A spectral element method for surface wave dispersion and adjoints

Rhys Hawkins

Research School of Earth Sciences, Australian National University, Canberra ACT 0200, Australia. E-mail: rhys.hawkins@anu.edu.au

Accepted 2018 July 6. Received 2018 March 2; in original form 2018 June 29

A spectral element method for modelling surface wave dispersion of Love and Rayleigh waves is presented. This method uses standard Gauss–Lobatto–Legendre polynomials coupled with a Gauss–Laguerre–Legendre element to represent a half-space, and in doing so, improves both the efficiency and accuracy of the calculation of phase and group velocities, particularly at lower frequencies. It is demonstrated that this method is able to directly represent 1-D earth models with smoothly varying structure, layered structure or combinations thereof. The method is both efficient and accurate, while the solution error can be tuned across a range of frequencies to balance the trade-off with computational cost. In addition, the adjoint technique is used to develop an efficient algorithm for the calculation of the gradient of arbitrary misfit functions with respect to earth model parameters, which is of prime interest in the inversion of surface waves. It is demonstrated that accurate gradients can be computed for misfit functions based on phase velocity, group velocity, and Rayleigh wave ellipticity. To demonstrate both the spectral element method and the adjoint method developed in this paper, two simulated inverse problems are presented using Love wave phase velocity observations and Rayleigh wave ellipticity observations.

Key words: Numerical methods; Inverse Theory; Surface waves; free oscillations.

1 INTRODUCTION

A common class of inference of the Earth’s near surface structure is the use of Love and Rayleigh waves. Observations of surface waves, obtained from earthquakes (Pilidou et al. 2005; Li & Detrick 2006), ambient noise cross-correlations (Lobkis & Weaver 2001; Campillo & Paul 2003; Shapiro & Campillo 2004; Larose et al. 2006) or Rayleigh wave ellipticity autocorrelations (Nakamura 1989) allow inference of Earth’s crustal structure to varying depths.

Knowledge of the shear wave velocities in the crust is important for seismic hazard assessment providing constraints on the properties of the upper crust including regolith, sediments and bedrock. Large variations can occur in near surface shear wave velocities including power law like profiles with depth (e.g. Molnár et al. 2010). Additionally, velocity contrasts across interfaces, such as the transition from sedimentary layering to bedrock, and the depths of such interfaces are important variables controlling site amplification and resonances that exacerbate the damaging effects of large earthquakes (Bard & Bouchon 1985; Anderson et al. 1996; Graves et al. 1998).

In the formulation of inverse problems for crustal scale surface wave dispersion, the observations consist of phase velocities, group velocities or Rayleigh wave ellipticity observations at a set of discrete frequencies. Lower frequency observations have increased sensitivities to depth, and it is this dependence of depth sensitivity with frequency that allows inference of subsurface structure. For problems at this scale, two approximating assumptions are commonly used: first, that the dispersion can be accurately modelled assuming a flat Earth, resulting in equations of motion defined with a Cartesian coordinate system (Aki & Richards 2002). At lower frequencies, this approximation becomes invalid and the near spherical nature of the Earth needs to be accounted for. Second, that the lateral heterogeneity is small or slowly varying with respect to the wavelengths considered. This allows a simplification whereby observations at a point on the Earth can be locally inverted for using a 1-D earth model. This approximation similarly breaks down at lower frequencies and recent studies have investigated 3-D sensitivities to local structure (Maupin 2017).

In a 1-D model of the Earth, Love waves are sensitive to density, and the horizontal and vertical shear wave velocities $v_{sH}$ and $v_{sV}$. Rayleigh waves are additionally sensitive to the horizontal and vertical $p$-wave velocities $v_{pH}$ and $v_{pV}$ (Aki & Richards 2002; Chapman 2004). For both types of surface waves however, the dominant sensitivity is to the shear wave velocity, and various parametrizations are used to reflect this. For example, solving directly for an isotropic shear wave velocity and using empirical relationships to compute density and $p$-wave velocity, or the use of fixed $v_{p}/v_{s}$ ratios (Bodin et al. 2012b).

Many methods have been proposed for numerical modelling of surface wave dispersion. Matrix propagator methods (Thomson 1950; Haskell 1953) remain a popular choice, however these are only strictly valid for an Earth consisting of a series homogeneous layers. Use of this approach for more complex velocity profiles is only approximate due to a necessary discretization into homogenous layers. Various improvements have been made to methods for calculation of dispersion for layered models (Schwab...
The numerical integration scheme of Takeuchi & Saito (1972) can incorporate arbitrary 1-D structure, however this approach requires multiple integrations using a numerical integration scheme from depth, for example Runge–Kutta, with a suitably chosen step size for convergence. This can be inefficient and in some cases numerically unstable. Methods using finite elements have also been proposed (e.g. Lysmer & Drake 1972; Haney & Douma 2011; Haney & Tsai 2017). These methods approximate the Earth as a series of gradient layers of shear wave velocity with depth. Due to the restriction of the approximation to linear functions, existing finite element approaches require many thin elements, approximately ten per wavelength, for accurate numerical simulation (Kausel & Roeser 1981). A recent advance has determined a method of reducing the number of elements by selecting an optimal element size as a function of depth (Haney & Tsai 2017). Rayleigh–Ritz methods, a similar variational approach that can use a fixed set of smooth basis functions has also been demonstrated for studies of the Earth at longer periods (Wiggins 1976). At the global scale and at much lower frequencies, surface wave modelling schemes need to account for the near spherical shape of the Earth and are represented using expansions of spherical harmonics (e.g. Woodhouse 1974; Tromp & Dahlen 1992a,b). The focus of this paper is restricted to range of frequencies where a flat Earth approximation is appropriate.

In this paper, a new accurate and efficient modelling approach for surface waves is proposed based on spectral elements that can simultaneously represent smoothly varying structure, layered structures, or combinations thereof. Furthermore, an adjoint method is developed for computing the gradient of arbitrary functions of phase velocity, group velocity or Rayleigh wave ellipticity, for example, a misfit function, with respect to earth model parameters. These developments are illustrated using two simple examples of a simulated inverse problem, involving gradient descent optimization.

2 A SPECTRAL ELEMENT METHOD FOR 1-D SURFACE WAVE DISPERSION

In recent years, the spectral element method (SEM; Patera 1984) has gained traction in the seismology community in simulating full waveform propagation through 1-D, 2-D and 3-D media at various scales (Komatitsch & Vilotte 1998; Komatitsch & Tromp 1999, 2002a,b; Fichtner et al. 2009). Its popularity owes much to its ability to easily incorporate the free-surface boundary conditions and its improved accuracy of the simulation of seismic wave propagation due to the incorporation of higher order elements.

The SEM is a general technique for the discrete solution of differential equations. The differential equations that govern surface wave propagation are summarized here from several sources (Love 1927; Anderson 1961; Takeuchi & Saito 1972; Kennett 1973; Achenbach 1975; Thomsen 1988; Chapman 2004; Fichtner 2011).

For the application of local inversion of surface wave observations, a 1-D approximation can be used that assumes lateral variation is small with respect to the wavelength of surface waves considered. In this 1-D case, material properties only vary in the radial direction and at most, transverse isotropy is resolvable. The 1-D earth model can then be represented as a function of the Love parameters $A$, $C$, $F$, $L$, $N$, and density $\rho$ with respect to depth (Love 1927). These parameters can be related to seismic wave speed parameters, specifically

\[ A = \rho v_{ph}^2 \]  

\[ C = \rho v_{ps}^2 \]  

\[ N = \rho v_{sh}^2 \]  

\[ L = \rho v_{sv}^2 \]  

where $v_{ph}$, $v_{ps}$, $v_{sh}$, $v_{sv}$ are the horizontal and vertical $P$-wave velocities and $V_{ph}$, $V_{ps}$ are the horizontal and vertical shear wave velocities respectively.

For the development here, a Cartesian coordinate system is used where $z$ is the radial direction into the Earth, $x$ along the surface of the Earth in the surface wave propagation direction and $y$ perpendicular along the surface. For the propagation of Love waves, the surface wave propagates with oscillations polarized in the $y$ direction. This oscillation varies with depth, or $z$, and is frequency and mode dependent. By representing this unknown function with \( V(\omega, k, z) \) (Aki & Richards 2002), the differential equation for Love waves can be expressed as

\[
\omega^2 \rho(z) V(\omega, k, z) - k^2 N(z) V(\omega, k, z)
+ \frac{d}{dz} \left[ L(z) \frac{dV(\omega, k, z)}{dz} \right] = 0.
\]  

(5)

For Rayleigh waves, particle motion is elliptical consisting of a vertical oscillation ($z$ direction) out of phase with an oscillation in the direction of propagation ($x$ direction). Labelling these two oscillation functions $U(\omega, k, z)$ and $W(\omega, k, z)$, the coupled differential equations for Rayleigh wave propagation are given by

\[
\omega^2 \rho(z) U(\omega, k, z) - k^2 A(z) U(\omega, k, z) - k F(z) \frac{dW(\omega, k, z)}{dz}
+ \frac{d}{dz} \left[ L(z) \left( \frac{dU(\omega, k, z)}{dz} - k W(\omega, k, z) \right) \right] = 0,
\]  

(6)

and

\[
\omega^2 \rho(z) W(\omega, k, z) - k^2 L(z) U(\omega, k, z) + k L(z) \frac{dU(\omega, k, z)}{dz}
+ \frac{d}{dz} \left[ C(z) \frac{dW(\omega, k, z)}{dz} + F(z) k U(\omega, k, z) \right] = 0.
\]  

(7)

The boundary conditions for both Love and Rayleigh require that the stress is zero at the surface and continuous throughout, and also, that the amplitude function decays to zero with depth. The single stress term for Love waves is

\[
\tau_{zz} = L(z) \frac{dV(\omega, k, z)}{dz}
\]  

(8)

and the two terms for Rayleigh waves are

\[
\tau_{xz} = L(z) \left( \frac{dU(\omega, k, z)}{dz} - k W(\omega, k, z) \right)
\]  

(9)

\[
\tau_{zx} = F(z) k U(\omega, k, z) + C(z) \frac{dW(\omega, k, z)}{dz}.
\]  

(10)

In order to maintain the continuity of the stresses as a function of depth, many modelling schemes restate each of the second-order eqs (5), (6) and (7) as two first-order equations by introducing additional variables representing the respective stresses as continuous functions of depth (Aki & Richards 2002).

In a spectral element formulation, the weak form of the equations is solved using a discretization of the 1-D model domain where each element uses polynomial basis functions. The weak
form of the equation implicitly enforces the surface boundary condition (Fichtner 2011., section 4.1.1). The polynomial approximations used within each element, the Gauss–Lobatto–Legendre polynomials in this formulation, are at least $C^1$ continuous meaning that within each element, the stress terms are continuous. At the element boundaries, the weak form of the equations allows trivial enforcement of the continuity of the stresses, in a similar manner to the zero surface stress condition. It is in part the ease of enforcing boundary conditions that has made the SEM an accurate and popular method for the simulation of seismic wave propagation.

Finally, the boundary condition at depth that requires the amplitude function of the oscillations decays to zero. Previous similar methods using finite elements approximated this boundary condition by extending the model to a sufficiently large depth where a rigid boundary condition was enforced (Lysmer 1970; Lysmer & Drake 1972; Haney & Douma 2011). The depth is chosen so that for the largest period considered, this artificial boundary condition does not adversely impact the solution. While capable of producing accurate dispersion curves, this approach increases the size of the model and in turn the computational cost of the solution. Recent advances (Haney & Tsai 2017) have shown how this extra computational cost can be reduced by optimally scaling the size of successive elements for a given frequency.

In the spectral element field, there have been various approaches for providing robust solutions to half-space problems using either a mapping from a finite to a half-space, for example $-1 \ldots 1$ to $0 \ldots \infty$ (Boyd 1987, 2001). This type of mapping technique has very recently been applied in modelling the gravitational potential in full waveform spectral element inversion to enable larger periods to be simulated more accurately (Gharti & Tromp 2017).

An alternative approach to the ‘finite to infinite’ domain transformation technique is the use of either Laguerre polynomials or Laguerre functions introduced by Mavriplis (1989). Laguerre polynomials and functions can be seamlessly incorporated into a Gauss–Lobatto–Legendre framework with the coupling of a finite domain to an infinite one (Shen 2000; Valenciano & Chaplain 2005). In a comparison of the mapping approach to Laguerre polynomials, Black (1998) found that the relative merits of one over the other were problem dependent.

Laguerre polynomials are defined recursively as

\begin{align*}
L_0(x) &= 1 \\
L_1(x) &= 1 - x \\
L_n(x) &= \frac{2n - 1 - x}{n} L_{n-1}(x) - \frac{n - 1}{n} L_{n-2}(x),
\end{align*}

and Laguerre functions

\begin{equation}
\hat{L}_n(x) = L_n(x) \exp \left\{ -\frac{x}{2} \right\}.
\end{equation}

For finite spectral elements, a physical part of the domain, $z$ is mapped to the $\zeta \in -1 \ldots 1$ interval. For an infinite spectral element, a half-space in the physical domain needs to be mapped from $z_{\min} \ldots \infty$ to $\zeta \in 0 \ldots \infty$. Unlike a finite domain element, the scaling is free to be chosen and allows the Laguerre element to optimally fit the problem at hand. For Love waves dispersion, in a constant layer over a half-space, the well-known analytical solution for the eigenfunction is given by

\begin{equation}
V(\omega, k, z) \propto \exp \left[ -\sqrt{k^2 - \frac{\omega^2 \rho}{L}} z \right].
\end{equation}

Hence when determining the transform from the physical domain to the spectral element domain, the transform

\begin{equation}
z(\zeta) = \frac{\zeta}{\sqrt{k^2 - \frac{\omega^2 \rho}{L}}},
\end{equation}

is optimal, and the Laguerre polynomials can perfectly fit the corresponding eigenfunction. Unfortunately there is a circularity here in that, in order to solve for $k$, the optimal scaling requires knowledge of $k$ at each frequency $\omega$.

