CHAPTER 2

Theory and Methods

2.1 Seismic Inversion Fundamentals

If we consider any physical system, the formulation of a forward problem would be predicting the result of a totally controllable experiment knowing the system characteristics. This can be mathematically expressed by the following equation:

$$d = Gm \tag{2-1}$$

where d is a set of observed measures, m is a model containing the system characteristics and G is a Kernel matrix connecting model parameters and observations.

In seismic data, calculating a synthetic seismogram is considered a forward problem as density and sonic data are parameters of the system that are used to predict the ideal seismic response (Russell, 1988). Unfortunately, due to many factors – errors in well log measurements, unknown wavelet, inaccurate migration, etc. – this prediction is never perfect. Still, forward problems are normally used to determine if a model fits a set of observations, as a manner of validating the assumptions made (Mosegaard and Tarantola, 2002). The inverse problem, on the other hand, is formulated when the system is a black box, but a set of observations is available. Thus, inferring the system parameters from a set of indirect measures is known as inversion. Besides the model (Earth properties), the source wavelet applied to the seismic experiment is also unknown. Seismic inversion attempts to infer both: wavelet effects need to be removed (deconvolved) from seismic data in order to obtain a reflectivity series; and the reflection series is a link to Acoustic Impedance (AI), which then can be used to estimate petrophysical properties of a reservoir. Many factors make the inverse problem underdetermined, i.e. *non-unique* (Jackson, 1972). For instance, the limited bandwidth of the source energy prevents us from resolving thin layers. Intrinsic complications regarding the experiment itself (noise, etc.) end up admitting many models to give satisfactory matches to the seismic traces.

Summing up, seismic inversion tries to remove the wavelet effect from the seismic trace in order to estimate the reflectivity series, and thus obtain the acoustic impedance. Properties of the geological layers cannot be inferred directly from seismic sections by any conventional seismic interpretation (Francis, 2005). Seismic Inversion, on the other hand, is able to increase the resolution of the seismic data and allows us to estimate rock properties by replacing the seismic signature with responses that correspond to acoustic impedance (Pendrel, 2006; Simm and Bacon, 2014).

Primarily, seismic inversion can be of two types: post-stack and pre-stack. Pre-stack Inversion methods consider the variations of amplitude with angle (AVA) through inverting simultaneously different angle gathers, while considering wavelet variation according to the angle range. But due to the nature of the present project, emphasis will be given to post-stack methods as the seismic data is available only as a full-stack volume. Lastly, this chapter will introduce the theoretical background for seismic inversion, including a brief explanation of the main methods to invert seismic data and different wavelet extraction approaches, until – finally – we present the mathematical formulation and virtues of the stochastic Gabor inversion (Naghadeh *et al.*, 2017) used in the current study.

2.1.1 The Convolutional Model

Every seismic experiment theoretically begins with the injection of a large amount of acoustic energy into the Earth's subsurface while the ongoings are monitored at the surface. This energy propagates as a wavefront through the system, which promptly responds to the impulse according to some basic physical laws. Every time the source impulse meets a boundary with a contrast in Acoustic Impedance, part of the energy is reflected back to the surface (Veeken, 2007). Acoustic Impedance (Z) is the product of bulk density and P-velocity of a particular stratum (Eq. 2-3). For rays hitting the interface perpendicularly, the amount of energy that is reflected when the source wavelet hits these boundaries with different AI is determined by the reflection series (r), as follows:

$$r_n = \frac{Z_n - Z_{n-1}}{Z_n + Z_{n-1}} \tag{2-2}$$

where,

$$Z = \rho V \tag{2-3}$$

and r_n stands for the reflection series of the nth layer, with layer $n-1$ immediately above layer n.

There also is an anelastic attenuation factor (absorption) that modifies the waveform of the source wavelet with depth. Thus, the wavelet in the recorded seismogram will vary with time; i.e., it is non-stationary (Margrave *et al.*, 2003). The convolutional model states that a seismic trace is the result of a convolution of the source wavelet with the reflectivity series, and can be expressed by:

$$s(t) = w(t) * r(t) + n(t) \tag{2-4}$$

where $*$ represents the convolution operator, $s(t)$ is the recorded seismogram, $w(t)$ is the propagation source wavelet, $r(t)$ is the reflectivity series (Rc) and $n(t)$ is the inherent noise associated with the experiment.

