Notes on the implementation of spectral element method (SEM) for numerical modeling and full waveform inversion

Pengliang Yang
Email: ypl.2100@gmail.com
March 1st, 2018

Abstract
This tutorial is dedicated to spectral element method (SEM) for (visco-)elastic seismic wave equation based on Komatitsch and Tromp (1999, 2002) and Fichtner (2011), Igel (2017). The software libc_SEM using message passing interface (MPI) and C programming language gives the practical implementation of SEM for numerical modeling and full waveform inversion in 2D.

Contents

1 Introduction

2 SEM in 1D [Igel, 2017]
   2.1 SEM/FEM for 1D elastic wave equation ........................................... 2
   2.2 Dividing computing domain into elements ........................................... 2
   2.3 From physical domain to reference space ........................................... 3
   2.4 Basis function using Lagrange polynomial ........................................... 3
   2.5 Integration point & Gauss-Lobatto-Legendre quadrature ....................... 4
   2.6 Global assembly ............................................................................... 5
   2.7 Time marching using Newmark scheme ............................................ 6

3 SEM in 2D and 3D
   3.1 Preliminary: Integral expansion by tensor product ................................ 6
   3.2 SEM for anisotropic elastic wave equation [Fichtner, 2011] .................. 7
   3.3 2D SEM for P-SV system ..................................................................... 8
   3.4 Attenuation ...................................................................................... 9
   3.5 CFL stability condition - Δt [Komatitsch et al., 2005] ....................... 10
   3.6 Resolution - Δx [Komatitsch, 1997] ............................................... 10

4 Full waveform inversion
   4.1 Wavelet estimation ......................................................................... 10
   4.2 FWI gradient with respect to parameters [Kamath and Tsvankin, 2016] .... 11

5 How to run

A Radiation pattern analysis [Kamath and Tsvankin, 2016]

1 Introduction
Spectral element method (SEM) is a specific form of finite element method (FEM), where the mass matrix can be trivially inverted thanks to its diagonal matrix structure, achieved by using a specific set of basis functions inside the elements—Lagrange polynomials—combined with an interpolation scheme based upon the Gauss–Lobatto–Legendre (GLL) collocation points. The use of Lagrange polynomial specified by GLL points leads to exact interpolation on collocation points with spectral convergence properties. The scheme can be explicitly and efficiently extrapolated just like finite-difference or pseudospectral implementations, without the need to solve a large linear system of equations.

1The materials extracted from different publications are slightly adapted for the consistency.
2 SEM in 1D (Igel, 2017)

2.1 SEM/FEM for 1D elastic wave equation

Let us start from 1D elastic wave equation:

\[ \rho \partial_t^2 u = \partial_x (\mu \partial_x u) + f \]  

(1)

where the displacement \( u \), external force \( f \), mass density \( \rho \), and shear modulus \( \mu \) depend on \( x \) and \( t \). The stress-free boundary condition (also referred to as free surface boundary condition) implies the vanishing of the traction perpendicular to the free surface \( \sigma_{ij}n_j = 0 (\sigma \cdot n = T = 0) \) with \( n_j \) being the normal vector. Due to stress-strain relation \( (\sigma_{ij} = \epsilon_{ijk}\epsilon_{kl}) \), we have

\[ \mu \partial_x u(x, t)|_{x=0,L} = 0 \]  

(2)

Multiplying a test function \( v(x) \) on the both sides of the wave equation and then integrate over computing domain \( \Omega = [0, L] \):

\[ \int_{\Omega} v \rho \partial_t^2 u \, dx - \int_{\Omega} v \partial_x (\mu \partial_x u) \, dx = \int_{\Omega} v f \, dx \]  

(3)

After integration by parts, it becomes

\[ \int_{\Omega} v \rho \partial_t^2 u \, dx + \int_{\Omega} \mu \partial_x v \partial_x u \, dx = \int_{\Omega} v f \, dx \]  

(4)

allowing for the boundary condition \( \partial_x u(x, t)|_{x=0} = \partial_x u(x, t)|_{x=L} = 0 \).

We approximate displacement field by a finite superposition of \( N+1 \) basis functions \( \phi_i(x) \) with \( i = 0, ..., N \) weighted by time-dependent coefficients \( u_i(t) \)

\[ u(x, t) \approx \bar{u}(x, t) = \sum_{i=0}^{N} u_i(t) \phi_i(x) \]  

(5)

In addition, we make another important step by using as test functions the same functions that are used to approximate our unknown fields (Galerkin principle, \( v(x) = \phi_j(x) \)), yielding

\[ \int_{\Omega} \phi_j \rho \partial_t^2 \bar{u} \, dx + \int_{\Omega} \mu \partial_x \phi_j \partial_x \bar{u} \, dx = \int_{\Omega} \phi_j f \, dx. \]  

(6)

That is,

\[ \sum_{i=0}^{N} \partial_t^2 u_i \int_{\Omega} \rho \phi_j \phi_i \, dx + \sum_{i=0}^{N} u_i \int_{\Omega} \mu \partial_x \phi_j \partial_x \phi_i \, dx = \int_{\Omega} \phi_j f \, dx. \]  

(7)

which reads in matrix

\[ M \partial_t^2 \mathbf{u}(t) + \mathbf{K} \mathbf{u}(t) = \mathbf{f}(t) \]  

(8)

with the matrix matrix \( \mathbf{M} \), stiffness matrix \( \mathbf{K} \) and the source vector \( \mathbf{f} \) being specified by

\[ M_{ji} = \int_{\Omega} \rho \phi_j \phi_i \, dx; \quad K_{ji} = \int_{\Omega} \mu \partial_x \phi_j \partial_x \phi_i \, dx; \quad f_j = \int_{\Omega} \phi_j f(x, t) \, dx \]  

(9)

2.2 Dividing computing domain into elements

We divide the domain \( \Omega \) into \( n_e \) subdomains \( \Omega_e \), i.e., \( \Omega = \bigcup_{e=1}^{n_e} \Omega_e \):

\[ \sum_{i=0}^{N} \partial_t^2 u_i \sum_{e=1}^{n_e} \int_{\Omega_e} \rho \phi_j \phi_i \, dx + \sum_{i=0}^{N} u_i \sum_{e=1}^{n_e} \int_{\Omega_e} \mu \partial_x \phi_j \partial_x \phi_i \, dx = \sum_{e=1}^{n_e} \int_{\Omega_e} \phi_j f \, dx. \]  

(10)

Instead of defining basis functions in \( \Omega \) we now restrict them to reside inside the elements \( \Omega_e \):

\[ \bar{u}(x, t)|_{x \in \Omega_e} = \sum_{i=0}^{N} u_i^e(t) \phi_i^e(x) \]  

(11)
Further, the derivative of
\[ \sum_{i=0}^{N} \frac{\partial^2}{\partial x^2} u_i \int_{\Omega_e} \sum_{i=0}^{N} \mu \partial_{x} \phi_i \partial_{x} \phi_i \, dx = \int_{\Omega_e} \phi_i f \, dx. \]  

with the matrix representation
\[ M^e \frac{\partial^2}{\partial x^2} u^e(t) + K^e u^e(t) = f^e(t) \]  

where \( u^e, K^e, M^e, \) and \( f^e \) are the coefficients of the unknown displacement inside the element, stiffness and mass matrices with information on the density and elastic parameters, and the forces, respectively. The sizes of the elemental vectors and matrices in this system are: \( u^e \in \mathbb{R}^{N+1}, K^e \in \mathbb{R}^{(N+1) \times (N+1)}, M^e \in \mathbb{R}^{(N+1) \times (N+1)}, f^e \in \mathbb{R}^{N+1}. \)

### 2.3 From physical domain to reference space

Thanks to the chain rule, for any function \( h(x) \) the integration can be computed by moving from the physical coordinate \( x \) to a reference coordinate \( \xi \) within the reference interval \([-1, 1] \):
\[ \int_{\Omega_e} h(x) \, dx = \int_{-1}^{1} h(x(\xi)) J_e \, d\xi, \]  

where \( x \) and \( \xi \) can be switched from one to the other
\[ x := x(\xi) = h_e \frac{\xi + 1}{2} + x_e, \quad \xi := \xi(x) = 2 \frac{x - x_e}{h} - 1. \]  