The approach advocated here is to iterate and use the result of a previous computation from which $k$ is known to have changed little. In most inversion problems using surface waves, the observations will consist of a phase velocity, group velocity or Rayleigh wave ellipticity for a set of frequencies. A simple approach for accurately scaling the Laguerre polynomial element is to order the observations from high frequency to low and then use an initial guess for $k$ for the highest frequency observations (where sensitivity to poor scaling of the half-space is lowest). After which, the $k$ from the previous frequency can be used to scale the Laguerre element for the next lower frequency observation and so on. A more computationally expensive approach would be to initially guess $k$ and then iterate until satisfactory convergence. However, in our experience the approximate $k$ from a higher frequency result is generally sufficient.

The attraction of the Laguerre functions is that their formulation in a spectral element solution is entirely compatible with Gauss–Lobatto–Legendre polynomials, with only the quadrature weights and differential operators varying, greatly simplifying the incorporation of an infinite domain element. In the formulation used in this study, an earth model will consist of some number of Gauss–Lobatto–Legendre spectral elements from the surface to the half-space, then a single Gauss–Laguerre–Legendre element to model the half-space. This results in an efficient solution for larger periods, similar to more advanced finite element approaches utilizing optimally sized elements (Haney & Tsai 2017), which is a distinct improvement over other approaches where the addition of many thin elements are needed in order to extend the modelled domain to a rigid basement. In Appendix B we show differences between inverting with and without the Laguerre element and given the ease of its use, and the low computational overhead, suggest that the addition of the Laguerre element should be the norm.

The detailed derivation of the spectral element formulation is outlined in Appendix A and follows a similar treatment to introductory texts (Pozrikidis 2005; Koprich 2009). In summary, for Love waves, the application of the SEM to the solution of (5), results in a linear eigensystem of the form

\begin{equation}
[\omega^2 A - k^2 B + C] V = 0,
\end{equation}

where the matrix $A$ is a diagonal matrix with terms that are functions of model densities, $B$ is a diagonal matrix with terms dependent on $N(z)$ and $C$ a block diagonal matrix with terms depending on $L(z)$, and the vector $V$ represents the polynomial approximation of the eigenfunction $V(\omega, k, z)$.

The Love wave matrix equation is already in a general eigenvalue problem form although we include an additional matrix pre-scaling which improves numerical accuracy (Ward 1981). The matrix equation (15) can then be solved with standard generalized eigen solving routines, for example, the routines provided by the LAPACK library (Anderson et al. 1999). Depending on application, other methods could be used that only compute selected eigenvalues (Lehoucq et al. 2011, section 4.1.1). The polynomial approximations used within each element, the Gauss–Lobatto–Legendre polynomials in this formulation, are at least $C^1$ continuous meaning that within each element, the stress terms are continuous. At the element boundaries, the weak form of the equations allows trivial enforcement of the continuity of the stresses, in a similar manner to the zero surface stress condition. It is in part the ease of enforcing boundary conditions that has made the SEM an accurate and popular method for the simulation of seismic wave propagation.
et al. 1998) as advocated by Haney & Tsai (2017) to improve efficieny. Alternatively, GPU eigensolver routines could be used for additional performance.

The eigenvalue obtained is $k^2$ hence complex or negative eigenvalues obtained from the solution of (15) represent modes that do not propagate (Lysmer & Drake 1972; Chang et al. 2010). The largest positive eigenvalue corresponds to the fundamental mode, and successively smaller positive eigenvalues are overtones.

For Rayleigh waves, the corresponding solution of (6) and (7) results in a quadratic eigensystem of the form

$$\begin{bmatrix} \omega^2 \mathbf{A}_x & 0 & 0 \\ 0 & \mathbf{B}_x & \mathbf{C}_x \\ 0 & 0 & \mathbf{A}_z \end{bmatrix} + k^2 \begin{bmatrix} \mathbf{B}_x & 0 & 0 \\ 0 & \mathbf{C}_x & \mathbf{A}_z \\ 0 & 0 & \mathbf{D}_z \end{bmatrix} + k \begin{bmatrix} \mathbf{C}_x & 0 & 0 \\ 0 & \mathbf{A}_z & \mathbf{D}_z \\ 0 & 0 & \mathbf{D}_z \end{bmatrix} \begin{bmatrix} u \\ \mathbf{U} \\ \mathbf{W} \end{bmatrix} = 0,$$

(16)

where the vectors $\mathbf{U}$ and $\mathbf{W}$ represent the spectral element polynomial approximations of the eigenfunctions of the horizontal and vertical oscillations respectively.

The Rayleigh matrix equation is a quadratic eigenproblem that has numerical stability issues (Tisseur & Meerbergen 2001) and is generally solved by restating the problem in one of the many companion forms to reduce the problem back to a general eigenproblem, as in existing Finite Element solutions (Haney & Tsai 2017). This is at the cost of doubling the size of the matrices.

The approach followed here is that advocated by Hammarling et al. (2013) that includes additional pre-conditioning of the matrices in addition to rearrangement into a companion form. In this approach, two scaling coefficients, $\gamma$ and $\delta$, are first computed using

$$\gamma = \sqrt{\sqrt{\frac{\mathbf{D}_x \mathbf{0} \mathbf{0}}{\mathbf{B}_x \mathbf{0} \mathbf{0}}} - \omega^2 \mathbf{A}_x \mathbf{0} \mathbf{0}}_2},$$

(17)

and

$$\delta = 2 \sqrt{\frac{\mathbf{D}_x \mathbf{0} \mathbf{0}}{\mathbf{B}_x \mathbf{0} \mathbf{0}}} - \omega^2 \mathbf{A}_x \mathbf{0} \mathbf{0}}_2 + \gamma \sqrt{\mathbf{0} \mathbf{C}_x \mathbf{0} \mathbf{0}}_2^1},$$

(18)

Then the Rayleigh wave eigensystem is rearranged with the scaling terms into second companion form

$$\begin{bmatrix} 0 & \delta \gamma \mathbf{C}_x & -I & 0 \\ \delta \gamma \mathbf{C}_x & 0 & 0 & -I \\ 0 & \delta \mathbf{D}_x - \omega^2 \mathbf{A}_x & 0 & 0 \\ 0 & 0 & -\gamma^2 \delta \mathbf{B}_x & 0 \\ 0 & 0 & 0 & -I \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \mathbf{W} \\ \mathbf{P}_x \\ \mathbf{P}_x \\ \mathbf{P}_x \end{bmatrix} = \alpha \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

(19)

The solution then can then be found for the eigenvalue $\alpha$, which gives the wave number

$$k = \gamma \alpha.$$

(20)

The addition of the pre-condition terms, $\gamma$ and $\delta$ have no physical meaning yet provide additional accuracy as shown in Appendix B. The right eigenvector components for the horizontal and vertical displacements, $\mathbf{U}$ and $\mathbf{W}$ are in the upper half of the eigenvector of the solution as indicated.

Both positive and negative real values of $k$ are valid propagating modes and represent propagation in the positive and negative $x$ direction. Similar to the Love wave case, the eigenvalue with the largest magnitude constitutes the fundamental mode and successively smaller magnitudes are higher order modes. There is a caveat in this case where it is common for repeated roots in the solution, that is, positive and negative eigenvalues with the same magnitudes (to machine precision) that correspond to the same mode travelling in opposite directions. These duplicate solutions should be considered as a single mode.

Given a solution for the wave number $k$ to the generalized eigenproblem for Love or Rayleigh waves, we can compute the phase velocity using

$$c(\omega) = \frac{\omega}{k}.$$

(21)

Similarly, group velocity is defined as

$$v_g(\omega) = \frac{\partial \omega}{\partial k},$$

(22)

and this can be computed from matrix equations (15) and (16) that express the relationship between $\omega$ and $k$.

For Love waves, the matrix eq. (15) can be reduced to a scalar equation by left multiplying by $\mathbf{V}^T$ to obtain

$$\omega \mathbf{V}^T \mathbf{A} - k^2 \mathbf{V}^T \mathbf{B} \mathbf{V} + \mathbf{V}^T \mathbf{C} \mathbf{V} = 0,$$

(23)

which can then be implicitly differentiated to obtain

$$v_g,\text{Love} = \frac{k \tilde{B}}{\omega \tilde{A}},$$

(24)

where $\tilde{B} = \mathbf{V}^T \mathbf{B} \mathbf{V}$ and $\tilde{A} = \mathbf{V}^T \mathbf{A}$.

For Rayleigh waves, once the eigenvalue and eigenvectors $\mathbf{U}$ and $\mathbf{W}$ are obtained, the complete solution of (16) is known in matrix form. Taking the same approach as for Love wave group velocity, Rayleigh wave group velocity can be computed by pre-multiplying (16) by the transposed right eigenvector to obtain a scalar equation and then differentiating to obtain

$$v_g,\text{Rayleigh} = \frac{2k \tilde{B} + \tilde{C}}{2\omega \tilde{A}},$$

(25)

where

$$\tilde{C} = [\mathbf{U}^T \mathbf{W}^T] \begin{bmatrix} \mathbf{0} & \mathbf{C}_x & \mathbf{0} \\ \mathbf{C}_x & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U},$$

(26)

$$\tilde{A} = [\mathbf{U}^T \mathbf{W}^T] \begin{bmatrix} \mathbf{0} & \mathbf{A}_z & \mathbf{0} \\ \mathbf{A}_z & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{W},$$

(27)

with $\tilde{B}$ having an expression of the same form as $\tilde{A}$.

Lastly, for Rayleigh wave ellipticity, the ratio of the horizontal amplitude over the vertical amplitude at the surface, often abbreviated to $H/V$, is simply computed from the first component of the horizontal and vertical eigenvectors, that is, given $\mathbf{U}$ and $\mathbf{W}$,

$$\frac{H}{V} = \frac{U_1}{W_1}.$$

(28)

In Fig. 1 we show plots of the log10 relative error between results obtained with the spectral element formulation presented here to the known analytical solution in the case of Love waves, and to the Thomson–Haskell approach in the case of Rayleigh waves. The horizontal dash red and blue lines are indicators of 5 and 1 per cent errors respectively and as can be seen from these summary plots, the accuracy is high for both Love and Rayleigh wave calculations. More detailed information on the earth models used and further detailed numerical analyses are shown in Appendix B.
Figure 1 The $\log_{10}$ relative error of the spectral element solution to a reference solution are shown for (a) Love wave phase velocity, (b) Love wave group velocity, (c) Rayleigh wave phase velocity, (d) Rayleigh wave group velocity and (e) Rayleigh wave ellipticity. In each of the plots, the inset graph shows the corresponding calculated dispersion curve or ellipticity in (e). The red and blue horizontal dashed lines represent 5 and 1 per cent error levels, respectively. For Love waves, the reference solution is an analytical solution and for Rayleigh waves the reference solution is computed using the Thomson–Haskell method.

In summary, we have developed a novel spectral element formulation for the solution of Love and Rayleigh surface wave dispersion problems. The benefits of this approach are three fold: first there is much greater flexibility in the representation of the earth model. The interpolation properties of the spectral elements allow exact or highly accurate representation of arbitrarily varying structure, and interface discontinuities. For example, in Fig. 2, an earth model is shown in (a) consisting of different order polynomials of isotropic
Figure 2 An arbitrary earth model consisting of a mixture of smoothly varying features over a half-space in (a), can be accurately represented in (b) with a spectral element mesh. Spectral element nodes are indicated with dots, Gauss–Lobatto–Legendre and Gauss–Laguerre–Legendre cardinal polynomials with faint lines, and the resulting interpolated earth model with heavy line. In this case the representation is exact (to machine precision). Similarly, in (c) the solution for the polynomial approximation of the amplitude function is well represented by the interpolation polynomials.

shear wave velocity in three 10km thick layers over a half-space with. In (b), we show the equivalent spectral element mesh where element boundaries are indicated with horizontal red dotted lines, the interpolation nodes are indicated with coloured dots for each element, and cardinal polynomials with faint lines, and the polynomial approximation in solid line. The interpolated representation in this case is exact to machine precision. The amplitude function is similarly approximated with the spectral element representation in (c). This flexibility allows us to accurately represent both layered models and smoothly varying models in a single framework.

Recent studies (Zion et al. 2018) using parametrizations allowing for polynomial variation of structure with depth have revealed mid-lithospheric velocity gradients with important implications for thermo-chemical composition, hence accurate modelling of interfaces and varying structure made available by the SEM presented here could be important in future surface wave studies.

While it is possible to approximate smooth earth models with fine layered models, this may require a large number of fine layers. As an example, in Fig. 3 we show Rayleigh wave ellipticity computed for a model with smoothly varying layers using the SEM and the Thomson–Haskell method with varying degrees of approximation. As can be seen from the Fig. 3(a), for the Thomson–Haskell method, the convergence is slow and a large number of fine layers are required to obtain an accurate solution. In contrast, convergence of the spectral element solution is far more rapid owing to the rapid convergence properties of SEMs (Pozrikidis 2005; Fichtner 2011). In examination of the layered approximations shown in parts (b), (c) and (d), we can see that a reasonable parametrization, for example (c), would lead to an error level of more than 10 per cent from more accurate approximations. This suggests that using an incorrect model parametrization in Rayleigh wave ellipticity inversion could result in poor estimates of the Earth.

Second, the method is accurate and stable with additional numerical tests that can be performed to verify correctness. The accuracy of the result can be improved with the use of higher order elements and subdivision of elements (commonly referred to as p and h refinement respectively in the Finite Element community) allowing this method to be confidently used in a broad range of problems. Examples of p refinement are shown in Fig. 4, where the representation of the first layer has used an increased order polynomial (i.e. larger number of interpolation nodes) which can be useful for improving the accuracy at higher frequency. The addition of the Gauss–Laguerre–Legendre element to represent the half-space is similarly useful for correctly modelling dispersion at lower frequencies. In Fig. 5 we show the adaptive Gauss–Laguerre–Legendre element for two different periods where the nodal spacing adapts to support smooth decay of the amplitude function with depth and therefore improving accuracy at lower frequencies. In summary, the spectral element mesh can be adapted to the velocity model and range of frequencies required in an efficient and stable manner.
Figure 3 In (a) we show the convergence of the spectral element method (solid black) relative to the Thomson–Haskell (dashed red) in calculating Rayleigh wave ellipticity for a smoothly varying model for a given number of parameters. The spectral element solution converges rapidly owing to its spectral convergence properties whereas the Thomson–Haskell method converges very slowly. Red dots along the Thomson–Haskell indicate the positions where examples of the layered model approximations are shown in (b), (c) and (d), which show successively better layered approximations to the true model, shown with a red curve in (b).

