{u}, \mathbf{u})| \tag{21}$$

that is, if $\widetilde{a}(\mathbf{u}, \mathbf{u})$ is small enough, $a(\mathbf{u}, \mathbf{u})$ is coercive. In [3], we show that

$$\begin{aligned} |\widetilde{a}(\mathbf{u}, \mathbf{v})| &\leq d(\alpha_1 + \beta)\|\mathbf{u}\|_{\mathbf{1}} \|\mathbf{v}\|_{\mathbf{0}} + \mathbf{d}\alpha_{\mathbf{2}} \|\mathbf{u}\|_{\mathbf{0}} \|\mathbf{v}\|_{\mathbf{0}} \\ &\leq d(\alpha_1 + \beta + \alpha_2)\|\mathbf{u}\|_{\mathbf{1}} \|\mathbf{v}\|_{\mathbf{1}}, \end{aligned} \tag{22}$$

where d denotes the number of spatial dimensions and

- $|b_i(\mathbf{x}) \leq \alpha_{\mathbf{1}}|$, $i = 1, \ldots, d$,
- $|\nabla \cdot \mathbf{b}| \leq \alpha_{\mathbf{2}}$, and
- $|\mathbf{c}| \leq \beta$.

By combining Equation (20) - (22), we see that for

$$\mu K_1 - d(\alpha_1 + \beta + \alpha_2) > 0, \tag{23}$$

$a(\mathbf{u}, \mathbf{v})$ is coercive.

Unfortunately, in practice the Korn constant K_1 is not known, and the estimate in Equation (23) is not very useful. Therefore, we limit ourselves to the investigation of the coercivity of $a(\mathbf{u},\mathbf{v})$ on the kernel of $b(\cdot,\cdot)$, that is, the space of divergence-free functions

$$V^0 = \{v \in H^1 : \nabla \cdot v = 0\}.$$

After a reformulation of $a(\mathbf{u},\mathbf{v})$ as

$$a(\mathbf{u},\mathbf{v}) = \int_{\Omega} \mu\varepsilon(\mathbf{u}) : \varepsilon(\mathbf{v}) + \int_{\mathbf{\Omega}} (\mathbf{u}\cdot\mathbf{b})(\nabla\cdot\mathbf{v}) + \int_{\mathbf{\Omega}} (\nabla\cdot\mathbf{u})(\mathbf{c}\cdot\mathbf{v})\mathbf{d\Omega}$$
$$- \int_{\Gamma} (\mathbf{u}\cdot\mathbf{b})(\mathbf{n}\cdot\mathbf{v})\,\mathbf{d\Gamma},$$

we see that it reduces to

$$a(\mathbf{u},\mathbf{v}) = \int_{\mathbf{\Omega}} \mu\varepsilon(\mathbf{u}) : \varepsilon(\mathbf{v}) - \int_{\mathbf{\Gamma}} (\mathbf{u}\cdot\mathbf{b})(\mathbf{n}\cdot\mathbf{v})\,\mathbf{d\Gamma}, \text{ for } \mathbf{u},\mathbf{v} \in \mathbf{V^0}. \quad (24)$$

Clearly, $a(\mathbf{u},\mathbf{v})$ is coercive for some special (combinations of) conditions on $\mathbf{b}$, $\mathbf{u}$, and $\mathbf{v}$ on Γ such as

- $\mathbf{b} = \mathbf{0}$ on Γ,
- $\mathbf{u} = \mathbf{0}$ on Γ,
- $\mathbf{u} \perp \mathbf{b}$ on Γ,
- and $\mathbf{v} \perp \mathbf{n}$ on Γ.

Those conditions are, with an exception for the Dirichlet condition $\mathbf{u} = \mathbf{0}$, fairly unphysical. A more realistic set of boundary conditions is,

$$\begin{cases} \mathbf{u} &= 0 & \text{on } \Gamma_0 \\ \sigma\mathbf{n} &= 0 & \text{on } \Gamma_1 \\ \sigma\mathbf{n} &= \mathbf{l} & \text{on } \Gamma_{\mathbf{l}} \end{cases}, \quad (25)$$

where $\mathbf{l}$ is an external load. Note that Equation (25) is a specification of the situation in Equation (8).

