Chapter 1

Introduction

1.1 Exploration seismics

The object of exploration seismics is obtaining structural subsurface information from seismic data, i.e., data obtained by recording elastic wave motion of the ground. The main reason for doing this is the exploration for oil or gas fields (hydro-carbonates). In exploration seismics this wave motion is excitated by an active source, the seismic source, e.g. for land seismics (onshore) dynamite. From the source elastic energy is radiated into the earth, and the earth reacts to this signal. The energy that is returned to the earth's surface, is then studied in order to infer the structure of the subsurface. Conventionally, three stages are discerned in obtaining the information of the subsurface, namely data *acquisition*, *processing* and *interpretation*.

In *seismic data acquisition*, we concern ourselves only with the data gathering in the field, and making sure the data is of sufficient quality. In seismic acquisition, an elastic wavefield is emitted by a seismic source at a certain location at the surface. The reflected wavefield is measured by receivers that are located along lines (2D seismics) or on a grid (3D seismics). After each such a *shot record* experiment, the source is moved to another location and the measurement is repeated. Figure 1.1 gives an illustration of seismic acquisition in a land (onshore) survey. At sea (in a marine or offshore survey) the source and receivers are towed behind a vessel. In order to gather the data, many choices have to be made which are related to the physics of the problem, the local situation and, of course, to economical considerations. For instance, a choice must made about the seismic source being used: on land, one usually has the choice between dynamite and vibroseis; at sea, air guns are deployed. Also on the sensor side, choices have to be made, mainly with respect to their frequency characteristics. With respect to the recording equipment, one usually does not have a choice for each survey but one must be able to exploit its capabilities as much as possible.
Figure 1.2 shows two raw seismic recordings, made on land and at sea. The land shot record (Figure 1.2a) shows immediately the strong noisy events, which are referred to as ground-roll or surface waves (they propagate along the surface). The reflection events (i.e. more flat events at e.g. 0.9 and 1.6 seconds) are hidden by this strong noise. The

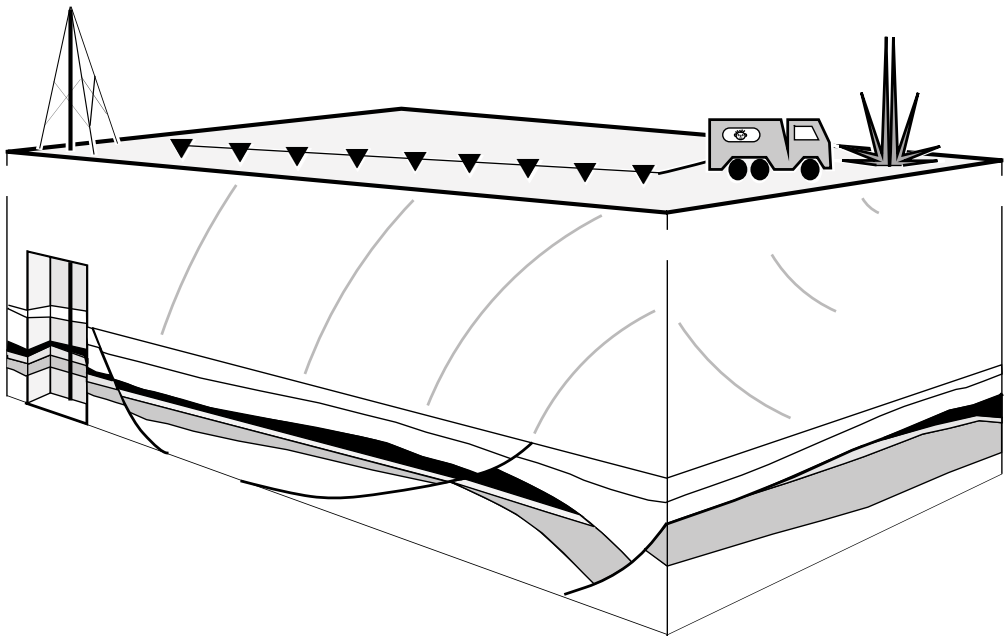


Figure 1.1: Seismic acquisition on land using a dynamite source and a cable of geophones.

marine shot record (Figure 1.2b) is more clean, as we measure in a water layer, which is a good conductor for sound. However, shear waves cannot propagate through water, and are therefore not measured for marine data. Note that the first 0.4 seconds of data are empty (except for the direct wave from source to receivers), meaning that the water bottom is approximately 300 meters depth (water velocity is approximately 1470 m/s). Note also in this shot record the strong multiples of the water bottom reflection at 0.8, 1.2, 1.6 and 2.0 seconds. This indicates a major problem in marine data: the surface reflects all energy back into the subsurface (reflection coefficient is $\approx$-1, which produces a large amount of multiply reflected events). In fact almost all events we see below 0.8 seconds are due to multiples.

So far, the choices are made about the hardware, but for a geophysicist the choice of the parameters of the survey are of prime importance. Relevant questions are such as: where do we want to put our shot positions, and where our sensor positions? How often do we want to shoot? What are the distances between the sensors themselves? These are questions which are considered in seismic survey design.

In *seismic processing,* we want to manipulate our gathered data such that we obtain an accurate image of the subsurface. To do this properly, we have to understand the physical processes that are involved in seismic experiments. For instance, the seismic source puts a certain signal (i.e. the source wavelet) into the earth and the structure of the subsurface does not depend on the signal we put in. Therefore, we have to remove this source signal from our data before we start the imaging of the subsurface. This process is called signature deconvolution. Another undesired effect for land data is the surface wave (see also Figure 1.2a). Since they travel along the earth's surface they do not contain any information of

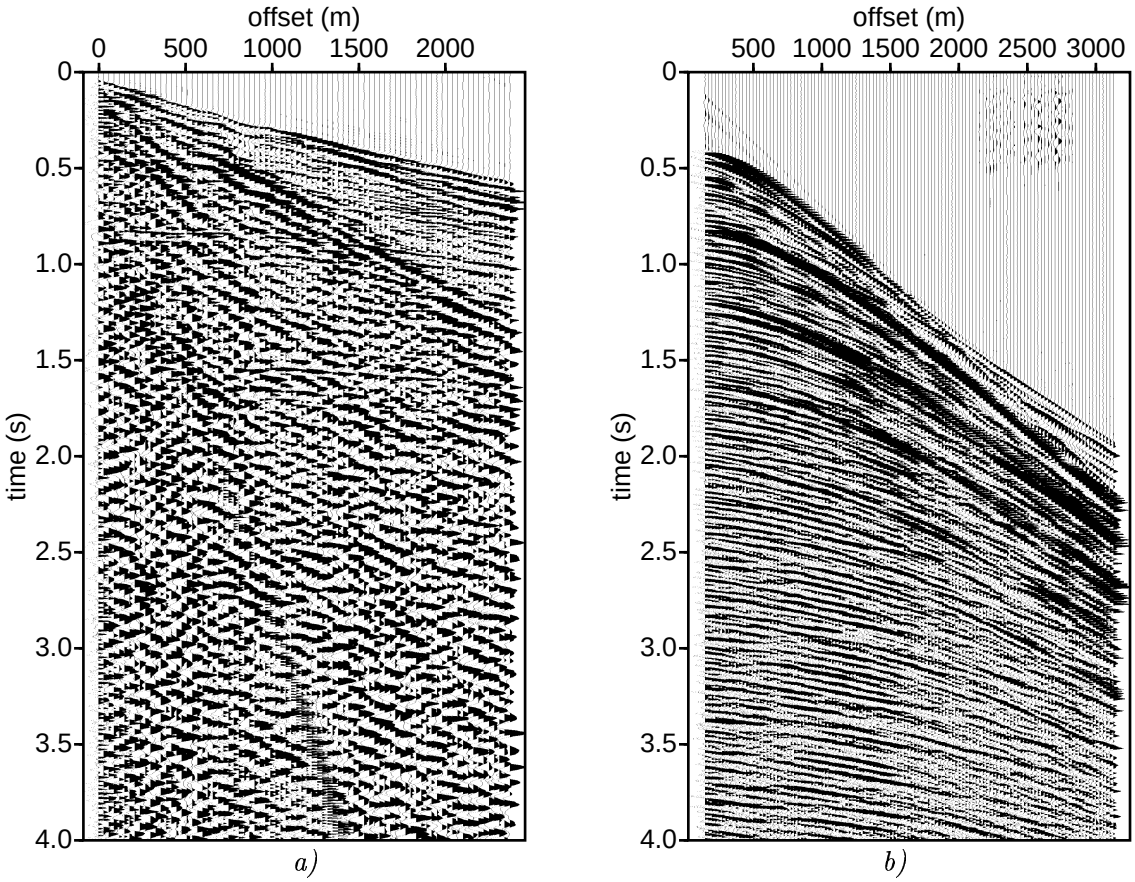


Figure 1.2: Raw seismic measurements. a) Shot record from a land survey. b) Shot record from a marine survey.

the deeper interior of the earth. We would like to remove these events by filtering, often done with two-dimensional filters, so-called (f, k_x) filtering. Again for land data, there is an effect of surface and near-surface topography which can have a tremendous effect on the total response as the near-surface conditions can vary strongly from location to location. This effect is corrected for in a so-called static correction procedure. As mentioned, for marine data the multiples (related to the water surface) are a major problem, which should be removed or suppressed in one of the first processing steps.

So far, we considered quite deterministic aspects in the seismic experiment. However, an important feature of seismic data is that the noise level is too high to obtain an accurate picture of the subsurface using only one shot record experiment. Therefore, in seismic acquisition we make a multiple-fold coverage of the subsurface in order to be able to increase the signal to noise ratio. It is because of this noise that already in our field procedure we take account of the fact that we want to add signals together during processing (stacking).

The final stage of the seismic procedure is the *seismic interpretation.* In general, it is in this stage that we translate the seismic information into the information we hoped

to obtain when we started out to do our survey. Often, this is geological information, but can also be information to a civil engineer, or someone from another field. When we look at a completely processed section from the point of view of the geologist, he hopes to see structures or features which he can relate to geological phenomena, such as faults, anticlines, etc. From the way the structures are built up, he hopes to be able to infer with what kind of rocks we are dealing, in what kind of environment and in which times the rocks were formed, which is all a very non-unique process. In general, we cannot infer any geology from seismics alone: too many different rocks and environments can give rise to (almost) the same seismic response. We need a lot of external geological information in order to make the interpretation less non-unique.

1.2 Structure of lecture notes

We start these lecture notes with some basic notions from signal processing, such as the Fourier transform, Nyquist criterion, aliasing, discretisation, etc. This is done in Chapter 2. We assume the reader has some acquaintance with these concepts such that we can keep the discussion rather short.

In the next chapter we will discuss a very basic processing sequence, which makes the reader acquainted with backbone of the more complete seismic processing sequence. In this basic processing sequence we first discuss the common-mid point concept, why we use it and how it is obtained. We then discuss what methods can be used to determine seismic velocities which are needed when we are adding traces to increase the signal to noise ratio. As said earlier, this latter process is called stacking. Finally, we discuss the simplest scheme of migration which takes account of the fact that we are dealing with a wave phenomenon, and this entails that we have to remove in some way these wave effects in order to arrive at a seismic section which looks similar to a real vertical slice through the subsurface. The migration, resulting in a time image, needs to be converted to depth using a time-to-depth procedure, which is the last section in this chapter.