Lastly, this method naturally solves for all existing higher order modes simultaneously. Other techniques such as the Thomson–Haskell matrix method or numerical integration schemes require a separate solution for each mode. In this approach, the eigenvalues are obtained by solving either (15) or (16) and consist of all modes capable of being represented by the discretization.

We previously mentioned Love waves with imaginary eigenvalues as examples of modes that do not propagate, however, additional modes that represent surface waves that do not decay with depth are possible for both Love and Rayleigh waves. We have taken a simple heuristic approach to detecting non-existing modes in that the phase velocity of a mode cannot exceed the largest shear wave velocity. This criterion correctly detects the non-existence in homogeneous media of Love waves and of higher mode Rayleigh waves. This is a quick and generally effective strategy for removal of non-physical modes but it may not be sufficient and can potentially suffer from problems with numerical precision.

In Haney & Tsai (2017), an additional criteria based upon the decay rate of the eigenfunction was used and this is a potential solution that could equally be applied in the spectral element formulation developed here. An analogue in this formulation is the extent of the Laguerre element due to the automatic scaling. This can be used to give a measure of the decay rate of the wave and a practical choice can be made for some threshold beyond which we assume the mode is non-physical. To give an example, for the solution of Love wave dispersion in nearly homogeneous models (with a surface shear wave velocity within 0.1 per cent of the half-space velocity), the Laguerre element adapts and extends to a large depth, beyond the radius of the Earth. This is clearly unphysical and could be used as an additional criterion for eliminating potential unphysical modes.
Figure 4 An arbitrary earth model consisting of a mixture of smoothly varying features over a half-space in (a) can be accurately represented in (b) with a spectral element mesh, in this case with a higher order element near the surface to more accurately represent near surface features. Spectral element nodes are indicated with dots, Gauss–Lobatto–Legendre and Gauss–Laguerre–Legendre cardinal polynomials with faint lines, and the resulting interpolated earth model with heavy line. In this case the representation is exact (to machine precision). Similarly, in (c) the solution for the polynomial approximation of the amplitude function is well represented by the interpolation polynomials.

An additional caveat is that the accuracy of the higher order modes will require greater refinement of the mesh than the fundamental mode. The accuracy of the solution is limited by the ability of the spectral polynomial approximation to adequately represent the eigenfunction(s). As the mode number increases, the eigenfunction(s) become more oscillatory and hence the spectral polynomial approximation of the eigenfunction becomes less accurate. For brevity, the focus in this manuscript is on the fundamental mode, however the method and strategies for increasing the accuracy of the fundamental mode can be similarly be applied for higher order modes, allowing selected higher order modes to be modelled simultaneously in an inverse problem.

3 SURFACE WAVE INVERSION

In local Love or Rayleigh surface wave observations, a set of surface wave observations at discrete frequencies are used to solve for a plausible 1-D earth model. There are many different strategies for the solution of inverse problems, for example, probabilistic sampling (Sambridge & Mosgaard 2002) and gradient descent (Menke 1989; Aster et al. 2005). The focus of this section is on using an adjoint technique for computing the gradient for the later class of inversion.

In a common form of geophysical inversion, there will be a misfit function, \( L \), that requires minimization. The gradient descent method iteratively adjusts model parameters using the gradient, that is \( \frac{dL}{dm} \), where \( m \) is a vector of model parameters, until some termination condition is met. Examples include terminating when a misfit threshold is reached, or the incremental change in misfit becomes smaller than some threshold, or simply a fixed number of iterations have been performed.

For calculating the gradients of the misfit function with respect to the model parameters, we describe our approach in terms of a generalized eigensystem. In both the solution of Love and Rayleigh waves, the equation to solve, (15) or (19), can be written as a generalized eigensystem of the form

\[
A'\phi_r - \alpha B'\phi_r = 0, \tag{29}
\]

with matrices \( A' \) and \( B' \) functions of a set of earth model parameters of interest \( m \). For solutions of Love wave dispersion, the matrices \( A' \) and \( B' \) are defined as

\[
\alpha = k^2 \tag{30}
\]

\[
A' = \omega^2 A + C \tag{31}
\]

\[
B' = B \tag{32}
\]

\[
\phi_r = V, \tag{33}
\]

with the terms of the right-hand side defined in (15).
SEM for surface wave dispersion

Figure 5 The Gauss–Laguerre–Legendre element, starting from 30 km depth and representing the half-space can be scaled to optimally represent the decay of the amplitude function with depth. In (a) the eigenfunction for the solution to the model in Fig. 2 is shown for a period of 20 s. In (b) the same solution is shown for a period of 30 s. In the plots, the scaling of the Laguerre element interpolation nodes (red dots) is apparent. For example, in (a) the first nodal point is after the boundary at 30 km is at just over 40 km. In (b) this first nodal point is located closer to 60 km depth to represent the slower amplitude decay with depth of the longer period Love wave.

Similarly, for solutions of Rayleigh wave dispersion, we have

\[
A' = \begin{bmatrix}
0 & \delta y C_x & -I & 0 \\
\delta y C_z & 0 & 0 & -I \\
\delta (D_x - \omega^2 A_x) & 0 & 0 & 0 \\
0 & \delta (D_z - \omega^2 A_z) & 0 & 0
\end{bmatrix} \begin{bmatrix}
U \\
W \\
P_x \\
P_z
\end{bmatrix} \quad (34)
\]

\[
B' = \begin{bmatrix}
-\gamma^2 \delta B_x & 0 & 0 & 0 \\
0 & -\gamma^2 \delta B_z & 0 & 0 \\
0 & 0 & -I & 0 \\
0 & 0 & 0 & -I
\end{bmatrix} \quad (35)
\]

again with terms on the right-hand side defined in (19).

In Appendix C we provide the detailed derivation of the general method for computing the gradient of an arbitrary function of the solution of the surface wave dispersion problem. This can be summarized as follows: given a misfit function \( L \), we first compute its derivatives with respect to the general eigensolution of (29), that is \( \frac{\partial L}{\partial \phi} \) and \( \frac{\partial L}{\partial \alpha} \). These derivatives are often easily obtained algebraically. Automatic differentiation is also an option for complicated
misfit functions (Griewank 2000), although there has been limited success of automatic differentiation in seismic applications (Sambridge et al. 2007; Cau & Liao 2015).

In broad terms, the adjoint calculation proceeds with first solving the general eigenproblem in (29) using the Schur decomposition. The same Schur decomposition is reused to solve the linear system of equations

\[(A^T - \alpha B^T)\lambda_{1,N} = (1 - \phi_i \phi_i^T) \frac{dC}{d\phi_i}, \quad (37)\]

for vector \(\lambda_{1,N}\) while ensuring the solution is orthogonal to \(\phi_i\), the left eigenvector. Next compute the orthogonal scalar term \(\beta\)

\[\beta = -\frac{\frac{db}{dm} + \phi_i^TB\lambda_{1,N}^{\prime}}{\phi_i^TB\phi_i^{\prime}}. \quad (38)\]

Finally, the gradient can be computed from

\[\frac{dC}{dm} = -(\lambda_{1,N}^{\prime} + \beta \phi_i)^T \left( \frac{dA}{dm} - \sigma^2 \frac{dB}{dm} \right) \phi_i. \quad (39)\]

A major benefit of the approach is that only a single decomposition of the eigensystem is required for both the solution of the surface wave dispersion problem and adjoints. The additional computational cost of computing the gradient of model parameters with respect to a misfit function is negligible relative to the cost of the forward model. This is illustrated in Table 1 where we show the computation cost of 1000 solutions for the forward model alone, the forward model plus gradient with phase velocity, group velocity and H/V misfit functions. By directly computing the Schur decomposition rather than using DGGEV, the general eigensystem routine from LAPACK (Anderson et al. 1999), additional computational cost savings are made through omitting unnecessary calculations.

As a demonstration of the adjoint technique we have developed for surface wave dispersion, we show a simple numerical example. Using the reference model as defined in Table B1, a synthetic true observation was generated at 0.1 Hz for Love and Rayleigh phase velocity, Love and Rayleigh group velocity and Rayleigh H/V ratio. The spectral element model was parameterized using an isotropic shear wave velocity model with density and isotropic p-wave velocity obtained using an empirical model (Brocher 2005). The shear wave velocity in the first layer was then varied from 0.5 km s\(^{-1}\) to 4.5 km s\(^{-1}\) (with the true value being 3.0 km s\(^{-1}\)) with steps of 1 m s\(^{-1}\). At each value of shear wave velocity in the first layer, the misfit and the gradient of the misfit relative to the velocity in the first layer was computed using the adjoint technique outlined above. Additionally a numerical approximation of the gradient was calculated using finite differences. The results of these experiments are shown in Fig. 6, where the left column shows the gradient computed with the adjoint technique (black line) with the finite difference approximation of the gradient shown with red circles (every hundredth point for clarity). Shown in the right-hand column are the corresponding misfit functions as the shear wave velocity is varied in the top most layer.

As is evident, the adjoint method compares well to the numerical approximation giving confidence in its correct formulation. It also shows the non-linearity of the problems considered, as both Rayleigh wave group velocity (g) and ellipticity (i) have local minima distant from the true value at 3 km s\(^{-1}\). Additional comparisons are shown in Appendix C comparing the numerical accuracy of the adjoint method presented here to existing techniques that use a layered model approximation.

### 3.1 Inversion of Love wave phase velocity

For a first demonstration of the spectral element surface wave dispersion method and adjoint in a realistic inverse problem, some synthetic Love wave phase velocity data are created between the periods of 3 and 30 s and Gaussian noise is added with a standard deviation of 50 m s\(^{-1}\).

A least squares misfit function of the form

\[L(m) = \sum_j \frac{(G(m)_{ij} - d_j)^2}{2\sigma_j^2} + \sum_i M_i \frac{(m_i - \bar{m}_i)^2}{2\sigma_m^2}. \quad (40)\]

is used where \(m\) is the model vector of isotropic shear wave velocities with depth expressed as polynomials of arbitrary order, \(G(m)\) is the forward model prediction for the \(i\)th observed phase velocity \(d_i\), \(\sigma_j\) is the independent Gaussian standard deviation of the errors on the data, \(\bar{m}_i\) is a reference model used in the inversion and \(\sigma_m\) is the standard deviation of expected deviations from the reference model. This formulation is analogous to Tikhonov regularization schemes of the form \(\min_0 ||G(m) - d||_2^2 + \lambda ||m - \bar{m}||_2^2\) with the term \(\lambda = \frac{2\sigma_j^2}{\sigma_m^2}\). Similarly, from a Bayesian perspective, this corresponds to Gaussian likelihood function with a Gaussian prior about a mean reference model. Then \(L(m)\) becomes the negative log of the posterior up to an additive constant.

It should be stressed here, that the inversion approach used here is in no way meant to represent the optimal choice of scheme for the solution of these type of problems, merely as a demonstration of the incorporation of the forward model and gradient calculation into a general optimization framework. Determining starting models for inversions (Xia et al. 1999; Haney & Tsai 2015, 2017), appropriate regularization constraints to incorporate prior knowledge, and choosing gradient step sizes and stopping criteria are all important components of geophysical inversion and areas of active research. Inversion of measured field data should include advances from these areas of research but the simple approach applied here suffices for a demonstration of the new forward modelling and adjoint gradient calculation methods.

The choice of prior variance \(\sigma_m^2\) could have a significant bearing on the result, as this represents a tuneable damping parameter varying the relative importance of observations to prior constraint. In the results presented three values are used representing tight (\(\sigma_m = 125m s^{-1}\)), medium (\(\sigma_m = 250m s^{-1}\)) and wide (\(\sigma_m = 500m s^{-1}\)) bounds, with the expectation that a well-tuned choice would fall somewhere within this range. To demonstrate the flexibility of the spectral element surface wave dispersion method, both a polynomial model and a layered model are inverted for. The order and number of layers are chosen so they have the same total number of model parameters (five in the examples to follow). In non-linear problems, the starting model can greatly influence the final optimal model and this is especially the case where there are local minima in which a gradient descent algorithm could easily become trapped. To more fully explore the solution, ten individual inversions are performed.

<table>
<thead>
<tr>
<th></th>
<th>DGGEV</th>
<th>Phase</th>
<th>Group</th>
<th>H/V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Love</td>
<td>0.184</td>
<td>0.190</td>
<td>0.194</td>
<td>N/A</td>
</tr>
<tr>
<td>Rayleigh</td>
<td>2.835</td>
<td>2.758</td>
<td>2.785</td>
<td>2.762</td>
</tr>
</tbody>
</table>
In a simple test, we compute the gradient for a misfit function of a single observation as shear wave velocity in the upper most layer is varied. In the left column the gradient is computed using the method presented here and shown with a black line. The red circles represent the numerical approximation using a simple finite difference approximation. In the right column is shown the corresponding misfit function as a function of the shear wave velocity in the first layer. Love wave phase velocity is shown in (a) and (b), Rayleigh wave phase velocity in (c) and (d), Love wave group velocity in (e) and (f), Rayleigh wave group velocity in (g) and (h), and Rayleigh wave ellipticity in (i) and (j).

Starting from a randomized initial earth model. This is generated by sampling from a Gaussian about the reference model, ̂m, which corresponds to the second term in (40).

Fig. 7 shows the results of inversions with a polynomial parametrization with each row corresponding to progressively weaker damping (or wider priors). The left column shows the final models obtained in varying coloured solid lines with the reference model plotted with a green dotted line and the true model plotted with a red dotted line. The right column shows the noisy data with red points and the fits to the data with coloured lines for each inversion.

