Geophysical Journal International

Geophys. J. Int. (2014) **196,** 34–77
doi: 10.1093/gji/ggt395
Advance Access publication 2013 October 22

Sensitivity kernels for viscoelastic loading based on adjoint methods

David Al-Attar[1] and Jeroen Tromp[2,3]

[1]*Bullard Laboratories, Department of Earth Sciences, University of Cambridge, Madingley Road, Cambridge,* CB3 OEZ, *UK. E-mail: da380@cam.ac.uk*
[2]*Department of Geosciences, Princeton University, Princeton,* NJ 08544, *USA*
[3]*Program in Applied & Computational Mathematics, Princeton University, Princeton,* NJ 08544, *USA*

Accepted 2013 September 26. Received 2013 September 26; in original form 2013 May 27

SUMMARY

Observations of glacial isostatic adjustment (GIA) allow for inferences to be made about mantle viscosity, ice sheet history and other related parameters. Typically, this inverse problem can be formulated as minimizing the misfit between the given observations and a corresponding set of synthetic data. When the number of parameters is large, solution of such optimization problems can be computationally challenging. A practical, albeit non-ideal, solution is to use gradient-based optimization. Although the gradient of the misfit required in such methods could be calculated approximately using finite differences, the necessary computation time grows linearly with the number of model parameters, and so this is often infeasible. A far better approach is to apply the 'adjoint method', which allows the exact gradient to be calculated from a single solution of the forward problem, along with one solution of the associated adjoint problem. As a first step towards applying the adjoint method to the GIA inverse problem, we consider its application to a simpler viscoelastic loading problem in which gravitationally self-consistent ocean loading is neglected. The earth model considered is non-rotating, self-gravitating, compressible, hydrostatically pre-stressed, laterally heterogeneous and possesses a Maxwell solid rheology. We determine adjoint equations and Fréchet kernels for this problem based on a Lagrange multiplier method. Given an objective functional J defined in terms of the surface deformation fields, we show that its first-order perturbation can be written $\delta J = \int_{M_S} K_\eta \delta \ln \eta \, \mathrm{d}V + \int_{t_0}^{t_1} \int_{\partial M} K_{\dot\sigma} \delta\dot\sigma \, \mathrm{d}S \, \mathrm{d}t$, where $\delta \ln \eta = \delta\eta/\eta$ denotes relative viscosity variations in solid regions M_S, $\mathrm{d}V$ is the volume element, $\delta\dot\sigma$ is the perturbation to the time derivative of the surface load which is defined on the earth model's surface ∂M and for times $[t_0, t_1]$ and $\mathrm{d}S$ is the surface element on ∂M. The 'viscosity kernel' K_η determines the linearized sensitivity of J to viscosity perturbations defined with respect to a laterally heterogeneous reference earth model, while the 'rate-of-loading kernel' $K_{\dot\sigma}$ determines the sensitivity to variations in the time derivative of the surface load. By restricting attention to spherically symmetric viscosity perturbations, we also obtain a 'radial viscosity kernel' $\overline{K}_\eta$ such that the associated contribution to δJ can be written $\int_{I_S} \overline{K}_\eta \delta \ln \eta \, \mathrm{d}r$, where I_S denotes the subset of radii lying in solid regions. In order to illustrate this theory, we describe its numerical implementation in the case of a spherically symmetric earth model using a 1-D spectral element method, and calculate sensitivity kernels for a range of realistic observables.

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

1 INTRODUCTION

Observations of glacial isostatic adjustment (GIA) allow inferences to be made about a range of parameters, including mantle viscosity, lithospheric thickness, past ice sheet thickness and other forms of surface loading, such as lakes and sedimentation (see Peltier 2004 for a review). Typically, these GIA inverse problems are formulated in terms of minimizing the misfit between a set of GIA observations and synthetic data calculated using an appropriate set of model parameters.

One approach to solve such problems is to employ the so-called gradient-based optimization methods in which the gradient of the misfit with respect to the model parameters is used to iteratively update the model in such a way that the misfit is lowered (e.g. Nocedal & Wright 2006). Although gradient-based methods certainly have shortcomings (e.g. convergence to local minima), they can provide a computationally feasible solution to large-scale optimization problems, particularly, when the calculation of synthetic data is computationally demanding. The simplest method for computing the required gradients of the misfit are finite difference schemes based upon a given parametrization of the model (see e.g. Mitrovica & Peltier 1991 and Peltier 2004 for the application of such techniques to the GIA problem). Unfortunately, the computational cost of such finite difference derivatives scales like the number of model parameters, and so becomes impractical if there are many model parameters and if solution of the forward problem is time consuming (e.g. Fichtner *et al.* 2006; Tape *et al.* 2007). For the GIA problem, this issue will be most acutely seen if calculations are performed in laterally heterogeneous earth models that require large-scale numerical solutions (e.g. Zhong *et al.* 2003; Latychev *et al.* 2005).

A better approach to calculating the required gradients is the 'adjoint method', which arose within the study of optimal control problems (e.g. Lions 1970; Tröltzsch 2005) and has been successfully applied in a number of geophysical applications, including seismic tomography (e.g. Tromp *et al.* 2005; Liu & Tromp 2008; Tape *et al.* 2009; Fichtner *et al.* 2010; Zhu *et al.* 2012). Using this method, calculation of the exact gradient requires the forward problem to be solved once, along with one solution of the associated 'adjoint problem'. The computational cost of the adjoint method is, therefore, independent of the number of model parameters, and so provides a tremendous computational saving when compared to finite difference calculations. There are, moreover, extensions of the adjoint method which allow for the calculation of higher order derivatives of the misfit with respect to the model parameters, and such terms can be usefully employed in the quantification of the obtainable model resolution (Fichtner & Trampert 2011).

Accurate modelling of GIA requires the solution of a complex system of equations incorporating the viscoelastic relaxation of the mantle, the response of the fluid outer core, gravitationally self-consistent ocean loading and rotational feedbacks (e.g. Farrel & Clark 1976; Mitrovica & Milne 2003; Peltier 2004). In order to simplify the presentation of the theory, we have decided to break the problem into smaller parts, and within this paper we shall neglect the effects of ocean loading and the Earth's rotation. The resulting 'viscoelastic loading problem' considered here will, however, form the basis for later work in which these additional features are incorporated. We note that the present theory can also with minor modification be applied to the study of post-seismic deformation.

The aim of this paper is to present an application of the adjoint method to the viscoelastic loading problem, and so to provide a computationally efficient approach for calculating the linearized sensitivity of observations to both mantle viscosity and ice load history. Extension of the method to include other model parameters, such as lithospheric thickness, is possible, but has yet to be developed. See Liu & Tromp (2008) for the application of adjoint methods to an analogous seismological problem incorporating boundary perturbations. In particular, we consider the surface loading of a laterally heterogeneous earth model, and our theory could readily be implemented using a new generation of fully numerical codes (e.g. Zhong *et al.* 2003; Latychev *et al.* 2005).

In our applications of the theory, we have, however, performed calculations only in spherically symmetric earth models. This has been done only to simplify the necessary calculations, and the results presented are hopefully sufficient to illustrate the potential of the method. The advantages of the adjoint method over other techniques for computing these gradients are very substantial, and do not depend in any way upon the numerical methods used. Numerical calculations are included in the paper only to illustrate the basic ideas, and so we do not feel it necessary nor appropriate to give details about computation times, etc. Instead, the main contribution of this paper is the derivation of the adjoint problem along with the associated sensitivity kernels. We again stress that this theory is valid in laterally heterogeneous earth models, and it should be quite simple to implement given access to a suitable code.

1.1 Background on gradient-based optimization

To better motivate our use of the adjoint method, it will be helpful to consider some basic aspects of gradient-based optimization. Let us write $\mathbf{m}$ for an element of the appropriate model space in the viscoelastic loading problem, which we assume to be a normed vector space. In practice, this model vector will comprise a pair of functions $\{\sigma, \eta\}$ with σ the time-dependent surface load, and η the mantle viscosity. Associated with a given value of $\mathbf{m}$, we can solve the forward viscoelastic loading problem, and so calculate the misfit between our synthetic predictions and the observed data. Writing $\hat{J}$ for the misfit function, we can then consider the scalar-valued mapping

$$\mathbf{m} \mapsto \hat{J}(\mathbf{m}), \tag{1.1}$$

which associates a real number to each possible model vector.

We seek the model vector $\tilde{\mathbf{m}}$, say, which minimizes $\hat{J}$. As is well known, a necessary condition for $\tilde{\mathbf{m}}$ to be such a minima is that the derivative of $\hat{J}$ vanishes here. Of course, this condition is not sufficient, and such an $\tilde{\mathbf{m}}$ could, for example, be simply a local minima of $\hat{J}$. We denote the value of this derivative at a point $\mathbf{m}$ by $D\hat{J}(\mathbf{m})$, and recall that it is a linear functional on the model space defined through the relation

$$\hat{J}(\mathbf{m} + \delta\mathbf{m}) = \hat{J}(\mathbf{m}) + \langle D\hat{J}(\mathbf{m}), \delta\mathbf{m}\rangle + O(\|\delta\mathbf{m}\|^2), \tag{1.2}$$

where $\delta\mathbf{m}$ is an arbitrary element of the model space, we have written $\langle D\hat{J}(\mathbf{m}), \delta\mathbf{m}\rangle$ for the action of the derivative on the given model perturbation, and $O(\|\delta\mathbf{m}\|^2)$ denotes a term that is of second order in the magnitude of $\delta\mathbf{m}$ (see e.g. Parker 1994 for further details). If the model space is finite dimensional, the above definition reduces to that of the gradient in multivariable calculus. Importantly, however, this definition also applies when $\mathbf{m}$ is an element of a normed function space, and this additional generality is required for the problem at hand. If we assume that the model space is in fact a Hilbert space (e.g. a space of square integrable functions), we can identify the derivative $D\hat{J}(\mathbf{m})$ with an element of the model space, and interpret $\langle \cdot\,, \cdot \rangle$ as an inner product.

With these notations, the necessary condition for $\tilde{\mathbf{m}}$ to be a local minimum of $\hat{J}$ can be written

$$D\hat{J}(\tilde{\mathbf{m}}) = \mathbf{0}, \tag{1.3}$$

where $\mathbf{0}$ denotes the zero vector in the model space. A practical approach for obtaining such an $\tilde{\mathbf{m}}$ is through the use of gradient-based optimization. The simplest method is the 'steepest descent algorithm', which iteratively constructs a sequence of model vectors $\mathbf{m}_0$, $\mathbf{m}_1$, $\mathbf{m}_2 \ldots$ through

$$\mathbf{m}_{i+1} = \mathbf{m}_i - \alpha_i D\hat{J}(\mathbf{m}_i), \tag{1.4}$$

with the step length $\alpha_i \geq 0$ obtained by minimizing

$$\alpha \mapsto \hat{J}[\mathbf{m}_i - \alpha D\hat{J}(\mathbf{m}_i)]. \tag{1.5}$$

For this algorithm to make sense, the derivatives $D\hat{J}(\mathbf{m}_i)$ must, of course, be identified with elements of the model space. We will see below that this condition need not hold for all choices of model parameters in the viscoelastic loading problem. Under mild assumptions on $\hat{J}$, and given enough iterations, the steepest descent algorithm will converge to a local minimum of the misfit function (e.g. convexity of J is sufficient). There are, of course, a number of similar but more sophisticated iterative schemes that possess better convergence properties (e.g. Nocedal & Wright 2006). Nonetheless, all such gradient-based methods require repeated evaluations of the derivative of $\hat{J}$ with respect to the model parameters.

The simplest way to compute these derivatives is to use a finite difference approximation. For example, to first order in $\|\delta\mathbf{m}\|$ we have

$$\langle D\hat{J}(\mathbf{m}), \delta\mathbf{m}\rangle \approx \frac{\hat{J}(\mathbf{m} + h\,\delta\mathbf{m}) - \hat{J}(\mathbf{m})}{h}, \tag{1.6}$$

with h a small number. Letting $\{\mathbf{m}^1, \ldots, \mathbf{m}^n\}$ be basis vectors in a suitable finite-dimensional subparametrization of the model space, we can use the above finite difference formula to approximate the n directional derivatives $\langle D\hat{J}(\mathbf{m}), \mathbf{m}^i\rangle$ for $i = 1, \ldots, n$, and so estimate the derivative of $\hat{J}$ defined with respect to the given model parametrization. Given a method for solving the forward problem, the advantage of such finite difference calculations is the ease of implementation. However, each calculation requires $n + 1$ evaluations of the misfit function, and so the forward viscoelastic loading problem must be solved $n + 1$ times. If the number of parameters is large, and if the computational cost of solving the forward problem is even moderately high, such finite difference approximations cannot be practically useful.

To illustrate this point, let us suppose that the lateral viscosity structure of the earth model's mantle is parametrized using spherical harmonics up to degree L, and with n_d radial basis functions. A simple calculation shows that this requires a total of

$$n = (L + 1)^2 n_d, \tag{1.7}$$

parameters to specify the viscosity structure. To use plausible numbers, we take as an example the shear velocity model S20RTS (Ritsema *et al.* 1999), which has $L = 20$, $n_d = 21$ and so would lead to a total of 9261 viscosity parameters. From the above discussion, we know that calculating the finite difference derivatives of the misfit functional with respect to this viscosity parametrization would require 9262 solutions of the forward problem. Furthermore, these derivatives would need to be calculated at each stage of the iterative scheme, and perhaps multiple times per iteration depending on the algorithm used. In laterally heterogeneous earth models, however, the viscoelastic loading problem must be solved using fully numerical methods (e.g. Zhong *et al.* 2003; Latychev *et al.* 2005), and each such calculation will require quite substantial time (details depending, of course, upon available resources). Clearly, in such a case, determining gradients using finite differences is simply not an option.

We note that a variant of the above finite difference approach is to calculate the directional derivatives $\langle D\hat{J}(\mathbf{m}), \mathbf{m}^i\rangle$ exactly using first-order perturbation theory. Although more accurate than finite difference approximations, this method still requires n solutions of the perturbed forward problem to obtain the derivative with respect to an n-dimensional model parametrization, and so is subject to the same computational limitations.

In order to apply gradient-based algorithms to such large-scale optimization problems, the only feasible way to calculate the required derivatives is the adjoint method. As we will see below, with the adjoint method we can calculate the exact derivative $D\hat{J}(\mathbf{m})$ from one solution of the forward problem, and one solution of the associated adjoint problem. Importantly, adjoint problems are closely related to the linearization of the forward problem, and so are typically no harder to solve (e.g. Tröltzsch 2005). In fact, for the viscoelastic loading problem considered in this paper, the adjoint problem is shown to be essentially identical in form to the forward problem, and so can be solved using the same numerical method. Furthermore, when using the adjoint method we do not simply determine the components of the derivative along the pre-defined basis vectors, but instead obtain the entire derivative $D\hat{J}(\mathbf{m})$ (in practice, of course, the viscoelastic loading problem will be solved numerically using a suitable discretization, and the derivative obtained will be discretized accordingly). The derivative $D\hat{J}(\mathbf{m})$ can then be trivially projected onto the basis vectors of any desired model parametrization, or alternatively, could be used directly to update the model.

To clearly see the great practical advantage of the adjoint method, we return to our example based on an S20RTS-type parametrization for mantle viscosity. We have seen that using finite differences it would take 9262 solutions of the forward problem to determine an approximate derivative of the misfit. Using the adjoint method, however, this derivative could be obtained exactly at a cost equivalent to only two solutions of the forward problem—an increase in efficiency by a factor of over 4600. For inverse problems involving computationally expensive forward calculations, such increases in efficiency are vital, and can make the difference between a problem being insurmountable or quite routine. To conclude, we repeat that in the context of gradient-based optimization, adjoint methods are often the only viable way to compute the required derivatives. In fact, finite difference schemes are equally efficient only in the uninteresting case of a single-model parameter. Importantly, these conclusions are not dependent in any way on the particular method used to perform the forward calculations.

2 THE FORWARD VISCOELASTIC LOADING PROBLEM

2.1 Equations of motion

We begin by recalling the equations of motion for the quasi-static deformation of a non-rotating, self-gravitating, hydrostatically pre-stressed, compressible and laterally heterogeneous earth model following Dahlen (1974), Dahlen & Tromp (1998) and Tromp & Mitrovica (1999a). Let the earth model occupy a bounded open volume $M \subseteq \mathbb{R}^3$ with smooth external boundary ∂M. Points in $\mathbb{R}^3$ will be denoted by their position vectors $\mathbf{x}$ relative to a fixed Cartesian coordinate system which, for definiteness, has its origin located at the centre of mass of M. The components of all vector and tensor fields will also be specified relative to this coordinate system. We assume that the earth model is further subdivided into a number of solid and fluid subregions that are separated by smooth, non-intersecting, closed surfaces called internal boundaries. The union of the solid regions will be denoted M_{S} and that of all fluid regions M_{F}. We write Σ for the union of all internal and external boundaries, with the internal boundaries being split into four subsets Σ_{SS}, Σ_{SF}, Σ_{FS} and Σ_{FF}, where here the first subscript denotes whether the region on the inside of the boundary is solid (S) or fluid (F), while the second subscript specifies whether the region on the outside of boundary is solid or fluid.

Prior to deformation we assume that the earth model is in hydrostatic equilibrium, this condition being expressed by the equation

$$\nabla p + \rho \nabla \Phi = \mathbf{0}, \tag{2.1}$$

for $\mathbf{x} \in M$, where p denotes the equilibrium pressure, ρ is the density of the earth model and Φ is the gravitational potential. The pressure is subject to the boundary condition

$$[p]_-^+ = 0, \tag{2.2}$$

for $\mathbf{x} \in \Sigma$, where $[\cdot]_-^+$ denotes the jump in a quantity on crossing a boundary in the direction of the outward unit normal vector $\hat{\mathbf{n}}$, and it is understood that $p = 0$ outside of M. The gravitational potential Φ is itself a solution of Poisson's equation

$$(4\pi G)^{-1} \nabla^2 \Phi = \rho, \tag{2.3}$$

for $\mathbf{x} \in \mathbb{R}^3$, where G is the gravitational constant, subject to the boundary conditions

$$[\Phi]_-^+ = 0, \quad [\hat{\mathbf{n}} \cdot \nabla \Phi]_-^+ = 0, \tag{2.4}$$

for $\mathbf{x} \in \Sigma$, along with the condition that $\Phi \to 0$ as $\|\mathbf{x}\| \to \infty$. In a non-rotating hydrostatic earth model, it is well known (e.g. Dahlen & Tromp 1998) that the level surfaces of ρ, p and Φ are concentric spheres. In particular, this condition requires that ρ, p and Φ are constants on either side of internal and external boundaries, and that on such boundaries

$$\nabla \Phi = g \hat{\mathbf{n}}. \tag{2.5}$$

This assumption of hydrostatic pre-stress cannot be strictly valid if the earth model possesses lateral variations in its density structure or if there is a spherical topography on any internal or external boundaries. It is, nonetheless, expected that the departure from a hydrostatic pre-stress field due to realistic lateral variations will be small, and we shall simply neglect additional terms in the equations of motion associated with deviatoric pre-stress (e.g. Dahlen & Tromp 1998, section 3.11.1). The above discussion carries over essentially unchanged to a rotating earth model, except in this case the level surfaces of the geopotential, density and pressure are ellipsoids determined through the solution of Clairaut's equation (e.g. Dahlen & Tromp 1998, section 14.1.1).

We now consider the response of such an earth model to a time-dependent surface load $\sigma(\mathbf{x}, t)$. The resulting deformation is described by the displacement field $\mathbf{u}(\mathbf{x}, t)$ and perturbed gravitational potential $\phi(\mathbf{x}, t)$; for convenience, in what follows, we shall refer to the pair $\{\mathbf{u}, \phi\}$ as being the 'deformation field' of the earth model. As shown by Dahlen (1974), the displacement field $\mathbf{u}$ is only well defined in solid portions of the earth model, while in fluid regions the deformation can be completely characterized by the perturbed gravitational potential. Prior to the onset of loading at time $t = t_0$, we assume that the earth model is at rest, which gives the initial conditions

$$\mathbf{u}(\mathbf{x}, t_0) = \mathbf{0}, \quad \phi(\mathbf{x}, t_0) = 0, \tag{2.6}$$

for all $\mathbf{x} \in M_{\mathrm{S}}$ and $\mathbf{x} \in \mathbb{R}^3$, respectively. The subsequent evolution of $\mathbf{u}$ and ϕ is governed by the quasi-static momentum equation (e.g. Dahlen 1974; Tromp & Mitrovica 1999)

$$-\nabla \cdot \mathbf{T} + \nabla(\rho \mathbf{u} \cdot \nabla \Phi) - \nabla \cdot (\rho \mathbf{u}) \nabla \Phi + \rho \nabla \phi = \mathbf{0}, \tag{2.7}$$

for $\mathbf{x}$ in the solid earth M_{S}, where $\mathbf{T}$ denotes the incremental Lagrangian–Cauchy stress tensor, and the 'hydrostatic Poisson equation'

$$(4\pi G)^{-1}\nabla^2\phi = \begin{cases} -\nabla\cdot(\rho\mathbf{u}) & \mathbf{x}\in M_{\mathrm{S}} \\ g^{-1}\phi\partial_n\rho & \mathbf{x}\in M_{\mathrm{F}} \\ 0 & \mathbf{x}\in\mathbb{R}^3\setminus M^{\mathrm{cl}} \end{cases}, \tag{2.8}$$

where ∂_n is the directional derivative along the outward normal to level surfaces of ρ, and M^{cl} denotes the closure of M. Here, we have made use of the identities

$$p_1 = -\rho\phi, \quad \rho_1 = g^{-1}\phi\partial_n\rho, \quad \mathbf{x}\in M_{\mathrm{F}}, \tag{2.9}$$

due to Dahlen (1974) and Crossley & Gubbins (1975), which follow from the hydrostatic condition in fluid regions and the assumption that the mass of the core remains constant during deformation. The boundary conditions for the problem are

$$\hat{\mathbf{n}}\cdot\mathbf{T} = -\sigma\nabla\Phi, \quad \mathbf{x}\in\partial M, \tag{2.10}$$

$$[\hat{\mathbf{n}}\cdot\mathbf{T}]_-^+ = \mathbf{0}, \quad \mathbf{x}\in\Sigma_{\mathrm{SS}}, \tag{2.11}$$

$$\hat{\mathbf{n}}\cdot\mathbf{T}^+ = \rho^-[\mathbf{u}^+\cdot\nabla\Phi + \phi]\hat{\mathbf{n}}, \quad \mathbf{x}\in\Sigma_{\mathrm{FS}}, \tag{2.12}$$

$$\hat{\mathbf{n}}\cdot\mathbf{T}^- = \rho^+[\mathbf{u}^-\cdot\nabla\Phi + \phi]\hat{\mathbf{n}}, \quad \mathbf{x}\in\Sigma_{\mathrm{SF}}, \tag{2.13}$$

$$[\mathbf{u}]_-^+ = \mathbf{0}, \quad \mathbf{x}\in\Sigma_{\mathrm{SS}}, \tag{2.14}$$

$$[\phi]_-^+ = 0, \quad \mathbf{x}\in\Sigma, \tag{2.15}$$

$$\left[(4\pi G)^{-1}\hat{\mathbf{n}}\cdot\nabla\phi\right]_-^+ - \rho^-\hat{\mathbf{n}}\cdot\mathbf{u}^- = \sigma, \quad \mathbf{x}\in\partial M, \tag{2.16}$$

$$\left[(4\pi G)^{-1}\hat{\mathbf{n}}\cdot\nabla\phi + \rho\hat{\mathbf{n}}\cdot\mathbf{u}\right]_-^+ = 0, \quad \mathbf{x}\in\Sigma_{\mathrm{SS}}, \tag{2.17}$$

$$\left[(4\pi G)^{-1}\hat{\mathbf{n}}\cdot\nabla\phi\right]_-^+ + [\rho]_-^+\,\hat{\mathbf{n}}\cdot\mathbf{u}^+ = 0, \quad \mathbf{x}\in\Sigma_{\mathrm{FS}}, \tag{2.18}$$

$$\left[(4\pi G)^{-1}\hat{\mathbf{n}}\cdot\nabla\phi\right]_-^+ + [\rho]_-^+\,\hat{\mathbf{n}}\cdot\mathbf{u}^- = 0, \quad \mathbf{x}\in\Sigma_{\mathrm{SF}}, \tag{2.19}$$

where a superscript $\pm$ denotes whether a term is evaluated on the upper $(+)$ or lower $(-)$ side of a discontinuity, and we have the decay condition at infinity

$$\lim_{\|\mathbf{x}\|\to\infty}\phi = 0. \tag{2.20}$$

2.2 Linear viscoelasticity with internal variables

2.2.1 Hooke's Law and Boltzmann's superposition principle

To complete the specification of the loading problem, we require a constitutive equation which relates the stress tensor to the deformation. It will be useful to first recall the case of an isotropic linear elastic solid for which this relationship is given by

$$\mathbf{T}(t) = \kappa\nabla\cdot\mathbf{u}(t)\,\mathbf{I} + 2\mu\mathbf{d}(t), \tag{2.21}$$

where for clarity we have suppressed the spatial dependence of the various terms. In this equation, κ and μ are, respectively, the bulk and shear moduli, $\mathbf{I}$ is the identity tensor and

$$\mathbf{d} = \mathbf{e} - \frac{1}{3}\mathrm{tr}(\mathbf{e})\mathbf{I}, \tag{2.22}$$

is the deviatoric part of the linearized strain tensor

$$\mathbf{e} = \frac{1}{2}\left[\nabla\mathbf{u} + (\nabla\mathbf{u})^T\right], \tag{2.23}$$

where the superscript T denotes the transpose of a second-order tensor.

The constitutive relation for a linear viscoelastic material generalizes eq. (2.21) to include a causal dependence on past deformation, and can be written in the form of Boltzmann's superposition principle

$$\mathbf{T}(t) = \kappa\nabla\cdot\mathbf{u}(t)\mathbf{I} + \int_{-\infty}^{t} 2\mu(t-t')\,\dot{\mathbf{d}}(t')\,\mathrm{d}t', \tag{2.24}$$

where μ is the time-dependent 'shear relaxation function' for the material, and $\dot{\mathbf{d}}$ is the time-derivative of $\mathbf{d}$. In writing eq. (2.24), we have again restricted attention to isotropic materials, and have, furthermore, neglected bulk viscoelasticity, these simplification being usually thought sufficient for GIA modelling (e.g. Wu & Peltier 1982; Latychev *et al.* 2005). Note that we are assuming a finite value for the bulk modulus, and so take into account the compressibility of the solid earth. Integrating by parts, we alternatively obtain

$$\mathbf{T}(t) = \kappa\nabla\cdot\mathbf{u}(t)\mathbf{I} + 2\mu_0\mathbf{d}(t) + \int_{-\infty}^{t} 2\dot{\mu}(t-t')\,\mathbf{d}(t')\,\mathrm{d}t', \tag{2.25}$$

where $\mu_0 = \mu(0)$ is known as the 'unrelaxed shear modulus' of the material. Written in this form, the first two terms on the right-hand side can be identified with the instantaneous elastic response of the material, while the final term is associated with viscoelastic relaxation.

2.2.2 *Maxwell solid rheologies*

In modelling GIA, the most commonly used shear relaxation function is a Maxwell solid for which

$$\mu(t) = \mu_0 \mathrm{e}^{-\frac{t}{\tau}}, \tag{2.26}$$

where $\mu_0 > 0$ is the unrelaxed shear modulus, and $\tau > 0$ the 'Maxwell relaxation time' (e.g. Wu & Peltier 1982; Latychev *et al.* 2005). It is well known that for deformations occurring rapidly relative to τ, a Maxwell solid responds like an isotropic elastic solid with bulk and shear moduli given by κ and μ_0, respectively. Conversely, for deformations slow relative to τ, a Maxwell solid responds like a compressible Newtonian fluid with viscosity

$$\eta = \mu_0\tau. \tag{2.27}$$

Substituting eq. (2.26) into eq. (2.25), we obtain

$$\mathbf{T}(t) = \kappa\nabla\cdot\mathbf{u}(t)\mathbf{I} + 2\mu_0\mathbf{d}(t) - \frac{2\mu_0}{\tau}\int_{t_0}^{t} \mathrm{e}^{-\frac{(t-t')}{\tau}}\,\mathbf{d}(t')\,\mathrm{d}t', \tag{2.28}$$

where we have made use of the initial conditions to fix the lower limit of integration in the final term. We now define the 'internal variable'

$$\mathbf{m}(t) = \frac{1}{\tau}\int_{t_0}^{t} \mathrm{e}^{-\frac{(t-t')}{\tau}}\,\mathbf{d}(t')\,\mathrm{d}t', \tag{2.29}$$

so that we can rewrite eq. (2.28) as

$$\mathbf{T} = \kappa\nabla\cdot\mathbf{u}\mathbf{I} + 2\mu_0(\mathbf{d} - \mathbf{m}). \tag{2.30}$$

As $\mathbf{d}$ is a trace-free tensor, the same is true of $\mathbf{m}$, and the deviatoric stress $\boldsymbol{\tau}$ in the material can be written

$$\boldsymbol{\tau} = 2\mu_0(\mathbf{d} - \mathbf{m}). \tag{2.31}$$

Differentiating eq. (2.29) with respect to time we see that $\mathbf{m}$ is a solution of the differential equation

$$\dot{\mathbf{m}} + \frac{1}{\tau}(\mathbf{m} - \mathbf{d}) = \mathbf{0}, \tag{2.32}$$

subject to the initial conditions

$$\mathbf{m}(t_0) = \mathbf{0}. \tag{2.33}$$

Taken together, eqs (2.30) and (2.32) are completely equivalent to eq. (2.28) but do not display an explicit dependence on the past deformation of the earth model. We note that this internal variable method can be extended in a simple manner to more complex linear and non-linear viscoelastic rheologies (e.g. Simo & Hughes 1998). Although such extensions in the GIA problem have been considered previously (e.g. Wu *et al.* 2010; Steffen *et al.* 2012) and may be relevant, we will not discuss them further.

2.3 Rate formulation of the viscoelastic loading problem

In Sections 2.1 and 2.2, we described a complete formulation of the viscoelastic loading problem in a Maxwell earth model. Our approach is equivalent to earlier time-domain studies (e.g. Hanyk *et al.* 1995; Zhong *et al.* 2003; Latychev *et al.* 2005), though it differs slightly due to the explicit introduction of internal variables. Following the method of Hanyk *et al.* (1995), numerical solutions of the problem could be obtained through the explicit introduction of an Euler time-stepping scheme. Moreover, the derivation of the adjoint loading problem in Section 3.2 could also be carried out using the above formulation of the problem as a starting point.

