Chapter 4

Deconvolution of reflectivity series

4.1 Modeling normal incidence seismograms

In this chapter we will study the problem of computing reflection and transmission coefficients for a layered media when plane waves propagate in the earth with angle of incidence \( i = 0 \) (Normal incidence). We will use these concepts to derive a model for normal incidence seismograms.

4.1.1 Normal incidence

Consider a plane wave impinging at angle of propagation \( i = 0 \) with respect to the normal (see Figure (4.1)). In this case we have three waves:

- Incident wave: \( \downarrow \) in medium 1
- Reflected wave: \( \uparrow \) in medium 1
- Transmitted wave: \( \downarrow \) in medium 2

Let us assume that the amount of incident wave is equal to 1, the amount of reflected wave is given by \( r \), and the amount of transmitted wave is denoted by \( t \). At the boundary the following condition should be satisfied (continuity of displacements)

\[ 1 + r = t \]
This equation has two unknowns, to compute the $r$ and $t$ we need an extra equation. We well consider conservation of energy. In the acoustic (vertical incidence case) conservation of energy leads to the following equation:

$$I_1 \times 1 + I_1 \times r^2 = I_2 \times t^2.$$  

The quantities $I_1$ and $I_2$ are called **acoustic impedances**

$$I_1 = \rho_1 v_1$$  
$$I_2 = \rho_2 v_2$$

where $\rho_1$ and $\rho_2$ are the densities of the material above and below the interface and $v_1$ and $v_2$ the P-velocities, respectively. After combining the equations of continuity of displacement and conservation of energy we obtain the following expressions

$$r = \frac{I_2 - I_1}{I_2 + I_1} \text{Reflection coefficient} \quad (4.1)$$  
$$t = \frac{2I_1}{I_2 + I_1} \text{Transmission coefficient} \quad (4.2)$$

The above analysis is valid for an incident plane wave propagating downwards. Let’s consider the case of an incident wave propagating upwards (Figure (4.2)).

- Incident wave: $\uparrow$ in medium 2
- Reflected wave: $\downarrow$ in medium 2
- Transmitted wave: $\uparrow$ in medium 1

In this case the reflection and transmission coefficients are given by

$$r' = \frac{I_1 - I_2}{I_2 + I_1} \quad (4.3)$$  
$$t' = \frac{2I_2}{I_2 + I_1} \quad (4.4)$$

From the above equations it is clear that

$$r' = -r \quad (4.5)$$
4.1.2 Impulse response

Let's assume that we run a zero offset experiment in a stratified earth composed of 4 layers plus a half-space of impedances given by $I_1$, $I_2$, $I_3$, $I_4$, $I_5$. (Figure (4.3)). At $t = 0$ a delta-like source emits energy into Earth. The energy is transmitted and reflected from the layers. If we do not consider multiples reflection, our sismogram will be composed of 4 arrivals (4 reflections).

To simplify the problem I will show how to compute the amplitude of the wave recorded at the surface of the earth generated (reflected) at the interface 4. First we have to compute the amount of amplitude transmitted to each one of the layers until reaching layer 4. This is given by the product of the transmission coefficients of each layer. In Figure (4.3) the transmission coefficients $t$ are replaces by their equivalent expression $(1 + r)$.

The amplitude of the wave when reaches the layer 4 is

Figure 4.1: P-wave normal incidence. The incident wave is propagating downwards.
CHAPTER 4. DECONVOLUTION OF REFLECTIVITY SERIES

1 + r' = t'  \[ r = -r' \]

Figure 4.2: P-wave normal incidence. The incident wave is propagating upwards.

\[
1 \times t_1 \times t_2 \times t_3 = (1 + r_1)(1 + r_2)(1 + r_3)
\]

does not depend on the wave's propagation direction.

when the wave is reflected in the layer 4 the total amplitude at that point (last expression) needs to be multiplied by the reflection coefficient of interface 4,

\[
1 \times t_1 \times t_2 \times t_3 \times r_4 = (1 + r_1)(1 + r_2)(1 + r_3)t_4
\]

Note that now the wave (reflected wave) is propagating upwards, therefore, the transmission coefficients are given by terms of the form

\[
1 + r' = 1 - r
\]

The final amplitude after propagating the wave to the surface of the Earth (this is what
4.1. MODELING NORMAL INCIDENCE SEISMOGRAMS

The final expression for the amplitude of the wave reflected in the interface 4 can be written down as follows

\[
\frac{(1 + r_1)(1 + r_2)(1 + r_3)}{1 - r_3} \times \frac{r_4}{(1 - r_1)(1 - r_2)(1 - r_3)}
\]

Figure 4.3: Amplitude of a wave plane wave propagating in a layered medium. Analysis of the wave reflected in the interface 4.
CHAPTER 4. DECONVOLUTION OF REFLECTIVITY SERIES

\[(1 - r_1^2)(1 - r_2^2)(1 - r_3^2)r_4.\]

It is clear that reflections occur at all the layers:

Amplitude of the reflection generated at intergace 1

\[A_1 = r_1\]

Amplitude of the reflection generated at intergace 2

\[A_2 = (1 - r_1^2)r_2\]

Amplitude of the reflection generated at intergace 3

\[A_3 = (1 - r_1^2)(1 - r_2^2)r_3\]

Amplitude of the reflection generated at intergace 4

\[A_4 = (1 - r_1^2)(1 - r_2^2)(1 - r_3^2)r_4\]

We can write a general expression for the amplitude of a reflection generated at the \(k\)-th interface:

\[A_1 = r_1\]

\[A_k = \prod_{i=1}^{k-1} (1 - r_i^2)r_k \quad k = 2, 3, 4, \ldots\]

How to interpret these results? If we assume that the earth is excited with a delta function, and neglecting the presence of multiples, our zero-offset seismogram will be a collection of delta functions (spikes) at arrival times given by the two-way travel time formula. The strength of each arrival will be proportional to the amplitude \(A_k\).

However, in real exploration seismology, it is impossible to have a source that resembles a delta function. The source signature is called a wavelet. This is a finite length time function that we will denote as \(w(t)\). In this case, the seismogram is represented by a superposition of wavelets arriving at different times and with amplitude proportional to \(A_k\).

In our model with 4 interfaces (Figure (4.3)) we will have 4 arrivals of amplitude \(A_1, A_2, A_3\) and \(A_4\). The seismogram can be expressed as follows
4.2. DECONVOLUTION OF REFLECTIVITY SERIES

\[ s(t) = A_1 w(t - t_1) + A_2 w(t - t_2) + A_3 w(t - t_3) + A_4 w(t - t_4) \]  
\[ (4.6) \]

where \( t_1, t_2, t_3 \) and \( t_4 \) are the arrival times of each reflection \(^1\)

Notice that if we neglect transmission effects, the amplitude \( A_i \) can be replaced by the reflection coefficient \( r_i \). In general we will assume an Earth model that consists of micro-layers. In this case we can write the seismic trace model a convolution between two time series: a wavelet and the reflectivity sequence:

\[ s_n = w_n \ast q_n. \]  
\[ (4.7) \]

4.2 Deconvolution of reflectivity series

So far we have discuss the problem of designing a deconvolution operator for a seismic wavelet. We have also examined a “toy” example, the minimum phase dipole (Chapter 2).

In general, the convolutional model is a very well accepted model to describe a seismic trace. In this model we say that the seismic trace (in general a zero-offset trace) can be written down as a convolution of two signals: a seismic wavelet (this is the source function) and the reflectivity series.

The reflectivity series is our “geological” unknown. In fact, the reflectivity is a sequence of spikes (reflectors) that indicates the position (in time) of layers in the subsurface, the strength or amplitude of each spike is an indicator of how much energy is reflected back to the receivers during the seismic experiment. Let’s write the seismogram as a simple convolution between a wavelet \( w_n \) and a reflectivity sequence \( q_n \):

\[ s_n = w_n \ast q_n. \]  
\[ (4.8) \]

In this simple model we have neglected the noise, in general we will assume that deterministic noise (multiples and ground roll) has been attenuated and therefore what is left is random noise:

\[ s_n = w_n \ast q_n + n_n \]  
\[ (4.9) \]

\(^1\)Notice that \( w(t - \tau) \) is \( w(t) \) after being delayed \( \tau \) seconds.
It is clear from the above equation that one has a problem with one equation (one observable) and two unknowns (the wavelet and the reflectivity). Therefore, the seismic deconvolution problem involves the solution of two subproblems:

- **Wavelet Estimation**
- **Operator design**

By wavelet estimation we refer to methods to estimate the seismic source from the seismic trace. In general these methods are statistical techniques that explode some properties of the remaining unknown (the reflectivity). We do have deterministic techniques based on the wave equation to estimate the seismic source in the marine case. These methods are beyond the scope of this course.

### 4.2.1 The autocorrelation sequence and the white reflectivity assumption

We have seen that the design of a Wiener filter involves the inversion of an autocorrelation matrix with Toeplitz structure. This matrix arises from the fact that we have represented our convolution model as a matrix times vector multiplication. To clarify the problem, let us assume that we have a 3 point wavelet and we compute the autocorrelation matrix. We first write down the convolution matrix:

\[
C = \begin{pmatrix}
w_0 & 0 \\
w_1 & w_0 \\
w_2 & w_1 \\
0 & w_2
\end{pmatrix}.
\]  

(4.10)

The autocorrelation matrix is given by

\[
R = C^T C = \begin{pmatrix}
r_0 & r_1 \\
r_1 & r_0
\end{pmatrix}.
\]  

(4.11)

Now we can try to write the autocorrelation coefficients in terms of the sample of the wavelet \( w_n \), in this case we have:

\[2\text{This is the matrix you would have used to design a 2 points spiking or Wiener filter}\]
The first coefficient is the zero-lag correlation coefficient, this is also a measure of the energy of the wavelet. The second coefficient \( r_1^w \) is the first lag of the correlation sequence.