In the tightest σ_m, the model recovery is very good except for the top and bottom of the first layer where the optimized models tend toward the reference model (shown in green). This indicates an insensitivity of the observations to the upper and lower most parts of the domain and the model is driven toward the reference model for these extremal points. This is most apparent at the 10 km interface in Fig. 7(a).

In the second row, where the medium level of damping is used, the results are much the same except one inversion has chosen a different class of model with high shear wave velocity at the surface. The fits to the observations shown in (f) are nonetheless still acceptable. Lastly, in the final row where the weakest constraint is used, there is a more even split between the two classes of final model, that is, those close to the true model and those with high shear wave velocity at the surface. The level of damping plays two...
Figure 7 The results of inversion of Love wave phase velocity observations with a smooth model parametrization over a half-space using the gradient descent approach. Ten inversions are performed starting from randomized models, and the left column [i.e. (a), (c) and (e)] shows the final models obtained in varying colour solid lines. The reference model is plotted with a green dotted line, and the true model with a red dotted line. The right column [i.e. (b), (d) and (f)] shows the final optimal model predictions with varying coloured lines with data plotted with red points and error bars. The results in the first row use narrowest prior on deviations from the reference model, the next is the medium and the final is the widest constraint.
Figure 8 The results of inversion of Love wave phase velocity observations with a layered model parametrization over a half-space using the gradient descent approach. Ten inversions are performed starting from randomized models, and the left column [i.e. (a), (c) and (e)] shows the final models obtained in varying colour solid lines. The reference model is plotted with a green dotted line, and the true model with a red dotted line. The right column [i.e. (b), (d) and (f)] shows the final optimal model predictions with varying coloured lines with data plotted with red points and error bars. The results in the first row use the narrowest prior on deviations from the reference model, the next is the medium and the final is the widest constraint.
Figure 9 The results of inversion of Rayleigh wave ellipticity observations with a smooth model parametrization over a half-space using the gradient descent approach. Ten inversions are performed starting from randomized models, and the left column [i.e. (a), (c) and (e)] shows the final models obtained in varying colour solid lines. The reference model is plotted with a green dotted line, and the true model with a red dotted line. The right column [i.e. (b), (e) and (f)] shows the final optimal model predictions with varying coloured lines with data plotted with red points and error bars. The results in the first row use the narrowest prior on deviations from the reference model, the next is the medium and the final is the widest constraint.
Figure 10  The results of inversion of Rayleigh wave ellipticity observations with a layered model parametrization over a half-space using the gradient descent approach. Ten inversions are performed starting from randomized models, and the left column [i.e. (a), (c) and (e)] shows the final models obtained in varying colour solid lines. The reference model is plotted with a green dotted line, and the true model with a red dotted line. The right column [i.e. (b), (d) and (f)] shows the final optimal model predictions with varying coloured lines with data plotted with red points and error bars. The results in the first row use the narrowest prior on deviations from the reference model, the next is the medium and the final is the widest constraint.
parts in this inversion, first as a constraint on the solution as is typical, but second for choosing the initial random model. As the \( \sigma_m \) is increased, the starting model will be further away from the truth, and this can increase the likelihood of finding local minima.

If a layered model is inverted for, instead of a smooth polynomial model, the results are similar as shown in Fig. 8. The fits to the data for the tightest and medium damping are indistinguishable; all solutions have converged to a similar model that is a faithful representation of the true model shown in red. For the weakest constraint, Fig. 8 shows a similar bi-modality between a models approximating the true synthetic model and another class of models with high velocity near the surface that was observed when using the smooth parametrization. In the fits for the weakest constraint, both class of models fit the observations to within the noise level.

### 3.2 Rayleigh wave ellipticity

In a second example, synthetic observations of Rayleigh wave ellipticity are created for the same periods as the Love wave example, that is between 3 and 30 s. Gaussian noise with a standard deviation of 0.02 is added which equates to a noise level of approximately 5 per cent. The misfit function is the same form as (40) from the previous Love wave phase velocity example, and the true model remains the same.

We show the results from inverting using a smooth parametrization in Fig. 9. In each of the model plots in the left column, the shape of the structure is recovered with increasing spread as the damping is weakened. Unlike the Love phase velocity example, the shear wave velocity at the surface is closer to the true model than the reference model, indicating greater sensitivity to the near surface. In all inversions, as shown in the right column, the fits to the observations in the right column are all within the noise level with little spread.

For a comparison, we show the result of inverting using the same observations and method with a layered parametrization in Fig. 10. The results are similar in character where the models obtained from 10 different inversions are similar, with the spread increasing as damping is weakened. Of particular interest is the fits seem to have greater variability at higher frequencies and may indicate that a layered model approximating smoothly varying structure may perform poorly in Rayleigh ellipticity inversions.

We have demonstrated with some simple synthetic examples an accurate and efficient method for both modelling and computing the gradients of a misfit function with respect to model parameters in Rayleigh wave ellipticity problems. This could be of benefit in the inversion of Rayleigh wave ellipticity observations to provide better understanding of the near surface structure of the Earth.

### 4 CONCLUSIONS

This paper has presented a derivation of an SEM for the solution of 1-D Love and Rayleigh surface wave propagation using a flat Earth approximation. This new formulation allows greater flexibility in the modelling of proposed earth models to include both smoothly varying structure, homogeneous layers with discontinuities, and any combination thereof. In addition, this derivation introduced the use of Laguerre polynomials coupled with the more common Gauss–Lobatto–Legendre spectral elements to improve the accuracy of the solution at longer periods. Several examples validating the accuracy of the method are discussed in Appendices A–C. A key point of the new method is that accuracy can be reliably tuned across a range of frequencies by the common adaptive refinement techniques of spectral elements, namely subdivision of elements and increasing element polynomial order. In summary, this method provides a new robust, accurate and flexible method for modelling Love and Rayleigh surface waves.

In addition to the general forward modelling method, we also developed the adjoint method for calculation of the gradient of a function of the solution to the eigenproblem with respect to the model parameters. In our implementation, this has been flexibly implemented with C++ templates to support a variety of 1-D earth model parametrizations, such as, isotropic \( v_p \) with empirical relationships for \( v_p \) and density to give one example. A single Schur decomposition is performed that can be used to solve both the generalization eigenforward model, and the linear system of equations for the adjoint. This provides a solution to the gradient of a misfit function with respect to model parameters almost at no overhead in addition to the forward model solution.

Two synthetic inversion problems were used to illustrate the application of the gradient calculation based on fitting of Love wave phase velocity and Rayleigh wave ellipticity observations. These examples demonstrate the effectiveness of the approach developed here for computing gradient for the optimization of earth models to fit dispersion and ellipticity observations. The methods presented here may have additional utility in computing spectral element crustal models that approximate observed surface wave dispersion in the crust to improve full waveform inversion efficiency (Fichtner & Igel 2008).

An aspect not explored in this work is the inversion for the location of discontinuities. In principle there is no impediment to extending the approach developed here to incorporate gradient calculations on the location of interfaces. Alternatively, these methods could be coupled to an Markov chain Monte Carlo (MCMC) inversion where locations of discontinuities are perturbed or trans-dimensional inversions where the number of layers is variable (Bodin et al. 2012a; Dettmer et al. 2012). Lastly, extension to 2-D or 3-D models could be possible and may allow more accurate modelling to account for lateral heterogeneity and sloping layers, although the computational expense may prove prohibitive.

### ACKNOWLEDGEMENTS

The author would like to thank Malcolm Sambridge for his support and encouragement throughout the gestation of this work and for feedback on earlier versions of this manuscript.

The author would like to thank Brian L. N. Kennett and Andrew Valentine for feedback on an earlier version of this manuscript. Stephen Roberts is thanked for his advice in aspects of the derivation.

An earlier version of this manuscript benefitted from constructive reviews by Matthew Haney and an anonymous reviewer for which we are grateful.

The author is supported by an Australian Government Research Training Program (RTP) Scholarship.

Aspects of this research were supported under Australian Research Council Discovery grant scheme, project DP110102098. The author also acknowledges contributions from Geoscience Australia through the Data science in solid Earth geophysics project.
REFERENCES


APPENDIX A: SPECTRAL ELEMENT METHOD FOR LOVE AND RAYLEIGH SURFACE WAVES

This appendix provides the detailed derivation of the new spectral element solution to 1-D surface wave dispersion for Love and Rayleigh waves introduced in Section 2.

A1 Love waves

Restating (5)

$$\omega^2 \rho(z) V(\omega, k, z) - k^2 N(z) V(\omega, k, z) + \frac{d}{dz} \left[ L(z) \frac{dV(\omega, k, z)}{dz} \right] = 0.$$  \hfill (A1)

The domain is divided into $E$ elements from the surface, $z = 0$ to the half-space $z = z_{\text{half-space}}$ and then forming the weak form of the equation by multiplying by a smooth test function $v$ and integrating to obtain

$$\omega^2 \sum_{e=1}^{E} \int_{z_{e-1}}^{z_{e}} \rho(\omega, k, z) v \, dz - k^2 \sum_{e=1}^{E} \int_{z_{e-1}}^{z_{e}} N(z) V(\omega, k, z) v \, dz + \sum_{e=1}^{E} \int_{z_{e-1}}^{z_{e}} \frac{d}{dz} \left[ L(z) \frac{dV(\omega, k, z)}{dz} \right] v \, dz = 0.$$  \hfill (A2)

The function $v$ is arbitrary except for the fact that it must obey the boundary conditions equally imposed on the solution for the displacement $V(\omega, k, z)$.

The last term can then be integrated by parts to obtain

$$\omega^2 \sum_{e=1}^{E} \int_{z_{e-1}}^{z_{e}} \rho(\omega, k, z) v \, dz - k^2 \sum_{e=1}^{E} \int_{z_{e-1}}^{z_{e}} N(z) V(\omega, k, z) v \, dz - \sum_{e=1}^{E} \int_{z_{e-1}}^{z_{e}} L(z) \frac{dV(\omega, k, z)}{dz} v \, dz + \sum_{e=1}^{E} L(z) \frac{dV(\omega, k, z)}{dz} v \bigg|_{z_{e-1}}^{z_{e}} = 0.$$  \hfill (A3)

Recall that the displacement $V(\omega, k, z)$ and the traction $L(z) \frac{dV(\omega, k, z)}{dz}$ must be continuous everywhere requiring that the last term perfectly cancels except for the surface and basement values leaving

$$\omega^2 \sum_{e=1}^{E} \int_{z_{e-1}}^{z_{e}} \rho(\omega, k, z) v \, dz - k^2 \sum_{e=1}^{E} \int_{z_{e-1}}^{z_{e}} N(z) V(\omega, k, z) v \, dz - \sum_{e=1}^{E} \int_{z_{e-1}}^{z_{e}} L(z) \frac{dV(\omega, k, z)}{dz} v \, dz + L(z_{\text{half-space}}) \frac{dV(\omega, k, z)}{dz} v(z_{\text{half-space}}) - L(0) \frac{dV(\omega, k, 0)}{dz} v(0) = 0.$$  \hfill (A4)

This can be reduced further by using the fact that the traction must vanish at the surface, hence $L(0) \frac{dV(\omega, k, 0)}{dz} = 0$, and

$$\omega^2 \sum_{e=1}^{E} \int_{z_{e-1}}^{z_{e}} \rho(\omega, k, z) v \, dz - k^2 \sum_{e=1}^{E} \int_{z_{e-1}}^{z_{e}} N(z) V(\omega, k, z) v \, dz - \sum_{e=1}^{E} \int_{z_{e-1}}^{z_{e}} L(z) \frac{dV(\omega, k, z)}{dz} v \, dz + L(z_{\text{half-space}}) \frac{dV(\omega, k, z)}{dz} v(z_{\text{half-space}}) = 0.$$  \hfill (A5)

The boundary condition at depth is replaced with a Laguerre element that satisfies the decay of $V(\omega, k, z)$ to zero as $z$ tends to infinity. A key component of the SEM is the use of Lagrange nodal interpolation polynomials with nodal points arranged at the zeros of Lobatto polynomials. The domain for these Gauss–Legendre–Lobatto nodes in 1-D is $-1 \ldots 1$ so each of the element integrals requires transformation to this domain. This is achieved with a simple linear transformation so that

$$z_\epsilon(\zeta) = z_{e-1} + \frac{\zeta + 1}{2} (z_e - z_{e-1}) = z_{e-1} + \frac{(\zeta + 1) \Delta z_e}{2},$$  \hfill (A6)

where $\zeta$ is the spectral element coordinate. The substitutions for the required change of variables are

$$dz = \frac{\Delta z_e}{2} d\zeta,$$

$$d\zeta = \frac{2}{\Delta z_e} dz.$$  \hfill (A7)

(A8)

With this linear transform, the system of equations becomes

$$\omega^2 \sum_{e=1}^{E} \int_{-1}^{1} \rho(\omega, k, z) v \, d\zeta - k^2 \sum_{e=1}^{E} \int_{-1}^{1} N(z) V(\omega, k, z) v \, d\zeta - \sum_{e=1}^{E} \int_{-1}^{1} L(z) \frac{dV(\omega, k, z)}{dz} v \, d\zeta = 0.$$  \hfill (A9)

If each of the varying functions of $\zeta$ are approximated with Gauss–Legendre–Lobatto polynomial approximations, that is

$$f(\zeta) \approx \sum_{i=0}^{P} f_i \Lambda_i^{(P)}(\zeta),$$  \hfill (A10)
where $\ell^{(p)}_i$ is the $P$th-order Lagrange interpolating polynomial, $\xi_i$ is the $i$th Lobatto collocation point for and $f_i$ is short hand for $f(\xi_i^{(p)})$.

The integrals of functions of $\xi$ can then be approximated

$$\int_{-1}^{1} f(\xi) \, d\xi \approx \sum_{i=0}^{P} f_i \omega_i^{(p)},$$

(A11)

where $\omega_i^{(p)}$ is the quadrature weight.