The Jacobian \( J_e \) and its inverse \( J_e^{-1} \) can then be easily computed, that is,
\[ J_e = \frac{dx}{d\xi} = \frac{h_e}{2}, J_e^{-1} = \frac{d\xi}{dx} = 2 \frac{h_e}{x}. \]  

We then have
\[ M^e_{ji} = \int_{-1}^{1} \rho(x(\xi)) \phi_j(x(\xi)) \phi_i(x(\xi)) \frac{dx}{d\xi} \, d\xi, \]  

\[ K^e_{ji} = \int_{-1}^{1} \mu(x(\xi)) \partial_{\xi} \phi_j(x(\xi)) \partial_{\xi} \phi_i(x(\xi)) \left( \frac{d\xi}{dx} \right)^2 \frac{dx}{d\xi} \, d\xi, \]  

\[ f^e_{j} = \int_{-1}^{1} \phi_j(x(\xi)) f(x(\xi), t) \frac{dx}{d\xi} \, d\xi. \]  

### 2.4 Basis function using Lagrange polynomial

We choose the Lagrange polynomial as the basis function \( \phi_i \):
\[ \phi_j(\xi) = \prod_{i=0, j \neq i}^{N} \frac{\xi - \xi_i}{\xi_j - \xi_i}, \quad i, j = 0, \ldots, N \]  

where \( \xi_i \) are arbitrary fixed points within \([-1, 1] \), while \( N \) indicates the order of the Lagrange polynomial. Note that \( \phi_j(\xi_i) = 0 (j \neq i) \) and \( \phi_j(\xi_j) = 1 \), the Lagrange polynomial has the cardinal interpolation property
\[ \phi_j(x_i) = \delta_{ij}. \]  

Further, the derivative of \( j^{th} \) Lagrange polynomial at \( i^{th} \) point \( \phi_j(\xi_i) \) can be computed
\[ \phi_j(\xi_i)(i \neq j) = \frac{1}{\xi_j - \xi_i} \sum_{k=0, k \neq i}^{N} \frac{\xi_i - \xi_k}{\xi_j - \xi_k} \phi_j(\xi_j) = \sum_{k=0}^{N} \frac{1}{\xi_j - \xi_k} \]  

\[ (20) \]
2.5 Integration point & Gauss-Lobatto-Legendre quadrature

The polynomials $p_n(x)$ (of degree $n$) and $p_m(x)$ (of degree $m$) are said to be orthogonal when their mutual projections satisfies

$$\int_a^b w(x)p_n(x)p_m(x)dx = A_n\delta_{mn}, \quad x \in [a, b]$$  \hfill (21)

where $w(x)$ denotes a positive weighting function, and $A_n$ is a normalisation constant. The two families that are most relevant in the context of the spectral-element method are the Legendre polynomials and the Lobatto polynomials.

Legendre polynomials, denoted by $L_n(x)$, are orthogonal with respect to the flat integration weight $w(x) = 1$ and the integration interval $[-1, 1]$, i.e.,

$$\int_{-1}^{1} L_n(x)L_m(x)dx = A_n\delta_{mn}, \quad L_n(x) = \frac{1}{2^n n!} \frac{d}{dx}(x^2 - 1)^n$$  \hfill (22)

Legendre polynomial can be defined by the following recursion relation with $L_0(x) = 1$ and $L_1(x) = x$:

$$L_{n+1}(x) = \frac{1}{n+1}((2n+1)xL_n(x) - nL_{n-1}(x))$$  \hfill (23)

Lobatto polynomials, $L_{o_n}(x)$, are defined in terms of the Legendre polynomials $L_n(x) = \frac{d}{dx}L_{n+1}(x)$. The Lobatto polynomials are the family that is orthogonal with respect to the integration weight $w(x) = 1 - x^2$ and the integration interval $[-1, 1]$, i.e.,

$$\int_{-1}^{1} (1 - x^2)L_{o_n}(x)L_{o_m}(x)dx = A_n\delta_{mn}.$$  \hfill (24)

SEM employs Gauss-Lobatto-Legendre (GLL) integration points for Lagrange polynomial, leading to the diagonal mass matrix for efficient matrix inversion through explicit time integration. GLL points $\xi_j (0 \leq j \leq N)$ are roots of the polynomial:

$$(1 - \xi^2)L_N(\xi) = (1 - \xi^2)L_{o_N}(\xi) = 0$$  \hfill (25)

A big advantage of GLL points is the application of Gauss-Lobatto-Legendre quadrature for spatial integration:

$$\int_{-1}^{1} f(\xi)d\xi \approx \int_{-1}^{1} f(\xi_j)\phi_j(\xi)d\xi = \sum_{j=0}^{N} f(\xi_j) \int_{-1}^{1} \phi_j(\xi)d\xi = \sum_{j=0}^{N} f(\xi_j)\omega_j$$  \hfill (26)

where $\omega_j = \int_{-1}^{1} \phi_j(\xi)d\xi$ denote the weights associated with the GLL point $\xi_j$. Both GLL points and the associated weights have analytic solutions:

$$x_0 = -1, x_N = 1, x_i = \text{roots of } L_N(x); \quad \omega_j = \frac{2}{N(N + 1)} \frac{1}{[L_N(\xi_j)]^2} \quad (\forall \xi_j \neq \pm 1); \quad \omega_j = \frac{2}{N(N + 1)} \quad (\xi_j = \pm 1)$$  \hfill (27)

Thus, equation (17) reads:

$$M_{ji}^s = \int_{-1}^{1} \rho(x)\phi_j(x)\phi_i(x)\frac{dx}{d\xi}d\xi = \omega_j\rho(\xi)\delta_{ij}\frac{dx}{d\xi}\bigg|_{\xi=\xi_j},$$

$$K_{ji}^s = \int_{-1}^{1} \rho(x)\partial_t\phi_j(x)\partial_t\phi_i(x)(\frac{dx}{d\xi})^2\frac{dx}{d\xi}d\xi = \sum_{k=0}^{N} \omega_k\rho(\xi)\partial_t\phi_j(\xi)\partial_t\phi_i(\xi)(\frac{dx}{d\xi})^2\frac{dx}{d\xi}\bigg|_{\xi=\xi_k},$$

$$f_j^s = \int_{-1}^{1} \phi_j(x)f(x,t)\frac{dx}{d\xi}d\xi = \omega_jf(x,t)\frac{dx}{d\xi}\bigg|_{\xi=\xi_j},$$  \hfill (28)

where the cardinal interpolation point of the Lagrange polynomial $\phi_j(\xi) = \delta_{ij}$ has been utilized. Due to property of Kronecker delta, $M_{ji}^s = 0$ for all $j \neq i$. In other words, the mass matrix is diagonal, and the diagonal terms are $M_{jj}^s = \omega_j\rho(\xi)\frac{dx}{d\xi}|_{\xi=\xi_j}$. Note that the stiffness matrix is not diagonal. Keep in mind that the source is injected at specific location, that is, $f(x,t) = s(t)\delta(x - x_s)$, where $s(t)$ is the source time function. When transferring from physical
coordinate $x$ to reference coordinate $\xi$, we have $\delta(x - x_s) = \frac{1}{J_e} \delta(\xi - \xi_s)$, leading to

$$f_j^g = \int_{-1}^{1} \phi_j[x(\xi)] f(x(\xi), t) J_e d\xi = s(t) \int_{-1}^{1} \phi_j[x(\xi)] \frac{1}{J_e} \delta(\xi - \xi_s) J_e d\xi$$

$$= s(t) \phi_j[x(\xi_s)] \int_{-1}^{1} \delta(\xi - \xi_s) d\xi = w_j^s s(t)$$

(29)

where the weight $w_j^s$ must be computed through Lagrange interpolation evaluated at source location $x_s$.

The accuracy of spectral-element solutions is controlled by both the size of the elements and the degree of the Lagrange polynomials. Decreasing the size of the elements and increasing the polynomial degree will mostly lead to higher accuracy. However, the maximum degree is very limited by the CFL stability condition and the available computational resources. This is because the distance between the first two GLL points decreases as $O(N - 2)$, i.e. quadratically with increasing degree. Thus, choosing $N \geq 8$ usually results in unreasonably small step lengths. On the other hand, by choosing $N < 4$ one sacrifices much of the accuracy of the spectral-element method. A compromise based on experience, is to use polynomial degrees between 4 and 7. One should use at least 5 grid points per minimum wavelength in order to obtain accurate solutions when the propagation distance is on the order of 10 to 50 wavelengths.

