Originally published as:


DOI: http://doi.org/10.1093/gji/ggu378
The forward sensitivity and adjoint-state methods of glacial isostatic adjustment

Zdeněk Martinec,1,2 Ingo Sasgen3,4 and Jakub Velímský2

1 Dublin Institute for Advanced Studies, 5 Merrion Square, Dublin 2, Ireland. E-mail: zdenek@cp.dias.ie
2 Department of Geophysics, Faculty of Mathematics and Physics, Charles University, V Holešovičkách 2, 180 00 Prague 8, Czech Republic
3 Department of Geosciences, The Pennsylvania State University, University Park, PA, USA
4 Department of Geodesy and Remote Sensing, German Research Centre for Geosciences, D-14473, Potsdam, Germany

SUMMARY
In this study, a new method for computing the sensitivity of the glacial isostatic adjustment (GIA) forward solution with respect to the Earth’s mantle viscosity, the so-called the forward sensitivity method (FSM), and a method for computing the gradient of data misfit with respect to viscosity parameters, the so-called adjoint-state method (ASM), are presented. These advanced formal methods complement each other in the inverse modelling of GIA-related observations. When solving this inverse problem, the first step is to calculate the forward sensitivities by the FSM and use them to fix the model parameters that do not affect the forward model solution, as well as identifying and removing redundant parts of the inferred viscosity structure. Once the viscosity model is optimized in view of the forward sensitivities, the minimization of the data misfit with respect to the viscosity parameters can be carried out by a gradient technique which makes use of the ASM. The aim is this paper is to derive the FSM and ASM in the forms that are closely associated with the forward solver of GIA developed by Martinec. Since this method is based on a continuous form of the forward model equations, which are then discretized by spectral and finite elements, we first derive the continuous forms of the FSM and ASM and then discretize them by the spectral and finite elements used in the discretization of the forward model equations. The advantage of this approach is that all three methods (forward, FSM and ASM) have the same matrix of equations and use the same methodology for the implementation of the time evolution of stresses. The only difference between the forward method and the FSM and ASM is that the different numerical differencing schemes for the time evolution of the Maxwell and generalized Maxwell viscous stresses are applied in the respective methods. However, it requires only a little extra computational time for carrying out the FSM and ASM numerically. An straightforward approach to compute the gradient of the data misfit is the brute-force method, whereby the partial derivatives of the misfit with respect to model parameters are approximated by the centred difference of two forward model runs. Although the brute-force method is useful for computing the gradient of the data misfit with respect to a small number of model parameters, it becomes expensive for a viscosity model with a large number of parameters. The ASM offers an efficient alternative for computing the gradient of the misfit since the computational time of the ASM is independent of the number of viscosity parameters. The ASM is thus highly efficient for calculating the gradient of the misfit for models with large numbers of parameters. However, the forward-model solution for each time step must be stored, hence the memory demands scale linearly with the number of time steps. This is the main drawback of the ASM.

Key words: Numerical solutions; Inverse theory; Sea level change; Rheology: mantle.

1 INTRODUCTION
Studies on glacial isostatic adjustment (GIA) provide, among other things, valuable information about the viscosity of the Earth’s interior. Although viscosity depends on the temperature, pressure and chemical composition of the Earth’s interior with many degrees of freedom, we are always forced to parametrize the viscosity by only a few parameters so that the inverse modelling (e.g. Tushingham & Peltier 1992;
et al. (1998) can be carried out with a certain degree of uniqueness. A major difficulty in the choice of model parametrization is to introduce those parameters that are most important for interpreting the data. Strictly, this information cannot be known a priori, but can be inferred from the analysis of the sensitivities, that is partial derivatives of the output field variables of a forward model with respect to the viscosity parameters. Once the sensitivities to all parameters are available, they can be used for ranking the relative importance of the viscosity parameters for a forward-modelled response, for refining an initial viscosity model to improve the fit to the observed data, and for assessing the uncertainty inherent in the inverse-modelled viscosity distribution due to the propagation of errors contaminating the data.

A straightforward approach to compute the sensitivities is the so-called brute-force method, whereby the partial derivatives with respect to model parameters are approximated by the centred difference of two forward model runs. Although the brute-force method is not particularly elegant, it is useful for computing the sensitivities with respect to a small number of model parameters, or for testing the accuracy of faster algorithms (which is the case in this paper). However, it becomes impractical for a viscosity model with a large number of parameters.

The forward sensitivity method (FSM) is an advanced formal alternative technique for calculating the required sensitivities. In the FSM, the forward model is differentiated with respect to model parameters and the resulting forward sensitivity equations are solved for the partial derivatives of the field variables. If there are $M$ model parameters, then the solutions of $M$ forward sensitivity equations are required. Although this is an excellent method when $M$ is small, it becomes computationally expensive for larger values of $M$.

The inverse modelling of GIA-related observations for estimating the viscosity structure can be formulated as a minimization of the misfit between the observed and synthetic data. One approach to find a misfit minimum is to employ a gradient method in which the gradient of the misfit with respect to the model parameters is used to iteratively update the model. The adjoint-state method (ASM) is an advanced formal technique for calculating the gradient of a misfit with respect to the model parameters. In this method, the adjoint-state equations are solved by making use of the nearly identical forward modelling code, but running it backwards in time. The gradient of the misfit to all model parameters are then obtained by a subsequent integration of the product of the forward and adjoint solutions. Thus, there is no need to solve repeatedly the forward model equations, or the forward sensitivity equations.

The first step of the inverse modelling of GIA consists of calculating the forward sensitivities by the FSM and using them to fix the model parameters that do not affect greatly the forward model solution, as well as identifying and removing redundant parts of viscosity structure. Once the viscosity model is optimized in view of the forward sensitivities, the minimization of the misfit can be carried out by a gradient technique which makes use of the ASM. In this sense, the FSM and ASM complement each other. It should be noted that the FSM sensitivities can also be used to calculate the gradient of the misfit with respect to the viscosity parameters. This may, however, be computationally expensive for a large number of viscosity parameters. Vice versa, the ASM only gives the gradient of the misfit, but not the forward sensitivities (e.g. Plessix 2006).

The mathematical foundations of the FSM and ASM for linear and non-linear dynamical systems has been described, for example, by Morse & Feshbach (1953), Lanczos (1961), Marchuk (1995), Cacuci (2003), Sandu et al. (2003) and Tarantola (2005), and its application to the surface loading of a viscoelastic earth model has recently been demonstrated by Al-Attar & Tromp (2014). The latter formulated the adjoint methods for calculating the gradient of the least-squares misfit of the data with respect to the Earth’s viscosity model and the time derivative of a surface load. To make the problem tractable, they did not take into account gravitationally self-consistent ocean loading and rotational feedbacks.

The aim of this paper is to derive the FSM and ASM in the forms that are closely associated with the forward solver of GIA developed by Martinec (2000). Since this method is based on a weak formulation, that is a continuous form of the forward model equations, which are then discretized by spectral and finite elements, we proceed in a similar way for the FSM and ASM. We first derive the continuous forms of the FSM and ASM and then discretize them by spectral and finite elements used for the discretization of the forward method. The advantage of this approach is that all three methods, that is the forward, FSM and ASM, have the same matrix of the system of equations and use the same methodology for the implementation of the time evolution of stresses. The only difference between the forward method and the FSM and ASM involves the different numerical differencing schemes for the time evolution of the Maxwell and generalized Maxwell viscous stresses are used in the respective methods. However, it requires only a little extra computational time to carry out the FSM and ASM numerically.

We formulate the FSM and ASM independently of Al-Attar & Tromp (2014) and extend their results for (i) deriving the strong (differential) and weak (integral) forms of the FSM and ASM; (ii) taking into account gravitationally self-consistent ice and ocean-water loading with a time-varying ocean geometry; (iii) deriving the existence and uniqueness conditions of the FSM and ASM and (iv) developing a numerically convenient time-differencing scheme for calculating the gradient of the viscous stress tensor with respect to viscosity parameters which is applicable to the adjoint-state calculations of the viscosity gradient of data misfit.

The outline of this paper is as follows. In Section 2, we briefly summarize the formulation of the forward method of GIA considering gravitationally self-consistent ice and ocean-water loading with a time-varying ocean geometry. In Section 3, we introduce the parametrization of the Earth’s mantle viscosity and formulate the FSM by the differentiation of the forward model equations with respect to the viscosity parameters. Particular effort is devoted to the differentiation of the sea level equation (Appendix A) and the character of the free-boundary value problem for the FSM in comparison to the forward method. We also give the uniqueness conditions for solving the FSM. In Section 4, we introduce the least-square data misfit and discuss various ways of weighting the data residuals. In Section 5, we summarize the formulation of the ASM, the details of which are given in Appendix B. In Appendix C, we derive the convolution identity, which forms the most important step in formulating the ASM. The solvability conditions of the ASM are discussed in the second part of Section 5. In Appendix D, we derive the weak forms of the FSM and ASM and show that the numerical methods of solving the FSM and ASM can easily be obtained by modifying
the numerical method of solving the forward model equations. In Section 6, we introduce three numerical differencing schemes for the time evolution of the viscous stresses in the forward method. Although Martinec (2000) uses the explicit Euler scheme, the aim of testing of two more schemes is whether similar schemes can be applied for the time evolution of the generalized Maxwell stresses arising in the FSM and ASM. This possibility is studied in detail in Section 7. Both sections make use of the property of the Maxwell viscoelastic rheology, reviewed in Appendix E. In Section 8, we briefly summarize the main results of our study.

2 FORWARD METHOD OF GIA

We are dealing with the viscoelastic response of a self-gravitating, deformable earth model to a surface mass load and intend to reformulate the traditional theory of viscoelastic relaxation (e.g. Farrell 1972; Wu & Peltier 1982) for the case where the deformations and gravity changes are prescribed at the Earth’s surface as additional data to the mass load. We first briefly recall the forward method of GIA in this section.

We consider a viscoelastic earth model \( B \) loaded at \( t = 0 + \) by a time-varying mass load with surface density \( \sigma \) that is placed at the external surface \( \partial B \). Neglecting inertial forces, the viscoelastic response of model \( B \) to the surface mass load is governed by the equation of linear-momentum conservation and by Poisson’s equation for small perturbations of a hydrostatically pre-stressed and self-gravitating continuum in a non-rotating reference frame,

\[
\begin{align*}
\text{div } & \tau - \varrho_0 \text{ grad } \phi_i - \text{div} (\varrho_0 \mathbf{u}) g_0 + \text{grad} (\varrho_0 \mathbf{u} \cdot g_0) = 0, \\
\frac{1}{4\pi G} & \nabla^2 \phi_1 + \text{div} (\varrho_0 \mathbf{u}) = 0,
\end{align*}
\]

where \( \tau (r, t) \) is the Lagrangian increment of the Cauchy stress tensor (abbreviated as the stress tensor hereafter), \( \mathbf{u}(r, t) \) is the displacement, \( \phi_i(r, t) \) is the gravitational potential increment and \( G \) is Newton’s gravitational constant. Eqs (1) are derived by applying the incremental field theory (e.g. Wu & Peltier 1982; Wolf 1991), which allows us to adopt the spherical approximation assumption. This means that (i) the Earth is represented by a sphere \( B \) with the unit normal to the external surface \( \partial B \) and an internal discontinuity \( \Sigma \) coinciding with the unit radial vector \( \mathbf{e}_r \), \( \mathbf{n} = \mathbf{e}_r \), and (ii) the unperturbed mass density \( \varrho_0 \) is radially dependent only, \( \varrho_0 = \varrho_0(r) \), which implies that the unperturbed gravitation \( g_0 \) has only the radial component, \( g_0 = -g_0(r) \mathbf{e}_r \).

The earth model \( B \) is assumed to be constituted of a Maxwell linear viscoelastic material. The constitutive equation for the stress tensor is composed of the elastic and viscous stress tensors, respectively,

\[
\begin{align*}
\tau(t) &= \tau^e(t) + \tau^v(t),
\end{align*}
\]

where

\[
\begin{align*}
\tau^e(t) &= m(0) : \mathbf{e}(t), \\
\tau^v(t) &= \int_0^t M(t') : \mathbf{e}(t - t') dt'.
\end{align*}
\]

Here, \( \mathbf{e} \) is the strain tensor, that is the symmetric part of grad \( \mathbf{u} \), that is \( \mathbf{e} = (\text{grad} \, \mathbf{u} + \text{grad} \, \mathbf{u}^T) / 2 \), \( m(t) \) and \( M(t) \) are the viscoelastic relaxation tensor and the specific relaxation tensor, respectively (see eqs. E11 and E13 in Appendix E, which describe the elastic and viscous constitutive properties of model \( B \)). The symbol : denotes the double-dot product of second-order tensors.

Depending on the time evolution of the surface-mass load, the boundary conditions on the external surface \( \partial B \) vary over time and are of the form (Longman 1963; Farrell 1972)

\[
\begin{align*}
\tau^- \cdot \mathbf{e}_r &= -g_0 \sigma \mathbf{e}_r, \\
[\phi_1]^+ &= 0, \\
\frac{1}{4\pi G} & \text{grad } \phi_1^+ \cdot \mathbf{e}_r - \varrho_0 (\mathbf{u}^- \cdot \mathbf{e}_r) = \sigma,
\end{align*}
\]

where \( \tau^- \), \( \varrho_0^- \) and \( \mathbf{u}^- \) denote the stress tensor, the unperturbed density and the displacement on the interior side of \( \partial B \), respectively, \( a \) is the radius of \( \partial B \), \( g_0 = g_0(a) \), and the symbol \( [f]^+ \) denotes the jump of the quantity \( f \) across \( \partial B \). Eqs (1) are yet subject to the homogeneous initial condition in \( B \),

\[
\begin{align*}
\mathbf{u}(r, 0) &= 0, \\
\phi_1(r, 0) &= 0.
\end{align*}
\]

The surface load \( \sigma \) consists of a combined ice and ocean-water load,

\[
\sigma = \begin{cases} 
\sigma^\text{ice} & \text{over continents}, \\
\varrho^g s & \text{over oceans},
\end{cases}
\]
where $\sigma^{\text{se}} = \sigma^{\text{se}}(\Omega, t)$ is a prescribed spatial and time-varying continental ice load, $\Omega = (\vartheta, \varphi)$ denotes a pair of angular spherical coordinates, that is spherical colatitude $\vartheta$ and longitude $\varphi$, $\rho^0$ is the ocean-water density and $s = s(\Omega, t)$ is the change of relative sea level height over the ocean domain $\mathcal{O}(\Omega, t)$ with respect to time $t = 0$. The quantity $s$ is determined by the sea level equation (Farrell & Clark 1976), which yields the gravitationally consistent evolution of sea level following the viscoelastic response of the solid Earth and the redistribution of water masses according to continental ice mass loss and coastline migration.

The ocean load $\varrho^0$'s, see eq. (A1) in Appendix A, and the ocean function $\mathcal{O}$, see eq. (A15), depend on the field variables searched for by solving the forward model equations, which means that the forward method of GIA is represented by a free boundary-value problem. More explicitly, in addition to the displacement $u(r, t)$ and the gravitational potential increment $\phi_i(r, t)$, the domain $\mathcal{O}(\Omega, t)$ occupied by ocean at time $t$ must also be determined. Besides the equation of linear-momentum conservation, Poisson’s equation and the initial conditions, the problem description also requires the initial configuration $\mathcal{O}^0(\Omega)$ of the ocean domain and the initial topography $\zeta^0(\Omega)$, as well as the conditions which hold along the free boundary.

We now consider the existence of a solution to the initial boundary-value problem (1)–(6). In the simplified case of a non self-gravitating sphere, that is, when $\text{div} \, \tau$ is the only term considered in the linear-momentum equation, and Poisson’s equation and two boundary conditions for $\phi_i$ are not considered in the formulation of the forward problem, the existence of a forward solution is only guaranteed if the load $\sigma$ satisfies the condition (Rektorys 1980; Nečas & Hlaváček 1981)

$$\int_{\mathcal{S}} \mathbf{r} \, \sigma(r, t) d\mathcal{S} = 0. \quad (7)$$

This condition means that the load of spherical harmonic degree 1 must be removed from the surface loading. Cathles (1971) has, however, observed that the consequence of self-gravitation is that, at a fixed time, a solution to the viscoelastic self-gravitating problem (1)–(6) exists. This means that the existence condition (7) is not required to be satisfied by surface loading. Consequently, the combined ice and water-ocean load that generally has a degree-one component and will cause a shift of the centre of mass of the Earth, does not need to be modified by removing the degree-one component since the conservation of linear momentum is not violated.

