Notes on the meaning of adjoint wave equation derived via Lagrange multiplier approach

Zhinan Xie\(^1\), Wanjun Ma\(^1\), Yonglu Zheng\(^1\) and Xubin Zhang\(^1\)

\(^1\) Institute of Engineering Mechanic, China Earthquake Administration, Harbin, Helongjiang, 150080, China
\(^\ast\) E-mail: wla_2012@163.com

Abstract. Full-wave adjoint inversion method can be formulated via the Lagrange multiplier approach, taking the forward seismic wave equation as an additional independent constraint on synthetic wavefield. The adjoint wave equation is explained as the Lagrange multiplier where the local minimum existed, in case the model parameters are unperturbed. Such an explanation is unsatisfactory, due to the fact that the variation of synthetic wavefield equals to zero in case the model parameters are unperturbed. To avoid such a contradiction, we give a more clear explanation about the meaning of the adjoint wave equation, that it serves for converting the gradient of misfit function with respect to synthetic wavefield to the gradient of misfit function with respect to the model parameter.

1. Introduction
Since the fundamental work of Tarantola (1984), full-wave adjoint inversion method has gained great success in seismic tomography (Tarantola 1987; Dahlen et al. 2000; Tape et al. 2009; Zhu & Tromp 2013), ocean tomography (Bennett 2005; Hermand et al. 2006; Wunsch & Heimbach 2007), optimal design (Bendsøe & Sigmund 2004), meteorology (Talagrand & Courtier 1987; Errico 1997) and et al. It provides an efficient, accurate as well as a (more) general and robust way to infer the real medium properties by iteratively minimizing the least square difference between the synthetic data and the real data. The whole process of its application can be summarized as following, (i) obtain the forward wavefield (synthetics) by solving a model that governs the propagation of the waves given a suitable source; (ii) chose a misfit function to quantitatively describe the difference between the measurement and simulated wavefield on receiver arrays; (iii) simulate the adjoint wavefield by re-using the same model in forward simulation albeit with a modified source, which is located at the receivers and determined solely by the choice of misfit function; (iv) combine the synthetic and adjoint wavefields by temporal convolution to establish the gradient (Fréchet derivative or sensitivity) or/and the Hessian of the misfit function with respect to material properties and/or the location of free and/or interface boundary; (v) feed gradient information into a suitable global or local gradient(or/and Hessian)-based optimization technique to update material and/or the location of free and/or interface boundary that improve the misfit function; (vi) repeat the five steps above iteratively until the minimum of misfit function has been reached. The role of the gradients is to provide the mapping from the misfit function to the place of the material properties or the location of free and/or interface boundary needed to be changed. The main advantages of adjoint inversion method are as the following: (i) its efficiency. Since the evaluation of the gradient or/and Hessian of misfit function is independent of the type or the number of medium properties; (ii) its accuracy and robustness. It allows the change of misfit functions during the inversion process, thus the trap in local extrema could be avoided (Luo et al. 2014).
Typically, in seismic tomography, two approaches have been used for adjoint inversion method formulation. They are the variational approach and the Lagrange multiplier approach. Elaborate description of the variational approach can be found in Tromp et al. (2005) together with its relationship with other inversion methods, such as the time reversal inversion method. According to our knowledge, Liu & Tromp (2006) firstly proposed the Lagrange multiplier approach for adjoint inversion method formulation, taking the forward seismic wave equation as an additional independent constraint on synthetic wavefield. The adjoint wave equation is explained as the Lagrange multiplier where the local minimum existed, in case the model parameters are unperturbed. Such an explanation is unsatisfactory, due to the fact that the variation of synthetic displacement equals to zero when the model parameters are unperturbed. In this note, we give a clear explanation about the meaning of the adjoint wave equation, that it serves for converting the gradient of misfit function with respect to synthetic field to the gradient of misfit function with respect to the model parameter. For complete, in following section, we first give the set of wave equation governing the seismic wave propagation in infinite-domain. Then, we discuss the Lagrange multiplier approach in detail to explain of the meaning of the adjoint wave equation.