2.6 Global assembly

The classic FEM/SEM assumes continuity of the solution fields at the element boundaries. Therefore, we simply need to add up the elemental solutions at the corresponding boundary collocation points. In the spectral-element, method each element boundary thus only has one value. For a system with $n_e$ elements and given polynomial order $N$, the global number of collocation points $n_g$ of our system is $n_g = n_e \times N + 1$. Assume $n_e = 3$, $N = 2$. Since the diagonal elements of the mass matrix are allowed to be stored in a vector, we obtain the global mass matrix vector as

$$M = \begin{bmatrix} M_{1,1}^e & 0 & 0 & 0 & M_{1,1}^e \\ M_{2,2}^e & M_{2,2}^e & 0 & 0 & M_{2,2}^e \\ M_{3,3}^e & M_{3,3}^e & 0 & 0 & M_{3,3}^e \\ 0 & 0 & M_{3,3}^e & 0 & 0 \\ 0 & 0 & 0 & M_{3,3}^e & 0 \\ 0 & 0 & 0 & 0 & M_{3,3}^e \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} M_{1,1}^e + M_{1,1}^g \\ M_{2,2}^e + M_{2,2}^g \\ M_{3,3}^e + M_{3,3}^g \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

(30)

Similarly, we obtain the global stiffness matrix

$$K = \begin{bmatrix} K_{1,1}^e & K_{1,2}^e & K_{1,3}^e & 0 & 0 & 0 & 0 \\ K_{2,1}^e & K_{2,2}^e & K_{2,3}^e & 0 & 0 & 0 & 0 \\ K_{3,1}^e & K_{3,2}^e & K_{3,3}^e + K_{1,1}^e & K_{3,2}^e & K_{3,3}^e & 0 & 0 \\ 0 & 0 & K_{1,1}^e & K_{2,2}^e & K_{2,3}^e & K_{3,3}^e & 0 \\ 0 & 0 & K_{3,1}^e & K_{3,2}^e & K_{3,3}^e + K_{1,1}^e & K_{3,2}^e & K_{3,3}^e \\ 0 & 0 & 0 & 0 & K_{3,1}^e & K_{3,2}^e & K_{3,3}^e \end{bmatrix}$$

(31)

and the source vector

$$f = \begin{bmatrix} f_1^g \\ f_2^g + f_1^g \\ f_3^g + f_1^g \\ f_1^g + f_2^g \\ f_2^g + f_3^g \\ f_3^g \end{bmatrix}$$

(32)

We end up with a system of equations for $n_g = n_e \times N + 1$ coefficients for the displacement $u_g$, where $N$ is the interpolation order and $n_e$ is the number of elements. The matrices $M$ and $K$ have dimensions $n_g \times n_g$. Due to its diagonal structure, $M$ is stored as a vector including $n_g$ diagonal elements to save memory in computer. The force vector $f$ also has $n_g$ elements. A simple centred finite difference discretization for second order time derivative leads to the wave extrapolation

$$u^{n+1} = 2u^n - u^{n-1} + \Delta t^2 M^{-1}(f^n - Ku^n)$$

(33)

where the efficient implementation is achieved thanks to the diagonal mass matrix $M$. 

5
2.7 Time marching using Newmark scheme

A complete matrix expression of SEM including the absorbing boundary condition reads

\[ Ma + Cv + Ku = f \quad \text{with} \quad a = \ddot{u} = \partial_{tt}u, \quad v = \dot{u} = \partial_t u. \]  

(34)

where \( C \) is the impedance. Here we consider the explicit time marching using the Newmark scheme with \( \beta = 0, \gamma = 1/2 \) involving three phases/steps:

- **Prediction phase**
  \[ u^{n+1/2} = u^n + 0.5\Delta t v^n \]  
  (35)

- **Resolution phase**
  \[ Ma^{n+1/2} = Ku^{n+1/2} - Cv^{n+1/2} + f^{n+1/2} \]  
  (36)

where the term \( Cv^{n+1/2} \) does not vanish only at the absorbing boundary. Since \( v^{n+1/2} = v^n + 0.5\Delta t a^{n+1/2} \) at the boundary the mass matrix \( M \) at the boundary also needs to be modified.

\[ (M + 0.5\Delta t(C))a^{n+1/2} = -Ku^{n+1/2} - Cv^n + f^{n+1/2} \]  

- **Correction phase**
  \[ v^{n+1} = v^n + 0.5\Delta t a^{n+1/2} \]  
  (37)

\[ u^{n+1} = u^{n+1/2} + 0.5\Delta t v^{n+1} \]  

(38)

While being more complicated than [33], the Newmark scheme has the advantageous property of conserving linear and angular momentum [Fichtner, 2011]. Another updating is equivalent to above as follows

- **Prediction phase**
  \[ v^{n+1/2} = v^n + 0.5\Delta t a^n \]  
  (39)

\[ u^{n+1} = u^n + \Delta t v^n + 0.5\Delta t^2 a^n = u^n + \Delta t v^{n+1/2} \]  

(40)

- **Resolution phase**
  \[ Ma^{n+1} = -Ku^{n+1} - Cv^{n+1} + f^{n+1} \]  
  (41)

Since \( v^{n+1} = v^{n+1/2} + 0.5\Delta t a^{n+1} \) at the boundary the mass matrix \( M \) at the boundary also needs to be modified.

\[ (M + 0.5\Delta t(C))a^{n+1} = -Ku^{n+1} - Cv^{n+1/2} + f^{n+1} \]  

- **Correction phase**
  \[ v^{n+1} = v^{n+1/2} + 0.5\Delta t a^{n+1} \]  
  (42)

3 SEM in 2D and 3D

3.1 Preliminary: Integral expansion by tensor product

The GLL points in 2D or 3D are simply defined based on tensorial products of GLL points in 1D thanks to the tensorial property of quadrangular and hexahedral elements. In reference space, they can be represented as tensorial products of \([-1,+1]\) interval: \( \Omega_c = [-1,+1] \times [-1,+1] \) in 2D and \( \Omega_c = [-1,+1] \times [-1,+1] \times [-1,+1] \) in 3D.

In 1D, the two Lagrange polynomials of degree 1 with two control points, \( \xi = -1 \) and \( \xi = 1 \), are \( \phi_0^{(1)}(\xi) = (\xi + 1)/2 \) and \( \phi_1^{(1)}(\xi) = (1 - \xi)/2 \), and the three Lagrange polynomials of degree 2 with three control points, \( \xi = -1, \xi = 0 \) and \( \xi = 1 \), are \( \phi_0^{(2)}(\xi) = \xi(\xi + 1)/2 \), \( \phi_1^{(2)}(\xi) = \xi(1 - \xi)/2 \) and \( \phi_2^{(2)}(\xi) = 1 - \xi^2 \).

In 2D, the four shape functions associated with the four-anchor quadrilateral element are products of degree 1 Lagrange polynomials: \( N_1(\xi, \eta) = \phi_0^{(1)}(\xi)\phi_0^{(1)}(\eta) \), \( N_2(\xi, \eta) = \phi_1^{(1)}(\xi)\phi_0^{(1)}(\eta) \), \( N_3(\xi, \eta) = \phi_0^{(1)}(\xi)\phi_1^{(1)}(\eta) \) and \( N_4(\xi, \eta) = \phi_1^{(1)}(\xi)\phi_1^{(1)}(\eta) \). Similarly, the shape functions of nine-anchor quadrilateral elements are products of degree 2 Lagrange polynomials. Define a set of \( n_a \) shape functions \( N_a(\xi, \eta) \) and control points \( x_a = (x_a, y_a) \). a point \( x = (x, y) \) in the reference space \( (\xi, \eta) \) can be written as

\[ x = \sum_{a=1}^{n_a} N_a(\xi, \eta) x_a, \]  

(43)
leading to
\[
\begin{aligned}
x_\xi &= \sum_{a=1}^{n_a} \frac{\partial}{\partial \xi} N_a(\xi, \eta)x_a, \\
y_\xi &= \sum_{a=1}^{n_a} \frac{\partial}{\partial \eta} N_a(\xi, \eta)y_a.
\end{aligned}
\] (44)

Then the Jacobian is given by
\[
J = \frac{\partial(x, y)}{\partial(\xi, \eta)} = \left| \begin{array}{cc}
\frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} \\
\frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta}
\end{array} \right| = x_\xi y_\eta - x_\eta y_\xi.
\] (45)