Moreover, Cathles (1971) also observed that the centre of mass of the Earth plus the load is fixed in space and there is no restriction on its position. This means that the uniqueness of the solution to the problem (1)–(6) is not guaranteed. In fact, at a fixed time, the problem (1)–(6) with the homogeneous boundary condition $\sigma = 0$ has a trivial solution for the strain tensor $\varepsilon$, that is $\varepsilon = 0$ in $\mathcal{B}$. However, this solution does not imply that the displacement $u$ also vanishes in $\mathcal{B}$ because of the Neumann-type of boundary condition (4). To ensure that the displacement field is uniquely determined by the solution $\varepsilon = 0$, we impose two additional constraints on a solution (Rektorys 1980; Martinec & Hagedoorn 2005):

$$\int_{\mathcal{B}} \varrho^0 \mathbf{u}(r, t) dV + \int_{\partial \mathcal{B}} \mathbf{r} \, \sigma(r, t) d\mathcal{S} = 0 \quad \text{and} \quad \int_{\mathcal{B}} \varrho^0 [\mathbf{r} \times \mathbf{u}(r, t)] dV = 0, \quad (8)$$

where $\times$ denotes the cross product of vectors. In summary, the solution of the original boundary-value problem for a fixed time unconditionally exists, but it is determined uniquely up to a rigid-body translation and rotation. The practical implication of this fact is that, if the linear momentum balance (1), is represented in terms of spherical harmonics, three equations for harmonic degree one are linearly dependent and two of them must be replaced by eqs (8). The spherical-harmonic representation of (8) is given by Martinec & Hagedoorn (2005, section 5.2).

### 3 FSM of GIA

The first step of the GIA inverse modelling consists of calculating the forward sensitivities, that is the partial derivatives of the forward solution with respect to the model parameters. As discussed in the introduction, they are then used to fix the model parameters which do not significantly affect the forward model solution, and to identify and remove overparametrized parts of the viscosity structure.

Let us consider the radially and laterally inhomogeneous shear viscosity $\eta$ of the viscoelastic sphere $\mathcal{B}$ to be represented in terms of an $M$-dimensional system of $r$-dependent base functions, and let the expansion coefficients of this representation be $\eta_1$, $\eta_2$, ..., $\eta_M$. Defining the parameter vector $\hat{\eta} := (\eta_1, \eta_2, \ldots, \eta_M)$, the dependence of the viscosity $\eta$ on the viscosity parameters $\hat{\eta}$ can be made explicit as

$$\eta = \eta(r, \hat{\eta}). \quad (9)$$

As introduced, the forward sensitivity analysis computes the sensitivities of the forward solution with respect to the viscosity parameters (often termed the forward sensitivities), that is the partial derivatives $\partial \mathbf{u} / \partial \eta_m$ and $\partial \phi_i / \partial \eta_m$, where $m = 1, \ldots, M$. To abbreviate the following notation, the partial derivatives with respect to the viscosity parameters are ordered in the gradient operator in the $M$-dimensional parameter space,

$$\nabla_{\hat{\eta}} := \left( \frac{\partial}{\partial \eta_1}, \frac{\partial}{\partial \eta_2}, \ldots, \frac{\partial}{\partial \eta_M} \right). \quad (10)$$

To form the forward sensitivity equations, also called the linear tangent equations of the model (e.g. McGillivray et al. 1994; Cacuci 2003; Sandu et al. 2003), let us consider the viscosity model (9) in the forward model equations and differentiate them with respect to the viscosity parameters $\hat{\eta}$. We begin with the differentiation of the constitutive equations (E4) and (E5),

$$\nabla_{\hat{\eta}} \mathbf{r}(t) = \mathbf{T}(t) + \mathbf{T}_{\hat{\eta}}(t), \quad (11)$$
where the tensors $T$ and $\bar{T}$ stand for
\[
T(t) := m(0) \cdot E(t) + \int_{0}^{t} M(t') \cdot E(t - t') \, dt',
\]
\[
\bar{T}(t) := \int_{0}^{t} \nabla_{\bar{\eta}} M(t') \cdot \varepsilon(t - t') \, dt',
\] (12)
and $E$ denotes the gradient of the strain tensor $\varepsilon$ with respect to the viscosity parameters $\bar{\eta}$,
\[
E := \nabla_{\bar{\eta}} \varepsilon.
\] (13)
Differentiating the linear-momentum equation and Poissons equation in eq. (1) with respect to the viscosity parameters $\bar{\eta}$ and substituting eq. (11) in the result yields
\[
\text{div} \ T - \frac{1}{G} \text{grad} \Phi_1 - \text{div} (\bar{\eta} U) = 0,
\] (14)
The new field variables, the so-called forward sensitivities, that are searched for by solving these equations inside the viscoelastic sphere $B$ are
\[
U := \nabla_{\bar{\eta}} u,
\]
\[
\Phi_1 := \nabla_{\bar{\eta}} \phi_1.
\] (15)
The differential eqs (14) need to be supplemented by the boundary conditions on the external surface $\partial B$. These result from the differentiation of the boundary conditions in eq. (4) with respect to the viscosity parameters $\bar{\eta}$,
\[
T \cdot e_r = -T_{\bar{\eta}} \cdot e_r - g_0 \nabla_{\bar{\eta}} \sigma \cdot e_r.
\]
[1 \cdot \Phi_1]_{b_{\bar{\eta}}} = 0,
\]
\[
\frac{1}{4\pi G} \text{grad} \Phi_1 \cdot e_r - \frac{\bar{\eta}}{g_0} (U_r \cdot e_r) = \nabla_{\bar{\eta}} \sigma.
\] (16)
Eqs (14) are yet subject to the homogeneous initial condition in $B$,
\[
U(r, 0) = 0,
\]
\[
\Phi_1(r, 0) = 0
\] (17)
It remains now to determine the gradient of the surface load $\sigma$ with respect to the viscosity parameters $\bar{\eta}$, that is the quantity $\nabla_{\bar{\eta}} \sigma$. The application of the gradient operator $\nabla_{\bar{\eta}}$ to eq. (6) yields
\[
\nabla_{\bar{\eta}} \sigma = \begin{cases} 
0 & \text{over continents,} \\
\frac{g}{G} \nabla_{\bar{\eta}} \sigma & \text{over oceans.}
\end{cases}
\] (18)
Note that $\nabla_{\bar{\eta}} \sigma = 0$ over the continents since we assume that the continental ice load is input information given independently of the viscosity structure of the Earth. The derivation of the viscosity gradient of the sea level equation is presented in Appendix A, resulting in
\[
\nabla_{\bar{\eta}} \sigma_\Omega(t) = \left[ E(\Omega, t) - U(\Omega, t) - (E - U)(t) \right] \Omega(t),
\] (19)
where $U = U \cdot e_r$ and $E = -\Phi_1/g_0$ (both $U$ and $\Phi_1$ are considered at the Earth’s surface) and $E - U$ is the ocean average value of the difference $E - U$, see eq. (A9).

The differentiation of the forward model equations with respect to the viscosity parameters $\bar{\eta}$ does not change the solvability conditions of the boundary-value problem. Hence, the existence of a solution of the forward sensitivity equations is unconditionally guaranteed. The uniqueness conditions (8) are then transformed to the form
\[
\int_{B} \bar{\eta} U(r, t) \, dV + \int_{\partial B} r \nabla_{\bar{\eta}} \sigma(r, t) \, dS = 0 \quad \text{and} \quad \int_{B} \bar{\eta} [r \times U(r, t)] \, dV = 0.
\] (20)
In the forward sensitivity analysis, the forward model equations are solved for each parameter $\eta_m$. The forward solutions $u(\eta_m)$ and $\phi_1(\eta_m)$ and the associated viscous stress tensor $\tau_{\bar{\eta}}(\eta_m)$ are used to compute a new source term $T_{\bar{\eta}}$ according to theory given in the following Section 6. Consequently, the forward sensitivity equations (14)–(19) are solved for the partial derivative $\partial u/\partial \eta_m$ and $\partial \phi_1/\partial \eta_m$.

An important feature of the forward sensitivity equations is that the load $E - U - \overline{E - U}$ in eq. (19) is applied over the ocean area defined by the ocean function $\Omega$, which in turn is determined by GIA forward modelling. Hence, when solving the forward sensitivity equations, there is no need to search for a new ocean function, but, at each time step, the shape of ocean area is defined by the GIA forward
modelled ocean function. On top of that, after computing the deformations fields $U$ and $E$, the sensitivity of the ocean function $O$ with respect to the viscosity parameters $\tilde{\eta}$, that is the viscosity gradient $\nabla_\tilde{\eta} O$, can be calculated by eq. (A20),

$$
\nabla_\tilde{\eta} O(\tilde{\eta}, t) = \left[ E(\tilde{\eta}, t) - U(\tilde{\eta}, t) - (E - U)(t) \right] \delta(C),
$$

(21)

where $\delta(C)$ is the Dirac delta function with the support on coastline $C$, see eq. (A14). This relation determines the sensitivity of the coastal regions to the viscosity structure and can be used for ranking the relative importance of the sea level indicator data at particular locations along the coast.

4 MISFIT FUNCTION AND ITS GRADIENT

The GIA forward problem formulated in Section 2 enables the modelling of the time evolution of the vertical and horizontal components of displacement $\mathbf{u}(\tilde{\eta})$ and gravitational potential increment $\phi_i(\tilde{\eta})$. These predicted data evaluated at the Earth’s surface can be compared with the observations $\mathbf{u}^{\text{(obs)}}$ and $\phi_i^{\text{(obs)}}$. The $\mathbf{u}^{\text{(obs)}}$ data can be obtained by interpreting GPS and tide-gauge observations (Davis & Mitrovica 1996; Braun et al. 2008), terrestrial gravity, satellite radar altimetry (Lee et al. 2008; Sæseng 2008; Steffen & Wu 2011). The gravitational potential signal $\phi_i^{\text{(obs)}}$ related to GIA has been identified and evaluated in the trends of the GRACE gravity fields over many previously, and currently glaciated regions including North America, (e.g. Tamisiea et al. 2007; Sæseng et al. 2012), Antarctica (e.g. Sæseng et al. 2007; Riva et al. 2009; Sæseng et al. 2013) and Fennoscandia, (e.g. Steffen et al. 2008; Steffen & Wu 2011).

To interpret observations $\mathbf{u}^{\text{(obs)}}$ and $\phi_i^{\text{(obs)}}$, let us assume that the forward sensitivity analysis for $\mathbf{u}(\tilde{\eta})$ and $\phi_i(\tilde{\eta})$ has been carried out (see Section 3) and the viscosity parameters that have minimal effect on the surface values of $\mathbf{u}(\tilde{\eta})$ and $\phi_i(\tilde{\eta})$ have been fixed and the viscosity parametrization (9) is optimally designed in view of the forward sensitivities. The differences between the observations $\mathbf{u}^{\text{(obs)}}$ and $\phi_i^{\text{(obs)}}$ made on the external surface $\partial B$ for times $t \in (0, T)$ and the predicted data $\mathbf{u}(\tilde{\eta})$ and $\phi_i(\tilde{\eta})$ can be used as a misfit for the GIA inverse modelling. The least-squares misfit is then defined as

$$
\chi^2(\tilde{\eta}) := \frac{1}{2} \int_0^T \int_{\partial B} \left\{ w_u^2 \left[ \mathbf{u}(\tilde{\eta}) - \mathbf{u}^{\text{(obs)}} \right]^2 + w_\phi^2 \left[ \phi_i(\tilde{\eta}) - \phi_i^{\text{(obs)}} \right]^2 \right\} \, dS \, dt,
$$

(22)

where the weighting factors $w_u = w_u(\tilde{\eta}, t)$ and $w_\phi = w_\phi(\tilde{\eta}, t)$ are chosen to be dimensionless such that the expression in the curly brackets has the unit of length squared. To ensure this, the second term is scaled by the square of surface gravity. Scaling by the factor $\varrho \Delta_1 \phi$ guarantees that the boundary-value data for the adjoint stress tensor and adjoint gravitational-potential increment, see eq. (B23) in Appendix B, have the same physical dimensions as those for the forward method, that is $[\varrho_s] = \text{kg} \, \text{m}^{-2}$ and $[\Delta_1 \phi] = \text{m} \, \text{s}^{-2}$, respectively. If the observations $\mathbf{u}^{\text{(obs)}}$ and $\phi_i^{\text{(obs)}}$ contain random errors which are statistically independent, the statistical variance of the observations may be substituted for the reciprocal value of $w_u^2$ and $w_\phi^2$ (e.g. Bevington 1969). The dependence of $w_u$ and $w_\phi$ on the angular coordinates $\tilde{\eta}$ allows the elimination of data from regions which are contaminated by other signals, while the dependence of $w_u$ and $w_\phi$ on time allows the elimination of the data for time epochs when data are missing, or, for example, account for the larger uncertainties associated with the geological evidence of surface-deformation and sea level change the further in the past one looks. The weighting factors can also account for the spatial and temporal varying sensitivities of $\mathbf{u}(\tilde{\eta})$ and $\phi_i(\tilde{\eta})$ over $\partial B$ determined by the forward sensitivity analysis (see Section 3). This may, for instance, be applied in the case where near-field and far-field surface data (i.e. data from around glaciated areas and from more distant locations, respectively) are combined in the misfit (22).

The GIA inverse modelling of a viscosity structure minimizes the misfit $\chi^2$ with respect to the viscosity parameters $\tilde{\eta}$ (e.g. Martinec & Wolf 2005). One approach to find the minimum of the misfit is to employ a gradient method, in which the gradient of $\chi^2$ with respect to the parameters $\tilde{\eta}$ is used to iteratively update the viscosity model such that the misfit is iteratively minimized.

Applying the gradient operator (10) to both sides of eq. (22) and realizing that the observations $\mathbf{u}^{\text{(obs)}}$ and $\phi_i^{\text{(obs)}}$ are independent of the viscosity parameters $\tilde{\eta}$, that is $\nabla_\tilde{\eta} \mathbf{u}^{\text{(obs)}} = \nabla_\tilde{\eta} \phi_1^{\text{(obs)}} = 0$, the gradient of $\chi^2(\tilde{\eta})$ is

$$
\nabla_{\tilde{\eta}} \chi^2 = \int_0^T \int_{\partial B} \left( \Delta \mathbf{u}(\tilde{\eta}) \cdot \nabla_\tilde{\eta} \mathbf{u} + \Delta \phi_i(\tilde{\eta}) \nabla_\tilde{\eta} \phi_i \right) \, dS \, dt,
$$

(23)

where $\Delta \mathbf{u}(\tilde{\eta})$ and $\Delta \phi_i(\tilde{\eta})$ are the weighted residuals of the surface displacement vector and gravitational potential increment, respectively,

$$
\Delta \mathbf{u}(\tilde{\eta}) := \varrho_0 g_0 w_u^2 \left[ \mathbf{u}(\tilde{\eta}) - \mathbf{u}^{\text{(obs)}} \right],
$$

$$
\Delta \phi_i(\tilde{\eta}) := \frac{\varrho_0}{g_0} w_\phi^2 \phi_i(\tilde{\eta}) - \phi_i^{\text{(obs)}}.
$$

(24)

The straightforward approach to finding $\nabla_{\tilde{\eta}} \chi^2$ is to approximate $\partial \chi^2/\partial \eta_m$ by a numerical differentiation of forward model runs, the so-called brute-force method (e.g. Bevington 1969). Alternatively, after completing the forward sensitivity analysis (see Section 3) for $m = 1, \ldots, M$, the forward sensitivities $\nabla_\tilde{\eta} \mathbf{u}$ and $\nabla_\tilde{\eta} \phi_i$ can be substituted into eq. (23) and $\nabla_{\tilde{\eta}} \chi^2$ is evaluated for the residuals $\Delta \mathbf{u}$ and $\Delta \phi_i$. However, it should be emphasized that the primary purpose of the FSM is to identify the redundant parts of viscosity parametrization, not to compute $\nabla_{\tilde{\eta}} \chi^2$. 


