CHAPTER 2

General Theory

2.1 General Theory of Full Waveform Inversion

Full waveform inversion (FWI) is a method that uses the recorded seismic data to estimate parameters of subsurface model (Yong Ma, 2010), such as seismic wave velocity by solving the inverse problem. It begins from the simple assumption that the wave equation can be solved numerically of the form below.

$$F(m) = p \tag{2.1}$$

where m is parameter describing the subsurface model, p is seismic wave field, and F is the function that describes how to calculate a seismic wave field from the given model. FWI is an algorithmic approach that uses the repeated application of forward problem was expressed in equation (2.1) to solve the non-linear inverse problem that can be expressed as

$$F^{-1}(d) = m' \tag{2.2}$$

where d contains the observed field data, m' is an estimated subsurface model, and the inverse of F function in equation (2.2) expresses the idea that if the model m' is placed back to equation (2.1), then the predicted data p' should be compatible with the original observed data d.

Unfortunately, most of the geophysical problems are non-linear. To solve that inverse problem, it is usually formulated as a least-squares optimization, which computes a model that minimises the difference between observed data and generated data from initial model or misfit function which can be expressed as follow.

$$E = \frac{1}{2}\sum_s (d_{obs} - F(m))^2 \tag{2.3}$$

where E is a misfit function, d_{obs} is the observed data, $F(m)$ is the generated data from model m, and s is number of data.

For solving an inverse problem, if the non-linear between the data and model is not too strong, it could be possible to linearise the problem and it can be solved iteratively by local optimisation methods. Then the new model (m) will be described by the linear combination of the initial or current model (m_0) and perturbed model or model correction (δm).

$$m = m_0 + \delta m \tag{2.4}$$

The problem is now to find a model correction (δm) to generate a new model which reduces the difference between observed dataset and calculated data from the current model or misfit function toward zero.

2.2 Gradient of misfit function

Solving a non-linear problem, gradient descent can also be used to solve a system of nonlinear equations and, it is more practical way to calculate the perturbation model (δm_{k}) by minimising the misfit function. The method begins with current model (m_{k}), then uses the gradient of the misfit function to evaluate the current model and estimate the perturbation model (δm_{k}), in equation (2.4) that reduce the data misfit function.

The gradient of misfit function represents the direction in which the misfit function is increasing most rapidly or the steepest direction. The misfit function can always be reduced by pursuing the negative of this direction (Pratt et al., 1998). Therefore, the estimation of the new model in equation (2.4) to reduce the misfit function by iteratively updating can be written as follow

$$m_{k+1} = m_k - \alpha_k \nabla_m E_k \tag{2.5}$$

where $\nabla_m E_k$ is the gradient of misfit function, α_k is the constant scaling for iteration number k. This scaling is chosen to indicate that how far the model can be updated to minimise the misfit function in the direction given by the gradient of misfit function. The role of the scaling can be also thought of as converting the units of the gradient vector to the model dimensions (Pratt et al., 1998).

The gradient of the misfit function ($\nabla_m E$) can be obtained by taking the first derivative of the misfit function, described in equation (2.3) with respect to each element of model m_i. Subsequently the gradient of misfit function should be:

$$\nabla_m E = \frac{\partial \vec{E}}{\partial m} = \frac{\partial}{\partial m}\left(\frac{1}{2}\sum(d_{obs} - F(m))^2\right) \tag{2.6}$$

$$\nabla_m E = -\sum \frac{\partial F(m)}{\partial m}(d_{obs} - F(m)) \tag{2.7}$$

$$\nabla_m E = -\sum \frac{\partial F(m)}{\partial m}\delta d \tag{2.8}$$

where $\frac{\partial F(m)}{\partial m}$ is the partial derivative of the data with respect to the model parameter, and $\delta d = \left(d_{obs} - F(m)\right)$ represents the residual data or the difference between the observed data and generated data from the model. Hence, the gradient of misfit function can be interpreted as a product of time correlation between the partial derivative wave field and the residual data in time domain. This process is aimed to pick missing information on the initial model and uses it to calculate the perturbation model from the optimisation algorithm. (Operto et al., 2013)

2.3 Partial derivative wave field

To calculate the gradient of misfit function, the partial derivative wave fields were required and used to detect the missing information on the initial model from true model. It begins with solving the acoustic wave equation. The full wave field can be calculated at any place and time in the current velocity model. The 2D acoustic wave equation can be written as follows.