The inverse Jacobian matrix is
\[
\frac{\partial(\xi, \eta)}{\partial(x, y)} = \left[ \begin{array}{cc}
\frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} \\
\frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta}
\end{array} \right] = \frac{1}{J} \left[ \begin{array}{cc}
\frac{\partial y}{\partial \xi} & -\frac{\partial x}{\partial \eta} \\
-\frac{\partial y}{\partial \eta} & \frac{\partial x}{\partial \xi}
\end{array} \right]
\] (46)
leading to
\[
\xi_x = \frac{y_\eta}{J}, \quad \xi_y = -\frac{x_\eta}{J}, \quad \eta_x = \frac{x_\xi}{J}, \quad \eta_y = -\frac{y_\xi}{J}.
\] (47)

3.2 SEM for anisotropic elastic wave equation (Fichtner, 2011)

The stress-strain relation in 3D anisotropic elastic medium is governed by \(\sigma_{ij} = c_{ijkl}\epsilon_{kl}\) with the strain \(\epsilon_{kl} = 1/2(\partial_k u_l + \partial_l u_k)\). We may write in matrix
\[
\left[ \begin{array}{c}
s_{xx} \\
s_{yy} \\
s_{zz} \\
s_{xy} \\
s_{xz} \\
s_{yz}
\end{array} \right]
= \left[ \begin{array}{cccccc}
c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} \\
c_{12} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} \\
c_{13} & c_{23} & c_{33} & c_{34} & c_{35} & c_{36} \\
c_{14} & c_{24} & c_{34} & c_{44} & c_{45} & c_{46} \\
c_{15} & c_{25} & c_{35} & c_{45} & c_{55} & c_{56} \\
c_{16} & c_{26} & c_{36} & c_{46} & c_{56} & c_{66}
\end{array} \right] \left[ \begin{array}{c}
\epsilon_{xx} \\
\epsilon_{yy} \\
\epsilon_{zz} \\
\epsilon_{xy} \\
\epsilon_{xz} \\
\epsilon_{yz}
\end{array} \right]
\] (48)

The 3D anisotropic elastic wave equation is given by \(\rho \partial_{tt}^2 u = \nabla \cdot \sigma + f\), which reads componentwisely
\[
\rho \partial_{tt}^2 u_p = \partial_t \sigma_{pq} + f_p, \quad p \in \{x, y, z\}, \quad \partial_t \sigma_{pq} = \partial_x \sigma_{px} + \partial_y \sigma_{py} + \partial_z \sigma_{pz}
\] (49)

where the Einstein summation rule for repeated indices has been applied. Apply the tensor product of 3 1D basis function as the 3D basis function \(\phi_{ijk}(\xi) = \phi_i(\xi)\phi_j(\xi_2)\phi_k(\xi_3)\), we have
\[
\int_{\Omega} \rho \phi_{ijk} \partial_{tt}^2 u_p d\mathbf{x} = \int_{\Omega} \phi_{ijk} \partial_t \sigma_{pq} d\mathbf{x} + \int_{\Omega} \phi_{ijk} f_p d\mathbf{x}
\] (50)

where the free surface boundary condition has been utilized through integration by parts. We may expand the displacement and the stress in the following
\[
u_p(\mathbf{x}, t) \approx \vec{u}_p(\mathbf{x}, t) = \sum_{i,j,k=0}^{N} \nu_p^{ijk}(t) \phi_{ijk}(\mathbf{x}), \quad \sigma_{pq}(\mathbf{x}, t) \approx \sigma_{pq}(\mathbf{x}, t) = \sum_{i,j,k=0}^{N} \sigma_{pq}^{ijk}(t) \phi_{ijk}(\mathbf{x})
\] (51)

- On each element, the first term on the left hand side is then
\[
\begin{aligned}
\int_{\Omega} \rho \phi_{ijk} \partial_{tt}^2 u_p d\mathbf{x} &= \sum_{q,r,s=0}^{N} \partial_t^2 u_p^{qrs}(t) \int_{\Omega} \rho(\mathbf{x}) \phi_{ijk}(\mathbf{x}) \phi_{qrs}(\mathbf{x}) d\mathbf{x} \\
&= \sum_{q,r,s=0}^{N} \partial_t^2 u_p^{qrs}(t) \int_{\Omega} \rho(\mathbf{\xi}) \phi_{ijk}(\mathbf{\xi}) \phi_{qrs}(\mathbf{\xi}) \frac{\partial \mathbf{\xi}}{\partial \xi_j} d\mathbf{\xi} \\
&= \sum_{q,r,s=0}^{N} \partial_t^2 u_p^{qrs}(t) \sum_{l,m,n=0}^{N} \omega \omega_m \omega_n \rho^{lmn} \phi_{ijkl}^{qrs} \phi_{qrs}^{lmn} \\
&= \sum_{q,r,s=0}^{N} \partial_t^2 u_p^{qrs}(t) \sum_{l,m,n=0}^{N} \omega \omega_m \omega_n \rho^{lmn} \delta_{qrs}^{lmn} \delta_{ijkl}^{qrs} \\
&= \omega \omega_j \omega_k \rho_{ijkl}^{qrs} \partial_t^2 u_p^{qrs}
\end{aligned}
\] (52)
The second term on the left hand side is
\[
\int_{\Omega_e} \partial_{\xi jk} \sigma_{pq} \xi_j \mathrm{d}\xi = \sum_{q=1}^{3} \int_{\Omega_e} \partial_{\xi jk} \sigma_{pq} \xi_j \mathrm{d}\xi = \sum_{q,a=1}^{3} \int_{\Omega_e} \frac{\partial \xi_a}{\partial x_q} \frac{\partial \xi_j}{\partial x_a} \phi_{ijk} \sigma_{pq} \xi_j \mathrm{d}\xi
\]
\[
= \sum_{q,a=1}^{3} \int_{\Omega_e} \frac{\partial \xi_a}{\partial x_q} \frac{\partial \xi_j}{\partial x_a} [\phi_{ij} (\xi_1) \phi_{j2} (\xi_2) \phi_{k} (\xi_3) \sigma_{pq} \xi_j] \mathrm{d}\xi
\]
\[
= \sum_{q=1}^{3} \int_{\Omega_e} \frac{\partial \xi_a}{\partial x_q} \frac{\partial \xi_j}{\partial x_a} [\phi_{ij} (\xi_1) \phi_{j2} (\xi_2) \phi_{k} (\xi_3) \sigma_{pq} \xi_j] \mathrm{d}\xi + \int_{\Omega_e} \frac{\partial \xi_j}{\partial x_q} \phi_{ij} (\xi_1) \phi_{j2} (\xi_2) \phi_{k} (\xi_3) \sigma_{pq} \xi_j \mathrm{d}\xi + \int_{\Omega_e} \frac{\partial \xi_j}{\partial x_q} \phi_{ij} (\xi_1) \phi_{j2} (\xi_2) \phi_{k} (\xi_3) \sigma_{pq} \xi_j \mathrm{d}\xi
\]
\[
= \sum_{q=1}^{3} \sum_{l,m,n=0}^{N} \omega_{l} \omega_{m} \omega_{n} \frac{\partial \xi_l}{\partial x_q} \frac{\partial \xi_j}{\partial x_a} \phi_{ij} (\xi_1) \phi_{j2} (\xi_2) \phi_{k} (\xi_3) \sigma_{pq} \xi_j \mathrm{d}\xi + \sum_{l,m,n=0}^{N} \omega_{l} \omega_{m} \omega_{n} \frac{\partial \xi_l}{\partial x_q} \frac{\partial \xi_j}{\partial x_a} \phi_{ij} (\xi_1) \phi_{j2} (\xi_2) \phi_{k} (\xi_3) \sigma_{pq} \xi_j \mathrm{d}\xi
\]
\[
= \sum_{q=1}^{3} \sum_{l=0}^{N} \omega_{l} \omega_{m} \omega_{n} \frac{\partial \xi_l}{\partial x_q} \frac{\partial \xi_j}{\partial x_a} \phi_{ij} (\xi_1) \phi_{j2} (\xi_2) \phi_{k} (\xi_3) \sigma_{pq} \xi_j \mathrm{d}\xi + \sum_{l=0}^{N} \omega_{l} \omega_{m} \omega_{n} \frac{\partial \xi_l}{\partial x_q} \frac{\partial \xi_j}{\partial x_a} \phi_{ij} (\xi_1) \phi_{j2} (\xi_2) \phi_{k} (\xi_3) \sigma_{pq} \xi_j \mathrm{d}\xi
\]
\[
= \sum_{q=1}^{3} [\omega_{l} \omega_{m} \omega_{n} \frac{\partial \xi_l}{\partial x_q} \frac{\partial \xi_j}{\partial x_a} \phi_{ij} (\xi_1) \phi_{j2} (\xi_2) \phi_{k} (\xi_3) \sigma_{pq} \xi_j \mathrm{d}\xi] + \sum_{l=0}^{N} \omega_{l} \omega_{m} \omega_{n} \frac{\partial \xi_l}{\partial x_q} \frac{\partial \xi_j}{\partial x_a} \phi_{ij} (\xi_1) \phi_{j2} (\xi_2) \phi_{k} (\xi_3) \sigma_{pq} \xi_j \mathrm{d}\xi
\]
\[
= \frac{3}{\omega_{l} \omega_{m} \omega_{n} \frac{\partial \xi_l}{\partial x_q} \frac{\partial \xi_j}{\partial x_a} \phi_{ij} (\xi_1) \phi_{j2} (\xi_2) \phi_{k} (\xi_3) \sigma_{pq} \xi_j \mathrm{d}\xi] + \sum_{l=0}^{N} \omega_{l} \omega_{m} \omega_{n} \frac{\partial \xi_l}{\partial x_q} \frac{\partial \xi_j}{\partial x_a} \phi_{ij} (\xi_1) \phi_{j2} (\xi_2) \phi_{k} (\xi_3) \sigma_{pq} \xi_j \mathrm{d}\xi
\]
\[
= \frac{3}{\omega_{l} \omega_{m} \omega_{n} \frac{\partial \xi_l}{\partial x_q} \frac{\partial \xi_j}{\partial x_a} \phi_{ij} (\xi_1) \phi_{j2} (\xi_2) \phi_{k} (\xi_3) \sigma_{pq} \xi_j \mathrm{d}\xi] + \sum_{l=0}^{N} \omega_{l} \omega_{m} \omega_{n} \frac{\partial \xi_l}{\partial x_q} \frac{\partial \xi_j}{\partial x_a} \phi_{ij} (\xi_1) \phi_{j2} (\xi_2) \phi_{k} (\xi_3) \sigma_{pq} \xi_j \mathrm{d}\xi
\]