2. Seismic wave propagation in infinite-domain

In elastic solid medium, the seismic wave propagation is governed by the second-order elastic wave equation in displacement

\[ \rho \partial^2_t u = \nabla \cdot \sigma + f, \]

where \( u \) is the displacement vector, \( \sigma \) is the symmetric second-order stress tensor, \( \rho \) is the mass density, and \( f \) is an external force representing the seismic source. The ‘\( \nabla \)’ is gradient operator, the ‘\( \cdot \)’ represents dot-product operation, the scalar product of the gradient operator with a tensor field represents its divergence, the ‘\( : \)’ represents a double tensorial contraction operation, and a dot over a symbol represents its derivative with respect to time. The model parameters are described by constitutive tensor \( c \) and the density \( \rho \). The constitutive tensor \( c \) is symmetric, with minor and major symmetries, and retains positive definiteness. In the isotropic case the component of \( c \) can be written as

\[ c_{ijkl} = \left( \kappa - \frac{2}{3} \mu \right) \delta_{ij} \delta_{kl} + \mu \left( \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} \right) = \lambda \delta_{ij} \delta_{kl} + \mu \left( \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} \right); i, j, k, l = 1, 2, 3, \]

where \( \kappa \) is bulk modulus, \( \mu \) is shear modulus and \( \lambda \) is Lamé’s first parameter. The initial disturbance is assumed to be zero

\[ u(x, 0) = 0; \quad \partial_t u(x, 0) = 0. \]

On surface of earth, denoted by \( \Theta \), the traction equals to zero

\[ \sigma(x, t) \cdot n = 0, \]

where \( n = n(x) \) denotes the spatially dependent outward surface normal. In order to simulate the seismic wave propagation in regional domain, artificial boundaries (denoted as \( \Gamma \)) are needed. We use the one-way first-order Stacey boundary condition to simulate the wave transmission along the artificial boundaries, which can be written as

\[ \sigma(x, t) \cdot n = \rho \left[ c_p n n + c_s (I - n n) \right] \cdot \partial_n u = B \cdot \partial_t u, \]

where \( n \) is the unit outward normal \( n \) to the artificial boundary, \( c_p \) the P wave velocity, \( c_s \) the S wave velocity, \( I \) is the unit matrix.

3. Formulation of adjoint inversion method

In this section, we show the derivation of adjoint wave equation with a simple waveform misfit
function. It is defined as differences between three-component waveform data \( \mathbf{d}(x_r, t) \) recorded at \( N \) stations \( x_r, r = 1, \ldots, N \), and the corresponding synthetics displacement \( \mathbf{u}(x_r, t, \mathbf{m}) \) for a given model vector \( \mathbf{m} \). Using least square norm to characterize the differences, the waveform misfit function may be written as

\[
\chi(\mathbf{m}) = \frac{1}{2} \sum_{r=1}^{N} \int_{0}^{T} \left\| \mathbf{u}(x_r, t, \mathbf{m}) - \mathbf{d}(x_r, t) \right\|^2 dt, \quad (6)
\]

where the M-dimensional vector \( \mathbf{m} \) characterizes the current model. In practice, both the data \( \mathbf{d} \) and the synthetics \( \mathbf{u} \) will be windowed, filtered, and possibly weighted on time interval \([0, T]\).

The purpose of inversion is to find the point \( \mathbf{m} \) in model space, based on which the value of the misfit function is the smallest. In seismic tomography the model vector \( \mathbf{m} \) is long, and therefore it is impractical and generally not feasible with limited computational resources, to perform a brute-force calculation of the gradient of misfit function, which can be written as

\[
\delta\chi(\mathbf{m}) = \sum_{r=1}^{N} \int_{0}^{T} \left[ \mathbf{u}(x_r, t, \mathbf{m}) - \mathbf{d}(x_r, t) \right] \cdot \delta\mathbf{u}(x_r, t, \mathbf{m}) dt = \sum_{r=1}^{N} \int_{0}^{T} \left[ \mathbf{u}(x_r, t, \mathbf{m}) - \mathbf{d}(x_r, t) \right] \cdot \nabla \cdot \delta\mathbf{m} dt. \quad (7)
\]

The main task in formulation of adjoint inversion methods is to give the explicit and efficient formula for gradient computation. Typically, it can be done based upon variation approach and the Lagrange multiplier approach. The latter approach is more convenient and easier for understanding. In Lagrange multiplier approach, instead of seeking to minimize the misfit function, we seek to minimize the augmented Lagrange misfit function

\[
\chi_L = \frac{1}{2} \sum_{r=1}^{N} \int_{0}^{T} \left\| \mathbf{u}(x_r, t) - \mathbf{d}(x_r, t) \right\|^2 dt - \int_{0}^{T} \int_{\Omega} \lambda \cdot \left( \rho \nabla^2 \mathbf{u} - \nabla \cdot \mathbf{f} \right) d^3x dt, \quad (8)
\]

where the synthetic displacement field \( \mathbf{u} \) satisfying the seismic wave equation is taken as a constraint and the vector Lagrange multiplier \( \lambda \) remains to be determined. It is noteworthy here that the constraint is trivial

\[
-\int_{0}^{T} \int_{\Omega} \lambda \cdot \left( \rho \nabla^2 \mathbf{u} - \nabla \cdot \mathbf{f} \right) d^3x dt = 0. \quad (9)
\]