5 ADJOINT-STATE METHOD OF GIA

The ASM provides an efficient alternative to the brute-force method and the FSM for evaluating $\nabla u$ without explicit knowledge of $\nabla u$ and $\nabla \phi$. Hence, the ASM is efficient for problems involving a large number of model parameters.

Al-Attar & Tromp (2014) derive the weak form of the ASM for the viscoelastic loading of a sphere. However, they do not provide the strong differential form of the ASM, do not specify the existence and uniqueness conditions of its solvability and do not take into account gravitationally self-consistent ice and ocean-water loading with a time-varying ocean geometry. We will here focus on solving these items. Hence, the ASM presented in this section is directly applicable to GIA inverse modelling.

A detailed derivation of the differential form of the ASM is presented in Appendix B which includes the gravitationally consistent implementation of a combined ice and ocean-water load with a time-varying coastline geometry. To summarize the results of Appendix B, we will now discuss the solvability conditions of the ASM since the derivation of the ASM in Appendix B does not provide them. In contrast to the forward method, a solution to the adjoint-state problem (25)–(29) does not, in general, exist. In other words, it may happen that the adjoint ocean load $\hat{u}(r, t)$ and the adjoint gravitational potential increment $\hat{\phi}(r, t)$ in $B$ by solving the adjoint-state problem:

$$\text{div}\ 1 \frac{\hat{\phi}}{4\pi G} + \text{div}(\hat{0}_0 \hat{u} + \text{grad}(\hat{0}_0 \hat{u})) = 0,$$

with the boundary condition on $\partial B$,

$$e_0 \cdot \hat{\tau} - e_0 = -g_0 \hat{\sigma} - \hat{\sigma}_1,$$

$$e_0 \cdot \hat{\tau} - (I - e_0 \otimes e_0) = -\hat{\sigma}_2,$$

and the homogeneous initial conditions in $B$,

$$\hat{u}(r, 0) = 0,$$

$$\hat{\phi}(r, 0) = 0.$$

The adjoint ocean load $\hat{\sigma}$ is given by

$$\hat{\sigma}(t) = e^t \left[ \hat{\tau}(t) - \hat{u}(t) - \hat{u} \right] \otimes \hat{e}(t),$$

while the loads $\hat{\sigma}_1$, $\hat{\sigma}_2$ and $\hat{\sigma}_3$ are the weighted differences between the modelled and observed quantities,

$$\hat{\sigma}_1(t) := \Delta u(T - t) \cdot e_0,$$

$$\hat{\sigma}_2(t) := \Delta u(T - t) \cdot (I - e_0 \otimes e_0),$$

$$\hat{\sigma}_3(t) := \Delta \phi(T - t),$$

where $I$ is the second-order identity tensor and the symbol $\otimes$ denotes the dyadic product of vectors. Comparing the formulations of the forward method and the ASM, we can see that both methods are governed by the same differential equations applied to the sphere $B$ with the same initial conditions in $B$. However, they differ in terms of the loads applied to the boundary $\partial B$, dependent upon which observations are taken into account for solving the ASM. Hence, the solvability conditions for the adjoint-state equations differ, in general, from those for the forward method.

We will now discuss the solvability conditions of the ASM since the derivation of the ASM in Appendix B does not provide them. In contrast to the forward method, a solution to the adjoint-state problem (25)–(29) does not, in general, exist. In other words, it may happen that the adjoint surface loads $\hat{\sigma}_1$, $\hat{\sigma}_2$ and $\hat{\sigma}_3$ are not in the image of the adjoint operator. To guarantee the existence of a solution, the adjoint surface loads must satisfy the existence condition (e.g. Nečas & Hlaváček 1981). Due to the same reasoning as for the forward method of GIA (Cathles 1971), the existence of an adjoint solution is guaranteed if $\hat{\sigma}_1$ and $\hat{\sigma}_3$ satisfy the condition

$$\int_{\partial B} r \hat{\sigma}_1(r, t) dS = \int_{\partial B} r \hat{\sigma}_3(r, t) dS,$$

and the surface traction $\hat{\sigma}_2$ satisfies the condition (Rektorys 1980; Nečas & Hlaváček 1981)

$$\int_{\partial B} r \times \hat{\sigma}_2(r, t) dS = 0.$$
The first condition means that the surface loads \( \hat{\sigma}_1 \) and \( \hat{\sigma}_3 \) cannot be chosen arbitrary, but are such that their surface net translations are equal to each other. The second condition means that the surface horizontal traction \( \hat{\sigma}_2 \) must be chosen such that its surface net torque vanishes. Note that the adjoint ocean load \( \hat{\sigma} \) must also satisfy conditions (30) and (31). However, \( \hat{\sigma} \) occurs in the boundary conditions (26); and (26); ‘symmetrically’ and, in addition, \( \hat{\sigma} \) has no component in the boundary condition (26);. Therefore, \( \hat{\sigma} \) satisfies both the existence conditions identically, and there is no need to remove the load of angular degree 1 from \( \hat{\sigma} \). This corresponds to the case of combined ice and ocean-water load in the forward method.

As a matter of fact, the observations of vertical displacement or its time rate associated with GIA phenomenon are, in general, independent of the observations of gravitational potential changes. Hence, the condition (30) can only be satisfied if

\[
\int_{\partial B} r \hat{\sigma}_1(r,t) dS = \int_{\partial B} r \hat{\sigma}_3(r,t) dS = 0. \tag{32}
\]

Representing the adjoint loads \( \hat{\sigma}_1, \hat{\sigma}_3 \) and \( \hat{\sigma}_2 \) in terms of scalar and vector spherical harmonics, respectively, the existence conditions can be expressed as (Martinec & Hagedoorn 2005; Martinec 2009)

\[
[\hat{\sigma}_1(t)]_{\text{m}} = [\hat{\sigma}_3(t)]_{\text{m}} = 0 \quad \text{and} \quad [\hat{\sigma}_2(t)]_{\text{m}} = 0 \quad m = 0, \pm 1, \tag{33}
\]

where the index \( m \) stands for the order of spherical harmonics, \([\hat{\sigma}_1]_{\text{m}} \) and \([\hat{\sigma}_3]_{\text{m}} \) are first-degree vertical-spheroidal vector spherical harmonics of loads \( r\hat{\sigma}_1 \) and \( r\hat{\sigma}_3 \), respectively, and \([\hat{\sigma}_2]_{\text{m}} \) are the first-degree toroidal vector spherical harmonics of load \( \hat{\sigma}_2 \). The existence of an adjoint solution is thus guaranteed if the following nine scalar conditions are satisfied. (i) The first-degree scalar spherical harmonics are removed from the residuals of vertical surface displacement \( \Delta u \cdot e_n \); (ii) the first-degree toroidal vector spherical harmonics are removed from the residuals of surface displacement \( \Delta u \) and (iii) the first-degree scalar spherical harmonics are removed from the residuals of the incremental gravitational potential \( \Delta \Phi \). This means that all of these spherical harmonics must be excluded from definition (22) of the least-squares misfit \( \chi^2(\hat{n}) \). Note that there is no existence condition imposed on the first-degree horizontal-spheroidal vector spherical harmonics of \( \Delta u \).

Having assumed the existence conditions (31) and (32) are satisfied, the adjoint displacement \( \hat{u} \) resulting from solving the adjoint-state initial- and boundary-value problem is determined uniquely up to a rigid-body translation and rotation. To guarantee the uniqueness of a solution, we impose two additional constraints on an adjoint solution. Similarly to the forward method, we require that

\[
\int_{\partial B} \hat{\Phi}_{\text{u}} \hat{u}(r,t) dV + \int_{\partial B} r \hat{\sigma}(r,t) dS = 0 \quad \text{and} \quad \int_{\partial B} \hat{\Phi}_{\text{u}}[r \times \hat{u}(r,t)] dV = 0. \tag{34}
\]

where \( \hat{\sigma} \) is the adjoint ocean load given by eq. (28).

Having solved the adjoint-state equations, the gradient of the misfit \( \chi^2 \) with respect to the viscosity parameters is expressed as

\[
\nabla_{\hat{n}} \chi^2 = \int_0^T \int_0^t T_{\hat{n}} : \hat{\epsilon} \, dV \, dt
\]

\[
= \int_{t_0}^T \int_{t_0}^t \hat{\epsilon}(T - t') : \nabla_{\hat{n}} M(t') : \mathbf{e}(t' - t) \, dV \, dt' \, dt.
\tag{35}
\]

The importance of this equation is that, once the forward problem (1)–(6) is solved and the misfit \( \chi^2 \) is evaluated from eq. (22), the gradient \( \nabla_{\hat{n}} \chi^2 \) may be evaluated for little more than the cost of a single solution of the adjoint-state equations (25)–(34) and two double dot products in eq. (35), regardless of the dimension of the viscosity vector \( \hat{n} \). This is compared to the brute-force method for calculating \( \nabla_{\hat{n}} \chi^2 \) that requires solving the forward model equations (1)–(6) twice per component of \( \hat{n} \).

We now explain the specific steps involved in the adjoint-state computations. First, the forward solutions \( \{u(t), \phi(t), \Pi(t), \mathcal{O}(t)\} \) are calculated at discrete time levels \( 0 = t_0 < t_1 < \cdots < t_n = T \) by solving the forward problem (1)–(6), and each solution \( \{u(t_i), \phi(t_i), \Pi(t_i), \mathcal{O}(t_i)\} \) must be stored. Then, the adjoint solutions \( \{\hat{u}(t_i), \phi(t_i), \hat{\Pi}(t_i)\}, i = 0, \ldots, n \), are calculated, proceeding again forward in time, starting with homogeneous values. In this step, the adjoint ocean load is calculated as

\[
\hat{\sigma}(\Omega, t_{i+1}) := \mathcal{O}^T \left[ \hat{\sigma}(\Omega, t_i) - \hat{u}(\Omega, t_i) - \mathbf{\hat{e}}(t_i) \right] \mathcal{O}(\Omega, T - t_{i+1}). \tag{36}
\]

What is important is that the adjoint-state calculations use the ocean function \( \mathcal{O}(t) \) which is determined by solving the forward method. Hence, there is no need to search for a new ocean function \( \hat{\mathcal{O}}(t) \) by solving the adjoint-state equations, but the shape of ocean area at each time step is determined by the forward-modelled ocean function, that is \( \hat{\mathcal{O}}(t) = \mathcal{O}(T - t) \). In other words, the ocean function for the adjoint-state equations is prescribed \textit{a priori} and the adjoint-state problem is solved for free-boundary function values that are, however, prescribed over an ocean area with a fixed geometry. As each adjoint solution \( \{\hat{u}(t_i), \phi(t_i), \hat{\Pi}(t_i)\} \) is computed, the misfit and its derivative are updated according to eq. (22) and (35), respectively. When \( \{\hat{u}(T), \phi(T), \hat{\Pi}(T)\} \) have finally been calculated, both \( \chi^2 \) and \( \nabla_{\hat{n}} \chi^2 \) are known. The forward solutions \( \{u(t_i), \phi(t_i), \Pi(t_i), \mathcal{O}(t_i)\} \) are stored because eqs (29), (35) and (36) depend on them for the adjoint-state calculations. As a result, the numerical algorithm has memory requirements that are linear to the number of time steps. This is the main drawback of the ASM.
6 TIME-DIFFERENCING SCHEME FOR VISCOS STRESS TENSOR

Comparing the formulations of the forward method, the FSM and ASM, we can see that all three are described by the coupled set of two second-order differential equations with the same differential operators. The methods differ by the applied volume forcing or surface loading. We can also see that the methodology for the implementation of the viscous stress tensor \( \tau^V \) in the forward method is the same as the implementation of the viscous gradient of the viscous stress tensor, that is the term \( T_\eta \), in the FSM and ASM. The implementations, however, differ in the numerical schemes for an optimal time evolution of the respective tensor. Although Martinec (2000) uses the explicit differencing scheme for the time evolution of the viscous stress tensor \( \tau^V \), which performs numerically well, in this section we propose two alternative schemes that are more computationally efficient than the explicit differencing scheme for computing \( T_\eta \).

The rheological model of the Earth used for the modelling of GIA is introduced in Appendix E. To find an approach to compute the viscous stress tensor \( \tau^V (t) \) numerically, let us first express the integral in eq. (E15) in an alternative form. Substituting \( \tau = t - t' \) for \( t' \) in eq. (E15) yields

\[
\tau(t') = -\frac{2}{\tau_{\text{relax}}} \int_0^{t'} m_2(t - t') \varepsilon(t') dt',
\]

where \( t' \) has again been used for the variable of integration. The differentiation of the last expression with respect to time gives

\[
\frac{d\tau(t')}{dt} = -\frac{2}{\tau_{\text{relax}}} \left[ m_2(0) \varepsilon(t) + \int_0^{t'} \frac{dm_2(t - t')}{dt} \varepsilon(t') dt' \right].
\]

To express the integral on the right-hand side, we differentiate eq. (E9) with respect to time,

\[
\frac{dm_2(t - t')}{dt} = -\frac{m_2(t - t')}{\tau_{\text{relax}}},
\]

and consider \( m_2(0) = \mu \). Eq. (38) then becomes

\[
\frac{d\tau^V(t)}{dt} + \frac{\tau^V(t)}{\tau_{\text{relax}}} = -\frac{\sigma^E(t)}{\tau_{\text{relax}}},
\]

where the elastic shear stress tensor \( \sigma^E \) has been introduced to abbreviate the notation,

\[
\sigma^E(t) := 2\mu \varepsilon(t).
\]

Eq. (40) can be viewed as the first-order ordinary differential equation for the time evolution of the viscous stress tensor \( \tau^V (t) \), provided that the function \( \sigma^E(t) \) and the initial condition \( \tau^V (0) \) are given. Eq. (37) implies that \( \tau^V (0) = 0 \) since the integrand in eq. (37) is bounded for \( t \to 0 \). Various modifications of the Runge–Kutta method can be applied to find an approximate solution of eq. (40). Here, we will apply the explicit and implicit Euler methods, which are the simplest versions of the Runge–Kutta method.

We choose a value \( \Delta t \) for the time step and set discrete time levels \( t^{i+1} = t^i + \Delta t \), \( i = 0, 1, \ldots \), with \( t^0 = 0 \). The constitutive equation (E4) for the viscoelastic stress tensor \( \tau \), when referred to the current time level \( t^{i+1} \), is given by

\[
\tau^{i+1} = \tau^{E,i+1} + \tau^{V,i+1},
\]

where the elastic stress tensor at the current time level, \( \tau^{E,i+1} \), is expressed in terms of the strain tensor at the current time level, \( \varepsilon^{i+1} \), according to eq. (E14) considered at \( t = t^{i+1} \). As introduced, the viscous stress tensor \( \tau^V \) is given by a numerical solution of eq. (40). Let a solution of this equation at time \( t' \) be \( \tau^{V,i} \), then in the subsequent (current) time level \( t^{i+1} \), the explicit Euler method applied to eq. (40) yields

\[
\tau^{V,i+1} = (1 - \alpha) \tau^{V,i} - \alpha \sigma^{E,i},
\]

whereas the implicit Euler method provides

\[
\tau^{V,i+1} = \frac{1}{1 + \alpha} \tau^{V,i} - \frac{\alpha}{1 + \alpha} \sigma^{E,i+1},
\]

where the dimensionless parameter \( \alpha \) is defined as

\[
\alpha := \frac{\Delta t}{\tau_{\text{relax}}}.\]

There is now the important question of which time-differencing scheme is most appropriate for the numerical solution of the forward model equations formulated in the preceding sections. The numerical solution for \( \tau^{V,i+1} \) by the implicit Euler method requires knowledge of the elastic shear stress tensor \( \sigma^E \) at the current time level \( t^{i+1} \). Consequently, the term \( \text{div} \ \tau \) in the linear momentum equation (1), which is referred to the current time \( t^{i+1} \), will consist of two current-time contributions, the elastic stress tensor \( \tau^{E,i+1} \) and the contribution \( -\frac{1}{\tau_{\text{relax}}} \sigma^{E,i+1} \) due to the implicit Euler scheme for \( \tau^{V,i+1} \) in eq. (44). This is the consequence of the implicit character of the implicit Euler method, meaning

\[\footnote{Note that the trace of tensor \( \sigma^E \) vanishes only for an incompressible material, which is considered here.}\]
that the new value of $\tau_{E,i+1}$ can only be found by solving the linear momentum equation for the sum of the elastic stress tensor $\tau_{E,i+1}$ and its shear part $\sigma_{E,i+1}$ multiplied by $\tau_{E,i}$. There is a serious drawback of this backward substitution to the linear momentum equation. Under the assumption that the shear viscosity $\eta$, occurring in $\tau_{\text{relax}}$, varies in both the radial and horizontal directions, it implies that the function $\alpha$, defined by eq. (45), also varies in both directions. This means that there is an angular coupling between the elastic shear stress tensor $\sigma_{E,i+1}$ and function $\alpha$ in the implicit Euler scheme in eq. (44). Applying the operator $\text{div}$ on $-\frac{\alpha}{1+\alpha} \sigma_{E,i+1}$ which is required in the linear momentum equation (1) results in an angularly coupled term. Consequently, the Galerkin matrix associated with this angularly coupled term is fully populated without any symmetries. This makes numerical computations significantly costly.