Assume a point force source \( f(x, t) = s(t) \delta(x - x_s) \), the source term on the right hand side is
\[
\int_{\Omega_e} \phi_{ij}(x) f_p(x, t) \mathrm{d}x = s(t) \int_{\Omega_e} \phi_{ij}(\xi) \frac{1}{\xi_j} \delta(\xi - \xi_s) J_\xi \mathrm{d}\xi = s(t) \int_{\Omega_e} \phi_{ij}(\xi_1) \delta(\xi - \xi_s) \mathrm{d}\xi
\]
\[
= \phi_{ij}(\xi_1) s(t) = \phi_{ij}(\xi_1) \phi_{j2} (\xi_2) \phi_{k} (\xi_3) s(t) = \omega_{l} \omega_{m} \omega_{n} \phi_{ij} (\xi_1) \phi_{j2} (\xi_2) \phi_{k} (\xi_3) s(t).
\]

where \( \omega_{l} = \phi_{ij} (\xi_1) \phi_{j2} (\xi_2) \phi_{k} (\xi_3) \) must be evaluated through Lagrange interpolation.

The paraxial approximation absorbing boundary condition proposed by [Clayton and Engquist, 1977] relates traction to velocity [Komatitsch and Tromp, 1999] via
\[
\sigma \cdot n = \rho \{ v_n \cdot \partial_t u + v_1 (t_1 \cdot \partial_t u) t_1 + v_2 (t_2 \cdot \partial_t u) t_2 \}
\]
where \( t_1 \) and \( t_2 \) are orthogonal unit vectors tangential to the absorbing boundary with unit outward normal \( n \), \( v_n \) is the quasi-P wave speed of waves travelling in the \( n \) direction, \( v_1 \) is the quasi-S wave speed of waves polarized in the \( t_1 \) direction, and \( v_2 \) is the quasi-S wave speed of waves polarized in the \( t_2 \) direction. On each element the boundary condition is then
\[
\int_{\partial \Omega} \phi_{ij} \sigma_{pq} \cdot \mathrm{d}\Sigma = \int_{\partial \Omega} \rho \{ v_n \cdot \partial_t u + v_1 (t_1 \cdot \partial_t u) t_1 + v_2 (t_2 \cdot \partial_t u) t_2 \} \phi_{ij} \cdot \mathrm{d}\Sigma
\]

### 3.3 2D SEM for P-SV system

In 2-D P-SV system, the stress-strain relation in [48] reduces to
\[
\begin{bmatrix}
\sigma_{xx} \\
\sigma_{yy} \\
\sigma_{xy}
\end{bmatrix} = \begin{bmatrix}
c_{11} & c_{13} & 0 \\
c_{13} & c_{33} & 0 \\
0 & 0 & c_{55}
\end{bmatrix} \begin{bmatrix}
\partial_x u_x \\
\partial_y u_y \\
\partial_x u_y + \partial_y u_x
\end{bmatrix}
\]
\[
\Rightarrow \begin{bmatrix}
\sigma_{xx} = c_{11} \partial_x u_x + c_{13} \partial_y u_x \\
\sigma_{yy} = c_{13} \partial_x u_y + c_{33} \partial_y u_y \\
\sigma_{xy} = c_{55} (\partial_x u_y + \partial_y u_x)
\end{bmatrix}
\]