In Chapter 4 we will deal with more extended parts of the seismic processing sequence which can or have to be applied in order to arrive at a decent seismic section. Amongst these processes are static corrections, deconvolution, (f, k_x) filtering, Dip Move-Out (DMO) and depth migration. We will make an effort to discuss all these subjects in a homogeneous manner.

Chapter 2

Basic signal analysis

2.1 Introduction

In this chapter we will discuss some basic notions from signal analysis. As the basis for signal analysis we use the Fourier transform. We shall define it here. In applications on the computer we do not deal with a continuous transform but with a discrete one, which introduces some special effects such as aliasing. We shall not go into much detail and assume the reader has some acquaintance with signal processing. As an extension of the one-dimensional Fourier transform we will also discuss the two-dimensional transform, which in the seismic case is a transformation with respect to time and space. We will put some more effort in looking at specific features in this domain. As applications we will look at some filtering techniques, not only in the time-transformed variable, but also in the space-transformed variable. Already in seismic data acquisition, we apply both kind of filters so, when processing, we must be aware of this.

2.2 The Fourier transform

Definitions

The Fourier transform is the transformation of a function into weights of sines and cosines of certain frequencies. In this transformation the sines and cosines are the basis functions into which the function is decomposed. Let us consider a function *g(t)* in which t denotes time and transform this function into the transform-variable domain, the frequency-domain, by:

$$G(f) = \int_{-\infty}^{+\infty} g(t) \exp(-2\pi i f t)\, dt, \tag{2.1}$$

in which i is the imaginary unit, i.e, $i = \sqrt{(}-1)$, and f is the transform variable frequency. In our notation, we use capital letters to denote the fact that the function is represented in the frequency domain.

The inverse transform, the reconstruction from the time signal out of its frequency

components, can be expressed as:

$$g(t) = \int_{-\infty}^{+\infty} G(f) \exp(2\pi i f t) df. \tag{2.2}$$

Convolution theorem

We will not go into detail about all the properties of the Fourier transform; for that we would like to refer to the standard book of [Bracewell, 1978]. However, we would like to mention a few properties which will be used in these notes. One very important one is the convolution theorem, which states that a convolution in the time domain is a multiplication in the Fourier domain. Mathematically, when we convolve a function *h(t)* with another function *g(t)*, we obtain a multiplication of the spectra of $h(t)$ and $g(t)$, i.e.:

$$\mathcal{F}_t \left(\int_{-\infty}^{+\infty} h(t') g(t - t') dt' \right) = H(f) G(f), \tag{2.3}$$

in which $\mathcal{F}_t$ denotes the Fourier transform with respect to the subscript variable t. Similarly, a convolution in the Fourier domain, is a multiplication in the time domain. We shall frequently make use of this property.

To prove this, we first write out the Fourier transform of $h(t)$ and $g(t)$ in full:

$$\mathcal{F}_t[h(t) * g(t)] = \int_{\infty}^{\infty} \left[\int_{\infty}^{\infty} g(a) h(t - a) da \right] \exp(-2\pi i f t) dt. \tag{2.4}$$

Then the integrand is multiplied by $1 = \exp(2\pi i f a) \exp(-2\pi i f a)$, and the order of integration is changed:

$$\mathcal{F}_t[h(t) * g(t)] = \int_{\infty}^{\infty} g(a) \left[\int_{\infty}^{\infty} h(t - a) \exp\{-2\pi i f (t - a)\} dt \right] \exp(-2\pi i f a) da, \tag{2.5}$$

in which the Fourier transform of $h(t)$ may now be recognized in the square brackets, and thus:

$$\mathcal{F}_t[h(t) * g(t)] = \int_{\infty}^{\infty} g(a) H(f) \exp(-2\pi i f a) da. \tag{2.6}$$

$H(f)$ may now be taken outside the integral as it is independent of the integration variable a. Then, the resulting integral can be recognized at the Fourier transform of g(t). This completes the proof.

A simple example : boxcar-function

A simple application of this property is windowing, or bandpass filtering. In general, in the time domain it is called windowing, while in the frequency domain windowing is called bandpass filtering. Let us consider the window $h(t)$ defined as follows:

$$h(t) = \begin{cases} 0 & \text{if } t \leq -a \\ 1 & \text{if } -a < t < a \\ 0 & \text{if } t \geq a. \end{cases} \tag{2.7}$$

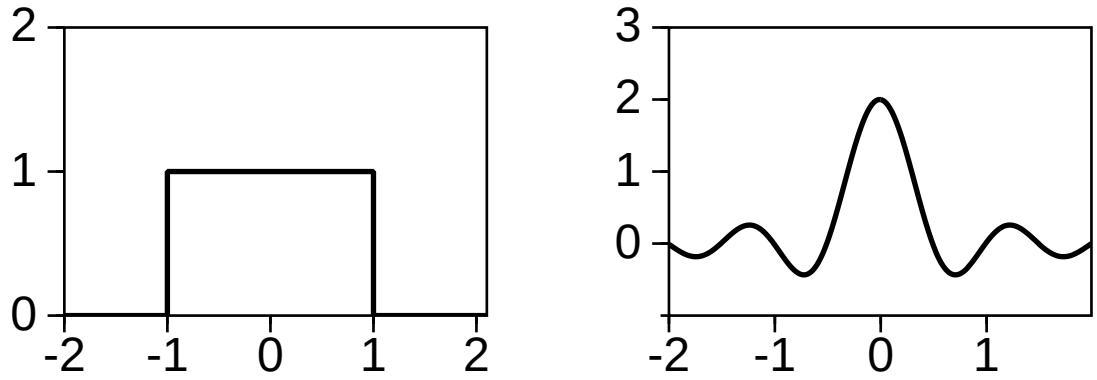


Figure 2.1: A window function in the time domain (left) corresponds to a sinc function in the frequency domain (right); the width of the window is chosen as 2, so $a = 1$.

When we apply this window in the time domain (i.e. multiply with the window function in the time domain), we convolve with its Fourier spectrum in the frequency domain. We therefore have to calculate the Fourier transform of the window function, which is:

$$H(f) = \int_{-a}^{a} \exp(-2\pi i f t) dt \tag{2.8}$$

$$= \frac{\exp(-2\pi i f t)}{-2\pi i f}\Big|_{t=-a}^{a} \tag{2.9}$$

$$= \frac{\sin(2\pi f a)}{\pi f}. \tag{2.10}$$

This function is drawn in figure 2.1. It is a scaled version of the so-called sinc-function and has many side-lobes with large amplitudes. When we multiply with the window function $h(t)$ in the time domain, and thus convolve with function $H(f)$ in the frequency domain, we will obtain a result which is not optimal. This shows that when we apply a window, we better smooth the edges of the window.

There are many ways to smooth the sides, such as a cosine roll-off (see figure 2.2):

$$h(t) = \begin{cases} 0 & \text{if } t < t_1 \\ \cos^2(\frac{t_2 - t}{t_2 - t_1}\frac{\pi}{2}) & \text{if } t_1 < t < t_2 \\ 1 & \text{if } t_2 < t < t_3 \\ \cos^2(\frac{t - t_3}{t_4 - t_3}\frac{\pi}{2}) & \text{if } t_3 < t < t_4 \\ 0 & \text{if } t_4 < t. \end{cases} \tag{2.11}$$

Another much used window is the gaussian:

$$h(t) = \exp(-a(t - t_0)^2) \qquad \text{for } a > 0, \tag{2.12}$$

for which its Fourier transform is also a gaussian. This is depicted in figure 2.3. These relations are important in seismic applications where very often windowing is applied, and thus the effect of the window must be well understood.

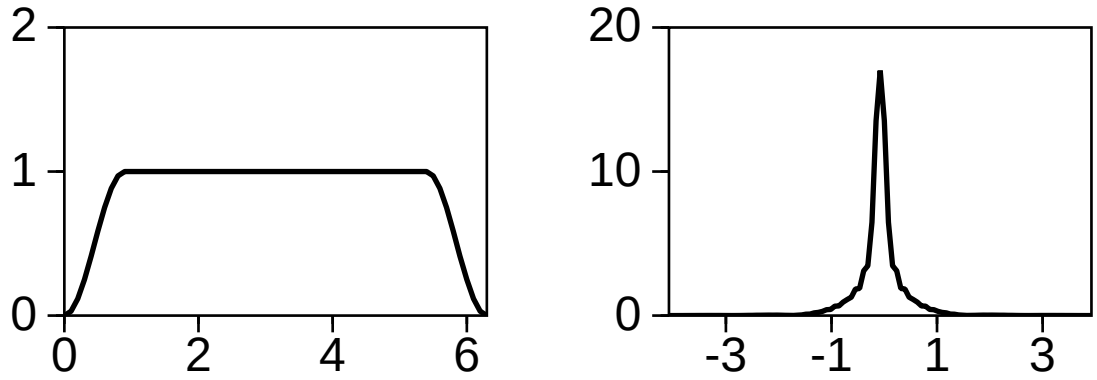


Figure 2.2: Tapering of a window function in the time domain (left) and its effect in the frequency domain (right) for a cosine roll-off.

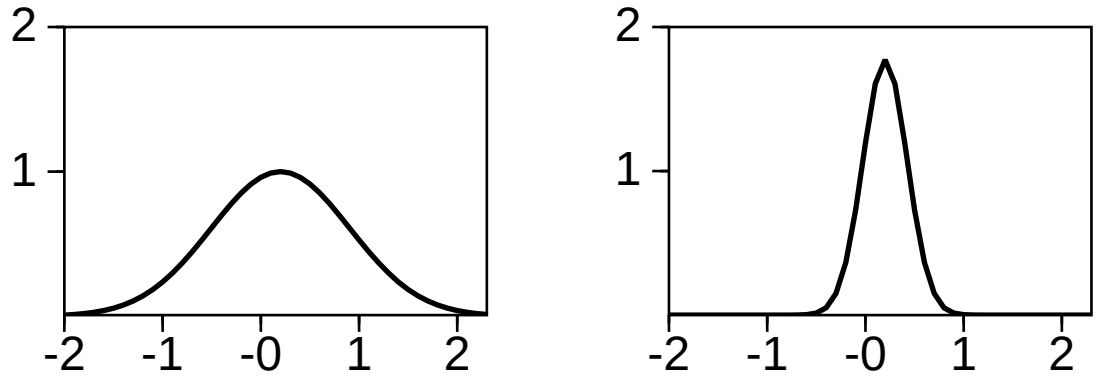


Figure 2.3: Tapering of a window function in the time domain (left) and its effect in the frequency domain (right) for a gaussian tapering.