But the convolutional model (Figure 2.1) works under a series of assumptions, including source wavelet stationarity (Yilmaz, 1987). If all the assumptions were valid and the source wavelet was infinitesimally short, i.e. containing all frequencies, the response to the seismic experiment would be the reflection series itself, or in other words, the response just from the geological medium. In practice, however, the source wavelet is band-limited and consequently the seismic data is band-limited as well (Tarantola, 1987). Deconvolution is commonly used to remove source wavelet effects from the seismic trace and consequently increases resolution. If stationary reflectivity inversion methods are used to obtain reflection series, the result is just a rough approximation due to inaccurate assumptions (Naghadeh and Morley, 2016a). Thus, in order to approximate the inversion results from the real reflection coefficient function (geological signature), it is important to consider the seismic signal as non-stationary.

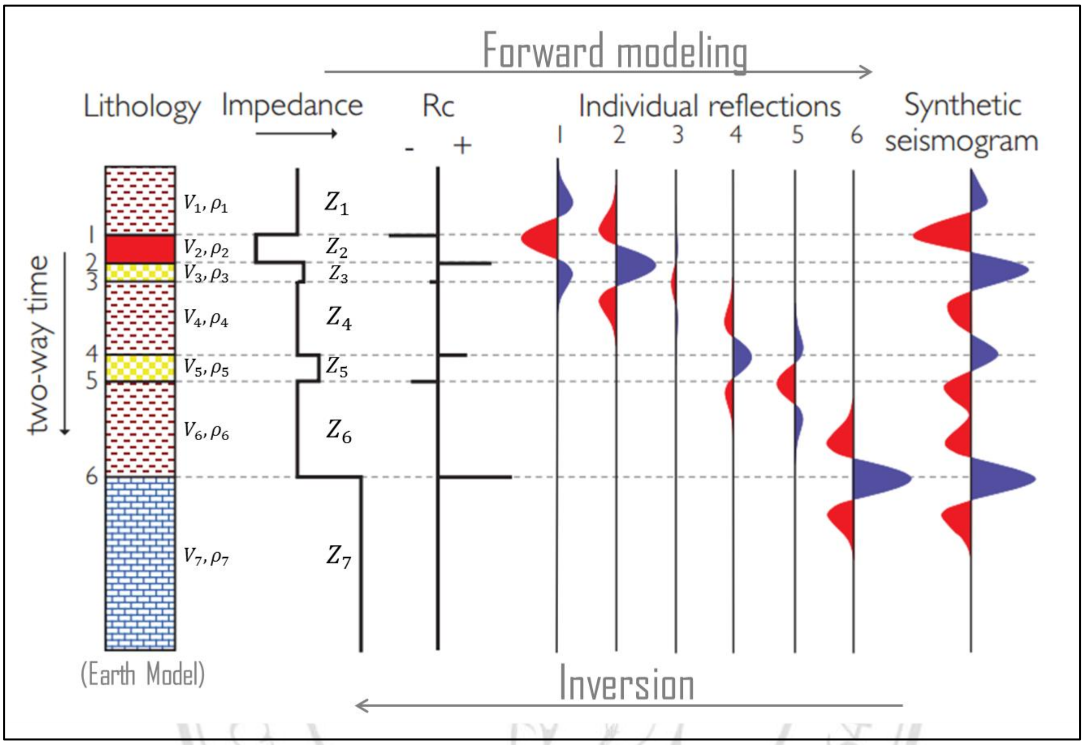

Figure 2.1: The convolutional model showing the path of forward modeling (model to data) and the inverse problem (data to model). A synthetic seismogram is created from a 1D model (well log data) through the convolution of a RC with the source wavelet (a zero-phase wavelet in this example, modified from Simm and Bacon, 2014)

2.2 Seismic Inversion Methods

2.2.1 Deterministic Inversion Methods

Seismic inversion can be of two types, deterministic and stochastic. Deterministic inversion algorithms try to minimize the difference between a modelled seismic trace and the actual seismic trace. One example is the recursive inversion, which seeks to obtain absolute impedance values by inverting Equation 2.2 (Russell, 1988):

$$Z_{n+1} = Z_n \left[\frac{1+r_n}{1-r_n}\right] \tag{2-5}$$

This approach involves scaling the seismic section to reflectivity and then adding a low-frequency model coming from the well logs or from the stacking velocities. The wavelet is not addressed by the recursive inversion, but it illustrates well the idea of deterministic methods (Simm and Bacon, 2014).