The correlation coefficients can be written using the following expression:

\[
r_j^w = \sum_k w_k w_{k+j}, \quad j = 0, \pm 1, \pm 2 \ldots .
\] (4.14)

In the Wiener filter the matrix \( R \) is an \( N \times N \) matrix where \( N \) is the length of the filter, in this case we will need to compute \( N \) autocorrelation coefficients:

\[
r_j^w, j = 0, N - 1 .
\]

In order to design a Wiener or spiking filter, we first need to know the wavelet. Unfortunately, the seismic wavelet is unknown. To solve this problem we use the white reflectivity assumption. Under this assumption the seismic reflectivity (the “geology”) is considered a zero mean white process (Robinson and Treitel, 1980).

A zero-mean white process is an uncorrelated process, in other words if \( r_j^q \) is the autocorrelation function of the reflectivity, then

\[
r_j^q = \begin{cases} P_q & j = 0 \\ 0 & j = \pm 1, \pm 2, \pm 3, \ldots \end{cases}.
\] (4.15)

The autocorrelation is a measure of similarity of a time series with itself. The zero lag coefficient measure the power of the signal \( P_q \), the first coefficient \( (j = 1) \) measures the similarity of the signal with a one-sample shifted version of itself.

If the reflectivity is a zero-mean white noise process, the following remarkable property is true:

---

\(^3\)Please, note that the supra-script \( w \) is used to stress that this is the autocorrelation of the wavelet
In other words: the autocorrelation function of the trace is an estimate (within a scale factor) of the autocorrelation of the wavelet. It is clear that now we can estimate the autocorrelation of the wavelet from the autocorrelation of our observable: the seismic trace.

We have managed to compute the autocorrelation function of the wavelet, but what about the wavelet. It turns out that the Z-transform of the autocorrelation sequence of the wavelet can be used to compute the seismic wavelet. In this case we need to make a new assumption, we will assume that the wavelet is a minimum phase wavelet. In general, this is a good assumption to deal with sources generated by explosions (dynamite).

It is easy to show that the z transform of the autocorrelation sequence can be decomposed as follows:

\[ R_w(z) = \sum_j r^w_j z^{-j} = W(z) W(z^{-1}) \]  

(this is for a real wavelet). In this case, the autocorrelation function provides information about the wavelet but cannot define the phase of the wavelet. After factorizing the above equation, one can select the zeros that lie outside the unit circle (the minimum phase dipoles!!). In this way we can recover a wavelet with minimum phase features consistent with a spectrum given by \( R_w(z) \).

The estimation of a minimum phase wavelet from the autocorrelation function is often referred as the spectral factorization problem. It can be solved using different techniques, in particular the Hilbert Transform provides a fast algorithm to compute the minimum phase wavelet.

### 4.2.2 What to do with the noise?

We start with our noisy seismogram:

\[ s_n = w_n * q_n + n_n \]  

The goal of the deconvolution process is to recover \( q_n \) from the data, \( s_n \). In order to achieve this objective, a filter \( f_k \) must be computed such that \( f_k * w_k = \delta_k \). Generally,
of course, we can only compute an estimate of the filter \( \hat{f}_k \) and \( \hat{f}_k \ast w_k = a_k \), where \( a_k \) is called the averaging function and resembles a delta function only in the ideal case. Applying \( \hat{f}_k \) to both sides of equation (4.18), yields the estimated output of the deconvolution process

\[
\hat{q}_k = a_k \ast q_k + \hat{f}_k \ast n_k \\
= q_k + (a_k - \delta_k) \ast q_k + \hat{f}_k \ast n_k .
\]  

(4.19)

Since our main requirement is to estimate a reliable model \( \hat{q}_t \) which is close to the true reflectivity, it is important to design a filter such that the error terms in equation (4.19) are as small as possible. Or in other words, one seeks a solution with the following properties:

\[
a_k = w_k \ast f_k \approx \delta_k ,
\]  

(4.20)

and

\[
f_k \ast n_k \approx 0
\]  

(4.21)

The last two expression can also be written in matrix form

\[
C_w f \approx d
\]  

(4.22)

and

\[
C_n f \approx 0
\]  

(4.23)

where \( C_w \) and \( C_n \) denote the convolution matrices for the wavelet and the noise, respectively. Both equations are honored when we minimize the following objective function:

\[
J = ||C_w f - d||^2 + \beta ||C_n f||^2 ,
\]  

(4.24)

where \( \beta \) is a tradeoff parameter. The second term in the last equation can be written as
where the matrix $\mathbf{C}_n^T \mathbf{C}_n$ is the noise autocorrelation matrix. If the noise is uncorrelated, we can replace $\mathbf{C}_n^T \mathbf{C}_n$ by its estimator

$$E[\mathbf{C}_n^T \mathbf{C}_n] = \sigma_n^2 \mathbf{I}.$$  

(4.26)

where $\sigma_n^2$ is an estimate of the variance of the noise. Now the objective function $J$ is given by,

$$J = ||\mathbf{C}_w f - d||^2 + \mu ||f||^2,$$

(4.27)

where $\mu = \sigma_n^2 \times \beta$. This is the objective function used to design an inverse filter, and the solution given by

$$\mathbf{f} = (\mathbf{R}_w + \mu \mathbf{I})^{-1} \mathbf{C}_w^T \mathbf{d}.$$  

(4.28)

In Figures (4.4), (4.5) and (4.6) we test the performance of the least squares inversion when dealing with noise free and noisy data. It is clear, that the pre-whitening parameter plays a key role in the deconvolution of noisy data.
4.2. DECONVOLUTION OF REFLECTIVITY SERIES

Figure 4.4: Deconvolution of a "clean" seismogram.

Figure 4.5: Deconvolution of a "noisy" seismogram. The tradeoff parameter is too small; the result is too unstable.
Figure 4.6: Deconvolution of a "noisy" seismogram. The tradeoff has been increased to stabilize the solution.
4.2. DECONVOLUTION OF REFLECTIVITY SERIES

4.2.3 Deconvolution in the frequency domain

A procedure similar to the one outlined in the previous section can be used to deconvolved data in the frequency domain. Taking the discrete Fourier Transform of equation (4.19) yields

\[ \hat{Q}_k = Q_k + (A_k - 1)Q_k + \hat{F}_k N_k. \]  

(4.29)

Since \( a_k \) should be a good approximation to a delta function, it turns out that the filter should be designed satisfying the following requirement

\[ W_k F_k = A_k \approx 1 \quad \forall k. \]  

(4.30)

Furthermore, in order to maintain the noise at a small level

\[ F_k N_k \approx 0 \quad \forall k. \]  

(4.31)

We can combine these two requirements into a single one. In order to achieve this, let us construct the following objective or cost function

\[ J = \sum_k |A_k - 1|^2 + \alpha \sum_k |F_k N_k|^2. \]  

(4.32)

Minimizing the objective function with respect to the filter coefficients leads to

\[ \hat{F}_k = \frac{W^*_k}{|W_k|^2 + \alpha |N_k|^2}. \]  

(4.33)

Finally, the reflectivity estimate is given by

\[ \hat{Q}_k = D_k \frac{W^*_k}{|W_k|^2 + \alpha |N_k|^2} \]

\[ = D_k \frac{W^*_k}{|W_k|^2 + \mu}. \]  

(4.34)

Since the noise has a flat spectrum (\(|N_k|^2 = \sigma_n^2\)) we can replace \( \alpha |N_k|^2 \) by another constant \( \mu \). An estimate of the variance of the reflectivity estimate in the frequency domain is given by
CHAPTER 4. DECONVOLUTION OF REFLECTIVITY SERIES

\[ Var(\hat{Q}_k) = |\hat{F}_k|^2 \sigma_N^2. \] (4.35)

after a few manipulation we end up with

\[ Var(\hat{Q}_k) = \frac{|W_k|^2 \sigma_N^2}{(|W_k|^2 + \mu)^2}. \] (4.36)

When \( \mu = 0 \) the variance can be extremely high at the frequencies at which the wavelet power is small. Similarly we can find a expression for the norm of the reflectivity estimate in the frequency domain

\[ N = \frac{\sum_k |\hat{Q}_k|^2}{\sigma_N^2} \sum_k |S_k|^2 Var(\hat{Q}_k) \] (4.37)

The misfit function is

\[ \Phi = \sum_k |S_k - W_k \hat{Q}_k|^2 \]
\[ = \frac{1}{\sigma_N^2} \sum_k |S_k|^2 \frac{\mu}{(|W_k|^2 + \mu)^2} \] (4.38)

**Regularization error and noise magnification**

If \( E_k \) denotes the deviation of the filter from the true inverse filter, defined by

\[ E_k = 1 - \hat{F}_k W_k \] (4.39)

we can write equation (4.12) as follows

\[ \hat{Q}_k = \hat{F}_k W_k Q_k + \hat{F}_k N_k \]
\[ = (1 - E_k) Q_k + \frac{1-E_k}{W_k} N_k \] (4.40)

then, the difference between the true reflectivity \( Q_k \) and the reflectivity estimate \( \hat{Q}_k \) is given by
4.2. DECONVOLUTION OF REFLECTIVITY SERIES

\[ \hat{Q}_k - Q_k = \frac{E_k Q_k}{RE} + \frac{1 - E_k}{W_k NAE} \]  

(4.41)

where \( RE \) stands for regularization error and \( NAE \) for noise amplification error. The NAE is independent of the data and can be expressed as a function of the wavelet:

\[ \frac{W_k^*}{|W_k|^2 + \mu} \]

It is clear that the more the filter resembles the inverse the wavelet \( W_k^{-1} \), the larger this error will be. The RE introduce data-dependent degradation (i.e., ringing).
4.3 Sparse deconvolution and Bayesian analysis

The deconvolution operator is usually stabilized by adding a small perturbation to the diagonal of the autocorrelation matrix (Robinson and Treitel, 1980). The latter is equivalent to solving the problem by means of zero order quadratic regularization.