In this example, we used a simple function to show the effect of filtering, and were able to show the effect of the filter in analytical terms. When we deal with the seismic experiment, the filters can most of the time not be calculated analytically, and we have to determine them in another way. Of course, filtering takes place in many ways, due to many different processes. In the seismic experiment, we put a signal into the earth which has a certain frequency content. A simple example of the spectrum of a dynamite source signature is given in figure 2.4. When we assume that the earth is convolutional, the output is just a multiplication of the spectrum of the earth response with the spectrum of the source signal. But also on the sensor side, we have a sensor response, filtering once more the input signal. The typical spectrum of a sensor on land, a geophone, is given in figure 2.5, which has been taken from [Pieuchot, 1984]. As we can see, the amplitudes below the resonance frequency are damped, due to the coil system of the geophone. It acts as a filter to the data, and the total signal can be regarded as a multiplication of the spectra of the source signal, the earth response, and the geophone response.

In a seismic recording system, we often have other filters available in the system but it is often a matter of choice whether they are activated or not. There is however one filter

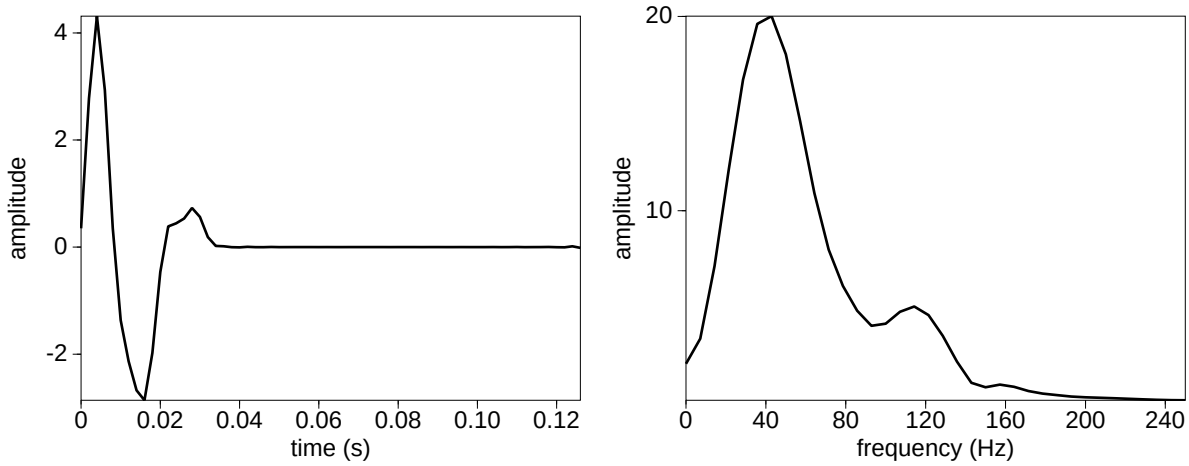


Figure 2.4: A dynamite wavelet in the time domain (left) and in the frequency domain (right).

which is mandatory when we deal with discretised data and that is the anti-alias filter. This high-cut frequency filter is also an analog filter but the reason why we use it will become clear in the next section, when we discuss the discrete Fourier transform.

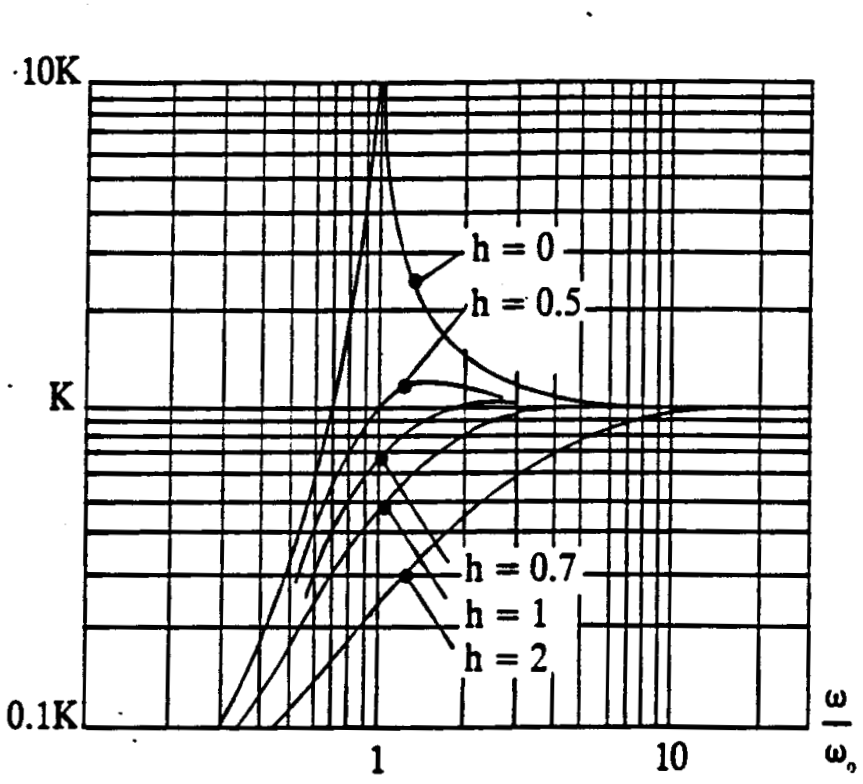


Figure 2.5: A typical geophone response curve (from Pieuchot, 1984).

Correlation

Another property, which is related to the convolution theorem, is one with respect to correlation. Let us first define the time reverse $b_{\rm rev}(t)$ of a signal $b(t)$ by:

$$b_{\rm rev}(t) = b^*(-t), \tag{2.13}$$

where $b(t)$ is allowed to be complex and the asterisk denotes complex conjugate. Normally, of course, we deal with real time signals. However, by allowing these signals to be complex it is easier to see their symmetry properties. When we apply a Fourier transformation to $b(t)$, and take the complex conjugate of each side, we obtain:

$$\begin{aligned} B^*(f) &= \left[\int_{-\infty}^{\infty} b(t)\exp(-2\pi i f t)dt\right]^* \\ &= \int_{-\infty}^{\infty} b^*(t)\exp(2\pi i f t)\,dt \\ &= \int_{-\infty}^{\infty} b^*(-t)\exp(-2\pi i f t)\,dt \\ &= \mathcal{F}_t[b^*(-t)] \\ &= \mathcal{F}_t[b_{\rm rev}(t)], \end{aligned} \tag{2.14}$$

which is the Fourier transform of $b_{\rm rev}(t)$. When we define the cross-correlation function ϕ_{ab} of two signals $a(t)$ and $b(t)$ as:

$$\phi_{ab}(\tau) = \int_{-\infty}^{\infty} a(t)b^*(t-\tau)dt, \tag{2.15}$$

then this can be recognized as a convolution of $a(t)$ with the time reverse of $b(t)$:

$$\begin{aligned} \phi_{ab}(\tau) &= \int_{-\infty}^{\infty} a(t)b_{\rm rev}(\tau - t)\,dt \\ &= a(\tau) * b_{\rm rev}(\tau) \end{aligned} \tag{2.16}$$

So then we obtain in the frequency domain:

$$\Phi_{ab}(f) = A(f)B^*(f). \tag{2.17}$$

The function $\Phi_{ab}(f)$ is known as the *cross-spectrum*. It can be seen that the correlation of $a(t)$ with $b(t)$ is not necessarily the same as the correlation of $b(t)$ with $a(t)$, but we still have the symmetry that:

$$\phi_{ab}(\tau) = \phi_{ba}(-\tau) \tag{2.18}$$

and thus

$$\Phi_{ab}(f) = \Phi^*_{ba}(f). \tag{2.19}$$

We have defined the cross-correlation, but in the same we can define the autocorrelation, when we substitute $a(t)$ for $b(t)$ in the above definitions. A special characteristic of the autocorrelation with respect to the cross-correlation is that the autocorrelation exhibits symmetry in time when the time signal is real. This is due to the fact that $\Phi_{aa}(f) = A(f)A^*(f)$ is real and consequently its inverse Fourier transform symmetric around $t = 0$.

2.3 The discrete Fourier transform

The discussion in the last section was focussed on the continuous representation of the Fourier transform with some applications. So what they had in common is that the filters were considered to be analog. Examples of filters are springs and dashpots in mechanical terms, or coils, capacitors and resistances in electrical terms. When we do seismic processing, we work on discretised data right from the start, because we need to use the computer to work on the data, and computers are digital these days. When using digital computers, filters can also be defined and are then by definition digital filters. But before discussing digital filters, we would first like to discuss some basic elements of digital signal analysis.

Again, we shall not go in much detail into how the discretised version of the Fourier transform is derived, but summarize the most important results from it. In appendix A we give the full derivation of the discrete form of the Fourier transform. Discretising means that we choose certain time instances, and certain frequency instances to sample our continuous signal, thus we take times $t = k\Delta t$ and $f = n\Delta f$. The most important consequence of discretising the data in one domain, is that the data in the other domain becomes *periodic*. Thus discretising in the time domain means that the spectrum becomes periodic, and vice versa.

Another important aspect is that in real life we cannot sample a signal until times at infinity. Therefore we can only measure the signal with a finite amount of samples. Therefore we always make an error, the error becoming smaller when we take more samples. This means that we can reconstruct the continuous signal only up to a certain degree *but never exactly.* Only when we use an infinite number of samples can we reconstruct our signal perfectly.

With these consequences in mind, we can write the discrete form of the Fourier transform as :

$$G_l = \Delta t \sum_{k=0}^{K-1} g_k \exp(-2\pi i k l/K) \qquad \text{for } l = 0, 1, ..., K-1, \tag{2.20}$$

in which summation is over the time samples k, and the spectral coefficients are given for frequency samples l. Also, K is the number of samples used to sample the continuous function, and Δt is the sampling rate. The inverse Fourier transform is then given by:

$$g_k = \Delta f \sum_{l=0}^{K-1} G_l \exp(2\pi i k l/K) \qquad \text{for } k = 0, 1, ..., K-1, \tag{2.21}$$

in which we now sum over the frequency components, and Δf is the sampling rate of the frequencies. A simple relation connected to this Fourier pair is that we have the equality:

$$K\Delta t\Delta f = 1. \tag{2.22}$$

When we sample continuous data, we must make sure we take enough samples, that is, we must take at least two samples for the maximum frequency in the data. Put in another

way, with a chosen Δt the maximum frequency which is represented properly is:

$$f_N = \frac{1}{2\Delta t}, \tag{2.23}$$

where the maximum frequency is called the Nyquist frequency, denoted by f_N.

Discretising in such a way does not seem to be a problem, but there is one snag: how do we know on beforehand what our highest frequency in our signal will be? We do not know and therefore we have to filter the continuous data before we digitize it. This means that we must include a high-cut filter that makes sure that the signal level is damped below the noise level at the Nyquist frequency and then the data can be digitized properly. This is an analog filter, commonly called the alias filter. This filter has always been applied to the data when the data arrives at the processing centre.