On each element, the first and the third terms are
\[
\int_{\Omega_e} \rho \phi_{ij} \partial_t^2 u_p \xi_j \mathrm{d}x = \omega_{l} \omega_{m} \omega_{n} \partial_t^2 u_p \xi_j \mathrm{d}x \quad \text{and} \quad \int_{\Omega_e} \phi_{ij}(x) f_p(x, t) \mathrm{d}x = \omega_{l} \omega_{m} \omega_{n} \phi_{ij}(x_1) \phi_{j2}(x_2) \phi_{k}(x_3) s(t).
\]
The second term associated with the stiffness matrix applied to vector is given by
\[
\int_{\Omega_e} \varphi_{ij} \sigma_{pq} \, dx = \int_{\Omega_e} \left( \sigma_{px} \partial_x \varphi_{ij} + \sigma_{pz} \partial_z \varphi_{ij} \right) \, dx
\]
\[
= \int_{\Omega_e} \left( \sigma_{px} \left( \partial_x \varphi_{ij} \xi_x + \partial_z \varphi_{ij} \eta_x \right) + \sigma_{pz} \left( \partial_x \varphi_{ij} \xi_z + \partial_z \varphi_{ij} \eta_z \right) \right) J_e \, d\xi
\]
\[
= \int_{\Omega_e} \left( \partial_x \varphi_{ij} \left( \sigma_{px} \xi_x + \sigma_{pz} \xi_z \right) + \partial_z \varphi_{ij} \left( \sigma_{px} \eta_x + \sigma_{pz} \eta_z \right) \right) J_e \, d\xi
\]
\[
= \sum_{r,s=0}^{N} \omega_{r,s} J_e^s \left[ \partial_x \varphi_{ij} \left( \sigma_{px} \xi_x + \sigma_{pz} \xi_z \right) + \partial_z \varphi_{ij} \left( \sigma_{px} \eta_x + \sigma_{pz} \eta_z \right) \right] \tau^r, \quad p \in \{x, z\}.
\]
where \( \xi = (\xi_1, \xi_2) = (\xi, \eta) \). Since
\[
\partial_x \varphi_{ij}^s = \varphi_{ij}^s(\xi^r) \varphi_{ij}^s(\eta^r) = \varphi_{ij}^s(\xi^r) \varphi_{ij}^s(\eta^r) = \varphi_{ij}^s(\eta^r) \delta_{ir},
\]
then
\[
\int_{\Omega_e} \varphi_{ij} \sigma_{pq} \, dx = \sum_{r=0}^{N} \omega_{r} J_e^r \varphi_{ij}^r(\xi^r) \left( \sigma_{px}^{r} \xi_x^{r} + \sigma_{pz}^{r} \xi_z^{r} \right) + \sum_{s=0}^{N} \omega_{s} J_e^s \varphi_{ij}^s(\eta^s) \left( \sigma_{px}^{s} \eta_x^{s} + \sigma_{pz}^{s} \eta_z^{s} \right), \quad p \in \{x, z\}
\]
Speaking componentwise, we have
\[
\int_{\Omega_e} \varphi_{ij} \sigma_{pq} \, dx = \sum_{k=0}^{N} \omega_k \varphi_{ij}^k(\xi) \omega_k J_e^k \left( \sigma_{px}^{k} \xi_x^{k} + \sigma_{pz}^{k} \xi_z^{k} \right) + \sum_{k=0}^{N} \omega_k \varphi_{ij}^k(\eta) \omega_k J_e^k \left( \sigma_{px}^{k} \eta_x^{k} + \sigma_{pz}^{k} \eta_z^{k} \right)
\]
where we know from the chain rule
\[
\sigma_{px}^{rs} = c_{11} \partial_x \xi_x^{rs} + c_{12} \partial_x \xi_z^{rs} = c_{11} \left( \xi_x^{rs} \partial_x \varphi_{ij}^s + \eta_x^{rs} \partial_x \varphi_{ij}^s \right) + c_{12} \left( \xi_z^{rs} \partial_x \varphi_{ij}^s + \eta_z^{rs} \partial_x \varphi_{ij}^s \right)
\]
\[
\sigma_{pz}^{rs} = c_{13} \partial_x \xi_x^{rs} + c_{33} \partial_x \xi_z^{rs} = c_{13} \left( \xi_x^{rs} \partial_x \varphi_{ij}^s + \eta_x^{rs} \partial_x \varphi_{ij}^s \right) + c_{33} \left( \xi_z^{rs} \partial_x \varphi_{ij}^s + \eta_z^{rs} \partial_x \varphi_{ij}^s \right)
\]
\[
\sigma_{zx}^{rs} = c_{44} \partial_x \xi_x^{rs} + \partial_x \xi_z^{rs} = c_{44} \left( \xi_x^{rs} \partial_x \varphi_{ij}^s + \eta_x^{rs} \partial_x \varphi_{ij}^s \right)
\]
in which
\[
\partial_x \xi_x^{rs} = \sum_{m,n=0}^{N} u_x^{mn}(t) \varphi_m^{rs}(\xi^r) \varphi_n^{rs}(\eta^r) = \sum_{m,n=0}^{N} u_x^{mn}(t) \varphi_m^{rs}(\eta^r),
\]
\[
\partial_x \xi_z^{rs} = \sum_{m,n=0}^{N} u_z^{mn}(t) \varphi_m^{rs}(\xi^r) \varphi_n^{rs}(\eta^r) = \sum_{m,n=0}^{N} u_z^{mn}(t) \varphi_m^{rs}(\eta^r),
\]
\[
\partial_x \eta_x^{rs} = \sum_{m,n=0}^{N} u_x^{mn}(t) \varphi_m^{rs}(\xi^r) \varphi_n^{rs}(\eta^r) = \sum_{m,n=0}^{N} u_x^{mn}(t) \varphi_m^{rs}(\eta^r),
\]
\[
\partial_x \eta_z^{rs} = \sum_{m,n=0}^{N} u_z^{mn}(t) \varphi_m^{rs}(\xi^r) \varphi_n^{rs}(\eta^r) = \sum_{m,n=0}^{N} u_z^{mn}(t) \varphi_m^{rs}(\eta^r).
\]

### 3.4 Attenuation

By introducing memory variables, the generalized Maxwell body (equivalent to standard linear solid model) gives constitutive relation \[\text{Yang} \text{ et al., 2016a}\]
\[
\sigma_{ij}(x, t) = c_{ijkl}(x) \left( \epsilon_{kl}(x, t) - \sum_{l=1}^{L} \gamma_{ijkl}(x) \epsilon_{kl}(x, t) \right), \quad \partial_t \epsilon_{kl}(x, t) + \omega \epsilon_{kl}(x, t) = \omega \epsilon_{kl}(x, t).
\]
After discretization, the memory variables should be updated through \(\text{Yang} \text{ et al., 2016b}\), see Appendix
\[
\epsilon_{kl}^{n+1} = e^{-\omega \Delta t} \epsilon_{kl}^{n} + (1 - e^{-\omega \Delta t}) \epsilon_{kl}^{n+\frac{1}{2}}, \quad l = 1, \ldots, L.
\]
To preserve the second order accuracy for the time discretization, the following approximation is useful \(\text{Moczo} \text{ et al., 2007a,b}\):
\[
\epsilon_{kl}^{n+\frac{3}{2}} \approx \frac{1}{2} \left( \epsilon_{kl}^{n} + \epsilon_{kl}^{n+1} \right), \quad l = 1, \ldots, L.
\]
3.5 CFL stability condition - $\Delta t$ (Komatitsch et al., 2005)

To guarantee accurate simulation of SEM, one needs to follow the CFL condition

$$C = \frac{\Delta t \max(V_p)}{\Delta x} < C_{\text{max}}$$  \hspace{1cm} (68)

where $\Delta x$ is the minimum distance between two neighboring points in SEM mesh, $\max(V_p)$ is the maximum velocity of the P-wave velocity. The heuristic rule of thumb that we use in practice is that for regular meshes $C_{\text{max}} = 0.5$, while for very irregular meshes with distorted elements and/or very heterogeneous media $C_{\text{max}}$ reduces to approximately 0.3 to 0.4. Typically, we restrict ourselves to lower order polynomial due to the computational cost by using higher order polynomials, i.e., $N = 4, 5$. Assume the element size in $x$ direction is $h$. In this case, the condition becomes

$$N = 4, \Delta x = 0.173h, \Delta t < 0.173 \times C_{\text{max}} \times \frac{h}{\max(V_p)}$$ \hspace{1cm} (69)

3.6 Resolution - $\Delta x$ (Komatitsch, 1997)

The minimum wavelength is defined as

$$\lambda_{\text{min}} = \frac{\min(V_s)}{f_{\text{max}}},$$ \hspace{1cm} (70)

where $V_s$ is the S-wave velocity and $f_{\text{max}}$ the highest frequency you would like to resolve, e.g. the maximum frequency at which the source spectrum has significant power (for a Ricker wavelet $f_{\text{max}} = 2.5 f_0$). For an element of size $h$ and polynomial order $N$, the number of nodes per shortest wavelength $\lambda_{\text{min}}$ should be larger than 4.5 - 5, i.e.,

$$G = \frac{\lambda_{\text{min}}}{h/N} \geq 5.$$ \hspace{1cm} (71)

4 Full waveform inversion

4.1 Wavelet estimation

The initial guess of the source wavelet might be different from the true source wavelet. In order to compensate for the inaccuracy of the initial source wavelet, an unknown matching filter is required to be estimated. In the presence of a matching filter $f_m$, the objective function is written as

$$\chi = \frac{1}{2} \sum_{s,r} \int_0^T dt (f_m * u - d)^2$$ \hspace{1cm} (72)

where $u$ is the synthetic data based on an initial source wavelet $w_i$, $d$ is the observed data. It implies that the observed is the true wavelet convolved with Green’s function, $d = w * g$; while the synthetic data corresponds to initial wavelet convolved with Green’s function $u = w_i * g$. We intend to find the true wavelet $w$ which is the matching filter convolved with the initial wavelet, that is,

$$w = w_i * f_m.$$ \hspace{1cm} (73)

Due to the linear property of convolution operator, we can rewrite the convolution between synthetic data $u$ the filter $f_m$ through a matrix expression, i.e., $f_m * u = D_u f_m$ :

$$\chi = \frac{1}{2} \sum_{s,r} \int_0^T dt (D_u f_m - d)^2$$ \hspace{1cm} (74)

To find the solution $f_m$, the normal equation is obtained by

$$\frac{\partial \chi}{\partial f_m} = 0 \leftrightarrow D_u^t D_u f_m = D_u^t d$$ \hspace{1cm} (75)

Keep in mind that the adjoint of convolution is cross-correlation. In time domain, we simply build the convolution operator and the adjoint given the input data $u$ and a filter $f_m$, then solve the problem using least-squares linear solver, for example, Claerbout’s conjugate gradient algorithm. Once the filter is found, we obtain the estimated wavelet according to (73).
normal equation is then
\[ \sum_{s,r} F[u]^*F[f]F[f_m] = \sum_{s,r} F[u]^*F[d] \Rightarrow f_m = F^{-1} \left[ \frac{\sum_{s,r} F[u]^*F[d]}{\sum_{s,r} F[u]^*F[u] + \epsilon} \right] \] (76)

where \( \epsilon \) is stabilization factor to avoid division by zero; \( F^{-1} \) stands for inverse Fourier transform.