Since the seismic wave equation is satisfied by \( \mathbf{u} \) everywhere in interior of the whole domain \( \Omega \), which implicate

\[
\chi_L = \chi, \quad (10)
\]

for all solution \( \mathbf{u} \) to seismic wave equation. Taking the variation of (8), we obtain

\[
\delta\chi_L = \int_{0}^{T} \int_{\Omega} \sum_{r=1}^{N} \left[ \mathbf{u}(x_r, t) - \mathbf{d}(x_r, t) \right] \cdot \delta \mathbf{D}(x - x_r) \cdot \mathbf{u}(x_r, t) d^3x dt
\]

\[
- \int_{0}^{T} \int_{\Omega} \lambda \cdot \left( \delta \rho \nabla^2 \mathbf{u} - \nabla \cdot (\delta \mathbf{c} : \nabla \mathbf{u}) - \delta \mathbf{f} \right) d^3x dt - \int_{0}^{T} \int_{\Omega} \lambda \cdot \left( \rho \nabla^2 \delta \mathbf{u} - \nabla \cdot (\mathbf{c} : \nabla \delta \mathbf{u}) \right) d^3x dt.
\]

where we also assume that the known sources are close enough to represent the real seismic sources, which implicate \( \delta \mathbf{f} = 0 \). Moreover, variational terms associated with \( \delta \lambda \) equal to zero since \( \mathbf{u} \) satisfies seismic wave equation in \( \Omega \). Upon integrating the terms involving spatial and temporal derivatives of both \( \mathbf{u} \) and the variation \( \delta \mathbf{u} \) by parts, we obtain after some algebra.
\[
\delta \chi = \int_0^T \int_{\Omega} \sum_{r=1}^N [\mathbf{u}(\mathbf{x}_r,t) - \mathbf{d}(\mathbf{x}_r,t)] \mathbf{D}(\mathbf{x} - \mathbf{x}_r) \cdot \delta \mathbf{u}(\mathbf{x},t) d^3x dt \\
- \int_0^T \int_{\Omega} (\delta \rho \lambda \cdot \nabla \lambda : \delta \mathbf{c}) d^3x dt \\
- \int_0^T \int_{\Omega} (\rho \delta \lambda \cdot \nabla \cdot (\mathbf{c} \cdot \nabla \lambda)) \cdot \delta \mathbf{u} d^3x dt \quad - \int_0^T \int_{\Omega} (\rho(\lambda \cdot \partial_t \mathbf{u} - \partial_t \lambda \cdot \delta \mathbf{u}))_t d^3x \\
+ \int_0^T \int_{\Omega} (\lambda \cdot [\mathbf{n} \cdot (\delta \mathbf{c} \cdot \nabla \mathbf{u} + \mathbf{c} \cdot \nabla \delta \mathbf{u})] - \mathbf{n} \cdot (\mathbf{c} \cdot \nabla \lambda) \cdot \delta \mathbf{u} d^3 \Omega dt \\
\]

where \( \delta \Omega = \Theta \cup \Gamma \), notation \([f]_0^T\) means \(f(T) - f(0)\), for any function \(f\). Perturbing the free-surface boundary condition implies \(\mathbf{n} \cdot (\delta \mathbf{c} \cdot \nabla \mathbf{u} + \mathbf{c} \cdot \nabla \delta \mathbf{u}) = 0\) on \(\delta \Omega\); perturbing the initial condition implies that \(\delta \mathbf{u}(\mathbf{x},0) = 0\) and \(\partial_t \delta \mathbf{u}(\mathbf{x},0) = 0\); perturbing the ABC implies that \(\mathbf{n} \cdot (\delta \mathbf{c} \cdot \nabla \mathbf{u} + \mathbf{c} \cdot \nabla \delta \mathbf{u}) = \partial \mathbf{B} \cdot \partial_t \delta \mathbf{u} + \partial \mathbf{B} \cdot \partial_t \delta \mathbf{u}\). Thus, we obtain