2.4 The spatial Fourier transform

For years, seismic recordings have taken place at discrete receiver stations, either with single geophones or with geophone patterns. However, it has not always been realized by practicing seismologists that we were sampling in the horizontal direction where we must also take account of the Nyquist criterion. This criterion is the same as for sampling in time, but only the parameters must be interpreted differently. For sampling in one spatial coordinate, say x, we define the following transform:

$$\tilde{G}(k_x) = \int_{-\infty}^{+\infty} G(x) \exp(2\pi i k_x x) dx, \tag{2.24}$$

where k_x denotes the spatial frequency and now the forward transform from x to k_x has the opposite sign in the exponential compared to the time-frequency case. Note also that we put a tilde $\tilde{}$ on top of a quantity to denote that it is represented in the k_x domain. The original x–domain signal can be synthesized by integrating over all k_x components, i.e.:

$$G(x) = \int_{-\infty}^{+\infty} \tilde{G}(k_x) \exp(-2\pi i k_x x) dk_x. \tag{2.25}$$

This type of transformation can thus be applied to any kind of data sampled in space. For instance, this transformation can be applied to data which represent the topography of the surface. This type of transformation is also often applied to gravimetric and magnetometric data, which represent the gravity and magnetic field at the surface, respectively. It should therefore be remarked that this transformation can be applied to many data, not at all from a geophysical nature, but for instance for image processing, economics, or whatever else.

Again, we have only defined the spatial Fourier transform for continuous signals but in seismics we deal with discrete samples in space so we have to make the transforms suitable for discrete signals as well. This discretisation goes in the same manner as for the temporal Fourier transform (see also appendix A), and we obtain the pair:

$$\tilde{G}_m = \Delta x \sum_{n=0}^{N-1} G_n \exp(2\pi i m n/N) \qquad \text{for } m = 0, 1, ..., N-1 \tag{2.26}$$

$$G_n = \Delta k_x \sum_{m=0}^{N-1} \tilde{G}_m \exp(-2\pi i m n/N) \qquad \text{for } n = 0, 1, ..., N-1, \tag{2.27}$$

where we have N spatial samples and we again have the relation:

$$N \Delta x \Delta k_x = 1. \tag{2.28}$$

The spatial Nyquist criterion is given by:

$$k_N = \frac{1}{2\Delta x}, \tag{2.29}$$

i.e., we must sample the maximum horizontal wavenumber at least two times. The same remarks as for the temporal Fourier transform are valid: discretising in x makes the k_x–spectrum periodic, and vice versa. An example of aliasing in space is given in figures 2.6 and 2.7 which can be recognized by the "folding" back of the events in the (f, k_x) domain if the dip in the space-time domain becomes too large. Also, we have to avoid steep slopes when windowing in order to avoid side-lobe leakage in the other domain.

2.5 The two-dimensional Fourier transform (in relation to waves)

In the previous sections we discussed the one-dimensional Fourier transform, but when we deal with seismic data, we deal with two-dimensional data. At one sensor position, we record the signal in time, and we record at many sensor positions, so we sample the time and the horizontal space coordinate. In obtaining the seismic image of the subsurface, we heavily depend on the two-dimensional nature of our data. Some operations take place directly in the (t, x) domain, but other processes are done in the so-called (f, k_x) domain where the f and k_x are the transform variables with respect to time t and space x, respectively. As well known, f is the temporal frequency, and k_x is the spatial frequency in the horizontal direction where we now specifically denote the subscript x to denote that we are sampling in one horizontal Cartesian coordinate direction, namely x. The two-dimensional Fourier transform of signal $s(x, t)$ is defined as:

$$\tilde{S}(k_x, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(x, t) \exp[2\pi i (k_x x - f t)] dt dx. \tag{2.30}$$

For seismic applications, the two-dimensional transformation is not a simple representation since the quantities on the axes are different: space and time. But, space and time are related, due to the fact that we deal with a wave phenomenon: a wavefront is propagating through the earth, and we record it at the earth's surface. Therefore, the quantity k_x should be interpreted also in terms of waves. For a sinusoidal signal, the frequency f is the number of cycles per unit time; the same is valid for the spatial frequency k_x. In the same way, k_x is the number of wavelengths per space unit, so k_x is related to the horizontal wavelength λ_x by:

$$k_x = \frac{1}{\lambda_x}. \tag{2.31}$$

When we substitute the relation:

$$c_x = f\lambda_x, \tag{2.32}$$

in which c_x denotes the wave speed in the x–direction, we obtain for the spatial wavenumber k_x :

$$k_x = \frac{f}{c_x}. \tag{2.33}$$

Let us look at the following example to illustrate some of the above. Suppose we have a linear event in x-t as depicted in figure 2.7. In general, the event can be defined as:

$$s(x,t) = s(t - x/c_x), \tag{2.34}$$

in which $s(t)$ is some characteristic function for the source signal. Applying a two-dimensional Fourier transform, i.e. a Fourier transformation with respect to time as well as the space-coordinate x :

$$\tilde{S}(k_x, f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} s(t - x/c_x)\exp[2\pi i(k_x x - ft)]dtdx. \tag{2.35}$$

Putting $\tau = (t - x/c_x)$, we have that $dt = d\tau$, at $t = -\infty, \tau = -\infty$, and at $t = +\infty, \tau = +\infty$. From this it follows that:

$$\tilde{S}(k_x, f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} s(\tau)\exp[-2\pi i f\tau]\exp[-2\pi i x(f/c_x - k_x)]d\tau dx. \tag{2.36}$$

The latter expression can be considered as the product of two integrals, the first one being a Fourier transform with respect to the variable τ only, i.e.:

$$S(f) = \int_{\tau=-\infty}^{\infty} s(\tau)\exp[-2\pi i f\tau]d\tau, \tag{2.37}$$

and the second integral being a Fourier transform with respect to the variable x only, so that the total expression (2.36) can be written as:

$$\tilde{S}(k_x, f) = S(f)\int_{x=-\infty}^{\infty} \exp[2\pi i x(k_x - f/c_x)]dx. \tag{2.38}$$

Now, because the delta-function is the inverse transform of the constant value 1, we have:

$$\delta(k_x) = \int_{x=-\infty}^{\infty} \exp[2\pi i x k_x]dx. \tag{2.39}$$

The integral transform in (2.38) can be recognized as a delayed delta function of k_x. Therefore it can be expressed as:

$$\tilde{S}(k_x, f) = S(f)\delta(k_x - f/c_x) \tag{2.40}$$

This expression can be recognized as the product of a complex frequency spectrum $S(f)$ (which is a function of the frequency f only!) and a "razor-blade"–like two-dimensional δ–function: $\delta(k_x - f/c_x)$, which is zero everywhere in the $k_x - f$ plane, except on the line $k_x - f/c_x = 0$, where it assumes infinite amplitude. Note also, that the dip of the event

in the (f, k_x) domain is reciprocal to the one in the time domain. For the linear event in the (t, x) domain, we obtain:

$$\frac{\Delta t}{\Delta x} = \frac{1}{c_x}, \tag{2.41}$$

while in the (f, k_x) domain we obtain:

$$\frac{\Delta f}{\Delta k_x} = c_x. \tag{2.42}$$

In the above example we saw that the wave $s(t - x/c_x)$ mapped onto the function $S(f)\delta(k_x - f/c_x)$ in the (f, k_x) domain. But as we have seen in the definition of the spatial wavenumber (eq.(2.33)), a frequency component f is included. This is the reason why often a substitution of $k_x = fp$ is done, in which p is called the horizontal slowness or ray parameter. The latter name is often used because it relates to the parameter that is constant across an interface when using Snell's law : $1/c_x = \sin\theta/c$. When using $k_x = fp$, the forward transformation (eq.(2.35)) reads:

$$\tilde{S}(fp, f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} s(t - x/c_x)\exp[2\pi i f(px - t)]dtdx. \tag{2.43}$$

where we now recognize the linear event $s(t - x/c_x)$ from above in the exponent. The linear event maps in the (p, f) domain as:

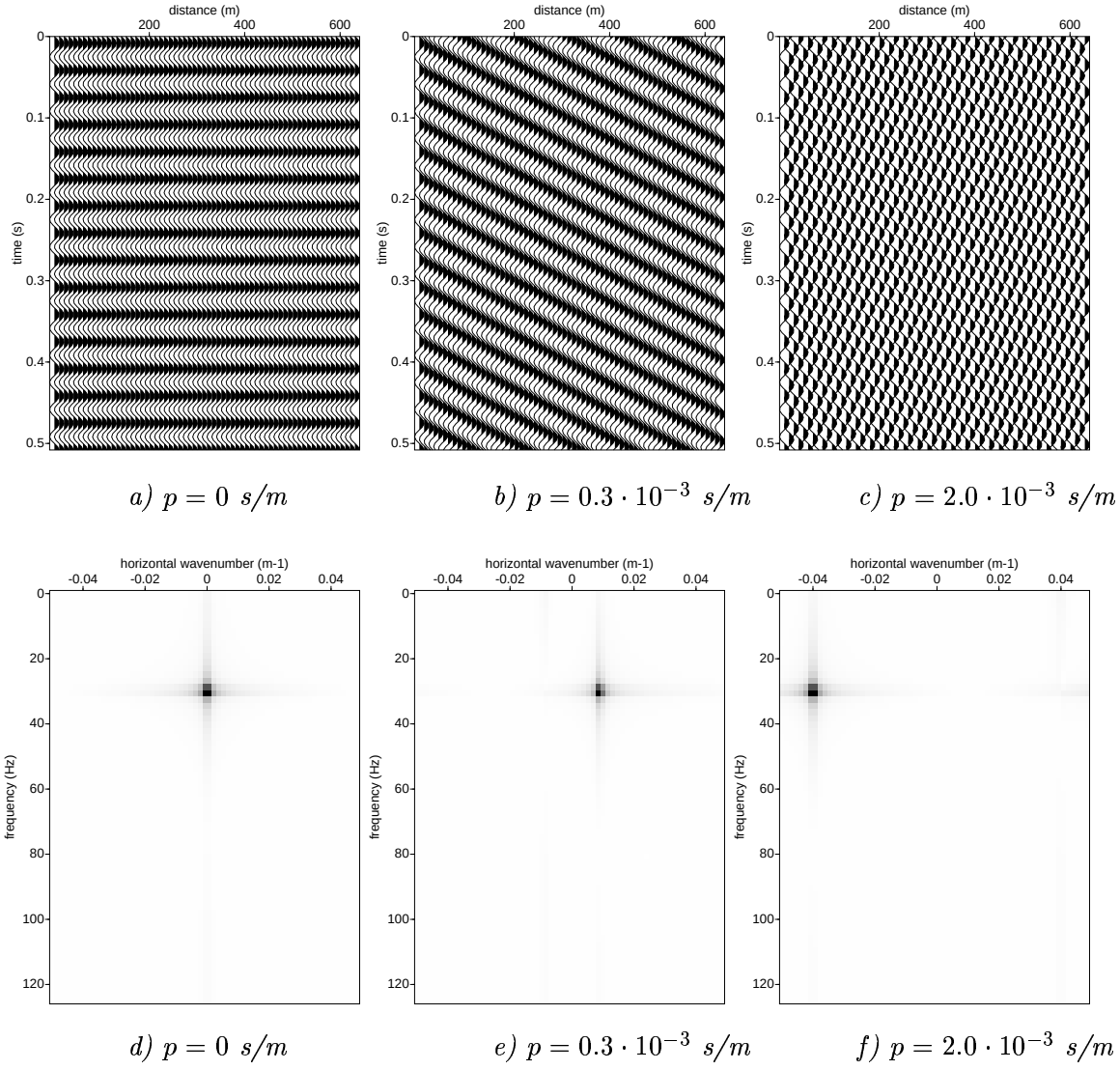
$$\tilde{S}(fp, f) = S(f)\delta(p - 1/c_x) \qquad (f \neq 0) \tag{2.44}$$

In this domain, the wave $s(t-x/c_x)$ becomes only non-zero for the constant ray parameter $p = 1/c_x$. The type of transformation described here is often used in the so-called Radon transformation. The Radon transformation exploits the wave character of our seismic data.