Deterministic inversion can be thought to be the convolutional model in reverse. Deconvolution is applied to the seismic trace, removing the wavelet effect and leaving the reflectivity series. There are numerous deterministic methods, but probably the most popular is known as model-based inversion (Russell and Hampson, 1991). One of its assumptions is that the wavelet is known. Well data (or a velocity model) is used to build an initial low-frequency model that is iteratively updated and checked against the seismic data. The aim is to modify the starting model until a minimum error between the synthetic and the seismic is obtained (Cooke and Schneider, 1983). The iterative comparison procedure used by the model-based method is a mathematical effort to minimize this function:

$$J(R) = \lambda_1 \|S - WR\|_2^2 + \lambda_2 \|M - HR\|_2^2 \tag{2-6}$$

where S is the seimic trace, W is the source wavelet, R is the final reflectivity, M the currently initial model, H the integration operator to change RC to AI and $*$ is the convolution operator. Note how the first term models the seismic trace and the second updates the low-frequency model. In deterministic inversion, λ_2 assumes a value of zero where the calculated impedance is within a range accepted by the constraints.

In order to guide the algorithm to an acceptable minimised-error solution, the model-based method makes use of constraints. It prevents the solutions from reaching ranges outside of geologically possible impedance values, for example. Thus, the solution is an impedance trace which respects the constraints whilst minimizing the error between the synthetic and the seismic data. Due to the non-unique character of inverse problems, there are always several impedance model solutions that provide an admissible match (Backus and Gilbert, 1970). A source of concern regards the algorithm used, or more specifically, if the algorithm finds the lowest possible synthetic error or stops its iterations after an acceptable error is achieved. The ideal would be to adopt an optimized algorithm that does not stop iterating at a so-called local minimum solution, but that searches for the global minimum, i.e., the lowest possible error instead (Simm and Bacon, 2014).

2.2.2 Stochastic Inversion Methods

A stochastic approach to the inverse problem can be thought initially to be from the same formulation introduced for the model-based deterministic inversion (Equation 2-6). The sum of the weighting factors (λ_1, λ_2) is equal to one ($\lambda_1 + \lambda_2 = 1$). Thus, assuming that:

$$R \approx 0.5\, \Delta \ln(AI) \tag{2-7}$$

and,

$$\begin{cases} \Delta = D \\ \ln(AI) = L_z \end{cases} \tag{2-8}$$

where D is the differential matrix and L_z is the logarithm of the Acoustic Impedance. This assumption is valid for RC's ≤ 0.3, which is normally the case (Velis, 2008).

Substituting Equation 2-8 into Equation 2-6, it is now:

$$J(L_z) = \lambda_1 \|S - 0.5WDL_z\|_2^2 + \lambda_2 \|M - 0.5HDL_z\|_2^2 \tag{2-9}$$

The first term (misfit term) minimizes the difference between the seismic and synthetic traces, while the second term forces a solution to honor the initial model (model penalty function). Clearly, the values assigned to the the weighting factors (λ_1, λ_2) control the importance given to one term or the other. It is highly subjective and will depend on the quality of the data. With $\lambda_2 = 0$, the objective function becomes the constrained model-based inversion and "hard" constraints are used, i.e. limited to a maximum fractional change in impedance from the initial-model values. If $\lambda_2 \neq 0$ the objective function leads the solution to be stochastic and bounded by "soft" constraints, i.e. the initial low-frequency model is considered as a separate component which is added to the seismic trace (Doyen, 2007).

In order to minimize the objective function, the following expression is established:

$$\frac{\partial J}{\partial L_z} = -\lambda_1 W^T D^T (S - 0.5WDL_z) - \lambda_2 H^T D^T (M - 0.5HDL_z) \qquad (2\text{-}10)$$

$\therefore$

$$0 = -\lambda_1 W^T D^T S + 0.5\lambda_1 W^T D^T W D L_z - \lambda_2 H^T D^T M + 0.5\lambda_2 H^T D^T H D L_z$$

(2-11)

to finally get L_z :

$$L_z = \frac{\lambda_1 W^T D^T S + \lambda_2 H^T D^T M}{0.5\lambda_1 W^T D^T W D + 0.5\lambda_2 H^T D^T H D}$$

(2-12)

This approach allows the stochastic inversion to include a low-frequency content coming typically from the well logs. It also makes it possible to add high-frequency content (Doyen, 2007). These high frequencies – which are outside of the seismic bandwidth – come from variogram modeling using the well data (Figure 2.2). Thus, close to wells, high resolution can be reasonably inferred and away from the wells the absence of a simplicity term along with the statistical conditioning still makes it possible to achieve resolution beyond deterministic inversion methods (Bosch et al., 2010).