The regularization technique is used to estimate a unique and stable solution by introducing some type of prior information. In this part of the course, I will examine different regularization strategies that may be used to improve the deconvolution of seismic records. Specifically, I will use the Huber and the Cauchy criteria to retrieve a sparse reflectivity sequence. These criteria are related to "long tail" probability distributions (Huber, 1981) which, in a Bayesian context, are used as a prior distribution for the reflectivity sequence. The seismic wavelet is assumed to be known or to be well approximated by an estimator.

Re-weighting strategies have been used in conjunction with least squares type estimators to diminish the influence of outliers in inverse problems (Scales et al., 1988). In robust statistics the influence function (Huber, 1981) that measures the influence of the residuals on the estimators is constructed so as to attenuate outliers. A similar manipulation can be applied to the regularization function. In this case the goal is to attenuate the side-lobe artifacts which are introduced in the deconvolution process. In this context, the regularization strategy is used to drive the output of the deconvolution to a prescribed part of the model space, quite to the contrary to the classical application in robust statistics where the chief goal is to attenuate the influence of gross errors. When the problem is properly regularized the resolution of close seismic arrivals can be significantly enhanced. The described procedure is used to overcome the poor resolution associated to quadratic regularization strategies. We have to realize, however, that only when the seismogram is composed of a finite superposition of seismic wavelets (sparse reflectivity assumption) that these techniques may provide a substantial improvement with respect to conventional deconvolution.

4.3.1 Norms for sparse deconvolution

The deconvolution problem may be stated as follows. Consider

\[ s_k = \sum_j w_j q_{k-j}, \quad k = 1, n_y \]  \hspace{1cm} (4.42)

where \( q_k, \ k = 1, n_q \) and \( s_k, \ k = 1, n_s \) are the input and the output series to the convolu-
tion process, respectively. The series \( w_k, \ k = 1, n_w \) is the “blurring function” or source wavelet. In time domain deconvolution, the goal is to find \( \hat{x}_k \) such that the residuals \( \epsilon_k \) are minimized

\[
\epsilon_k = s_k - \sum_j w_j \hat{q}_{k-j}, \quad k = 1, n_s. \tag{4.43}
\]

In the least squares approach the following objective function of the residuals is minimized:

\[
J = \sum_k \rho_1 \left( \frac{\epsilon_k}{\sigma_k} \right), \tag{4.44}
\]

where

\[
\rho_1(u) = \frac{1}{2} u^2. \tag{4.45}
\]

The residuals are weighted according to the data accuracy that is given by the inverse of the standard error of each observation \( \sigma_k \). For simplicity we shall assume that \( \sigma_k = \sigma_n, \ k = 1, n_s \). The minimization of \( J \) is accomplished by solving the following system:

\[
\frac{\partial J}{\partial q_l} = \sum_k \psi \left( \frac{\epsilon}{\sigma_n} \right) w_{k-l} = 0, \quad l = 1, n_q \tag{4.46}
\]

where \( \psi(u) = \frac{d\rho(u)}{du} \). In robust statistics the function \( \psi \) is called the influence function. This function measures the influence of the residuals on the parameter estimators. The minimization of \( J \) leads to the following system of normal equations

\[
\sum_k \sum_j w_{k-j} w_{k-l} q_j = \sum_k w_{k-l} s_k, \tag{4.47}
\]

or in matrix form

\[
Rq = g. \tag{4.48}
\]
Equation (4.48) is stabilized by adding a small perturbation to the diagonal of the matrix \( R \). This is equivalent to minimize the following modified objective or cost function

\[
J = J + J_q,
\]

where \( J_q = \sum_k \rho_1 \left( \frac{q_k}{\sigma_q} \right) \). The term \( J_q \) is the regulariser of the problem. This particular procedure is called zero order quadratic regularization. The cost function \( J \) which I will denote \( J_{11} \) is given by

\[
J_{11} = \sum_k \rho_1 \left( \frac{\epsilon_k}{\sigma_n} \right) + \sum_i \rho_1 \left( \frac{q_i}{\sigma_q} \right). \tag{4.50}
\]

The minimum of equation (4.50) is reached at the point

\[
\hat{q} = (R + \mu I)^{-1} g, \tag{4.51}
\]

where \( \mu = \sigma_n^2 / \sigma_q^2 \) is the damping parameter of the problem. In filter theory, \( \mu \) is also called the pre-whitening parameter (Robinson and Treitel, 1980).

### 4.3.2 Modifying \( J_q \)

A standard procedure in robust statistics is based on redesigning the influence function in order to attenuate the influence of outliers. A similar modification can be used to design the regularization function, \( J_q \).

The data misfit is modeled using the functional \( \rho_1 \), while for the regularization term I will introduce the following modification (Huber, 1981)

\[
\rho_2(u) = \begin{cases} 
  u^2 / 2 & \text{if } |u| \leq a \\
  a |u| - a^2 / 2 & \text{if } |u| > a 
\end{cases}. \tag{4.52}
\]

The deconvolution problem is solved by minimizing the following cost function

\[
J_{12} = \sum_k \rho_1 \left( \frac{\epsilon_k}{\sigma_n} \right) + \sum_i \rho_2 \left( \frac{q_i}{\sigma_q} \right). \tag{4.53}
\]
The influence function for $\rho_2$ becomes:

$$
\psi_2(u) = \begin{cases} 
  u & \text{if } |u| \leq a \\
  a \, \text{sign}(u) & \text{if } |u| > a 
\end{cases}
$$

(4.54)

The function $\rho_2$ behaves identically to $\rho_1$ for small values of $u$. When $|u| > a$, $\rho_2$ defines a line and its associated influence function becomes a constant.

We can define another function with similar behavior

$$
\rho_3(u) = \ln\left(\frac{u^2}{2} + 1\right).
$$

(4.55)

When $u$ is small $\rho_3 \rightarrow \rho_1$. The influence function corresponding to $\rho_3$ is given by

$$
\psi_3(u) = \frac{u}{\frac{u^2}{2} + 1}.
$$

(4.56)

If $\rho_3$ is adopted the deconvolution problem is solved by minimizing a cost function designated by $J_{13}$

$$
J_{13} = \sum_k \rho_1\left(\frac{e_k}{\sigma_n}\right) + \sum_i \rho_3\left(\frac{q_i}{\sigma_q}\right).
$$

(4.57)

The cost function $J_{13}$ can be derived using Bayes’ rule by assuming Gaussian errors and a Cauchy prior probability to model the unknown parameters (Sacchi and Ulrych, 1995).

Figure (4.7) shows the functions $\rho_1$, $\rho_2$ and $\rho_3$. In Figure (4.8) the corresponding influence functions $\psi_1$, $\psi_2$ and $\psi_3$ are displayed. The functions $\rho_2$ and $\psi_2$ were calculated for $a = 1$ and $a = 2$. When the parameter $a$ is small compared to the normalized variable $u$ the width of the transition zone $-a < u < a$ becomes very narrow. In this case the cost function $\sum_i \rho_2(u_i)$ behaves like a $l_1$ norm, $l_1 = \sum_i |u_i|$.

### 4.3.3 Iterative solution

The solution to the least squares problem with zero order quadratic regularization expressed in equation (4.51) can be modified to introduce into the regularization the functionals $\rho_2$ (Huber criterion) and/or $\rho_3$ (Cauchy criterion). In this case the system of nor-
mal equations is obtained by equating to zero the gradient of the objective function, \( J_{12} \) or \( J_{13} \),

\[
(R + \mu Q)q = g. \tag{4.58}
\]

If the problem is regularized with \( \rho_2 \), the matrix \( Q \), which I will call \( Q_2 \), has the following diagonal elements

\[
Q_{2_{ii}} = \begin{cases} 
1 & \text{if } |\frac{q_i}{\sigma_q}| \leq a \\
\frac{a}{\sigma_q} & \text{if } |\frac{q_i}{\sigma_q}| > a
\end{cases} \tag{4.59}
\]

When \( \rho_3 \) is adopted, the matrix \( Q \) in equation (4.58), which I will denote \( Q_3 \), has the following diagonal elements

\[
Q_{3_{ii}} = \frac{1}{1 + \frac{q_i^2}{2\sigma_q^2}}. \tag{4.60}
\]

We can make an analogy with the zero order quadratic regularization to understand the effect of \( Q \) on solving the system expressed by equation (4.58). In the zero order quadratic regularization the damping term \( \mu \) in equation (4.51) corresponds to the ratio of two variances, \( \mu = \sigma_n^2/\sigma_q^2 \). When \( \rho_2 \) or \( \rho_3 \) are used, the damping term becomes a ratio of the variance of the noise to a model dependent variance which I will designate \( \sigma^2(q_i) \). This variable takes the following form when the problem is regularized with \( \rho_2 \):

\[
\sigma^2(q_i) = \begin{cases} 
\sigma_q & \text{if } |\frac{q_i}{\sigma_q}| \leq a \\
\frac{\sigma_q^2}{\sigma_q^2} & \text{if } |\frac{q_i}{\sigma_q}| > a
\end{cases} \tag{4.61}
\]

The last equation shows that above a threshold the variance of \( x_i \) is proportional to \( |x_i| \).

A similar argument leads to the variance for the regularization with \( \rho_3 \)

\[
\sigma^2(q_i) = \sigma_q^2 + \frac{q_i^2}{2}, \tag{4.62}
\]

in which case the variance has a parabolic growth with the amplitude of \( q \).