For the further analysis of the coercivity of $a(\cdot,\cdot)$ on V^0, we observe that

$$\varepsilon(\mathbf{v}) = \mu\nabla(\nabla\cdot\mathbf{v}) + \mu\nabla\times\nabla\times\mathbf{v}, \quad (26)$$

and that the first term of the right hand side vanishes on V^0. Hence, we get that

$$a(\mathbf{u},\mathbf{v}) = \int_{\mathbf{\Omega}} \mu\nabla\times\mathbf{u}\cdot\nabla\times\mathbf{v} + \int_{\mathbf{\Gamma_L}} \mu\mathbf{v}\cdot\mathbf{n}\times\nabla\times\mathbf{u} - \int_{\mathbf{\Gamma_L}} (\mathbf{u}\cdot\mathbf{b})(\mathbf{n}\cdot\mathbf{v}), \quad (27)$$

where $\Gamma_L = \Gamma_1 \cup \Gamma_\mathbf{l}$.

Development of the second integral (the first boundary integral) in Equation (27) yields

$$\int_{\Gamma_L} \mu \mathbf{v} \cdot \mathbf{n} \times \nabla \times \mathbf{u} =$$

$$\int_{\Gamma_L} \mu \mathbf{v} \cdot \{\nabla(\mathbf{u}\cdot\mathbf{n}) - (\mathbf{n}\cdot\nabla)\mathbf{u} - (\mathbf{u}\cdot\nabla)\mathbf{n} - \mathbf{u}\times(\nabla\times\mathbf{n})\}$$

$$= \int_{\Gamma_L} \mu v_j \{\partial_j(u_i n_i) - (n_l\partial_l)u_j - (u_k\partial_k)n_j - \epsilon_{jkl}\epsilon_{lmn}u_k\partial_m n_n\}$$

$$= \int_{\Gamma_L} \mu v_j \{u_i\partial_j n_i - (u_k\partial_k)n_j + n_i\partial_j u_i - (n_l\partial_l)u_j - \epsilon_{jkl}\epsilon_{lmn}u_k\partial_m n_n\}$$

$$= \int_{\Gamma_L} \mu \mathbf{v} \cdot \left\{(\mathbf{n}\cdot\nabla\mathbf{u}^\mathbf{T} - \mathbf{n}\cdot\nabla\mathbf{u}) + (\mathbf{u}\cdot\nabla\mathbf{n}^\mathbf{T} - \mathbf{u}\cdot\nabla\mathbf{n}) - \mathbf{u}\times(\nabla\times\mathbf{n})\right\}$$

$$= 2\int_{\Gamma_L} \mu \mathbf{v} \cdot \left\{-\mathbf{n}\cdot\mathbf{W}(\mathbf{u}) + \mathbf{u}\cdot(\nabla\mathbf{n}^\mathbf{T} - \nabla\mathbf{n}) - \mathbf{u}\times(\nabla\times\mathbf{n})\right\},$$

where ϵ_{ijk} is the Levy-Chevita symbol and the summation convention $u_i v_i = \sum_i u_i v_i$ is assumed. The tensor $W(\mathbf{v}) = \mathbf{0.5}(\nabla\mathbf{v} - \nabla\mathbf{v}^\mathbf{T})$, the *spin tensor*, is the skew-symmetric part of the infinitesimal displacement tensor. The strain tensor $\varepsilon(\mathbf{v})$ is the symmetric part.