In this subsection, however, we describe an alternative 'rate formulation' of the viscoelastic loading problem that forms the basis for our numerical calculations and also for the derivation of the adjoint viscoelastic loading problem. Within the context of this paper, the advantages of this rate formulation are somewhat marginal, with the most important being (i) greater flexibility in the choice of time-stepping algorithm, and (ii) the ability to determine rates of change of the deformation fields directly without recourse to numerical differentiation. The true utility of this rate formulation will, however, only become apparent in a later paper where we consider the incorporation of gravitationally consistent sea level changes. Although a full description of this approach is beyond our present scope, it will be useful to discuss some complications associated with the incorporation of sea level changes into the GIA problem, and so motivate our formulation of the problem in terms of rates of change.

When considering loading due to both ice sheets and the ocean, the surface load σ is a sum of two terms

$$\sigma = \rho_{\mathrm{i}} I + \rho_{\mathrm{w}} S, \tag{2.34}$$

with ρ_{i} the density of ice, I the ice thickness, ρ_{w} the density of water and S the ocean height (e.g. Farrell & Clark 1976; Mitrovica & Milne 2003). In the forward GIA problem the ice thickness is given, but the evolution of the ocean height must be determined from the requirement that it is an equipotential surface of the gravitational potential and that the total H_2O mass is conserved between the oceans and ice sheets. If it is assumed that the geometry of the ocean basins does not vary with time, then these conditions lead to a simple linear relationship between variations in ocean height and the deformation fields $\mathbf{u}$ and ϕ (e.g. Farrel & Clark 1976). Variations in ocean basin geometry do, however, constitute an important feature of the GIA problem, and their inclusion leads to a complex non-linear relationship between ocean height and the deformation fields that cannot be solved explicitly. At present, the effects of shoreline migration is incorporated into the GIA modelling through iterative numerical schemes (e.g. Mitrovica & Milne 2003). While effective for forward calculations, such iterative schemes are not well suited to the development of adjoint methods, and it is desirable to have a non-iterative formulation of the GIA problem that incorporates changes in ocean basin geometry. In future work, we will present such a formulation of the GIA problem based upon the rate formulation of the viscoelastic loading problem described below. The key step in developing this approach is the observation that it is not the load σ itself but its time-derivative $\dot{\sigma}$ that occurs in the rate formulation of the problem. From eq. (2.34), we have

$$\dot{\sigma} = \rho_{\mathrm{i}} \dot{I} + \rho_{\mathrm{w}} \dot{S}, \tag{2.35}$$

and so see that in this case we instead require an equation relating the rate of change of the ocean height $\dot{S}$ to the deformation fields $\mathbf{u}$ and ϕ. It may be shown that such a relation can be obtained in closed form even when ocean basin geometry is allowed to vary with time. Moreover, this equation is linear in the time derivatives $\dot{\mathbf{u}}$ and $\dot{\phi}$ of the deformation fields, with the non-linearity of the problem being expressed solely through a dependence on the instantaneous configuration of the ocean basins. Combining this result with the rate formulation of the viscoelastic loading problem, we obtain a fully explicit-coupled system of quasi-linear evolution equations, which incorporate variations in the geometry of ocean basins. These evolution equations are, furthermore, in a form suitable for the application of adjoint methods, and the results of Section 3 of this paper can be extended to the full GIA problem in a relatively simple manner.

2.3.1 Strong form of the loading problem

To obtain the rate formulation of the viscoelastic loading problem, we first differentiate eq. (2.30) with respect to time and make use of eq. (2.32) to obtain the identity

$$\dot{\mathbf{T}} = \kappa \nabla \cdot \dot{\mathbf{u}}\, \mathbf{I} + 2\mu_0 \dot{\mathbf{d}} + \frac{2\mu_0}{\tau}(\mathbf{m} - \mathbf{d}). \tag{2.36}$$

If we then differentiate eq. (2.7) with respect to time and substitute in eq. (2.36), we arrive at

$$-\nabla \cdot \left(\kappa \nabla \cdot \dot{\mathbf{u}}\, \mathbf{I} + 2\mu_0 \dot{\mathbf{d}}\right) + \nabla(\rho \dot{\mathbf{u}} \cdot \nabla \Phi) - \nabla \cdot (\rho \dot{\mathbf{u}}) \nabla \Phi + \rho \nabla \dot{\phi} - \nabla \cdot \left[\frac{2\mu_0}{\tau}(\mathbf{m} - \mathbf{d})\right] = \mathbf{0}, \tag{2.37}$$

while differentiating eq. (2.8) we find

$$(4\pi G)^{-1} \nabla^2 \dot{\phi} = \begin{cases} -\nabla \cdot (\rho \dot{\mathbf{u}}) & \mathbf{x} \in M_{\mathrm{S}} \\ g^{-1} \dot{\phi} \partial_n \rho & \mathbf{x} \in M_{\mathrm{F}} \\ 0 & \mathbf{x} \in \mathbb{R}^3 \setminus M^{\mathrm{cl}} \end{cases}, \tag{2.38}$$

and proceeding similarly with the boundary conditions we obtain

$$\hat{\mathbf{n}} \cdot \left(\kappa \nabla \cdot \dot{\mathbf{u}}\,\mathbf{I} + 2\mu_0 \dot{\mathbf{d}}\right) + \frac{2\mu_0}{\tau}\hat{\mathbf{n}} \cdot (\mathbf{m} - \mathbf{d}) = -\dot{\sigma}\nabla\Phi, \quad \mathbf{x} \in \partial M, \tag{2.39}$$

$$\left[\hat{\mathbf{n}} \cdot \left(\kappa \nabla \cdot \dot{\mathbf{u}}\,\mathbf{I} + 2\mu_0 \dot{\mathbf{d}}\right) + \frac{2\mu_0}{\tau}\hat{\mathbf{n}} \cdot (\mathbf{m} - \mathbf{d})\right]_-^+ = \mathbf{0}, \quad \mathbf{x} \in \Sigma_{\mathrm{SS}}, \tag{2.40}$$

$$\hat{\mathbf{n}} \cdot \left(\kappa^+ \nabla \cdot \dot{\mathbf{u}}^+\,\mathbf{I} + 2\mu_0^+ \dot{\mathbf{d}}^+\right) + \frac{2\mu_0^+}{\tau^+}\hat{\mathbf{n}} \cdot (\mathbf{m}^+ - \mathbf{d}^+) = \rho^-[\dot{\mathbf{u}}^+ \cdot \nabla\Phi + \dot{\phi}]\hat{\mathbf{n}}, \quad \mathbf{x} \in \Sigma_{\mathrm{FS}}, \tag{2.41}$$

$$\hat{\mathbf{n}} \cdot \left(\kappa^- \nabla \cdot \dot{\mathbf{u}}^-\,\mathbf{I} + 2\mu_0^- \dot{\mathbf{d}}^-\right) + \frac{2\mu_0^-}{\tau^-}\hat{\mathbf{n}} \cdot (\mathbf{m}^- - \mathbf{d}^-) = \rho^+[\dot{\mathbf{u}}^- \cdot \nabla\Phi + \dot{\phi}]\hat{\mathbf{n}}, \quad \mathbf{x} \in \Sigma_{\mathrm{SF}}, \tag{2.42}$$

$$[\dot{\mathbf{u}}]_-^+ = \mathbf{0}, \quad \mathbf{x} \in \Sigma_{\mathrm{SS}}, \tag{2.43}$$

$$\left[\dot{\phi}\right]_-^+ = 0, \quad \mathbf{x} \in \Sigma, \tag{2.44}$$

$$\left[(4\pi G)^{-1}\hat{\mathbf{n}} \cdot \nabla\dot{\phi}\right]_-^+ - \rho^-\hat{\mathbf{n}} \cdot \dot{\mathbf{u}}^- = \dot{\sigma}, \quad \mathbf{x} \in \partial M, \tag{2.45}$$

$$\left[(4\pi G)^{-1}\hat{\mathbf{n}} \cdot \nabla\dot{\phi} + \rho\hat{\mathbf{n}} \cdot \dot{\mathbf{u}}\right]_-^+ = 0, \quad \mathbf{x} \in \Sigma_{\mathrm{SS}}, \tag{2.46}$$

$$\left[(4\pi G)^{-1}\hat{\mathbf{n}} \cdot \nabla\dot{\phi}\right]_-^+ + [\rho]_-^+\,\hat{\mathbf{n}} \cdot \dot{\mathbf{u}}^+ = 0, \quad \mathbf{x} \in \Sigma_{\mathrm{FS}}, \tag{2.47}$$

$$\left[(4\pi G)^{-1}\hat{\mathbf{n}} \cdot \nabla\dot{\phi}\right]_-^+ + [\rho]_-^+\,\hat{\mathbf{n}} \cdot \dot{\mathbf{u}}^- = 0, \quad \mathbf{x} \in \Sigma_{\mathrm{SF}}, \tag{2.48}$$

along with the condition that $\dot{\phi} \to 0$ as $\|\mathbf{x}\| \to \infty$. Taken together, eqs (2.37) through to (2.48) constitute a linear boundary value problem whose solution gives the instantaneous values of $\dot{\mathbf{u}}$ and $\dot{\phi}$ from knowledge of the current values of these fields along with those of the internal variable $\mathbf{m}$ and the time derivative of the applied surface load $\dot{\sigma}$. Moreover, this boundary value problem takes precisely the same form as a linear elastostatic loading problem but for the occurrence of additional body and surface forces associated with viscoelastic relaxation. Finally, combining this boundary value problem with eq. (2.32), we obtain the desired coupled system of evolution equations for the variables $\mathbf{u}$, ϕ and $\mathbf{m}$. The above derivation can also be extended to the case where the load σ possesses a finite number of discontinuities, and the necessary jump conditions can be found in Appendix A.

2.3.2 Weak form of the loading problem

The time-domain loading problem described above is said to be in the 'strong form' due to the equations being defined pointwise in space and in time. We now consider an alternative 'weak form' of the problem which involves global integrals over the spatial domain. This formulation provides the basis for the numerical solution of the loading problem using finite-element-like methods, and also plays a central role in the derivation of the adjoint loading problem below.

To describe the weak form of the viscoelastic loading problem, we introduce suitably regular time-independent test functions $\mathbf{u}'$, ϕ' and $\mathbf{m}'$, the first and third of which are defined in M_{S}, while the second is defined in $\mathbb{R}^3$. We also require that $\mathbf{u}'$ and ϕ' satisfies the kinematic boundary conditions

$$\left[\mathbf{u}'\right]_-^+ = \mathbf{0}, \quad \mathbf{x} \in \Sigma_{\mathrm{SS}}, \tag{2.49}$$

$$\left[\phi'\right]_-^+ = 0, \quad \mathbf{x} \in \Sigma, \tag{2.50}$$

and that $\phi' \to 0$ as $\|\mathbf{x}\| \to \infty$. We then say that $\mathbf{u}$, ϕ and $\mathbf{m}$ are solutions of the weak form of the viscoelastic loading problem if

$$\mathcal{A}(\dot{\mathbf{u}}, \dot{\phi}\,|\,\mathbf{u}', \phi') - \int_{M_{\mathrm{S}}} 2\mu_0 \left[\dot{\mathbf{m}} : \mathbf{m}' + \frac{1}{\tau}(\mathbf{d} - \mathbf{m}) : (\mathbf{d}' - \mathbf{m}')\right] \mathrm{d}V + \int_{\partial M} (\nabla\Phi \cdot \mathbf{u}' + \phi')\dot{\sigma}\,\mathrm{d}S = 0, \tag{2.51}$$

holds for all test functions $\mathbf{u}'$, ϕ' and $\mathbf{m}'$, where $\mathbf{d}'$ is the deviatoric strain tensor associated with $\mathbf{u}'$, the notation ':' denotes the contraction of second-order tensors, and we have defined the bilinear form

$$\begin{aligned}\mathcal{A}(\mathbf{u},\phi\,|\,\mathbf{u}',\phi') &= \int_{M_S}\kappa\nabla\cdot\mathbf{u}\,\nabla\cdot\mathbf{u}'\,\mathrm{d}V+\int_{M_S}2\mu\mathbf{d}:\mathbf{d}'\,\mathrm{d}V+\frac{1}{2}\int_{M_S}\rho\left[\nabla(\mathbf{u}\cdot\nabla\Phi)\cdot\mathbf{u}'+\nabla(\mathbf{u}'\cdot\nabla\Phi)\cdot\mathbf{u}\right]\mathrm{d}V\\ &-\frac{1}{2}\int_{M_S}\rho\left(\nabla\cdot\mathbf{u}\nabla\Phi\cdot\mathbf{u}'+\nabla\cdot\mathbf{u}'\nabla\Phi\cdot\mathbf{u}\right)\mathrm{d}V+\int_{M_S}\rho(\nabla\phi\cdot\mathbf{u}'+\mathbf{u}\cdot\nabla\phi')\,\mathrm{d}V+\frac{1}{4\pi G}\int_{\mathbb{R}^3}\nabla\phi\cdot\nabla\phi'\,\mathrm{d}V\\ &+\int_{M_F}g^{-1}\phi\phi'\partial_n\rho\,\mathrm{d}V+\int_{\Sigma_{FS}}\rho^-g\hat{\mathbf{n}}\cdot\mathbf{u}\,\hat{\mathbf{n}}\cdot\mathbf{u}'\,\mathrm{d}S-\int_{\Sigma_{SF}}\rho^+g\hat{\mathbf{n}}\cdot\mathbf{u}\,\hat{\mathbf{n}}\cdot\mathbf{u}'\,\mathrm{d}S+\int_{\Sigma_{FS}}\rho^-(\phi\mathbf{u}'+\mathbf{u}\phi')\cdot\hat{\mathbf{n}}\,\mathrm{d}S\\ &-\int_{\Sigma_{SF}}\rho^+(\phi\mathbf{u}'+\mathbf{u}\phi')\cdot\hat{\mathbf{n}}\,\mathrm{d}S,\end{aligned} \tag{2.52}$$

which is symmetric in the sense that

$$\mathcal{A}(\mathbf{u},\phi\,|\,\mathbf{u}',\phi')=\mathcal{A}(\mathbf{u}',\phi'\,|\,\mathbf{u},\phi). \tag{2.53}$$

Given that $\mathbf{u}$, ϕ and $\mathbf{m}$ are solutions of the strong form of the loading problem, it may be shown that eq. (2.51) holds for arbitrary test functions through a lengthy but simple argument whose details can be found in Appendix B. Conversely, if $\mathbf{u}$, ϕ and $\mathbf{m}$ are solutions of the weak form of the loading problem, then this argument can be reversed to show that they also satisfy the strong form of the loading problem.

In the weak form of the viscoelastic loading problem given in eqs (2.51) and (B39), the time evolution of the fields $\mathbf{u}$, ϕ and $\mathbf{m}$ is defined pointwise. This approach is well suited to numerical solution using finite-element-like methods coupled to a suitable time-stepping algorithm. In developing the adjoint method, it will be useful to obtain an alternative 'time-integrated weak form' of the viscoelastic loading problem in which the time evolution of the system is also expressed in a variational manner. Having done so, we can then take the time-integrated weak form as the sole constraint when defining an appropriate Lagrangian functional in Section 3.2, and need not explicitly incorporate the initial or jump conditions into the PDE-constrained optimization problem (PDE is an abbreviation of partial differential equation). To obtain this result, we let $[t_0, t_1]$ be the time interval of interest, and suppose that the test functions $\mathbf{u}'$, ϕ' and $\mathbf{m}'$ are time dependent, that they satisfy the terminal conditions

$$\mathbf{u}'(t_1)=\mathbf{0},\quad \phi'(t_1)=0,\quad \mathbf{m}'(t_1)=\mathbf{0}, \tag{2.54}$$

and are continuous at all times when σ undergoes a jump discontinuity. Integrating eq. (2.51) with respect to time we find

$$\int_{t_0}^{t_1}\left\{\mathcal{A}(\dot{\mathbf{u}},\dot{\phi}\,|\,\mathbf{u}',\phi')-\int_{M_S}2\mu_0\left[\dot{\mathbf{m}}:\mathbf{m}'+\frac{1}{\tau}(\mathbf{d}-\mathbf{m}):(\mathbf{d}'-\mathbf{m}')\right]\mathrm{d}V+\int_{\partial M}(\nabla\Phi\cdot\mathbf{u}'+\phi')\dot{\sigma}\,\mathrm{d}S\right\}\mathrm{d}t=0, \tag{2.55}$$

and using integration by parts obtain

$$\int_{t_0}^{t_1}\left\{-\mathcal{A}(\mathbf{u},\phi\,|\,\dot{\mathbf{u}}',\dot{\phi}')-\int_{M_S}2\mu_0\left[-\mathbf{m}:\dot{\mathbf{m}}'+\frac{1}{\tau}(\mathbf{d}-\mathbf{m}):(\mathbf{d}'-\mathbf{m}')\right]\mathrm{d}V-\int_{\partial M}(\nabla\Phi\cdot\dot{\mathbf{u}}'+\dot{\phi}')\sigma\,\mathrm{d}S\right\}\mathrm{d}t=0, \tag{2.56}$$

where we have made use of the continuity and terminal conditions on the test functions and have employed the jump condition in eq. (B39). This argument shows that if $\mathbf{u}$, ϕ and $\mathbf{m}$ are solutions of the viscoelastic loading problem, then eq. (2.56) holds for any such test functions. Conversely, if eq. (2.56) holds for all suitable test functions we can reverse the above argument to conclude that $\mathbf{u}$, ϕ and $\mathbf{m}$ satisfy the evolution equation given in eq. (2.51), the jump conditions in eq. (B39) and the initial conditions for the problem.

3 THE INVERSE VISCOELASTIC LOADING PROBLEM

We now turn attention to the 'inverse viscoelastic loading problem', where we recall that by this we mean the simplified version of the inverse GIA problem in which the effects of gravitationally self-consistent ocean loading and rotational feedbacks have been neglected. In this problem, the data comprise observations of the deformation of a viscoelastic earth model due to surface loading, and from these observations we wish to make inferences about both the load σ and the mantle viscosity η. In this problem the observations do, of course, also depend on the geometry and elastic structure of the earth model, but we shall suppose that these parameters are known. In particular, we assume that the earth model has an elastic lithosphere of a known, but potentially laterally variable, thickness. Extension of the adjoint method to include the lithospheric thickness as an additional model parameter is possible, but has yet to be carried out.

For a set of model parameters $\{\sigma, \eta\}$, we can solve the viscoelastic loading problem to determine synthetic deformation fields $\mathbf{u}$ and ϕ, and so compare the predictions of the model with observations. Typically, the agreement between the observed and synthetic deformation fields is quantified in terms a misfit functional, which may also incorporate some form of model regularization. In this manner, we can reduce the inverse problem to finding the model which minimizes the misfit functional. Such optimization problems can be practically solved using gradient-based algorithms (Nocedal & Wright 2006; Tape *et al.* 2007), and our aim in this section will be the efficient calculation of the gradients required by such methods.

3.1 An example objective functional

To motivate the development of the adjoint method, we now consider an idealized, but realistic, example of the inverse viscoelastic loading problem. Let $\mathbf{x}_i \in \partial M$ for $i = 1, \ldots, N_r$ be a set of surface locations at which we have measured vertical displacements d_{ij} at the times t_{ij} for $j = 1, \ldots, N_t^i$. Excepting recent global positioning system (GPS)-derived uplift measurements (e.g. Milne *et al.* 2001), such observations cannot be readily made in practice. However, within our simplified viscoelastic loading problem, these displacement measurements should be regarded as analogs for those of past sea level (e.g. Peltier 2004). Any such data set will, of course, be subject to errors in both the recorded displacements and times. For simplicity, we suppose that the errors are uncorrelated, zero-mean, Gaussian random errors, such that t_{ij} and d_{ij} have variances given by Δt_{ij}^2 and Δd_{ij}^2, respectively.

For given values of $\{\sigma, \eta\}$, we write $\mathbf{u}$ for the corresponding synthetic displacement field, and quantify the misfit between the given data and our model predictions through the conventional least-squares misfit

$$J = \frac{1}{2N_r} \sum_{i=1}^{N_r} \frac{1}{N_t^i} \sum_{j=1}^{N_t^i} \left\{ \frac{1}{\Delta d_{ij}^2} [\hat{\mathbf{n}} \cdot \mathbf{u}(\mathbf{x}_i, t'_{ij}) - d_{ij}]^2 + \frac{1}{\Delta t_{ij}^2} (t'_{ij} - t_{ij})^2 \right\}, \tag{3.1}$$

where t'_{ij} are 'model times' that have been introduced into the problem to account for the presence of timing errors in the data. In the interest of generality, we shall refer to J as the 'objective functional' for the problem, and regard it as a function of the surface values of $\mathbf{u}$ and ϕ. As the synthetic deformation fields $\mathbf{u}$ and ϕ depend on the model parameters $\{\sigma, \eta\}$, the same is implicitly true of J, and we can define a 'reduced objective functional' through the relation

$$\hat{J}(\sigma, \eta) = J(\mathbf{u}, \phi), \tag{3.2}$$

where it is understood that $\mathbf{u}$ and ϕ are the deformation fields obtained by solving the viscoelastic loading problem corresponding to the given model parameters. In practice, the objective functional may also display an explicit dependence on $\{\sigma, \eta\}$ due to model regularization. We need not, however, consider further any explicit dependence of $\hat{J}$ on the model parameters, as the resulting contributions to the derivatives can be directly calculated.

Given perturbations to $\{\sigma, \eta\}$, our aim is to write the resulting first-order perturbation to $\hat{J}$ in the form

$$\delta\hat{J} = \int_{M_S} K_\eta \delta \ln \eta \, \mathrm{d}V + \int_{t_0}^{t_1} \int_{\partial M} K_\sigma \delta\sigma \, \mathrm{d}S \, \mathrm{d}t, \tag{3.3}$$

where K_η is 'viscosity sensitivity kernel' that is defined for $\mathbf{x} \in M_S$, while K_σ is the 'load sensitivity kernel', which is defined on $\partial M \times [t_0, t_1]$, with $[t_0, t_1]$ an appropriate time interval, and where we have, for convenience, normalized the perturbation $\delta \ln \eta = \delta\eta/\eta$ to the viscosity by its references value. In fact, we will actually find it preferable to use $\dot{\sigma}$ and not σ as the loading parameter, and so will also obtain the alternate expression

$$\delta\hat{J} = \int_{M_S} K_\eta \delta \ln \eta \, \mathrm{d}V + \int_{t_0}^{t_1} \int_{\partial M} K_{\dot{\sigma}} \delta\dot{\sigma} \, \mathrm{d}S \, \mathrm{d}t, \tag{3.4}$$

where $K_{\dot{\sigma}}$ is the 'rate-of-load sensitivity kernel', which is defined on $\partial M \times [t_0, t_1]$.

The viscosity kernel introduced in eq. (3.3) expresses the linearized sensitivity of $\hat{J}$ to a laterally heterogeneous viscosity perturbation $\delta \ln \eta$ defined with respect to a laterally heterogeneous reference model. It is, however, of some interest to restrict attention to spherically symmetric viscosity perturbations defined with respect to a spherically symmetric background models. In this case, we can define a 'radial viscosity kernel' $\overline{K}_\eta$ through the relation

$$\int_{I_S} \overline{K}_\eta \delta \ln \eta \, \mathrm{d}r = \int_{M_S} K_\eta \delta \ln \eta \, \mathrm{d}V, \tag{3.5}$$

where I_S denotes the subset of radii r in the earth model lying in solid regions, and it is understood that here η and $\delta\eta$ are spherically symmetric. It follows that $\overline{K}_\eta$ is given by

$$\overline{K}_\eta = \int_\Omega K_\eta r^2 \, \mathrm{d}S, \tag{3.6}$$

where Ω denotes the unit two-sphere.

3.2 The adjoint viscoelastic loading problem

We wish to determine the derivative of the reduced objective functional $\hat{J}$ defined in eq. (3.2) with respect to the model parameters $\{\sigma, \eta\}$. As, however, $\hat{J}$ only depends on the model parameters implicitly, its derivatives are most efficiently calculated using the method of PDE-constrained optimization (e.g. Tromp *et al.* 2005; Tröltzsch 2005, Liu & Tromp 2008). To do so, we introduce the Lagrangian functional

$$L = J - \int_{t_0}^{t_1} \left\{ \mathcal{A}(\mathbf{u}, \phi \,|\, \dot{\mathbf{u}}', \dot{\phi}') + \int_{M_\mathrm{S}} 2\mu_0 \left[-\mathbf{m} : \dot{\mathbf{m}}' + \frac{1}{\tau}(\mathbf{d} - \mathbf{m}) : (\mathbf{d}' - \mathbf{m}') \right] \mathrm{d}V + \int_{\partial M} (\nabla\Phi \cdot \dot{\mathbf{u}}' + \dot{\phi}')\sigma \, \mathrm{d}S \right\} \mathrm{d}t, \tag{3.7}$$

where the first term on the right-hand side is the objective functional of interest, and the second term is the time-integrated weak form of the viscoelastic loading problem given in eq. (2.56). Here, the test functions $\mathbf{u}'$, ϕ' and $\mathbf{m}'$ act as Lagrange multiplier fields associated with the constraint that $\mathbf{u}$, ϕ and $\mathbf{m}$ are solutions of the viscoelastic loading problem. Note that, for consistency, in what follows we shall persist in referring to $\mathbf{u}'$, ϕ' and $\mathbf{m}'$ as test functions.

This Lagrangian could have equivalently been obtained by starting from the strong form of the viscoelastic loading problem given in Section 2.3 and adding to J each of the equations of motion and boundary conditions as constraints using appropriate Lagrange multiplier fields. After substantial manipulation of the resulting expression, this Lagrangian would then reduce to the form given above. Although this alternative approach is, perhaps, more clearly motivated, we have preferred to first obtain the weak form of the viscoelastic loading problem as this result is both useful in numerical applications, and leads to a simple derivation of the adjoint loading problem.

Following the method of Lagrange multipliers, the first-order perturbation $\delta\hat{J}$ to the reduced objective functional with respect to variations in $\{\sigma, \eta\}$ can be determined by calculating the corresponding first-order perturbation δL to the Lagrangian functional subject to the condition that L is stationary with respect to variations in both the forward fields $\mathbf{u}$, ϕ and $\mathbf{m}$ and the test functions $\mathbf{u}'$, ϕ' and $\mathbf{m}'$ (e.g. Tröltzsch 2005). That L be stationary with respect to variations in the test functions simply implies that $\mathbf{u}$, ϕ and $\mathbf{m}$ are solutions of the viscoelastic loading problem. From the condition that L be stationary with respect to variations in $\mathbf{u}$, ϕ and $\mathbf{m}$, we obtain

$$\int_{t_0}^{t_1} \left\{ -\mathcal{A}(\delta\mathbf{u}, \delta\phi \,|\, \dot{\mathbf{u}}', \dot{\phi}') - \int_{M_\mathrm{S}} 2\mu_0 \left[-\delta\mathbf{m} : \dot{\mathbf{m}}' + \frac{1}{\tau}(\delta\mathbf{d} - \delta\mathbf{m}) : (\mathbf{d}' - \mathbf{m}') \right] \mathrm{d}V + \int_{\partial M} \left(\dot{\mathbf{h}} \cdot \delta\mathbf{u} + \dot{h}\delta\phi \right) \mathrm{d}S \right\} \mathrm{d}t = 0, \tag{3.8}$$

for all $\delta\mathbf{u}$, $\delta\phi$ and $\delta\mathbf{m}$. Here, for both convenience and generality we have written the first-order perturbation to J with respect to $\mathbf{u}$ and ϕ in the form

$$\delta J = \int_{t_0}^{t_1} \int_{\partial M} \left(\dot{\mathbf{h}} \cdot \delta\mathbf{u} + \dot{h}\delta\phi \right) \mathrm{d}S \, \mathrm{d}t, \tag{3.9}$$

where $\dot{\mathbf{h}}$ is the Fréchet kernel of J with respect to $\mathbf{u}$, and $\dot{h}$ is the Fréchet kernel of J with respect to ϕ. Equivalently, this implies that for all time-independent $\delta\mathbf{u}$, $\delta\phi$ and $\delta\mathbf{m}$, we have

$$-\mathcal{A}(\delta\mathbf{u}, \delta\phi \,|\, \dot{\mathbf{u}}', \dot{\phi}') - \int_{M_\mathrm{S}} 2\mu_0 \left[-\delta\mathbf{m} : \dot{\mathbf{m}}' + \frac{1}{\tau}(\delta\mathbf{d} - \delta\mathbf{m}) : (\mathbf{d}' - \mathbf{m}') \right] \mathrm{d}V + \int_{\partial M} \left(\dot{\mathbf{h}} \cdot \delta\mathbf{u} + \dot{h}\delta\phi \right) \mathrm{d}S = 0. \tag{3.10}$$

The Fréchet kernels $\dot{\mathbf{h}}$ and $\dot{h}$ introduced above have been written as time derivatives to simplify the resulting form of the adjoint problem. This does, however, lead to an ambiguity in their definition up to the addition of a constant in time, which we remove by requiring that $\mathbf{h} = \mathbf{0}$, $h = 0$ for $t \geq t_1$. For the particular objective functional given in eq. (3.1), we readily find

$$\mathbf{h}(\mathbf{x}, t) = -\frac{1}{N_r} \sum_{i=1}^{N_r} \frac{1}{N_t^i} \sum_{j=1}^{N_t^i} \frac{1}{\Delta d_{ij}^2} [\hat{\mathbf{n}} \cdot \mathbf{u}(\mathbf{x}_i, t'_{ij}) - d_{ij}] \delta(\mathbf{x} - \mathbf{x}_i) H(t'_{ij} - t) \hat{\mathbf{n}}, \tag{3.11}$$

$$h(\mathbf{x}, t) = 0, \tag{3.12}$$

where $\delta(\mathbf{x})$ is the Dirac delta function on ∂M, $H(t)$ the Heaviside step function, we have recalled the identity

$$\dot{H}(t' - t) = -\delta(t - t') \tag{3.13}$$

and the time dependence of $\mathbf{h}$ has been fixed using the convention described above.