On the other hand, the explicit Euler scheme in eq. (43) has an explicit character, meaning that the new value of $\tau_{F,i+1}$ is expressed in terms of the quantities that are already known from the previous time step. In this case, there is no necessity to make a backward substitution from the explicit Euler scheme to the linear momentum equation, and the operator $\text{div}$ in eq. (1) at the current time $t^{i+1}$ is applied only to the elastic stress tensor $\tau_{E,i+1}$. Assuming that the elastic rigidity $\mu$ varies only in the radial direction, which is the common assumption adopted in GIA modelling, the Galerkin matrix associated with the term $\text{div} \tau_{E,i+1}$ has a block-diagonal structure, which is a great advantage for the numerical solution of the Galerkin system of equations. This is the reason for the implementation chosen by Hanyk et al. (1996) and Martinec (2000). We should, however, note that the explicit Euler scheme is not unconditionally stable. To ensure that the explicit Euler scheme in eq. (43) is contractive (and thus stable), the parameter $\alpha$ must be less than 1. Eq. (45) then dictates that the time step $\Delta t$ cannot be chosen arbitrarily to be large, but must be less than the shortest Maxwell relaxation time associated with the earth model.

There now arises a natural question, with a view of the solution of the FSM and ASM, whether there is a modification possible to the implicit Euler method that retains its numerical stability, but has the advantage of a simple numerical implementation of the explicit Euler method. The above two Euler methods can be combined such that the second term on the left-hand side of eq. (40) is numerically approximated by the implicit Euler method, while the term on the right-hand side of eq. (40) is approximated by the explicit Euler method. This combined, or semi-implicit Euler method has the form

$$\tau_{F,i+1} = \frac{1}{1+\alpha} \tau_{F,i} - \frac{\alpha}{1+\alpha} \sigma_{E,i}.$$  \tag{46}$$

Evidently, this scheme has the advantage of the explicit Euler method, that is, the Galerkin matrix associated with the linear momentum equation has a block-diagonal structure. It also has the advantage of the implicit Euler method such that the time step $\Delta t$ is not constrained by the condition $\alpha < 1$. The only concern is with the numerical accuracy of the time sequence $\tau_{F,i+1}$.

The accuracy of the explicit and semi-implicit Euler methods will be checked by comparing their performances against the analytical and semi-analytical approaches to the explicit modelling of the viscoelastic relaxation of an incompressible, spherically layered, viscoelastic sphere loaded by a unit mass as a Heaviside step function in time. The response of the sphere can be represented in the form (e.g. Spada et al. 2011)

$$y(r, t) = y^E(r) - \sum_{k=1}^{K} y^E_k(r) \frac{1 - e^{\alpha t}}{s_k} t \geq 0,$$  \tag{47}$$

where the symbol $y$ stands for the viscoelastic load Love numbers $h$, $\ell$ and $k$ for the vertical and horizontal displacements and the incremental gravitational potential, respectively. The elastic amplitudes $y^E(r)$ and viscous amplitudes $y^E_k(r)$ depend on the radial distance $r$ from the centre of the sphere. The quantities $s_k$ are characteristic frequencies, or negative inverse relaxation times. The load Love numbers $y$ are characterized by the harmonic degree $j$ of a surface load, which is assumed implicit in eq. (47).

The first numerical test is performed for a homogeneous viscoelastic sphere with the parameters of Model 1 (Wu & Ni 1996, table 1). The viscoelastic relaxation of a homogeneous sphere is carried by a single relaxation mode $M_0$, which is associated with the density discontinuity at the external surface. The negative inverse relaxation time $s_0$ and associated elastic and viscous amplitudes can be expressed analytically (Wu & Peltier 1982). The Maxwell relaxation time associated with this model is $\tau_{\text{relax}} = 739$ yr.

Fig. 1 shows the time evolution of the Love number $h_j(r)$ for harmonic degrees $j = 2, 10$ and 30 at the surface of the sphere ($r = a$) after the onset of a unit mass load at $t = 0$. The explicit and semi-implicit Euler schemes (dashed lines) are compared with the analytical solution (47) (solid lines) for three time steps, $\Delta t = 3, 30$ and 300 yr. The dashed and solid curves in the left-hand panels show an excellent agreement for all chosen time steps, which means that the explicit Euler scheme is sufficiently accurate for computing the viscoelastic response of the earth model and the three chosen time steps. On the contrary, the dashed and solid curves in the right-hand panels coincide for time step $\Delta t = 3$ yr, slightly differ for $\Delta t = 30$ yr, and significantly differ for $\Delta t = 300$ yr. This means that the accuracy of the semi-implicit Euler method decreases with increasing time step much more than for the explicit Euler method. We can conclude that the accuracy of the explicit Euler method is sufficient for all chosen time steps, while the accuracy of the semi-implicit Euler method is sufficient when using time steps of 3 and 30 yr, but insufficient for a time step of 300 yr.

The second test is performed for the two-layer, incompressible viscoelastic model with the parameters of Model 3 (Wu & Ni 1996, table 1). (There are two Maxwell relaxation times associated with this model, $\tau_{\text{relax}} = 1296$ and 6265 yr.) The viscoelastic response of this model consists of two buoyancy modes, $M_0$ and $M_1$, associated with the density discontinuity at the external surface and at the solid-solid interface between the outer solid shell and the inner solid sphere, respectively. In addition to the buoyancy modes, there are $V_1$ and $V_2$ viscoelastic modes associated with the jumps in viscosity and elastic rigidity at the internal interface. We use the propagator-matrix approach
Forward sensitivity and adjoint-state methods

Figure 1. Time relaxation of the Love number $h_j(a, t)$ for harmonic degrees $j = 2$ (orange), $j = 10$ (magenta) and $j = 30$ (green), respectively, after the onset of a unit mass load applied at the north pole of a sphere as a Heaviside step function in time. The explicit (the left-hand panels) and semi-implicit (the right-hand panels) Euler schemes (dashed lines) are compared with an analytical solution (solid lines) for three times steps, $\Delta t = 3, 30$ and 300 yr. The results apply to a homogeneous sphere with a radius $a = 6.371 \times 10^6$ m, elastic shear modulus $\mu = 1.5 \times 10^{11}$ Pa, viscosity $\eta = 3.5 \times 10^{21}$ Pa s, and unperturbed density $\rho_0 = 5.52 \times 10^3$ kg m$^{-3}$. In all figures, the dashed lines are black if they are barely distinguishable from the solid lines.

The associated elastic and viscous amplitudes are then computed analytically. Fig. 2 shows the results of similar tests as that carried out in Fig. 1, but now for the two-layer sphere. From Fig. 2, we can draw similar conclusions as for the homogeneous sphere. The numerical accuracy of both Euler schemes decreases with the increasing size of the time step. However, the accuracy of the explicit Euler scheme (45) is still sufficient for a larger time step (i.e. $\Delta t = 300$ yr), for which the semi-implicit Euler scheme (46) results in considerably different relaxation curves than for the analytical solution. We can conclude that the explicit Euler scheme (45) is more convenient for numerical computations of the viscous stress $\tau^V$ due to its higher numerical accuracy than the semi-implicit scheme (46).

7 A GENERALIZED MAXWELL VISCOELASTIC MODEL

To set up the source term $\delta F_\delta^i$ in the FSM, or to calculate the gradient of the misfit $\nabla_\delta \chi^2$ in the ASM, the gradient of the viscous stress tensor $\tau^V$ with respect to the viscosity parameters $\tilde{\eta}$, that is the term $T_{ij}$, must be computed and evolved over time. This section shows that $T_{ij}$ consists of two components, the viscous stress tensor $\tau^V$ and the generalized Maxwell stress tensor $\sigma^V$, and designs a numerically convenient time-differencing scheme for $\sigma^V$.

7.1 The gradient of viscous stress tensor

Let us first deal with the partial derivatives of the specific relaxation tensor $M$ with respect to the viscosity parameters $\tilde{\eta}$. To express the derivatives analytically, the parametrization of viscosity $\eta$ in eq. (9) must be specified. The weak formulation of eq. (D5) assumes that the
as a square-integrable function in $\mathcal{B}$, that is $\eta \in L^2(\mathcal{B})$. This allows $\eta$ to be approximated by piecewise constant functions,

$$
\eta(r) = \eta_m,
$$

(48)

where $\eta_m$ is the value of $\eta$ at a gridpoint $r = r_m$, $m = 1, 2, \ldots, M$. Likewise, let $\mu_m$ be the value of elastic rigidity $\mu$ at the gridpoint $r_m$. Then, the value of the Maxwell relaxation time (E10) at the gridpoint $r_m$ is

$$
\tau_{\text{relax},m} = \eta_m/\mu_m.
$$

(49)

This piecewise-constant-function discretization implies that the discretized eqs (48) and (49) are of the same forms as the original eqs (9) and (E10) where $\eta$ and $\mu$ are spatially varying functions. This allows us to drop the grid index $m$ at $\eta_{\text{int}}, \mu_{\text{int}}$ and $\tau_{\text{relax},,m}$ and abbreviate the following derivations. The differentiation of $M$ with respect to a viscosity parameter $\eta$ is given by

$$
\frac{\partial M}{\partial \eta} = \frac{\partial M}{\partial \tau_{\text{relax}}} \frac{\partial \tau_{\text{relax}}}{\partial \eta} = \frac{1}{\mu} \frac{\partial M}{\partial \tau_{\text{relax}}} = -\frac{1}{\mu \tau_{\text{relax}}} \left( M + \frac{\partial m_2}{\partial \tau_{\text{relax}}} I_2 \right) = -\frac{1}{\eta} \left( 1 - \frac{t}{\tau_{\text{relax}}} \right) M.
$$

(50)

where the steps are justified, in order, by the chain rule of differentiation, equations (E10), (E13) and (E9). Note that the grid index $m$ at $\eta$, $\tau_{\text{relax}}$ and $\mu$ is dropped.
Second, using successively equations (12)$\gamma$ and (50), the $m$th component of $T_\parallel$ in the viscosity-parameter space can be expressed in terms of $\tau^V$ and an additional stress tensor $\sigma^V$ as follows,

$$
[T_\parallel(t)]_m = \int_0^t \frac{\partial M}{\partial \eta} : \varepsilon(t - t')dt' \\
= -\frac{1}{\eta} \int_0^t \left(1 - \frac{t'}{\tau_{relax}}\right) M(t') : \varepsilon(t - t')dt' \\
= -\frac{1}{\eta} \int_0^t M(t') : \varepsilon(t - t')dt' - \frac{1}{\tau_{relax}} \int_0^t t' M(t') : \varepsilon(t - t')dt' .
$$

(51)

In view of eq. (E5)$\gamma$, the first term in the square brackets is the viscous stress tensor $\tau^V$. Denoting the second term as

$$
\sigma^V(t) := -\frac{1}{\tau_{relax}} \int_0^t t' M(t') : \varepsilon(t - t')dt',
$$

eq. (51) may be written in a short form,

$$
[T_\parallel(t)]_m = -\frac{1}{\tau_{relax}} (\tau^V(t) + \sigma^V(t)) .
$$

(53)

where the grid index $m$ at the viscosity is dropped and the tensors $\tau^V$ and $\sigma^V$ are evaluated for the values of the rheological parameters at the $m$th gridpoint. It is convenient to change the variable of integration in eq. (52). Substituting $\tau = t - t'$ for $t'$ yields

$$
\sigma^V(t) = -\frac{1}{\tau_{relax}} \int_0^t (t - t') M(t - t') : \varepsilon(t')dt',
$$

(54)

where $t'$ has again been used for the variable of integration.

7.2 A generalized Maxwell model

To interpret the tensor $\sigma^V$ and find a numerical method for its calculation, let us differentiate eq. (54) with respect to time and use successively equations (E5)$\gamma$, (E13) and (39),

$$
\frac{d\sigma^V(t)}{dt} = -\frac{1}{\tau_{relax}} \int_0^t \left( t - t' \right) M(t - t') : \varepsilon(t')dt' + \int_0^t M(t - t') : \varepsilon(t')dt' + \int_0^t (t - t') \frac{dM(t - t')}{dt} : \varepsilon(t')dt' .
$$

(55)

In view of (54), the last term in the square brackets is $\sigma^V$ and eq. (55) is written as

$$
\frac{d\sigma^V(t)}{dt} + \frac{\sigma^V(t)}{\tau_{relax}} = -\frac{\tau^V(t)}{\tau_{relax}} .
$$

(56)

The initial value of $\sigma^V$ is zero, that is $\sigma^V(0) = 0$, for the same reason as for $\tau^V(0) = 0$.

Let us differentiate eq. (56) one more time and substitute for the time derivative of $\tau^V$ from eq. (40), and for $\tau^V$ from eq. (56). We then obtain

$$
\frac{d^2\sigma^V(t)}{dt^2} + \frac{2}{\tau_{relax}} \frac{d\sigma^V(t)}{dt} + \frac{\sigma^V(t)}{\tau_{relax}^2} = \frac{\sigma^K(t)}{\tau_{relax}^2} .
$$

(57)

This is the stress–strain relation of a generalized Maxwell model for linear viscoelastic material consisting of two parallely connected Maxwell viscoelastic models (e.g. Ben-Menahem & Singh 1981). The tensor $\sigma^V$ is then identified as the stress tensor of this generalized Maxwell model.

The homogeneous equation associated with eq. (57) has the characteristic equation $\lambda^2 + 2\lambda/\tau_{relax} + 1/\tau_{relax}^2 = 0$, which has one repeated double eigenvalue $\lambda = -1/\tau_{relax}$. A general solution for the eigenvector is $a_1 e^{-t/\tau_{relax}} + a_2 t e^{-t/\tau_{relax}}$, where $a_1$ and $a_2$ are time-independent coefficients. The initial condition $\sigma^V(0) = 0$ implies that $a_1 = 0$. The solution of the inhomogeneous eq. (57) is then given by the Boltzmann superposition principle (e.g. Ben-Menahem & Singh 1981),

$$
\sigma^V(t) = \frac{1}{\tau_{relax}} \int_0^t t' e^{-t'/\tau_{relax}} \sigma^K(t - t')dt' .
$$

(58)
The first-order ordinary differential eq. (56) can be used to compute the generalized Maxwell stress tensor $\sigma^V(t)$ numerically. Comparing it with eq. (40) for $\tau^V$, we can see that they are of the same form. Hence, the numerical time-differencing schemes used to calculate $\tau^V$ can now be applied to $\sigma^V$. We assume that the viscous stress tensor $\tau^V(t)$ has been computed at the current time level $i + 1$ and its numerical approximation $\tau^{V,i+1}$ is provided. In analogy to eqs (E15), the explicit time-differencing scheme for the numerical solution of eq. (56) is

$$\sigma^{V,i+1} = (1 - \alpha) \sigma^{V,i} - \alpha \tau^{V,i+1},$$

while the semi-implicit time-differencing scheme is

$$\sigma^{V,i+1} = \frac{1}{1 + \alpha} \sigma^{V,i} - \frac{\alpha}{1 + \alpha} \tau^{V,i+1},$$

where $\sigma^{V,i}$ is a numerical solution of eq. (56) at the previous time level $i$ and $\alpha$ is given by eq. (45). Both schemes have an explicit character, meaning that a new value $\sigma^{V,i+1}$ is computed from $\sigma^{V,i}$, which is known from the previous time step and $\tau^{V,i+1}$ computed at the current time step. As far as the stability of the schemes is concerned, the explicit scheme (61) is contractive (and thus stable) only when the parameter $\alpha < 1$, whereas the semi-implicit scheme (62) is contractive for any value of $\alpha$, and thus unconditionally stable. As in the case of the time-differencing schemes for $\tau^V$, the only matter of concern is the numerical accuracy of the two schemes.