Now taking the first two terms of (A9) in turn, the first two are

$$\omega^2 \sum_{e=1}^{E} \frac{\Delta z_e}{2} \int_{-1}^{1} \rho V(\omega, k, z) v \, d\xi \approx \omega^2 \sum_{e=1}^{E} \frac{\Delta z_e}{2} \sum_{i=0}^{P} v_i \rho_i \omega_i,$$

(A12)

and

$$-k^2 \sum_{e=1}^{E} \frac{\Delta z_e}{2} \int_{-1}^{1} N(\xi)V(\omega, k, z) v \, d\xi \approx -k^2 \sum_{e=1}^{E} \frac{\Delta z_e}{2} \sum_{j=0}^{P} v_j N_j \omega_j.$$

(A13)

For the third term, the derivatives of the functions add a complication. Given the approximation of the functions of $\xi$ with polynomials, the derivatives of said functions are given by

$$\frac{df(\xi)}{d\xi} \approx \sum_{i=0}^{P} f_i \frac{d\ell^{(p)}_i}{d\xi}(\xi_i^{(p)}),$$

(A14)

and the third term can be written

$$-k^2 \sum_{e=1}^{E} \frac{\Delta z_e}{2} \int_{-1}^{1} L(\xi) \frac{dW}{d\xi} \, d\xi \approx \left[-k^2 \sum_{e=1}^{E} \frac{\Delta z_e}{2} \sum_{j=0}^{P} v_j L_j \left( \sum_{k=0}^{P} V_k \frac{d\ell^{(p)}_k}{d\xi}(\xi_k^{(p)}) \right) \left( \sum_{l=0}^{P} \frac{d\ell^{(p)}_l}{d\xi}(\xi_l^{(p)}) \right) \right].$$

(A15)

in which the ordering of summation can be rearranged to

$$\sum_{e=1}^{E} \frac{2 \Delta z_e}{\Delta x_e} \sum_{i=0}^{P} v_i \left[ \sum_{k=0}^{P} V_k \left( \sum_{l=0}^{P} w_l L_j \left( \frac{d\ell^{(p)}_k}{d\xi}(\xi_k^{(p)}) \right) \frac{d\ell^{(p)}_l}{d\xi}(\xi_l^{(p)}) \right) \right].$$

(A16)

Combining these, the full equation becomes

$$\sum_{e=1}^{E} \sum_{i=0}^{P} v_i \left[ \omega^2 \frac{\Delta z_e}{2} V_i \rho_i w_i - k^2 \frac{\Delta z_e}{2} N_i V_i w_i - \frac{2 \Delta z_e}{2} \sum_{k=0}^{P} V_k \left( \sum_{l=0}^{P} w_l L_j \left( \frac{d\ell^{(p)}_k}{d\xi}(\xi_k^{(p)}) \right) \frac{d\ell^{(p)}_l}{d\xi}(\xi_l^{(p)}) \right) \right] = 0.$$  

(A17)

Now since the interpolation polynomials are orthogonal, the $v_i$‘s are linearly independent with the exception that at element boundaries the condition $\omega^2 = \omega^2 v_i$ must be met for continuity of the test function $v$. This means that internal collocation points must satisfy

$$\left[ \omega^2 \frac{\Delta z_e}{2} V_i \rho_i w_i - k^2 \frac{\Delta z_e}{2} N_i V_i w_i - \frac{2 \Delta z_e}{2} \sum_{k=0}^{P} V_k \left( \sum_{l=0}^{P} w_l L_j \left( \frac{d\ell^{(p)}_k}{d\xi}(\xi_k^{(p)}) \right) \frac{d\ell^{(p)}_l}{d\xi}(\xi_l^{(p)}) \right) \right] = 0.$$  

(A18)

These independent equations can be assembled into a matrix equation

$$\left[ \omega^2 A - k^2 B - C \right] \mathbf{V} = 0,$$

(A19)

where the matrix $A$ has elements defined by (A12), matrix $B$ has elements defined by (A13), matrix $C$ has elements defined by (A16), and vector $\mathbf{V}$ is the polynomial approximation of the amplitude function $V(\omega, k, z)$.

### A2 Rayleigh waves

For brevity, only the derivation for the horizontal displacement equation is shown with the derivation for the vertical displacement equation being exactly the same with some terms interchanged. Restating (6)

$$\rho \omega^2 U(\omega, k, z) - k^2 A(\xi) U(\omega, k, z) - k F(\xi) \frac{dW(\omega, k, z)}{dz} + \frac{d}{dz} \left[ \left( \frac{dU(\omega, k, z)}{dz} - k W(\omega, k, z) \right) \right] = 0.$$  

(A20)

The domain is divided into $E$ unevenly sized elements from the surface, $z = 0$, to the basement, $z = z_{\text{half-space}}$, and then forming the weak version of the equation by multiplying by a smooth test function, $v$, and integrating. After these steps the equation becomes

$$\omega^2 \int_{e=1}^{E} \rho U(\omega, k, z) v \, dz - k^2 \int_{e=1}^{E} \int_{z_e}^{z_{e+1}} A(z) U(\omega, k, z) v \, dz - k \int_{e=1}^{E} \int_{z_e}^{z_{e+1}} F(z) \frac{dW(\omega, k, z)}{dz} v \, dz$$

$$+ \frac{d}{dz} \left[ L(z) \left( \frac{dU(\omega, k, z)}{dz} - k W(\omega, k, z) \right) \right] v \, dz = 0.$$  

(A21)
Integrating the last term by parts gives

\[
\omega^2 \sum_{e=1}^{E} \int_{z_{e-1}}^{z_e} \rho U(\omega, k, z)v \, dz - k^2 \sum_{e=1}^{E} \int_{z_{e-1}}^{z_e} A(z)U(\omega, k, z)v \, dz - k \sum_{e=1}^{E} \int_{z_{e-1}}^{z_e} F(z) \frac{dW(\omega, k, z)}{dz} v \, dz \\
- \sum_{e=1}^{E} \int_{z_{e-1}}^{z_e} L(z) \left( \frac{dU(\omega, k, z)}{dz} - kW(\omega, k, z) \right) \frac{dv}{dz} \, dz + \sum_{e=1}^{E} L(z) \left( \frac{dU(\omega, k, z)}{dz} - kW(\omega, k, z) \right) \bigg|_{z_{e-1}}^{z_e} = 0. \tag{A22}
\]

The term \( L(z) \left( \frac{dU(\omega, k, z)}{dz} - kW(\omega, k, z) \right) \) is the amplitude of the stress \( \tau_{zz} \) which must be continuous across element boundaries. This means that most of these terms cancel will each other in the summation leaving only the end points at \( z = 0 \) and \( z = z_{\text{half-space}} \), that is

\[
\omega^2 \sum_{e=1}^{E} \int_{z_{e-1}}^{z_e} \rho U(\omega, k, z)v \, dz - k^2 \sum_{e=1}^{E} \int_{z_{e-1}}^{z_e} A(z)U(\omega, k, z)v \, dz - k \sum_{e=1}^{E} \int_{z_{e-1}}^{z_e} F(z) \frac{dW(\omega, k, z)}{dz} v \, dz \\
- \sum_{e=1}^{E} \int_{z_{e-1}}^{z_e} L(z) \left( \frac{dU(\omega, k, z)}{dz} - kW(\omega, k, z) \right) \frac{dv}{dz} \, dz + \sum_{e=1}^{E} L(z) \left( \frac{dU(\omega, k, z)}{dz} - kW(\omega, k, z) \right) \bigg|_{z_{e-1}}^{z_e} = 0. \tag{A23}
\]

At \( z = 0 \) the traction \( \tau_{zz} \) must vanish and therefore the term \( L(z) \left( \frac{dU(\omega, k, z)}{dz} - kW(\omega, k, z) \right) \) must be zero. For now it is assumed the basement will be treated separately, leaving

\[
\omega^2 \sum_{e=1}^{E} \int_{z_{e-1}}^{z_e} \rho U(\omega, k, z)v \, dz - k^2 \sum_{e=1}^{E} \int_{z_{e-1}}^{z_e} A(z)U(\omega, k, z)v \, dz - k \sum_{e=1}^{E} \int_{z_{e-1}}^{z_e} F(z) \frac{dW(\omega, k, z)}{dz} v \, dz \\
- \sum_{e=1}^{E} \int_{z_{e-1}}^{z_e} L(z) \left( \frac{dU(\omega, k, z)}{dz} - kW(\omega, k, z) \right) \frac{dv}{dz} \, dz = 0. \tag{A24}
\]

Lastly, the right-hand side is rearranged to collect polynomial terms of \( k \) giving

\[
\omega^2 \sum_{e=1}^{E} \int_{z_{e-1}}^{z_e} \rho U(\omega, k, z)v \, dz - k^2 \sum_{e=1}^{E} \int_{z_{e-1}}^{z_e} A(z)U(\omega, k, z)v \, dz - k \sum_{e=1}^{E} \int_{z_{e-1}}^{z_e} F(z) \frac{dW(\omega, k, z)}{dz} v \, dz \\
- \sum_{e=1}^{E} \int_{z_{e-1}}^{z_e} L(z) \frac{dU(\omega, k, z)}{dz} - kW(\omega, k, z) \frac{dv}{dz} \, dz = 0. \tag{A25}
\]

A change of variable is required to adjust the integrals in each element to span from \(-1 \ldots 1\). The affine transform from the \( \zeta \) to \( z \) coordinate in each element is given by

\[
z_e(\zeta) = z_{e-1} + \frac{\zeta + 1}{2} (z_e - z_{e-1}) = z_{e-1} + \frac{(\zeta + 1) \Delta z_e}{2}. \tag{A26}
\]

From these, the substitutions for the change of variables are

\[
dz = \frac{\Delta z_e}{2} d\zeta
\]

\[
\frac{d}{dz} = \frac{2}{\Delta z_e} \frac{d}{d\zeta}
\]

Substituting these into the equation gives

\[
\omega^2 \sum_{e=1}^{E} \int_{-1}^{1} \rho U^e \psi^e \, d\zeta - k^2 \sum_{e=1}^{E} \int_{-1}^{1} A^e U^e \psi^e \, d\zeta - k \sum_{e=1}^{E} \int_{-1}^{1} F^e \frac{dW^e}{d\zeta} \psi^e - L^e W^e \frac{dv^e}{d\zeta} \, d\zeta - \sum_{e=1}^{E} \frac{2}{\Delta z_e} \int_{-1}^{1} L^e \frac{dv^e}{d\zeta} d\zeta = 0. \tag{A29}
\]

where the superscript \( e \) represents the piece wise function in element \( e \). In a similar manner to the Love wave derivation, integrals can be replaced with their approximate Gauss–Lobatto–Legendre quadrature rule equivalents, that is

\[
\Phi(\zeta) \approx \sum_{j=0}^{N} \phi_j(\zeta) \Phi(\zeta_j), \tag{A30}
\]

where \( \Phi \) is the interpolated function, \( \phi_j \) is the \( j \)th Lagrange cardinal polynomial and \( \zeta_j \) is the \( j \)th Lobatto point. The substitutions of each of the terms of (A29) follow.

The first two terms are straight forward substitutions

\[
\int_{-1}^{1} \rho U^e \psi^e \, d\zeta \approx \sum_{j=0}^{P} w_j \rho_j U_j \psi_j, \tag{A31}
\]
and
\[ \int_{-1}^{1} A^e U^e v^e \, d\xi \approx \sum_{j=0}^{P} w_j A_j U_j v_j. \]  
(A32)

The third term has derivative terms that can be computed from the polynomial representation. It is useful to arrange the summations with the outer summation over the smooth test function \( v \), for this reason the integral is initially split the integral where individual summations are computed the rearranged to
\[ \int_{-1}^{1} F^e \frac{dW^e}{d\xi} v^e - L^e W^e \frac{dv^e}{d\xi} \, d\xi \approx \sum_{j=0}^{P} w_j v_j \sum_{k=0}^{P} W_k I_k^e(\xi_j) - \sum_{k=0}^{P} \sum_{j=0}^{P} w_k L_k W_k \sum_{j=0}^{P} v_j I_j^e(\xi_k) \]
\[ \approx \sum_{j=0}^{P} v_j \sum_{k=0}^{P} \left[ F_j w_j W_k I_k^e(\xi_j) - L_k w_k I_k^e(\xi_j) v_j \right] \approx \sum_{j=0}^{P} \sum_{k=0}^{P} W_k \left[ F_j w_j I_k^e(\xi_j) - L_k w_k I_k^e(\xi_j) v_j \right] \]  
(A33)

The last term similarly has derivatives the summation is re-ordered so that the outer summation is over the smooth test function \( v \)
\[ \int_{-1}^{1} L^e \frac{dU^e}{d\xi} v^e \, d\xi \approx \sum_{j=0}^{P} w_j L_j \sum_{k=0}^{P} U_k \sum_{l=0}^{P} v_l I_l^e(\xi_j) \approx \sum_{j=0}^{P} \sum_{k=0}^{P} w_j L_j \sum_{l=0}^{P} U_k \sum_{j=0}^{P} v_l I_l^e(\xi_j) . \]  
(A34)

Now since the interpolation polynomials are orthogonal, the \( v^e \)'s are linearly independent with the exception that at element boundaries the condition \( v^e = v_0^{e+1} \) must be met for continuity of the test function \( v \). This means that when combining all substitutions above, the internal collocation points must satisfy
\[ \sum_{j=0}^{P} \sum_{k=0}^{P} \left[ -\omega^2 \frac{\Delta \xi^e}{2} w_j \rho_j^e U_j^e + k^2 \frac{\Delta \xi^e}{2} w_j L_j^e + k \sum_{k=0}^{P} r_1^e \left( w_j F_k I_k^e(\xi_j) - w_k L_j I_k^e(\xi_j) \right) + \frac{2}{\Delta \xi^e} \sum_{k=0}^{P} U_k^e \sum_{l=0}^{P} I_l^e(\xi_j) r_1^e \right] w_j = 0. \]  
(A35)

The derivation of the vertical displacement equation is omitted and only the final result presented here as the procedure is the same as for the horizontal equation. The internal collocation points for the vertical equation similarly must satisfy
\[ \sum_{j=0}^{P} \sum_{k=0}^{P} \left[ -\omega^2 \frac{\Delta \xi^e}{2} w_j \rho_j^e r_2^e + k^2 \frac{\Delta \xi^e}{2} w_j L_j r_2^e + k \sum_{k=0}^{P} r_2^e \left( F_k w_j I_k^e(\xi_j) - w_k L_j I_k^e(\xi_j) \right) + \frac{2}{\Delta \xi^e} \sum_{k=0}^{P} r_2^e \sum_{l=0}^{P} I_l^e(\xi_j) r_2^e \right] w_j = 0. \]  
(A36)

Both (A35) and (A36) can then be combined into a single coupled matrix equation to give
\[ \begin{bmatrix} \omega^2 - A_1 & 0 \\ 0 & A_2 \end{bmatrix} + k^2 \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix} + k \begin{bmatrix} C_1 & 0 \\ 0 & C_2 \end{bmatrix} + \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix} \begin{bmatrix} U \\ W \end{bmatrix} = 0, \]  
(A37)

where matrix \( A_1 \) has terms given by (A31) matrix \( B \), has terms given by (A32), matrix \( C \) has terms given by (A33), matrix \( D \) has terms given by (A34), and the vectors \( U \) and \( W \) represent the polynomial approximation of the amplitude of the horizontal and vertical oscillations respectively. Expressions for the elements of matrices \( A_1 \), \( B \), \( C \), and \( D \) can be derived from (A36).