At this point, it will be useful to define 'adjoint fields' through

$$\mathbf{u}^\dagger(t) = \mathbf{u}'(t_1 - t + t_0), \tag{3.14}$$

$$\phi^\dagger(t) = \phi'(t_1 - t + t_0), \tag{3.15}$$

$$\mathbf{m}^\dagger(t) = \mathbf{m}'(t_1 - t + t_0), \tag{3.16}$$

and, similarly, the 'adjoint loads'

$$\mathbf{h}^\dagger(t) = \mathbf{h}(t_1 - t + t_0), \tag{3.17}$$

$$h^\dagger(t) = h(t_1 - t + t_0). \tag{3.18}$$

We see, in particular, that the terminal conditions given in eq. (2.54) for the time-dependent test functions are transformed into initial conditions

$$\mathbf{u}^\dagger(t_0) = \mathbf{0}, \quad \phi^\dagger(t_0) = 0, \quad \mathbf{m}^\dagger(t_0) = \mathbf{0}, \tag{3.19}$$

for the adjoint fields. If we replace the test functions in eq. (3.10) by the corresponding adjoint variables, we then obtain

$$\mathcal{A}(\dot{\mathbf{u}}^\dagger, \dot{\phi}^\dagger \,|\, \delta\mathbf{u}, \delta\phi) - \int_{M_S} 2\mu_0 \left[\dot{\mathbf{m}}^\dagger : \delta\mathbf{m} + \frac{1}{\tau}(\mathbf{d}^\dagger - \mathbf{m}^\dagger) : (\delta\mathbf{d} - \delta\mathbf{m}) \right] \mathrm{d}V - \int_{\partial M} \left(\dot{\mathbf{h}}^\dagger \cdot \delta\mathbf{u} + \dot{h}^\dagger \delta\phi \right) \mathrm{d}S = 0, \tag{3.20}$$

which is to hold for all $\delta\mathbf{u}$, $\delta\phi$ and $\delta\mathbf{m}$, and where we have made use of the symmetry of $\mathcal{A}$ given in eq. (2.53). We shall say that eq. (3.20) is the weak form of the 'adjoint viscoelastic loading problem', and require that the adjoint fields satisfy this equation. This adjoint viscoelastic loading problem is of precisely the same form as the viscoelastic loading problem given in eq. (2.51), with the only difference being in the force terms. Consequently, the same numerical method can be used for the solution of both the forward and adjoint viscoelastic loading problems.

With the test functions $\mathbf{u}'$, ϕ' and $\mathbf{m}'$, determined through the solution of the adjoint viscoelastic loading problem, we can equate the first-order perturbation to the reduced objective functional $\hat{J}$ with that of L and so obtain

$$\begin{aligned}\delta\hat{J} = &\int_{t_0}^{t_1}\int_{M_S} \frac{2\mu_0}{\tau}[\mathbf{d}(t) - \mathbf{m}(t)] : [\mathbf{d}^\dagger(t_1 - t + t_0) - \mathbf{m}^\dagger(t_1 - t + t_0)]\frac{\delta\eta}{\eta}\,\mathrm{d}V\,\mathrm{d}t \\ &- \int_{t_0}^{t_1}\int_{\partial M} \left[\nabla\Phi \cdot \dot{\mathbf{u}}^\dagger(t_1 - t + t_0) + \dot{\phi}^\dagger(t_1 - t + t_0)\right]\delta\sigma(t)\,\mathrm{d}S\,\mathrm{d}t,\end{aligned} \tag{3.21}$$

where for clarity we have explicitly included time arguments. From this expression we can identify the desired sensitivity kernels as being

$$K_\eta = \int_{t_0}^{t_1} \frac{1}{2\eta}\boldsymbol{\tau}(t) : \boldsymbol{\tau}^\dagger(t_1 - t + t_0)\,\mathrm{d}t, \tag{3.22}$$

$$K_\sigma(t) = -\nabla\Phi \cdot \dot{\mathbf{u}}^\dagger(t_1 - t + t_0) - \dot{\phi}^\dagger(t_1 - t + t_0), \tag{3.23}$$

where we have recalled the expression in eq. (2.31) for the deviatoric stress.

The adjoint viscoelastic loading problem given in weak form in eq. (3.20) along with the sensitivity kernels in eqs (3.22) and (3.23) are the main results of this paper. These sensitivity kernels represent the derivative of the reduced objective functional, and could be used in a gradient-based optimization algorithm to iteratively update the model parameters $\{\sigma, \eta\}$ so as to better fit the given rebound curves in our example inverse problem. To calculate these sensitivity kernels, we must:

(1) Solve the forward viscoelastic loading problem once to obtain $\mathbf{u}$, ϕ and $\boldsymbol{\tau}$.
(2) Calculate the adjoint loads $\mathbf{h}^\dagger$ and $h^\dagger$ from $\mathbf{u}$ and ϕ using eq. (3.11).
(3) Solve the adjoint viscoelastic loading problem once to obtain $\mathbf{u}^\dagger$, $\phi^\dagger$ and $\boldsymbol{\tau}^\dagger$.
(4) Calculate the sensitivity kernels using eqs (3.22) and (3.23).

In this process, we note that the objective functional enters the problem only through the adjoint loads $\mathbf{h}^\dagger$ and $h^\dagger$. It follows that given any objective functional J, we can similarly calculate the associated sensitivity kernels, and to do so we need only determine the appropriate form of the adjoint loads. In Appendix C, we illustrate the versatility of the adjoint method by describing a number of other possible objective functionals. In particular, we note that such objective functionals need not represent a misfit between data and synthetic predictions. Rather, the objective functional can simply represent some property of the deformation fields of interest, and the sensitivity kernels show, in a linearized manner, how this property depends on both the load and viscosity.

During the implementation of the adjoint method, only one solution of the forward viscoelastic loading problem is required, and then one further solution of the adjoint viscoelastic loading problem. It is this property that gives the adjoint method such great computational efficiency. Moreover, we saw above that the adjoint viscoelastic loading problem is actually the same as the forward problem except for the force term. This means that a code for solving the forward viscoelastic loading problem can with only very minor modification be used to solve the adjoint problem as well. Finally, it is noteworthy that the expression for the viscosity sensitivity kernel given in eq. (3.22) does not depend upon the internal variable formulation of viscoelasticity used in this paper. This means that any code capable of determining the deviatoric stresses associated with the forward viscoelastic loading problem could be used to calculate viscosity kernels. In particular, this should be the case with existing fully numerical codes for solving the viscoelastic loading problem in laterally heterogeneous earth models (e.g. Zhong *et al.* 2003; Latychev *et al.* 2005).

3.3 Rate-of-loading kernel

If the adjoint loads undergo a finite jump discontinuity, then the same is true of the adjoint deformation fields $\mathbf{u}^\dagger$, and $\phi^\dagger$, and the size of this jump can be obtained through solution of the static loading problem give in eq. (B40). Such jumps in the adjoint load occur, in particular, whenever the objective functional depends upon the point value of the deformation fields at a given time. For example, we see that the adjoint load $\mathbf{h}^\dagger$ corresponding to eq. (3.11) is discontinuous at each of the times $t_1 - t_{ij} - t_0$ for $i = 1, \ldots, N_r$ and $j = 1, \ldots, N_t^i$. In eq. (3.23), it is the time derivative of the adjoint deformation fields that occur in the expression for the load kernel, and so we see that for each discontinuity in the adjoint loads at a time t' there corresponds a delta-function-like singularity in K_σ at the reversed time $t_1 - t' - t_0$. In such cases, the derivative of $\hat{J}$ with respect to σ cannot be identified with a possible load, and gradient-based methods, such as the steepest descent algorithm, are inapplicable.

This problem can be circumvented through a method for 'quelling' singular kernels described by Backus (1970). To do so, we regard $\dot{\sigma}$ and not σ as the relevant model parameter, and using integration by parts obtain

$$-\int_{t_0}^{t_1}\int_{\partial M}\left[\nabla\Phi\cdot\dot{\mathbf{u}}^{\dagger}(t_1-t+t_0)+\dot{\phi}^{\dagger}(t_1-t+t_0)\right]\delta\sigma(t)\,\mathrm{d}S\,\mathrm{d}t=\int_{t_0}^{t_1}\int_{\partial M}\left[\nabla\Phi\cdot\mathbf{u}^{\dagger}(t_1-t+t_0)+\phi^{\dagger}(t_1-t+t_0)\right]\delta\dot{\sigma}(t)\,\mathrm{d}S\,\mathrm{d}t, \tag{3.24}$$

where we have used the initial conditions on the adjoint deformation fields and the load perturbation. Substituting this result into eq. (3.21), we can then identify the rate-of-loading sensitivity kernel as

$$K_{\dot{\sigma}}(t)=\nabla\Phi\cdot\mathbf{u}^{\dagger}(t_1-t+t_0)+\phi^{\dagger}(t_1-t+t_0). \tag{3.25}$$

In the case that the adjoint loads are discontinuous at a time t', we now see that the rate-of-loading kernel is non-singular, though discontinuous, at the reversed time $t_1-t'-t_0$. Because of this property, $K_{\dot{\sigma}}$ is suitable for use with methods like the steepest descent algorithm, and we take $\dot{\sigma}$ as our preferred loading parameter. Moreover, when adding an update to $\dot{\sigma}$ that is discontinuous in time, the corresponding update to σ is, of course, continuous in time, and it is only such loads that are physically acceptable.

4 APPLICATION TO SPHERICALLY SYMMETRIC EARTH MODELS

To illustrate potential applications of the adjoint method, we now present some calculations performed in a spherically symmetric earth model possessing a solid elastic inner core, an inviscid fluid outer core, a viscoelastic mantle with Maxwell solid rheology, and an elastic lithosphere. Details of the numerical methods can be found in Appendix D, and here we simply outline the computational steps in calculating both the rate-of-loading and viscosity kernels. In these examples we have considered spherically symmetric earth models to simplify the numerical calculations, but this is in no way a limitation of the theory. Similarly, we note that, though all spherical harmonic expansions occurring below are truncated at degree 80, this was an arbitrary choice, and calculations including higher harmonic degrees can, and have, been performed. For a number of the examples, we refer the reader to Appendix C where discussion of the various objective functionals J can be found.

The earth model's elastic structure has been taken to be isotropic preliminary reference earth model (PREM) of Dziewonski & Anderson (1981) with the ocean replaced by crustal material. In order to assess the effects of the background viscosity structure on the sensitivity kernels, we have considered three viscosity models of increasing complexity shown in Fig. 1. The third and most complex of these models is a modified version of VM2 of Peltier (2004) in which the piecewise constant viscosity structure has been smoothed using a two-point running average, except the major discontinuity at 670 km depth has been retained. We use this smoothed model to emphasize that such time-domain

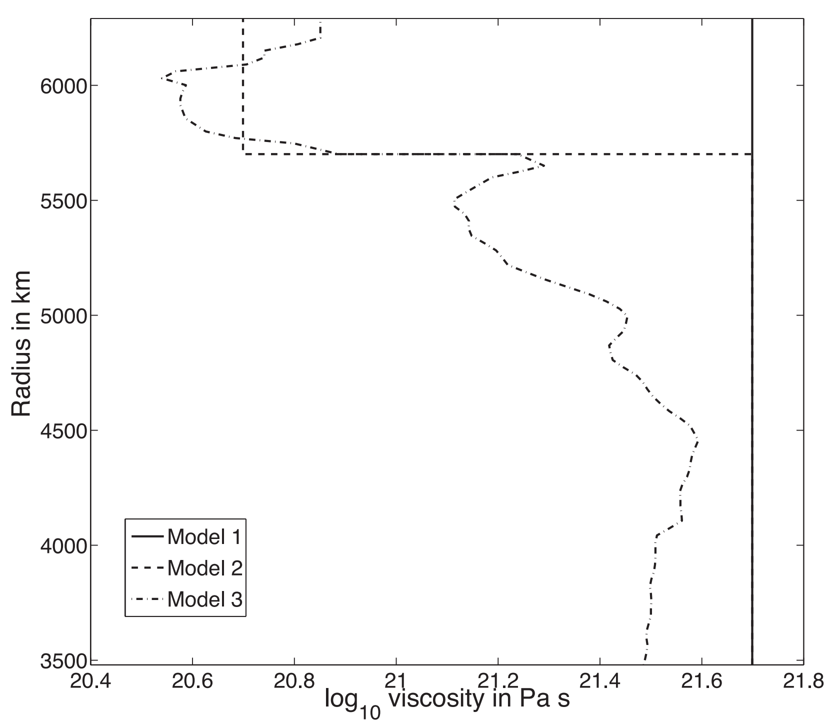


Figure 1. Radial viscosity profiles of mantle viscosity for the three earth models used in the range 3480–6251 km. In each case, the elastic structure of the model is that of isotropic PREM (Dziewonski & Anderson 1981) and there is a 120-km thick elastic lithosphere.

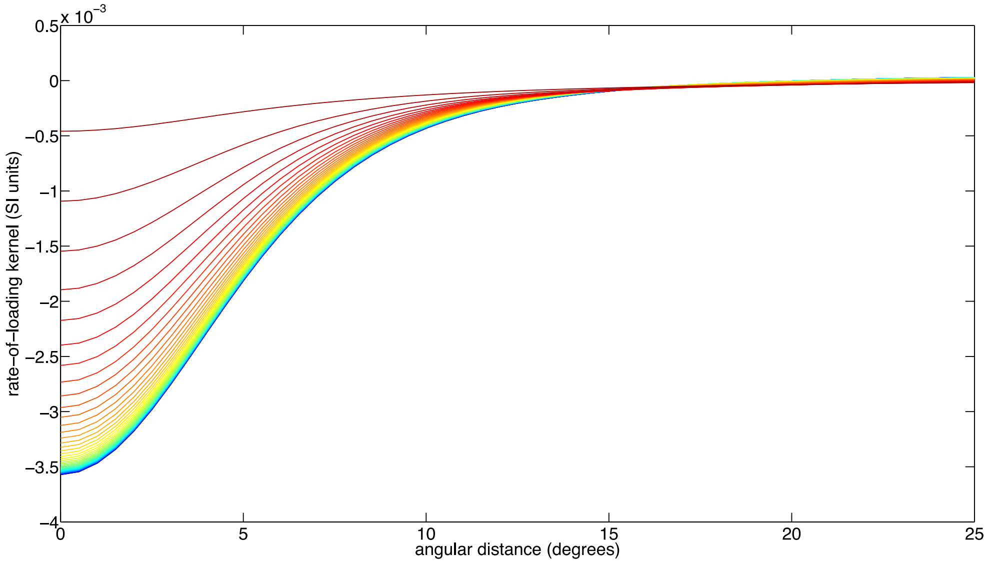


Figure 2. The rate-of-loading kernel calculated in viscosity model 2 corresponding to a point measurement of vertical displacement plotted as a function of angular distance from the observation point. Each line on the plot shows the rate-of-loading kernel at a value of $t - t'$ in the range -50 to 0 kyr in 1 kyr intervals. The line for -50 kyr is plotted in darkest blue and intersects the y-axis at the lowest point, while the line for 0 kyr is plotted in darkest red and intersects the y-axis at the highest point. The colours of the lines corresponding to intermediate times grade smoothly between those of the endpoints.

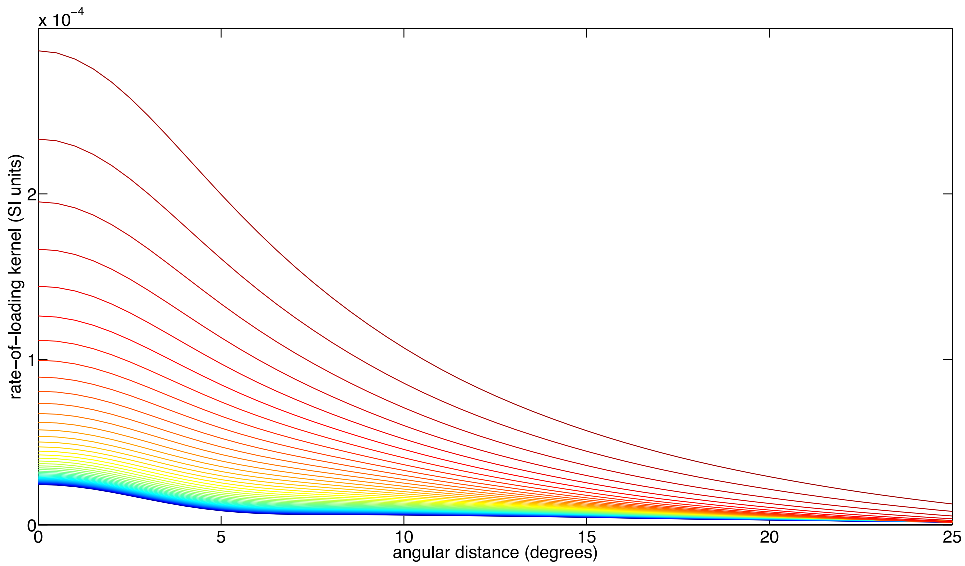


Figure 3. The rate-of-loading kernel corresponding to a point measurement of geoid anomaly plotted as a function of angular distance from the observation point. Each line on the plot shows the rate-of-loading kernel at a value of $t - t'$ in the range -50 to 0 kyr in 1 kyr intervals. The line for -50 kyr is plotted in darkest blue and intersects the y-axis at the lowest point, while the line for 0 kyr is plotted in darkest red and intersects the y-axis at the highest point. The colours of the lines corresponding to intermediate times grade smoothly between those of the endpoints.

methods can be applied with continuously varying radial viscosities. This situation is in contrast with the 'viscoelastic normal mode method' that can only be rigorously applied in piecewise constant viscosity structures (e.g. Fang & Hager 1995; Han & Wahr 1995; Al-Attar 2011).

4.1 Rate-of-loading sensitivity kernels

4.1.1 Computational details

The rate-of-loading kernel $K_{\dot{\sigma}}$ corresponding to a given objective functional J may be calculated in the following manner:

(1) Solve the forward viscoelastic loading problem to determine the deformation fields $\mathbf{u}$ and ϕ.
(2) From $\mathbf{u}$ and ϕ calculate the adjoint loads $\mathbf{h}^{\dagger}$ and $h^{\dagger}$ associated with J through eqs (3.9), (3.17) and (3.18).
(3) Using these adjoint loads, solve the adjoint viscoelastic loading problem to obtain the adjoint deformation fields $\mathbf{u}^{\dagger}$ and $\phi^{\dagger}$.
(4) Compute the rate-of-loading sensitivity kernel from $\mathbf{u}^{\dagger}$ and $\phi^{\dagger}$ using eq. (3.25).

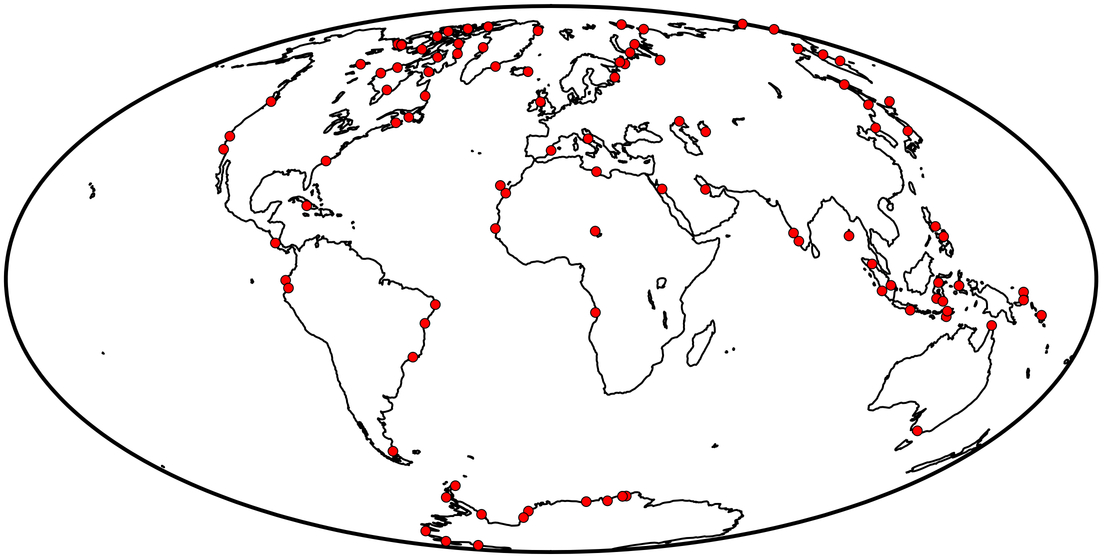


Figure 4. Location distribution used in the generation of a synthetic data set of vertical uplift measurements.

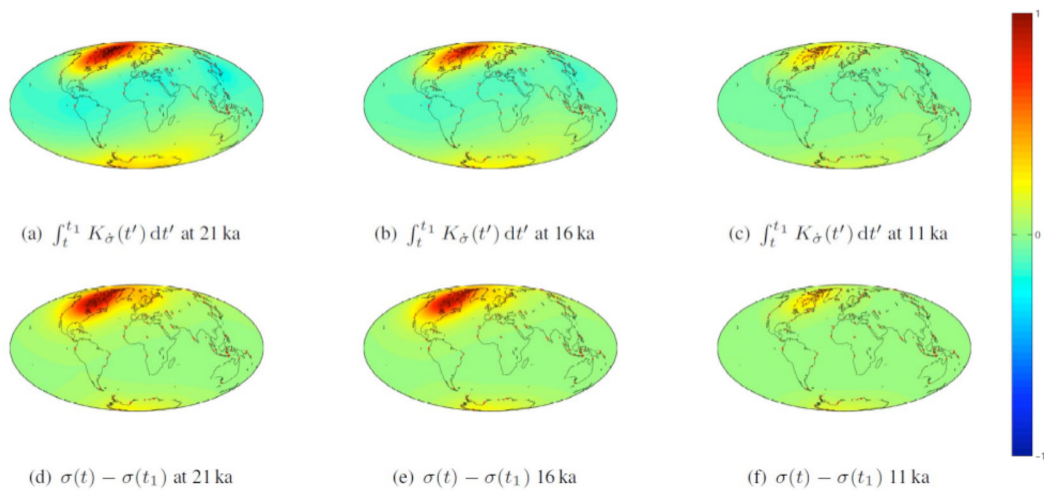


Figure 5. Comparison of the unscaled first ice load update $\int_t^{t_1} K_{\dot\sigma}(t')\,\mathrm{d}t'$ with true ice load $\sigma(t) - \sigma(t_1)$ used to generate the synthetic data set at times 21, 16 and 11 ka. As we are showing the model update prior to determining the scaling factor α in eq. (4.6), we do not show the absolute amplitudes of either of the ice loads. However, the relative amplitude of the loads at each of the three time slices is plotted on the same scale and so is directly comparable.

We note that the dependence of $K_{\dot\sigma}$ on the forward deformation fields $\mathbf{u}$ and ϕ arises solely from that of the adjoint loads. Consequently, simultaneous access to both the forward and adjoint deformations fields is not required in the calculation of the rate-of-loading kernel.

4.1.2 Rate-of-loading kernels for point observations

To illustrate the basic properties of the rate-of-loading kernels, we first consider point measurements in space and time of the deformation fields. In doing so, we let the objective functional be either

$$J(\mathbf{u}, \phi) = \hat{\mathbf{n}} \cdot \mathbf{u}(\mathbf{x}', t'), \tag{4.1}$$

or

$$J(\mathbf{u}, \phi) = -g^{-1}\phi(\mathbf{x}', t'), \tag{4.2}$$

which correspond, respectively, to the values of the vertical displacement and geoid anomaly at the point $\mathbf{x}' \in \partial M$ and time t'. Expressions for the adjoint loads $\mathbf{h}^\dagger$ and $h^\dagger$ corresponding to the objective functional in eq. (4.1) can be trivially found from eq. (C2) to be

$$\mathbf{h}^\dagger(\mathbf{x}, t) = -\delta(\mathbf{x} - \mathbf{x}')H(t' - t_1 + t - t_0)\hat{\mathbf{n}}, \quad h^\dagger(\mathbf{x}, t) = 0, \tag{4.3}$$

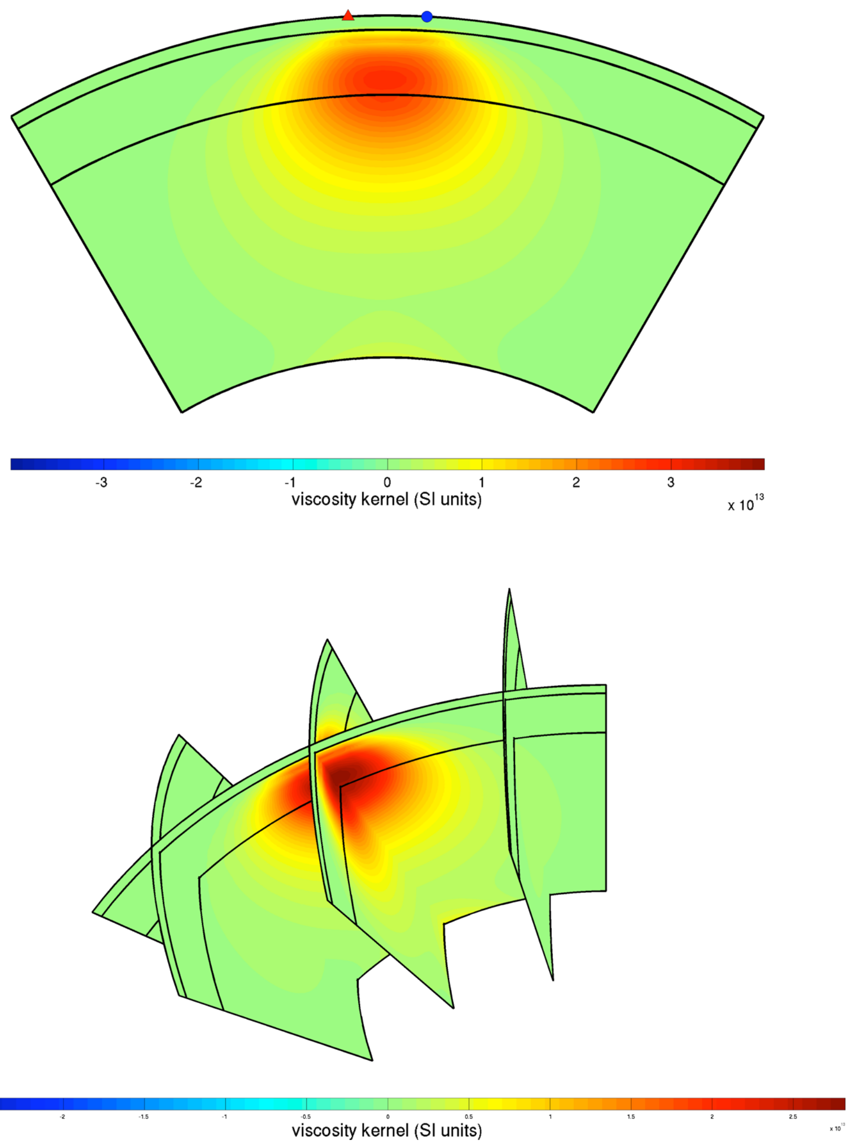


Figure 6. Viscosity sensitivity kernel corresponding to a point measurement of vertical displacement. In this calculation, the load was a delta function in space located at the red triangle seen on the upper figure, while its time dependence was that of a Heaviside step function with unit amplitude. The vertical displacement measurement was made 5000 yr after the onset of loading at the point indicated by the blue circle on the upper figure, which is separated from the loading point by an angular distance of 6°. For this calculation viscosity model 1 was used, and the calculations performed using spherical harmonic expansions up to degree 80. In the upper figure a radial section through the kernel is shown along the great circular arc between the load and observation point, with radii going from the CMB to the earth model's surface. The 670 km discontinuity and base of the elastic lithosphere are also displayed. In the lower figure, this radial section is plotted again along with three transverse radial sections that allow for the 3-D shape of the kernel to be better understood.

while those corresponding to the geoid measurement in eq. (4.2) are

$$\mathbf{h}^{\dagger}(\mathbf{x}, t) = \mathbf{0}, \quad h^{\dagger}(\mathbf{x}, t) = g^{-1}\delta(\mathbf{x} - \mathbf{x}')H(t' - t_1 + t - t_0). \tag{4.4}$$

In both cases the adjoint loads are independent of the forward deformation fields, and so calculation of the associated rate-of-loading kernels requires only the solution of the adjoint viscoelastic loading problem. Due to the spherical symmetry of the earth model, it is clear that for such measurements the rate-of-loading kernel will vary spatially only with angular distance from the observation point $\mathbf{x}'$. Similarly, the time dependence of $K_{\dot{\sigma}}$ is only through the difference $t - t'$, and due to causality we have $K_{\dot{\sigma}} = 0$ for $t > t'$.

In Fig. 2, we plot the rate-of-loading kernel corresponding to a point measurement of vertical displacement as a function of angular distance from the observation point $\mathbf{x}'$ for values of $t - t'$ between -50 and 0 kyr. At any particular time, it is seen that the sensitivity of the observation to the rate-of-loading $\dot{\sigma}$ is greatest directly beneath the observation point, and that this sensitivity decreases in amplitude rapidly with angular distance. At a fixed angular distance, we see from Fig. 2 that as $t - t'$ decreases the amplitude of the kernel increases at an exponentially decreasing rate, tending to a constant value as $t - t' \to -\infty$. This observation can readily be understood on physical grounds by considering the transition from elastic to viscous behaviour that occurs in the earth model.

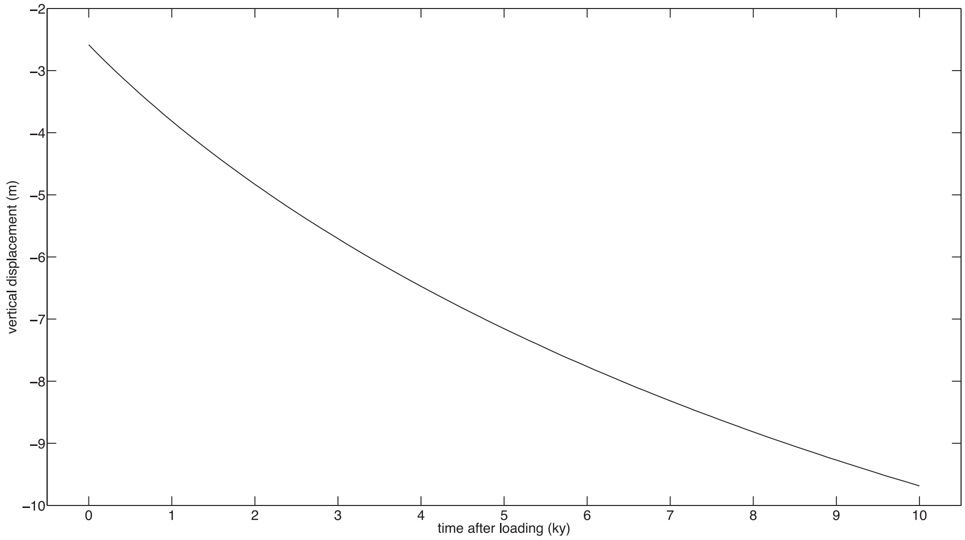


Figure 7. Vertical displacement recorded at the observation point for loading problem described in the caption of Fig. 6.