\[
\delta \chi_L = \int_0^T \int_{\Omega} \sum_{r=1}^N [\mathbf{u}(\mathbf{x}_r,t) - \mathbf{d}(\mathbf{x}_r,t)] \mathbf{D}(\mathbf{x} - \mathbf{x}_r) \cdot \delta \mathbf{u}(\mathbf{x},t) d^3x dt \\
- \int_0^T \int_{\Omega} (\delta \rho \lambda \cdot \nabla \lambda : \delta \mathbf{c}) d^3x dt \\
- \int_0^T \int_{\Omega} (\rho \delta \lambda \cdot \nabla \cdot (\mathbf{c} \cdot \nabla \lambda)) \cdot \delta \mathbf{u} d^3x dt \quad - \int_0^T \int_{\Omega} (\rho(\lambda \cdot \partial_t \mathbf{u} - \partial_t \lambda \cdot \delta \mathbf{u}))_t d^3x \\
- \int_0^T \int_{\Omega} (\lambda \cdot [\mathbf{n} \cdot (\mathbf{c} \cdot \nabla \lambda) \cdot \delta \mathbf{u} + \mathbf{B} \cdot \partial_t \lambda \cdot \delta \mathbf{u})]_t d^3 \Omega dt + \int_1 \frac{\partial \mathbf{B} \cdot \delta \mathbf{u}}{\partial \mathbf{B} \cdot \partial_t \delta \mathbf{u}} d\Gamma \\
\]

where the notation \([f]_0^T\) means \(f(T)\).

According to Liu & Tromp (2006), in the absence of perturbations in the model parameters \(\delta \rho\), \(\delta \mathbf{c}\), and \(\delta \mathbf{f}\), the local minimum existed condition of \(\chi_L\) deduces to the variation in the action is stationary with respect to perturbations \(\delta \mathbf{u}\). Thus, the Lagrange multiplier \(\lambda\) should satisfy the equation

\[
\rho \delta \lambda \cdot \nabla \cdot (\mathbf{c} \cdot \nabla \lambda) = \sum_{r=1}^N [\mathbf{u}(\mathbf{x}_r, T-t) - \mathbf{d}(\mathbf{x}_r, T-t)] \delta(\mathbf{x} - \mathbf{x}_r),
\]

subjecting to the adjoint free boundary condition

\[
\mathbf{n} \cdot \mathbf{c} \cdot \nabla \lambda = 0 \quad \text{on} \ \Theta,
\]

together with adjoint ABC

\[
\mathbf{n} \cdot \mathbf{c} \cdot \nabla \lambda = \mathbf{B} \cdot \partial_t \mathbf{u} \quad \text{on} \ \Gamma,
\]

and the end conditions

\[
\lambda(\mathbf{x}, T) = 0, \quad \partial_t \lambda(\mathbf{x}, T) = 0.
\]

Moreover, introduce the adjoint wave field \(\mathbf{u}_\lambda^*(\mathbf{x}, T-t) = \lambda(\mathbf{x}, T-t)\). The variation of \(\chi_L\) equals to

\[
\delta \chi_L = \int_\Omega (\delta \rho \lambda^*_\rho + \delta \mathbf{c} : \mathbf{K}_c) d\Omega,
\]

where \(\mathbf{K}_c = -\int_0^T \mathbf{u}_\lambda^*(\mathbf{x}, T-t) \cdot \partial_t \mathbf{u}(\mathbf{x},t) dt\) and \(\mathbf{K}_c = -\int_0^T \nabla \mathbf{u}_\lambda^*(\mathbf{x}, T-t) \cdot \nabla \mathbf{u}(\mathbf{x},t) dt\).

The adjoint wave field \(\mathbf{u}_\lambda^*\) is determined by exactly the same set of wave equation, boundary conditions, and initial boundary conditions as the forward wavefield \(\mathbf{u}\) with the exception of the
source term: the forward wavefield is determined by the source, whereas the adjoint wave field is generated by using the time-reversed differences between the synthetics $u$ and the data $d$ at the receivers as simultaneous sources.