Straightforward vector calculus reveals that, at least in two space dimensions, $(\mathbf{u}\cdot\nabla\mathbf{n}^\mathbf{T} - \mathbf{u}\cdot\nabla\mathbf{n}) - \mathbf{u}\times(\nabla\times\mathbf{n}) = \mathbf{0}$, and we can therefore establish that

$$a(\mathbf{u},\mathbf{u}) \geq \mathbf{C}\|\mathbf{u}\|_\mathbf{1} \quad \forall \mathbf{u} \in \mathbf{V_0},$$

where C is a positive constant, holds when

$$2\mu\mathbf{n}\cdot\mathbf{W}(\mathbf{u}) + \mathbf{n}(\mathbf{u}\cdot\mathbf{b}) = \mathbf{0} \qquad \text{on } \mathbf{\Gamma_L}. \tag{28}$$

That is, to ensure the coercivity of $a(\cdot,\cdot)$ on V^0, an additional boundary condition, which ensures that the amount of stress that is carried to the boundary by the pre-stress advection is balanced by a rotational stress on the boundary must be applied to the system .

4.2 Algebraic Problem

After the FE discretization, the algebraic problem corresponding to Equation (7) takes the form,

$$\mathcal{A}\mathbf{x} = \begin{bmatrix} M & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix}, \tag{29}$$

where $M \in \mathbb{R}^{N_u \times N_u}$ is non-symmetric, sparse and in general indefinite, and $C^{N_p \times N_p}$ is positive semi-definite.

The algebraic problem in Equation 29 is solved with an iterative solution method, preconditioned by the block-lower triangular matrix

$$\mathcal{D} = \begin{bmatrix} D_1 & 0 \\ B & D_2 \end{bmatrix}, \tag{30}$$

where the first diagonal block D_1 approximates M, and the second diagonal block, D_2, approximates the negative Schur complement of $\mathcal{A}$, $S = C + BM^{-1}B^T$. See for example [5] for further details on preconditioners and solution techniques for saddle point problems.

The construction of D_2 is inspired by the assembly of $\mathcal{A}$. The global stiffness matrix is assembled from element stiffness matrices

$$\mathcal{A} = \sum_E \mathcal{A}_E \quad \text{where} \quad \mathcal{A}_E = \begin{bmatrix} M_E & B_E^T \\ B_E & -C_E \end{bmatrix},$$

and for non-singular M_E, we compute the exact Schur complement of $\mathcal{A}_E$, $S_E = C_E + B_E M_E^{-1} B^T$. We then assemble these local contributions to form D_2, $D_2 = \sum_E S_E$. See [3] for details.

In order to ensure the non-singularity of M_E on elements away from a Dirichlet boundary, M_E is regularized by a small positive number,

$$\widetilde{M}_E = M + \epsilon h^2 I, \tag{31}$$

where $\epsilon > 0$ is an arbitrary parameter and h is the discretization parameter. The nonsingular $\widetilde{M}_E$ is used to to compute the modified local Schur complements $\widetilde{S}_E$, $C_E + B_E \widetilde{M}_E^{-1} B^T$, which are used in the assembly of D_2, $D_2 = \sum_E \widetilde{S}_E$.

When $\mathcal{D}$ is applied, D_1 and D_2 are solved with an inner iterative solution method, preconditioned by a robust multilevel preconditioner, see [4] for details.

In this Approach, we use a code developed by the first author. This code is based on the open source packages `PETSc` [1] and `deal.II` [2].

5 Experiments

The numerical experiments are performed on the following problem setting:

Problem 5.1 *A 2D flat Earth model, which is symmetric with respect to $x = 0$, is subjected to a Heaviside load of a 1000 km wide and 2 km thick ice sheet. The size of the domain is 10 000 km width and*

4000 km depth and the boundary conditions are homogeneous Dirichlet conditions on the boundary $y = -4000$ km and symmetry conditions on the boundary $x = 0$. Homogeneous Neumann conditions on the boundary $x = 10000$ km and on the boundary segment $y = 0$, $x > 1000$ km. The Young's modulus of the solid is 400 GPa, and its density ρ_r is 3000 kg m^{-3}. The density of the ice is 981 kg m^{-3}. The Poisson ratio $\nu \in [0.2, 0.5]$.