For these particular calculations, we used viscosity model 2, and expanded the adjoint loads in spherical harmonics up to degree $L = 80$ using the smoothed delta function expansion described in Appendix E. The value of maximum spherical harmonic degree does influence the spatial form of the rate-of-loading kernel, with higher values of L leading to kernels that are more narrowly distributed about the observation point. As explained in Appendix E, the choice of L amounts an *a priori* assumption on the spatial smoothness of the load σ.

In a similar manner, Fig. 3 plots the rate-of-loading kernel corresponding to a point measurement of geoid anomaly as a function of angular distance from the observation point $\mathbf{x}'$ for values of $t - t'$ between -50 and 0 kyr. It is again seen that at any particular time the sensitivity of the observation to the rate-of-loading $\dot{\sigma}$ is greatest directly beneath the observation point, and that this sensitivity decreases in amplitude with angular distance. At a fixed angular distance, however, we see that as $t - t'$ decreases the amplitude of the rate-of-loading kernel decreases at an exponentially decreasing rate. Again, this observation can be readily understood from the increasing degree to which the load is compensated during the transition from elastic to viscous behaviour in the earth model.

4.1.3 Rate-of-loading kernels for a vertical displacement data set

To examine the rate-of-loading kernels in a more elaborate example, we consider the objective functional defined in eq. (3.1). We recall that this objective functional measures the least-squares misfit between a data set comprising vertical displacement measurements and a corresponding set of synthetic vertical displacements calculated using given viscosity and load models. To generate such a data set, we performed a forward calculation over the period 21 to 0 ka using viscosity model 2 along with the ice sheet model ICE-5G of Peltier (2004) expanded in spherical harmonics up to degree 80. The vertical displacements were recorded at 100 randomly distributed coastal locations $\mathbf{x}_i$ shown in Fig. 4, and at each such location a randomly generated set of times t_{ij} for $j = 1, \ldots, N_t^i$ was constructed such that the average time interval between samples was 500 yr. To each of these data points d_{ij}, we added Gaussian random errors with zero mean and standard deviation equal to 0.1 m. Similarly, we added zero mean Gaussian random errors to each of the times t_{ij} with standard deviation equal to 100 yr.

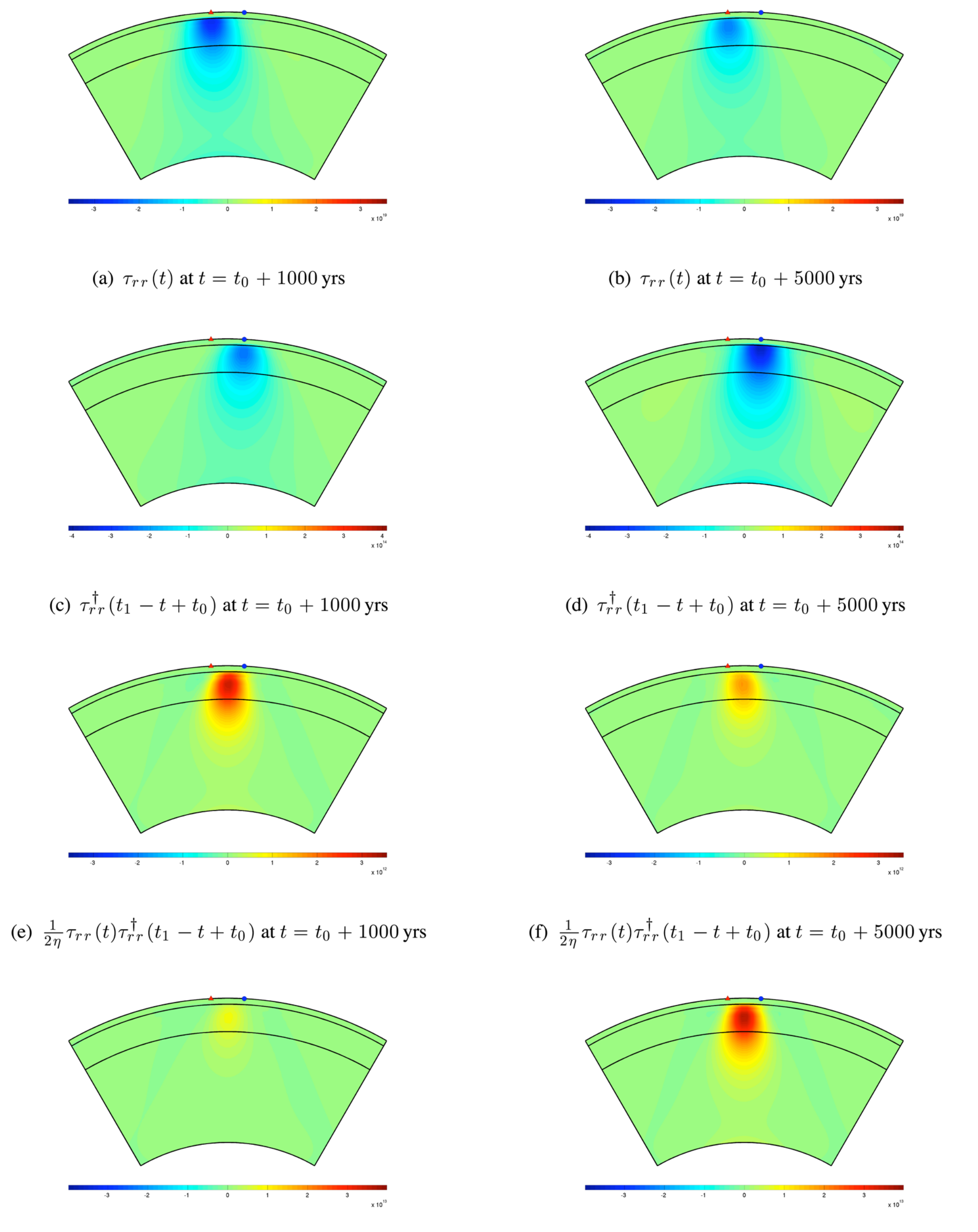
Having constructed such a data set, we can formulate an inverse problem to recover the ice load history, where for the moment we assume that the viscosity model is exactly known. As a starting point, we assume that the model times t'_{ij} coincide with the observed times t_{ij}, and that the initial ice load σ_0 is given by the final time slice of ICE-5G. With these assumptions, the synthetic data $\hat{\mathbf{n}} \cdot \mathbf{u}(\mathbf{x}_i, t'_{ij})$ are all equal to zero, and we can calculate the associated value of objective functional J and the adjoint loads $\mathbf{h}^\dagger$ and $h^\dagger$. We then solve the adjoint viscoelastic loading problem, and so obtain the rate-of-loading kernel $K_{\dot{\sigma}}$.

To interpret this kernel it will be useful, however, to consider how we would construct the first update to the ice load using the steepest descent algorithm. Within this method, the first update to the rate-of-loading is given by

$$\dot{\sigma}_1 = -\alpha K_{\dot{\sigma}}, \tag{4.5}$$

where α is some positive constant whose value is determined using a suitable line search algorithm (e.g. Nocedal & Wright 2006). The physical interpretation of $\dot{\sigma}_1$ is not, however, straightforward, and it is preferable to examine at the actual increment to the load

$$\sigma_1(t) = \sigma_0 + \alpha \int_t^{t_1} K_{\dot{\sigma}}(t')\,\mathrm{d}t', \tag{4.6}$$



(a) $\tau_{rr}(t)$ at $t = t_0 + 1000$ yrs

(b) $\tau_{rr}(t)$ at $t = t_0 + 5000$ yrs

(c) $\tau^{\dagger}_{rr}(t_1 - t + t_0)$ at $t = t_0 + 1000$ yrs

(d) $\tau^{\dagger}_{rr}(t_1 - t + t_0)$ at $t = t_0 + 5000$ yrs

(e) $\frac{1}{2\eta}\tau_{rr}(t)\tau^{\dagger}_{rr}(t_1 - t + t_0)$ at $t = t_0 + 1000$ yrs

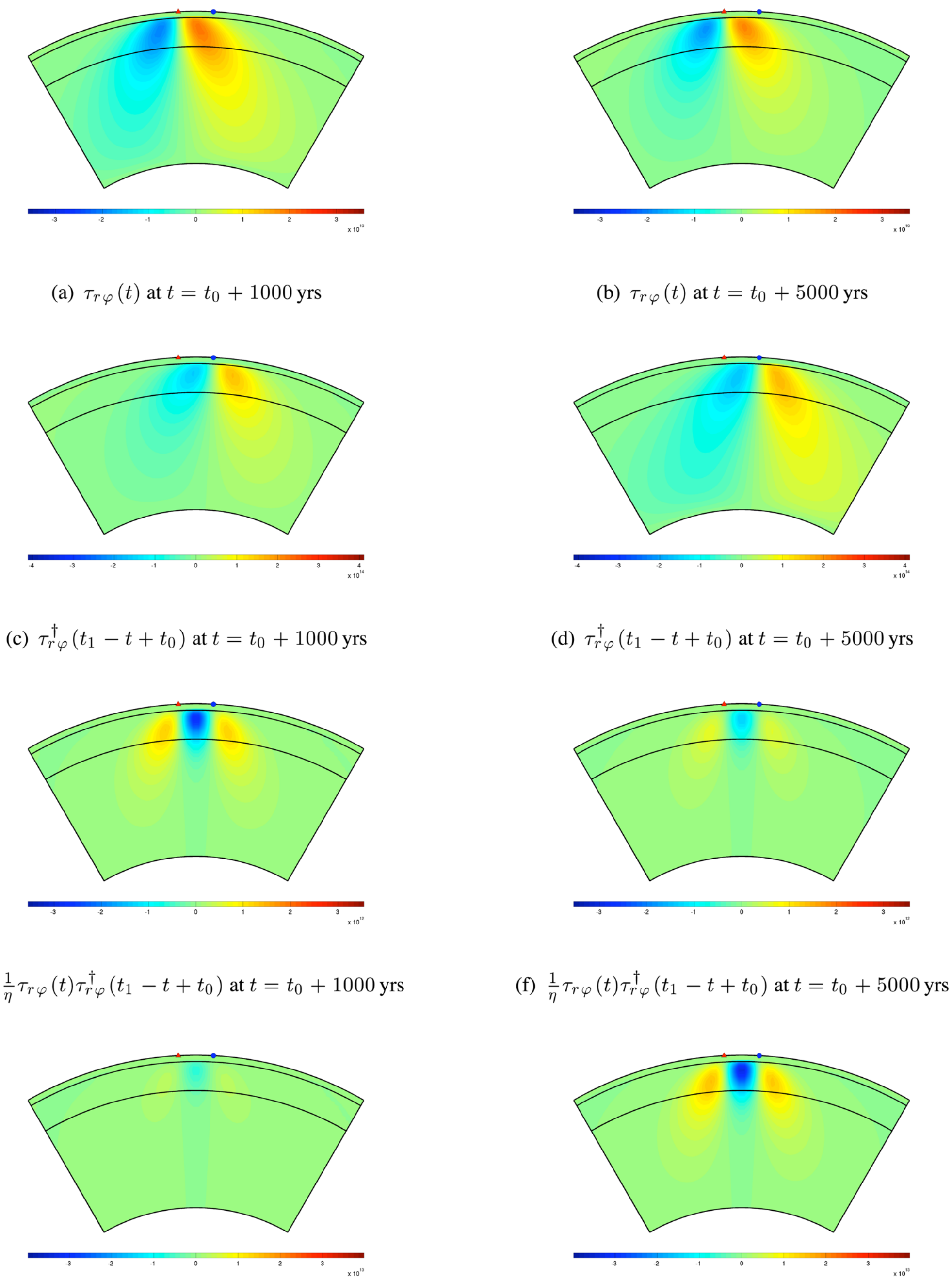
(f) $\frac{1}{2\eta}\tau_{rr}(t)\tau^{\dagger}_{rr}(t_1 - t + t_0)$ at $t = t_0 + 5000$ yrs

(g) $\int_{t_0}^{t} \frac{1}{2\eta}\tau_{rr}(t')\tau^{\dagger}_{rr}(t_1 - t' + t_0)\,\mathrm{d}t'$ at $t = t_0 + 1000$ yrs

(h) $\int_{t_0}^{t} \frac{1}{2\eta}\tau_{rr}(t')\tau^{\dagger}_{rr}(t_1 - t' + t_0)\,\mathrm{d}t'$ at $t = t_0 + 5000$ yrs

Figure 8. Stages in the construction of the viscosity sensitivity kernel shown in Fig. 6. In each of these figures, the load's position is denoted by a red triangle, while that of the observation point is shown by a blue circle. In the uppermost two figures, the τ_{rr} component of the deviatoric stress due to forward deformation field is shown at times (a) 1000 yr and (b) 5000 yr after the onset of loading. In these two figures the colour scale remains fixed so that relative changes in amplitude can be seen (this also the case for the other pairs of figures plotted below). On the next row is shown the $\tau^{\dagger}_{rr}$ component of the deviatoric stress due to the adjoint deformation field at the corresponding reversed times (c) 4000 yr and (d) 0 yr. The product of these two deviatoric stress fields weighted by $\frac{1}{2\eta}$ is then plotted on the third row, this field being equal to r–r contribution to the integrand in the expression for the viscosity kernel given in eq. (4.9). The lowermost pair of figures shows the partial contribution of the r–r component to the total viscosity obtained by numerically evaluating the integral in eq. (4.9) up to the specified time limits. Note that the colour scale in these final two figures is the same as in Fig. 6.

where we have used the condition $\sigma_1(t_1) = \sigma_0$ to fix the constant of integration. Writing σ for ice load used in generating the data set, we can then visually compare $\sigma(t) - \sigma(t_1)$ with $\sigma_1(t) - \sigma_1(t_1)$ to see how successful the first iteration has been in recovering the true ice load. In fact, we can simply ignore the scaling factor α and compare $\sigma(t) - \sigma(t_1)$ with $\int_t^{t_1} K_{\dot{\sigma}}(t')\,\mathrm{d}t'$, so long as we only consider the relative amplitudes of the two ice loads. Such a comparison can be seen in Fig. 5, where time slices of these two loads are plotted at 21, 16 and 11 ka. We note that within these figures the colour scale remains fixed between the subsequent plots of $\sigma(t) - \sigma(t_1)$ so that relative changes in amplitude



(a) $\tau_{r\varphi}(t)$ at $t = t_0 + 1000$ yrs

(b) $\tau_{r\varphi}(t)$ at $t = t_0 + 5000$ yrs

(c) $\tau^{\dagger}_{r\varphi}(t_1 - t + t_0)$ at $t = t_0 + 1000$ yrs

(d) $\tau^{\dagger}_{r\varphi}(t_1 - t + t_0)$ at $t = t_0 + 5000$ yrs

(e) $\frac{1}{\eta}\tau_{r\varphi}(t)\tau^{\dagger}_{r\varphi}(t_1 - t + t_0)$ at $t = t_0 + 1000$ yrs

(f) $\frac{1}{\eta}\tau_{r\varphi}(t)\tau^{\dagger}_{r\varphi}(t_1 - t + t_0)$ at $t = t_0 + 5000$ yrs

(g) $\int_{t_0}^{t} \frac{1}{\eta}\tau_{r\varphi}(t')\tau^{\dagger}_{r\varphi}(t_1 - t' + t_0)\,\mathrm{d}t'$ at $t = t_0 + 1000$ yrs

(h) $\int_{t_0}^{t} \frac{1}{\eta}\tau_{r\varphi}(t')\tau^{\dagger}_{r\varphi}(t_1 - t' + t_0)\,\mathrm{d}t'$ at $t = t_0 + 5000$ yrs

Figure 9. Stages in the construction of the viscosity sensitivity kernel shown in Fig. 6. In each of these figures, the load's position is denoted by a red triangle, while that of the observation point is shown by a blue circle. In the uppermost two figures, the $\tau_{r\varphi}$ component of the deviatoric stress due to forward deformation field is shown at times (a) 1000 yr and (b) 5000 yr after the onset of loading. In these two figures the colour scale remains fixed so that relative changes in amplitude can be seen (this also the case for the other pairs of figures plotted below). On the next row is shown the $\tau^{\dagger}_{r\varphi}$ component of the deviatoric stress due to the adjoint deformation field at the corresponding reversed times (c) 4000 yr and (d) 0 yr. The product of these two deviatoric stress fields weighted by $\frac{1}{2\eta}$ is then plotted on the third row, this field being equal to r–φ contribution to the integrand in the expression for the viscosity kernel given in eq. (4.9). The lowermost pair of figures shows the partial contribution of the r–φ component to the total viscosity obtained by numerically evaluating the integral in eq. (4.9) up to the specified time limits. Note that the colour scale in these final two figures is the same as in Fig. 6.

can be identified [and similarly for the plots of $\int_{t}^{t_1} K_{\dot{\sigma}}(t')\,\mathrm{d}t'$]. Given that we are looking at only the first update of an iterative scheme, the correspondence with the true model is actually quite good, with ice being added to the model in broadly correct locations. Furthermore, we see that the relative change in amplitude with time between the two models is of about the same magnitude. Had we continued this iterative process using steepest descent or some similar method, we would expect to gradually converge to an ice load that produces a good fit between

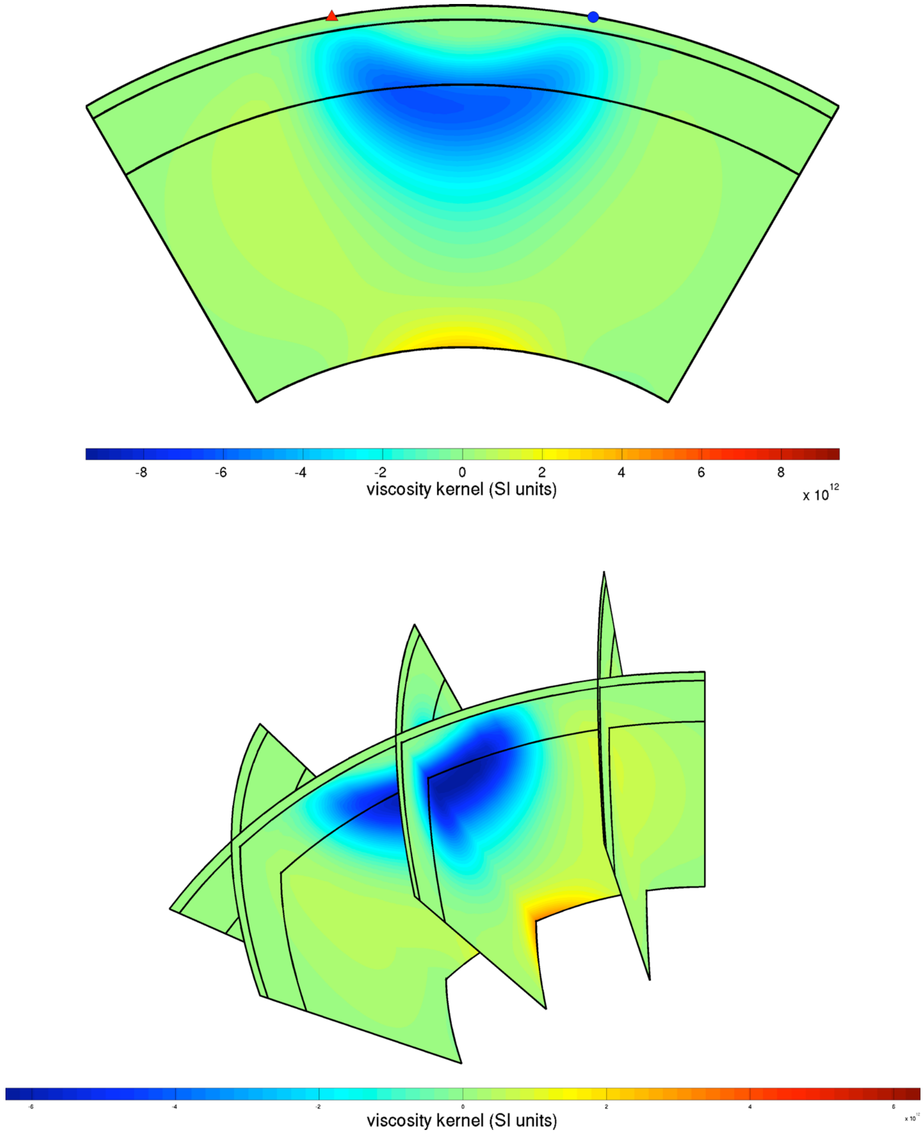


Figure 10. Viscosity sensitivity kernel corresponding to a point measurement of vertical displacement. In this calculation, the load was a delta function in space located at the red triangle seen on the upper figure, while its time dependence was that of a Heaviside step function with unit amplitude. The vertical displacement measurement was made 5000 yr after the onset of loading at the point indicated by the blue circle on the upper figure, which is separated from the loading point by an angular distance of 20°. For this calculation, viscosity model 1 was used, and the calculations performed using spherical harmonic expansions up to degree 80. In the upper figure, a radial section through the kernel is shown along the great circular arc between the load and observation point, with radii going from the CMB to the earth model's surface. The 670 km discontinuity and base of the elastic lithosphere are also displayed. In the lower figure, this radial section is plotted again along with three transverse radial sections that allow for the 3-D shape of the kernel to be better understood.

the given data and our synthetic predictions. Whether such an ice model would resemble our input model in detail is, of course, a more difficult question depending on the resolution in space and time obtainable from the given observations.

4.2 Viscosity sensitivity kernels

4.2.1 Computational details

The viscosity kernel K_η corresponding to a given objective functional J may be calculated in the following manner:

(1) Solve the forward viscoelastic loading problem to determine the deformation fields $\mathbf{u}$ and ϕ, and compute the associated deviatoric stress tensor $\boldsymbol{\tau}$.

(2) From $\mathbf{u}$ and ϕ calculate the adjoint loads $\mathbf{h}^\dagger$ and $h^\dagger$ associated with J through eqs (3.9), (3.17) and (3.18).

(3) Using these adjoint loads, solve the adjoint viscoelastic loading problem to obtain the adjoint deformation fields $\mathbf{u}^\dagger$ and $\phi^\dagger$, and compute the associated deviatoric stress tensor $\boldsymbol{\tau}^\dagger$.

(4) Compute the viscosity sensitivity kernel from $\boldsymbol{\tau}$, and $\boldsymbol{\tau}^\dagger$ using eq. (3.22).

In contrast to the rate-of-loading kernel, computation of the viscosity kernel requires simultaneous access to both the forward and adjoint deformation fields. In the calculations we have performed, it has been possible to store the entire time history of both the forward and adjoint deformation fields in memory, and so computation of the viscosity kernel using eq. (3.22) has been straightforward. Looking towards calculations in laterally heterogeneous earth models, memory requirements for storing the necessary stress fields will, of course, be higher. Given the relatively small number of time steps required in these calculations (as, say, compared to elastic wave propagation problems), it should, however, still be possible to store the required fields in memory (Hom-Nath Gharti, personal communication, 2013).

4.2.2 Viscosity kernels for point measurements of vertical displacement

To illustrate the basic properties of the viscosity kernel we again consider the objective functional

$$J(\mathbf{u}, \phi) = \hat{\mathbf{n}} \cdot \mathbf{u}(\mathbf{x}', t'), \tag{4.7}$$

which corresponds to the value of the vertical displacement at the point $\mathbf{x}' \in \partial M$ and time t'. In contrast to the case of the rate-of-loading kernel, to calculate the viscosity kernel we must explicitly specify the load that produced the forward deformation field. To keep the problem as simple as possible, we take this load to be

$$\sigma(\mathbf{x}, t) = \delta(\mathbf{x} - \mathbf{x}_s) H(t - t_s), \tag{4.8}$$

where $\mathbf{x}_s \in \partial M$ is the loading point, and t_s is the time at which the load is applied. Due to the spherical symmetry of the earth model, the form of the viscosity kernel does not depend upon the absolute positions of $\mathbf{x}'$ and $\mathbf{x}_s$, but only upon their angular separation. We emphasize, however, that even within spherical symmetric earth models the viscosity kernels are not spherically symmetric functions.

In Fig. 6, we plot such a viscosity kernel, with the angular separation between the load and observation points being equal to 6°, and the vertical displacement measurement made 5000 yr after the onset of loading. For this calculation we used viscosity model 1, and expanded all fields up to spherical harmonic degree 80. It is seen that the viscosity kernel has a broad region of positive sensitivity located in the upper portion of the mantle between the load and observation points and that its amplitude decays to zero away from this region. Moreover, the kernel is smooth, displaying no clear dependence upon discontinuities in the elastic structure of the earth model. In Fig. 7, we plot the vertical displacement at the observation point for 10 kyr following the onset of loading. It is seen that initially there is a small elastic subsidence which increases in amplitude at an exponentially decreasing rate due to viscous flow in the mantle. The positivity of the viscosity kernel implies that if we increased the mantle viscosity uniformly, then the vertical displacement at the observation point would be larger (i.e. less negative). This result makes physical sense, as a higher mantle viscosity would slow down the viscous relaxation, and so diminish the magnitude of the subsidence that could have accumulated prior to the observation time.

To develop intuition about the viscosity kernel in Fig. 6, it will be useful to consider how it is constructed through the combination of the forward and adjoint deviatoric stress fields $\boldsymbol{\tau}$ and $\boldsymbol{\tau}^{\dagger}$, respectively. Expressing these tensor fields in terms of their spherical polar components, we can rewrite eq. (3.22) as

$$K_{\eta} = \int_{t_0}^{t_1} \frac{1}{2\eta} \left(\tau_{rr}\tau^{\dagger}_{rr} + 2\tau_{r\theta}\tau^{\dagger}_{r\theta} + 2\tau_{r\varphi}\tau^{\dagger}_{r\varphi} + \tau_{\theta\theta}\tau^{\dagger}_{\theta\theta} + 2\tau_{\theta\varphi}\tau^{\dagger}_{\theta\varphi} + \tau_{\varphi\varphi}\tau^{\dagger}_{\varphi\varphi} \right) \mathrm{d}t, \tag{4.9}$$

where it is understood that the forward fields are evaluated at the time t, while the adjoint fields are evaluated at the reversed time $t_1 - t + t_0$. In Fig. 8, we have plotted the r–r components of $\boldsymbol{\tau}(t)$ and $\boldsymbol{\tau}^{\dagger}(t_1 - t + t_0)$ for $t = t_0 + 1000$ yr and $t = t_0 + 5000$ yr. Also plotted is the associated contribution to the integrand in eq. (4.9) at these times, along with the result of evaluating this integral up until the specified time limits. In Fig. 9, we similarly plot the contribution of the r–φ components to the viscosity kernel. In these figures we see that the stress components have high amplitudes in the upper mantle beneath their respective loads, and that their amplitude decays quite rapidly with increasing depth. The load and observation points are located quite close to one another in this example, and, as a result, the product of the stress components have highest amplitudes at shallow depths, and are concentrated in the area between load and observation points.

As a second example, we repeat the above calculations but now place the observation point at an angular distance of 20° from the load. The resulting sensitivity kernel is plotted in Fig. 10, and the vertical displacement time-series at the observation point can be seen in Fig. 11. The viscosity kernel is again seen to be concentrated in a region lying between the load and observation points, but in this case the sign of the kernel in the upper portion of the mantle is reversed and there is a diminished sensitivity to the viscosity structure in the uppermost mantle beneath the lithosphere. Furthermore, we see that there is a small positive sensitivity within the lower mantle, including an area lying just above the core–mantle boundary (CMB). The sign change of the kernel in the upper mantle can be understood from Fig. 11, where we see that the vertical displacement at the observation point has a more complex evolution in time. In particular, at the observation time of 5000 yr the rate of change of the vertical displacement is positive. Increasing the upper-mantle viscosity would slow the viscous relaxation, and so would in turn decrease the observed displacement. An interesting feature of this example is that the larger separation of the load and observation points leads the products of the stress components take more complex form, and to possess increased amplitudes at greater depths. In particular, we see how sensitivity to viscosity structures near the CMB can be generated even though neither the forward nor adjoint stress fields have large amplitudes in this region.

In the two examples described above, the calculations were performed using viscosity model 1 which, we recall, has a simple uniform mantle viscosity. Because the dependence of the deformation fields on the viscosity is non-linear, these kernel are influenced by the choice of

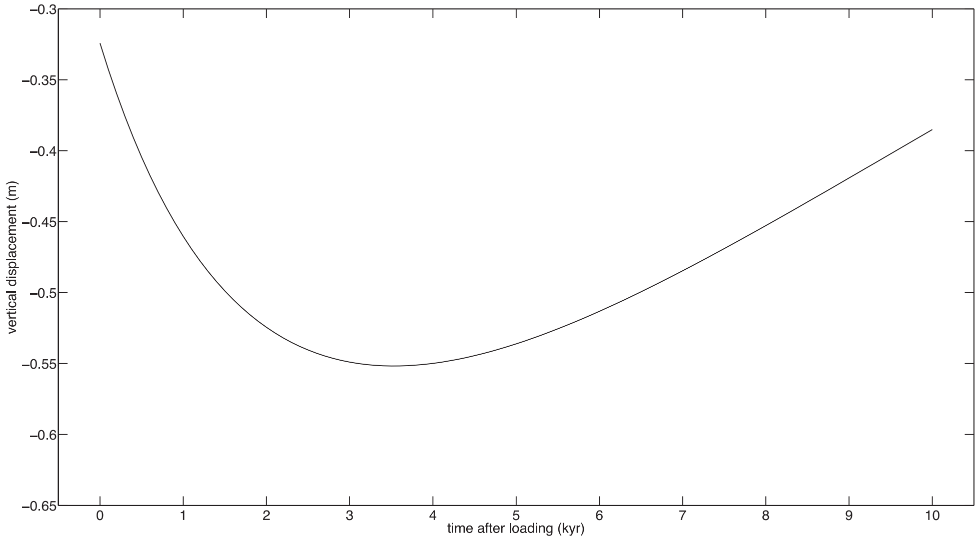
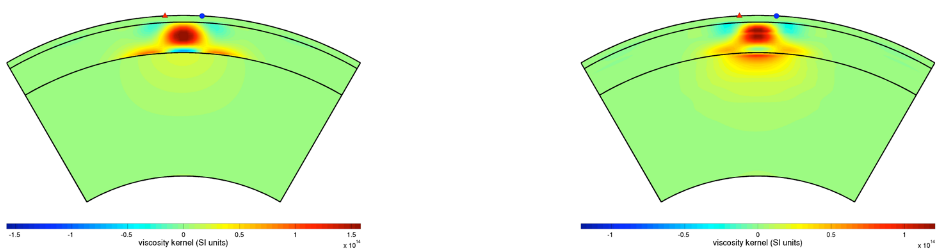
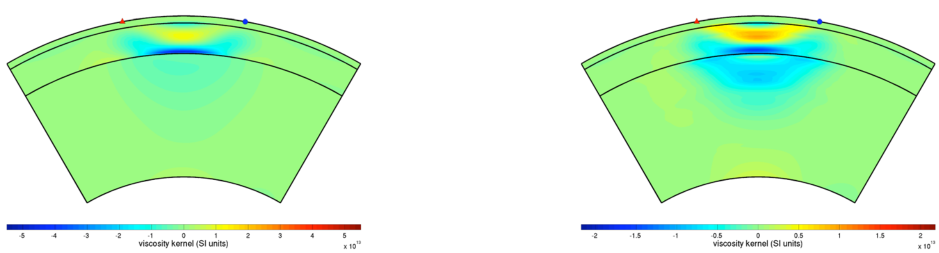


Figure 11. Vertical displacement recorded at the observation point for loading problem described in the caption of Fig. 10.