Equation (4.58) is solved using the following iterative scheme:
1. Start with an initial reflectivity sequence $q^0$

2. Select the hyperparameters of the problem $\sigma_n$, $\sigma_q$, and $\alpha$ (Huber criterion) or $\sigma_n$ and $\sigma_q$ (Cauchy criterion).

3. Compute $\mu = \sigma_n^2/\sigma_q^2$, $Q^{(0)}$, and the source autocorrelation matrix $R$.

4. Iteratively solve equation (4.58) using the following algorithm

$$q^{(k)} = (\mu Q^{(k-1)} + R)^{-1}g$$

where $k$ is the iteration number.

5. The procedure is stopped when the following tolerance criterion is satisfied

$$\frac{|J^{(k)} - J^{(k-1)}|}{(|J^{(k)}| + |J^{(k-1)}|)/2} < \text{tolerance}$$

where $J = J_{1.2}$ or $J_{1.3}$ depending on the regularization criterion.

6. Compute the data misfit. Select new hyperparameters if the misfit is not satisfactory. The strategy for hyperparameter selection is discussed in the following section.

Each iteration demands one matrix inversion. The procedure can be accelerated by using an iterative solver like the conjugate gradient (CG) algorithm. The advantage of using CG is that an approximate solution can be computed by truncating the number of iterations.

The effect of $Q$ can be summarized as follows. In each iteration, the non-linearity produces a solution that has the minimum amount of structure or maximum sparseness. The validity of this type of solutions is subject to the validity of the sparse reflectivity assumption.

### 4.3.4 Hyperparameter selection

The determination of the parameter $\mu$ in equation (4.58) is crucial, but unfortunately it cannot be determined a priori. A wrong selection of $\mu$ may yield a solution that is unreasonable. We will assume that the variance of the noise $\sigma_n^2$ is known. If the Cauchy criterion is adopted, only one independent parameter $\sigma_x$ must be determined ($\mu = \sigma_n^2/\sigma_q^2$).
When the Huber criterion is used, two independent parameters are needed: $\sigma_q$ and $a$. The parameter $a$ is assigned as follows $a = c \times \sigma_q$, where $c$ is a scalar ($c = 0.1 - 1$). If $c$ is large ($c > 2$) the Huber criterion behaves like the standard quadratic form $\rho_1$.

We adopt the discrepancy principle which determines the parameter $\sigma_q$ from the requirement that the data misfit matches the power of the noise. Since we have assumed that the noise is normally distributed the data misfit obeys a $\chi^2$ statistics

$$
\chi^2 = \frac{1}{\sigma_n^2} \sum_{k=1}^{n_s} \epsilon_k^2.
$$

The expected value of the $\chi^2$ statistic is used as a target misfit, $E[\chi^2] = n_s$ ($n_s$ is the number of observations), where the largest acceptable value at 99% confidence limit is $\approx n_s + 3.3\sqrt{n_s}$.

Figure (4.9a) portrays a simulated reflectivity impulse response for a simple earth model, the seismogram generated by convolving the impulse response with the source, and the seismic source. Gaussian noise was added to the synthetic seismogram with standard deviation $\sigma_n = 5 \times 10^{-2}$. This represents a relative noise amplitude of 17.5%. The relative noise magnitude is specified by a percentage of the maximum noise-free seismogram amplitude that its standard deviation represents ($\sigma_n / \max(s_k) \times 100$).

The deconvolution was carried out using zero order quadratic regularization (minimizing $J_{11}$), the Huber criterion $\rho_2$ (minimizing $J_{12}$) and the Cauchy criterion $\rho_3$ (minimizing $J_{13}$). The estimated impulse responses are displayed in Figures 4.9b, c, and d together with the reconstructed seismograms and residuals (original minus reconstructed data). The parameter $\sigma_q$ was selected according to the $\chi^2$ criterion.

The solution with $\rho_1$ (zero order quadratic regularization) does not allow us to properly identify each arrival. The re-weighted strategy yields highly resolved estimates of the position and amplitude of each seismic reflection.

In general, about 5 – 10 iterations are sufficient to find a good approximation to the minimum of the cost function.

The portion of stacked seismic section shown in Figure (4.10a) is used to test the performance of the re-weighted deconvolution procedure when dealing with field data. The stacked section is obtained by applying normal moveout correction, and summing traces from common mid point (CMP) gathers. The data consist of 24 traces which delineate several seismic horizons. In particular, we are interested in the coherent events at $\approx 0.95$sec which may represent a thin layer. The seismic wavelet (4.11) was extracted using a cepstrum-cumulant approach. The wavelet is retrieved using two different techniques which are radically different in formulation: a cepstrum decomposition and
4.3. SPARSE DECONVOLUTION AND BAYESIAN ANALYSIS

Figure 4.7: Cost functions. a) $\rho_1(u)$, b) $\rho_2(u) a = 1$, c) $\rho_2(u) a = 2$, and d) $\rho_3(u)$.

a fourth-order cumulant matching approach. Since the recovered wavelets were very similar, the confidence in the wavelet estimators increases. It is important to stress that, unlike in many deconvolution scenarios, in seismic deconvolution the kernel function or source wavelet is unknown. The deconvolved data are shown in Figure (4.10). In this example, the problem is regularized by means of the Cauchy criterion ($\rho_3$). The $\chi^2$ criterion was used to estimate the parameter $\sigma_x$. A relative noise amplitude of 2\% was assumed. The latter was used to estimate the standard error of the noise $\sigma_n$. Similar results were obtained using Huber's weights.
Figure 4.8: Influence functions. a) $\psi_1(u)$, b) $\psi_2(u) a = 1$, c) $\psi_2(u) a = 2$, and d) $\psi_3(u)$. 
Figure 4.9: a) Synthetic impulse response (left), seismogram (center), and source wavelet (right). The seismogram was contaminated with Gaussian noise. b) Deconvolution using zero order quadratic regularization: estimated impulse response (left), reconstructed seismogram (center), and residuals (original minus reconstructed data). c) Deconvolution by means of the Huber criterion \((a = 1)\). d) Deconvolution by means of the Cauchy criterion.
Figure 4.10: (a) Original seismic section. (b) Deconvolved seismic section using the Cauchy criterion to regularize the inversion. The source wavelet was retrieved using a combined cepstrum-cumulant approach.
Figure 4.11: Seismic wavelet corresponding to the data inverted in the previous Figure.
4.4 Bayesian inversion of impedance

In this section we discuss the problem of estimating a "blocky" impedance model from normal incidence data. We will adopt again the convolution model

\[ s_k = w_k * q_k . \]  

(4.66)

The reflectivity is denoted by \( q_k \) and the wavelet by \( w_k \). The goal is to recover \( q_k \) from a noisy version of \( s_k \). We assume that impedance constraints are provided at different time levels. After writing the usual logarithmic approximation for the impedance

\[ \xi_k = \frac{1}{2} \ln(\tilde{z}_k / \tilde{z}_0) = \sum_{i=1}^{N_k} x_i , \]  

(4.67)

we are now in condition of writing equations (4.66) and (4.67) as two linear system of equations. In matrix notation

\[ Wq = s + n \]  

(4.68)

\[ Cx = \xi + \epsilon , \]  

(4.69)

where the matrices \( W \) and \( C \) correspond to the wavelet matrix and to the impedance constraint matrix, respectively. The matrix \( W \) contains the wavelet properly padded with zeros in order to express discrete convolution in matrix form. The matrix \( C \) is a simple integrator operator. Bayes’ rule is used to incorporate the prior probability of the unknown reflectivity \( q \) into the problem. In many applications we want to estimate a blocky impedance profile. In this case, a ‘long tailed’ distribution may be used to estimate a sparse reflectivity sequence.

Noise in the trace is modelled by means of the usual Gaussian assumption. The uncertainties of the constraints are also assumed Gaussian (note that we are using the variable \( \xi \) which can take positive and negative values and not \( z \) which is strictly positive).

Bayes’ rule is used to define the posteriori probability of the reflectivity sequence. The solution is computed by maximizing the posteriori probability (MAP solution). This is equivalent to minimize the following cost function:

\[ J = \alpha J_q + \frac{1}{2} \left| \left| \mathbb{I}_2 \left( \frac{1}{\sigma} (Wq - s) \right) \right| \right|^2 + \frac{1}{2} \left| \left| S^{-1} (Cq - \xi) \right| \right|^2 \]  

(4.70)
where \( \sigma^2 \) is the variance of the noise in the seismic trace and the matrix \( S \) is a diagonal matrix with the following elements

\[
S_{ii} = \sigma_{c_i}^2.
\]  

(4.71)

In equation (4.70) we are specifying three different features that the solution must satisfied:

- 1 - The solution must be sparse.
- 2 - The solution must honour the seismic trace.
- 3 - The solution must honour a set of impedance constraints.