**APPENDIX B: NUMERICAL EXAMPLES**

To evaluate this new method of computing phase velocity from a 1-D earth model, we show comparisons between analytic results and solutions obtained using existing techniques. Arguably the most common approach used by practitioners for the calculation of phase velocity for a given earth model and frequency is the Thomson–Haskell propagator matrices approach (Thomson 1950; Haskell 1953). It is a fast, efficient and generally stable method although for Rayleigh waves there are some numerical stability issues (Aki & Richards 2002; Ke et al. 2011).

**B1 Love waves**

For Love waves, the earth model used is from Aki & Richards (2002) and consists of an homogeneous 10 km layer over a half-space with material properties shown in Table B1.

<table>
<thead>
<tr>
<th>Layer</th>
<th>( \rho ) (kg m(^{-3}))</th>
<th>( V_s ) (m s(^{-1}))</th>
<th>( V_p ) (m s(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layer</td>
<td>2800</td>
<td>3000</td>
<td>5000</td>
</tr>
<tr>
<td>Half-space</td>
<td>3200</td>
<td>5000</td>
<td>8000</td>
</tr>
</tbody>
</table>
These plots show dispersion curves computed with different methods for the model from Aki & Richards (2002). The methods are (a) the Thomson–Haskell propagator matrix method, (b) the spectral element method with fixed boundary, (c) the spectral element method with a Laguerre boundary with fixed scale, and (d) the spectral element method with automatic scaling. In each plot the computed dispersion curve is shown in solid black and the analytical solution with a red dotted line. The grey shaded region represents the frequency range generally of interest to ambient noise studies. While (b) and (c) have artefacts at low frequencies, (d) is accurate across all frequencies.

The phase velocity over a frequency range of near 0 to 1 Hertz was computed for the various methods. The comparison is performed between the results of a Thomson–Haskell calculation, a spectral element solution with a fixed boundary condition, a spectral element solution with a Laguerre boundary condition with fixed scaling, and a spectral element solution with a Laguerre boundary condition with automatic scaling. Each of these methods is compared to the known analytical solution derived in Aki & Richards (2002).

For the spectral element solution with a fixed boundary condition, the domain is extended to 100 km depth with additional elements to approximate the half-space. For the spectral element solution with a Laguerre half-space, a fixed scaling of $1.0 \times 10^{-4}$ was used and for the automatic scaling, the phase velocity is computed from high frequency (where incorrect scaling has little impact) to low and the Laguerre scaling term is updated using the wave number of the previous result. This last auto-scaling method is the approach used in the inversion examples. For all spectral element solutions, fifth-order polynomials are used except where otherwise indicated.

The results of the calculations are plotted in Fig. B1. It can be seen from the plots that in general the match between the analytical result and computed dispersion curve is visually identical with the exception of (b) and (c). In (b), as frequency approaches zero, there is divergence from the analytic result. This is because at low frequencies, the wave number becomes sensitive to the fixed boundary and causes the phase velocity to become unrealistic. In (c) there is a similar effect except that in this case it is poor scaling of the Laguerre spectral element representing the half-space. In (c), the Laguerre scale is tuned so that performance is good in the range of frequencies of interest to ambient noise studies. In (d), with the automatic updating of Laguerre scaling, the accuracy of the lower frequencies is far better.

In Fig. B2, the log$_{10}$ of the relative error compared to the analytic result for each of the calculations is shown. In each of the plots, the horizontal blue dashed line represents a one percent relative error and a reasonable threshold for accuracy. In (b) the spectral element solution with a fixed boundary can satisfactorily compute the dispersion across the range of frequencies of interest by sufficiently padding the model (down to 100 km depth in this case). Similarly in (c), a fixed Laguerre scale can ensure accuracy over a frequency range of interest. However, the most accurate method in this case is the SEM with the automatic updating of the Laguerre scaling term in (d). It additionally outperforms the Thomson–Haskell method at higher frequencies.
Figure B2 These plots show the log10 relative error of the various methods tested. The methods are the Thomson–Haskell propagator matrix method (a), the spectral element method with fixed boundary (b), the spectral element method with a Laguerre boundary with fixed scale (c) and the spectral element method with automatic scaling (d). In each plot, the horizontal red and blue dashed lines represent 5 and 1 per cent relative errors, respectively. The grey shaded region represents the frequency range generally of interest to ambient noise studies.

Table B2. The earth model parameters for the calculation of Rayleigh wave phase velocity. The earth model consists of two simple homogeneous 10 km thick layers over a half-space.

<table>
<thead>
<tr>
<th>Layer</th>
<th>(\rho) (kg m(^{-3}))</th>
<th>(V_s) (m s(^{-1}))</th>
<th>(V_p) (m s(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layer 1</td>
<td>2800</td>
<td>3000</td>
<td>5100</td>
</tr>
<tr>
<td>Layer 2</td>
<td>3000</td>
<td>4000</td>
<td>6800</td>
</tr>
<tr>
<td>Half-space</td>
<td>3200</td>
<td>5000</td>
<td>8500</td>
</tr>
</tbody>
</table>

B2 Rayleigh waves

Similar tests are performed with the Rayleigh wave solvers for a two layer model over a half-space with material properties shown in Table B2. In the case of Rayleigh waves, there is an analytic results available for an acoustic layer over a half-space, but the current formulation only supports elastic layers. Hence, the Thomson–Haskell solution is used for comparison purposes.

In Fig. B3 the results of the spectral element calculations are shown. In (a) is the spectral element solution with a fixed boundary condition, (b) with a Laguerre boundary condition and fixed scale, and (c) the Laguerre boundary condition with automatic scaling. Across the range of frequencies of interest, the match is visually good with the exception of (a) at the low end of the grey shaded region.

The relative error plot, compared to a Thomson–Haskell solution, is shown in Fig. B4. Both of the solutions using the Laguerre boundary element are well below the one percent error line across the range of frequencies highlighted. One disconcerting feature is that in this case the error increases as frequency increases. This is caused by the inability of the 5th order spectral polynomial to accurately represent higher frequency eigenfunctions in the first 10km of the model.

If higher frequencies are required, then as with all spectral element solutions, the grid can be made finer or the order of the polynomial in each element can be increased. These are known as \(h\)-refinement and \(p\)-refinement in the broader finite element research community.
Figure B3 These plots show the dispersion curve computed with different methods. The methods are the spectral element method with fixed boundary (a), the spectral element method with a Laguerre boundary with fixed scale (b) and the spectral element method with automatic scaling (c). In each plot the computed dispersion curve is shown in solid black and the Thomson–Haskell solution with a red dotted line. The grey shaded region represents the frequency range generally of interest to ambient noise studies.

(Pozrikidis 2005; Efendiev & Hou 2009; Kopriva 2009). As an example, if the order of the polynomial is increased from 5 to 10, the results shown in Figs B5(a) and (b) are obtained where the Thomson–Haskell solution is matched across the entire frequency range. The downside of increasing the order is that this increases the size of the matrices in the generalized eigenproblem and therefore the computational cost. In this particular example, the computational time resulting from a doubling of the order from 5th to 10th was approximately a four times increase.

Alternatively, using a similar approach to Schwab & Knopoff (1970), the spectral element mesh can be refined near the surface using an approximation of the wavelength as guide to provide the needed accuracy and this approach is shown in Figs B5(c) and (d). The approximate wavelength can be estimated using an initial guess or estimated from a higher frequency similar to the calculation of the optimal scaling for the Laguerre element. This is a more targeted refinement of the mesh to provide accuracy for higher frequencies and in this case only results in a 1.7 times increase in computational cost.

In Fig. B6 the comparison plots of the SEM with Laguerre automatic scaling to the analytic Love group velocity is shown in (a) with the relative error plot in (b). Fig. B6(c) shows the comparison between the Thomson–Haskell method where the agreement is good as evident in the relative error plot in (d). Lastly in B6(e) the Rayleigh wave ellipticity H/V ratio is compared between the Thomson–Haskell approach and the spectral element solution with Laguerre element automatic scaling.

B3 Pre-conditioning

For the Rayleigh wave solution, we use the recommendations of Hammarling et al. (2013) and include matrix scaling terms to pre-condition the eigenvalue computation as shown in (19). In Fig. B7 we show a comparison between the calculation of the Rayleigh wave phase velocity (a) with and (b) without pre-conditioning the matrices. As can be seen from the two plots, when the pre-conditioning scaling terms are removed, the errors become larger, particularly at higher and lower frequencies.
Figure B4 These plots show the log10 of the relative errors with respect to the Thomson–Haskell solution for the different methods. The methods are the spectral element method with fixed boundary (a), the spectral element method with a Laguerre boundary with fixed scale (b) and the spectral element method with automatic scaling (c). The grey shaded region represents the frequency range generally of interest to ambient noise studies.

B4 Eigenfunctions

A final verification is to check that the boundary conditions of the differential equations are satisfied. Recall that for Love waves, the displacement and the traction, dictated by the leading term \( L \frac{\partial u}{\partial z} \), must both be continuous functions of depth. Additionally the traction must be zero at the free surface and decay to zero at depth. In Fig. B8 the displacement is shown in (a), and the traction shown in (b), for three representative periods normalized so the maximum amplitude is one. In this figure, both displacement and traction are continuous with depth, and the traction is zero at the surface. In the figures, the red dashed line indicates the half-space boundary. The Laguerre element, which represents the eigenfunctions below this point, accurately represents the exponential decay. Careful inspection of the 10 s period curve in (a) reveals an extremely subtle oscillation which is a result of Laguerre scale mismatch. This could be remedied through increasing the order of the Laguerre function, or through further iterative refinement of the Laguerre scaling term, but this oscillation is negligible.

For Rayleigh waves, two tractions, \( \tau_{xz} \) and \( \tau_{zz} \), must be zero at the surface and decay to zero at depth. Similarly, the displacements must be continuous and decay to zero at depth. In Fig. B9 the horizontal displacement is shown in (a) and vertical displacement in (b). The traction \( \tau_{xz} \) is shown in (c) and \( \tau_{zz} \) (d). In this model, there is an extra layer compared, and the first layer boundary is represented with a black dashed line and the half-space boundary with a red dashed line. Visual inspection reveals that all boundary conditions and continuity requirements are satisfied for the Rayleigh wave tests.

In summary, the SEM presented here can accurately compute phase and group velocity and compares favourably with analytic and existing forward modelling methods. An important point is that the presented SEM is flexible in its solution accuracy and there are many strategies that can be used to trade-off computational effort and accuracy (Patera 1984; Pozrikidis 2005).

B5 Computational cost

Table B3 shows the relative time taken for the dispersion curve points computed in the accuracy test in Appendix B1. In the table, the newly developed SEM is comparable in terms of cost with the Thomson–Haskell method. The computational cost is dictated by the size
If the order of the polynomials is increased in the solution, the accuracy is improved at higher frequency at the cost of higher computational effort. In (a) is the dispersion curve computed with 10th order elements compared to the Thomson–Haskell method, and (b) shows the log₁₀ of the relative error. The time to compute is approximately 5 times that of Fig. B3 which used fifth-order elements. An alternative to simply globally increasing the order of all elements is to refine elements near the surface at higher frequency. The result of refining the elements when necessary is the improved accuracy as shown in the comparison in (c) and the relative error plot in (d). This refinement results in only a 1.7 times penalty in computational time.

of the matrix which in turn is controlled by the number of layers and the polynomial order used in the elements. Therefore, the SEM with Laguerre boundaries is marginally quicker than the fixed boundary case because extra elements need to be added to the model to approximate a half-space and this padding increases the matrix dimension of the solution.

Table B4 shows a similar comparison between Thomson–Haskell and the SEM for the calculation of Rayleigh waves. In this case there is a significant extra cost in the spectral element method, approximately a factor of six times slower. The primary cause of this is that the matrix is a factor of four times larger than an equivalent Love wave spectral element solver. In the current implementation, the matrix equation is reconstructed for every calculation, so it is possible that this could be further optimized. Nonetheless, for typical dispersion inversions involving Rayleigh waves, a forward model computation would cost less than a second on modern hardware and is therefore feasible for use in an inversion.
In (a) the analytic result for Love wave group velocity shown in red dashed line is compared to the calculation using the new spectral element method with the correspond log_{10} plot of the relative error in (b). In (c) the Thomson–Haskell approach for Rayleigh group velocity is compared and the log_{10} plot of the relative error in (d). Finally in (e) the Rayleigh wave ellipticity ratio H/V is compared between the Thomson–Haskell approach and the spectral element approach with the log_{10} plot of the relative error in (f). In the relative error plots, the horizontal red and blue dash lines represent 5 and 1 per cent relative error levels, respectively.
Figure B7 A comparison between the eigensolutions for Rayleigh waves using pre-conditioning in (a) and without in (b). The error level is generally higher and less consistent, that is, greater variation, when the pre-conditioning is not used.