(a) Viscosity kernel of Fig.6 recalculated using viscosity model 2 (b) Viscosity kernel of Fig.6 recalculated using viscosity model 3

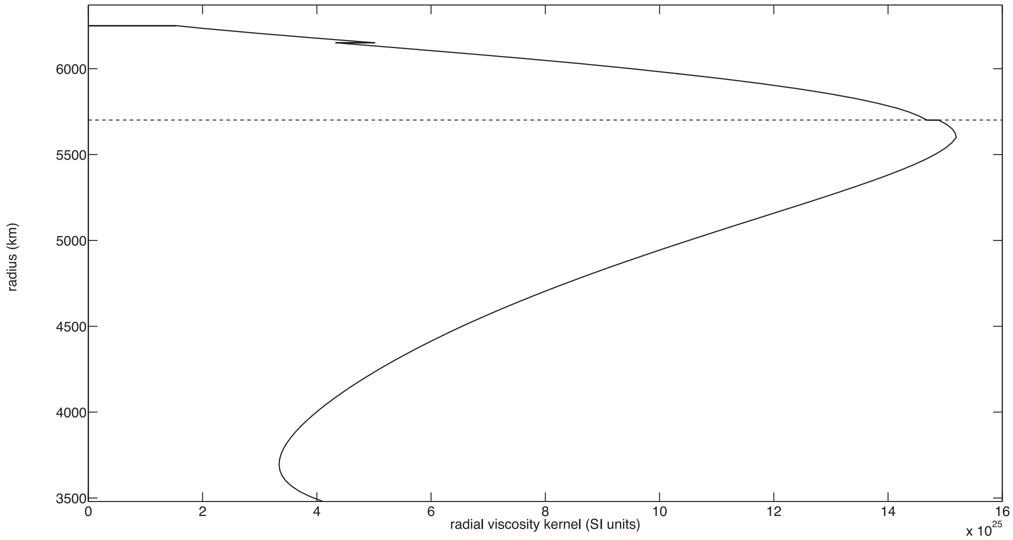


(c) Viscosity kernel of Fig. 10 recalculated using viscosity model 2 (d) Viscosity kernel of Fig.10 recalculated using viscosity model 3

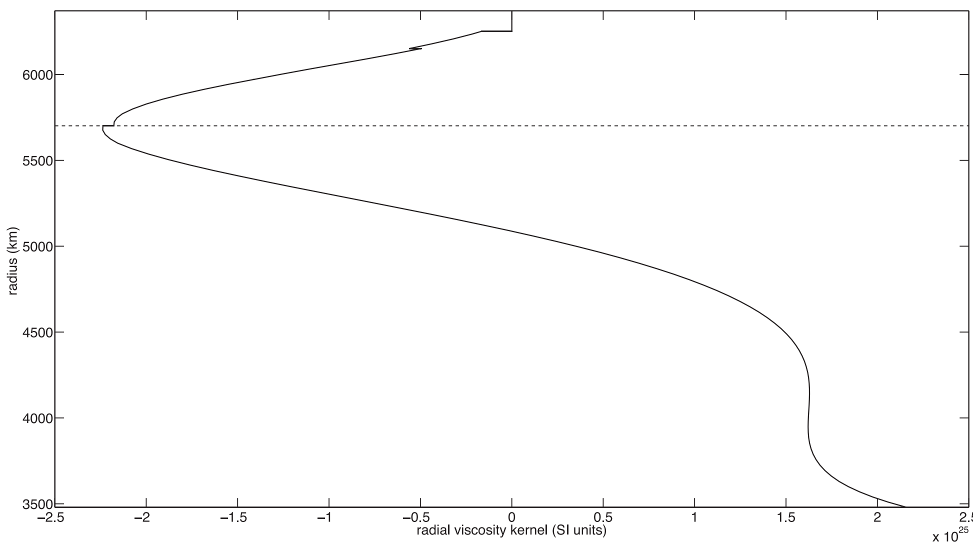
Figure 12. The strong dependence of the viscosity sensitivity kernel on the background viscosity model can be seen by recalculating the kernel corresponding to a particular measurement using different viscosity models. For these calculations, we replicated exactly the steps used to produced the kernels seen in Figs 6 and 10, with the only difference being the choice of viscosity model.

reference viscosity model. To conduct a preliminary investigation of this effect we have repeated the calculations done to produce the kernels in Figs 6 and 10 with all parameters remaining the same but for the viscosity model, which was taken to be either model 2 or model 3, as shown in Fig. 1. The viscosity kernels resulting from these calculations can be seen in Fig. 12. An important feature of both viscosity models 2 and 3 is the presence of a discontinuity at 670 km, which separates a low-viscosity upper mantle from a higher viscosity lower mantle. Physically, low viscosities within the upper mantle tend to concentrate flow within this region, and so leads to the increased sensitivity of surface observations to the upper-mantle viscosity structure, which can clearly be seen in Fig. 12.

In addition to the 3-D viscosity kernels K_η shown above, we can also calculate the corresponding radial viscosity kernels $\overline{K}_\eta$. For these calculations we make use of eq. (F3), which allows the radial kernels to be calculated in a very efficient manner. In Fig. 13, we plot the radial viscosity kernels corresponding to the volumetric kernels in Figs 6 and 10. The basic properties of these radial kernels can be readily understood by comparison with the volumetric kernels. Such radial viscosity kernels could be used in the construction of radial viscosity models from GIA observations as has been done previously using approximate kernels calculated using finite difference methods



(a) $\overline{K}_{\eta}$ corresponding to the kernel shown in Fig.6



(b) $\overline{K}_{\eta}$ corresponding to the kernel shown in Fig.10

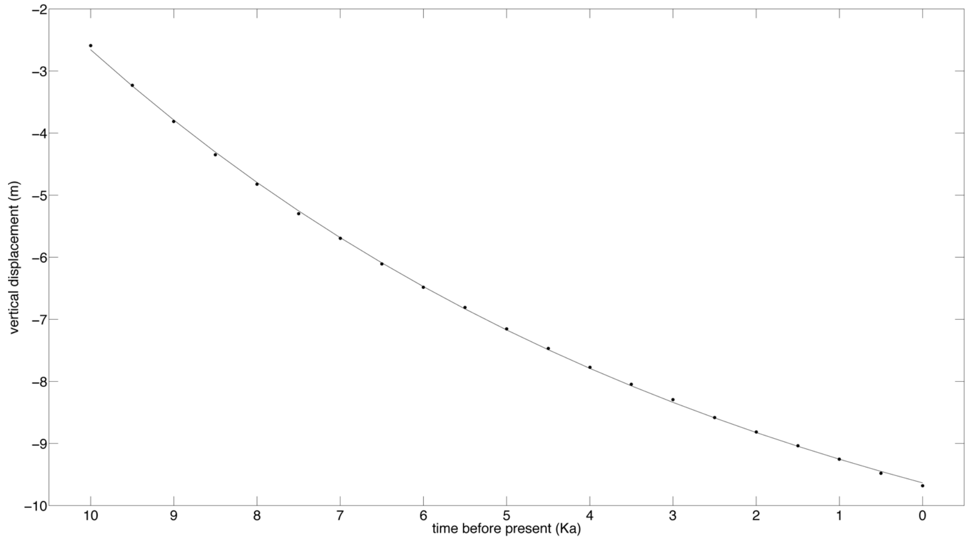
Figure 13. Radial viscosity sensitivity kernels corresponding to the volumetric kernels shown in Figs 6 and 10. These radial kernels were calculated from eq. (F3) using viscosity model 1 and with all fields again expanded in spherical harmonics up to degree 80.

(e.g. Mitrovica & Peltier 1991; Peltier 2004). The use of the adjoint method does, however, offer the advantage that the kernels are exact and can be calculated in an efficient manner requiring only one simulation of forward loading problem and one further simulation of the adjoint loading problem. This is in comparison to the calculation of finite difference kernels, which require repeated solution of the forward problem.

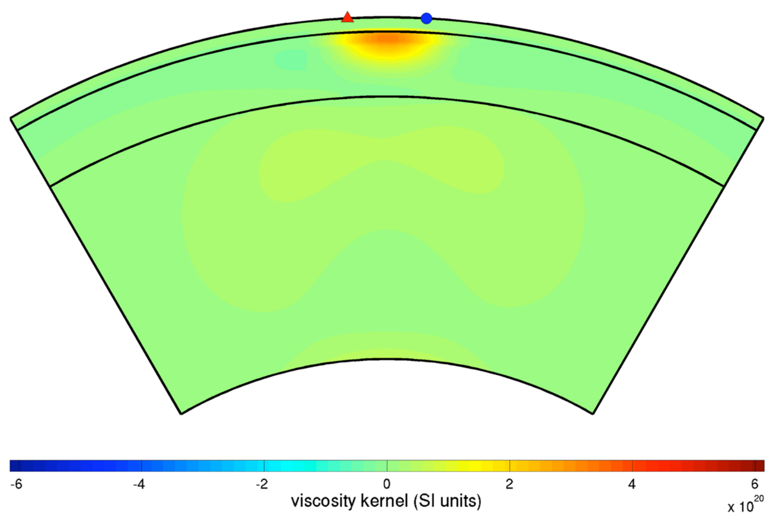
4.2.3 Kernels for relaxation times

As a final example, we consider viscosity kernels corresponding to measurements of relaxation times from rebound curves as described in example (5) of Appendix C. To do so, we make use of the point load defined in eq. (4.8) and generate the time-series of vertical displacements shown in Fig. 14(a). Here, the observation point located 6° away from the load, and the displacements are sampled every 500 yr for 10 000 yr after the onset of loading. In the calculation we used viscosity model 1, and again expanded all fields up to spherical harmonic degree 80.

From this time-series, we determined the best-fitting relaxation curve in the form of eq. (C12) using the method described in Appendix C. The resulting curve is also plotted on Fig. 14(a), and is found to have a relaxation time of $1/s = 8255$ yr (note that negligibly small random errors where added to the synthetic data, but these have essentially no effect on the obtained parameters a, b and s). Making use of eq. (C26),



(a) Measured vertical displacements and best-fitting exponential curve

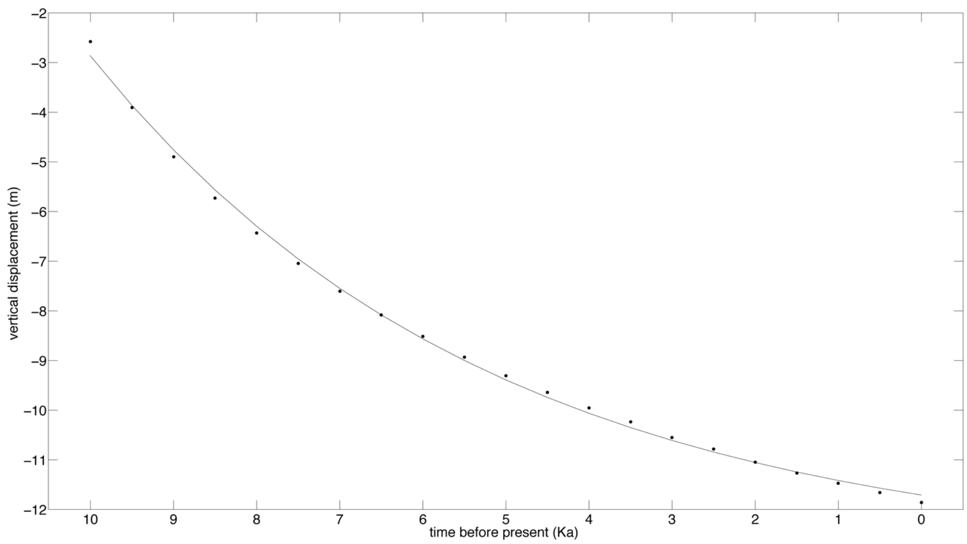


(b) Viscosity sensitivity kernel corresponding to the measured relaxation-time

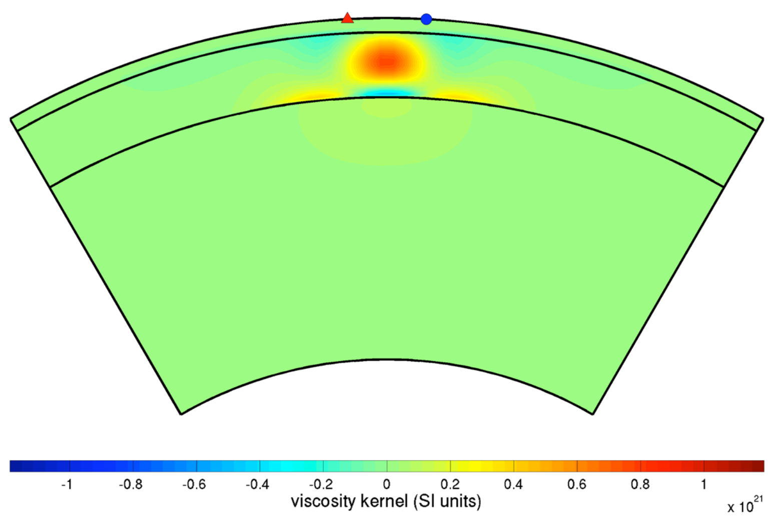
Figure 14. Viscosity sensitivity kernel corresponding to a relaxation time measurement calculated using viscosity model 1. For this calculation, the load was a delta function in space located at the red triangle seen on the lower figure, while its time dependence was that of a Heaviside step function with unit amplitude. Vertical displacements were recorded every 500 yr for 10 000 yr after the onset of loading at the point indicated by the blue circle on the lower figure, and are plotted as black circles on the uppermost figure. A best-fitting decaying exponential given by eq. (C12) was obtained, and the corresponding relaxation curve is shown on the upper figure with a solid line. For this problem, the best-fitting parameter were: $a = 2.96$ m, $b = -12.59$ m and $1/s = 8255$ yr. In the lower figure, the viscosity sensitivity kernel corresponding to the relaxation time measurement is plotted.

we then calculated the adjoint load for this measurement, and so determined the corresponding viscosity kernel, which is shown in shown in Fig. 14(b). It is seen that the kernel for the relaxation time measurement differs substantially from that for a point measurement of displacement shown in Fig. 6. In particular, this kernel is concentrated in the uppermost few hundred kilometres of the sublithospheric mantle. We also performed a similar calculation using viscosity model 2, with the results shown in Fig. 15. In both cases, the viscosity kernel is either solely or largely positive in sign, this being consistent with the physical expectation that larger viscosities increase the timescale over which the earth model responds to surface loading.

As discussed by Lambeck (2006), by fitting parametric curves to such observations we are not actually incorporating the physics of the problem, and so need not learn anything of value. However, a great virtue of the adjoint method is that it allows us to rigorously associate sensitivity kernels with complex data measurements for which there exists no simple physical relation to the underlying model. In the present case, these kernels show, in a linearized sense, what information such relaxation time measurements truly contain about the viscosity structure.



(a) Measured vertical displacements and best-fitting exponential curve



(b) Viscosity sensitivity kernel corresponding to the measured relaxation-time

Figure 15. As with Fig. 14, but here the calculations were performed in viscosity model 2, and the best-fitting parameters were: $a = 1.28$ m, $b = -12.99$ m and $1/s = 4834$ yr.

5 CONCLUSIONS

We have presented an application of the adjoint method to the surface loading of viscoelastic earth models. Using this method we can efficiently calculate sensitivity kernels that express the linearized sensitivity of a given objective functional to both mantle viscosity and the time derivative of the surface load. Calculation of these sensitivity kernels requires one solution of the forward viscoelastic loading problem and one solution of the adjoint viscoelastic loading problem. This is in stark contrast to finite difference approximations of the derivatives, which rapidly become impractical as the number of model parameters increases. Moreover, this adjoint problem is of exactly the same form as the forward problem and so can be solved using the same numerical methods. The derivation of the adjoint viscoelastic loading problem in eq. (3.20), and the corresponding expressions for the viscosity kernel K_η in eq. (3.22), and the rate-of-loading kernel $K_{\dot{\sigma}}$ in eq. (3.25) constitute our main results. With access to a suitable numerical code, this adjoint method could readily be applied to viscoelastic loading in laterally heterogeneous earth models.

As we have not taken into account gravitationally self-consistent ocean loading, nor rotational feedbacks, the theory presented in this paper is not directly applicable to the GIA inverse problem. An extension of the adjoint methods to this more complex problem is possible, and will be presented in a future paper. In carrying out this extension, a major role is played by the rate formulation of the viscoelastic loading problem employed here. In fact, it was the consideration of the full GIA problem that motivated the development of this rate formulation, and it is in this context that its full advantages will be seen.

ACKNOWLEDGEMENTS

During the majority of this project DA-A was supported through a Junior Research Fellowship at Merton College, Oxford. He would like to gratefully acknowledge advice from John Woodhouse, Jerry Mitrovica and Hom-Nath Gharti. Furthermore, John Woodhouse kindly provided the routine used to calculate generalized spherical harmonics.

REFERENCES

Adam, D., 2002. Gravity measurement: amazing grace, *Nature,* **416,** 10–11.

Al-Attar, D., 2011. Theoretical problems in global seismology and geodynamics, *DPhil thesis,* University of Oxford.

Backus, G.E., 1970. Inference from inadequate and inaccurate data, II, *Proc. Nat. Acad. Sci.,* **65,** 281–287.

Crossley, D.J. & Gubbins, D., 1975. Static deformation of the Earth's liquid core, *Geophys. Res. Lett.,* **2,** 1–4.

Dahlen, F.A., 1974. On the static deformation of an earth model with a fluid core, *Geophys. J. R. astr. Soc.,* **36,** 461–485.

Dahlen, F.A. & Tromp, J., 1998, *Theoretical Global Seismology,* Princeton University Press.

Dziewonski, A. & Anderson, D.L., 1981. Preliminary reference Earth model, *Phys. Earth planet. Inter.,* **25,** 297–356.

Fang, M. & Hager, B.H., 1995. The singularity mystery associated with a radially continuous Maxwell viscoelastic structure, *Geophys. J. Int.,* **123,** 849–865.

Farrell, W.E. & Clark, J.A., 1976. On post glacial sea level, *Geophys. J. R. astr. Soc.,* **46,** 647–667.

Fichtner, A. & Trampert, J., 2011. Resolution analysis in full waveform inversion, *Geophys. J. Int.,* **187,** 1604–1624.

Fichtner, A., Bunge, H.-P. & Igel, H., 2006. The adjoint method in seismology: I. Theory, *Phys. Earth. planet. Inter.,* **157,** 86–104.

Fichtner, A., Kennett, B.L.N., Igel, H. & Bunge, H.-P., 2010. Full waveform tomography for radially anisotropic structure: new insights into present and past states of the Australasian upper mantle, *Earth planet. Sci. Lett.,* **290,** 270–280.

Han, D. & Wahr, J., 1995. The viscoelastic relaxation of a realistically stratified earth, and a further analysis of postglacial rebound, *Geophys. J. Int.,* **120,** 287–311.

Hanyk, L., Moser, J., Yuen, D.A. & Matyska, C., 1995. Time-domain approach for the transient responses in stratified viscoelastic Earth models, *Geophys. Res. Lett.,* **22,** 1285–1288.

Lambeck, K., 2006. Hyperbolic tangents are no substitute for simple classical physics, *GFF,* **128**(4), 349–350.

Latychev, K., Mitrovica, J.X., Tromp, J., Tamisiea, M.E., Komatitsch, D. & Christara, C.C., 2005. Glacial isostatic adjustment on 3-D Earth models: a finite-volume formulation, *Geophys. J. Int.,* **161,** 421–444.

Lions, J.L., 1970. *Optimal Control of Systems Governed by Partial Differential Equations,* Springer-Verlag.

Liu, Q. & Tromp, J., 2008. Finite-frequency sensitivity kernels for global seismic wave propagation based upon adjoint methods, *Geophys. J. Int.,* **174,** 265–286.

Milne, G.A., Davis, J.L., Mitrovica, J.X., Schenrneck, H.-G., Johansson, J.M., Vermeer, M. & Koivula, H., 2001. Space-geodetic constraints on glacial isostatic adjustment in Fennoscandia, *Science,* **291,** 2381–2385.

Mitrovica, J.X. & Milne, G.A., 2003. On post-glacial sea level: I. General theory, *Geophys. J. Int.,* **154,** 253–267.

Mitrovica, J.X. & Peltier, W.R., 1991. A complete formalism for the inversion of post-glacial rebound data: resolving power analysis, *Geophys. J. Int.,* **104,** 267–288.

Mitrovica, J.X. & Peltier, W.R., 1993. The inference of mantle viscosity from an inversion of the Fennoscandian relaxation spectrum, *Geophys. J. Int.,* **114,** 45–62.

Nocedal, J. & Wright, S.J., 2006. *Numerical Optimization,* Springer-Verlag.

Parker, R.L., 1994. *Geophysical Inverse Theory,* Princeton University Press.

Peltier, W.R., 2004. Global glacial isostasy and the surface of the ice-age earth: the ICE-5G (VM2), *Annu. Rev. Earth planet. Sci.,* **32,** 111–149.

Phinney, R.A. & Burridge, R., 1973. Representation of the elastic-gravitational excitation of a spherical earth model by generalized spherical harmonics, *Geophys. J. R. astr. Soc.,* **34,** 451–487.

Press, W.H., Flannery, B.P., Teukolsky, S.A. & Vetterling, W.T., 1986. *Numerical Recipes,* Cambridge University Press.

Ritsema, J., van Heijst, H.J. & Woodhouse, J.H., 1999. Complex shear velocity structure imaged beneath Africa and Iceland, *Science,* **286,** 1925–1928.

Steffen, H., Wu, P. & Wang, H., 2012. Optimal locations for absolute gravity measurements and sensitivity of GRACE observations for constraining glacial isostatic adjustment on the northern hemisphere, *Geophys. J. Int.,* **190,** 1483–1494.

Tape, C., Liu, Q. & Tromp, J., 2007. Finite-frequency tomography using adjoint methods—methodology and examples using membrane surface waves, *Geophys. J. Int.,* **168,** 1105–1129.

Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2009. Adjoint tomography of the Southern California Crust, *Science,* **325,** 988–992.

Tröltzsch, F., 2005. *Optimal Control of Partial Differential Equations,* Springer-Verlag.

Tromp, J. & Mitrovica, J.X., 1999. Surface loading of a viscoelastic earth—I. General theory, *Geophys. J. Int.,* **137,** 847–855.

Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.,* **160,** 195–216.

Simo, J.C. & Hughes, T.J.R., 1998. *Computational Inelasticity,* Springer-Verlag.

Wu, P. & Peltier, W.R., 1982. Viscous gravitational relaxation, *Geophys. J. Int.,* **70,** 435–485.

Wu, P., Steffen, H. & Wang, H., 2010. Optimal locations for GPS measurements in North America and northern Europe for constraining glacial isostatic adjustment, *Geophys. J. Int.,* **181,** 653–664.

Zhong, S., Paulson, A. & Wahr, J., 2003. Three-dimensional finite-element modelling of Earth's viscoelastic deformation: effects of lateral variations in lithospheric thickness, *Geophys. J. Int.,* **155,** 679–695.

Zhu, H., Bozdağ, E., Peter, D. & Tromp, J., 2009. Structure of the European upper mantle revealed by adjoint tomography, *Nat. Geosci.,* **5,** 493–498.

APPENDIX A: JUMP CONDITIONS FOR THE RATE FORMULATION OF THE VISCOELASTIC LOADING PROBLEM

In the derivation of the rate formulation of the viscoelastic loading problem given in Section 2.3, we assumed that σ is a piecewise continuously differentiable function of time. In practice, it is, however, often useful to consider surface loads that possesses a finite number of jump discontinuities. Physically, we expect that the response of the earth model to such a step load will be instantaneously elastic, and this is indeed found to be the case. Letting t be a time at which σ undergoes the finite jump

$$\Delta\sigma(t) = \lim_{h\to 0}\left[\sigma(t+h) - \sigma(t-h)\right], \tag{A1}$$

we write $\Delta\mathbf{u}$, $\Delta\phi$ and $\Delta\mathbf{m}$, for the corresponding jumps in $\mathbf{u}$, ϕ and $\mathbf{m}$. Through a limiting argument it may be shown that

$$\Delta\mathbf{m} = \mathbf{0}, \tag{A2}$$

while $\Delta\mathbf{u}$ and $\Delta\phi$ must satisfy the static loading problem

$$-\nabla\cdot\left(\kappa\nabla\cdot\Delta\mathbf{u}\,\mathbf{I} + 2\mu_0\dot{\mathbf{d}}\right) + \nabla(\rho\Delta\mathbf{u}\cdot\nabla\Phi) - \nabla\cdot(\rho\Delta\mathbf{u})\nabla\Phi + \rho\nabla\Delta\phi = \mathbf{0}, \tag{A3}$$

and

$$(4\pi G)^{-1}\nabla^2\Delta\phi = \begin{cases} -\nabla\cdot(\rho\Delta\mathbf{u}) & \mathbf{x}\in M_{\mathrm{S}} \\ g^{-1}\Delta\phi\partial_n\rho & \mathbf{x}\in M_{\mathrm{F}} \\ 0 & \mathbf{x}\in\mathbb{R}^3\setminus M^{\mathrm{cl}} \end{cases}, \tag{A4}$$

which is subject to the boundary conditions

$$\hat{\mathbf{n}}\cdot(\kappa\nabla\cdot\Delta\mathbf{u}\,\mathbf{I} + 2\mu_0\Delta\mathbf{d}) = -\Delta\sigma\nabla\Phi, \quad \mathbf{x}\in\partial M, \tag{A5}$$

$$[\hat{\mathbf{n}}\cdot(\kappa\nabla\cdot\Delta\mathbf{u}\,\mathbf{I} + 2\mu_0\Delta\mathbf{d})]_-^+ = \mathbf{0}, \quad \mathbf{x}\in\Sigma_{\mathrm{SS}}, \tag{A6}$$

$$\hat{\mathbf{n}}\cdot\left(\kappa^+\nabla\cdot\Delta\mathbf{u}^+\,\mathbf{I} + 2\mu_0^+\Delta\mathbf{d}^+\right) = \rho^-[\Delta\mathbf{u}^+\cdot\nabla\Phi + \phi]\hat{\mathbf{n}}, \quad \mathbf{x}\in\Sigma_{\mathrm{FS}}, \tag{A7}$$

$$\hat{\mathbf{n}}\cdot\left(\kappa^-\nabla\cdot\Delta\mathbf{u}^-\,\mathbf{I} + 2\mu_0^-\Delta\mathbf{d}^-\right) = \rho^+[\Delta\mathbf{u}^-\cdot\nabla\Phi + \phi]\hat{\mathbf{n}}, \quad \mathbf{x}\in\Sigma_{\mathrm{SF}}, \tag{A8}$$

$$[\Delta\mathbf{u}]_-^+ = \mathbf{0}, \quad \mathbf{x}\in\Sigma_{\mathrm{SS}}, \tag{A9}$$

$$[\Delta\phi]_-^+ = 0, \quad \mathbf{x}\in\Sigma, \tag{A10}$$

$$\left[(4\pi G)^{-1}\hat{\mathbf{n}}\cdot\nabla\Delta\phi\right]_-^+ - \rho^-\hat{\mathbf{n}}\cdot\Delta\mathbf{u}^- = \Delta\sigma, \quad \mathbf{x}\in\partial M, \tag{A11}$$

$$\left[(4\pi G)^{-1}\hat{\mathbf{n}}\cdot\nabla\Delta\phi + \rho\hat{\mathbf{n}}\cdot\Delta\mathbf{u}\right]_-^+ = 0, \quad \mathbf{x}\in\Sigma_{\mathrm{SS}}, \tag{A12}$$

$$\left[(4\pi G)^{-1}\hat{\mathbf{n}}\cdot\nabla\Delta\phi\right]_-^+ + [\rho]_-^+\,\hat{\mathbf{n}}\cdot\Delta\mathbf{u}^+ = 0, \quad \mathbf{x}\in\Sigma_{\mathrm{FS}}, \tag{A13}$$

$$\left[(4\pi G)^{-1}\hat{\mathbf{n}}\cdot\nabla\Delta\phi\right]_-^+ + [\rho]_-^+\,\hat{\mathbf{n}}\cdot\Delta\mathbf{u}^- = 0, \quad \mathbf{x}\in\Sigma_{\mathrm{SF}}, \tag{A14}$$

along with the condition that $\Delta\phi \to 0$ as $\|\mathbf{x}\| \to \infty$. By combining these jump conditions with the evolution equations given in Section 2.3, we obtain the desired time-domain formulation of the viscoelastic loading problem in the case that the surface load is piecewise continuous.