However, it is not satisfactory that the derivation of adjoint wave field, given by Liu & Tromp (2006), based on the local minimum existed condition in absence of perturbations in the model parameters $\delta \rho$, $\delta c$, and $\delta f$. Since in absence of $\delta \rho$, $\delta c$ and $\delta f$, $\delta u$ is also equal to zero, which can be easily found via Born approximation. Keeping in mind that, in $\Omega$ we have

$$\rho \frac{\partial^2 u}{\partial t^2} - \nabla \cdot (c \nabla u) = f,$$

and correspondingly

$$\delta \rho \frac{\partial^2 u}{\partial t^2} - \nabla \cdot (\delta c \nabla u) - \delta f = \rho \frac{\partial^2 \delta u}{\partial t^2} - \nabla \cdot (c \delta \nabla u).$$

Combining the seismic wave equation and (20), we can easily find the last two integration terms in (11) equal to zero, also the last three integration terms in (12) together with the last six integration terms in (13). For any choice of $\lambda$, the variation of Lagrange misfit function is always equal to the variation of the original misfit function, while the particular choice of $\lambda$ to satisfy the adjoint wave equation is only due to the conversion of the gradient of misfit function with respect to synthetic field to the gradient of misfit function with respect to the model parameter perturbation, i.e.,

$$\delta \lambda = \int_0^T \int_\Omega \sum_{r=1}^N \left[ u(x_r,t) - d(x_r,t) \right] \delta u(x_r,t) \cdot \delta c \cdot \nabla u \cdot d^3x dt = \int_0^T \int_\Omega \left( \delta \rho \lambda \cdot \frac{\partial^2 u}{\partial t^2} - \nabla \lambda \cdot \delta c \cdot \nabla u \right) d^3x dt .$$

With the latter expression of the variation of the misfit function, we are able to iteratively update the model parameter to minimize the misfit function. Furthermore, it is noteworthy here that the point in model space, corresponding to the local minimum of misfit function, can be explicitly found after solving the linear system arose in

$$-\int_0^T \int_\Omega \left( \delta \rho \lambda \cdot \frac{\partial^2 u}{\partial t^2} - \nabla \lambda \cdot \delta c \cdot \nabla u \right) d^3x dt = 0 .$$

Based on that, the corresponding point in model space is $m + (\delta \rho \lambda \cdot \delta c)$. However in practice, such a linear system may be difficult to solve due to the generally huge size of the system.

4. Conclusions and future work

In this note, we give a clear explanation of the Lagrange multiplier approach used for adjoint inversion formulation that it serves for converting the gradient of misfit function with respect to synthetic wavefield to the gradient of misfit function with respect to the model parameter perturbation. Though we do not improve the adjoint inversion method, we believe that the clear understanding could help understanding the latest improvement of adjoint inversion methods with consideration of the inversion simultaneous in data and model space.

Acknowledgments

This study was funded by the National Key R&D Program of China (No. 2018YFC1504004), National Natural Science Foundation of China (No.51678539). The authors also thank Professor Liao Zhenpeng for his continuous support.

References
[1] Bendsøe, M. P. Sigmund, O. (2004). Topology optimization: theory, methods and applications. Springer-Verlag, 2nd edn, New York.
[2] Bennett, A. F. (2005). Inverse modeling of the ocean and atmosphere. Cambridge University Press, New York.
[3] Dahlen, F. A., Hung, S. H., & Nolet, G. (2000). Fréchet kernels for finite-frequency traveltimes—I. Theory. Geophysical Journal International, 141(1), 157-174.
[4] Errico, R. M. (1997). What is an adjoint model? Bulletin of the American Meteorological Society, 78(11):2577-2592.
[5] Hermand, J. P., Meyer, M., Asch, M., Berrada, M. (2006). Adjoint-based acoustic inversion for the physical characterization of a shallow water environment. The Journal of the Acoustical Society of America, 119(6):3860-3871.
[6] Liu, Q., & Tromp, J. (2006). Finite-frequency kernels based on adjoint methods. Bulletin of the Seismological Society of America, 96(6): 2383-2397.
[7] Talagrand, O., Courtier, P. (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. Quarterly Journal of the Royal Meteorological Society, 113(478):1311-1328.
[8] Tape, C., Liu, Q., Maggi, A., Tromp, J. (2009). Adjoint tomography of the southern California crust. Science, 325(5943):988-992.
[9] Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. Geophysics, 49(8):1259-1266.
[10] Tarantola, A., (1987). Inverse problem theory: methods for data fitting and model parameter estimation, Elsevier Science Publishers, Amsterdam, Netherlands.
[11] Tromp, J., Tape, C., Liu, Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. Geophysical Journal International, 160(1):195-216.
[12] Wunsch, C., & Heimbach, P. (2007). Practical global oceanic state estimation. Physica D: Nonlinear Phenomena, 230(1-2):197-208.
[13] Zhu, H., Tromp, J. (2013). Mapping tectonic deformation in the crust and upper mantle beneath Europe and the North Atlantic Ocean. Science, 341(6148):871-875.