The parameter \( \alpha \) is the weighting parameter or hyperparameter that determines the relative amount of sparseness that can be brought into the inversion. The term \( J_q \) is derived using four different priors which induce the associated regularization criteria for sparse spike inversion. The four sparseness criteria that we have studied are the \( L_p \) criterion, the \textit{Cauchy} criterion, the \textit{Sech} criterion, and \textit{Huber} criterion. These criteria were all used in robust statistics to diminish the influence of outliers in estimation problems (Huber, 1981). In our application, these criteria are used to impose sparseness into the reflectivity estimate. \( J_q \) in equation (4.70) is given by one of the following regularization terms

\[
J_p = \frac{1}{p} \sum_i |q_i|^p
\]  

(4.72)

\[
J_{\text{Cauchy}} = \frac{1}{2} \sum_i \ln \left( 1 + \frac{q_i^2}{\sigma_q^2} \right)
\]  

(4.73)

\[
J_{\text{Sech}} = \sum_i \ln \left( \cosh \frac{q_i}{\sigma_q} \right)
\]  

(4.74)

\[
J_{\text{Huber}} = \sum_i \begin{cases} 
\frac{q_i^2}{2} & \text{if } |q_i| \leq q_c \\
a |q_i| - \frac{q_c^2}{2} & \text{if } |q_i| > q_c
\end{cases}
\]  

(4.75)

A non-linear conjugate gradient algorithm is used to minimize the cost function. The parameter \( \alpha \) is continuously adapted to satisfy a misfit criterion. As an example we show the inversion of two traces with the \( L_p \) \((p = 1.2) \) prior and with the \textit{Cauchy} criterion. The hyperparameters of the problem were fitted using a \( \chi^2 \) criterion. The same misfit target was used in both examples, \( \chi^2 = \text{ number of traces + number of constraints} \).
In Figure (4.12) we portray a window of a seismic section pre-processed for impedance inversion. The inversion was carried out using the *Huber* and $L_p$ criteria ($p = 1.1$). The reflectivities estimates are shown in Figures (4.13) and (4.15). The constrained inversion of the impedance profile is depicted in Figures (4.14) and (4.16).
Figure 4.13: Reflectivity inversion using the $L_p$ norm, $p = 1.1$
Figure 4.14: Constrained impedance inversion using the $L_p$ norm, $p = 1.1$. 
Figure 4.15: Reflectivity inversion \textit{Huber} norm.
Figure 4.16: Constrained impedance inversion using $H_{uber}$ norm
4.5 Linear programming impedance inversion

In this section I will discuss the classical approach to impedance inversion proposed by Oldenburg et al. (1983).

I will also provide a subroutine to perform the $L_1$ inversion of a seismogram using the a Linear programming solver.

The subroutine $11_{-inv}$ is designed to perform 1-D inversion of the acoustic impedance. The algorithm permits the incorporation of impedance constraints at different times in the form of upper and lower bounds. The convolution model is adopted:

$$s_k = w_k * q_k.$$  \hspace{1cm} (4.76)

The wavelet is assumed to be known or to be well approximated by an estimate. The reflectivity is denoted by $q_k$ and the wavelet by $w_k$. The goal is to recover $q_k$ from a noisy version of $s_k$. The algorithm also assumes that impedances constraints are provided at different time levels. In such a case we can write the usual logarithmic approximation for the impedance

$$\xi_k = \frac{1}{2} \ln(z_k/z_0) = \sum_{i=1}^{N_k} q_i,$$  \hspace{1cm} (4.77)

In matrix notation we can write

$$Wq = s + e$$  \hspace{1cm} (4.78)

$$Cq = \xi$$  \hspace{1cm} (4.79)

where the matrices $W$ and $C$ correspond to the wavelet matrix and to the impedance constraint matrix, respectively. The matrix $W$ contains the wavelet properly padded with zeros in order to express discrete convolution in matrix form. The matrix $C$ is a simple integrator operator.

The procedure proposed to recover the reflectivity is based on the minimization of the $l_1$ cost function of the problem. Instead of using a Conjugate Gradient technique to minimize the cost function (like in the Bayesian approach discussed in the previous section.) we adopt the linear programming approach.

\footnote{Let me know if you want to have the complete f77 source code}
4.5.1 Constrained minimization using linear programming

The cost function of the problem is defined as

\[ J = \alpha |q_1| + |e_1| \]  

(4.80)

and the constraint minimization problem is set as follows

Minimize \[ J = \alpha |q_1| + |e_1| \]  
subject to \[ Wq = s + e \]  
and \[ \xi_l < Cq < \xi_u \]  

(4.81) \hspace{3cm} (4.82) \hspace{3cm} (4.83)

The last problem is solved using the linear programming approach. We first specify a tableau (subroutine tableau) where we load the objective function, the data constraint and the inequality constraints. The Linear programming solution is then invoked to retrieve the solution (subroutine c11).

The parameter \( \alpha \) is the tradeoff parameter of the problem. In the program a1pha is given in percentage of the \( l_1 \) norm of the wavelet. This is analogous to the pre-whitening parameter in spiking deconvolution which is given as percentage of the zero lag autocorrelation coefficient.

Usually, \( 1 - 10\% \) should be enough to stabilize the inversion. The parameter also controls the sparseness of the solution (number of nonzero reflectivity samples). When \( \alpha \to 0 \) the inversion is too rough (unstable). When \( \alpha \) is too large (50\%) the inversion is too sparse.

4.5.2 Example

A sparse reflectivity series convolved with a 20Hz Ricker wavelet is used to test the algorithm. The length of the trace is 150 samples. The additive noise is white and Gaussian (SNR=20). In this example we retrieved impedance bounds every 10 samples. The total number of impedance bounds to be honoured is \( nic = 15 \). The results are shown in Figure (4.17).

4.5.3 Linear programming code

The subroutine 11_inv is the basic program that you need for sparse inversion. Some variables were set with default values. The subroutine tableau is used to load the objec-
Figure 4.17: (a) Seismic reflectivity. (b) Input seismic trace. (c) $L_1$ inversion of the reflectivity. (d) Predicted trace (convolution of the wavelet with the inverted reflectivity). (e) True impedance profile. The bars indicate the impedance bounds (hard constraints). (f) Impedance recovered with the $L_1$ inversion, note that the bounds are properly honoured.

tive function of the problem, the data constraints and the impedance bounds.

The code is too long, this is why I only print the part containing the comments. If you wish to try with this code, I have a ready-to-use version in a disk.

```
subroutine l1_inv(t,n1,w,nw,
   * nic,ncon,z_upper,z_lower,
   * alpha,r,tp,iter,e1,e2)
  
c This subroutine does the l1 inversion of a seismogram
```
c using the linear programming approach

c INPUT

c n1 : number of samples of the trace
c t(n1) : the seismic trace (properly scaled)
c nw : number of points of the wavelet
c w(nw) : the wavelet (properly scaled)
c nic : number of impedance constraints
c ncon(nic) : position of each impedance constraint in samples
c z_upper(nic) : upper impedance constraint at the position ncon
c z_lower(nic) : lower impedance constraint at the position ncon
    the impedance bounds are for the variable
    0.5ln(z/z0) (see text)
c alpha : regularization parameter.
    alpha is a percentage of the l1 norm of the wavelet.
    In spiking deconv. (12) is a percentage of the main diagonal
    of the autocorrelation matrix. The strategy that I have
    adopted is similar. Use 1-10%.
c

c OUTPUT

c r(n1) : the reflectivity estimate
c tp(n1) : the predicted seismic trace, i.e., r(t)*w(t)
c iter : number of iterations that the linear programing routine
    used to find a feasible solution
c e1 : the l1 error
c e2 : the rms error
    e1,e2 are the misfit figure.


c SUBR:  tableau : make the tableau for the lprog.
c         cl1 : solve the lprog. problem

c NOTES: the variable KODE serves to check is the lprog. routine
    has found a feasible solution. KODE=0 (always) before
    calling to cl1 (The linear prog). KODE=0 (always) after calling cl1 for normal execution.

    parameter (nrd=400,nw=101,ntd=500,ncd=400,
4.5. **LINEAR PROGRAMMING IMPEDANCE INVERSION**

* klm=nr+2*nt+2*ncd,
* klm2=nr+2*nt+2*ncd+2,
* nklm=2*(nt+nr)+nt+2*ncd,
* n2d =nt+nr+2)

real * 8 q(klm2d,n2d), x(n2d), res(klm)
real * 8 cu(2,nkld)
integer iu(2,nklmd), s(klm)

real * 8 w(nwd),r(nrd),t(ntd),tp(ntd)
real * 8 z_upper(nrd),z_lower(nrd)
integer ncon(nrd)

real * 8 toler
real * 8 alpha

real * 8 e1,e2

nr=n1
nt=n1+nw-1
do i=nr+1,nt
t(i)=0.
endo

c Prepare the tableau for the simplex
    call tableau(nw,w, nr, nt, t, alpha, nic, ncon, zcon,
    #       z_upper, z_lower, q,
    #       k, l, m, n)
4.6  Non-minimum phase wavelet estimation

In this section we will analyse the problem of estimating a non-minimum phase wavelet for seismic deconvolution/inversion. The wavelet we have used to process our data in the preceding section was estimated using a cumulant approach. We will discuss here how to estimate the seismic wavelet using the aforementioned procedure.

One of the early techniques for nonminimum phase signal decomposition has been homomorphic filtering based on the complex cepstrum (Ulrych, 1971) or differential cepstrum (Bednar and Watt, 1985). Anyone exposed to homomorphic filtering has probably experienced the gap that exits between theory and practice. The approach is elegant and general (Hilbert transformation and spectral factorization can be implemented by means of the homomorphic transform.) However, practical problems have provided serious obstacles for the cepstral approach in the deconvolution arena. In particular cepstral decomposition assumes that the reflectivity is an impulsive train that can be completely separated from the wavelet in the cepstral domain (Ulrych, 1971). In fact, the technique does not even consider the reflectivity sequence as a stationary process. Another problem with the cepstrum is that the homomorphic system does not allow for the presence of additive noise in the formulation of the problem. In fact, the highly nonlinear nature of the cepstrum complicates the effect of additive noise.

The definition of the cepstrum in terms of higher order cumulants enables us to retrieve the cepstrum of the source wavelet (analytically) when the reflectivity does not consist of a train of spikes. Since the bispectrum and the trispectrum conserve the phase characteristic of the wavelet it is evident that the cepstrum derived from these polyspectra will also conserve phase information.

4.6.1  Non-minimum phase system identification

The classical system identification literature has been primarily dominated by least squares approaches. Least squares based techniques are attractive since they yield maximum likelihood estimates of parameters where the observable are a linear combination of a white sequence. These identification procedures use the autocorrelation function of the data. However, the autocorrelation function annihilates the phase of the system (c.f. the wavelet), and therefore these techniques are useful primarily to identify minimum phase systems, or systems where the phase can be specified a priori i.e., a zero phase assumption.