Figure B8 The normalized eigenfunctions for Love waves of selected periods for the synthetic model. (a) is the horizontal displacement, (b) is $J_{\ell} \frac{dl}{H}$. 
Figure B9. The normalized eigenfunctions for Rayleigh waves of selected periods for the synthetic model. (a) is the horizontal displacement, (b) is the vertical displacement, (c) is $L \left( \frac{dU}{dz} - kW \right)$ and (d) is $FkU + C \frac{dU}{dz}$.

Table B3. A comparison of the computational time for the methods of computing Love wave dispersion curves. The times quoted represent the time to evaluate 512 different frequencies and as can be seen all non-analytic methods are comparable in terms of computational cost.

<table>
<thead>
<tr>
<th>Method</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytic</td>
<td>0.010</td>
</tr>
<tr>
<td>Thomson–Haskell</td>
<td>0.150</td>
</tr>
<tr>
<td>Spectral element (fixed)</td>
<td>0.200</td>
</tr>
<tr>
<td>Spectral element (Laguerre)</td>
<td>0.140</td>
</tr>
<tr>
<td>Spectral element (auto-Laguerre)</td>
<td>0.160</td>
</tr>
</tbody>
</table>
In order to solve (C6) a further constraint is required and this can be remedied by recognizing that (C1) has an infinite number of model parameters, and the state space vector can be written

\[ x = \begin{bmatrix} \phi_r \\ \phi_r' \\ \alpha \end{bmatrix}. \]  

(C2)

Given a misfit function \( L(m) \), the derivative can be calculated using the chain rule to give

\[ \frac{dL}{dm} = \frac{dL}{dx} \frac{dx}{dm}. \]  

(C3)

The last term is found by differentiating (C1)

\[ \frac{df}{dx} + \frac{df}{dm} = 0, \]  

(C4)

substitution into (C3) allows the gradient of the likelihood with respect to model parameters \( m \) to be written

\[ \frac{dL}{dp} = \frac{dL}{dx} \left( \frac{df}{dx} \right)^{-1} \frac{df}{dm}. \]  

(C5)

A problem with computing (C5) directly is that, in the case of eigenproblems, \( \frac{df}{dm} \) is a rank 3 tensor terms of dimensions \( M \) by \( N + 1 \) by \( N + 1 \), whereas \( \frac{df}{dx} \) is an \( N + 1 \) by \( N + 1 \) matrix, where \( M \) is the number of parameters in vector \( m \) and \( N \) is the size of the eigenproblem or length of the eigenvector \( \phi_r \). The computational cost of the tensor/matrix multiplication increases rapidly as \( N \) and \( M \) increases.

To reduce this cost, first, once a solution is found to \( f \), rather than computing the rank 3 tensor terms, we compute the matrices \( \frac{df}{dx} \) and \( \frac{df}{dm} \). This is more efficient due to the fact that matrices \( A' \) and \( B' \) are sparse. In this formulation, \( B' \) is diagonal for both Love and Rayleigh waves, and \( A' \) is block diagonal for Love waves and composed of block diagonal matrices for Rayleigh waves.

Second, the adjoint technique combines the first two terms of (C5) by letting \( \lambda \) be the vector of adjoint state variables that is solved for using

\[ \frac{df^T}{dx} \lambda = \frac{dL^T}{dx}. \]  

(C6)

then the misfit function gradient can be computed with

\[ \frac{dL}{dm} = -\lambda \frac{df}{dm}. \]  

(C7)

The state space vector defined in (C2) consists of the right eigenvector and the eigenvalue or \( N + 1 \) unknowns whereas in (C1), there are \( N \) equations. In order to solve (C6) a further constraint is required and this can be remedied by recognizing that (C1) has an infinite number of solutions since for any given eigenvalue, a multiple of the eigenvector will also be a solution. A simple addition is to add a constraint to the norm of the eigenvector so the augmented \( f' \) becomes

\[ f' = \begin{bmatrix} A' \phi_r - aB' \phi_r \\ \phi_r^T \phi_r - 1 \end{bmatrix}. \]  

(C8)

This then results in \( N + 1 \) equations, with \( N \) being the size of the square matrices \( A' \) and \( B' \).
A solution then is required for the system
\[
\frac{d \mathbf{r}^T}{dt} \left[ \lambda_{1,N} \mathbf{r} \right] = \frac{d \mathcal{L}}{dt}
\]
and substitution of the general eigensystem (C8) results in the two equations
\[
(\mathbf{A}^T - \alpha \mathbf{B}^T)\lambda_{1,N} = \frac{d \mathcal{L}}{d \phi_r} - 2 \phi_l \lambda_{N+1}
\]
(C10)
\[
-(\mathbf{B} \phi_r)^T \lambda_{1,N} = \frac{d \mathcal{L}}{d \alpha}.
\]
(C11)

The term \( \lambda_{N+1} \) can be solved for by pre-multiplying (C10) by \( \phi_r^T \) which reduces the left-hand side to zero giving
\[
\lambda_{N+1} = \frac{1}{2} \phi_r \frac{d \mathcal{L}}{d \phi_r}.
\]
(C12)
and (C10) becomes
\[
(\mathbf{A}^T - \alpha \mathbf{B}^T) \lambda_{1,N} = (1 - \phi_r \phi_r^T) \frac{d \mathcal{L}}{d \phi_r}.
\]
(C13)

The key to solving eqs (C13) and (C11) for \( \lambda_{1,N} \) is to recognize that the matrix in the linear system in (C13) has a null-space of \( \phi_r \), that is, the left eigenvector. Hence for any valid solution, \( \lambda'_{1,N} \) to (C13), \( \lambda_{1,N} + \beta \phi_l \) is also a solution.

Once a solution to (C13) is obtained, \( \beta \) can be solved for using (C11) where
\[
-(\mathbf{B} \phi_r)^T (\lambda'_{1,N} + \beta \phi_l) = \frac{d \mathcal{L}}{d \alpha},
\]
(C14)
giving
\[
\beta = -\frac{\frac{d \mathcal{L}}{d \phi_r} + \phi_r^T \mathbf{B} \lambda'_{1,N}}{\phi_l \phi_r^T}.
\]
(C15)

Finally, this can be substituted into (C5) to obtain
\[
\frac{d \mathcal{L}}{d \mathbf{m}} = -(\lambda'_{1,N} + \beta \phi_l)^T \left( \frac{d \mathbf{A}}{d \mathbf{m}} - \alpha \frac{d \mathbf{B}}{d \mathbf{m}} + \phi_l \right) \phi_r.
\]
(C16)

In summary, the steps for computing the gradient of a function of the general eigensystem are as follows:

(i) Solve the general eigensystem for \( \alpha, \phi_l, \phi_r \) using a criteria for choosing the eigenvalue, for example, the largest Eigenvalue corresponding to the fundamental mode.
(ii) Compute \( \frac{d \mathbf{A}}{d \mathbf{m}} \phi_r \) and \( \left[ \frac{d \mathbf{r}}{d \mathbf{m}} \right] \phi_r \).
(iii) Compute \( \frac{d \mathcal{L}}{d \phi_r} \) and \( \frac{d \mathcal{L}}{d \alpha} \).
(iv) Solve (C13) for \( \lambda'_{1,N} \) and ensure the solution is orthogonal to \( \phi_r \).
(v) Compute \( \beta \) using (C15).
(vi) Compute the gradient using (C16).

This is a general solution for eigenvalue problems of the form encountered for Love and Rayleigh waves in the previous section. For problems where the misfit function is a function of phase or group velocity or Rayleigh wave ellipticity, expressions for \( \frac{d \mathcal{L}}{d \alpha} \) and \( \frac{d \mathcal{L}}{d \phi_r} \) are required and these are outlined in the following sections.

### C1 Love and Rayleigh phase velocity

In the case where the gradient of a function of the phase velocity is being sought, which is a function only of \( \alpha \), we have \( \frac{d \mathcal{L}}{d \phi_r} = 0 \) and so
\[
\frac{d \mathcal{L}}{d \mathbf{m}} = \frac{d \mathcal{L}}{d \alpha} \left( \phi_r \left( \frac{d \mathbf{A}}{d \mathbf{m}} - \alpha \frac{d \mathbf{B}}{d \mathbf{m}} \right) \phi_r \right).
\]
(C17)

For Love waves, in the general form, we have \( k^2 = \alpha \) and given a simple function of \( \mathcal{L} = c_{Love} \), using the chain rule
\[
\frac{d c_{Love}}{d \alpha} = \frac{-\omega}{2k^3},
\]
(C18)
and
\[
\frac{d c_{Love}}{d \mathbf{m}} = \frac{-\omega}{2k^3 \phi_r \phi_r^T} \left( \phi_r \left( \frac{d \mathbf{A}}{d \mathbf{m}} - k^2 \frac{d \mathbf{B}}{d \mathbf{m}} \right) \phi_r \right).
\]
(C19)
In the Love wave case, the matrices are symmetric and hence $\phi_i = \phi_i$. If we assume that the model parameters are the Lamé parameter $\mu$ and $\rho$ we can restate the above equation in terms of the integrals the spectral solution approximates

$$
\frac{dc_{\text{Love}}}{dm} = \frac{-\omega}{2k^3} \int_0^\infty \mu(z) \phi_i(z)^2 \, dz \left[ \int_0^\infty \omega^2 \phi_i(z)^2 \delta \mu \, dz - k^2 \int_0^\infty \phi_i(z)^2 \delta \mu \, dz \right].
$$

(C20)

This shows a direct correspondence between the result using a variational principle (eq. 7.71, Aki & Richards 2002) and the adjoint method with spectral elements which is a good validation of the method.

For Rayleigh waves, in the general form we have $k = \gamma \alpha$ where $\gamma$ is the scaling parameter to improve the accuracy of the quadratic eigenproblem and given a simple function of $L = c_{\text{Rayleigh}}$, using the chain rule

$$
\frac{dc_{\text{Rayleigh}}}{d\alpha} = -\frac{\gamma \omega}{k^2}.
$$

(C21)

and

$$
\frac{dc_{\text{Rayleigh}}}{dm} = -\frac{\gamma \omega}{k^2} \phi_i^T B \phi_i \left[ \phi_i^T (dA' \frac{d\alpha}{dm} + m dB - \alpha dB') + \phi_i^T \frac{d\alpha}{dm} \right].
$$

(C22)

Unlike the Love wave case, comparing the Rayleigh wave result to published variational methods is difficult owing to the companion form that the matrices are rearranged into.

A key result here is that the solution to the adjoint is simple and the gradient can be computed directly from the full solution. This means for phase velocity inverse problems, the gradient of the misfit with respect to model parameters can be efficiently computed with a handful of dot products because $B'$ is a diagonal matrix.

This leads to the possibility that rather than use the adjoint method, we could simply compute the model parameter derivatives with respect to phase directly, sometimes called Frechet derivatives. This is a possibility for phase velocity but doesn’t hold for group velocity or Rayleigh wave ellipticity. We stress that the adjoint method is more general and allows for different misfit functions and is likely more efficient as the complexity of the misfit function increases.

### C2 Love group velocity

For Love wave group velocity, (24) was derived and has dependence on both the eigenvalue, right eigenvector, and matrices $A'$ and $B'$. Differentiation of (24) gives

$$
\frac{d\nu_{g, \text{Love}}}{dm} = \frac{d\nu_{g, \text{Love}}}{d\alpha} \frac{d\alpha}{dm} + \frac{d\nu_{g, \text{Love}}}{dV} \frac{dV}{dm} + \frac{d\nu_{g, \text{Love}}}{dA'} \frac{dA'}{dm} + \frac{d\nu_{g, \text{Love}}}{dB'} \frac{dB'}{dm}.
$$

(C23)

Since the last two terms can be computed directly from the model parameters, the remaining derivation focuses on the solution of the first two terms. The first of these upon rearrangement is

$$
\frac{d\nu_{g, \text{Love}}}{d\alpha} = \frac{B}{2k\omega B},
$$

(C24)

and the second upon rearrangement becomes

$$
\frac{d\nu_{g, \text{Love}}}{d\phi_i} = \frac{2k(\hat{A}B\phi_i - \hat{B}A\phi_i)}{\omega A^2}.
$$

(C25)

So for the gradient of Love group velocity with respect to model parameters, the last two terms of (C23) are computed directly and combined with the adjoint method outlined above using (C24) and (C25).