$$\frac{\partial^2 P(x,z,\omega)}{\partial x^2} + \frac{\partial^2 P(x,z,\omega)}{\partial z^2} + \frac{\omega^2}{c^2}P(x,z,\omega) = S(x,z,\omega) \tag{2.9}$$

where ω is the angular frequency, c is the velocity, $P(x,z,\omega)$ is the pressure field and $S(x,z,\omega)$ is the source function. Then, this equation can be written in the following simple matrix form (Shin et al., 2001)

$$\vec{W}\vec{P} = \vec{S} \tag{2.10}$$

where $\vec{W}$ is a complex impedance matrix ($\vec{W} = \frac{\omega^2}{c^2} + \nabla^2$ where ∇^2 is the Laplacian operator), $\vec{P}$ is the pressure field and $\vec{S}$ is the source function. Taking the partial derivative of previous equation with respect to the model parameter m_i on both sides, the result will be equation (2.9) and because source function is independent of model parameter, then it becomes zero. (Shin et al., 2001)

$$\overrightarrow{W}\frac{\partial \vec{P}}{\partial m_i} + \vec{P}\frac{\partial \overrightarrow{W}}{\partial m_i} = 0 \tag{2.11}$$

or

$$\overrightarrow{W}\frac{\partial \vec{P}}{\partial m_i} = -\vec{P}\frac{\partial \overrightarrow{W}}{\partial m_i} \tag{2.12}$$

This equation can be rearranged to:

$$\overrightarrow{W}\frac{\partial \vec{P}}{\partial m_i} = \vec{f}^{(i)} \tag{2.13}$$

Where $\vec{f}^{(i)} = -\vec{P}\frac{\partial \overrightarrow{W}}{\partial m_i}$ is a virtual source term at location of ith parameter (Shin et al., 2001), which represents the interaction (or scattering) of a predicted wave field $\vec{P}$ with model m_i as shown in Figure 2.1 (Pratt et al., 1998). Therefore, $\frac{\partial \vec{P}}{\partial m_i}$ was referred as the partial derivative wave field from location ith, which can be described as follow

$$\frac{\partial \vec{P}}{\partial m_i} = \overrightarrow{W}^{-1}\vec{f}^{(i)} \tag{2.14}$$

From the above equation, the partial derivative wave field with respect to the velocity model can be interpreted as the wave field emitted at surface source, scattered by diffraction point located at m_i and recorded by surface receivers. (Operto et al., 2013)

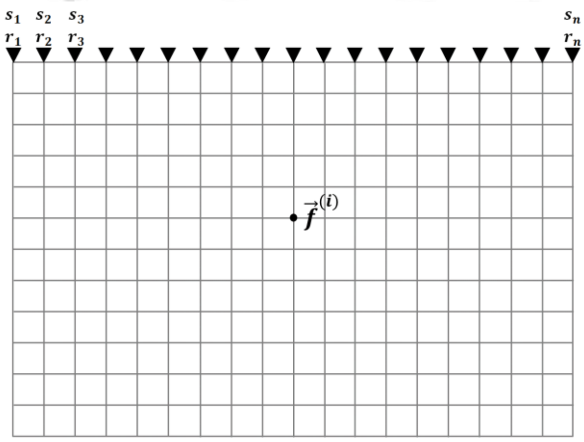

Figure 2.1: $N_x \times N_z$ grid of earth model with n sources and receivers distributed along surface, $\vec{f}^{(i)}$ indicates a virtual source according to node i^{th} model parameter (Shin et al., 2001)

2.4 Finite difference modelling for acoustic wave

In this study, the synthetic datasets were generated based on a constant density solution to the wave equation using Matlab code in AFD (Acoustic Finite Difference) package developed by CREWES (CREWES, 2015). This algorithm solves the wave equation over a discrete set of grid points or model elements using the central finite difference schemes to approximate the second derivative to the scalar wave equation. (Youzwishen et al, 1999)

$$\frac{\partial^2 \emptyset(x,z,t)}{\partial t^2} = v^2(x,z)\nabla^2\emptyset(x,z,t) \tag{2.15}$$

where $\emptyset$ is a wave function, v is a velocity and the Laplacian, ∇^2 is given by

$$\nabla^2\emptyset = \frac{\partial^2\emptyset}{\partial x^2} + \frac{\partial^2\emptyset}{\partial z^2} \tag{2.16}$$