The accuracy of the explicit and semi-implicit time-differencing schemes will be checked by comparing their performances against the analytical and semi-analytical solutions for an incompressible, spherically layered, viscoelastic sphere loaded by a unit mass as a Heaviside step function in time. The differentiation of eq. (47) for the time evolution of the load Love number $y(t)$ with respect to the viscosity parameter $\eta_m$ yields

$$\frac{\partial y(t)}{\partial \eta_m} = -\sum_{k=1}^{K} \left[ \frac{\partial y_k^V}{\partial \eta_m} \frac{1 - e^{\beta_k t}}{s_k} - \frac{h_k^V}{s_k} \left( 1 + ts_k - e^{\beta_k t} \frac{\partial s_k}{\partial \eta_m} \right) \right] t \geq 0.$$
Forward sensitivity and adjoint-state methods

Figure 4. Time evolution of the sensitivities of the Love number $h_j(a, t)$ with respect to viscosity for harmonic degrees $j = 2$ (orange), $j = 10$ (magenta) and $j = 30$ (green), respectively, after the onset of a unit mass load at the north pole of a sphere as a Heaviside step function in time. The explicit (the left-hand panels) and semi-implicit (the right-hand panels) Euler schemes (dashed lines) are compared with an analytical solution (solid lines) for three times steps, $\Delta t = 3, 30$ and $300$ yr. The results apply to the homogeneous sphere with the parameters given in the caption of Fig. 1. In all figures, sensitivities have the physical dimension of $\text{Pa}^{-1} \text{s}^{-1}$ and are multiplied by a constant factor of $10^{21}$ to simplify the plotting.

The partial derivatives of the characteristic frequencies and associated viscous amplitudes can be evaluated analytically for a homogeneous viscoelastic sphere. After some algebraic manipulation, we find that

$$\frac{\partial y(t)}{\partial \eta} = -\frac{V}{\eta} t e^{s M_0} t \geq 0,$$

(64)

where $s^{M_0}$ is the negative inverse relaxation time of the $M_0$ buoyancy relaxation mode and $\eta$ is the shear viscosity of a homogeneous sphere.

The analytical expressions for the inverse relaxation times and the elastic and viscous amplitudes of the relaxation modes for a multilayer sphere were found by Amelung & Wolf (1994) and Wu & Ni (1996). Since the resulting expressions are algebraically complicated, and analytical differentiation would result in even more algebraically complex expressions, the sensitivities $\frac{\partial y}{\partial \eta_m}$ will be computed by direct numerical differentiation of the load Love numbers (47), the so-called brute-force method (e.g. Bevington 1969), in which the partial derivative of $y$ with respect to $\eta_m$ at the point $\hat{\eta}_m^0$ is approximated by the second-order accuracy, centred difference of two forward-model runs,

$$\left[ \frac{\partial y}{\partial \eta_m} \right] \approx \frac{y(\eta_1^0, \ldots, \eta_m^0 + \varepsilon, \ldots, \eta_M^0) - y(\eta_1^0, \ldots, \eta_m^0 - \varepsilon, \ldots, \eta_M^0)}{2\varepsilon},$$

(65)

where $\varepsilon$ refers to a perturbation applied to the nominal value of $\eta_m^0$.

The first numerical test is performed for a homogeneous viscoelastic sphere with the parameters of Model 1 (Wu & Ni 1996). Fig. 4, similarly to Figs 1 and 2, shows the time evolution of the sensitivities $\frac{\partial h_j}{\partial \eta}$ for harmonic degrees $j = 2, 10$ and $30$ after the onset of a unit mass load at $t = 0 +$. The explicit (the left-hand panels) and semi-implicit (the right-hand panels) Euler schemes (dashed lines) are compared...
with an analytical solution (solid lines) given by eq. (64) for three times steps, $\Delta t = 3, 30$ and $300$ yr. The dashed and solid curves in the right-hand panels, that is the semi-implicit case, show excellent agreement for all chosen times steps, which means that the semi-implicit Euler scheme is sufficiently accurate for this model and the three time steps. On the contrary, the dashed and solid curves in the left-hand panels, that is the explicit case, coincide for time step $\Delta t = 3$ yr, slightly differ for $\Delta t = 30$ yr and considerably differ for $\Delta t = 300$ yr. This means that the accuracy of the explicit Euler method decreases with increasing time step, hence the accuracy of the explicit method is insufficiently accurate for the time step $\Delta t = 300$ yr. Note that it was the other way around for Figs 1 and 2.

The second test is performed for the two-layer, incompressible viscoelastic model with the parameters of Model 3 (Wu & Ni 1996). The explicit and semi-implicit Euler schemes are now checked against the brute-force differentiation. We first vary $\varepsilon$ in the viscosity parameters in eq. (65) to find the value for which the approximation error of the brute-force method is sufficiently small for a meaningful comparison with the Euler differencing schemes. We found (not shown here) that the approximation error of the brute-force method decreases when $\varepsilon$ decreases and the value $\varepsilon \approx 0.001 \times \tilde{\eta}$ is small enough to run the benchmark against the Euler differentiating scheme.

Figs 5 and 6 show the accuracy tests of the Euler schemes for the two-layer sphere. The sensitivities of the vertical-displacement Love number $h_j$ with respect to the lower- and upper-mantle viscosities, $\eta_{\text{LM}}$ and $\eta_{\text{UM}}$, respectively, that is the partial derivatives $\partial h_j / \partial \eta_{\text{LM}}$ and $\partial h_j / \partial \eta_{\text{UM}}$ for $j = 2, 10$ and $30$, are plotted as functions of time after the onset of a unit mass load at $t = 0^+$. The explicit and semi-implicit Euler schemes (dashed lines) are checked against the brute-force differentiation (65) (solid line) for the same three times steps as in the previous tests. From these figures, we can deduce similar conclusions as for a homogeneous sphere. The numerical accuracy of both Euler schemes decreases with increasing size of the time step. However, the accuracy of the semi-implicit Euler scheme (62) is still sufficient for a larger time step (e.g. $\Delta t = 300$ yr) to be used, for which the application of the explicit Euler scheme (61) results in significantly different relaxation curves than the brute-force differentiation. We can conclude that the semi-implicit Euler scheme (62) is more convenient for
Figure 6. As for Fig. 5, but for the sensitivities of $h_j(a, t)$ with respect to the upper-mantle viscosity.

Numerical computations of the generalized Maxwell stress $\sigma^V$ due to its better numerical accuracy than the explicit scheme (61). These figures also show that, after the onset of a surface load, the induced surface vertical displacement is influenced by the lower-mantle viscosity mainly at longer spatial-wavelength scales (i.e. lower degrees $j$), while the upper-mantle viscosity predominantly influences shorter spatial wavelength scales (i.e. higher degrees $j$). We can also see that the upper-mantle sensitivities are monotonically increasing functions of harmonic degree $j$, while it is an opposite case for the lower-mantle sensitivities.

For the sake of completeness, Fig. 7 shows the sensitivities $\partial h_j / \partial \eta_{LM}$ and $\partial h_j / \partial \eta_{UM}$ as functions of the radial distance $r$ at the last computational time, $t = 12$ kyr. We can see that the lower-mantle sensitivities (the top left-hand panel in Fig. 7) monotonically increase with radius in the lower mantle up to the 670 km discontinuity. Then in the upper mantle, the lower-mantle sensitivities decrease for degrees $j = 10$ and $30$. This means that observations of the vertical displacement at the surface are less sensitive to the lower-mantle viscosity than 'hypothetical' observations of vertical deformations of the 670 km discontinuity. The upper-mantle sensitivities (the top right-hand panel in Fig. 7) are concentrated in the upper mantle and nearly vanish in the lower mantle. In addition, the upper-mantle sensitivity for degree $j = 2$ is effectively equal to zero in the upper mantle. We can conclude that, in contrast to the lower-mantle sensitivities, the Earth’s surface is the place where vertical displacements of medium and shorter spatial wavelength scales are most sensitive to the upper-mantle viscosity, which is in fact well-known in the GIA community.

The middle panels in Fig. 7 are the same as for the upper panels, but for the sensitivities of the horizontal-displacement load Love number $\ell_j$ with respect to the lower- and upper-mantle viscosities, respectively, that is the partial derivatives $\partial \ell_j / \partial \eta_{LM}$ and $\partial \ell_j / \partial \eta_{UM}$. We can see that the lower-mantle sensitivities (the middle left-hand panel) also decrease with increasing degree $j$, but faster than those for the vertical displacements (the top left-hand panel). For instance, the lower-mantle sensitivities for horizontal displacements are effectively zero for angular degrees $j > 10$. Comparing the vertical scales on the middle left-hand and middle right-hand panels in Fig. 7 shows that the observations of the surface horizontal displacement are predominantly sensitive to the upper-mantle viscosity.
Figure 7. The sensitivities of the vertical-displacement load Love number \( h_j(r, t) \) (top panels), the horizontal-displacement load Love number \( \ell_j(r, t) \) (middle panels) and the incremental-gravity-potential load Love number \( k_j(r, t) \) (bottom panels) with respect to the lower (the left-hand panels) and upper (the right-hand panels) mantle viscosities for \( j = 2 \) (orange), \( j = 10 \) (magenta) and \( j = 30 \) (green) as functions of the radial distance \( r \) at computational time \( t_1 = 12 \) kyr. The results apply to the two-layer sphere with the parameters given in the caption of Fig. 2.

Finally, the bottom panels in Fig. 7 show the sensitivities of the gravitational–potential load Love number \( k_j \) with respect to the lower (the left-hand panel) and upper (the right-hand panel) mantle viscosities, respectively. Since the gravitational potential changes and the surface vertical deformations are used for constraining GIA models, let us have a look at the similarities and the differences between the sensitivities for \( k_j \) and \( h_j \). In contrast to the lower-mantle sensitivities for \( h_j \), the lower-mantle sensitivities for \( k_j \) monotonically decrease with increasing degree \( j \). In addition, these sensitivities do not vary significantly in the upper mantle (the bottom left-hand panel) which is not the case for the lower-mantle sensitivities for \( h_j \) (the top left-hand panel of Fig. 7). As far as the upper-mantle sensitivities for \( k_j \) are concerned, they are concentrated in the upper mantle, as are those for \( h_j \), but the amplitudes of \( \partial k_j / \partial n_{\text{LM}} \) do not increase monotonically with increasing degree \( j \).

Fig. 7 demonstrates the limits in the use of surface-observations of GIA to constrain mantle viscosity, and consequently shows where other \textit{a priori} information, for example, from compositional modelling, would help most.

It is important to realize that the sensitivities themselves depend on the values of the inferred viscosity and layering that is used in the models. This has an important consequence for GIA inverse modelling. When the viscosity parameters are iteratively updated such that a data misfit is minimized, the forward sensitivities change for the updated viscosity parameters. It is thus necessary to recompute the forward sensitivities after several iteration steps to check whether they have not changed significantly. If this is the case, the viscosity parametrization must be redesigned and the GIA inversion performed from the beginning.

The results of the numerical tests carried out for the viscous stress tensor \( \tau^V \) in Section 2.3 and for the generalized Maxwell stress tensor \( \sigma^V \) performed in this section can therefore be summarized as follows. The explicit Euler method is sufficiently accurate for resolving the time evolution of the viscous stress tensor \( \tau^V \), but insufficiently accurate for the time evolution of the generalized Maxwell stress tensor \( \sigma^V \). On the contrary, the semi-implicit Euler method is insufficiently accurate for time evolution of the viscous stress tensor \( \tau^V \), but sufficiently accurate for the generalized Maxwell stress tensor \( \sigma^V \). This conclusion holds for the case where the size of the time step is comparable to the
shortest Maxwell relaxation time. If the time step is significantly shorter (e.g. by an order in magnitude) than the shortest Maxwell relaxation time, then both Euler schemes are sufficiently accurate for computing the time evolution of both stress tensors $\tau^V$ and $\sigma^V$.

8 CONCLUSIONS

The presented FSM and ASM for interpreting GIA-related observations can mathematically be classified as free boundary-value problems, however, with a different character than a free boundary-value problem for the forward method of GIA. The forward method of GIA implements an ocean load as a free boundary-value function over an ocean area with a free geometry, that is, an ocean load and the shape of ocean itself are being sought in addition to the deformation and gravitational-potential increment fields, by solving the forward method. The FSM and ASM apply the adjoint ocean load also as a free boundary-value function, but over an ocean area with a fixed geometry given by the ocean function determined by the forward method.

The solvability conditions of free boundary-value problems also differ whether the forward sensitivity or the adjoint-state equations are solved. The forward method and the FSM are unconditionally solvable, regardless of whether or not a combined ice and ocean-water load or the adjoint ocean load contains the first-degree spherical harmonics. On the contrary, the existence of a solution of the ASM is guaranteed only if nine conditions on the misfit between the given GIA-related data and the forward model predictions are fulfilled. The six uniqueness conditions, needed to be supplied to all three boundary-value problems can be defined by the requirement that a rigid-body translation and a rigid-body rotation, expressed in the corresponding field variables, are to vanish.

The forward method of GIA requires the viscous stress tensor $\tau^V$ to evolve over time, while the FSM and ASM also require the calculation of the time evolution of the generalized Maxwell stress tensor $\sigma^V$. The numerical tests carried out in this study show that the explicit time-differencing scheme is sufficiently accurate for the time evolution of the viscous stress tensor $\tau^V$, confirming an earlier study by Hanyk et al. (1996). However, this scheme is insufficiently accurate for the time evolution of the generalized Maxwell stress tensor $\sigma^V$. On the contrary, the semi-implicit time-differencing scheme is insufficiently accurate for the time evolution of the viscous stress tensor $\tau^V$, but sufficiently accurate for the evolution of the generalized Maxwell stress tensor $\sigma^V$. This holds for the case where the size of the time step is comparable to the shortest Maxwell relaxation time. If the step size is, however, significantly less than the shortest Maxwell relaxation time (e.g. by an order in magnitude), then both Euler schemes are adequate to compute the time evolution of both stress tensors. In other words, if the explicit time-differencing scheme is used for the time evolution of $\sigma^V$ and a sufficient accuracy of $\sigma^V$ is required, the time step must be chosen rather short. Then, the computation of $\sigma^V$ takes long computational time.

When establishing the FSM and ASM of GIA, our aim was to derive such approaches that are closely associated with the forward solver of GIA developed by Martinec (2000). Since this method, as well as the approach of Al-Attar & Tromp (2014), is based on a continuous form of the forward model equations, which are then discretized by spectral and finite elements, we proceeded in a similar way for the FSM and ASM. The advantage of this approach is that all three methods (forward, FSM and ASM) have the same matrix of the system of equations and use the same methodology for the implementation of the time evolution of the stresses. The only difference between the forward method and the FSM and ASM is that the different numerical differencing schemes for the time evolution of the Maxwell and generalized Maxwell viscous stresses are used in the respective methods. However, it requires only a little extra computational time to carry out the FSM and ASM numerically. It should be mentioned an alternative approach to derive the so-called discrete FSM and ASM (e.g. Chavent 2009) is based on the discrete form of the forward model equations. This approach may deserve attention in future studies on GIA inverse modelling.