We begin with Wold’s decomposition theorem. According to this celebrated theorem any discrete stationary process can be expressed as the sum of two independent processes, one purely deterministic, and another purely non-deterministic. The decompo-
sition theorem also states that the purely non-deterministic part can be written in terms of a linear transformation of a white process, \( \varepsilon_k \). In other words as long as the process is stationary, there is always a representation of the process given by

\[
x_t = \sum_k \varepsilon_k h_{t-k},
\]

(4.84)

where for simplicity we have omitted the deterministic part of the process which is not needed in the seismic deconvolution scenario. Clearly, the Wold decomposition theorem dictates that a decomposition exists where the wavelet \( h_k \) is causal and minimum phase. It is clear that one of the consequences of this theorem is that the description of the process is non-unique. One can generate a stationary nonlinear process and according to the theorem this process will have a MA representation given by equation (4.84).

We will use the following model to describe the seismic trace

\[
x_k = r_k * w_k \\
s_k = q_k + n_k
\]

(4.85)

where \( x_k \) is the noise-free seismic trace, and \( s_k \) is the trace corrupted with noise. The source wavelet is denoted by \( w_k \) and the reflectivity sequence by \( q_k \). We will assume that \( q_k \) is white and iid. The noise \( n_k \) is considered zero-mean, Gaussian, and independent of \( x_k \). The transfer function of the system is given by

\[
W(z) = |W(z)| e^{i\Phi(z)}.
\]

(4.86)

The seismic deconvolution problem is to reconstruct the magnitude and the phase of the transfer function from the output data \( s_k \). Thus far, we have only considered that \( q_k \) is white. To continue with the analysis the following comments are in order:

1. If \( q_k \) is Gaussian and \( W(z) \) is minimum phase, autocorrelation based methods will correctly identify both the amplitude and the phase of the system.
2. If \( q_k \) is Gaussian and \( W(z) \) is nonminimum phase, no technique will correctly identify the phase of the system.
3. If \( q_k \) is non-Gaussian and \( W(z) \) is nonminimum phase, true magnitude and phase of the system transfer function can be recovered by knowing the actual distribu-
tion of $q_k$. For MA processes of order one, it has been demonstrated that a $L_1$ optimization provides an estimate of the amplitude and phase of the system when the driving noise of the process is non-Gaussian (Scargle, 1977).

The above statement (3) is very important. In fact, it suggests that we can still have faith in stochastic wavelet estimation procedures. It is also clear that non-Gaussianity plays a key role in nonminimum phase wavelet estimation. In fact, MED-type estimators are an example where departure from normality is postulated. The non-uniqueness expressed by Wold theorem is eliminated by restricting $q_k$ to be non-Gaussian.

Higher order spectra are defined in terms of higher order cumulants and contain information regarding deviation from Gaussianity. Quite contrary to power spectral density estimates, higher order spectra retain phase information which allows nonminimum phase system identification/estimation.

### 4.6.2 The bicepstrum

If $W(z)$ is stable and if $q_k$ is non-Gaussian, white, iid., with skewness $\beta \neq 0$ then the bispectrum of $x_k$ is given by

$$B_x(z_1, z_2) = \sum_m \sum_n r_x^{(3)}(m, n) z_1^m z_2^n$$

where

$$r_x^{(3)}(m, n) = E[x_k x_{k+m} x_{k+n}]$$

is the third order moment of the data. Since the third order moment of a Gaussian signal vanishes, we can write $B_s(z_1, z_2) = B_x(z_1, z_2)$. The bispectrum can be written in terms of the transfer function $W(z)$ as follows (Nikias and Raghuveer, 1987)

$$B_x(z_1, z_2) = \beta W(z_1) W(z_2) W(z_1^{-1} z_2^{-1})$$

Since the wavelet is, in general, nonminimum phase, the transfer function can be written as

$$W(z) = A z^L W^{\min}(z) W^{\max}(z^{-1})$$
where \( A \) is a constant, \( l \) an integer associated with a linear phase shift, and \( W^{\text{min}}(z) \), \( W^{\text{max}}(z^{-1}) \) are the minimum and maximum phase components of the wavelet, respectively. Substituting equation (4.90) into (4.89) the bispectrum becomes

\[
B_x(z_1, z_2) = \beta W^{\text{min}}(z_1)W^{\text{min}}(z_2) \\
W^{\text{max}}(z_1^{-1})W^{\text{max}}(z_2^{-1}) \\
W^{\text{min}}(z_1^{-1}z_2^{-1})W^{\text{max}}(z_1z_2). \tag{4.91}
\]

Now, we define the bicepstrum as

\[
\hat{b}(n, m) = Z^{-1}[\ln B_x(z_1, z_2)] \tag{4.92}
\]

where \( Z^{-1} \) stands for the inverse 2-D \( z \)-transform. We note that

\[
\ln[B_x(z_1, z_2)] = \ln[\beta] \\
+ \ln[W^{\text{min}}(z_1)] + \ln[W^{\text{min}}(z_2)] \\
+ \ln[W^{\text{max}}(z_1^{-1})] + \ln[W^{\text{max}}(z_2^{-1})] \\
+ \ln[W^{\text{min}}(z_1^{-1}z_2^{-1})] + \ln[W^{\text{max}}(z_1z_2)]. \tag{4.93}
\]

The inversion of equation (4.93) yields a complete separation into minimum phase and maximum phase components. Since the homomorphic transform maps minimum phase components into the positive semi-axis of the cepstrum and maximum phase components into the negative semi-axis, it is easy to verify that

\[
\hat{b}(m, 0) = \hat{w}^{\text{min}}(m), \quad m > 0 \tag{4.94}
\]

\[
\hat{b}(m, 0) = \hat{w}^{\text{max}}(m), \quad m < 0. \tag{4.95}
\]

where \( \hat{w}(n) \) indicates the cepstrum of the wavelet. It is clear that with only two semi-axes of the bicepstrum we can completely define the cepstrum of the wavelet (the value \( \hat{w}(0) \) cannot be recovered but represents only a scale factor).

We use a similar approach below to compute the tricepstrum (the cepstrum of the fourth order cumulant).
4.6.3 The tricepstrum

Defining the trispectrum as

\[
T_x(z_1, z_2, z_3) = \sum_m \sum_n \sum_l r_x^{(4)}(n, m, l) z_1^m z_2^n z_3^l
\]  

where \( r_x^{(4)}(n, m, l) \) is now the fourth order cumulant of \( x_k \). The trispectrum can be written in terms of the transfer function of the system as follows

\[
T_x(z_1, z_2, z_3) = \gamma A^4 W^{\min}(z_1) W^{\min}(z_2) W^{\min}(z_3) \\
W^{\max}(z_1^{-1}) W^{\max}(z_2^{-1}) W^{\max}(z_3^{-1}) \\
W^{\min}(z_1^{-1} z_2^{-1} z_3^{-1}) W^{\max}(z_1 z_2 z_3)
\]  

(4.97)

After taking logarithm of the trispectrum we end up with

\[
\ln[T_x(z_1, z_2, z_3)] = \ln[\gamma A^4] + \ln[W^{\min}(z_1)] + \ln[W^{\min}(z_2)] + \ln[W^{\min}(z_3)] \\
+ \ln[W^{\max}(z_1^{-1})] + \ln[W^{\max}(z_2^{-1})] + \ln[W^{\max}(z_3^{-1})] \\
\ln[W^{\min}(z_1^{-1} z_2^{-1} z_3^{-1})] + \ln[W^{\max}(z_1 z_2 z_3)]
\]  

(4.98)

The inversion of the last expression will map minimum and phase components into the tricepstrum domain as follows:

\[
\hat{t}(m, n, l) = \begin{cases} 
\ln[\gamma A^4] & m = n = l = 0 \\
\hat{w}_{\min}(m) & m > 0, \quad n = l = 0 \\
\hat{w}_{\min}(n) & n > 0, \quad m = l = 0 \\
\hat{w}_{\min}(l) & l > 0, \quad m = n = 0 \\
\hat{w}_{\max}(m) & m < 0, \quad n = l = 0 \\
\hat{w}_{\max}(n) & n < 0, \quad m = l = 0 \\
\hat{w}_{\max}(l) & l < 0, \quad m = n = 0 \\
\hat{w}_{\min}(m) & m = n = l < 0 \\
\hat{w}_{\max}(m) & m = n = l > 0.
\end{cases}
\]

The origin of tricepstrum, \( \hat{t}(0, 0, 0) = \ln(A^4 \gamma) \) represents a scale factor and, therefore, can be ignored.
4.6. NON-MINIMUM PHASE WAVELET ESTIMATION

In general, we estimate the wavelet from 2 semi-axis of the tricepstrum:

\[
\hat{t}(m, 0, 0) = \hat{\omega}_{\text{min}}(m), \quad m > 0 
\]  \hspace{1cm} (4.99)

\[
\hat{t}(m, 0, 0) = \hat{\omega}_{\text{max}}(m), \quad m < 0. 
\]  \hspace{1cm} (4.100)

4.6.4 Computing the bicepstrum and the tricepstrum

We first note that if \(z_1 = e^{-i\omega_1}\) and \(z_2 = e^{-i\omega_2}\) we have

\[
\hat{B}_x(\omega_1, \omega_2) = \ln[B_x(\omega_1, \omega_2)] = \ln |B_x(\omega_1, \omega_2)| + i[2k\pi + \Phi(\omega_1, \omega_2)].
\]  \hspace{1cm} (4.101)