APPENDIX B: DERIVATION OF THE WEAK FORM OF THE VISCOELASTIC LOADING PROBLEM

B1 Operator notation for the strong form of the loading problem

To obtain the weak form of the viscoelastic loading problem given in eq. (2.51), it will be useful to first restate the strong form of the problem in a concise operator notation. We begin by defining the partial differential operators

$$A_{11}\mathbf{u} = -\nabla\cdot(\kappa\nabla\cdot\mathbf{u}\,\mathbf{1} + 2\mu\mathbf{d}) + \nabla(\rho\mathbf{u}\cdot\nabla\Phi) - \nabla\cdot(\rho\mathbf{u})\nabla\Phi, \tag{B1}$$

$$A_{12}\phi = \rho\nabla\phi, \tag{B2}$$

$$A_{21}\mathbf{u} = \begin{cases} -\nabla\cdot(\rho\mathbf{u}) & \mathbf{x}\in M_{\mathrm{S}} \\ 0 & \mathbf{x}\notin M_{\mathrm{S}} \end{cases}, \tag{B3}$$

$$A_{22}\phi = \begin{cases} -(4\pi G)^{-1}\nabla^2\phi & \mathbf{x}\in \mathbb{R}^3\setminus M_{\mathrm{F}} \\ -(4\pi G)^{-1}\nabla^2\phi + g^{-1}\phi\partial_n\rho & \mathbf{x}\in M_{\mathrm{F}} \end{cases}, \tag{B4}$$

where $\mathbf{u}$ and ϕ are suitably regular vector and scalar fields defined, respectively, on M_{S} and $\mathbb{R}^3$ that satisfy the kinematic boundary conditions

$$[\mathbf{u}]_-^+ = \mathbf{0}, \quad \mathbf{x}\in\Sigma_{\mathrm{SS}}, \tag{B5}$$

$$[\phi]_-^+ = 0, \quad \mathbf{x}\in\Sigma, \tag{B6}$$

and $\mathbf{d}$ is the deviatoric strain tensor associated with $\mathbf{u}$. For such fields we also introduce the boundary operators

$$B_{11}\mathbf{u} = \begin{cases} \hat{\mathbf{n}}\cdot(\kappa\nabla\cdot\mathbf{u}\,\mathbf{1} + 2\mu\mathbf{d}) & \mathbf{x}\in\partial M \\ -\left[\hat{\mathbf{n}}\cdot(\kappa\nabla\cdot\mathbf{u}\,\mathbf{1}+2\mu\mathbf{d})\right]_-^+ & \mathbf{x}\in\Sigma_{\mathrm{SS}} \\ -\hat{\mathbf{n}}\cdot(\kappa^+\nabla\cdot\mathbf{u}^+\,\mathbf{1}+2\mu^+\mathbf{d}^+) + \rho^-\mathbf{u}^+\cdot\nabla\Phi\hat{\mathbf{n}} & \mathbf{x}\in\Sigma_{\mathrm{FS}} \\ \hat{\mathbf{n}}\cdot(\kappa^-\nabla\cdot\mathbf{u}^-\,\mathbf{1}+2\mu^-\mathbf{d}^-) - \rho^+\mathbf{u}^-\cdot\nabla\Phi\hat{\mathbf{n}} & \mathbf{x}\in\Sigma_{\mathrm{SF}} \end{cases}, \tag{B7}$$

$$B_{12}\phi = \begin{cases} \mathbf{0} & \mathbf{x}\in\partial M\cup\Sigma_{\mathrm{SS}} \\ \rho^-\phi\hat{\mathbf{n}} & \mathbf{x}\in\Sigma_{\mathrm{FS}} \\ -\rho^+\phi\hat{\mathbf{n}} & \mathbf{x}\in\Sigma_{\mathrm{SF}} \end{cases}, \tag{B8}$$

$$B_{21}\mathbf{u} = \begin{cases} \rho^-\mathbf{u}^-\cdot\hat{\mathbf{n}} & \mathbf{x}\in\partial M \\ -\left[\rho\mathbf{u}\cdot\hat{\mathbf{n}}\right]_-^+ & \mathbf{x}\in\Sigma_{\mathrm{SS}} \\ -\left[\rho\right]_-^+\mathbf{u}^+\cdot\hat{\mathbf{n}} & \mathbf{x}\in\Sigma_{\mathrm{FS}} \\ -\left[\rho\right]_-^+\mathbf{u}^-\cdot\hat{\mathbf{n}} & \mathbf{x}\in\Sigma_{\mathrm{SF}} \end{cases}, \tag{B9}$$

$$B_{22}\phi = -(4\pi G)^{-1}\left[\hat{\mathbf{n}}\cdot\nabla\phi\right]_-^+ \quad \mathbf{x}\in\Sigma, \tag{B10}$$

where, as above, superscripts $\pm$ indicate on which side a boundary the enclosed term is evaluated. Finally, we define the operator acting on symmetric second-order tensor fields by

$$R\mathbf{m} = -\nabla\cdot\left(\frac{2\mu_0}{\tau}\mathbf{m}\right), \tag{B11}$$

and also the associated boundary operator

$$S\mathbf{m} = \begin{cases} \frac{2\mu_0^-}{\tau^-}\hat{\mathbf{n}}\cdot\mathbf{m} & \mathbf{x}\in\partial M \\ -\left[\frac{2\mu_0}{\tau}\hat{\mathbf{n}}\cdot\mathbf{m}\right]_-^+ & \mathbf{x}\in\Sigma_{\mathrm{SS}} \\ \frac{2\mu_0^+}{\tau^+}\hat{\mathbf{n}}\cdot\mathbf{m}^+ & \mathbf{x}\in\Sigma_{\mathrm{FS}} \\ -\frac{2\mu_0^-}{\tau^-}\hat{\mathbf{n}}\cdot\mathbf{m}^- & \mathbf{x}\in\Sigma_{\mathrm{SF}} \end{cases}. \tag{B12}$$

With these notations, it is a simple matter to verify that the strong form of the time-domain viscoelastic loading problem given in eqs (2.37) and (2.38) can be written

$$A_{11}\dot{\mathbf{u}} + A_{12}\dot{\phi} + R(\mathbf{m}-\mathbf{d}) = \mathbf{0}, \tag{B13}$$

$$A_{21}\dot{\mathbf{u}} + A_{22}\dot{\phi} = 0, \tag{B14}$$

$$\dot{\mathbf{m}} + \frac{1}{\tau}(\mathbf{m}-\mathbf{d}) = \mathbf{0}, \tag{B15}$$

while the boundary conditions for the problem can be similarly expressed as

$$B_{11}\dot{\mathbf{u}} + B_{12}\dot{\phi} + S(\mathbf{m}-\mathbf{d}) = -\dot{\sigma}\nabla\Phi, \tag{B16}$$

$$B_{21}\dot{\mathbf{u}} + B_{22}\dot{\phi} = -\dot{\sigma}, \tag{B17}$$

where for convenience we have extended σ to equal zero for $\mathbf{x} \notin \partial M$.

B2 Green's identities

Letting $\mathbf{u}'$ and ϕ' be test functions also satisfying the kinematic boundary conditions given in eqs (B5) and (B6), we shall now establish the following Green's identity

$$\mathcal{A}(\mathbf{u}, \phi \,|\, \mathbf{u}', \phi') = \int_{M_\mathrm{S}} A_{11}\mathbf{u}\cdot\mathbf{u}'\,\mathrm{d}V + \int_{M_\mathrm{S}} A_{12}\mathbf{u}\,\phi'\,\mathrm{d}V + \int_{M_\mathrm{S}} A_{21}\phi\,\mathbf{u}'\,\mathrm{d}V + \int_{\mathbb{R}^3} A_{22}\phi\,\phi'\,\mathrm{d}V + \int_{\Sigma} B_{11}\mathbf{u}\cdot\mathbf{u}'\,\mathrm{d}S + \int_{\Sigma} B_{12}\mathbf{u}\,\phi'\,\mathrm{d}S$$
$$+ \int_{\Sigma} B_{21}\phi\,\mathbf{u}'\,\mathrm{d}S + \int_{\Sigma} B_{22}\phi\,\phi'\,\mathrm{d}S, \tag{B18}$$

where $\mathcal{A}$ is the symmetric form defined in eq. (2.52). To obtain this result, it is useful to consider each of the volume integrals on the right-hand side of this identity in turn. First, using the above definitions we have

$$\int_{M_\mathrm{S}} A_{11}\mathbf{u}\cdot\mathbf{u}'\,\mathrm{d}V = -\int_{M_\mathrm{S}} \nabla\cdot(\kappa\nabla\cdot\mathbf{u}\,\mathbf{I} + 2\mu\mathbf{d})\cdot\mathbf{u}'\,\mathrm{d}V + \int_{M_\mathrm{S}} [\nabla(\rho\mathbf{u}\cdot\nabla\Phi) - \nabla\cdot(\rho\mathbf{u})\nabla\Phi]\cdot\mathbf{u}'\,\mathrm{d}V. \tag{B19}$$

Integrating the first term on the right-hand side of this equation by parts we find

$$-\int_{M_\mathrm{S}} \nabla\cdot(\kappa\nabla\cdot\mathbf{u}\,\mathbf{I} + 2\mu\mathbf{d})\cdot\mathbf{u}'\,\mathrm{d}V = \int_{M_\mathrm{S}} \kappa\nabla\cdot\mathbf{u}\,\nabla\cdot\mathbf{u}'\,\mathrm{d}V + \int_{M_\mathrm{S}} 2\mu\mathbf{d}:\mathbf{d}'\,\mathrm{d}V - \int_{\partial M} \hat{\mathbf{n}}\cdot(\kappa\nabla\cdot\mathbf{u}\,\mathbf{I} + 2\mu\mathbf{d})\cdot\mathbf{u}'\,\mathrm{d}S$$
$$+ \int_{\Sigma_\mathrm{SS}} [\hat{\mathbf{n}}\cdot(\kappa\nabla\cdot\mathbf{u}\,\mathbf{I} + 2\mu\mathbf{d})]_-^+\cdot\mathbf{u}'\,\mathrm{d}S + \int_{\Sigma_\mathrm{FS}} \hat{\mathbf{n}}\cdot(\kappa\nabla\cdot\mathbf{u}\,\mathbf{I} + 2\mu\mathbf{d})\cdot\mathbf{u}'\,\mathrm{d}S$$
$$- \int_{\Sigma_\mathrm{SF}} \hat{\mathbf{n}}\cdot(\kappa\nabla\cdot\mathbf{u}\,\mathbf{I} + 2\mu\mathbf{d})\cdot\mathbf{u}'\,\mathrm{d}S, \tag{B20}$$

and recalling the definition of the boundary operator B_{11}, we can alternatively write the surface integral terms to obtain

$$-\int_{M_\mathrm{S}} \nabla\cdot(\kappa\nabla\cdot\mathbf{u}\,\mathbf{I} + 2\mu\mathbf{d})\cdot\mathbf{u}'\,\mathrm{d}V = \int_{M_\mathrm{S}} \kappa\nabla\cdot\mathbf{u}\,\nabla\cdot\mathbf{u}'\,\mathrm{d}V + \int_{M_\mathrm{S}} 2\mu\mathbf{d}:\mathbf{d}'\,\mathrm{d}V - \int_{\Sigma} B_{11}\mathbf{u}\cdot\mathbf{u}'\,\mathrm{d}S + \int_{\Sigma_\mathrm{FS}} \rho^- g\hat{\mathbf{n}}\cdot\mathbf{u}\,\hat{\mathbf{n}}\cdot\mathbf{u}'\,\mathrm{d}S$$
$$- \int_{\Sigma_\mathrm{SF}} \rho^+ g\hat{\mathbf{n}}\cdot\mathbf{u}\,\hat{\mathbf{n}}\cdot\mathbf{u}'\,\mathrm{d}S, \tag{B21}$$

where we have also made use of the relation $\nabla\Phi = g\hat{\mathbf{n}}$ for $\mathbf{x} \in \Sigma$. In dealing with the term

$$\int_{M_\mathrm{S}} [\nabla(\rho\mathbf{u}\cdot\nabla\Phi) - \nabla\cdot(\rho\mathbf{u})\nabla\Phi]\cdot\mathbf{u}'\,\mathrm{d}V, \tag{B22}$$

we first require the identity

$$\nabla(\rho\mathbf{u}\cdot\nabla\Phi) - \nabla\cdot(\rho\mathbf{u})\nabla\Phi = \rho\nabla(\mathbf{u}\cdot\nabla\Phi) + \nabla\rho\mathbf{u}\cdot\nabla\Phi - \rho\nabla\cdot\mathbf{u}\nabla\Phi - \nabla\rho\cdot\mathbf{u}\nabla\Phi$$
$$= \rho\left\{\nabla(\mathbf{u}\cdot\nabla\Phi) - \nabla\cdot\mathbf{u}\nabla\Phi\right\}, \tag{B23}$$

which follows from the hydrostatic condition that $\nabla\Phi$ and $\nabla\rho$ are parallel. From this result we obtain

$$[\nabla(\rho\mathbf{u}\cdot\nabla\Phi) - \nabla\cdot(\rho\mathbf{u})\nabla\Phi]\cdot\mathbf{u}' = \rho\,[\nabla(\mathbf{u}\cdot\nabla\Phi) - \nabla\cdot\mathbf{u}\nabla\Phi]\cdot\mathbf{u}'$$
$$= \nabla(\mathbf{u}\cdot\nabla\Phi)\cdot\rho\mathbf{u}' - \nabla\cdot\mathbf{u}\,\rho\nabla\Phi\cdot\mathbf{u}'$$
$$= \nabla\cdot\left(\rho\mathbf{u}\cdot\nabla\Phi\,\mathbf{u}' - \rho\mathbf{u}'\cdot\nabla\Phi\,\mathbf{u}\right) + \mathbf{u}\cdot\left[\nabla(\rho\mathbf{u}'\cdot\nabla\Phi) - \nabla\cdot(\rho\mathbf{u}')\nabla\Phi\right]$$
$$= \nabla\cdot\left(\rho\mathbf{u}\cdot\nabla\Phi\,\mathbf{u}' - \rho\mathbf{u}'\cdot\nabla\Phi\,\mathbf{u}\right) + \rho\mathbf{u}\cdot\left[\nabla(\mathbf{u}'\cdot\nabla\Phi) - \nabla\cdot\mathbf{u}'\nabla\Phi\right]. \tag{B24}$$

Integrating the first term of the final equality we find

$$\int_{M_\mathrm{S}} \nabla\cdot\left(\rho\mathbf{u}\cdot\nabla\Phi\,\mathbf{u}' - \rho\mathbf{u}'\cdot\nabla\Phi\,\mathbf{u}\right)\,\mathrm{d}V = -\int_{\Sigma} \left[\rho\mathbf{u}\cdot\nabla\Phi\,\mathbf{u}'\cdot\hat{\mathbf{n}} - \rho\mathbf{u}'\cdot\nabla\Phi\,\mathbf{u}\cdot\hat{\mathbf{n}}\right]_-^+\,\mathrm{d}S, \tag{B25}$$

but the hydrostatic condition implies that $\nabla\Phi = g\hat{\mathbf{n}}$, so that this integral vanishes, and we obtain

$$\int_{M_\mathrm{S}} [\nabla(\rho\mathbf{u}\cdot\nabla\Phi) - \nabla\cdot(\rho\mathbf{u})\nabla\Phi]\cdot\mathbf{u}'\,\mathrm{d}V = \int_{M_\mathrm{S}} \rho\mathbf{u}\cdot\left[\nabla(\mathbf{u}'\cdot\nabla\Phi) - \nabla\cdot\mathbf{u}'\nabla\Phi\right]\,\mathrm{d}V. \tag{B26}$$

We can write the integrand on the right-hand side as

$$\begin{aligned}\rho\mathbf{u}\cdot\left[\nabla(\mathbf{u}'\cdot\nabla\Phi)-\nabla\cdot\mathbf{u}'\nabla\Phi\right]&=\mathbf{u}\cdot\left[\rho\nabla(\mathbf{u}'\cdot\nabla\Phi)-\rho\nabla\cdot\mathbf{u}'\nabla\Phi\right]\\&=\mathbf{u}\cdot\left[\nabla(\rho\mathbf{u}'\cdot\nabla\Phi)-\nabla\rho\mathbf{u}'\cdot\nabla\Phi-\nabla\cdot(\rho\mathbf{u}')\nabla\Phi+\nabla\rho\cdot\mathbf{u}'\nabla\Phi\right]\\&=\mathbf{u}\cdot\left[\nabla(\rho\mathbf{u}'\cdot\nabla\Phi)-\nabla\cdot(\rho\mathbf{u}')\nabla\Phi\right],\end{aligned}\tag{B27}$$

where the final equality again follows from the hydrostatic condition. We, therefore, obtain the identity

$$\int_{M_S}[\nabla(\rho\mathbf{u}\cdot\nabla\Phi)-\nabla\cdot(\rho\mathbf{u})\nabla\Phi]\cdot\mathbf{u}'\,\mathrm{d}V=\int_{M_S}\mathbf{u}\cdot\left[\nabla(\rho\mathbf{u}'\cdot\nabla\Phi)-\nabla\cdot(\rho\mathbf{u}')\nabla\Phi\right]\mathrm{d}V,\tag{B28}$$

which shows that the integral on the left-hand side is, in fact, symmetric in $\mathbf{u}$ and $\mathbf{u}'$. We can, alternatively, write this integral in an explicitly symmetric form

$$\int_{M_S}\{\nabla(\rho\mathbf{u}\cdot\nabla\Phi)-\nabla\cdot(\rho\mathbf{u})\nabla\Phi\}\cdot\mathbf{u}'\,\mathrm{d}V=\frac{1}{2}\int_{M_S}\rho\left[\nabla(\mathbf{u}\cdot\nabla\Phi)\cdot\mathbf{u}'+\mathbf{u}\cdot\nabla(\mathbf{u}'\cdot\nabla\Phi)\right]\mathrm{d}V-\frac{1}{2}\int_{M_S}\rho\left[\nabla\cdot\mathbf{u}\,\nabla\Phi\cdot\mathbf{u}'+\nabla\Phi\cdot\mathbf{u}\,\nabla\cdot\mathbf{u}'\right]\mathrm{d}V,\tag{B29}$$

where we have used eq. (B26) along with the corresponding equation with $\mathbf{u}$ and $\mathbf{u}'$ transposed. Combining eqs (B21) and (B29), we can now write

$$\begin{aligned}\int_{M_S}A_{11}\mathbf{u}\cdot\mathbf{u}'\,\mathrm{d}V+\int_{\Sigma}B_{11}\mathbf{u}\cdot\mathbf{u}'\,\mathrm{d}S=&\int_{M_S}\kappa\nabla\cdot\mathbf{u}\,\nabla\cdot\mathbf{u}'\,\mathrm{d}V+\int_{M_S}2\mu\mathbf{d}:\mathbf{d}'\,\mathrm{d}V+\frac{1}{2}\int_{M_S}\rho\left[\nabla(\mathbf{u}\cdot\nabla\Phi)\cdot\mathbf{u}'+\mathbf{u}\cdot\nabla(\mathbf{u}'\cdot\nabla\Phi)\right]\mathrm{d}V\\&-\frac{1}{2}\int_{M_S}\rho\left(\nabla\cdot\mathbf{u}\,\nabla\Phi\cdot\mathbf{u}'+\nabla\Phi\cdot\mathbf{u}\,\nabla\cdot\mathbf{u}'\right)\mathrm{d}V.+\int_{\Sigma_{FS}}\rho^{-}g\hat{\mathbf{n}}\cdot\mathbf{u}\,\hat{\mathbf{n}}\cdot\mathbf{u}'\,\mathrm{d}S-\int_{\Sigma_{SF}}\rho^{+}g\hat{\mathbf{n}}\cdot\mathbf{u}\,\hat{\mathbf{n}}\cdot\mathbf{u}'\,\mathrm{d}S.\end{aligned}\tag{B30}$$

Turning now to the third term on the right-hand side of eq. (B18), we find using integration by parts

$$\begin{aligned}\int_{M_S}A_{21}\mathbf{u}\,\phi'\,\mathrm{d}V&=\int_{M_S}\rho\mathbf{u}\cdot\nabla\phi'\,\mathrm{d}V+\int_{\Sigma}[\rho\hat{\mathbf{n}}\cdot\mathbf{u}]_{-}^{+}\,\phi'\,\mathrm{d}S\\&=\int_{M_S}\mathbf{u}\cdot A_{12}\phi'\,\mathrm{d}V-\int_{\Sigma}B_{21}\mathbf{u}\,\phi'\,\mathrm{d}S+\int_{\Sigma}\mathbf{u}\cdot B_{12}\phi'\,\mathrm{d}S,\end{aligned}\tag{B31}$$

where in obtaining the final equality we have recalled the definition of the boundary operators B_{12} and B_{21}. Using this result we can now write

$$\begin{aligned}\int_{M_S}A_{21}\mathbf{u}\,\phi'\,\mathrm{d}V+\int_{\Sigma}B_{21}\mathbf{u}\,\phi'\,\mathrm{d}S+\int_{M_S}\mathbf{u}\cdot A_{12}\phi'\,\mathrm{d}V+\int_{\Sigma}\mathbf{u}\cdot B_{12}\phi'\,\mathrm{d}S=&\int_{M_S}\rho(\nabla\phi\cdot\mathbf{u}'+\mathbf{u}\cdot\nabla\phi')\,\mathrm{d}V+\int_{\Sigma_{FS}}\rho^{-}(\phi\,\mathbf{u}'+\mathbf{u}\,\phi')\cdot\hat{\mathbf{n}}\,\mathrm{d}S\\&-\int_{\Sigma_{SF}}\rho^{+}(\phi\,\mathbf{u}'+\mathbf{u}\,\phi')\cdot\hat{\mathbf{n}}\,\mathrm{d}S.\end{aligned}\tag{B32}$$

Finally, considering the fourth term on the right-hand side of eq. (B18) and using integration by parts we find

$$\int_{\mathbb{R}^3}A_{22}\phi\,\phi'\,\mathrm{d}V+\int_{\Sigma}B_{22}\phi\,\phi'\,\mathrm{d}S=\frac{1}{4\pi G}\int_{\mathbb{R}^3}\nabla\phi\cdot\nabla\phi'\,\mathrm{d}V+\int_{M_F}g^{-1}\phi\,\phi'\partial_n\rho\,\mathrm{d}V.\tag{B33}$$

Adding eqs (B30), (B32) and (B33), we then obtain the desired Green's identity.

B3 Derivation of the weak form

Using the above results, the derivation of the weak form of the viscoelastic loading problem proceeds as follows. From the Green's identity in eq. (B18), we can write

$$\begin{aligned}\mathcal{A}(\dot{\mathbf{u}},\dot{\phi}\,|\,\mathbf{u}',\phi')=&\int_{M_S}A_{11}\dot{\mathbf{u}}\cdot\mathbf{u}'\,\mathrm{d}V+\int_{M_S}A_{12}\dot{\mathbf{u}}\,\phi'\,\mathrm{d}V+\int_{M_S}A_{21}\dot{\phi}\,\mathbf{u}'\,\mathrm{d}V+\int_{\mathbb{R}^3}A_{22}\dot{\phi}\,\phi'\,\mathrm{d}V+\int_{\Sigma}B_{11}\dot{\mathbf{u}}\cdot\mathbf{u}'\,\mathrm{d}S+\int_{\Sigma}B_{12}\dot{\mathbf{u}}\,\phi'\,\mathrm{d}S\\&+\int_{\Sigma}B_{21}\dot{\phi}\,\mathbf{u}'\,\mathrm{d}S+\int_{\Sigma}B_{22}\dot{\phi}\,\phi'\,\mathrm{d}S,\end{aligned}\tag{B34}$$

where $\mathbf{u}$ and ϕ are solutions of the strong form of the viscoelastic loading problem, and $\mathbf{u}'$ and ϕ' are arbitrary test functions. Substituting eqs (B13) to (B17) into the right-hand side of eq. (B34) and making use of the readily verified identity

$$\int_{M_S}R(\mathbf{m}-\mathbf{d})\cdot\mathbf{u}'\,\mathrm{d}V+\int_{\Sigma}S(\mathbf{m}-\mathbf{d})\cdot\mathbf{u}'\,\mathrm{d}S=\int_{M_S}\frac{2\mu_0}{\tau}(\mathbf{m}-\mathbf{d}):\mathbf{d}'\,\mathrm{d}V,\tag{B35}$$

we obtain

$$\mathcal{A}(\dot{\mathbf{u}}, \dot{\phi} \,|\, \mathbf{u}', \phi') + \int_{M_S} \frac{2\mu_0}{\tau}(\mathbf{m} - \mathbf{d}) : \mathbf{d}' \, \mathrm{d}V + \int_{\partial M} (\nabla\Phi \cdot \mathbf{u}' + \phi')\dot{\sigma} \, \mathrm{d}S = 0. \tag{B36}$$

Considering now eq. (2.32), we see trivially that

$$\int_{M_S} 2\mu_0 \left[\dot{\mathbf{m}} + \frac{1}{\tau}(\mathbf{m} - \mathbf{d}) \right] : \mathbf{m}' \, \mathrm{d}V = 0, \tag{B37}$$

for any sufficiently regular test function $\mathbf{m}'$. Finally, we subtract eq. (B37) from eq. (B36) to obtain

$$\mathcal{A}(\dot{\mathbf{u}}, \dot{\phi} \,|\, \mathbf{u}', \phi') - \int_{M_S} 2\mu_0 \left[\dot{\mathbf{m}} : \mathbf{m}' + \frac{1}{\tau}(\mathbf{d} - \mathbf{m}) : (\mathbf{d}' - \mathbf{m}') \right] \mathrm{d}V + \int_{\partial M} (\nabla\Phi \cdot \mathbf{u}' + \phi')\dot{\sigma} \, \mathrm{d}S = 0, \tag{B38}$$

which must hold for all $\mathbf{u}'$, ϕ' and $\mathbf{m}'$, and constitutes the desired weak form of the viscoelastic loading problem. We note that though the inclusion of the factor of $2\mu_0$ in eq. (B37), and the choice of sign in combining eqs (B36) and (B37) were essentially arbitrary, they were made so as to maximize the symmetry of the equations, and, in particular, to insure that forward and adjoint loading problems have exactly the same form.

Repeating the above argument, we can also obtain a weak form of the jump conditions on $\mathbf{u}$ and ϕ given in Appendix A as

$$\mathcal{A}(\Delta\mathbf{u}, \Delta\phi \,|\, \mathbf{u}', \phi') + \int_{\partial M} (\nabla\Phi \cdot \mathbf{u}' + \phi')\Delta\sigma \, \mathrm{d}S = 0, \tag{B39}$$

which must hold for all $\mathbf{u}'$ and ϕ' at those times when the load σ undergoes a finite jump discontinuity. Similarly, in the adjoint viscoelastic loading problem described in Section 3.2 we note that if the adjoint loads $\mathbf{h}^\dagger$ or $h^\dagger$ undergo a finite jump discontinuities at a time t, then the weak form of this problem given in eq. (3.20) must be supplemented by $\Delta\mathbf{m}^\dagger = \mathbf{0}$ and the jump conditions

$$\mathcal{A}(\Delta\mathbf{u}^\dagger, \Delta\phi^\dagger \,|\, \delta\mathbf{u}, \delta\phi) - \int_{\partial M} (\Delta\mathbf{h}^\dagger \cdot \delta\mathbf{u} + \Delta h^\dagger \delta\phi) \, \mathrm{d}S = 0, \tag{B40}$$

which must hold for all $\delta\mathbf{u}$ and $\delta\phi$.

APPENDIX C: EXAMPLE OBJECTIVE FUNCTIONALS

The objective functional J used in the adjoint method can be very general. In this appendix, we describe a number of further examples, and give formulae for the associated Fréchet kernels $\mathbf{h}$ and h required in the construction of the adjoint loads.

(1) First, we consider a point measurement of vertical displacement made at a given location $\mathbf{x}' \in \partial M$ and time $t' \in [t_0, t_1]$. In this case, we have

$$J(\mathbf{u}, \phi) = \hat{\mathbf{n}} \cdot \mathbf{u}(\mathbf{x}', t'), \tag{C1}$$

and can then write δJ in the form of eq. (3.9) by taking

$$\mathbf{h}(\mathbf{x}, t) = -\delta(\mathbf{x} - \mathbf{x}')H(t' - t)\hat{\mathbf{n}}, \quad h(\mathbf{x}, t) = 0. \tag{C2}$$

(2) We now consider a point measurement of an arbitrary component of the surface velocity as could be measured using GPS networks deployed in formerly glaciated regions (e.g. Milne *et al.* 2001). In practice, we cannot, however, measure the velocity field directly, and must approximate its value using changes in displacement measurements over some finite time interval. Considering the simplest such approximation, we take as objective functional

$$J(\mathbf{u}, \phi) = \frac{1}{\Delta t} \hat{\boldsymbol{\nu}} \cdot \left[\mathbf{u}(\mathbf{x}', t' + \Delta t) - \mathbf{u}(\mathbf{x}', t') \right], \tag{C3}$$

where $\hat{\boldsymbol{\nu}}$ is a given unit vector in the measurement direction, $\mathbf{x}'$ is the observation point on ∂M, t' the time of the first displacement measurement and Δt the time interval between the two measurements. Corresponding to this objective functional, we obtain an expression for δJ in the form of eq. (3.9) by setting

$$\mathbf{h}(\mathbf{x}, t) = -\frac{1}{\Delta t}\delta(\mathbf{x} - \mathbf{x}') \left[H(t' + \Delta t - t) - H(t' - t) \right] \hat{\boldsymbol{\nu}}, \quad h(\mathbf{x}, t) = 0. \tag{C4}$$

Our use of a finite difference expression for the velocity in eq. (C3) is not only more representative of practical measurements, but is also convenient from a theoretical perspective. This is because the corresponding rate-of-loading kernel can be shown to be non-singular, while if we had instead included the actual point value of the velocity field in eq. (C3), the rate-of-loading kernel would possess a delta function like singularity at the observation time.

(3) Next, we consider two examples motivated by gravity measurements, such as those provided by the Gravity Recovery and Climate Experiment (GRACE) satellite (e.g. Adam 2002). First, we suppose that we have measured the spherical harmonic coefficient of degree l and order m of the perturbed gravitational potential ϕ at the time t'. We shall assume that this coefficient is defined with respect to the real

spherical harmonics $\mathcal{Y}_{lm}$ defined in appendix B.7 of Dahlen & Tromp (1998) as they are better suited to applications with real-valued fields. We can then write

$$\phi_{lm}(t') = \frac{1}{a^2}\int_{\partial M} \mathcal{Y}_{lm}(\theta,\varphi)\,\phi(a,\theta,\varphi,t')\,\mathrm{d}S, \tag{C5}$$

where, for simplicity, we have assumed that the surface of the earth model ∂M is a sphere with radius a. If we now take

$$J(\mathbf{u},\phi) = \phi_{lm}(t'), \tag{C6}$$

then we can write δJ in the form of eq. (3.9) by setting

$$\mathbf{h}(\mathbf{x},t) = \mathbf{0}, \quad h(\mathbf{x},t) = -\frac{1}{a^2}\mathcal{Y}_{lm}(\theta,\varphi)H(t'-t), \tag{C7}$$

where (θ,φ) are the usual spherical polar coordinates on ∂M.