In figure 2.6 three plane waves are displayed with increasing ray parameter, the time and space sampling being 4 ms and 10 m respectively. Clearly, the effect of aliasing can be observed both in the time domain (positive dip looks as negative dip) as well as the wavenumber domain (wrap around along the wavenumber axis).

In figure 2.7 we have created three linear events with different velocities. When we transform the time-axis to the frequency domain, we obtain an interference pattern from which we cannot recognize very much. the last picture gives the (f, k_x) spectrum by Fourier transforming with respect to the horizontal coordinate.

Later on, in chapter 4, we shall go into more detail in this domain, where we can make some nice classification based on its characteristics in this domain. There are also other applications of the (f, k_x) domain. For example, differential operators in the x domain may become simple algebraic operations in the k_x domain for some simple models, making the operation more easy to understand. Also, some operations may be more stable to compute in the k_x domain than in the x domain, or sometimes the other way around. In the next chapters we will see a few applications of this.



a) $p = 0\ s/m$ *b)* $p = 0.3 \cdot 10^{-3}\ s/m$ *c)* $p = 2.0 \cdot 10^{-3}\ s/m$

d) $p = 0\ s/m$ *e)* $p = 0.3 \cdot 10^{-3}\ s/m$ *f)* $p = 2.0 \cdot 10^{-3}\ s/m$

Figure 2.6: Three discretised plane waves of frequency 30 Hz with ray parameter $p = 0, p = 0.3 \cdot 10^{-3}$ and $p = 2.0 \cdot 10^{-3}$ s/m respectively in the time domain (upper part) and in the wavenumber domain (lower part). The time sampling is 4 ms and the spatial sampling is 10 m.

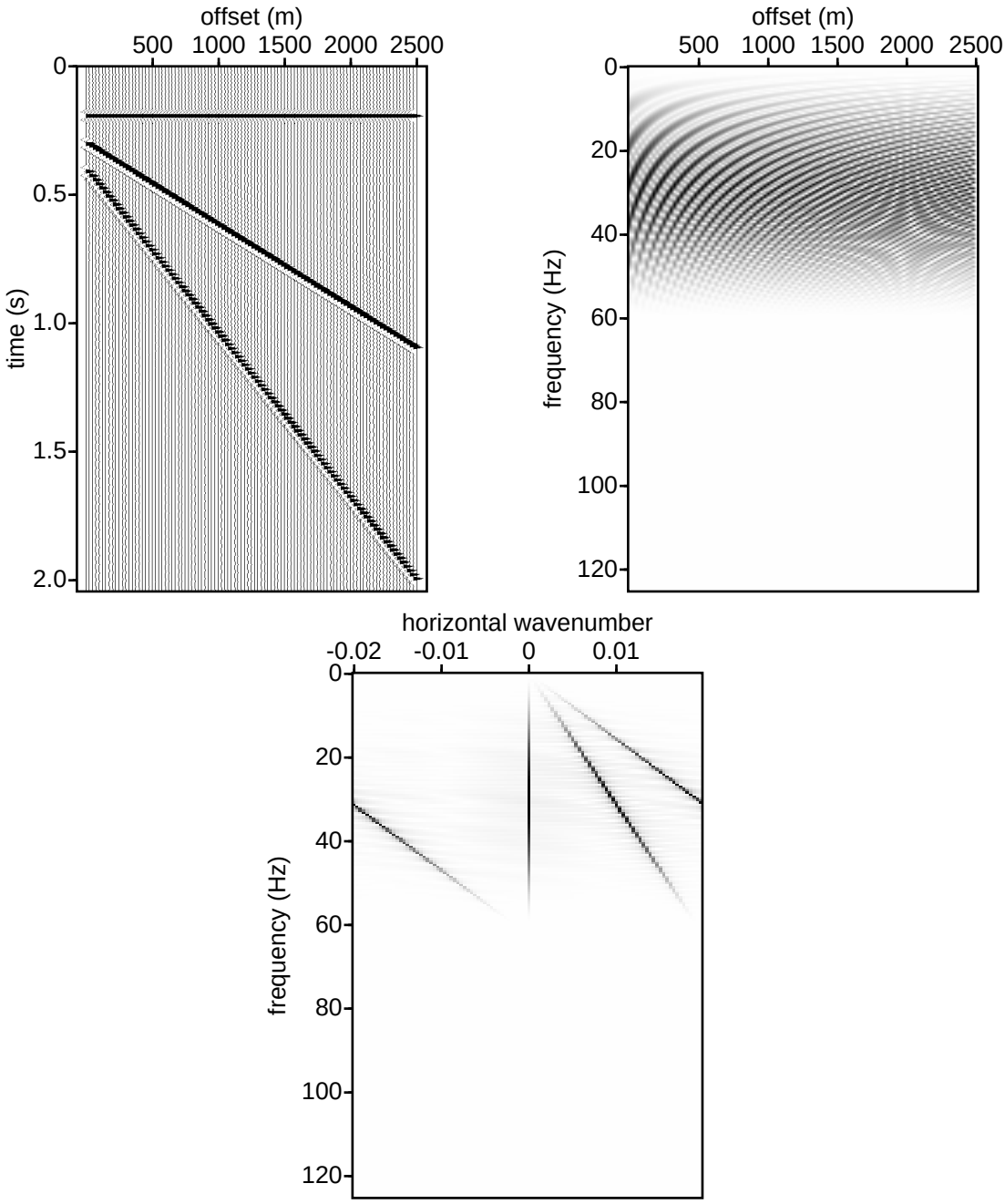


Figure 2.7: A linear event in (t, x) (upper left), (f, x) (upper right) and (f, k_x) (lower).

Chapter 3

A basic seismic processing sequence

3.1 Seismic processing and imaging

Wave propagation versus signal to noise ratio

In seismic processing we are going to manipulate our measured data, such that we obtain an accurate image of the subsurface. In fact the main problem is that the information we measure at the surface is a function of *time*, should be mapped to the correct position in *depth* in the subsurface. This means that reflection energy has to be repositioned, which is called *migration*. In figure 3.1 it is clearly visible that three reflections from totally different points in the subsurface are received at the same geophone position. We can consider two ways of introducing seismic processing to a newcomer.

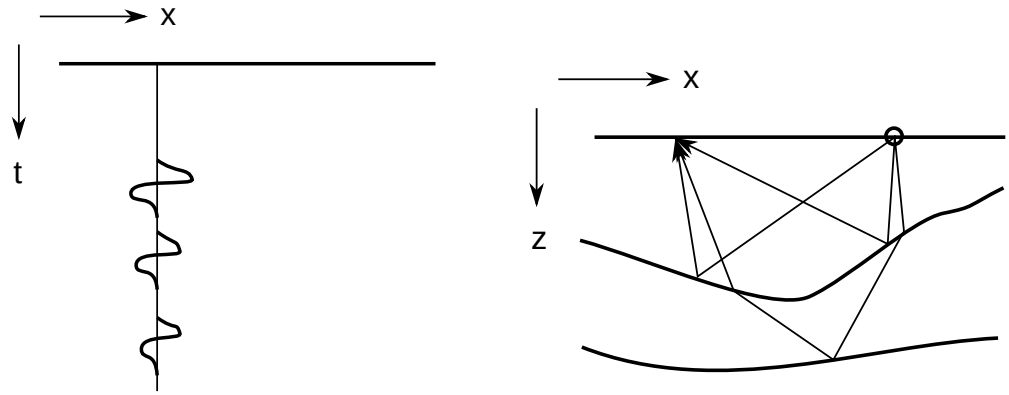


Figure 3.1: Reflections in time (a) and in depth (b).

One is in terms of wave theory. We have to understand the physical processes that are involved all the way from the seismic source, through the subsurface, until the seismic recording instrument. We have to try to obtain only those features which are due to the structure of the subsurface and not related to other features. For instance, we want to know the source signal we put into the earth because then we can remove it from our data later: the structure of the subsurface does not depend on the source signal. In this way we can remove or suppress certain unwanted features in the image we want to obtain.

Another way of introducing seismic processing to a newcomer is more in terms of the image we obtain: signal-to-noise ratio and resolution. In order to see the image we need to have at least a moderate signal-to-noise ratio. We would like this ratio to be as large as possible by trying to suppress unwanted features in the final image. Another aspect of the final seismic image is the resolution: we would like the image to be as "crisp" as possible. As you may know, these two aspects cannot be seen separately. Usually, given a certain data set, an increase in signal-to-noise ratio decreases the resolution (as information is stacked together), and also an increase in resolution (by correctly incorporating wave theory) has normally the consequence that the signal-to-noise ratio gets worse. In seismic processing we would like to obtain the optimum between the two: a good, although not perfect, signal-to-noise ratio with a good resolution.

In these notes we take the view of trying to understand each process in the wave problem, and try to find ways to cope with them. In this way we hope at least to increase the signal-to-noise ratio, perhaps at some costs with respect to resolution. This is perhaps a very important characteristic of raw seismic data: it has a very poor signal-to-noise ratio, and it needs a lot of cleaning up before the image of the subsurface can be made visible. It is along this line that we will discuss seismic processing: trying to understand the physical processes. Sometimes, we will refer to the effect it can have on the total signal in terms of signal-to-noise ratio and resolution.

With seismic processing, we have many physical processes we have to take into account. Actually, there are too many and this means that we must make simplifying assumptions. First, we only look at reflected energy, not at critically refracted waves, resonances, surface waves, etc. Of course, these types of waves contain much information of the subsurface (e.g. the surface waves contain information of the upper layers) but these waves are treated as noise. Also critically refracted waves contain useful information about the subsurface. That information is indeed used indirectly in reflection seismics via determining static corrections, but in the seismic processing itself, this information is thrown away and thus treated as noise. Another important assumption in processing is that the earth is not elastic, but acoustic. In conventional processing, we mostly look at P-wave arrivals, and neglect any mode-conversion to S-waves, and even if we consider S-waves, we do not include any conversions to P-waves. Some elastic-wave processing is done in research environments, but are still very rarely used in production. Money is better spent on 3-D "P-wave" seismics, rather than on 2-D "elastic" seismics; 3-D seismics with three-component sources and receivers are still prohibitively expensive in seismic data acquisition.