Let us now remark upon the computation time and memory required for calculating the forward sensitivities and the gradient of a data misfit. The forward model variables in the approach developed by Martinec (2000) are parametrized, for a fixed time, by tensor surface spherical harmonics in the angular direction and by piecewise linear finite elements in the radial direction with finite cut-off degrees, $j_{\text{max}}$ and $k_{\text{max}}$, respectively. For $j_{\text{max}} = 180$ and $k_{\text{max}} = 38$, the solution of the forward model equations requires 2.5 GB memory and takes about 10 hr of CPU time for 110 time steps on a computer with the AMD 64-bit architecture. The code scales well with an OpenMP parallelization such that the time consumption is reduced to 0.7 hr when using 20 CPUs (V. Klemann, personal communication).

The brute-force method requires two forward model runs for each viscosity parameter to approximate the forward sensitivity with respect to a viscosity parameter by the centred difference. Thus, the brute-force method requires twice the forward-model computation time. In the forward sensitivity analysis, the forward model equations are solved simultaneously with the forward sensitivity equations for each viscosity parameter, resulting in the forward sensitivity with respect to a viscosity parameter. Thus, the FSM requires little more than the forward-model computational time. The FSM is computationally feasible for ranking the relative importance of viscosity parameters for a forward-modelled response and refining an initial viscosity-model parametrization.

Importantly, the FSM provides, among other forward sensitivities, the sensitivity of the sea level indicator with respect to a viscosity stratification. This quantity can be used for ranking the relative importance of the sea level indicator data at particular locations along the coast when GIA inverse modelling is carried out to determine viscosity stratification.

The ASM does not provide forward sensitivities, but the gradient of a data misfit with respect to the model parameters and offers a very efficient tool for gradient-based optimization methods. The computational cost of the ASM is independent of the number of viscosity parameters, and one computation run requires about the same as the FSM computation time. The ASM is thus highly efficient for calculating the gradient of a data misfit for models with large numbers of parameters. However, the forward-model solution for each time step must be stored, hence the memory demands scale linearly with the number of time steps. This is the main drawback of the ASM.
ACKNOWLEDGEMENTS

The authors would like to thank Kevin Fleming and two anonymous reviewers for their comments on the manuscript that have helped to improve the manuscript. This work has been carried out within the framework of the research program 11/RFP.1/GEO/3309 financed by the Science Foundation of Ireland. ZM acknowledges this support.

REFERENCES


Bevington, P.R., 1969. Data Reduction and Error Analysis for the Physical Sciences, McGraw-Hill.


Morse, P.W. & Feshbach, H., 1953. Methods of Theoretical Physics, McGraw-Hill.


APPENDIX A: THE VISCOSITY GRADIENT OF THE SEA LEVEL EQUATION AND OCEAN FUNCTION

We follow the implementation of the sea level equation as described by Hagedoorn et al. (2007) who identified the uniform and non-uniform contributions to the change in relative sea level height,

\[ s(\Omega, t) = s^{\text{UF}}(\Omega, t) + s^{\text{NU}}(\Omega, t). \]  

(A1)

The respective contributions are given by the application of two principles: (i) mass is conserved in the ice-ocean system, and (ii) the sea surface coincides with the geoid. These principles result in

\[ s^{\text{UF}}(\Omega, t) = h^{\text{UF}}(t)\mathcal{O}(\Omega, t), \]

\[ s^{\text{NU}}(\Omega, t) = [\bar{e}(\Omega, t) - u(\Omega, t)]\mathcal{O}(\Omega, t) - \zeta^{(0)}(\Omega)[\mathcal{O}(\Omega, t) - \mathcal{O}^{(0)}(\Omega)]. \]  

(A2)

Here, \( u \) is the radial displacement, \( u = u \cdot \mathbf{e}_r \), \( e \) is the geoid height, \( e = -\phi_1/g_0 \) (both \( u \) and \( \phi_1 \) are considered at the Earth’s surface), \( \mathcal{O} \) is the ocean function, which is equal to 1 for oceans and 0 for continents, \( \zeta^{(0)} \) is the topography of the undeformed Earth and \( \mathcal{O}^{(0)} \) is the associated ocean function. Moreover,

\[ h^{\text{UF}}(t) = -\frac{1}{\mathcal{O}_0(t)}\left[ \frac{1}{\mathcal{O}^{(0)}}(0) + s^{\text{NU}}(t) \right], \]  

(A3)

where \( f_{00} \) denotes the zero-degree coefficient of the spherical harmonic representation of the function \( f(\Omega) \).

We aim here to find the expression for the viscosity gradient of the change in relative sea level height \( s(\Omega, t) \) (see eq. A1), meaning we are searching for \( \nabla s \). We start with the zero-degree coefficient \( s^{\text{NU}}(t) \) on the right-hand side of eq. (A3). The projection of eq. (A2) to the zero-degree spherical harmonic \( Y_{00}(\Omega) = 1/\sqrt{4\pi} \) is

\[ s^{\text{NU}}_{00}(t) = \int_{\Omega_0} \left[ \left[ \bar{e}(\Omega, t) - u(\Omega, t) - \zeta^{(0)}(\Omega) \right] \mathcal{O}(\Omega, t) + \zeta^{(0)}(\Omega)\mathcal{O}^{(0)}(\Omega) \right] Y_{00}(\Omega)d\Omega, \]  

(A4)

where \( \Omega_0 \) is the full solid angle and \( d\Omega \) is its infinitesimal element. When the viscosity structure of the earth model is changed, the deformation fields \( e \) and \( u \), and the ocean function \( \mathcal{O}(\Omega, t) \) are changed as a response. On the contrary, we assume that the topography/bathymetry and the ocean function of the undeformed Earth, \( \zeta^{(0)}(\Omega) \) and \( \mathcal{O}^{(0)}(\Omega) \), respectively, are defined independently of the viscosity stratification. Hence, the differentiation of eq. (A4) with respect to the viscosity parameters \( \eta \) yields

\[ \nabla s^{\text{NU}}_{00} = \int_{\Omega_0} (E - U)\mathcal{O}Y_{00}d\Omega + \int_{\Omega_0} (e - u - \zeta^{(0)}) \nabla_\Omega \mathcal{O} Y_{00}d\Omega, \]  

(A5)

where

\[ U := U(\mathbf{a}, \Omega, t) \cdot \mathbf{e}_r, \]

\[ E := -\Phi_1(\mathbf{a}, \Omega, t)/g_0, \]  

(A6)

and \( U \) and \( \Phi_1 \) are defined by eq. (15). To abbreviate the notation, we have omitted the arguments \( \Omega \) and \( t \) of functions; they are the same as those in eq. (A4). Scaling the last equation by the zero-degree coefficient of the ocean function,

\[ \mathcal{O}_{00} = \int_{\Omega_0} \mathcal{O}Y_{00}d\Omega, \]  

(A7)

we obtain

\[ \frac{1}{\mathcal{O}_{00}} \nabla s^{\text{NU}}_{00} = E - U + \frac{1}{\mathcal{O}_{00}} \int_{\Omega_0} (e - u - \zeta^{(0)}) \nabla_\Omega \mathcal{O} Y_{00}d\Omega, \]  

(A8)

where \( \mathcal{F} \) is the ocean average value of function \( F(\Omega) \),

\[ \mathcal{F} := \frac{\int_{\Omega_0} F(\Omega)\mathcal{O}(\Omega)d\Omega}{\int_{\Omega_0} \mathcal{O}(\Omega)d\Omega}. \]  

(A9)
In the next step, we apply the gradient operator $\nabla_h$ to eq. (A3),

$$\nabla_h h^{UF} = -\frac{1}{C_{00}} \nabla h^{NL}_0 - \frac{h^{UF}}{C_{00}} \nabla h_{00}.$$  \hfill (A10)

Differentiating eq. (A7) with respect to $\tilde{\eta}$, substituting this result and eq. (A8) into eq. (A10) and realizing that $h^{UF}$ does not depend on the angular coordinates $\Omega$ (but only on time $t$), we obtain

$$\nabla_h h^{UF} = E - U - \frac{Y_{00}}{C_{00}} \int_{\Omega_0} (e - u + h^{UF} - \xi^{(0)}) \nabla h \Omega d\Omega.$$  \hfill (A11)

Finally, we apply the gradient operator $\nabla_h$ to eq. (A1),

$$\nabla_h \xi = (E - U + \nabla_h h^{UF}) \Omega + (e - u + h^{UF} - \xi^{(0)}) \nabla h \Omega.$$  \hfill (A12)

Substituting for $\nabla_h h^{UF}$ from eq. (A11), we end up with

$$\nabla_h \xi = (E - U - E - U) \Omega + (e - u + h^{UF} - \xi^{(0)}) \nabla h \Omega - \frac{Y_{00}}{C_{00}} \int_{\Omega_0} (e - u + h^{UF} - \xi^{(0)}) \nabla h \Omega d\Omega.$$  \hfill (A13)

To recall the functional dependency, $h^{UF} = h^{UF}(t)$, $\xi^{(0)} = \xi^{(0)}(\Omega)$, and all other field variables and the ocean function are dependent on both $\Omega$ and $t$.

To derive the viscosity gradient of the ocean function, that is $\nabla_h \Omega$, we first define the coastline function $C$,

$$C(s, t) := \{ \Omega \in \Omega_0 \mid e(\Omega, t) - u(\Omega, t) + h^{UF}(t) - \xi^{(0)}(\Omega) = 0 \},$$  \hfill (A14)

where $s$ denotes the position on $C$. Clearly, $C$ is a subset of $\Omega_0$. In view of that, the ocean function $\Omega$ is defined as the Heaviside function $H$ with a unit step at the coastline $C$,

$$\Omega(\Omega, t) := H \{ e(\Omega, t) - u(\Omega, t) + h^{UF}(t) - \xi^{(0)}(\Omega) \}.$$  \hfill (A15)

This is the definition of the ocean function by Hagedoorn et al. (2007) which is simplified for the case when a floating-ice contribution is not considered in the implementation of the sea level equation. Applying the gradient operator $\nabla_h$ to the ocean function and using the chain rule of differentiation, we obtain

$$\nabla_h \Omega = \frac{\partial H(z)}{\partial z} \left( E - U + \nabla_h h^{UF} \right).$$  \hfill (A16)

where $z := e - u + h^{UF} - \xi^{(0)}$ is the argument of $H$, while $U$ and $E$ are defined by eq. (A6). To simplify the notation, we have again omitted the arguments $\Omega$ and $t$. The differentiation of the Heaviside function $H(z)$ with respect to $z$ gives the Dirac delta function $\delta(z)$. In our case, where the variable $z$ is located at the coastline $C$, we have

$$\nabla_h \Omega = \delta(C) \left( E - U + \nabla_h h^{UF} \right).$$  \hfill (A17)

where $\delta(C)$ is the Dirac delta function with support on the coastline $C$. Its action on a smooth function $f(\Omega)$ is given by a line integral along the coastline,

$$\int_{\Omega_0} f(\Omega) d\Omega = \int_C f(s) ds,$$  \hfill (A18)

where $ds$ is the arc length of $C$. Substituting for $\nabla_h h^{UF}$ from eq. (A11), we obtain

$$\nabla_h \Omega = \delta(C) \left[ E - U - E - U - \frac{Y_{00}}{C_{00}} \int_{\Omega_0} (e - u + h^{UF} - \xi^{(0)}) \nabla h \Omega d\Omega \right].$$  \hfill (A19)

The integration in the last term on the right-hand side is carried out along the coastline $C$ due to the multiplication by $\nabla_h \Omega$ in the argument. Moreover, the argument is proportional to the variable $z = e - u + h^{UF} - \xi^{(0)}$, which is equal to zero at $C$. Hence, the integral in eq. (A19) vanishes and we obtain a simple formula for $\nabla_h \Omega$,

$$\nabla_h \Omega = (E - U - E - U) \delta(C).$$  \hfill (A20)

To recall the functional dependency, $U$, $E$ and $\nabla_h \Omega$ are dependent upon $\Omega$ and $t$, while $E - U$ is only dependent on $t$. Due to the same arguments, the last two terms in eq. (A13) vanish and the viscosity gradient of the change in relative sea level height is reduced to

$$\nabla_h \xi = (E - U - E - U) \Omega.$$  \hfill (A21)
APPENDIX B: A DETAILED DERIVATION OF THE ADJOINT METHOD

Let us assume that the given GIA-related data and the forward model predictions are available over a time interval \( t \in (0, T) \). We introduce the reverse, or adjoint time \( \hat{t} \) by

\[
\hat{t} := T - t. 
\]  

Multiplying the forward sensitivity equations (14) and boundary conditions (16) by sufficiently smooth functions \( \hat{u} = \hat{u}(r, \hat{t}) \) and \( \hat{\phi}_1 = \hat{\phi}_1(r, \hat{t}) \), respectively, then integrating over the sphere \( B \) and its surface \( \partial B \), respectively, and over the time interval \( (0, T) \), and summing up the results, yields the multilinear form

\[
\int_0^T \int_B \left\{ \text{div} (T + T_\partial) - \varrho_0 \text{grad} \Phi_1 + F \right\} \cdot \hat{u} + \left[ \frac{1}{4\pi G} \nabla^2 \Phi_1 + \text{div} (\varrho_0 U) \right] \hat{\phi}_1 \, dV \, dt 
\]

\[
- \int_0^T \int_B \left[ \mathbf{e}_r \cdot (T^- + T^-_\partial) + \varrho_0 \nabla \sigma \mathbf{e}_r \right] \cdot \hat{u}^- \, dS \, dt 
\]

\[
+ \int_0^T \int_B \left( - \frac{1}{4\pi G} \text{grad} \Phi_1^+ \cdot \mathbf{e}_r - \varrho_0 (U^- \cdot \mathbf{e}_r) - \nabla \sigma \right) \hat{\phi}_1^- \, dS \, dt 
\]

\[
+ \int_0^T \int_B \frac{1}{4\pi G} (\text{grad} \Phi_1^+ \cdot \mathbf{e}_r) \hat{\phi}_1^- \, dS \, dt = 0, 
\]

(B2)

where the field variables \( U, \Phi_1 \) and \( T, T_\partial \) are given by eqs (15) and (12), respectively, and

\[
F := -\text{div} (\varrho_0 U) g_0 + \text{grad} (\varrho_0 U \cdot g_0). 
\]  

(B3)

The reason for introducing the reverse time \( \hat{t} \) in the argument of the adjoint functions \( \hat{u}(\hat{t}) \) and \( \hat{\phi}_1(\hat{t}) \) is the intention of making use of the convolution identity in eq. (B5). All other time-dependent variables in the multilinear form (B2) are, however, functions of ‘forward’ time \( t \). To abbreviate the notation, we drop the argument \( \hat{t} \) of the hatted functions, for example, \( \hat{u}(\hat{t}) \) is shortened to \( \hat{u} \). Likewise, we drop the argument \( t \) of the unhatted functions, hence, for example, \( U(t) \) is shortened to \( U \).

We next aim to transform the multilinear form (B2) such that \( \hat{u}(r, \hat{t}) \) and \( \hat{\phi}_1(r, \hat{t}) \) exchange their positions with \( U(r, t) \) and \( \Phi_1(r, t) \). To make the following derivation transparent, we split the multilinear form into several terms and make arrangements for each term separately.