The pre-stress is hydrostatic, i.e.

$$p_0 = -\rho_r g \mathbf{x} \cdot \mathbf{e_d}$$

where g=9.81 $kgms^{-2}$ is the gravity constant and e_d is the unit vector directed downwards.

Remark 5.1 *For Problem 5.1 the bilinear form $a(\mathbf{u}, \mathbf{v})$ can be shown to be coercive for incompressible material, but for compressible materials the situation is not clear. The numerical experiments on the other hand do not show any kind of unstable behaviour for $\nu < 0.5$, which indicates that the advection term is small enough, and the estimate in Equation* (23) *holds.*

The domain is discretized with uniform rectangular finite elements. In Approach I standard bilinear basis functions are used, whereas in Approach II, a pair of stable, modified Taylor–Hood ($Q1$ -iso $Q1$) bilinear basis functions are used. That is, the basis functions for the displacements $\mathbf{u}$ live on a mesh that is a uniform refinement of the mesh on which the pressure variables p live. The meshes are chosen such that the number of displacement degrees of freedom are the same in Approach I and Approach II.

In Approach II, the generalized conjugate gradient - minimized residual (GCG-MR) method is chosen as iterative scheme for the outer and the two inner iterative solver. For robustness and efficiency reasons the preconditioner D_1 is taken as a multilevel preconditioner for the scaled vector-valued Laplacian matrix

$$(D_1)_{ij} = \sum_{k=1}^{d} \nabla v_k^{(i)} \cdot \nabla v_k^{(i)},$$

instead for the matrix M, see [4] for details. The algebraic problem in Equation (29) is solved to an accuracy of six orders of magnitude relative to the initial residual, and the inner iterative solvers for D_1 and D_2 are solved to a relative accuracy of 0.5 and 0.1, respectively. This choice of these parameters, and the multilevel preconditioner for D_1, are found to give the smallest overall solution times.

N	$\mathbf{u(0.5,0.1)}$	$p(0.5,0.1)$	$\mathbf{u(1,0.1)}$	$p(1,0.1)$
1583	4e-07	1e-08	4e-07	2e-08
6043	4e-10	6e-09	4e-10	6e-09
23603	5e-08	9e-09	5e-08	9e-09
	$\mathbf{u(5,0.1)}$	$p(5,0.1)$	$\mathbf{u(10,0.1)}$	$p(10,0.1)$
1583	4e-07	6e-08	4e-07	1e-07
6043	4e-07	6e-08	4e-07	8e-08
23603	5e-08	9e-09	5e-08	9e-09

Table 1: The relative error in the solution, depending on the parameter ϵ in Equation (). The errors are measured in l^2-norm, that is $\mathbf{u}(\epsilon_\mathbf{1}, \epsilon_\mathbf{2}) = \|\mathbf{u}_{\epsilon_\mathbf{1}} - \mathbf{u}_{\epsilon_\mathbf{2}}\|_{\mathbf{l^2}} / \|\mathbf{u}_{\epsilon_\mathbf{2}}\|_{\mathbf{l^2}}$, and similarly for p. N denotes the problem size. The Poisson number $\nu = 0.3$.

The problem sizes N respited in the Tables 1 - 5 refer to the size of the algebraic problem in Approach II, which, due to the choice of finite elements, is approximately 11 % larger than the algebraic problem in Approach I.

5.1 Results

In Table 1, the dependency of the solution from Approach II on the parameter ϵ in Equation (31) for different problem sizes is shown. The errors in Table 1 are measured relative to the solution corresponding to the smallest regularization parameter ($\epsilon = 0.1$), and as is clear from the table, the impact of ϵ is negligible in the parameter range and for the problem sizes investigated.