### 4.2 FWI gradient with respect to parameters (Kamath and Tsvankin, 2016)

In the following we assume the source wavelet is known without matching filter and we focus on finding the gradient of
the least-squares misfit function with respect to each parameter. The least-squares misfit function is
\[ \chi = \| u_i - d_i \|^2 := \int_T \int_\Omega dtdx(u_i(x,t) - d_i(x,t))^2\delta(x - x_r) \] (77)

Using the method of Lagrange multipliers, the augmented Lagrangian is defined as
\[ L = \chi + \int_T \int_\Omega dtdx\bar{u}_i \left[ \rho\partial_t^2 u_i - \partial_j c_{ijkl}\partial_l u_k - f_i \right] \] (78)

where \( \bar{u}_i \) is the adjoint state variable.

The perturbation of the Lagrangian is given by
\[
\delta L = \frac{1}{2} \sum_i \left\| u_i + \delta u_i - d_i \right\|^2 + \int_T \int_\Omega dtdx(\bar{u}_i + \delta \bar{u}_i) \left[ (\rho + \delta \rho)\partial_t^2 (u_i + \delta u_i) - \partial_j (c_{ijkl} + \delta c_{ijkl})\partial_l (u_k + \delta u_k) - f_i \right]
- \frac{1}{2} \sum_i \sum_r \left\| u_i - d_i \right\|^2 - \int_T \int_\Omega dtdx\bar{u}_i \left[ \rho\partial_t^2 u_i - \partial_j c_{ijkl}\partial_l u_k - f_i \right]
= \int_T \int_\Omega dtdx \sum_i (u_i - d_i)\delta u_i \delta(x - x_r) + \int_T \int_\Omega dtdx\bar{u}_i \left[ \rho\partial_t^2 u_i - \partial_j c_{ijkl}\partial_l u_k \right]
+ \int_T \int_\Omega dtdx \delta \bar{u}_i \left[ \rho\partial_t^2 u_i - \partial_j c_{ijkl}\partial_l u_k - f_i \right] + \int_T \int_\Omega dtdx\bar{u}_i \left[ \delta \rho\partial_t^2 u_i - \partial_j c_{ijkl}\partial_l u_k \right]
= \int_T \int_\Omega dtdx \delta \bar{u}_i \left[ \rho\partial_t^2 u_i - \partial_j c_{ijkl}\partial_l u_k + \sum_r (u_i - d_i)\delta u_r \delta(x - x_r) \right]
\]
adjoint equation

+ \int_\Omega dx \left[ \rho \bar{u}_i \delta \bar{u}_i - \rho \partial_t \bar{u}_i \delta u_i \right] \bigg|_{IC/FC=0} + \int_T dt \left[ c_{ijkl} \bar{u}_i \partial_l u_k - c_{ijkl} \partial_l u_i \delta u_k \right] \bigg|_{BC=0}
+ \int_T \int_\Omega dtdx \delta \bar{u}_i \left[ \rho\partial_t^2 u_i - \partial_j c_{ijkl}\partial_l u_k - f_i \right] \bigg|_{BC=0}
+ \int_T \int_\Omega dtdx \delta \bar{u}_i \rho \partial_t^2 u_i \delta r + \int_T \int_\Omega dtdx \delta \bar{u}_i \partial_l u_k \partial_l c_{ijkl} - \int_T dt \left[ \delta c_{ijkl} \bar{u}_i \partial_l u_k \right] \bigg|_{BC=0}
\] (79)

where we require the initial condition (IC) for the state variables \((u_i \text{ and } \delta u_i)\), the final condition (FC) for the adjoint variables \((\bar{u}_i \text{ and } \delta \bar{u}_i)\) and the boundary condition (BC) for state and adjoint variables to vanish:
\[
IC : \delta u_i(x,t)\big|_{t=0} = \partial_t \delta u_i(x,t)\big|_{t=0} = 0; \quad BC : \delta u_i(x,t)\big|_{x=\partial\Omega} = 0
FC : \bar{u}_i(x,t)\big|_{t=T} = \partial_t \bar{u}_i(x,t)\big|_{t=T} = 0; \quad BC : \bar{u}_i(x,t)\big|_{x=\partial\Omega} = 0
\] (80)

The Lagrange is stationary with respect to the variables \(u_i, \bar{u}_i \text{ and the medium properties } \rho \text{ and } c_{ijkl} \), which leads to the adjoint equation
\[
\rho\partial_t^2 u_i - \partial_j c_{ijkl}\partial_l u_t + \sum_r (u_i - d_i)\delta u_r \delta(x - x_r) = 0
\] (81)
and the gradient expressions of the misfit (the Lagrangian is the same as the misfit when the state and adjoint equations are satisfied, i.e., $\delta X = \delta L$)

$$\frac{\partial X}{\partial \rho} = \int_T dt \bar{u}_i \partial^2_t u_i = - \int_T dt \partial \bar{u}_i \partial_t u_i,$$  
$$\frac{\partial X}{\partial c_{ijkl}} = \int_T dt \partial_j \bar{u}_i \partial_l u_k$$  

(82)

5 How to run

suwaveform type=ricker1 dt=0.0005 ns=2000 fpeak=8 |sustrip > fricker.bin

echo "
===============================================================================
mode=0 // 0, forward modeling; 1, FWI; 2, RTM; 3, source inversion; 4, FWI gradient
acquifile=acqui.txt
rhofile=models/rho.bin
vpfile=models/vp.bin
vsfile=models/vs.bin
wltfile=fricker.bin

nt=2000 //number of time steps
dt=0.0005 //temporal sampling
n1=141 //size of input FD model
n2=141 //size of input FD model
d1=5 //grid spacing of the input FD model
d2=5 //grid spacing of the input FD model
ne1=25 //number of SEM elements in z direction
ne2=25 //number of SEM elements in x direction
freqmax=20 //2.5*f0 is the maximum freq
dr=5 //decimation ratio from CFL to Nyquist

===============================================================================
niter=50 //number of iterations using l-BFGS
nls=20 //number of line search per iteration
npair=5 //memory length in l-BFGS
invpar=1,2 //0, invert rho; 1, invert vp; 2, invert vs

bound=0 //bound the inversion parameters
minpar=1200,700 //lower bound for inversion parameters indexed above
maxpar=3700,2400 //upper bound for inversion parameters indexed above

===============================================================================
wxwdat=100 // dx for data weighting
xwdat=0,0.3,0.7,1,1 // weights for dx increment

===============================================================================
muteopt=0 // 0, no mute; 1, front mute; 2, tail mute; 3, front and tail mute
ntaper=20 //number of points for taper
xmute1=50,775.8,2227.4
tmute1=0.1,0.4,1
xmute2=50,603.3,1954.3
tmute2=0.3,0.64,1

" > inputpar.txt

to run sem2d code:
mpirun -np 2 ../bin/sem2d $(cat inputpar.txt) > &out
The output messages will be stored in the file `out`.