As said previously, the conventional way of processing is to obtain an image of the primary P-wave reflectivity, so the image could be called the "primary P-wave reflectivity

image". All other arrivals/signals are treated as noise. As the name "primary P-wave reflectivity" suggests, multiples are treated as noise (as opposed to "primaries"); S-wave are treated as noise (as opposed to P-waves); refractions are treated as noise (as opposed to reflectivity). Therefore, we can define the signal-to-noise ratio as:

$$\frac{S}{N} = \frac{\text{Signal}}{\text{Noise}} = \frac{\text{Primary P-wave Reflection Energy}}{\text{All but Primary P-wave Reflection Energy}} \tag{3.1}$$

It can be seen now that processing of seismic data is to cancel out and/or remove all the energy which is not primary P-wave reflectivity energy, and "map" the reflectivity in depth from the time-recordings made at the surface. In terms of total impulse response of the earth $G(x, y, t)$, we want to obtain that part of the impuls response of the earth which is due to primary P-wave reflections:

$$G(x, y, t) \overset{\text{Processing}}{\rightarrow} G_{\text{primary,P-wave,reflectivity}}(x, y, z) \tag{3.2}$$

Overview of seismic processing steps

In this chapter, we will look at a basic processing sequence in order to see the backbone of a larger sequence. The steps which will be considered here are common to seismic processing of data gathered on land (on-shore) as well as at sea (off-shore). These are: CMP sorting, NMO correction and velocity analysis, stacking, migration and time to depth conversion. Although this is a basic processing sequence, it does not mean that this will always give a good image: on land statics problems can be the largest problem and has to be dealt with separately; also on land we have to deal with the surface waves which are often removed by (f, k_x) filtering; at sea the source wavelet is not always a clean one and one has to cancel this effect via signature deconvolution. Another problem for marine data are the strong multiples that are mainly generated in the water layer. We will certainly discuss these processes but leave them until the next chapter.

3.2 Sorting of seismic data

Common shot and common receiver gathers

When data is shot in the field, we record the shots sequentially. By a (shot) record we mean all the recordings from the sensors for a single shot experiment. Normally, the measurement for one source at one receiver location is called a trace, which is a time series of reflections. It is obvious that for each shot we will order these recordings (traces) by increasing (or decreasing) offset. The offset is defined as the distance from source to receiver. A simple simulated example of such a shot is given in figure 3.2. In this figure on the left hand side the ray paths from source to the receivers of the seismic waves are shown. Note that due to the different velocities in the different layers, the ray paths are bended according to Snell's law. For this record, one shot consists of the explosion from one charge of dynamite (supposed it is measured on land). The data is stored in the recording instrument and then put onto a magnetic tape, record by record.

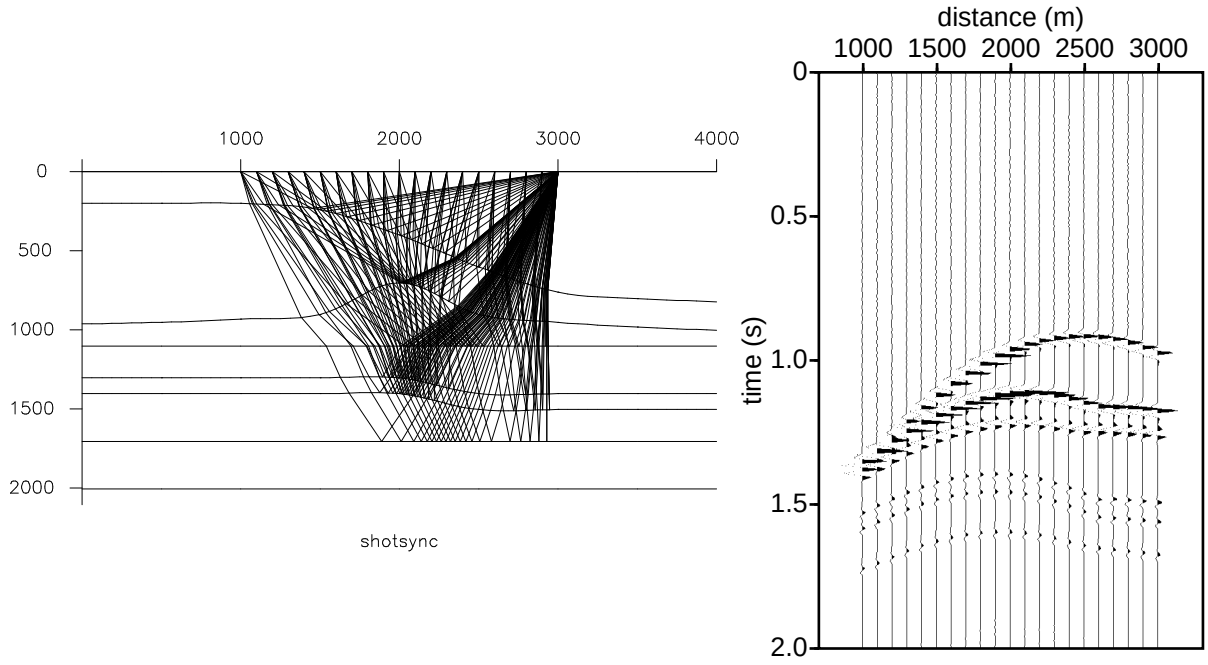


Figure 3.2: Shot gather measurement.

When the next shot is fired, we do the same, record with the instrument and then write the data onto tape. We say that the data is shot ordered. A section as shown in figure 3.2 is commonly called a *common-shot gather,* or common-shot panel: we show the recorded wave field for one shot.

It can be guessed that if we talk about shot ordered data, we could also have receiver ordered data. This is indeed the case. One could get all the shots together, of course in an increasing shot position, belonging to one receiver position. Such a gather is called a *common-receiver gather (or panel).* However, this assumes that during acquisition the same receiver position is covered by different shots. In practice, we often make use of reciprocity: interchanging source and receiver will give exactly the same response (if the directional properties of the source and receiver can be considered identical). In fact figure 3.2 can also be considered as a common receiver gather, where all ray paths from different shots come together at one receiver position.

Why should we need these distinctions? A nice feature about a common-shot gather is to see whether a receiver position has a higher elevation than its neighbors and thus gives an extra time shift in its record. This effect is called "statics". Therefore common-shot gathers are good for detecting geophone statics. In the same way, we can see on common receiver gathers whether a shot was set deeper than the neighboring shot positions, and therefore common-receiver gathers are good for detecting shot statics (see Chapter 4).

Common midpoint gathers

The way of organizing the data in common-shot gathers is just a consequence of the logistics in the field, but for some processing steps it is not a convenient sorting the data.

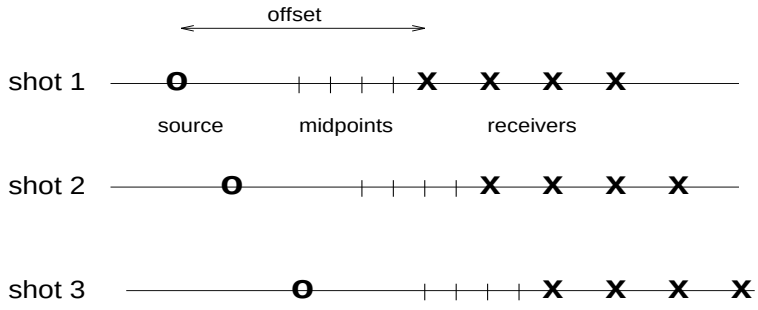


Figure 3.3: Midpoint definition in between sources and receivers.

A commonly used way of sorting the data is in *common-midpoint gathers.* A mid-point is here defined as the mid-point between source and receiver position. An illustration of the mid-point is given in figure 3.3. We gather those traces that have a certain midpoint in common, like in figure 3.3, the record from receiver 3 due to shot 1, and the record from receiver 1 due to shot 2. Once we have gathered all the traces with a common-midpoint (CMP) position, we have to decide how to order these records for one CMP, and the logical choice is to order them by increasing (or decreasing) offset. A gather for one mid-point position with the traces for increasing (or decreasing) offsets is called a common-midpoint gather (or panel). Figure 3.4 shows a CMP gather for the same subsurface model as figure 3.2.

For what reason is the common-midpoint gather convenient? The most important one

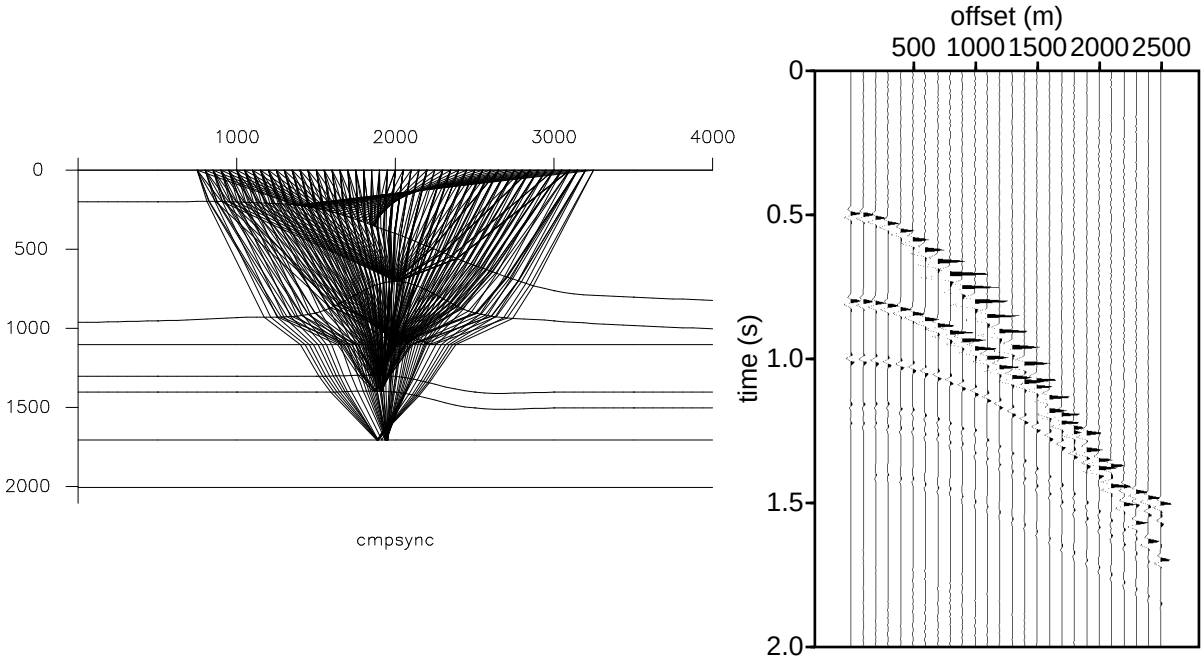


Figure 3.4: Common midpoint gather.