In Tables 2, 3, and 4, relative errors for different problem sizes and Poisson number are shown on two different depths in the lithosphere, on the surface (0 km) and in the vicinity of the transition region between the crust and the upper mantle (62.5 km), the Mohorovičić region. The experiments reported in Tables 2 and 3 are performed to verify that the two approaches deliver the same result, and therefore the advection of the pre-stress is omitted because the difference in the treatment of this term between Approach I and II should lead to different results.

Table 2 shows the relative error in the solutions from Approach I and II for two values of ν far from the incompressible limit, that is, in the region of the parameter range where Approach I should not suffer from locking effects. For increasing problem size, the error approaches the error approaches the precision in the output from `PETSc` (five digits).

ν	0.2		0.3	
depth(km)	0	62.5	0	62.5
N = 23603	0.00054601	0.00041437	0.00090089	0.00052987
N = 93283	0.00018977	9.4125e-05	0.00033034	0.00011011
N = 370883	3.5814e-05	3.5327e-05	5.4699e-05	3.7304e-05

Table 2: The relative error in l^2-norm between the displacement from the two solution approaches. N denotes the problem size, and ν denotes the Poisson number.

depth(km)	0	62.5
$\nu = 0.2$	3.5814e-05	3.5327e-05
$\nu = 0.3$	5.4699e-05	3.7304e-05
$\nu = 0.4$	0.00013178	4.3554e-05
$\nu = 0.45$	0.00018893	5.4707e-05
$\nu = 0.47$	0.00021319	5.7848e-05
$\nu = 0.49$	0.00023803	6.4616e-05
$\nu = 0.4999$	0.00024976	6.5966e-05
$\nu = 0.49999$	0.00024983	6.559e-05
$\nu = 0.499999$	0.0002501	6.5616e-05
$\nu = 0.4999999$	0.0002505	6.6535e-05

Table 3: $\|\mathbf{u_I} - \mathbf{u_{II}}\|_{\mathbf{l^2}} / \|\mathbf{u_{II}}\|_{\mathbf{l^2}}$ for different Poisson number ν, where $\mathbf{u_I}, \mathbf{i} = \mathbf{I}, \mathbf{II}$ is the solution from approach i. The pre-stress advection is omitted and the problem size is $N = 370883$.

depth(km)	0	62.5
$\nu = 0.2$	0.13425	0.13442
$\nu = 0.3$	0.10329	0.10294
$\nu = 0.4$	0.055008	0.054138
$\nu = 0.45$	0.032919	0.03229
$\nu = 0.47$	0.030324	0.029921
$\nu = 0.49$	0.032096	0.031792
$\nu = 0.4999$	0.033775	0.033453
$\nu = 0.49999$	0.033791	0.033467
$\nu = 0.499999$	0.033792	0.033468
$\nu = 0.4999999$	0.033789	0.033463

Table 4: $\|\mathbf{u_I} - \mathbf{u_{II}}\|_{l^2}/\|\mathbf{u_{II}}\|_{l^2}$ for different Poisson number ν, where $\mathbf{u_I}, \mathbf{i} = \mathbf{I}, \mathbf{II}$ is the solution from approach i. The pre-stress advection is included and the problem size is $N = 370883$.

Phase	Pre-processing		Solving	
Approach	I	II	I	II
$N = 6043$	1	0.65332	1.0977	1.1856 (0.51367)
$N = 23603$	3.326	2.6387	4.7178	4.7256 (1.7803)
$N = 93283$	13.018	10.5928	18.055	21.293 (8.9258)
$N = 370883$	50.537	42.6699	72.984	95.3594 (41.0166)
$N = 1479043$	269.06	171.891	317.48	686.564 (213.082)

Table 5: The CPU time spent on a pre-processing step and the solution of the problem for different problem sizes for the two approaches.The figures in the parentheses is the time spent by the iterative solver. The Poisson number $\nu = 0.3$

The behaviour of Approach I and II in the incompressible limit is shown in Table 3. The relative error between solutions $\mathbf{u_I}$ and $\mathbf{u_{II}}$ for all values of $\nu \leq 0.4999999$ is of the order of the precision of the output from `PETSc`, and no signs of locking effects for Approach I can be seen.