A  **Radiation pattern analysis** *(Kamath and Tsvankin, 2016)*

Consider frequency domain wave equation

\[-\omega^2 \rho u_t - \partial_j (c_{ijkl} \partial_l u_k) = f_i \tag{83}\]

Similarly, the perturbed wave field satisfy the same wave equation

\[-\omega^2 (\rho + \delta \rho) (u_t + \delta u_t) - \partial_j ((c_{ijkl} + \delta c_{ijkl}) \partial_l (u_k + \delta u_k)) = f_i \tag{84}\]

Under Born approximation, subtracting the above two equations gives

\[-\rho \omega^2 \delta u_t - \partial_j (c_{ijkl} \partial_l \delta u_k) = \omega^2 \delta \rho u_t + \partial_j (\delta c_{ijkl} \partial_l u_k) \tag{85}\]

The representation theorem using the Green’s function for the $n = 1, 2, 3$ component of the displacement field

\[u_n(x_r, \omega) = \int_{\Omega} d\mathbf{x}' G_{ni}(x_r, \mathbf{x}', \omega) h_i(\mathbf{x}', \omega) \tag{86}\]

yields

\[\delta u_n = \int_{\Omega} d\mathbf{x}' G_{ni} \omega^2 \delta \rho u_t + \partial_j (\delta c_{ijkl} \partial_l u_k) = \int_{\Omega} d\mathbf{x}' G_{ni} \omega^2 \delta \rho u_t + \int_{\partial \Omega} G_{ni} \delta c_{ijkl} \partial_l u_k n_j dS - \int_{\Omega} d\mathbf{x}' \partial_j G_{ni} \delta c_{ijkl} \partial_l u_k \tag{87}\]

We replace the displacement field by the product between the source at $x_s$ and the Green’s function $u_k(x', x_s, \omega) = f_m(x_s, \omega)G_{km}(x', x_s, \omega)$

\[\delta u_n = \int_{\Omega} d\mathbf{x}' G_{ni} \omega^2 \delta \rho u_t - \int_{\Omega} d\mathbf{x}' \partial_j G_{ni} \delta c_{ijkl} \partial_l u_k = \int_{\Omega} d\mathbf{x}' G_{ni} \omega^2 G_{im} f_m \delta \rho - \int_{\Omega} d\mathbf{x}' \partial_j G_{ni} \delta c_{ijkl} \partial_l u_k \tag{88}\]

Denote $\mathbf{p}^s$ and $\mathbf{p}^r$ are the unit slowness vector connecting the source $x_s$ and the receiver $x_r$ with the point scatter $\mathbf{x}'$.

**Applying the stationary phase method**, the asymptotic (high-frequency) representation of the Green’s function is given by *(Vavryčuk, 2007; Červený, 2001, eqn 2.5.75)*

\[G_{ni}(x_r, x', \omega) = \frac{1}{4\pi \rho \nu_g^2 \sqrt{K}} \frac{1}{R'} \exp\left(i \frac{\pi}{2} \sigma_0 + i\omega \nu_g^s \nu_g^r (x_r - x') \right); \quad G_{km}(x', x_s, \omega) = \frac{1}{4\pi \rho \nu_g^2 \sqrt{K}} \frac{1}{R} \exp\left(i \frac{\pi}{2} \sigma_0 + i\omega \nu_g^s \nu_g^r (x' - x_s) \right) \tag{89}\]

where $K = K_1 K_2$ is the Gaussian curvature of the slowness surface, $g_n$ and $g_l$ are two unit polarization vectors associated with $G_n^s$; $\nu_g^s$ and $\nu_g^r$ are the group velocity, and $\sigma_0 = 1 - \frac{3}{4} \pi m K_1 - \frac{3}{4} \pi m K_2$. Inserting the above asymptotic expression into *(88)* gives

\[\delta u_n = -\int_{\Omega} d\mathbf{x}' \nu_g^s \nu_g^r \omega^2 f_m A(x_s, x') A(x', x_s) g_n^s g_m^s (-g_l^s g_l^r \nu_g^s \nu_g^r \delta \rho + p_j^s g_l^s p_l^s \delta c_{ijkl}) \tag{90}\]

The radiation pattern is then given by

\[\Omega = -g_l^s g_l^r \nu_g^s \nu_g^r \delta \rho + p_j^s g_l^s p_l^s \delta c_{ijkl} \tag{91}\]

which only depends on the subsurface point perturbations and on the incoming (incident) and outgoing (diffracted) wave modes. It does not depend on the source type (force and moment) or the wave type recorded at the receiver (stresses or displacements).

Let us now consider 2D source-receiver geometry as shown in Fig.1. In transmission regime, the components of the unit slowness and the polarization vector will be the same for P-waves:

\[\mathbf{p}^s = (p_1^s, p_2^s) = (-\sin \theta, \cos \theta), \quad \mathbf{g}^s = (g_1^s, g_2^s) = (-\sin \theta, \cos \theta); \quad \mathbf{p}^r = (p_1^r, p_2^r) = (\sin \varphi, \cos \varphi), \quad \mathbf{g}^r = (g_1^r, g_2^r) = (\sin \varphi, \cos \varphi). \tag{92}\]
Then the coefficients in the front of the radiation pattern for \( c_{ij} \) are

\[
\begin{align*}
\Omega(c_{11}) &= p_{11}^i g_{11}^i g_{11}^r = \sin^2 \theta \sin^2 \varphi \\
\Omega(c_{33}) &= p_{33}^i g_{33}^i g_{33}^r = \cos^2 \theta \cos^2 \varphi \\
\Omega(c_{13}) &= p_{11}^i g_{13}^i g_{13}^r + p_{33}^i g_{33}^i g_{13}^r = \sin^2 \theta \cos^2 \varphi + \cos^2 \theta \sin^2 \varphi \\
\Omega(c_{55}) &= p_{11}^i g_{11}^i g_{55}^r + p_{33}^i g_{33}^i g_{55}^r + p_{13}^i g_{13}^i g_{55}^r + p_{31}^i g_{31}^i g_{55}^r = \sin 2\theta \sin 2\varphi
\end{align*}
\] (93)

For SV-waves, the unit slowness vector are orthogonal to the polarization vector:

\[
\begin{align*}
p^s &= (p_{11}^s, p_{33}^s) = (-\sin \theta, \cos \theta), \quad g^s = (g_{11}^s, g_{33}^s) = (\cos \theta, \sin \theta); \\
p^r &= (p_{11}^r, p_{33}^r) = (\sin \varphi, \cos \varphi), \quad g^r = (g_{11}^r, g_{33}^r) = (-\cos \varphi, \sin \varphi).
\end{align*}
\] (94)

Then the coefficients in the front of the radiation pattern for \( c_{ij} \) are

\[
\begin{align*}
\Omega(c_{11}) &= p_{11}^i g_{11}^i g_{11}^r = \sin \theta \cos \theta \sin \varphi \cos \phi \\
\Omega(c_{33}) &= p_{33}^i g_{33}^i g_{33}^r = \sin \theta \cos \theta \sin \varphi \cos \phi \\
\Omega(c_{13}) &= p_{11}^i g_{13}^i g_{33}^r + p_{33}^i g_{31}^i g_{13}^r = -2 \sin \theta \cos \theta \cos \varphi \sin \varphi \\
\Omega(c_{55}) &= p_{11}^i g_{11}^i g_{55}^r + p_{33}^i g_{33}^i g_{55}^r + p_{13}^i g_{13}^i g_{55}^r + p_{31}^i g_{31}^i g_{55}^r = \sin^2 \varphi - \cos^2 \varphi
\end{align*}
\] (95)

The parameterization \((\rho, V_h, V_p, V_s, V_{nmo})\) by

\[
c_{11} = \rho V_h^2; c_{33} = \rho V_p^2; c_{13} = \rho \sqrt{(V_p^2 - V_s^2)(V_{nmo}^2 - V_s^2)} - \rho V_s^2, c_{55} = \rho V_s^2
\] (96)
yields (under constant \(V_p/V_s\) ratio assumption in homogeneous isotropic medium, \(V_p = V_h = V_{nmo} = 2V_s\))

\[
\begin{align*}
\delta c_{11} &= \delta \rho + 2\rho V_h \delta V_h; \\
\delta c_{33} &= \delta \rho + 2\rho V_p \delta V_p; \\
\delta c_{13} &= \delta \rho + \rho V_p (\delta V_p + \delta V_{nmo} - \delta V_s); \\
\delta c_{55} &= \delta \rho + 2\rho V_s \delta V_s.
\end{align*}
\] (97)

leading to the following radiation patterns

\[
\begin{align*}
\Omega(V_p) &= 2\rho V_p \Omega(c_{33}) + \rho V_p \Omega(c_{13}) \\
\Omega(V_h) &= 2\rho V_h \Omega(c_{11}) \\
\Omega(V_{nmo}) &= \rho V_p \Omega(c_{13}) \\
\Omega(V_s) &= 2\rho V_s \Omega(c_{55}) - \rho V_p \Omega(c_{13}) \\
\Omega(\rho') &= \Omega(\rho) + \Omega(c_{11}) + \Omega(c_{13}) + \Omega(c_{33}) + \Omega(c_{55})
\end{align*}
\] (98)
References