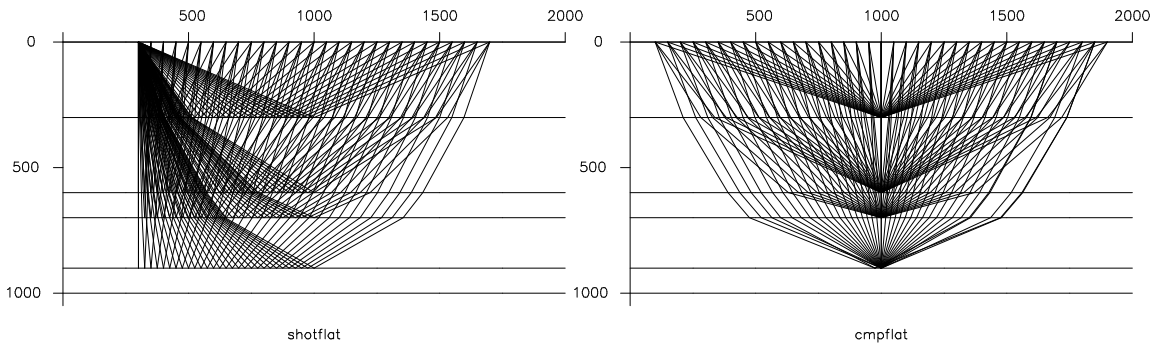


Figure 3.5: Common shot and common midpoint gather for horizontally layered earth.

is for stacking which we shall discuss in one of the next sections. Suppose the earth would consist of horizontal layers as depicted in figure 3.5. Then the geometrical arrival from shot to receiver all reflect right below the midpoint between the source and receiver, and thus the reflection points in the subsurface then only differ in depths. With other words, all the reflections measured at the different offsets in a CMP gather carry information on the same subsurface points (below the midpoint position). If we would make a correction for the offset dependence of the traveltime for each trace, the reflections from the same place would arrive at the same time for all the traces, and thus we could add the traces together to increase the signal-to-noise ratio. This process is called normal move-out (NMO) correction and stacking respectively, as will be discussed later. This argumentation is not valid for common-shot gathers since the reflection points in the subsurface do not coincide for each trace (for a horizontally layered earth). However, for a laterally varying medium, as shown in figure 3.4 the reflections within a CMP gather are coming still from a small region, and the stacking procedure may still give acceptable results. As a result, the resolution of the final image will be limited by the assumption that all energy in a CMP gather comes from the same subsurface points. In chapter 4 we will see that corrections can be included for dipping reflectors, such that this midpoint smear is corrected for (DMO).

Common offset gathers

As can be expected, we can also form a *common-offset gather,* a gather in which we collect all those source-receiver pairs that have a certain offset in common. Usually, we shoot with fixed distances between source and receivers, and so we will have as many traces in our common-offset gather as there are shots, thus often quite a large amount. For the model of figure 3.2 and figure 3.4 the zero offset configuration (i.e. source and receivers at the same positions) is shown in figure 3.6. Note that in the zero offset section the general structures can already be recognized. Common offset gathers are used in prestack migration algorithms since it can give a check on velocities. Migrating a common-offset gather for a small offset should give the same image as a migration of such a gather for a large offset, otherwise the velocity used in the migration is not the right one. The *zero* offset section takes a special place in the seismic processing, as a stacked section is supposed to resemble a zero offset section (see section 3.4 on stacking).

A graph combining all this information is given in figure 3.7. Here we assumed we have

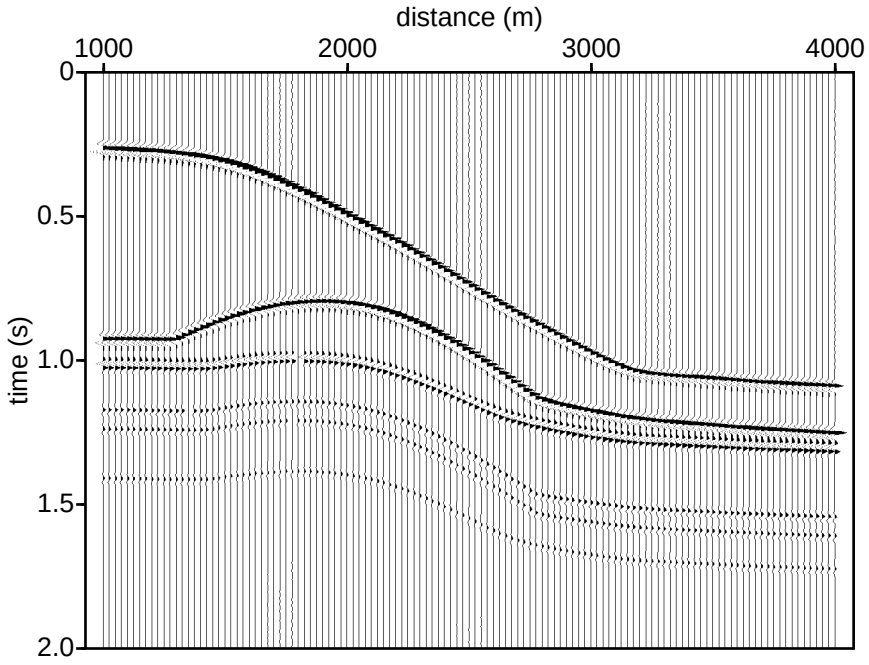


Figure 3.6: Zero offset gather.

recorded along a line in the field, which we call the x-direction. Also, we have assumed that we have 10 receiver positions with the first receiver at the source location (i.e. at zero offset). On the horizontal axis we have plotted the x–coordinate of the source (x_s), while on the vertical axis we have put the x–coordinate of the receiver (x_r). Then, each grid

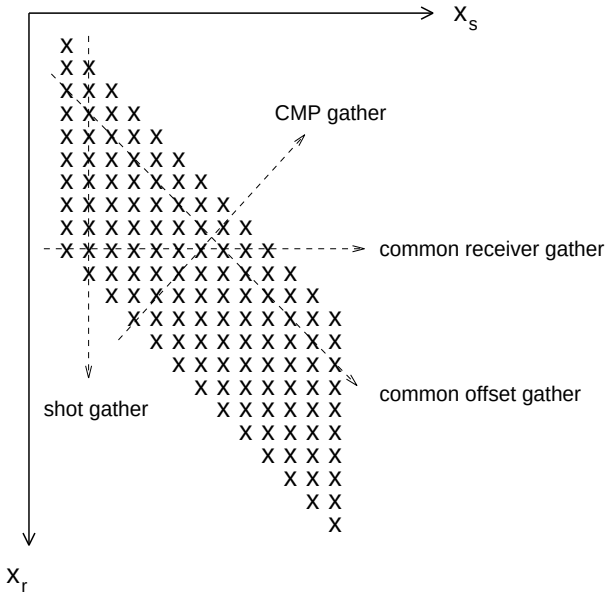


Figure 3.7: Relation between different sortings of seismic data.

point determines where a recording has taken place. In this graph a column represents a common-shot gather, and a horizontal line a common-receiver gather. A common-midpoint gather is given by the line $x_s + x_r = constant$, which is a line at 45 degrees with a negative slope. A common-offset gather is given by the line $x_s - x_r = constant$, which is a line of 45 degrees but now with a positive slope.

What can be noticed in the graph, is that we started out with 10 receiver positions for each shot, while the CMP gather contains only 5 traces. Why that so? This can be seen in figure 3.3. When we shift one source position to the next, we actually shift *two* CMP's because the distance between each CMP is half the source spacing. So a factor two is involved. On the other hand there are twice as many CMP gathers, as the total of traces in the survey is of course constant in any sorting domain. Because each CMP gather has half the number of traces it means that the distance between two traces is twice as large as between to receivers in a shot gather. With other words, each CMP gather has a twice as coarse spatial sampling compared to a common shot gather.

In figure 3.7 we assumed the spacing between the shot positions and the receiver positions were the same but this does not need to be so. This also influences the number of traces in a CMP gather. The number of traces in a CMP gather is called the *multiplicity* or the *fold*. It can be shown easily that the multiplicity M is:

$$M = \frac{N_r}{2\Delta x_s/\Delta x_r}, \tag{3.3}$$

in which N_r is the number of receivers per shot, Δx_s is the spacing between the shot positions, and Δx_r is the spacing between the receivers.

In the above argumentation there is still one assumption made, and that is that the earth is horizontally layered. When the earth is not like that, the reflection points do not coincide any more, see figure 3.4. Still, the results obtained with this assumption are very good, it only gets worse results when the dips of the layers of the earth are becoming steep. We will come to that in section 3.4 when discussing stacking.

3.3 Normal move-out and velocity analysis

NMO curve for single interface

The most important physical parameter needed for obtaining an accurate image of the subsurface, is the velocity of the medium. We record our data at the surface *in time*, and what we wish to obtain is an image of the subsurface *in depth*. The link between time and depth is of course the wave velocity, which varies in the earth from position to position (i.e. the earth is an inhomogeneous medium). Unfortunately, it is not so easy to obtain a good velocity model and this is often an iterative process. In this section we will discuss the effect of the velocity on the obtained data. We will first discuss some simple models in order to understand the features we can encounter in real data. As a consequence of this, we will discuss which operations have to be applied to the data in order to obtain the desired information. We assume here that we deal with a CMP gather.

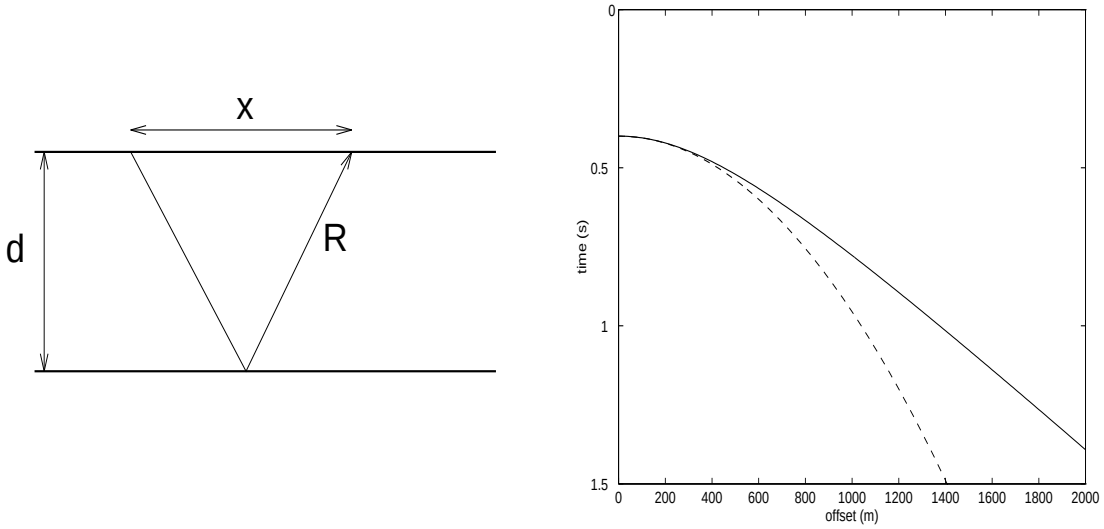


Figure 3.8: a) Distances in a subsurface model with one flat reflector. b) NMO curve for geometry of a) with depth 300 m and velocity of 1500 m/s. The dashed line is the parabolic approximation of the hyperbola.