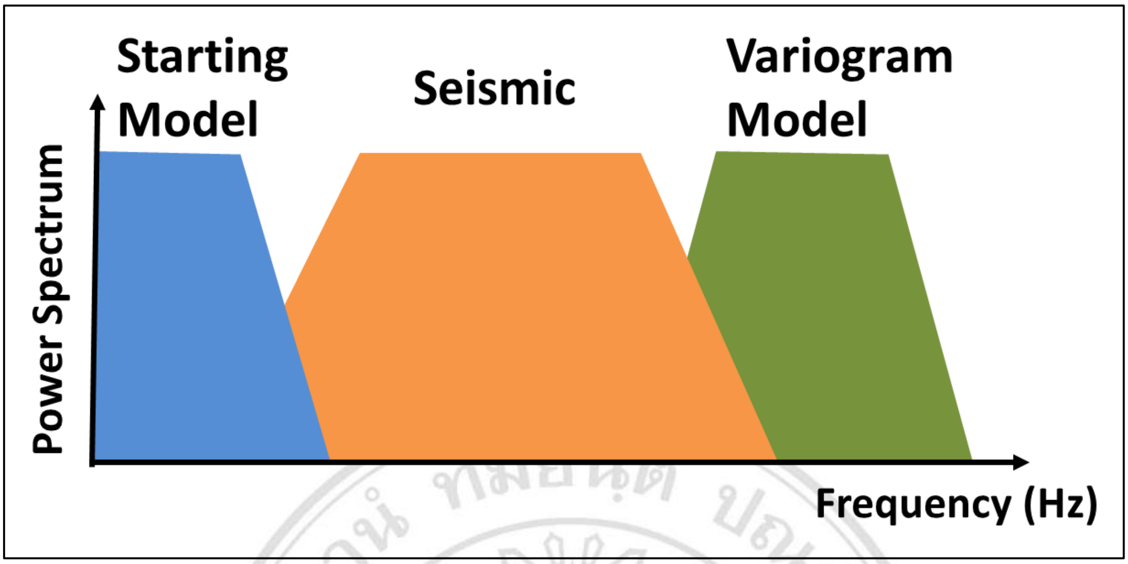

Figure 2.2: Schematic illustrating the bandwidth of a stochastic inversion result. A starting model adds low-frequency information. This low-frequency model comes from filtered well data, which is interpolated between well control points; seismic bandwidth determine the intermediate frequency content; and a vertical variogram model controls the high-frequency end of the spectrum (Doyen, 2007)

The wavelet is vital in the inversion process. In order to correctly estimate the reflectivity series from seismic amplitudes, so that the acoustic impedance values can be obtained, it is necessary to understand the wavelet. An assumption of the methods mentioned thus far is to consider that the wavelet is known (see section 2.2.1). Moreover, they assume the wavelet is stationary (Yilmaz, 1987), which – as discussed in section 2.1.1 – is not exactly a realistic assumption. Results from methods with such assumptions are consequently a rough approximation of the real Earth signature (i.e. the RC). So, to increase resolution by effectively removing the wavelet effect from seismic data, it is important to consider the non-stationary nature of the seismic signal (Naghadeh *et al.*, 2017).

In the early 2000's, Margrave *et al.* (2003) introduced a deconvolution method (known as Gabor Deconvolution) that compensates for attenuation and frequency dispersion effects, or in other words, accounts for non-stationarity. A Gabor transform confines a signal in a specific time range using a window function (Gaussian window) and then a Fourier Transform is applied to extract the signal's time-variable spectrum (Naghadeh and Morley, 2016b). The method splits the signal into multiple windows and uses a Hilbert transform of the logarithm of its amplitude spectrum to estimate the source wavelet properties. In the end, the non-stationary deconvolution is enabled by "mapping" the changes in wavelet properties (amplitude and phase) at each of the analyzed windows, i.e. analogous to its changes with time. These non-stationary source wavelet properties can then be included in a kernel matrix to be deconvolved with the seismic data allowing for amplitude compensation, zero-phase and temporal resolution enhancement.

2.2.2.1 Stochastic Gabor Inversion

Naghadeh *et al.* (2017) incorporated the Gabor deconvolution method into a stochastic inversion framework. This method promises to completely remove the source wavelet effects by extracting the time-variant wavelet properties, even from noisy datasets. The estimated reflectivity will then present higher resolution and the bias incorporated by the well-log information (low-frequency and variogram model) will lead the stochastic Gabor inversion to accurately estimate acoustic impedance values. The mathematical formulation to the stochastic Gabor method is summarized below. For a complete description, please refer to Naghadeh *et al.* (2017).