(4) We can build on the previous example by considering a GRACE-like data set comprising discrete time-series $\{\phi_{lm,j}\}$ of spherical harmonic coefficients in the range of angular degrees $2 \le l \le L$ and times t_j for $j = 1,\ldots,N_t$. Relative to the timescale of GIA, the timing of such GRACE measurements are precisely known and so the t_j can be taken as error-free. The measured coefficients will, however, still be subject to errors, which we suppose to be uncorrelated in time, have zero mean, and possess a time-independent non-singular covariance matrix with components $C_{lm,l'm'}$. We can take as objective functional the quadratic misfit between the observed gravity coefficients $\phi_{lm,j}$ and the synthetic predictions $\phi_{lm}(t_j)$ calculated in a given model

$$J(\mathbf{u},\phi) = \frac{1}{2N_t(L+1)^2}\sum_{j=1}^{N_t}\sum_{lml'm'}[\phi_{lm}(t_j)-\phi_{lm,j}](C^{-1})_{lml'm'}[\phi_{l'm'}(t_j)-\phi_{l'm',j}], \tag{C8}$$

where $(C^{-1})_{lml'm',j}$ are components of the inverse covariance matrix, and for brevity we have left the index ranges in the spherical harmonic sums implicit. Perturbing ϕ in this expression, we find

$$\delta J = \frac{1}{N_t(L+1)^2}\left\{\sum_{j=1}^{N_t}\sum_{lml'm'}[\phi_{lm}(t_j)-\phi_{lm,j}](C^{-1})_{lml'm'}\delta\phi_{l'm'}(t_j)\right\}, \tag{C9}$$

where we have made use of the symmetry of the covariance matrix. From the results of the previous example, we can write δJ in the form of eq. (3.9) by taking

$$\mathbf{h}(\mathbf{x},t) = \mathbf{0}, \tag{C10}$$

$$h(\mathbf{x},t) = \frac{-1}{N_t(L+1)^2a^2}\sum_{j=1}^{N_t}\left\{\sum_{lml'm'}[\phi_{lm}(t_j)-\phi_{lm,j}](C^{-1})_{lml'm'}\mathcal{Y}_{l'm'}(\theta,\varphi)\right\}H(t_j-t). \tag{C11}$$

(5) As a final example, we consider a more elaborate objective functional whose dependence on the deformation fields $\mathbf{u}$ and ϕ is defined through a subsidiary optimization problem. Let $\mathbf{x}'$ be a location on ∂M at which we have recorded vertical displacements d_j at times t_j for $j = 1,\ldots,N_t$. In formerly glaciated areas, it has been observed that such rebound curves can in some circumstance be approximated parametrically by decaying exponentials of the form

$$d_j = a\mathrm{e}^{-st'_j} + b, \quad j = 1,\ldots,N_t, \tag{C12}$$

where a and b are amplitude parameters, s is an 'inverse decay time', and t'_j are model times that have again been introduced to account for timing errors in the data (e.g. Mitrovica & Peltier 1993). See Lambeck (2006) for an interesting discussion on the appropriateness of such measurements.

The best-fitting parameters can be found by minimizing the quadratic misfit functional

$$\chi = \frac{1}{2N_t}\sum_{j=1}^{N_t}\left\{\frac{1}{\Delta d_j^2}(a\mathrm{e}^{-st'_j}+b-d_j)^2 + \frac{1}{\Delta t_j^2}(t'_j-t_j)^2\right\}, \tag{C13}$$

where, for simplicity, we have assumed that the data are subject to uncorrelated, zero-mean, random errors with t_j and d_j having variances equal to Δt_j^2 and Δd_j^2, respectively. Introducing for convenience a $2N_t$-dimensional 'data vector' $\mathbf{d}$ and a (N_t+3)-dimensional 'parameter vector' $\mathbf{q}$ by

$$\mathbf{d} = \begin{pmatrix} d_1 \\ \vdots \\ d_{N_t} \\ t_0 \\ \vdots \\ t_{N_r}, \end{pmatrix}, \quad \mathbf{q} = \begin{pmatrix} a \\ b \\ s \\ t'_0 \\ \vdots \\ t'_{N_t}, \end{pmatrix}. \tag{C14}$$

we can then state this optimization problem as finding the $\mathbf{q}$ that minimizes the function $\mathbf{q} \mapsto \chi(\mathbf{q}, \mathbf{d})$ for given $\mathbf{d}$. If $\tilde{\mathbf{q}}$ is the solution to this problem, then we must have

$$(\nabla_{\mathbf{q}}\chi)(\tilde{\mathbf{q}}, \mathbf{d}) = \mathbf{0}, \tag{C15}$$

where $\nabla_{\mathbf{q}}\chi$ denotes the gradient of χ with respect to the parameter vector, and the above notation means that this gradient is evaluated at the point $(\tilde{\mathbf{q}}, \mathbf{d})$. Moreover, for this point to be a local minimum of χ, the Hessian $\nabla_{\mathbf{q}}\nabla_{\mathbf{q}}\chi$ must be positive-definite at $\tilde{\mathbf{q}}$. Such a best-fitting parameter vector can be determined iteratively from an initial guess $\mathbf{q}_0$, say, using Newton's method

$$\mathbf{q}_{i+1} = \mathbf{q}_i - \gamma \left[(\nabla_{\mathbf{q}}\nabla_{\mathbf{q}}\chi)(\mathbf{q}_i, \mathbf{d})\right]^{-1} (\nabla_{\mathbf{q}}\chi)(\mathbf{q}_i, \mathbf{d}), \tag{C16}$$

to produce a sequence of parameter vectors that converges to the desired $\tilde{\mathbf{q}}$. Here, $0 < \gamma \leq 1$ is damping parameter used to stabilize the iterations.

Alternatively, we can apply this exact process to synthetic data calculated for given model parameters $\{\sigma, \eta\}$. We can then regard a, b and s as being functions of the synthetic deformation fields $\mathbf{u}$ and ϕ and, in turn, the model parameters $\{\sigma, \eta\}$. For objective functional, we now take

$$J(\mathbf{u}, \phi) = f(\tilde{\mathbf{q}}), \tag{C17}$$

with f a given function and $\tilde{\mathbf{q}}$ the best-fitting parameter vector. In particular, we could take $f(\tilde{\mathbf{q}}) = 1/\tilde{s}$, which is equal to the best-fitting relaxation time for the rebound curve. For such an objective functional we now derive an expression for its first-order perturbation of the form of eq. (3.9).

Letting $\mathbf{d}$ be the data vector corresponding to the given synthetic displacement vector $\mathbf{u}$, we again write $\tilde{\mathbf{q}}$ for the best-fitting parameter vector. If we now perturb the deformation field by $\delta\mathbf{u}$, we can write the associated perturbation to the data vector as

$$\delta\mathbf{d} = \begin{pmatrix} \hat{\mathbf{n}} \cdot \delta\mathbf{u}(\mathbf{x}', t_0) \\ \vdots \\ \hat{\mathbf{n}} \cdot \delta\mathbf{u}(\mathbf{x}', t_{N_t}) \\ 0 \\ \vdots \\ 0, \end{pmatrix}. \tag{C18}$$

where the lower half of this vector comprises N_t terms equal to zero due to the times t_j remaining unchanged. This expression can alternatively be written in the form

$$\delta\mathbf{d} = \int_{t_0}^{t_1}\int_{\partial M} \mathbf{P}(\mathbf{x}, t)\delta\mathbf{u}(\mathbf{x}, t)\,\mathrm{d}S\,\mathrm{d}t, \tag{C19}$$

where $\mathbf{P}$ is a matrix-valued function given by

$$\mathbf{P}(\mathbf{x}, t) = -\sum_{j=1}^{N_t} \delta(\mathbf{x} - \mathbf{x}')H(t_j - t)\hat{\mathbf{e}}_j \otimes \hat{\mathbf{n}}, \tag{C20}$$

with $\hat{\mathbf{e}}_j$ an $2N_t$-dimensional vector whose jth component is equal to 1, while all others vanish. Corresponding to the perturbation in $\mathbf{d}$, the best-fitting parameter vector is modified so that

$$(\nabla_{\mathbf{q}}\chi)(\tilde{\mathbf{q}} + \delta\tilde{\mathbf{q}}, \mathbf{d} + \delta\mathbf{d}) = \mathbf{0}. \tag{C21}$$

Expanding this equation to first order in perturbed quantities, we obtain

$$\delta\tilde{\mathbf{q}} = -[(\nabla_{\mathbf{q}}\nabla_{\mathbf{q}}\chi)(\tilde{\mathbf{q}}, \mathbf{d})]^{-1}[(\nabla_{\mathbf{q}}\nabla_{\mathbf{d}}\chi)(\tilde{\mathbf{q}}, \mathbf{d})]\delta\mathbf{d}, \tag{C22}$$

which together with eq. (C19) implies that

$$\delta\tilde{\mathbf{q}} = \int_{t_0}^{t_1}\int_{\partial M} \mathbf{Q}(\mathbf{x}, t)\delta\mathbf{u}(\mathbf{x}, t)\,\mathrm{d}S\,\mathrm{d}t, \tag{C23}$$

where we have defined

$$\mathbf{Q}(\mathbf{x}, t) = -[(\nabla_{\mathbf{q}}\nabla_{\mathbf{q}}\chi)(\tilde{\mathbf{q}}, \mathbf{d})]^{-1}[(\nabla_{\mathbf{q}}\nabla_{\mathbf{d}}\chi)(\tilde{\mathbf{q}}, \mathbf{d})]\mathbf{P}(\mathbf{x}, t). \tag{C24}$$

Using eq. (C23), the desired expression for δJ can be obtained by setting

$$\mathbf{h}(\mathbf{x}, t) = -[\nabla_{\mathbf{q}} f(\tilde{\mathbf{q}})]\cdot[(\nabla_{\mathbf{q}}\nabla_{\mathbf{q}}\chi)(\tilde{\mathbf{q}}, \mathbf{d})]^{-1}[(\nabla_{\mathbf{q}}\nabla_{\mathbf{d}}\chi)(\tilde{\mathbf{q}}, \mathbf{d})]\mathbf{P}(\mathbf{x}, t), \quad h(\mathbf{x}, t) = 0. \tag{C25}$$

In particular, for the choice $f(\tilde{\mathbf{q}}) = 1/\tilde{s}$, we obtain

$$\mathbf{h}(\mathbf{x}, t) = \frac{1}{\tilde{s}^2}\hat{\mathbf{e}}_3\cdot[(\nabla_{\mathbf{q}}\nabla_{\mathbf{q}}\chi)(\tilde{\mathbf{q}}, \mathbf{d})]^{-1}[(\nabla_{\mathbf{q}}\nabla_{\mathbf{d}}\chi)(\tilde{\mathbf{q}}, \mathbf{d})]\mathbf{P}(\mathbf{x}, t), \quad h(\mathbf{x}, t) = 0, \tag{C26}$$

for the corresponding Fréchet kernel, where $\hat{\mathbf{e}}_3$ is an $(N_t + 3)$-dimensional vector whose third component is equal to 1 and all others equal zero.

APPENDIX D: NUMERICAL CALCULATIONS IN SPHERICALLY SYMMETRIC EARTH MODELS

In this appendix we describe the numerical implementation of our theory in a spherically symmetric earth model possessing a solid elastic inner core, a compressible inviscid fluid outer core, a viscoelastic mantle with Maxwell-solid rheology and an elastic lithosphere. In doing so, our starting point is the weak form of the viscoelastic loading problem given in eqs (2.51) and (B39). In spherically symmetric earth models, this problem can be simplified substantially by expanding the various fields in generalized spherical harmonics, and so obtaining a 'reduced weak form' of the loading problem in which there is complete decoupling of the radial expansion functions for each spherical harmonic of degree l and order m. There is, moreover, a further decoupling of the system into spheroidal and toroidal subsystems, with solution of the latter system not being required in the GIA forward problem. We then briefly describe a simple and efficient numerical method for the solution of the reduced weak form of the viscoelastic loading problem. This approach is based on the spatial discretization using a 1-D spectral-element method, while the time evolution of the system is determined using an explicit second-order Runge–Kutta integration scheme. Using this numerical method, we can obtain solutions to both the GIA forward and adjoint problems, and so determine sensitivity kernels for both rate-of-loading and mantle viscosity. The numerical code developed for solving the viscoelastic loading problem has been benchmarked against results presented in the literature.

D1 Generalized spherical harmonic expansions

We let (r, θ, φ) denote the usual spherical polar coordinates, $(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\varphi}})$ the associated unit basis vectors and recall the definition

$$\hat{\mathbf{e}}_- = \frac{1}{\sqrt{2}}(\hat{\boldsymbol{\theta}} - \mathrm{i}\hat{\boldsymbol{\varphi}}), \tag{D1}$$

$$\hat{\mathbf{e}}_0 = \hat{\mathbf{r}}, \tag{D2}$$

$$\hat{\mathbf{e}}_+ = -\frac{1}{\sqrt{2}}(\hat{\boldsymbol{\theta}} + \mathrm{i}\hat{\boldsymbol{\varphi}}), \tag{D3}$$

of the 'canonical basis vectors' given by Phinney & Burridge (1973). We write u^-, u^0 and u^+ for the so-called 'contravariant components' of $\mathbf{u}$ with respect to the canonical basis vectors, and note that they are related to the spherical polar components through

$$u_r = u^0, \tag{D4}$$

$$u_\theta = \frac{1}{\sqrt{2}}(u^- - u^+), \tag{D5}$$

$$u_\varphi = -\frac{\mathrm{i}}{\sqrt{2}}(u^- + u^+). \tag{D6}$$

These contravariant components can be expanded in the form

$$u^\alpha = \sum_{lm} u^\alpha_{lm} Y^\alpha_{lm}, \tag{D7}$$

where Y^N_{lm} are the fully normalized generalized spherical harmonics defined in appendix C of Dahlen & Tromp (1998), and summation is over integer values for $0 \le l \le \infty$ and $-l \le m \le l$. It will be useful to define an alternative set of coefficients functions for this vector through

$$u^{\pm}_{lm} = \frac{1}{\sqrt{2}} k (V_{lm} \pm \mathrm{i} W_{lm}), \tag{D8}$$

$$u^{0}_{lm} = U_{lm}, \tag{D9}$$

where

$$k = \sqrt{l(l+1)}, \tag{D10}$$

which correspond to the toroidal and spheroidal combinations of Phinney & Burridge (1973). Given these definitions, it is a simple matter to show that the spherical polar components of the displacement vector can be written

$$u_r = \sum_{lm} U_{lm} Y^0_{lm}, \tag{D11}$$

$$u_\theta = \sum_{lm} \frac{1}{2} k \left[(Y^{-1}_{lm} - Y^{+1}_{lm}) V_{lm} - \mathrm{i}(Y^{-1}_{lm} + Y^{+1}_{lm}) W_{lm} \right], \tag{D12}$$

$$u_\phi = -\sum_{lm} \frac{1}{2} k \left[\mathrm{i}(Y^{-1}_{lm} + Y^{+1}_{lm}) V_{lm} + (Y^{-1}_{lm} - Y^{+1}_{lm}) W_{lm} \right]. \tag{D13}$$

We note that these expansions could alternatively be written in using the ordinary spherical harmonics Y_{lm} along with their derivatives with respect to θ and ϕ (e.g. Woodhouse & Deuss 2007). For our numerical calculations, however, we find it easier to work directly with generalized spherical harmonics. In a similar manner, we can expand the perturbed gravitational potential as

$$\phi = \sum_{lm} \phi_{lm} Y^0_{lm}, \tag{D14}$$

and the applied surface σ load in the form

$$\sigma = \sum_{lm} \sigma_{lm} Y^0_{lm}. \tag{D15}$$

Although we wish only to deal with real-valued loads and deformation fields, the generalized spherical harmonics and associated expansion coefficients are complex valued. Using the identity

$$Y^{N*}_{lm} = (-1)^{m+N} Y^{-N}_{l-m}, \tag{D16}$$

given in eq. (C109) of Dahlen & Tromp (1998), it may be shown that the requirement that σ is real valued implies

$$\sigma^*_{lm} = (-1)^m \sigma_{l-m}. \tag{D17}$$

It follows that in specifying the load σ we need only give the values of its expansion coefficients σ_{lm} for $0 \le m \le l$. Similarly, we can use eq. (D16) to obtain the identities

$$U^*_{lm} = (-1)^m U_{l-m}, \quad V^*_{lm} = (-1)^m V_{l-m}, \quad \phi^*_{lm} = (-1)^m \phi_{l-m}. \tag{D18}$$

In obtaining the reduced weak form of the viscoelastic loading problem, we will require expressions for a number of spatial derivatives of the fields $\mathbf{u}$ and ϕ. Making use of the rules of contravariant differentiation (Phinney & Burridge 1973; Dahlen & Tromp 1998), it may be shown that

$$\nabla \cdot \mathbf{u} = \sum_{lm} r^{-1} \left[r \partial_r U_{lm} + 2U_{lm} - k^2 V_{lm} \right] Y^0_{lm}, \tag{D19}$$

and that the contravariant components of the deviatoric strain $\mathbf{d}$ associated with $\mathbf{u}$ can be expanded as

$$d^{\alpha\beta} = \sum_{lm} d^{\alpha\beta}_{lm} Y^{\alpha+\beta}_{lm}, \tag{D20}$$

where the expansion coefficients are given by

$$d^{\pm\pm}_{lm} = \frac{1}{2} k \sqrt{k^2 - 2}\, r^{-1} (V_{lm} \pm \mathrm{i} W_{lm}), \tag{D21}$$

$$d^{00}_{lm} = \frac{2}{3} r^{-1} \left(r \partial_r U_{lm} - U_{lm} + \frac{1}{2} k^2 V_{lm} \right), \tag{D22}$$

$$d_{lm}^{0\pm} = \frac{1}{2\sqrt{2}} k r^{-1} \left(r \partial_r V_{lm} - V_{lm} + U_{lm} \right) \pm \frac{\mathrm{i}}{2\sqrt{2}} k r^{-1} \left(r \partial_r W_{lm} - W_{lm} \right), \tag{D23}$$

$$d_{lm}^{\pm\mp} = \frac{1}{3} r^{-1} \left(r \partial_r U_{lm} - U_{lm} + \frac{1}{2} k^2 V_{lm} \right). \tag{D24}$$

Here, we note that

$$d_{lm}^{00} - 2 d_{lm}^{-+} = 0, \tag{D25}$$

in accordance with the requirement that **d** is symmetric and has zero trace. Similarly, we find that the gradient of the perturbed gravitational potential can be written

$$(\nabla \phi)^{\alpha} = \sum_{lm} \phi_{lm}^{|\alpha} Y_{lm}^{\alpha}, \tag{D26}$$

where the coefficients $\phi_{lm}^{|\alpha}$ are given by

$$\phi_{lm}^{|\pm} = \frac{1}{\sqrt{2}} k r^{-1} \phi_{lm}, \tag{D27}$$

$$\phi_{lm}^{|0} = \partial_r \phi_{lm}. \tag{D28}$$

Turning now to the internal variable **m**, we expand its contravariant components as

$$m^{\alpha\beta} = \sum_{lm} m_{lm}^{\alpha\beta} Y_{lm}^{\alpha+\beta}. \tag{D29}$$

Motivated by the above expressions for $d_{lm}^{\alpha\beta}$, we introduce a set of new coefficient functions M_{lm}, N_{lm}, R_{lm}, S_{lm} and T_{lm} through

$$m_{lm}^{\pm\pm} = \frac{1}{2} k \sqrt{k^2 - 2} r^{-1} \left(M_{lm} \pm \mathrm{i} N_{lm} \right), \tag{D30}$$

$$m_{lm}^{00} = \frac{2}{3} r^{-1} R_{lm}, \tag{D31}$$

$$m_{lm}^{0\pm} = \frac{1}{2\sqrt{2}} k r^{-1} \left(S_{lm} \pm \mathrm{i} T_{lm} \right), \tag{D32}$$

$$m_{lm}^{\pm\mp} = \frac{1}{3} r^{-1} R_{lm}, \tag{D33}$$

which from eq. (D16) can be seen to satisfy

$$M_{lm}^{*} = (-1)^m M_{l-m}, \quad R_{lm}^{*} = (-1)^m R_{l-m}, \quad S_{lm}^{*} = (-1)^m S_{l-m}. \tag{D34}$$

The spherical polar components of **m** are related to the contravariant components by

$$m_{rr} = m^{00}, \tag{D35}$$

$$m_{r\theta} = \frac{1}{\sqrt{2}} (m^{0-} - m^{0+}), \tag{D36}$$

$$m_{r\varphi} = -\frac{\mathrm{i}}{\sqrt{2}} (m^{0-} + m^{0+}), \tag{D37}$$

$$m_{\theta\theta} = \frac{1}{2} (m^{--} + m^{++}) - m^{-+}, \tag{D38}$$

$$m_{\theta\varphi} = -\frac{\mathrm{i}}{2} (m^{--} - m^{++}), \tag{D39}$$

where we have only written down five of the components due to the symmetry of **m** and the trace-free condition

$$m_{rr} + m_{\theta\theta} + m_{\varphi\varphi} = 0, \tag{D40}$$

and from eq. (D29), we obtain

$$m_{rr} = \sum_{lm} \frac{2}{3} r^{-1} R_{lm} Y^{0}_{lm}, \tag{D41}$$

$$m_{r\theta} = \sum_{lm} \frac{1}{4} k r^{-1} \left[S_{lm}(Y^{-1}_{lm} - Y^{+1}_{lm}) - \mathrm{i} T_{lm}(Y^{-1}_{lm} + Y^{+1}_{lm}) \right], \tag{D42}$$

$$m_{r\phi} = -\sum_{lm} \frac{1}{4} k r^{-1} \left[\mathrm{i} S_{lm}(Y^{-1}_{lm} + Y^{+1}_{lm}) + T_{lm}(Y^{-1}_{lm} - Y^{+1}_{lm}) \right], \tag{D43}$$

$$m_{\theta\theta} = \sum_{lm} r^{-1} \left\{ \frac{1}{4} k \sqrt{k^2 - 2} \left[M_{lm}(Y^{-2}_{lm} + Y^{+2}_{lm}) - \mathrm{i} N_{lm}(Y^{-2}_{lm} - Y^{+2}_{lm}) \right] - \frac{1}{3} R_{lm} Y^{0}_{lm} \right\}, \tag{D44}$$

$$m_{\theta\varphi} = -\sum_{lm} \frac{1}{4} k \sqrt{k^2 - 2} r^{-1} \left[\mathrm{i} M_{lm}(Y^{-2}_{lm} - Y^{+2}_{lm}) + N_{lm}(Y^{-2}_{lm} + Y^{+2}_{lm}) \right]. \tag{D45}$$

In computing the viscosity kernel from eq. (4.9), we must evaluate the deviatoric stress tensor $\boldsymbol{\tau}$, and we note that its spherical polar components can be calculated using eqs (D41)–(D45) if we make the substitutions

$$M_{lm} \to 2\mu_0 (V_{lm} - M_{lm}), \tag{D46}$$

$$R_{lm} \to 2\mu_0 (r\partial_r U_{lm} - U_{lm} + \frac{1}{2} k^2 V_{lm} - R_{lm}), \tag{D47}$$

$$S_{lm} \to 2\mu_0 (r\partial_r V_{lm} - V_{lm} + U_{lm} - S_{lm}), \tag{D48}$$

$$N_{lm} \to 2\mu_0 (W_{lm} - N_{lm}), \tag{D49}$$

$$T_{lm} \to 2\mu_0 (r\partial_r W_{lm} - W_{lm} - T_{lm}). \tag{D50}$$

D2 Reduction of the weak form to decoupled radial equations

We now derive the 'reduced weak form' of the viscoelastic loading problem. To do so, we start from the weak form of the problem given in eq. (2.51), and assume that the fields $\mathbf{u}$, ϕ and $\mathbf{m}$ have been expanded in generalized spherical harmonics in the manner described in the previous subsection. It will, however, be useful to write eq. (2.51) in the slightly different form

$$\mathcal{A}(\dot{\mathbf{u}}, \dot{\phi} \,|\, \mathbf{u}'^{*}, \phi'^{*}) - \int_{M_\mathrm{S}} 2\mu_0 \left[\dot{\mathbf{m}} : \mathbf{m}' + \frac{1}{\tau}(\mathbf{d} - \mathbf{m}) : (\mathbf{d}'^{*} - \mathbf{m}'^{*}) \right] \mathrm{d}V + \int_{\partial M} (\nabla\Phi \cdot \mathbf{u}'^{*} + \phi'^{*}) \dot{\sigma} \, \mathrm{d}S = 0, \tag{D51}$$

where $*$ denotes complex conjugation. It is clear that eq. (D51) holding for all test functions is equivalent to our previous statement of the weak form of the problem. The presence of complex conjugates in eq. (D51) is, however, very convenient due to their occurrence in the generalized spherical harmonic orthonormality relation

$$\int_{\Omega} Y^{N*}_{lm} Y^{N}_{l'm'} \, \mathrm{d}S = \delta_{ll'} \delta_{mm'}, \tag{D52}$$

where Ω denotes the unit two-sphere (e.g. Dahlen & Tromp 1998, eq. C110). We now select particular values for (l, m) and suppose that the test functions $\mathbf{u}'$, ϕ' and $\mathbf{m}'$ have contravariant components given by

$$u'^{\alpha} = u'^{\alpha}_{lm} Y^{\alpha}_{lm}, \tag{D53}$$

$$\phi' = \phi'_{lm} Y^{0}_{lm}, \tag{D54}$$

$$m'^{\alpha\beta} = m'^{\alpha\beta}_{lm} Y^{\alpha+\beta}_{lm}. \tag{D55}$$

It will be useful to again introduce additional coefficient functions for $\mathbf{u}'$ by

$$u'^{\pm}_{lm} = \frac{1}{\sqrt{2}} k (V'_{lm} \pm \mathrm{i} W'_{lm}), \tag{D56}$$

$$u'^{0}_{lm} = U'_{lm}, \tag{D57}$$

and for $\mathbf{m}'$ by

$$m'^{\pm\pm}_{lm} = \frac{1}{2}k\sqrt{k^2-2}r^{-1}\left(M'_{lm} \pm \mathrm{i}N'_{lm}\right), \tag{D58}$$

$$m'^{00}_{lm} = \frac{2}{3}r^{-1}R'_{lm}, \tag{D59}$$

$$m'^{0\pm}_{lm} = \frac{1}{2\sqrt{2}}kr^{-1}\left(S'_{lm} \pm \mathrm{i}T'_{lm}\right), \tag{D60}$$

$$m'^{\pm\mp}_{lm} = \frac{1}{3}r^{-1}R'_{lm}. \tag{D61}$$

Substituting these expansions into the bilinear form $\mathcal{A}$ and making use of eq. (D52), we find after a simple but lengthy calculation that

$$\begin{aligned}\mathcal{A}(\mathbf{u},\phi\,|\,\mathbf{u}'^*,\phi'^*) = &\int_{I_S}\kappa(r\partial_r U_{lm}+2U_{lm}-k^2V_{lm})(r\partial_r U'^*_{lm}+2U'^*_{lm}-k^2V'^*_{lm})\,\mathrm{d}r + \frac{4}{3}\int_{I_S}\mu_0(r\partial_r U_{lm}-U_{lm}\\ &+\frac{1}{2}k^2V_{lm})\left(r\partial_r U'^*_{lm}-U'^*_{lm}+\frac{1}{2}k^2V'^*_{lm}\right)\mathrm{d}r + k^2\int_{I_S}\mu_0(r\partial_r V_{lm}-V_{lm}+U_{lm})(r\partial_r V'^*_{lm}-V'^*_{lm}+U'^*_{lm})\,\mathrm{d}r\\ &+k^2\int_{I_S}\mu_0(r\partial_r W_{lm}-W_{lm})(r\partial_r W'^*_{lm}-W'^*_{lm})\,\mathrm{d}r + k^2(k^2-2)\int_{I_S}\mu_0(V_{lm}V'^*_{lm}+W_{lm}W'^*_{lm})\,\mathrm{d}r\\ &+\int_{I_S}4\rho(\pi G\rho r-g)U_{lm}U'^*_{lm}r\,\mathrm{d}r + k^2\int_{I_S}\rho g(V_{lm}U'^*_{lm}+U_{lm}V'^*_{lm})r\,\mathrm{d}r + \int_{I_S}\rho(\partial_r\phi_{lm}U'^*_{lm}+U_{lm}\partial_r\phi'^*_{lm})r^2\,\mathrm{d}r\\ &+k^2\int_{I_S}\rho(V_{lm}\phi'^*_{lm}+\phi_{lm}V'^*_{lm})r\,\mathrm{d}r + \frac{1}{4\pi G}\int_{I}(\partial_r\phi_{lm}\partial_r\phi'^*_{lm}r^2+k^2\phi_{lm}\phi'^*_{lm})\,\mathrm{d}r + \frac{1}{4\pi G}(l+1)a\phi_{lm}(a)\phi'^*_{lm}(a)\\ &+\int_{I_F}g^{-1}\partial_r\rho\phi_{lm}\phi'^*_{lm}r^2\,\mathrm{d}r + \sum_{r\in d_{FS}}\rho^-(r)\left[g(r)U_{lm}(r)U'^*_{lm}(r)+\phi_{lm}(r)U'^*_{lm}(r)+U_{lm}(r)\phi'^*_{lm}(r)\right]r^2\\ &-\sum_{r\in d_{SF}}\rho^+(r)\left[g(r)U_{lm}(r)U'^*_{lm}(r)+\phi_{lm}(r)U'^*_{lm}(r)+U_{lm}(r)\phi'^*_{lm}(r)\right]r^2.\end{aligned} \tag{D62}$$