Since the phase of the complex variable \(\hat{B}_x(\omega_1, \omega_2)\) is undefined it appears that an unwrapping procedure is mandatory. Fortunately, the unwrapping can be omitted by defining the differential cepstrum, as follows

\[
\frac{d\hat{B}_x(\omega_1, \omega_2)}{d\omega_1} = \frac{1}{B_x(\omega_1, \omega_2)} \frac{dB_x(\omega_1, \omega_2)}{d\omega_1}
\]  \hspace{1cm} (4.102)

where the derivatives are obtained by

\[
\frac{dB_x(\omega_1, \omega_2)}{d\omega_1} = \mathcal{F}_2[-imr_{x_3}(m, n)]
\]  \hspace{1cm} (4.103)

and we can write (4.102) as

\[
\hat{b}(m, n) = \frac{1}{m} \mathcal{F}_2^{-1} \left[ \frac{\mathcal{F}_2[mr_{x_3}(m, n)]}{\mathcal{F}_2[r_{x_3}(m, n)]} \right].
\]  \hspace{1cm} (4.104)

A similar algorithm is used to estimate the tricepstrum

\[
\hat{t}(m, n, l) = \frac{1}{m} \mathcal{F}_3^{-1} \left[ \frac{\mathcal{F}_3[mr_{x_4}(m, n, l)]}{\mathcal{F}_3[r_{x_4}(m, n, l)]} \right].
\]  \hspace{1cm} (4.105)
4.6.5 Examples

In the first simulation we convolved a nonminimum phase wavelet with an exponentially distributed sequence with skewness $\beta = 1$. The additive noise is white and Gaussian. The standard deviation of the noise represents 1% of the maximum amplitude of the signal. We used 6 records of $600$ samples to estimate the third order cumulant. The bicepstrum was computed using equation (4.104). Figure (4.18) shows the true cepstrum of the wavelet and the cepstrum extracted form the bicepstrum using. Figure (4.19) shows the true and the estimated wavelets together with the associated minimum and maximum phase components. The maximum phase component has a zero close to the unit circle that is manifested in the negative semi-axis of the cepstrum as a ‘long’ decaying oscillatory signal.
Figure 4.19: (a) and (b): True minimum phase and maximum phase components of the wavelet (c). (d), (e), and (f): Estimators of (a), (b), and (c) computed form the biceps-strum.
CHAPTER 4. DECONVOLUTION OF REFLECTIVITY SERIES

Figure 4.20: (a) Cepstrum of the true wavelet. (b) Cepstrum of wavelet derived from the tricepstrum.

The same wavelet was convolved with a white, non-Gaussian, reflectivity with non vanishing kurtosis. The fourth order cumulant was computed from 4 records of 1000 samples each. The cepstrum of the wavelet was estimated from the tricepstrum. The results are shown in Figure (4.20). The time signature and its maximum and minimum phase components are displayed in Figure (4.21). The technique permits the recovery of the cepstral coefficients of the wavelet with an accuracy proportional to the accuracy of the estimation of the cumulant.

In this section we provide an heuristic analysis of the performance of the algorithm.

We simulate 20 realizations of a non-Gaussian (\(\gamma \neq 0\)) process which is convolved with the source wavelet portrayed in Figure (4.22). Figure (4.23) shows the wavelet retrieved from the tricepstrum. The fourth-order cumulant was estimated using 4 segments of 1000 samples each. In figures 4.24) and (4.25), we used 4 segments of 500 and 250 samples, respectively. These results indicate that a fairly good reconstruction of the amplitude and phase can be achieved for large data sets.

Figure (4.26) portrays a segment of a seismic section pre-processed for impedance inversion. The segment is composed of 24 traces of 300 samples each. The fourth order cumulant is estimated from each trace and the average cumulant is used to identify the wavelet. Figure (4.27) shows the cepstrum of the wavelet retrieved from one of the axis of the tricepstrum. The minimum and maximum phase components of the wavelet are shown in Figure (4.28).

The tricepstrum estimator of the wavelet is illustrated in Figure (4.29). For comparison we also show the estimator of the wavelet computed using a cumulant matching approach (Velis and Ulrych, 1996). The later uses a global optimization procedure (simulated annealing) to find the wavelet that best reproduces the cumulant of the data.
Figure 4.21: (a) and (b): True minimum phase and maximum phase components of the wavelet (c). (d), (e), and (f): Estimators of (a), (b), and (c) computed from the triceps-
Figure 4.22: Synthetic wavelet.
Figure 4.23: Wavelet estimation using the tricepstrum. The fourth order cumulant was estimated from 4 segments of 1000 samples each. The figures correspond to 20 realizations of the process.
Figure 4.24: Wavelet estimation using the tricepstrum. The fourth order cumulant was estimated from 4 segments of 500 samples each. The figures correspond to 20 realizations of the process.
Figure 4.25: Wavelet estimation using the tricepstrum. The fourth order cumulant was estimated from 4 segments of 250 samples each. The figures correspond to 20 realizations of the process.
CHAPTER 4. DECONVOLUTION OF REFLECTIVITY SERIES

Figure 4.26: Segment of seismic section pre-processed for impedance inversion.

Figure 4.27: Cepstrum of wavelet estimated from the tricepstrum of the data. An average fourth-order cumulant derived from 24 traces was used to retrieve the tricepstrum.
Figure 4.28: Minimum and maximum phase decomposition of the wavelet after cepstral lifting.

Figure 4.29: Wavelet estimates computed using the tricepstrum (left) and cumulant matching plus non-linear optimization (right).
Figure 4.30: Schematic representation of the tricepstrum for a non-Gaussian MA process.
4.7 Minimum entropy deconvolution

The MED technique (minimum entropy deconvolution) proposed by Wiggins (1978), offers a different approach to seismic deconvolution. While the classical methods such as spiking and predictive deconvolution (Robinson and Treitel, 1980) seek to whiten the spectra, the MED seeks the smallest number of large spikes which are consistent with the data.

Despite these differences, both methods constitute a linear approach to seismic deconvolution. The spiking and predictive filters are obtained by inverting the Toeplitz matrix; the MED filter is calculated in an iterative procedure in which the Toeplitz matrix is inverted at each step (Wiggins, 1978). These filters are quite different in nature, but as they are linear operators none of them can handle band-limited data properly. This limitation is difficult to overcome when dealing with noisy data.

In this work a frequency-domain version of the MED scheme is developed. This approach involves maximizing a generalized entropy norm with respect to the seismic reflectivity. The particular norm in which we are going to focus our attention (the logarithmic norm) has also been used for deconvolution and wavelet extraction by Postic et al. (1980) in an attempt to overcome the limitations of the classical MED method (Wiggins, 1978).

For band-limited signals the deconvolution can be achieved by reconstruction of the reflectivity spectrum. Two main procedures have been developed to reach the latter goal. The first method (Levy and Fullagar, 1981) is based on a linear programming (LP) approach. This method attempts to find the reflectivity series with minimum absolute norm remaining consistent with the data. The second approach (Lines and Clayton, 1977) fits a complex autoregressive (AR) model to the data spectrum and from the information available in the actual band attempts to extrapolate the missing low and high frequencies of the reflectivity. Both methods have been improved and expanded to cope with acoustic impedance inversion from band-limited reflection seismograms. In the LP approach, Oldenburg et al. (1983) incorporated impedance constraints to the problem.

Our method can be mainly compared with the LP approach. This is because both methods seek for an extremum point of a given norm which is a function of the underlying unknown function: the reflectivity. The main advantage of an entropy norm over the absolute norm ($l_1$) is that the minimization procedure leads to an easy to handle algorithm, avoiding the computationally expensive cost of linear programming routines. It must be stressed that the proposed method provides a unifying thread between the LP (Oldenburg et al., 1983) and the MED approach (Wiggins, 1978).
4.7.1 Minimum Entropy estimators

The normal incidence seismogram model can be expressed as the convolution between two basic components: the reflectivity \( y_t \) and the wavelet \( w_t \). If we denote the noise-free seismic trace by \( x_t \) then

\[
s_t = w_t * q_t
\]  

(4.106)

where \(*\) denotes discrete convolution. The goal of the deconvolution process is to recover \( q_t \) from \( s_t \). If we adopt a linear scheme, an operator \( f_t \) such that

\[
q_t = s_t * f_t
\]  

(4.107)

must be obtained. Note that if \( s_t \) is a band-limited signal, only a part of \( q_t \) can be recovered.

Usually the signal is contaminated with noise, then the normal incidence seismogram model is

\[
s_t = w_t * q_t + n_t.
\]  

(4.108)

We want to compute a filter \( f_t \) such that \( f_t * w_t = \delta_t \), but usually we have an estimate of the filter \( \hat{f}_t \). Then \( \hat{f}_t * w_t = a_t \), where \( a_t \) is called the averaging function which in the ideal case should resemble a delta function (Oldenburg, 1981). Operating with the filter \( \hat{f}_t \) on the seismic trace

\[
\hat{q}_t = a_t * q_t + \hat{f}_t * n_t
\]  

(4.109)

\[
= q_t + (a_t - \delta_t) * q_t + \hat{f}_t * n_t.
\]  

(4.110)

Equation (4.110) shows that the filter not only has to make \( a_t \) close to a delta, but also has to keep the noise level as small as possible. It follows that we are faced with the usual trade-off between resolution and statistical reliability.

At this point it must be said that even with the best seismic field and processing techniques, the band-pass nature of the "earth system" is always present, i.e. after removing the wavelet, only a portion of the reflectivity spectrum is available. In other words \( \hat{q}_t = q_t * a_t \) is band-limited. For further developments \( \hat{q}_t \) will be called the band-limited
reflectivity and $q_i$ the full-band reflectivity. We will assume that the wavelet has been removed, therefore $a_i$ has zero phase with constant amplitude in the frequency range $[\omega_L, \omega_H]$ and zero amplitude outside that interval.