B1 Arrangement of the terms containing \( T + T_\partial \)

Making use of Green’s theorem for a symmetric tensor \( \mathbf{r} \),

\[
\int_S (\text{div} \mathbf{r} \cdot \mathbf{u}) \, dV = \int_S (\mathbf{e}_r \cdot \mathbf{r}^- \cdot \mathbf{u}^-) \, dS - \int_S (\mathbf{r} : \mathbf{e}) \, dV, 
\]

and the convolution identity (C6) applied to \( T \) and \( \hat{t} \), that is

\[
\int_0^T T(t) : \hat{\mathbf{e}}(\hat{t}) \, dt = \int_0^T \hat{t}(\hat{t}) : E(t) \, dt, 
\]

(B5)

where \( E(t) \) is defined by eq. (13), \( \hat{\mathbf{e}}(\hat{t}) := [\text{grad} \, \hat{u}(\hat{t}) + \text{grad} \, \hat{t} \hat{u}(\hat{t})]/2 \) and \( \hat{t}(\hat{t}) \) is defined by eq. (E2) for the hatted variables,

\[
\hat{t}(\hat{t}) := m(0) : \hat{\mathbf{e}}(\hat{t}) + \int_0^\hat{t} M(t') : \hat{\mathbf{e}}(\hat{t} - t') \, dt', 
\]

(B6)

the terms in the multilinear form (B2) containing \( T + T_\partial \) can be arranged as

\[
\int_0^T \int_B \text{div} (T + T_\partial) \cdot \hat{u} \, dV \, dt - \int_0^T \int_B \mathbf{e}_r \cdot (T^- + T^-_\partial) \cdot \hat{u}^- \, dS \, dt 
\]

\[
= -\int_0^T \int_B (T + T_\partial) : \hat{\mathbf{e}} \, dV \, dt 
\]

\[
= -\int_0^T \int_B \hat{t} : E \, dV \, dt - \int_0^T \int_B T_\partial : \hat{\mathbf{e}} \, dV \, dt 
\]

\[
= \int_0^T \int_B \text{div} \hat{\mathbf{e}} \cdot U \, dV \, dt - \int_0^T \int_B \mathbf{e}_r \cdot \hat{t}^- \cdot U^- \, dS \, dt - \int_0^T \int_B T_\partial : \hat{\mathbf{e}} \, dV \, dt. 
\]

(B7)

Note again that the unhatted variables are functions of ‘forward’ time \( t \), while the hatted variables are of the reverse time \( \hat{t} \).
B2 Arrangement of the terms containing $\hat{\phi}_1$

Making use of Green’s theorems,

$$\int_B \nabla (\varrho_0 u) \hat{\phi}_1 \, dV = \int_{\partial B} \varrho_0 (e_r \cdot u^-) \hat{\phi}_1^- \, dS - \int_B \varrho_0 (u \cdot \nabla \hat{\phi}_1) \, dV,$$

$$\int_B \nabla^2 \Phi_1 \hat{\phi}_1 \, dV = \int_{\partial B} (e_r \cdot \nabla \Phi^-) \hat{\phi}_1^- \, dS - \int_B (\nabla \Phi_1 \cdot \nabla \hat{\phi}_1) \, dV.$$  \hspace{1cm} (B8)

we interchange the functions $\Phi_1$ and $\hat{\phi}_1$ and subtract the new equation from the original one and obtain Green’s second identity of the form

$$\int_B \nabla^2 \Phi_1 \hat{\phi}_1 \, dV - \int_{\partial B} (e_r \cdot \nabla \Phi^-) \hat{\phi}_1^- \, dS = \int_B \nabla^2 \Phi_1 \hat{\phi}_1 \, dV - \int_{\partial B} (e_r \cdot \nabla \Phi^-) \hat{\phi}_1^- \, dS.$$  \hspace{1cm} (B9)

The sum of the terms in the multilinear form (B2) containing function $\hat{\Phi}_1$ can now be arranged as

$$\int_B \left[ - \varrho_0 \nabla \Phi_1 \cdot \hat{u} + \nabla (\varrho_0 U) \hat{\Phi}_1 + \frac{1}{4\pi G} \nabla^2 \Phi_1 \hat{\Phi}_1 \right] \, dV - \int_{\partial B} e_r \cdot \left( \frac{1}{4\pi G} \nabla \Phi_1^- + \varrho_0 U^- \right) \hat{\phi}_1^- \, dS$$

$$= \int_B \left[ - \varrho_0 \nabla \hat{\phi}_1 \cdot U + \nabla (\varrho_0 \hat{u}) \Phi_1 + \frac{1}{4\pi G} \nabla^2 \hat{\phi}_1 \Phi_1 \right] \, dV - \int_{\partial B} e_r \cdot \left( \frac{1}{4\pi G} \nabla \hat{\phi}_1^- + \varrho_0 \hat{u} \right) \Phi_1^- \, dS.$$  \hspace{1cm} (B10)

We further require that $\hat{\phi}_1$ passes continuously through $\partial B$,

$$\left[ \hat{\phi}_1 \right]^+ = 0,$$  \hspace{1cm} (B11)

and $\Phi_1$ and $\hat{\phi}_1$ are harmonic outside $B$ and vanish at infinity. Under these assumptions and the continuity condition (16) for $\Phi_1$, the last integral in the multilinear form (B2) can be arranged as

$$\int_{\partial B} (\nabla \Phi_1^+ \cdot e_r) \hat{\phi}_1^- \, dS = \int_{\partial B} (\nabla \hat{\phi}_1^+ \cdot e_r) \Phi_1^- \, dS.$$  \hspace{1cm} (B12)

B3 Arrangement of the term containing $F$

Making use of Green’s theorem,

$$\int_B U \cdot \nabla (u \cdot g_0) \, dV = \int_{\partial B} (e_r \cdot U^-)(u^- \cdot g_0) \, dS - \int_B \nabla U (u \cdot g_0) \, dV,$$  \hspace{1cm} (B13)

and assuming that $\partial B$ is a sphere and $g_0$ has only a radial component on $\partial B$, the position $U$ and $u$ in the surface integral is interchangeable and we obtain the identity

$$\int_B U \cdot \nabla (u \cdot g_0) \, dV + \int_B \nabla U (u \cdot g_0) \, dV = \int_B u \cdot \nabla (U \cdot g_0) \, dV + \int_B \nabla u (U \cdot g_0) \, dV.$$  \hspace{1cm} (B14)

Moreover, under the assumptions that $\varrho_0 = \varrho_0(r)$ and $g_0 = -g_0(r)e_r$, the following differential identity holds

$$\nabla (\varrho_0 u) g_0 - \nabla (\varrho_0 u \cdot g_0) = \varrho_0 \left( \nabla u \cdot g_0 - \nabla (u \cdot g_0) \right).$$  \hspace{1cm} (B15)

The term in the multilinear form (B2) containing $F$ can now be arranged as

$$\int_B F \cdot \hat{u} \, dV = \int_B \left[ - \nabla (\varrho_0 U) g_0 + \nabla (\varrho_0 U \cdot g_0) \right] \cdot \hat{u} \, dV$$

$$= \int_B \varrho_0 \left[ - (\nabla \hat{u}) g_0 + \nabla (\hat{u} \cdot g_0) \right] \cdot U \, dV$$

$$= \int_B \left[ - (\varrho_0 \hat{u}) g_0 + \nabla (\varrho_0 \hat{u} \cdot g_0) \right] \cdot U \, dV$$

$$= \int_B \hat{f} \cdot U \, dV,$$  \hspace{1cm} (B16)

where

$$\hat{f} := \text{div} (\varrho_0 \hat{u}) g_0 + \text{grad} (\varrho_0 \hat{u} \cdot g_0).$$  \hspace{1cm} (B17)
B4 Arrangement of the term containing $\nabla_{\Sigma} \sigma$

The surface integral over $\partial B$ containing $\nabla_{\Sigma} \sigma$ occurring in the multilinear form (B2) can then be arranged as follows:

$$\int_{\partial B} (g_{0} \dot{u} + \dot{\phi}) \nabla_{\Sigma} \sigma \, dS$$

$$= e^{0} \int_{\partial B} (g_{0} \dot{u} + \dot{\phi}) \nabla_{\Sigma} \sigma \, dS$$

$$= e^{0} g_{0} \int_{\partial B} (\dot{u} - \dot{\phi}) \nabla_{\Sigma} \sigma \, dS$$

$$= -e^{0} g_{0} a^{2} \int_{\Omega_{0}} \left[ \dot{\varepsilon} - \dot{\bar{\varepsilon}} \right] \left[ E(\Omega) - U(\Omega) \right] \nabla_{\Sigma} \sigma \, d\Omega$$

$$= -e^{0} g_{0} a^{2} \int_{\Omega_{0}} \left[ \dot{\varepsilon} - \dot{\bar{\varepsilon}} \right] \left[ E(\Omega) - U(\Omega) \right] \nabla_{\Sigma} \sigma \, d\Omega$$

$$+ e^{0} g_{0} a^{2} \int_{\Omega_{0}} \left[ \dot{\varepsilon} - \dot{\bar{\varepsilon}} \right] \left[ E(\Omega) - U(\Omega) \right] \nabla_{\Sigma} \sigma \, d\Omega$$

$$= -e^{0} g_{0} a^{2} \int_{\Omega_{0}} \left[ \dot{\varepsilon} - \dot{\bar{\varepsilon}} \right] \left[ E(\Omega) - U(\Omega) \right] \nabla_{\Sigma} \sigma \, d\Omega$$

$$= -e^{0} g_{0} a^{2} \int_{\Omega_{0}} \left[ \dot{\varepsilon} - \dot{\bar{\varepsilon}} \right] \left[ E(\Omega) - U(\Omega) \right] \nabla_{\Sigma} \sigma \, d\Omega$$

$$= e^{0} \int_{\partial B} \left[ g_{0} U(\Omega) + \Phi_{1}(\Omega) \right] \left[ \dot{\varepsilon} - \dot{\bar{\varepsilon}} \right] \nabla_{\Sigma} \sigma \, dS$$

$$= \int_{\partial B} \left[ g_{0} U(\Omega) + \Phi_{1}(\Omega) \right] \sigma \nabla_{\Sigma} \sigma \, dS, \quad (B18)$$

where the steps are justified, in order, by eq. (18), the relation $\dot{\varepsilon} = -\dot{\phi}_{1}/g_{0}$, eqs (A21) and (A9), the exchange of integrations over $\Omega_{0}$ and $\Omega'_{0}$, and eqs (A9) and (A21). Moreover, we introduce the adjoint surface load as

$$\dot{\sigma}(\Omega, i) := e^{0} \left[ \dot{\varepsilon} - \dot{\bar{\varepsilon}} \right] \nabla_{\Sigma} \sigma \, dS, \quad (B19)$$

where for reasons of clarity, the dependency on time $\hat{t}$ is denoted explicitly.

B5 The adjoint method

The multilinear form (B2) now takes the form

$$\int_{0}^{T} \int_{B} \left\{ \text{div} \left( \dot{\varepsilon} - \Phi_{1} \right) \right\} \cdot U + \left[ \frac{1}{4\pi G} \nabla^{2} \Phi_{1} + \text{div} (g_{0} \dot{u}) \right] \Phi_{1} - \sigma \cdot \dot{\varepsilon} \right\} dV \, dt$$

$$- \int_{0}^{T} \int_{\partial B} \left[ e_{r} \cdot \dot{\varepsilon} - g_{0} \dot{\phi} \cdot e_{r} \right] \cdot U^{*} \, dS \, dt$$

$$+ \int_{0}^{T} \int_{\partial B} \left[ \frac{1}{4\pi G} \left[ \nabla^{2} \Phi_{1} \right]^{+} - e_{r} \cdot e_{r} \right] \Phi_{1}^{+} \, dS \, dt = 0. \quad (B20)$$

Remembering that $\nabla_{\Sigma} u$ and $\nabla_{\Sigma} \phi_{1}$ are the derivatives that we wish to eliminate from $\nabla_{\Sigma} \chi^{2}$, we subtract the homogeneous eq. (B20) from eq. (23) (note the physical units of eq. (B20) are the same as $\nabla_{\Sigma} \chi^{2}$, namely kg m s$^{-2}$/[η], provided that the physical units of the adjoint variables $\hat{u}$ and $\phi_{1}$ are the same as $u$ and $\phi_{1}$):

$$\nabla_{\Sigma} \chi^{2} = \int_{0}^{T} \int_{B} \left( \Delta u^{*} \cdot U^{*} + \Delta \phi_{1} \Phi_{1}^{+} \right) dS \, dt$$

$$- \int_{0}^{T} \int_{B} \left\{ \text{div} \left( \dot{\varepsilon} - \Phi_{1} \right) \right\} \cdot U + \left[ \frac{1}{4\pi G} \nabla^{2} \Phi_{1} + \text{div} (g_{0} \dot{u}) \right] \Phi_{1} - \sigma \cdot \dot{\varepsilon} \right\} dV \, dt$$

$$+ \int_{0}^{T} \int_{\partial B} \left[ e_{r} \cdot \dot{\varepsilon} - g_{0} \dot{\phi} \cdot e_{r} \right] \cdot U^{*} \, dS \, dt - \int_{0}^{T} \int_{\partial B} \left[ \frac{1}{4\pi G} \left[ \nabla^{2} \Phi_{1} \right]^{+} - e_{r} \cdot e_{r} \right] \Phi_{1}^{+} \, dS \, dt. \quad (B21)$$
The adjoint method may now be formulated as follows. The adjoint displacement \( \hat{\boldsymbol{u}}(r, t) \) and adjoint gravitational potential increment \( \hat{\phi}_1(r, t) \) in \( B \) for \( t \in (0, T) \) is to be found such that

\[
\text{div} \, \hat{\boldsymbol{r}} - \varrho_0 \text{grad} \, \hat{\phi}_1 + \hat{\boldsymbol{f}} = 0, \tag{B22}
\]

\[
\hat{\boldsymbol{f}} = -\text{div} (\varrho_0 \hat{\boldsymbol{u}}) \, \mathbf{g}_0 + \text{grad} (\varrho_0 \hat{\boldsymbol{u}} \cdot \mathbf{g}_0).
\]

Inspecting eq. (B6), the term in the square brackets is equal to

\[
\nabla \cdot \hat{\boldsymbol{t}}
\]

\[
= \int_{\partial B} (\hat{\boldsymbol{u}} \cdot \mathbf{n}) \, \text{d}S - \int_{\partial B} (\hat{\boldsymbol{t}} \cdot \mathbf{n}) \, \text{d}S,
\]

for \( \hat{\boldsymbol{t}} \) and replace the variable

\[
\hat{\phi}_1(0, t) = 0
\]

in \( B \). Finally the gradient of misfit function \( \chi^2 \) is

\[
\nabla_{\hat{\eta}} \chi^2 = \int_0^T \int_{\partial B} \hat{\boldsymbol{e}} \, \hat{\eta} \, \text{d}V \, \text{d}t
\]

\[
= \int_0^T \int_{t'} \hat{\boldsymbol{e}}(T - t') : \nabla_{\hat{\eta}} M(t') : \boldsymbol{e}(t - t') \, \text{d}V \, \text{d}t' \, \text{d}t.
\]

(B26)

**APPENDIX C: CONVOLUTION IDENTITY**

Since the functions \( \boldsymbol{m}(t) \) and \( E(t) \) are causal, the convolution integral (12) can be rewritten through the change of variable \( \tau = t - t' \) as

\[
T(t) = \boldsymbol{m}(0) : E(t) + \int_0^t M(t - \tau) : E(\tau) \, \text{d}\tau.
\]

(C1)

Multiplying eq. (C1) by a causal strain tensor \( \hat{\boldsymbol{e}}(t_1 - t) \), that is \( \hat{\boldsymbol{e}}(t_1 - t) = \mathbf{0} \) for \( t > t_1 \), and integrating the result over the interval \((0, t_1)\) yields

\[
\int_{t=0}^{t_1} T(t) : \hat{\boldsymbol{e}}(t_1 - t) \, \text{d}t = \int_{t=0}^{t_1} \hat{\boldsymbol{e}}(t_1 - t) : \int_{t=0}^{t_1} \boldsymbol{m}(0) : E(t) \, \text{d}t + \int_{t=0}^{t_1} \hat{\boldsymbol{e}}(t_1 - t) : \int_{t=0}^{t_1} M(t - \tau) : E(\tau) \, \text{d}\tau \, \text{d}t.
\]

(C2)

We interchange the order of integration over \( t \) and \( \tau \) in the second integral on the right-hand side,

\[
\int_{t=0}^{t_1} \int_{t'=0}^{t} \hat{\boldsymbol{e}}(t_1 - t) : \int_{t=0}^{t_1} \hat{\boldsymbol{e}}(t_1 - t') : M(t - \tau) : E(\tau) \, \text{d}\tau \, \text{d}t' = \int_{t=0}^{t_1} \int_{t'=0}^{t} \int_{t=0}^{t_1} \hat{\boldsymbol{e}}(t_1 - t') : M(t') : E(\tau) \, \text{d}\tau \, \text{d}t' \, \text{d}t,
\]

and replace the variable \( t \) by a new variable \( t' \) by the substitution \( t' = t - \tau \),

\[
\int_{t=0}^{t_1} \int_{t'=0}^{t} \hat{\boldsymbol{e}}(t_1 - t) : M(t - \tau) : E(\tau) \, \text{d}\tau \, \text{d}t' = \int_{t=0}^{t_1} \int_{t'=0}^{t} \int_{t=0}^{t_1 - \tau} \hat{\boldsymbol{e}}(t_1 - t' - \tau) : M(t') : E(\tau) \, \text{d}\tau \, \text{d}t' \, \text{d}t.
\]

(C4)

Making use of the symmetry \( M_{ijkl} = M_{klji} \) (Coleman 1964), we then have

\[
\int_{t=0}^{t_1} \int_{t'=0}^{t} \hat{\boldsymbol{e}}(t_1 - t) : M(t - \tau) : E(\tau) \, \text{d}\tau \, \text{d}t' = \int_{t=0}^{t_1} \int_{t'=0}^{t} \int_{t=0}^{t_1 - \tau} \hat{\boldsymbol{e}}(t_1 - t' - \tau) : M(t') : E(\tau) \, \text{d}\tau \, \text{d}t' \, \text{d}t.
\]

(C5)

Inspecting eq. (B6), the term in the square brackets is equal to \( \hat{\boldsymbol{t}}(t_1 - t) - \boldsymbol{m}(0) : \hat{\boldsymbol{e}}(t_1 - t) \). Substituting eq. (C5) into eq. (C3), we obtain the identity

\[
\int_{t=0}^{t_1} T(t) : \hat{\boldsymbol{e}}(t_1 - t) \, \text{d}t = \int_{t=0}^{t_1} \hat{\boldsymbol{t}}(t_1 - t) : E(t) \, \text{d}t.
\]

(C6)
APPENDIX D: THE WEAK FORMULATION OF THE FSM AND ASM

We will now show that the forward-sensitivity and adjoint-state equations derived in their differential forms in Sections for section missing, respectively, can be reformulated in a weak sense. The aim is to demonstrate that the spectral-finite element method for solving the forward model equations of GIA (Martinec 2000) can be easily modified for the numerical solution of the forward-sensitivity and adjoint-state equations. The below-presented formulations have all the advantages of a weak formulation of any boundary-value problem for the second-order differential equation (e.g. Křížek & Neittaanmäki 1990), and is particularly well suitable for the numerical implementation of the finite-element technique.