In obtaining this result, we have made use of Poisson's equation

$$\partial_r^2\Phi + 2r^{-1}\partial_r\Phi = 4\pi G\rho, \tag{D63}$$

to eliminate $\partial_r^2\Phi$, which arises from the terms $\nabla(\mathbf{u}\cdot\nabla\Phi)$ and $\nabla(\mathbf{u}'\cdot\nabla\Phi)$ in the definition of $\mathcal{A}$. We have also employed a Dirichlet-to-Neumann map to write

$$\frac{1}{4\pi G}\int_{\mathbb{R}^3}\nabla\phi\cdot\nabla\phi'^*\,\mathrm{d}V = \frac{1}{4\pi G}\int_M\nabla\phi\cdot\nabla\phi'^*\,\mathrm{d}V + \frac{1}{4\pi G}\sum_{lm}(l+1)a\phi_{lm}(a)\phi'^*_{lm}(a), \tag{D64}$$

and so circumvent the need to explicitly define ϕ and ϕ' outside of the earth model (Chaljub & Valette 2004). In a similar manner, we obtain

$$\int_{M_S}2\mu_0\dot{\mathbf{m}}:\mathbf{m}'^*\,\mathrm{d}V = \frac{4}{3}\int_{I_S}\mu_0\dot{R}_{lm}R'^*_{lm}\,\mathrm{d}r + k^2\int_{I_S}\mu_0(\dot{S}_{lm}S'^*_{lm}+\dot{T}_{lm}T'^*_{lm})\,\mathrm{d}r + k^2(k^2-2)\int_{I_S}\mu_0(\dot{M}_{lm}M'^*_{lm}+\dot{N}_{lm}N'^*_{lm})\,\mathrm{d}r, \tag{D65}$$

$$\begin{aligned}\int_{M_S}\frac{2\mu_0}{\tau}(\mathbf{d}-\mathbf{m}):(\mathbf{d}'^*-\mathbf{m}'^*)\,\mathrm{d}V = &\frac{4}{3}\int_{I_S}\frac{\mu_0}{\tau}\left(r\partial_r U_{lm}-U_{lm}+\frac{1}{2}k^2V_{lm}-R_{lm}\right)\left(r\partial_r U'^*_{lm}-U'^*_{lm}+\frac{1}{2}k^2V'^*_{lm}-R'^*_{lm}\right)\mathrm{d}r\\ &+k^2\int_{I_S}\frac{\mu_0}{\tau}(r\partial_r V_{lm}-V_{lm}+U_{lm}-S_{lm})(r\partial_r V'^*_{lm}-V'^*_{lm}+U'^*_{lm}-S'^*_{lm})\,\mathrm{d}r\\ &+k^2\int_{I_S}\frac{\mu_0}{\tau}(r\partial_r W_{lm}-W_{lm}-T_{lm})(r\partial_r W'^*_{lm}-W'^*_{lm}-T'^*_{lm})\,\mathrm{d}r\\ &+k^2(k^2-2)\int_{I_S}\frac{\mu_0}{\tau}(V_{lm}-M_{lm})(V'^*_{lm}-M'^*_{lm})\,\mathrm{d}r + k^2(k^2-2)\int_{I_S}\frac{\mu_0}{\tau}(W_{lm}-N_{lm})(W'^*_{lm}-N'^*_{lm})\,\mathrm{d}r,\end{aligned} \tag{D66}$$

$$\int_{\partial M} (\nabla\Phi \cdot \mathbf{u}'^{*} + \phi'^{*})\dot{\sigma}\,\mathrm{d}S = -[gU_{lm}'^{*}(a) + \phi_{lm}'(a)]a^2\dot{\sigma}_{lm}. \tag{D67}$$

In the above expressions, it is seen that the variables naturally decouple into the 'spheroidal subsystem'

$$\{U_{lm}, V_{lm}, \phi_{lm}, M_{lm}, R_{lm}, S_{lm}\}, \tag{D68}$$

and the 'toroidal subsystem'

$$\{W_{lm}, N_{lm}, T_{lm}\}. \tag{D69}$$

We can now obtain reduced weak form of the viscoelastic loading problem. To do so, we first consider the equations for the spheroidal subsystem, and set all toroidal test functions equal to zero. We then also set

$$U_{lm}' = V_{lm}' = \phi_{lm}' = 0, \tag{D70}$$

but leave M_{lm}', R_{lm}' and S_{lm}' arbitrary. From eq. (D51) along with eqs (D65) and (D66), we see that we must have

$$\dot{M}_{lm} + \frac{1}{\tau}M_{lm} = \frac{1}{\tau}V_{lm}, \tag{D71}$$

$$\dot{R}_{lm} + \frac{1}{\tau}R_{lm} = \frac{1}{\tau}\left(r\partial_r U_{lm} - U_{lm} + \frac{1}{2}k^2 V_{lm}\right), \tag{D72}$$

$$\dot{S}_{lm} + \frac{1}{\tau}S_{lm} = \frac{1}{\tau}\left(r\partial_r V_{lm} - V_{lm} + U_{lm}\right). \tag{D73}$$

We note that these equations can also be obtained directly from eq. (2.32). Next, we set

$$M_{lm}' = R_{lm}' = S_{lm}' = 0, \tag{D74}$$

but let U_{lm}', V_{lm}' and ϕ_{lm}' be arbitrary. Using eqs (D62), (D66) and (D67), we see that eq. (D51) implies

$$\mathcal{A}_l^S(\dot{U}_{lm}, \dot{V}_{lm}, \dot{\phi}_{lm} \,|\, U_{lm}'^{*}, V_{lm}'^{*}, \phi_{lm}'^{*}) = \mathcal{R}_l^S(U_{lm}, V_{lm}, M_{lm}, R_{lm}, S_{lm} \,|\, U_{lm}'^{*}, V_{lm}'^{*}) + \mathcal{F}_l^S(\dot{\sigma}_{lm} \,|\, U_{lm}'^{*}, \phi_{lm}'^{*}), \tag{D75}$$

which must hold for all radial test functions U_{lm}', V_{lm}' and ϕ_{lm}'. In this equation, $\mathcal{A}_l^S$ is the bilinear form

$$\begin{aligned}
\mathcal{A}_l^S(U_{lm}, V_{lm}, \phi_{lm} \,|\, U_{lm}', V_{lm}', \phi_{lm}') = {} & \int_{I_S} \kappa(r\partial_r U_{lm} + 2U_{lm} - k^2 V_{lm})(r\partial_r U_{lm}' + 2U_{lm}' - k^2 V_{lm}')\,\mathrm{d}r \\
& + \frac{4}{3}\int_{I_S} \mu_0\left(r\partial_r U_{lm} - U_{lm} + \frac{1}{2}k^2 V_{lm}\right)(r\partial_r U_{lm}' - U_{lm}' + \frac{1}{2}k^2 V_{lm}')\,\mathrm{d}r \\
& + k^2\int_{I_S} \mu_0(r\partial_r V_{lm} - V_{lm} + U_{lm})(r\partial_r V_{lm}' - V_{lm}' + U_{lm}')\,\mathrm{d}r + k^2(k^2-2)\int_{I_s} \mu_0 V_{lm}V_{lm}'\,\mathrm{d}r \\
& + \int_{I_S} 4\rho(\pi G\rho r - g)U_{lm}U_{lm}' r\,\mathrm{d}r + k^2\int_{I_S} \rho g(V_{lm}U_{lm}' + U_{lm}V_{lm}')r\,\mathrm{d}r \\
& + \int_{I_S} \rho(\partial_r\phi_{lm}U_{lm}' + U_{lm}\partial_r\phi_{lm}')r^2\,\mathrm{d}r + k^2\int_{I_S} \rho(V_{lm}\phi_{lm}' + \phi_{lm}V_{lm}')r\,\mathrm{d}r \\
& + \frac{1}{4\pi G}\int_I (\partial_r\phi_{lm}\partial_r\phi_{lm}' r^2 + k^2\phi_{lm}\phi_{lm}')\,\mathrm{d}r + \frac{1}{4\pi G}(l+1)a\phi_{lm}(a)\phi_{lm}'(a) + \int_{I_F} g^{-1}\partial_r\rho\phi_{lm}\phi_{lm}' r^2\,\mathrm{d}r \\
& + \sum_{r\in d_{FS}} \rho^-(r)\left[g(r)U_{lm}(r)U_{lm}'(r) + \phi_{lm}(r)U_{lm}'(r) + U_{lm}(r)\phi_{lm}'(r)\right]r^2 \\
& - \sum_{r\in d_{SF}} \rho^+(r)\left[g(r)U_{lm}(r)U_{lm}'(r) + \phi_{lm}(r)U_{lm}'(r) + U_{lm}(r)\phi_{lm}'(r)\right]r^2,
\end{aligned} \tag{D76}$$

which is symmetric in the sense that

$$\mathcal{A}_l^S(U_{lm}, V_{lm}, \phi_{lm} \,|\, U_{lm}', V_{lm}', \phi_{lm}') = \mathcal{A}_l^S(U_{lm}', V_{lm}', \phi_{lm}' \,|\, U_{lm}, V_{lm}, \phi_{lm}), \tag{D77}$$

while $\mathcal{R}_l$ is given by

$$\mathcal{R}_l^S(U_{lm}, V_{lm}, M_{lm}, R_{lm}, S_{lm} \mid U'_{lm}, V'_{lm}) = \frac{4}{3}\int_{I_S} \frac{\mu_0}{\tau}\left(r\,\partial_r U_{lm} - U_{lm} + \frac{1}{2}k^2 V_{lm} - R_{lm}\right)\left(r\,\partial_r U'_{lm} - U'_{lm} + \frac{1}{2}k^2 V'_{lm}\right) \mathrm{d}r$$
$$+ k^2 \int_{I_S} \frac{\mu_0}{\tau}(r\,\partial_r V_{lm} - V_{lm} + U_{lm} - S_{lm})(r\,\partial_r V'_{lm} - V'_{lm} + U'_{lm})\,\mathrm{d}r$$
$$+ k^2(k^2-2)\int_{I_S} \frac{\mu_0}{\tau}(V_{lm} - M_{lm})V'_{lm}\,\mathrm{d}r, \tag{D78}$$

and $\mathcal{F}_l$ by

$$\mathcal{F}_l^S(\dot{\sigma}_{lm} \mid U'_{lm}, \phi'_{lm}) = -[gU'_{lm}(a) + \phi'_{lm}(a)]a^2\dot{\sigma}_{lm}. \tag{D79}$$

Taken together, eqs (D75), (D71), (D72) and (D73) constitute a coupled system of evolution equations for the variables

$$\{U_{lm}, V_{lm}, \phi_{lm}, M_{lm}, R_{lm}, S_{lm}\}. \tag{D80}$$

In the case that σ_{lm} is discontinuous, this system of evolution equations must be supplemented by the jump conditions

$$\mathcal{A}_l^S(\Delta U_{lm}, \Delta V_{lm}, \Delta\phi_{lm} \mid U'^*_{lm}, V'^*_{lm}, \phi'_{lm}) = \mathcal{F}_l^S(\Delta\sigma_{lm} \mid U'^*_{lm}, \phi'^*_{lm}), \tag{D81}$$

and

$$\Delta M_{lm} = \Delta R_{lm} = \Delta S_{lm} = 0. \tag{D82}$$

Repeating the above derivation but now for the toroidal variables, we obtain

$$\mathcal{A}_l^T(\dot{W}_{lm} \mid W'^*_{lm}) = \mathcal{R}_l^T(W_{lm}, N_{lm}, T_{lm} \mid W'^*_{lm}), \tag{D83}$$

$$\dot{N}_{lm} + \frac{1}{\tau}N_{lm} = \frac{1}{\tau}W_{lm}, \tag{D84}$$

$$\dot{T}_{lm} + \frac{1}{\tau}T_{lm} = \frac{1}{\tau}(r\,\partial_r W_{lm} - W_{lm}), \tag{D85}$$

where

$$\mathcal{A}_l^T(W_{lm} \mid W'_{lm}) = k^2\int_{I_S} \mu_0(r\,\partial_r W_{lm} - W_{lm})(r\,\partial_r W'_{lm} - W'_{lm})\,\mathrm{d}r + k^2(k^2-2)\int_{I_S}\mu_0 W_{lm}W'_{lm}\,\mathrm{d}r, \tag{D86}$$

$$\mathcal{R}_l^T(W_{lm}, N_{lm}, T_{lm} \mid W'_{lm}) = k^2\int_{I_S}\frac{\mu_0}{\tau}(r\,\partial_r W_{lm} - W_{lm} - T_{lm})(r\,\partial_r W'_{lm} - W'_{lm})\,\mathrm{d}r + k^2(k^2-2)\int_{I_S}\frac{\mu_0}{\tau}(W_{lm} - N_{lm})W'_{lm}\,\mathrm{d}r. \tag{D87}$$

As these equations have no force terms, it follows from the initial conditions that

$$W_{lm} = N_{lm} = T_{lm} = 0, \tag{D88}$$

for all $t \in [t_0, t_1]$, and so we need not consider the toroidal subsystem explicitly when considering the GIA forward problem in spherically symmetric earth models.

The force term occurring in the adjoint viscoelastic loading problem stated in eq. (3.20) takes a more general form than that in the forward problem considered above. It will, therefore, be of use to generalize the above results slightly by making the substitution

$$\int_{\partial M}(\nabla\Phi\cdot\mathbf{u}'^* + \phi'^*)\dot{\sigma}\,\mathrm{d}S \to \int_{\partial M}(\dot{\mathbf{h}}\cdot\mathbf{u}'^* + \dot{h}\,\phi'^*)\,\mathrm{d}S, \tag{D89}$$

in eq. (D51) with $\mathbf{h}$ and h given functions on ∂M. We suppose that $\mathbf{h}$ and h have been expanded in generalized spherical harmonics

$$h^\alpha = \sum_{lm} h^\alpha_{lm} Y^\alpha_{lm}, \tag{D90}$$

$$h = \sum_{lm} h_{lm} Y^0_{lm}, \tag{D91}$$

where h^α denote the contraviant components of $\mathbf{h}$, and that the h^α_{lm} take the form

$$h^\pm_{lm} = \frac{1}{\sqrt{2}}k(G_{lm} \pm \mathrm{i}H_{lm}), \tag{D92}$$

$$h^0_{lm} = F_{lm}, \tag{D93}$$

which correspond to the toroidal and spheroidal combinations of Phinney & Burridge (1973). With the test functions $\mathbf{u}'$, and ϕ' defined in eqs (D53) and (D54), we obtain

$$\int_{\partial M} (\dot{\mathbf{h}} \cdot \mathbf{u}'^{*} + \dot{h}\phi'^{*})\, \mathrm{d}S = a^2 \dot{F}_{lm} U_{lm}^{'*}(a) + k^2 a^2 \left[\dot{G}_{lm} V_{lm}^{'*}(a) + \dot{h}_{lm} W_{lm}^{'*}(a)\right] + a^2 \dot{h}_{lm} \phi_{lm}^{'*}(a). \tag{D94}$$

It follows that for such a force term eq. (D75) for the spheroidal subsystem must be modified by making the replacement

$$\mathcal{F}_l^S(\dot{\sigma}_{lm} \mid U_{lm}^{'*}, \phi_{lm}^{'*}) \rightarrow \mathcal{F}_l^S(\dot{F}_{lm}, \dot{G}_{lm}, \dot{h}_{lm} \mid U_{lm}^{'*}, V_{lm}^{'*}, \phi_{lm}^{'*}), \tag{D95}$$

where the term on the right-hand side is defined by

$$\mathcal{F}_l^S(\dot{F}_{lm}, \dot{G}_{lm}, \dot{h}_{lm} \mid U_{lm}^{'*}, V_{lm}^{'*}, \phi_{lm}^{'*}) = a^2 \dot{F}_{lm} U_{lm}^{'*}(a) + k^2 a^2 \dot{G}_{lm} V_{lm}^{'*}(a) + a^2 \dot{h}_{lm} \phi_{lm}^{'*}(a). \tag{D96}$$

Similarly, eq. (D83) for the toroidal subsystem becomes

$$\mathcal{A}_l^T(\dot{W}_{lm} \mid W_{lm}^{'*}) = \mathcal{R}_l^T(W_{lm}, N_{lm}, T_{lm} \mid W_{lm}^{'*}) + \mathcal{F}_l^T(\dot{H}_{lm} \mid W_{lm}^{'*}), \tag{D97}$$

with

$$\mathcal{F}_l^T(\dot{H}_{lm} \mid W_{lm}^{'*}) = k^2 a^2 \dot{h}_{lm} W_{lm}^{'*}(a). \tag{D98}$$

We note that for this more general problem, the toroidal equations are not necessarily homogeneous. In particular, for certain choices of objective functional the adjoint deformation fields may have a non-zero toroidal component.

D3 Numerical solution of the viscoelastic loading problem in a spherical earth model

We now consider the numerical solution of the viscoelastic loading problem in a spherically symmetric earth model based upon the reduced weak form of the problem described above. In this approach, we have seen that there is complete decoupling between the equations for different spherical harmonics degrees and orders, and so we focus on one particular value of (l, m) and drop the spherical harmonic subscripts on the various fields for notational simplicity. There is, moreover, a further decoupling into the spheroidal and toroidal subsystems, and for definiteness we shall restrict attention to the more interesting spheroidal case.

It follows that we wish to determine the time evolution for $t \in [t_0, t_1]$ of the fields $\{U, V, \phi, M, R, S\}$, which are governed by the eqs (D71), (D72), (D73) and (D75) and subject to initial conditions

$$U = V = \phi = M = R = S = 0, \tag{D99}$$

at time t_0. Let us suppose that at some time $t \in [t_0, t_1]$ we know the values of $\{U, V, \phi, M, R, S\}$, this being, in particular, the case for $t = t_0$. We can then use eqs (D71), (D72) and (D73) to calculate the corresponding values of $\dot{M}$, $\dot{R}$ and $\dot{S}$ explicitly. Furthermore, eq. (D75) defines a 1-D boundary value problem whose solution yields the values of the time derivatives $\dot{U}$, $\dot{V}$ and $\dot{\phi}$. Supposing for the moment that this boundary value problem can be solved, the time evolution of the system can be approximated numerically using a suitable time-stepping algorithm. In our calculations, we have employed a second-order Runge–Kutta method (e.g. Press *et al.* 1986) with fixed time step equal to $\frac{1}{2}\tau_{\min}$, where $\tau_{\min}$ is the smallest value of the Maxwell relaxation time of the earth model. Numerical test have shown this integration scheme to be stable, and it seems to be more than sufficient for this problem. In the case that the applied surface load is only piecewise continuous, this time-stepping scheme must be supplemented with the jump condition in eq. (D81), which is of the same form as the boundary value problem arising in the calculation of the time derivatives $\dot{U}$, $\dot{V}$ and $\dot{\phi}$.

In describing the numerical solution of eq. (D75) it will simplify the exposition to instead consider the equation

$$\mathcal{A}_l^S(U_{lm}, V_{lm}, \phi_{lm} \mid U_{lm}^{'*}, V_{lm}^{'*}, \phi_{lm}^{'*}) = \mathcal{F}_l^S(\sigma_{lm} \mid U_{lm}^{'*}, \phi_{lm}^{'*}), \tag{D100}$$

which arises from the spherical harmonic reduction of a static elastic loading problem. The only non-notational difference between eqs (D75) and (D100) is that the force term in the former problem is slightly more complicated, and it should be clear how the method described can be extended to this case. To determine a numerical solution of eq. (D100), we employ a spectral element discretization to reduce the problem to a system of linear algebraic equations. In the following subsections, we briefly explain the details of this discretization and the solution of the resulting linear system.

D3.1 Radial mesh and Lagrange polynomial interpolation

The radial mesh comprises n_e elements $I_i = [r_i, r_{i+1}]$ with $r_i < r_{i+1}$ such that

$$I = \bigcup_{i=1}^{n_e} I_i, \tag{D101}$$

where we recall that $I = [0, a]$ with a the radius of the earth model. We assume that these elements are such that all radial discontinuities within the model lie on the boundary between adjacent elements. A surface load having spherical harmonic degree l has a characteristic length equal to $\frac{2\pi a}{(l+\frac{1}{2})}$, and this is also true of the resulting deformation. In order for our radial discretization to be sufficiently accurate, we require that each element of the mesh spans a radial distance less than one-tenth of this characteristic length, with the suitability of this choice having been verified numerically. Each of these elements can be mapped bijectively onto a reference element $\xi \in [-1, 1]$ through

$$\xi(r) = 2\frac{r - r_i}{r_{i+1} - r_i} - 1, \tag{D102}$$

and with inverse given by

$$r(\xi) = r_i + \frac{1}{2}(r_{i+1} - r_i)(\xi + 1). \tag{D103}$$

Using these mappings, it is then possible to represent any function defined on I_i by one defined on the reference element $[-1, 1]$. We also note from eq. (D103) the relation

$$\frac{\mathrm{d}r}{\mathrm{d}\xi} = \frac{1}{2}(r_{i+1} - r_i), \tag{D104}$$

which enables the transformation of derivatives and integrals between the two coordinate systems.

We now fix an integer $n_l \geq 2$, and let ξ_j be the $n_l + 1$ Gauss–Lobatto–Legendre (GLL) points on $[-1, 1]$, which are defined as roots of the equation

$$(1 - \xi^2)P'_{n_l}(\xi) = 0, \tag{D105}$$

where P_{n_l} is the n_lth Legendre polynomial, and we assume the ordering $\xi_j < \xi_{j+1}$ (e.g. Komatitsch & Tromp 1998). We note that $\xi_1 = -1$ and $\xi_{n_l+1} = 1$, so some GLL points always lie exactly on the boundary of the reference element. Associated with these GLL points, we define the $n_l + 1$ Lagrange polynomials

$$l_j(\xi) = \frac{\prod_{k=1}^{n_l+1}(\xi - \xi_k)}{\prod_{k\neq j}(\xi_j - \xi_k)}, \tag{D106}$$

which satisfy

$$l_j(\xi_k) = \delta_{jk}. \tag{D107}$$

Using eq. (D103), we can map these GLL points to a set of nodes

$$r_{ij} = r_i + \frac{1}{2}(r_{i+1} - r_i)(\xi_j + 1), \tag{D108}$$

lying in the ith element of the mesh and such that there always exist nodes at either end of the element.

Let f be a scalar function defined on I and write $f_{ij} = f(r_{ij})$ for its value at the jth node lying in ith element. We then define the Lagrange interpolation of f on I_i by

$$f(r(\xi)) \approx \sum_{j=1}^{n_l} f_{ij} l_j(\xi), \tag{D109}$$

which is seen from eq. (D107) to be exact at each r_{ij}, and is, in fact, everywhere exact if f is a polynomial of order less than or equal to $n_l + 1$. Corresponding to eq. (D109), we obtain an approximate expression for the radial derivative f on I_i in the form

$$\frac{\mathrm{d}f}{\mathrm{d}r}(r(\xi)) \approx \sum_{j=1}^{n_l} f_{ij} l'_j(\xi)\frac{\mathrm{d}\xi}{\mathrm{d}r} = \frac{2}{r_{i+1} - r_i}\sum_{j=1}^{n_l} f_{ij} l'_j(\xi), \tag{D110}$$

where l'_j is the derivative of the jth Lagrange polynomial with respect to its argument and in obtaining the final equality we have made use of eq. (D104). We can also approximate the integral of f over I_i using the GLL quadrature rule as

$$\int_{I_i} f(r)\,\mathrm{d}r = \int_{-1}^{1} f(\xi(r))\frac{\mathrm{d}r}{\mathrm{d}\xi}\,\mathrm{d}\xi \approx \frac{1}{2}(r_{i+1} - r_i)\sum_{j=1}^{n_l} f_{ij} w_j, \tag{D111}$$

with w_j integration weights that can be calculated numerically (Canuto *et al.* 1988, p. 61).

D3.2 Assembly and solution of the linear system

To discretize eq. (D100), we approximate U, V and ϕ using Lagrange interpolation on each of the nodes I_i of the radial mesh

$$U(r(\xi)) \approx \sum_{j=1}^{n_l} U_{ij} l_j(\xi), \tag{D112}$$

$$V(r(\xi)) \approx \sum_{j=1}^{n_l} V_{ij} l_j(\xi), \tag{D113}$$

$$\phi(r(\xi)) \approx \sum_{j=1}^{n_l} \phi_{ij} l_j(\xi), \tag{D114}$$

and so reduce the problem to determining the unknown interpolation coefficients U_{ij}, V_{ij} and ϕ_{ij}. As U and V are defined only in solid regions, U_{ij} and V_{ij} needed be introduced only for those elements $I_i \subseteq I_S$. Furthermore, continuity of U, V in I_S and of ϕ in I means that for adjacent radial elements we must have

$$U_{i\,n_l+1} = U_{i+1\,1}, \tag{D115}$$

$$V_{i\,n_l+1} = V_{i+1\,1}, \tag{D116}$$

$$\phi_{i\,n_l+1} = \phi_{i+1\,1}, \tag{D117}$$

where the first two of these conditions apply only if both I_i and I_{i+1} lie in solid regions. Subject to these constraints, we can construct a global numbering of the independent coefficients such that at each global node in the mesh there is either one or three degrees of freedom depending on whether the node lies, respectively, in fluid or solid regions of the model. We note, in particular, that if a node lies on a boundary between fluid and solid regions then it is associated with three degrees of freedom. We can similarly approximate the test functions U', V' and ϕ' in the above manner using Lagrange polynomials. Using these interpolation formulae, we can then reduce eq. (D100) into a system of linear algebraic equations for the interpolation coefficients U_{ij}, V_{ij} and ϕ_{ij}. In doing so, we make use of eqs (D110) and (D111) to approximately evaluate the various derivatives and integrals occurring in the definition of the bilinear form $\mathcal{A}_l^S$. With a suitable ordering of the independent interpolation coefficients, this process results in a banded symmetric linear system, with the bandwidth equal to $6(n_l - 1) + 1$. Even for fairly large values of l (e.g. less than 500) the size of the linear system is modest, and can readily be solved using lower-upper decomposition (e.g. Press *et al.* 1986).

APPENDIX E: SMOOTHED DELTA FUNCTION EXPANSIONS IN SPHERICAL HARMONICS

In the calculation of the adjoint load corresponding to certain objective functionals, it is necessary to form spherical harmonic expansions of delta functions $\delta(\mathbf{x} - \mathbf{x}')$ on the sphere. This is, in particular, the case for point measurements of the deformation field in space. The exact expansion of this function is readily found to be

$$\delta(\mathbf{x} - \mathbf{x}') = \frac{1}{a^2} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y_{lm}^{0*}(\theta', \varphi') Y_{lm}^{0}(\theta, \varphi), \tag{E1}$$

where a is the radius of ∂M, (θ, φ) are the spherical polar coordinates associated with $\mathbf{x}$ and (θ', φ') are those corresponding to $\mathbf{x}'$ and the convergence of this infinite sum must be understood in the distributional sense. In practice, we must, of course, work with a truncated version of this expansion, which includes non-zero terms only up to some maximum spherical harmonic degree L. The simplest such truncated expansion is given by

$$\delta(\mathbf{x} - \mathbf{x}') \approx \frac{1}{a^2} \sum_{l=0}^{L} \sum_{m=-l}^{l} Y_{lm}^{0*}(\theta', \varphi') Y_{lm}^{0}(\theta, \varphi), \tag{E2}$$

where we have removed all terms of degree higher than L. Use of this expansion does, however, lead to significant ringing due to the sharp cut-off of the spherical harmonic sum. Nonetheless, if we know *a priori* that the load σ, which produced the deformation field has negligible power in its spherical harmonic spectrum for $l > L$, then eq. (E2) is a suitable integral kernel for making spatial point measurements of the deformation field. Unfortunately, however, this ringing inherited by the adjoint deformation fields and, in turn, by the rate-of-loading and viscosity kernels. In the gradient-based optimization methods, these kernels are used to iteratively update the load and viscosity models, and so ringing in the kernels can lead to short wavelength artefacts in the models obtained. To circumvent this problem, we instead make use of the alternative truncated expansion

$$\delta(\mathbf{x} - \mathbf{x}') \approx \frac{1}{a^2} \sum_{l=0}^{L} \sum_{m=-l}^{l} \exp\left(-2\pi \frac{l+1}{L+\frac{1}{2}}\right) Y_{lm}^{0*}(\theta', \varphi') Y_{lm}^{0}(\theta, \varphi), \tag{E3}$$

which for $L \to \infty$ also converges in the sense of distributions. In this case, there is no undesired ringing away from the observation point due to the exponential damping of short wavelength terms in the expansion. Moreover, it may be shown that the expansion in eq. (E3) defines a function on ∂M that has its support approximately contained in a spherical cap about the observation point (θ', φ') whose radius scales like $2\pi a/(L + \frac{1}{2})$, and that the integral of this function over ∂M is approximately equal to 1. Because of these properties, eq. (E3) defines a suitable integral kernel for measuring the point values of deformation fields which are *a priori* known to be smooth over length scales shorter than $2\pi a/(L+\frac{1}{2})$. Use of eq. (E3) in the adjoint calculations leads to rate-of-loading kernels that are also smooth over length scales of $2\pi a/(L+\frac{1}{2})$. Within the context of gradient-based optimization, this means that updates to the load model cannot add shorter wavelength features, and so use of eq. (E3) for a given value of L amounts to an implicit form of regularization on the smoothness of the load model.

APPENDIX F: RADIAL VISCOSITY KERNELS

In Section 3.1, we introduced a radial viscosity kernel $\overline{K}_\eta$ through

$$\overline{K}_\eta = \int_\Omega K_\eta r^2 \, \mathrm{d}S. \tag{F1}$$

From eq. (3.22), we see that

$$\overline{K}_\eta = \frac{1}{2\eta}\int_{t_0}^{t_1}\int_\Omega r^2 \boldsymbol{\tau}(t) : \boldsymbol{\tau}^\dagger(t_1 - t + t_0)\,\mathrm{d}S\,\mathrm{d}t, \tag{F2}$$

and using the generalized spherical harmonic expansions of $\boldsymbol{\tau}$ and $\boldsymbol{\tau}^\dagger$ along with the orthonormality relation in eq. (D52) we obtain

$$\begin{aligned}\overline{K}_\eta = {} & \frac{4\mu_0^2}{3\eta}\sum_{lm}\int_{t_0}^{t_1}\left(r\,\partial_r U_{lm} - U_{lm} + \frac{1}{2}k^2 V_{lm} - R_{lm}\right)\left(r\,\partial_r U_{lm}^{\dagger *} - U_{lm}^{\dagger *} + \frac{1}{2}k^2 V_{lm}^{\dagger *} - R_{lm}^{\dagger *}\right)\mathrm{d}t \\ & + \frac{\mu_0^2}{\eta}\sum_{lm} k^2 \int_{t_0}^{t_1}(r\,\partial_r V_{lm} - V_{lm} + U_{lm} - S_{lm})(r\,\partial_r V_{lm}^{\dagger *} - V_{lm}^{\dagger *} + U_{lm}^{\dagger *} - S_{lm}^{\dagger *})\,\mathrm{d}t \\ & + \frac{\mu_0^2}{\eta}\sum_{lm} k^2(k^2-2)\int_{t_0}^{t_1}(V_{lm} - M_{lm})(V_{lm}^{\dagger *} - M_{lm}^{\dagger *})\,\mathrm{d}t,\end{aligned} \tag{F3}$$

where it is understood that the forward fields are evaluated at the time t, while the adjoint fields are evaluated at the reversed time $t_1 - t + t_0$. We note that in obtaining this expression we have made use of the identities given in eqs (D18) and (D34).