Estimating the full-band reflectivity from the band-limited reflectivity is a nonunique linear inverse problem. Neglecting the noise term, the last assessment may be easily confirmed taking the Fourier transform of equation (4.102)

$$\hat{Q}(\omega) = A(\omega) \cdot Q(\omega).$$

(4.111)

It is easy to see that $Q(\omega)$ can take any value at those frequencies at which $A(\omega)$ vanishes. The knowledge of $\hat{Q}(\omega)$ is not enough to estimate the portion of $Q(\omega)$ outside the non-zero band of $A(\omega)$. Hence, there exists an infinite number of models $q_i$ that satisfy equation (4.110). In other words, $\hat{Q}(\omega)$ gives no information about the parts of $Q(\omega)$ which belong to the null space. In the next analysis we will discuss how to limit the nonuniqueness of the problem.

### 4.7.2 Entropy norms and simplicity

Among all the possible solutions to the problem stated in the previous section, we will look for those particular solutions in which a reasonable feature of the reflectivity is reached. Usually, parsimony is a required feature of an acceptable model. “Minimum structure” or “simple solution” are terms often used for a model with parsimonious behaviour. In Wiggins’ original approach, the term “minimum entropy” is used as synonymous with “maximum order”. The term is appropriate to set up the main difference between MED and spiking or predictive deconvolution. While spiking or predictive deconvolution attempt to whiten the data (minimum order), the MED approach seeks for a solution consisting mainly of isolated spikes. Wiggins’ entropy was inspired by factor analysis techniques, and can be regarded as a particular member of a broad family of norms of the form

$$V(q) = \frac{1}{NF(N)} \sum_{i=1}^{N} q_i^{*} F(q_i),$$

(4.112)

where the vector $q$ represents the reflection coefficient series of length $N$, and $q_i^{*}$ an amplitude normalized measure given by
In formula (4.112), \( F(q_i) \) is a monotonically increasing function of \( q_i \), which is often called the entropy function. Having defined \( F(q_i) \), the following inequality can be established:

\[
F(1)/F(N) \leq V \leq 1.
\] (4.114)

The normalization factor in equation (6) guarantees the same upper limit for any entropy function. Note that for the most simple case, a series with all zeros and one spike, the norm reaches the upper bound \( V = 1 \). When all the samples are equal \( V \) reaches the lower bound.

The original MED norm is obtained when \( F(q_i') = q_i' \). In many synthetic examples we found that this norm is very sensitive to strong reflections. To avoid such inconvenient we have tested other norms concluding that better results are achieved with the logarithmic norm in which \( F(q_i) = \ln(q_i) \).

### 4.7.3 Wiggins’ algorithm

A trivial solution to the problem stated in equation (4.107) is

\[
\hat{f}_i = s_i^{-1},
\hat{q}_i = \delta_i.
\] (4.115)

Where \( s_i^{-1} \) stands for the inverse of \( s_i \) if it exists. To avoid such solution a fixed length must be imposed to the operator \( f_i \), then

\[
q_n = \sum_{l=1}^{LF} f_i \cdot s_{n-l}.
\] (4.116)

The criterion for designing the operator \( f_k \) may be set as

\[
\frac{\partial V}{\partial f_k} = 0, \quad k = 1, 2, ..., LF;
\] (4.117)
4.7. MINIMUM ENTROPY DECONVOLUTION

\[ \frac{\partial V}{\partial f_k} = \frac{1}{N_F(N)} \sum_i \left( F(q'_i) + q'_i \frac{\partial F(q'_i)}{\partial q'_i} \right) \frac{\partial q'_i}{\partial f_k}. \]

From equation (4.107) it follows that \( \frac{\partial q_n}{\partial f_k} = s_{n-k} \), and after some algebraic manipulations equation (7.19) becomes

\[ \sum_l f_l \sum_n s_{n-k}s_{n-l} = \sum_i b_i s_{i-k}, \quad (4.118) \]

where

\[ b_i = \frac{G(q'_i) q_i}{N \sum_j G(q'_j) q'_j} \quad (4.119) \]

and

\[ G(q'_i) = F(q'_i) + q'_i \frac{\partial F(q'_i)}{\partial q'_i}. \quad (4.120) \]

We can use two criteria \( F(q'_i) = q'_i \) (Wiggins’ entropy function) or \( F(q'_i) = \ln(q'_i) \) (logarithmic entropy function). Numerical experiments suggests that the second one appears to be a better choice for seismic deconvolution (Sacchi et. al, 1994).

Expression (4.118) corresponds to the system used to design the well known Wiener or shaping filter (Robinson and Treitel, 1980). This filter seeks to convert an input signal \( x \) into a desired output \( b \). In matrix notation:

\[ Rf = g(f), \quad (4.121) \]

where \( R \) is the Toeplitz matrix of the data and the vector \( g(f) \) is the crosscorrelation between \( b \) and \( s \). The system must be solved through an iterative algorithm:

\[ f^{(n)} = R^{-1} \cdot g(f^{(n-1)}), \quad (4.122) \]

where the upper index \( n \) denotes iteration number. In each iteration the system is solved with Levinson’s algorithm (Robinson and Treitel, 1980). The initial value for this system is \( f^{(0)} = (0, 0, 0, \cdots, 1, \cdots, 0, 0, 0) \). Note that in each iteration the system attempts to reproduce the features of the reflectivity series. If the proper length is chosen, the system leads to a useful maximum and the main reflections can be estimated (Wiggins, 1978).
4.7.4 Frequency domain algorithm (Sacchi et. al, 1994)

In the frequency domain, the maximization of the entropy is subjected to the following constraint:

\[ Q(\omega) = \hat{Q}(\omega), \quad \omega \in [\omega_L, \omega_H]. \tag{4.123} \]

For practical purposes let us define equation (4.110) by means of the discrete Fourier transform

\[ \hat{Q}_k = A_k \cdot Q_k, \tag{4.124} \]

where the lower index \( k \) denotes frequency sample. So the maximization of \( V \) with mid-band constraints can be written down as

\[ \text{Maximize} \quad V, \tag{4.125} \]

subjected to

\[ Q_k = \sum_{n=0}^{N-1} \hat{q}_n e^{-i2\pi nk/N}, \quad k = k_L, \ldots, k_H, \tag{4.126} \]

where \( k_L \) and \( k_H \) are the samples that correspond to \( \omega_L \) and \( \omega_H \), respectively. It is easy to see that the mid-band \([\omega_L, \omega_H]\) must be kept unchanged throughout the algorithm.

The solution of the last problem can be achieved solving the following system of equations:

\[
\frac{\partial V}{\partial Q_k} + \sum_{l=k_L}^{k_H} \lambda_l \frac{\partial Q_k}{\partial q_{nl}} (Q_l - \sum_{n=0}^{N-1} \hat{q}_n e^{-i2\pi nl/N}) = 0
\]

\[
Q_k - \sum_{n=0}^{N-1} \hat{q}_n e^{-i2\pi nk/N} = 0, \quad k = k_L, \ldots, k_H
\]

where \( \lambda_l \) are the Lagrange multipliers of the problem. Taking the derivative, inserting \( Q_k \) in the constraint and then the multipliers in equation (18), the following result is
obtained:
\[
Q_k = \begin{cases} 
\frac{\sum_{n=0}^{N-1} G(d_n)q_n e^{-j2\pi kn/N}}{\sum_j G(j)d_j/N}, & k \notin [k_L,k_H]; \\
\hat{Q}_k, & k \in [k_L,k_H].
\end{cases}
\]

From equations (4.119) and (4.123) it is easy to see that
\[
Q_k = \begin{cases} 
B_k, & k \notin [k_L,k_H]; \\
\hat{Q}_k, & k \in [k_L,k_H],
\end{cases}
\] (4.129)

where \(B_k\) is the discrete Fourier transform of \(b_t\) (4.119). Because \(b_k\) is a nonlinear function of \(q_t\), the problem must be solved as follows:

1. The algorithm is initialized by letting \(q_t = \hat{q}_t\).
2. \(b_t\) and \(B_k\) are computed.
3. The missing low and high frequencies are replaced by \(B_k\).
4. From the inverse Fourier transform, an estimate of the reflectivity is calculated. The norm \(V\) is also evaluated to check convergence and a new iteration starts in step 2.

The algorithm, as we can see, is very simple and can be efficiently coded used the FFT.

The subroutine to run FMED

We have developed a FMED library to reconsruct missing low and high frequencies after zero-phase deconvolution\(^5\).

The subroutine \texttt{fmed} is called as follows:

\begin{verbatim}
subroutine FMED(s,ns,fl,fh,dt,nfft,cw1,cw2,iter_max,th,io,fmax)
\end{verbatim}

where the following parameters are provided to the subroutine:

\(s\) (n1) A seismic trace of \(ns\) samples. The trace is a band-limited version of the reflectivity \(f1,fh\) min and max frequency that define the band-limited reflectivity (in Hz).
\(dt\) sampling rate in sec.

\(^5\)Let me know if you are interested in a copy of this library
Figure 4.31: True reflectivity (top), band-limited reflectivity (center), and inverted reflectivity using FMED (bottom)

\texttt{nfft} Number of freq. samples for the fft  
\texttt{cw1(nfft),cw2(nfft)} pre-computed tables of exponentials for the FFT. See subroutine \texttt{fftcw}.

\texttt{iter\_max} Maximum number of iterations ($\leq 10$)

\texttt{th} Values of reflectivity below \texttt{th} are set to zero, if you doubt \texttt{th}=0.

\texttt{io} If \texttt{io.eq.0} the algorithm completes all the high frequencies up to the the Nyquist frequency. If \texttt{io.ne.0} the algorithm will complete the high frequencies up to \texttt{fmax}.

\texttt{fmax} maximum frequency to complete when \texttt{io.ne.0}
4.8 References