The weak forms of all three methods (forward, FSM and ASM) allow the parametrization of a 3-D viscosity by the functions from the $L_2$ space, for instance, by piecewise-constant functions. This is a comfortable parametrization since one can define the viscosity structure ‘pointwise’, and the point values serve as the viscosity parameters. See Martinec (2000, eq. 98), or eq. (48) in the text.

D1 The weak formulation of the FSM

The initial, boundary-value problem (1)–(6) for the forward method of GIA has been formulated in a weak sense by Martinec (2000). Comparing eqs (1)–(6) with eqs (14), (16) and (17)–(20) for the FSM, we can find two differences between the two formulations. (i) The right-hand side of the linear momentum equation (14) contains an additional volume source term $T_{\Omega} \cdot e$, which occurs on the right-hand side of the boundary condition (16) for the surface traction. (ii) When solving the forward method of GIA, the coastline function $O(\Omega, t)$, see eq. (A14), and associated ocean function $O(\Omega, t)$, see eq. (A15), must be calculated at each time step by solving eqs (A1)–(A3) and (A15). As eq. (19) shows, this step is not necessary when solving the forward sensitivity equations since the coastline and ocean functions are those determined by solving the forward method. Moreover, instead of applying the combined ice and ocean-water load in the forward method, in the FSM we apply the load $\sigma^{FS} \nabla s$ over the ocean area defined by the forward-modelled ocean function $O(\Omega, t)$. We will follow the weak formulation for the forward method of GIA by Martinec (2000) and modify it, with respect to the two differences, for the FSM.

Following the considerations by Martinec (2000), for functions ($U, \Phi_1, \Pi$) from an appropriate functional space $V_{sol}$, whose detailed specification is given by eq. (27) (Martinec 2000), let us define the energy functional $E$ by summing the term associated with the pressure, $E_{press}$, elastic shear energy $E_{shear}$, gravity energy $E_{gravity}$ and the term associated with the uniqueness of a solution, $E_{uniq}$, as

$$E(U, \Phi_1, \Pi) := E_{press}(U, \Pi) + E_{shear}(U) + E_{gravity}(U, \Phi_1) + E_{uniq}(U),$$

where the energy constituents are defined by eqs (30)–(33) (Martinec 2000). Moreover, we introduce the linear functional $F^{+1}$ as the sum of the dissipative term taken at time $t$, the term associated with the boundary conditions (16) taken at time $t^{+1}$, and the term associated with the volume source term on the right-hand side of the linear momentum equation (14) taken at time $t$,

$$F^{+1}(U, \Phi_1) := F_{diss}(U) + F_{surf}(U, \Phi_1) + F_{\Omega}(U),$$

where the first two terms of the right-hand side are given by eqs (35) and (36) (Martinec 2000). The surface load $\sigma$ needed to be specified for evaluating $F_{surf}$ is now given by the load $\sigma^{FS} = \sigma^{FS} \nabla s$. Inspecting eq. (19), we can see that $\nabla s$ depends on $U \cdot e$, $\Phi_1$ and $O$, where $O$ is only dependent on the forward-modelled deformations variables $u \cdot e$, and $\phi_1$, that is, $O$ is independent of the forward-sensitivity fields $U \cdot e$ and $\Phi_1$. This means that the surface load $\sigma^{FS}$ of the FSM can be applied in a similar way as the sea level equation in the forward method for a particular case where the ocean function is prescribed a priori. This is, for instance, the case of a fixed-coastline geometry (Spada et al. 2012). Explicitly, the surface load for the FSM at the current time $t^{+1}$ is

$$\sigma^{FS}(U) = \sigma^{FS}\left[\frac{E(\Omega, t') - U(\Omega, t') - (E - U)(t')}{{\phi_1}(O, t^{+1})}\right].$$

As far as the last term on the right-hand side of eq. (D2) is concerned, it has the form

$$F_{\Omega}(U) = - \int_{\Omega} (T_{\Omega} \cdot E) \, dV.$$

The weak formulation of the forward sensitivity equations consists of finding fields ($U, \Phi_1, \Pi$) from the functional space $V_{sol}$ that fulfill the homogeneous initial conditions and, at a fixed time $t^{+1}$, satisfy the following variational equality,

$$\delta E(U^{+1}, \Phi_1^{+1}, \Pi^{+1}, \delta U, \delta \Phi_1, \delta \Pi) = \delta F^{+1}(\delta U, \delta \Phi_1) \quad \forall (\delta U, \delta \Phi_1, \delta \Pi) \in V_{sol},$$

where the variation of the energy constituents are given by eqs (40)–(43) (Martinec 2000). The variation of the linear function $F^{+1}$ is

$$\delta F^{+1}(\delta U, \delta \Phi_1) = \delta F_{diss}(\delta U) + \delta F_{surf}(\delta U, \delta \Phi_1) + \delta F_{\Omega}(\delta U),$$

where the first two terms on the right-hand side are given by eqs (45) and (46) (Martinec 2000) and the last term is

$$\delta F_{\Omega}(\delta U) = - \int_{\Omega} (T_{\Omega} \cdot \delta E) \, dV$$

with $\delta E := (\text{grad} \delta U + \text{grad} E \delta U)/2$. 
To show that the weak solution generalizes the differential solution to the problem expressed by eqs (14)–(17), let us temporarily assume that the weak solution \((U, \Phi_1, \Pi)\) is sufficiently smooth such that Green’s theorem (49) (Martinec 2000) can be applied to the integral on the right-hand side of eq. (D7). This yields
\[
\delta F_{0}^{*}(\delta U) = \int_{B} \left( \text{div} \mathbf{T}_{0}^{-} \cdot \delta U \right) dV - \int_{\partial B} \left( \mathbf{e}_{r} \cdot \mathbf{T}_{0}^{-} \cdot \delta U \right) dS + \int_{\Sigma} \left( \mathbf{e}_{r} \cdot \mathbf{T}_{0}^{-} \right) \cdot \delta U dS. \tag{D8}
\]

We modify the variational equality (52) (Martinec 2000) by (i) adding the term \(\delta F_{1}^{*}(\delta U)\) on the right-hand side and (ii) putting \(\sigma = \varphi_{0} \nabla_{\xi} s\) in the term \(\delta F_{1}^{*+}(\delta U, \delta \Phi_{1})\). Then, the implication (53) (Martinec 2000) applied to the modified variational equality provides the constitutive equation (11) and (12), the divergence-free constraint (E16) for \(U\), the linear momentum balance, Poisson’s equation in (14) and the interface conditions (16).

**D2 The weak formulation of the ASM**

We will now demonstrate that the ASM derived in a strong differential form can be reformulated in a weak sense. Comparing the boundary conditions (4) of the forward method with boundary conditions (26) of the ASM, we can find two differences between them. (i) The surface load \(\dot{\sigma}_{1}\) for the normal component of the adjoint surface traction differs from the surface adjoint load \(\hat{\sigma}_{1}\) for the adjoint gravitational-potential increment, and (ii) there is a non-zero surface adjoint load \(\hat{\sigma}_{2}\) for the tangential components of the adjoint surface traction. We will follow the weak formulation of the GIA forward method by Martinec (2000) and modify it with respect to these two differences for the ASM.

The weak formulation of the initial, boundary-value problem (25)–(31) consists of finding fields \((\hat{u}, \hat{\sigma}, \hat{\Phi})\) from the functional space \(V_{\text{sol}}\), whose detailed specification is given by eq. (27) (Martinec 2000), that fulfill homogeneous initial conditions and, at a fixed time \(t^{+}\), satisfy the following variational equality,
\[
\delta E \left( \hat{u}_{t^{+}}, \hat{\sigma}_{t^{+}}, \hat{\Phi}_{t^{+}}, \delta \hat{u}, \delta \hat{\sigma}, \delta \hat{\Phi} \right) = \delta F_{t^{+}}^{*+} \left( \delta \hat{u}, \delta \hat{\sigma}, \delta \hat{\Phi} \right) \quad \forall \left( \delta \hat{u}, \delta \hat{\sigma}, \delta \hat{\Phi} \right) \in V_{\text{sol}}, \tag{D9}
\]

where \(\delta E\) corresponding to the energy constituents and the linear functional are given by eqs (40)–(43) and (44)–(46) (Martinec 2000), respectively. The surface-load variables \(b_{0}\) and \(b_{1}\), defined for the forward method by eqs (37)–(38) (Martinec 2000), need to be redefined for the ASM as follows:
\[
b_{0} := - \left( \varphi_{0} \dot{\sigma} + \dot{\sigma}_{1} \right) \mathbf{e}_{r} + \dot{\sigma}_{2}, \\
b_{1} := \frac{1}{4\pi G} \left( \text{grad} \hat{\phi}^{+} \cdot \mathbf{e}_{r} \right) - \left( \dot{\sigma} + \dot{\sigma}_{1} \right), \tag{D10}
\]

where the adjoint loads \(\dot{\sigma}\) and \(\dot{\sigma}_{1}\) are given by eqs (28) and (29), respectively, and \(\hat{\phi}^{+}\) is the gravitational potential increment on the exterior side of \(\partial B\).

**APPENDIX E: RHEOLOGICAL MODEL**

A number of assumptions on the rheological model of the Earth are introduced for the modelling of GIA. We summarize them in this appendix.

First, the earth model \(B\) is assumed to be constituted of a linear viscoelastic material. For such a rheology, the stress at the current time \(t\) is assumed to depend upon the deformation history and viscoelastic material property at all past epochs through the convolution integral of the Stieltjes type (Gurtin & Sternberg 1962),
\[
\tau(t) = \int_{-\infty}^{t} \varepsilon(t - t') : \mathbf{m}(t') dt', \tag{E1}
\]

where \(\varepsilon\) is the strain tensor, that is the symmetric part of \(\text{grad} \mathbf{u}\), that is \(\varepsilon = \left( \text{grad} \mathbf{u} + \text{grad} \mathbf{u}^{\top} / 2 \right)\), \(\mathbf{m}\) is the ‘viscoelastic relaxation tensor’, and the symbol ‘:’ denotes the double-dot product of second-order tensors. The Stieltjes integral can be transformed to the Riemann integral, obtaining
\[
\tau(t) = \mathbf{m}(0) : \varepsilon(t) + \int_{0}^{t} \mathbf{M}(t') : \varepsilon(t - t') dt', \tag{E2}
\]

where \(\mathbf{M}(t)\) is the ‘specific relaxation tensor’ (e.g. Marques & Creus 2012), which is given by the time derivative of the relaxation tensor \(\mathbf{m}(t)\),
\[
\mathbf{M}(t) := \frac{\text{d} \mathbf{m}(t)}{\text{d}t}. \tag{E3}
\]

The constitutive equation (E2) is composed of the elastic and viscous stress tensors, respectively,
\[
\tau(t) = \tau^{E}(t) + \tau^{V}(t), \tag{E4}
\]
where
\[ \tau^E(t) = m(0) : \varepsilon(t), \]
\[ \tau^V(t) = \int_0^t M(t') : \varepsilon(t - t')dt'. \] (E5)

Second, the viscoelastic material of model B is assumed to be isotropic. The viscoelastic relaxation tensor \( m \) is then expressed in terms of two fourth-order isotropic tensors \( I_1 \) and \( I_2 \) in the form (Cathles 1975; Wolf 2003)
\[ m(t) = m_1(t)I_1 + m_2(t)I_2, \] (E6)
where \( m_1(t) \) and \( m_2(t) \) are the first and second (or, shear) relaxation functions, respectively, and the components of \( I_1 \) and \( I_2 \) are
\[ (I_1)_{ijkl} = \delta_{ij}\delta_{kl}, \quad (I_2)_{ijkl} = \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}. \] (E7)

Third, the time-relaxation of normal stresses are not considered. Then, the first relaxation function \( m_1 \) is constant in time and is equal to the first Lamé parameter \( \lambda \),
\[ m_1(t) = \lambda, \] (E8)
whereas the shear relaxation function \( m_2 \) is of the form
\[ m_2(t) = \mu e^{-t/\tau_{\text{relax}}}, \] (E9)
where \( \mu \) is the elastic rigidity,
\[ \tau_{\text{relax}} = \eta/\mu \] (E10)
is the Maxwell relaxation time and \( \eta \) is the shear viscosity. The viscoelastic relaxation tensor is then of the form
\[ m(t) = \lambda I_1 + m_2(t)I_2. \] (E11)
In particular,
\[ m(0) = \lambda I_1 + \mu I_2. \] (E12)
Substituting eq. (E11) into eq. (E3) and differentiating eq. (E9) with respect to time provides the relation between the specific relaxation tensor \( M \) and shear relaxation function \( m_2 \),
\[ M(t) = -\frac{m_2(t)}{\tau_{\text{relax}}} I_2. \] (E13)
By eqs (E12) and (E13), the elastic and viscous stress tensors, see eq. (E5), are expressed as,
\[ \tau^E(t) = \lambda \left( I_1 : \varepsilon(t) \right) + \mu \left( I_2 : \varepsilon(t) \right) \]
\[ = \lambda \text{div} u(t) I + 2\mu \varepsilon(t), \] (E14)
where \( I \) is the second-order identity tensor, and
\[ \tau^V(t) = -\frac{2}{\tau_{\text{relax}}} \int_0^t m_2(t')\varepsilon(t - t')dt', \] (E15)
respectively.

Fourth, if an incompressible material is additionally assumed, meaning when the divergence-free constraint of the displacement vector is required, that is
\[ \text{div} u = 0, \] (E16)
then \( \lambda \to \infty \), but the product \( \lambda \text{div} u \) in the elastic stress tensor remains bounded and is replaced by the perturbation \( \Pi \) of mechanical pressure,
\[ \tau^E(t) = \Pi(t)I + 2\mu \varepsilon(t). \] (E17)
The constitutive equation (E15) for the viscous stress tensor does not change in this case.