In order to estimate the time-variant properties of the source wavelet, Gabor deconvolution in the frequency domain is applied:

$$G_d(\tau,\omega) = \int_{-\infty}^{+\infty} d(t)g(t-\tau)e^{-2i\omega t}dt \tag{2-13}$$

where $G_d(\tau,\omega)$ is the Gabor transform of signal $d(t)$, windowed by $g(t)$, which is a Gaussian function centered at location τ. Angular frequency is represented by ω.

After splitting the signal into multiple windows, a time-frequency analysis is carried out at every single window until, eventually, the whole signal is analyzed. Amplitude and phase wavelet spectra at each step are obtained by:

$$A_W(\tau,\omega) = S_w \times abs\left(\int_{-\infty}^{+\infty} d(t)g(t-\tau)e^{-2i\omega t}dt\right) \tag{2-14}$$

$$Ph_W(\tau,\omega) = \text{Hilbert}\big(\log\big(A_W(\tau,\omega)\big)\big)$$

(2-15)

where S_w is the smoothing window, $Ph_W(\tau,\omega)$ is the phase spectrum and Hilbert is the Hilbert transform. In the frequency domain, for any time range of signal, the wavelet properties can be estimated by the following equation:

$$W_a(\tau,\omega) = A_W(\tau,\omega).\exp\big(sqrt(-1).Ph_W(\tau,\omega)\big)$$

(2-16)

The deconvolved signal in the frequency domain will be:

$$Dec_{signal} = \frac{G_d(\tau,\omega)}{W_a(\tau,\omega)}$$

(2-17)

And in time-domain:

$$Dec_{signal} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} Dec_{signal}(\tau,\omega)\chi(t-\tau)\,e^{-2i\omega t}\,dt\,d\omega$$

(2-17)

where $\chi(t-\tau)$ is a synthesis window. Gabor deconvolution can also be applied in the time domain, so to get the wavelet properties for each windowed signal in time:

$$W_a(t) = \text{IFFT}\big(W_a(\tau,\omega)\big)$$

(2-18)

where $W_a(\tau,\omega)$ is the extracted wavelet in a special time range and IFFT is the inverse of a fast Fourier Transform. Once one possesses the wavelet (with its properties) for different time ranges, a kernel matrix can be created to perform the non-stationary deconvolution in the time-domain:

$$d(t) = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}, \text{Kernel}_{Matrix} = \begin{bmatrix} W_{11} & \cdots & W_{1n} \\ \vdots & \ddots & \vdots \\ W_{n1} & \cdots & W_{nn} \end{bmatrix}$$

(2-19)

$$Dec_{signal}(t) = \text{Kernel}_{Matrix}{}^{-1}.\, d(t)$$
(2-20)

Due to the instability of the inverse solution, a regularization parameter is required:

$$Dec_{signal}(t) = (\text{Kernel}_{Matrix} + \varepsilon I)^{-1}.\, d(t) \tag{2-21}$$

where I is the identity matrix and ε is a small positive number (pre-whitening).

Thus, the attempt is to obtain the RC from deconvolution as close as possible to R_r, which is the reference RC at the well location. Or, mathematically:

$$R_r - R = e_r \cong 0 \tag{2-22}$$

with e_r corresponding to the amount of error due to noise affecting the wavelet extraction. Ideally $e_r = 0$. So, similar to Equations 2.10 and 2.11, to get a stochastic reflectivity inversion:

$$\frac{\partial J}{\partial R} = 0 \Rightarrow R = \frac{\lambda_1 W^T S + \lambda_2 W^T S_r}{W^T W + \mu \times \text{diag}((\text{abs}(R_r) + \varepsilon)^{-1})}$$
(2-23)

where S_r is the product of the convolution of the reference RC at the well location with the non-stationary wavelet estimated by the Gabor deconvolution. μ and ε are small regularization parameters which will lead the inversion result to be sparse.

Equation 2-8 is used to replace R with AI, so that the stochastic Gabor inversion is obtained through the following expression:

$$L_z = \frac{\lambda_1 D^T W^T S + \lambda_2 D^T W^T S_r}{0.5 D^T W^T W D + \mu \times \text{diag}((\text{abs}(R_r) + \varepsilon)^{-1})}$$
(2-24)