In Table 4 the relative error between the solutions from Approach I and II are shown the advection of the pre-stress is included in the model. As can be seen from the table, the different ways to treat this term leads to a relative error of between 3% and 13 %, depending on the value of the Poisson number.

Table 5 show the time spent on the pre-processing phase and the solution phase of the two approaches for different problem sizes N. For Approach II, the pre-processing is the assembly of the stiffness matrix, and for Approach I, it is probably the same. The solution phase for Approach II contains the construction of the preconditioner and the time spent by the iterative solution method. The time spent on the latter is reported within parentheses in the rightmost column of Table 5.

The problem size N in Table 5 refer to the size of the saddle-point problem in Equation (29). One shall bear in mind when comparing the times in the table that the algebraic problem to solve in Approach II is approximately 11% larger than the problem in Approach I because of the choice of the stable pair of mixed finite element spaces.

6 Conclusions

In this paper, we have reported on the current status of an ongoing project aimed at the development of an efficient and accurate code for simulation of the isostatic viscoelastic response of the lithosphere of the Earth to glaciation and deglaciation.

We have reported results from two different approaches to formulate and solve the equations of the purely elastic isostatic response. The approaches differ in the capability to handle the case of a purely incompressible limit, how the advection of pre-stress is treated, and how the algebraic problem arising after a finite element discretization is solved. In the case where we can compare the two methods, that is for not purely incompressible material and without advection of pre-stress, the two approaches give the same result, independently of the problem size, the Poisson number or the regularization parameter ϵ.

With the pre-stress advection added to the problem, the solutions given by Approach I and II differ, but this is what could be expected since this term is treated differently for the two cases.

Also included in this paper is derivation of a boundary condition that guarantees the solvability of the problem formulation in Approach II for the case of a purely incompressible solid and a general advection of the pre-stress. This boundary condition is not, however, necessary to invoke on the problem setting used for the reported numerical experiments because of the special geometry of the problem and the special form of the pre-stress.

7 Acknowledgements

The first author is grateful for the opportunity given to him by Prof. Svetozar Margenov to visit the Institute of Parallel Processing, The Bulgarian Academy of Sciences, within the framework of BIS 21++.

He is also greatly indebted to Prof. Raytcho Lazarov, Department of Mathematics, Texas A&M University, for the fruitful discussion on the weak and strong coercivity of the bilinear form $a(\cdot,\cdot)$.

References

[1] Satish Balay, Kris Buschelman, William D. Gropp, Dinesh Kaushik, Matthew G. Knepley, Lois Curfman McInnes, Barry F. Smith, and Hong Zhang. PETSc Web page, 2001. http://www.mcs.anl.gov/petsc.

[2] W. Bangerth, R. Hartmann, and G. Kanschat. `deal.II` *Differential Equations Analysis Library, Technical Reference*. IWR. `http://www.dealii.org`.

[3] E. Bängtsson and M. Neytcheva. Numerical simulations of glacial rebound using preconditioned iterative solution methods. *Applications of Mathematics*, 50(3), 2005.

[4] E. Bängtsson and M. Neytcheva. An agglomerate multilevel preconditioner for linear isostasy saddle point problems. *Lecture Notes in Computer Science*, 3743:113–120, 2006.

[5] M. Benzi, G.H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Mathematica*, pages 1–137, 2005.

[6] D. Braess. *Finite elements. Theory, fast solvers, and applications in solid mechanics.* Cambridge University Press, second edition, 2001.

[7] F. Brezzi and K-J Bathe. A discourse on the stability conditions for mixed finite element formulations. *Computer Methods in Applied Mechanics and Engineering*, 82:27–57, 1990.

[8] P. Wu. Using commercial finite element packages for the study of earth deformations, sea level and the state of stress. *Geophysical Journal International*, 158:401–408, 2004.