### C3 Rayleigh Group Velocity

The group velocity for Rayleigh waves was given by (25) was derived and differentiating this we can obtain

$$
\frac{d\nu_{g, \text{Rayleigh}}}{dm} = \frac{d\nu_{g, \text{Rayleigh}}}{d\alpha} \frac{d\alpha}{dm} + \frac{d\nu_{g, \text{Rayleigh}}}{d\phi_i} \frac{d\phi_i}{dm} + \frac{d\nu_{g, \text{Rayleigh}}}{dA} \frac{dA}{dm} + \frac{d\nu_{g, \text{Rayleigh}}}{dB} \frac{dB}{dm} + \frac{d\nu_{g, \text{Rayleigh}}}{dC} \frac{dC}{dm}.
$$

(C26)

Similar to the Love wave case, the last three terms can be directly computed and the first two terms computed with the adjoint method outlined above with

$$
\frac{d\nu_{g, \text{Rayleigh}}}{d\alpha} = \frac{\gamma B}{\omega A}.
$$

(C27)

and

$$
\frac{d\nu_{g, \text{Rayleigh}}}{d\phi_i} = \frac{(\hat{A}C - 2\hat{C}A - 4k\hat{B}A + 4k\hat{A}B) \phi_i}{2\omega A^2}.
$$

(C28)
C4 Rayleigh H/V adjoint

For Rayleigh ellipticity, in contrast to the phase velocity solution, a gradient of a functional that depends purely on the right eigenvector is sought. The components of the right eigenvector can be written

\[ \phi_r = \begin{bmatrix} U_1 & U_2 & \ldots & U_N & W_1 & W_2 & \ldots & W_N & \mathbf{P}_x & \mathbf{P}_y \end{bmatrix}^T, \tag{C29} \]

where \( U_r \) is the polynomial approximation of the horizontal oscillations at interpolation nodes and \( W_r \) is for the vertical oscillations. Hence the Rayleigh wave ellipticity at the surface is given by

\[ \frac{H}{V} = \frac{U_1}{W_1}. \tag{C30} \]

and, as a simple example where \( \mathcal{L} = \frac{U_1}{W_1} \), \( \frac{\partial \mathcal{L}}{\partial \phi} \) can be specified using

\[ \frac{\partial \mathcal{L}}{\partial \phi} = \begin{bmatrix} \frac{1}{W_1} & 0 & \ldots & 0 & \frac{1}{W_1} & 0 & \ldots \end{bmatrix}^T, \tag{C31} \]

and used in adjoint algorithm described earlier with \( \frac{\partial \mathcal{L}}{\partial \phi} \) set to zero. In this case, there is not a convenient algebraic solution for the gradient of the H/V ratio with respect to model parameters.

In previous perturbational techniques for Rayleigh wave ellipticity (Tanimoto & Alvizuri 2006), the gradient was computed in terms of the orthogonal basis formed from higher order mode eigenfunctions. It was found that this provided good convergence and a numerical scheme was used. In contrast, the adjoint approach only requires the eigenfunction at the mode of interest.

If we examine (C13) and (C16), the solution uses a projection of (C31) orthogonal to the right eigenvector. An explicit orthogonal projection is optimal in this case compared to the orthogonal basis formed from higher order mode eigenfunctions. Unlike the perturbational technique, we have found that the adjoint calculation of the gradient with respect to misfit functions to be accurate when compared to numerical approximations. The accuracy and efficiency of this approach for computing Rayleigh wave ellipticity gradients with respect to earth model parameters could be an important step forward for geophysical inference.

An additional note with Rayleigh wave ellipticity is that the sign of the horizontal and vertical components cannot be distinguished so typically the absolute value of the H/V ratio is used. In our implementation we check the sign of the H/V ratio before taking its absolute value and manually correct the gradient.

C5 Jacobian

As a useful comparison to other approaches, we also compute the gradient of the Rayleigh wave phase velocity with respect to layer shear wave velocity to a test model, MODX, proposed by Xia et al. (1999) and compare gradient values. The results are shown in Table C1.

C6 Implementation details

C7 Matrix derivatives

A key component of the algorithm is the projection of some parametrization of the earth model onto a spectral element mesh expressed in terms of the Love transverse isotropy parameters (Love 1927), \( \rho, A, C, F, L \) and \( N \). In the discussion on the numerical accuracy of the method, two strategies were introduced for improving accuracy, namely increasing the polynomial order in the elements, and subdivision of elements within a wavelength of the surface. This introduces a complexity in the computation of the gradient.

Fortunately, projection of lower order parameterizations to higher order polynomials is exact (to machine accuracy) as previously shown in Figs 2 and 4. The transformation to higher order is a linear operation and hence the Jacobian is straightforward to compute. Similarly for subdivision of elements, representation is exact if discontinuities in the earth model are preserved. In Fig. C1 this is demonstrated where the first layer of the model in (a) is subdivided near its centre and projected to a spectral element mesh in (b) while preserving the original function.

There are many different parametrizations that could be used to represent earth models, examples range from parametrizations using only an isotropic shear wave velocity varying with depth and computing \( p \)-wave velocity and density from empirical relationships, to a complete specification of anisotropic shear wave velocities, anisotropic \( p \)-wave velocities and densities with depth. This choice of parametrization will be problem dependent and in an effort to maintain some flexibility, this has been implemented using C++ templates that compute appropriate derivatives with respect to the Love parameters.

Finally, the generic matrix derivatives \( \frac{\partial \mathcal{L}}{\partial \mathbf{m}} \) are sparse rank three tensors. To save computational effort and computer memory, these are computed after a solution eigenvalue and right eigenvector are found and the terms \( \frac{\partial \mathcal{L}}{\partial \mathbf{m}} \) and \( \frac{\partial \mathcal{L}}{\partial \mathbf{m}} \) are computed directly. This is time and memory efficient as matrices \( \mathbf{A} \) are block diagonal or composed of block diagonal matrices and \( \mathbf{B} \) is diagonal.

C8 Computing the Inverse

As is common in adjoint solutions for the gradient, typically an additional computation is required of approximately the same computational complexity. For example, in classical full waveform adjoint tomography, the wave equations are solved both forwards and backwards in time in order to compute the gradient (Fichtner 2011). A similar cost exists here in the calculation of the gradient where the right eigenvector is
Table C1. The table presents a comparison of the Jacobian computed for the MOD-X model from Xia et al. (1999) to other existing techniques. In the last two columns, we show the result using 5th and 10th order polynomials using the spectral element formulation developed in this manuscript. Across the model parameters and range of frequencies, there is strong agreement between the differing methods.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5 Hz</td>
<td>1</td>
<td>0.018</td>
<td>0.01809</td>
<td>0.01809</td>
<td>0.018089331</td>
<td>0.018090823</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.018</td>
<td>0.01834</td>
<td>0.01834</td>
<td>0.018339149</td>
<td>0.018340717</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.022</td>
<td>0.02219</td>
<td>0.02219</td>
<td>0.022190540</td>
<td>0.022192493</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.021</td>
<td>0.02036</td>
<td>0.02036</td>
<td>0.020360282</td>
<td>0.020362191</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.017</td>
<td>0.01750</td>
<td>0.01750</td>
<td>0.017497181</td>
<td>0.017498979</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.872</td>
<td>0.87242</td>
<td>0.87242</td>
<td>0.872594173</td>
<td>0.87249910</td>
</tr>
<tr>
<td>10 Hz</td>
<td>1</td>
<td>0.130</td>
<td>0.13002</td>
<td>0.13002</td>
<td>0.130010451</td>
<td>0.130019665</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.106</td>
<td>0.10646</td>
<td>0.10646</td>
<td>0.106450266</td>
<td>0.106457798</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.062</td>
<td>0.06174</td>
<td>0.06174</td>
<td>0.061737902</td>
<td>0.061742870</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.025</td>
<td>0.02467</td>
<td>0.02467</td>
<td>0.024666697</td>
<td>0.024669556</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.022</td>
<td>0.02225</td>
<td>0.02225</td>
<td>0.022251230</td>
<td>0.022253742</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.766</td>
<td>0.76580</td>
<td>0.76580</td>
<td>0.765779892</td>
<td>0.765801069</td>
</tr>
<tr>
<td>15 Hz</td>
<td>1</td>
<td>1.067</td>
<td>1.06737</td>
<td>1.06766</td>
<td>1.067645911</td>
<td>1.067662896</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.925</td>
<td>0.92456</td>
<td>0.92490</td>
<td>0.924894950</td>
<td>0.924900414</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.313</td>
<td>0.31290</td>
<td>0.31304</td>
<td>0.313042845</td>
<td>0.313040967</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.034</td>
<td>0.03356</td>
<td>0.03359</td>
<td>0.033588727</td>
<td>0.033587845</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.017</td>
<td>0.01667</td>
<td>0.01665</td>
<td>0.016649822</td>
<td>0.016652700</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.262</td>
<td>0.26222</td>
<td>0.26204</td>
<td>0.262031385</td>
<td>0.262040224</td>
</tr>
<tr>
<td>20 Hz</td>
<td>1</td>
<td>0.155</td>
<td>0.15460</td>
<td>0.15460</td>
<td>0.154605014</td>
<td>0.154600341</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.037</td>
<td>1.03665</td>
<td>1.03665</td>
<td>1.036647585</td>
<td>1.036647263</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.967</td>
<td>0.96729</td>
<td>0.96729</td>
<td>0.967294198</td>
<td>0.967294246</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.457</td>
<td>0.45739</td>
<td>0.45739</td>
<td>0.457393419</td>
<td>0.457393578</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.145</td>
<td>0.14507</td>
<td>0.14507</td>
<td>0.145072741</td>
<td>0.145072903</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.040</td>
<td>0.04024</td>
<td>0.04024</td>
<td>0.040237181</td>
<td>0.040237283</td>
</tr>
<tr>
<td>25 Hz</td>
<td>1</td>
<td>0.293</td>
<td>0.29284</td>
<td>0.29284</td>
<td>0.292837905</td>
<td>0.292837972</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.072</td>
<td>1.07203</td>
<td>1.07203</td>
<td>1.072039907</td>
<td>1.072039892</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.517</td>
<td>0.51690</td>
<td>0.51690</td>
<td>0.516896612</td>
<td>0.516896600</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.102</td>
<td>0.10261</td>
<td>0.10261</td>
<td>0.102607201</td>
<td>0.102607198</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.012</td>
<td>0.01137</td>
<td>0.01137</td>
<td>0.011372179</td>
<td>0.011372178</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.001</td>
<td>0.00074</td>
<td>0.00074</td>
<td>0.000742087</td>
<td>0.000742087</td>
</tr>
<tr>
<td>30 Hz</td>
<td>1</td>
<td>0.520</td>
<td>0.52038</td>
<td>0.52024</td>
<td>0.520240911</td>
<td>0.520241004</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.923</td>
<td>0.92342</td>
<td>0.92354</td>
<td>0.923543352</td>
<td>0.923543232</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.202</td>
<td>0.20184</td>
<td>0.20196</td>
<td>0.201964381</td>
<td>0.201964373</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.016</td>
<td>0.01590</td>
<td>0.01592</td>
<td>0.015916957</td>
<td>0.015916954</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.000</td>
<td>0.00069</td>
<td>0.00060</td>
<td>0.000599531</td>
<td>0.000599530</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.000</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.000010650</td>
<td>0.000010650</td>
</tr>
</tbody>
</table>

involved, that is, for group velocity or Rayleigh wave ellipticity problems. First, the general eigenequation

\[ A' \phi = \alpha B' \phi, \tag{C32} \]

needs to be solved, followed by a solution to the linear system

\[ \left( A'^T - \alpha B'^T \right) \lambda = \left( 1 - \phi \phi^T \right) \frac{dC}{d\phi}. \tag{C33} \]

From this observation it would be expected that the computational cost of computing the gradient when the right eigenvector is involved would be approximately a factor of 2 more expensive.

This can be significantly ameliorated by examining the algorithm for the calculation of the solution to (C32) (Moler & Stewart 1973). In summary, the algorithm scales and permutes the matrices to condition and isolate “easy” eigenvalues. This step is unnecessary for the matrices constructed using the SEM in Section 2 as the matrices are already conditioned and are in a combination of diagonal and block diagonal form. The second step reduces \( A' \) and \( B' \) to upper Hessenberg form, a third step further reduces these matrices to a quasi-upper triangular matrix \( S \) and upper triangular matrix \( P \) so that

\[ A' = QSZ^T, \tag{C34} \]

and

\[ B' = QPZ^T. \tag{C35} \]
Figure C1  An arbitrary earth model consisting of a mixture of smoothly varying features over a half-space in (a) can be accurately represented in (b) with a spectral element mesh. In this example, the first layer in (a) is refined by subdivision into two elements in (b). The representation is still exact to machine precision allowing projection of models and accurate calculation of the Jacobians. Spectral element nodes are indicated with dots, Gauss–Lobatto–Legendre and Gauss–Laguerre–Legendre cardinal polynomials with faint lines, and the resulting interpolated earth model with heavy line. In this case the representation is exact (to machine precision). Similarly, in (c) the solution for the polynomial approximation of the amplitude function is well represented by the two interpolation polynomials.

where \( Q \) and \( Z \) are orthogonal matrices. Matrix \( S \) is upper triangular except for the case where there is a complex eigenvalue, in which case the matrix will have a single non-zero in the lower triangle for each complex eigenvalue pair. For example, in a simple 4x4 system, if the decomposition of \( A' \) was of the form

\[
S = \begin{bmatrix}
  s_{11} & s_{12} & s_{13} & s_{14} \\
  0 & s_{22} & s_{23} & s_{24} \\
  0 & s_{32} & s_{33} & s_{34} \\
  0 & 0 & 0 & s_{44}
\end{bmatrix}
\]

(C36)

Then the system has a two real eigenvalues partially determined by \( s_{11} \) and \( s_{44} \) and two complex eigenvalues partially determined by the \( 2 \times 2 \) system \( s_{22}, s_{23}, s_{24} \), and \( s_{33} \) (the corresponding terms in the matrix \( P \) are used in conjunction with \( S \) to fully determine the eigenvalues). Lastly, eigenvectors are computed from the eigenvalues and Schur decomposition of \( A' \) and \( B' \). The available routine for solving the generalized eigenproblem has no facility to return the orthogonal matrices \( Q \) and \( Z \) or to skip unnecessary scaling and perturbation tests.

Rather than solve (C33) directly using a common linear algebra routine, the approach taken here is to create a modified version of the standard LAPACK routine that skips unnecessary steps and returns the matrices \( Q \) and \( Z \) to efficiently solve using forward substitution owing to the pseudo-triangular nature of \( S^T - \alpha P^T \). For the forward substitution to be accurate, the general matrices \( A' \) and \( B' \) need to be well conditioned with respect to each other. The method outlined by Hammarling et al. (2013) achieves this for the Rayleigh wave case and a simplified scaling based on Ward (1981) is sufficient for the Love wave case.

Therefore, only a single decomposition is required for the solution of the eigenproblem and any gradient with respect to the eigenvalue or right eigenvector. It should be stressed that in the case of computing the gradient with respect to multiple modes, the decomposition still only needs to be performed once to compute the gradient with respect to all modes of interest, a potentially large computational savings in multi-mode studies.