The Laplacian operator and the second order time derivative in scalar wave equation can be estimated by the finite difference schemes using 5 points of grid respectively for the second order approximation, as shown in Figure 2.2. Therefore, the Laplacian operator and time derivative can be expressed as follow. (Youzwishen et al, 1999)

$$\nabla^2\emptyset_j^n \approx \frac{\emptyset_j^{n+1} - 2\emptyset_j^n + \emptyset_j^{n-1}}{\Delta x^2} + \frac{\emptyset_{j+1}^n - 2\emptyset_j^n + \emptyset_{j-1}^n}{\Delta z^2} \tag{2.17}$$

$$\frac{\partial^2\emptyset(t)}{\partial t^2} \approx \frac{\emptyset(t+\Delta t) - 2\emptyset(t) + \emptyset(t-\Delta t)}{\Delta t^2} \tag{2.18}$$

where n is the x coordinate and j is the z coordinate of the grid, as shown in Figure 2.2.

By substituting the approximation of Laplacian operator in (2.17) and time derivative in (2.18) into the scalar wave equation in (2.15), the estimation of the wave field at time $t + \Delta t$ can be solved as:

$$\emptyset_j^n(t+\Delta t) \approx \left(\Delta t^2\left({v_j^n}^2\right)\nabla^2 + 2\right)\emptyset_j^n(t) - \emptyset_j^n(t-\Delta t) \tag{2.19}$$

Equation (2.19) demonstrates that the wave field at time $t + \Delta t$ can be estimated by knowing the wave field at time t and $t - \Delta t$. This process is called time stepping or snapshot. (Youzwishen et al, 1999). Figure 2.3 illustrates a workflow for estimating the wave field at time $t + \Delta t$ time using stepping finaite difference in equation (2.19).

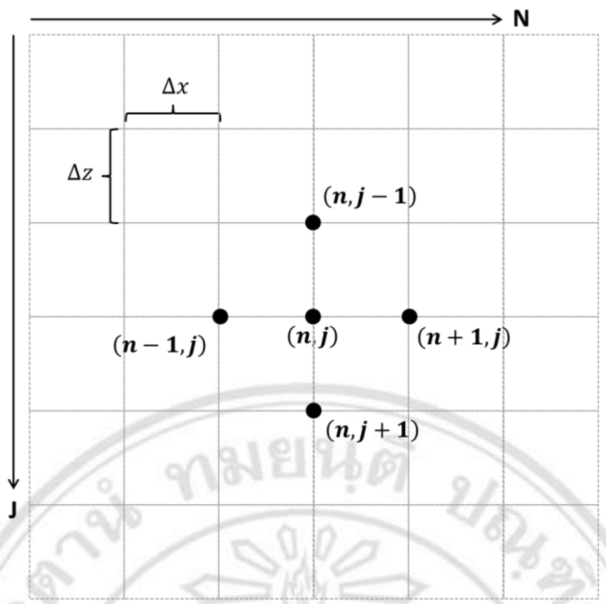

Figure 2.2: The computational grid of the second order finite difference approximation.

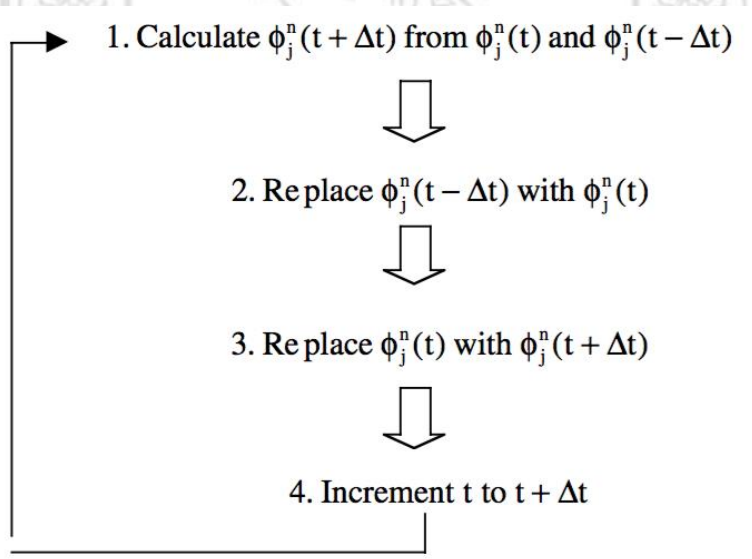

Figure 2.3: Time stepping Finite difference workflow (Youzwishen et al, 1999).