Let us first consider the reflection from a single interface as depicted in figure 3.8. The time for the geometrical ray from source to receiver is given by:

$$T = \frac{R}{c} = \frac{(4d^2 + x^2)^{1/2}}{c}, \tag{3.4}$$

in which x is the source-receiver distance, R is the total distance traveled by the ray, d is the thickness of the layer and c is the wave speed. When we write $2d/c$ as T_0, then we can rewrite this equation as:

$$T = T_0 \left(1 + \frac{x^2}{c^2 T_0^2}\right)^{1/2}. \tag{3.5}$$

Note that this function describes a hyperbola. We can see that we have an extra time delay due to the factor $x^2/(c^2 T_0)$. The extra time delay is called the Normal Move Out, abbreviated to NMO. This extra term is solely due to the extra offset of the receiver with respect to the source; at coincident source-receiver position this term is zero. Often, the square-root term in this equation is approximated by its one-term Taylor series expansion, i.e.:

$$T \simeq T_0 + \frac{x^2}{2c^2 T_0}. \tag{3.6}$$

Figure 3.8b shows the traveltime curve for a layer of 300 meter depth and a velocity of 1500 m/s. The dashed line in this figure shows the parabolic approximation according to equation (3.6).

In seismic processing we are not interested in the extra time delay due to the receiver position: the image of the subsurface should be independent of it. The removal of the extra time delay due to NMO is called the *NMO correction*.

NMO curve for more than one interface

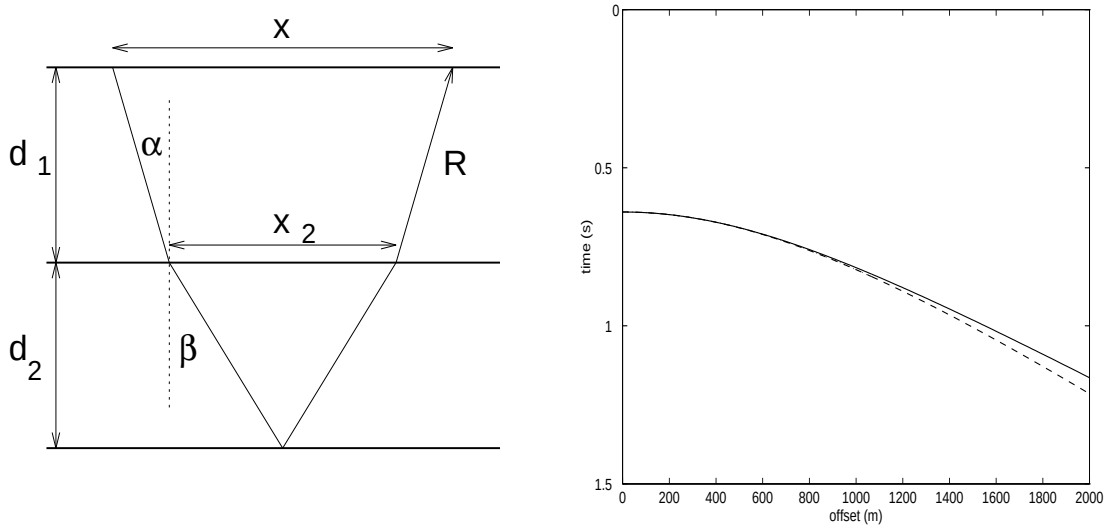


Figure 3.9: a) Distances in a subsurface model with two flat reflectors. b) NMO curve for second reflector with depth 300 m of each layer and velocities of 1500 m/s and 2500 m/s in the first and second layer respectively. The dashed line is the hyperbolic approximation of the traveltime curve.

Let us now move to a model with two interfaces, as depicted in figure 3.9. We call the source-receiver distance x, the horizontal distance the ray has traveled in the second layer x_2, the wave speed in the first layer c_1, and in the second c_2, the thickness of the first layer d_1, and of the second d_2. Then the traveltime from source to receiver is given by:

$$T = \frac{\left((x-x_2)^2+4d_1^2\right)^{1/2}}{c_1} + \frac{\left(x_2^2+4d_2^2\right)^{1/2}}{c_2} \tag{3.7}$$

$$= \frac{2d_1}{c_1}\left(1+\frac{(x-x_2)^2}{4d_1^2}\right)^{1/2} + \frac{2d_2}{c_2}\left(1+\frac{x_2^2}{4d_2^2}\right)^{1/2} \tag{3.8}$$

$$= T_1\left(1+\frac{x_1^2}{T_1^2c_1^2}\right)^{1/2} + T_2\left(1+\frac{x_2^2}{T_2^2c_2^2}\right)^{1/2}, \tag{3.9}$$

in which T_1 and T_2 are the zero-offset traveltimes through the first and second layer respectively, and $x_1 = x - x_2$. The problem with this formula is that, if we assume that c_1 and c_2, are known, we do not know x_2. Therefore we cannot directly use this expression to describe the move-out behaviour of this two-reflector model.

In order to tackle this, we first expand the square-root terms in equation (3.9) in a Taylor series expansion as we did for the one-interface case:

$$T \simeq T_1 + \frac{x_1^2}{2T_1c_1^2} + T_2 + \frac{x_2^2}{2T_2c_2^2}. \tag{3.10}$$

and we square this equation in order to obtain:

$$T^2 = (T_1+T_2)^2 + (T_1+T_2)\left(\frac{x_1^2}{T_1c_1^2} + \frac{x_2^2}{T_2c_2^2}\right) + O(x^4). \tag{3.11}$$

In this equation, we still have the distances x_1 and x_2 present. A relation between x_1 and x_2 can be found using Snell's law at the interface, being:

$$\frac{\sin\alpha}{c_1} = \frac{\sin\beta}{c_2}, \tag{3.12}$$

with α and β are the angles of the ray with the normal in layer 1 and 2 respectively, when crossing the first interface (see also figure 3.9). We make an approximation for small angles for which $\sin\alpha \approx \tan\alpha$ and $\sin\beta \approx \tan\beta$, such that equation (3.12) becomes:

$$\frac{x_1}{2d_1c_1} \approx \frac{x_2}{2d_2c_2}, \tag{3.13}$$

or

$$\frac{x_1}{T_1c_1^2} \approx \frac{x_2}{T_2c_2^2}. \tag{3.14}$$

Writing this as $x_2 = (T_2c_2^2)/(T_1c_1^2)x_1$ and substituting this in $x_1 + x_2 = x$, we have:

$$x_1 = x\frac{T_1c_1^2}{T_1c_1^2 + T_2c_2^2}, \tag{3.15}$$

Similarly for x_2, we obtain:

$$x_2 = x\frac{T_2c_2^2}{T_1c_1^2 + T_2c_2^2}. \tag{3.16}$$

We can use equations (3.15) and (3.16) in the quadratic form of eq.(3.11) to obtain:

$$T^2 \approx (T_1+T_2)^2 + (T_1+T_2)x^2\left(\frac{T_1c_1^2 + T_2c_2^2}{(T_1c_1^2 + T_2c_2^2)^2}\right) \tag{3.17}$$

$$\approx (T_1+T_2)^2 + \frac{(T_1+T_2)}{T_1c_1^2 + T_2c_2^2}x^2. \tag{3.18}$$

This equation is of the form:

$$T^2 = T_{tot}(0)^2 + \frac{x^2}{c_{rms}^2}. \tag{3.19}$$

with c_{rms} is what is called the root-mean-square velocity:

$$c_{rms}^2 = \frac{1}{T_{tot}(0)}\sum_{i=1}^{N} c_i^2 T_i(0), \tag{3.20}$$

in which $T_i(0)$ denotes the zero-offset traveltime through the i–th layer; $T_{tot}(0)$ denotes the total zero-offset time:

$$T_{tot}(0) = \sum_{i=1}^{N} T_i(0). \tag{3.21}$$

We see here that with the assumptions made, a hyperbolic move-out for the interfaces below the first one is obtained. The approximation however is a very good one at small and intermediate offsets (for horizontal layers) but becomes worse when the offset becomes

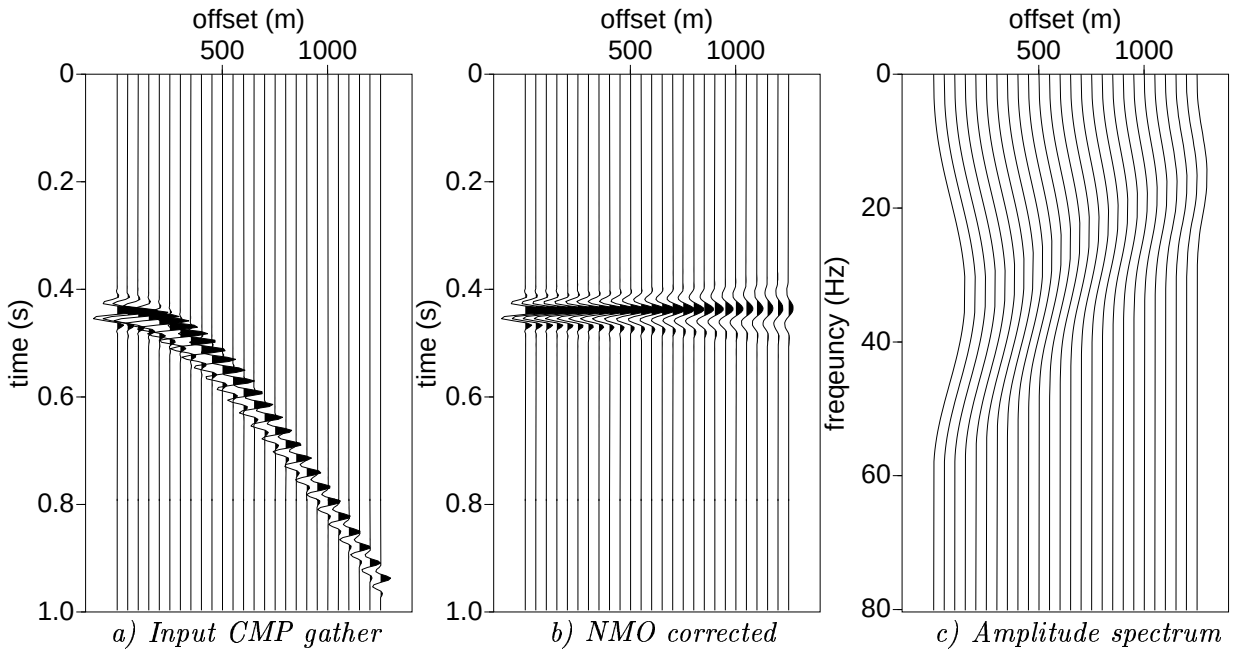


Figure 3.10: CMP gather with one reflection before (a) and after NMO correction (b). The amplitude spectrum in (c) of the NMO corrected gather shows the frequency reduction due to the stretching effect.

large. This effect can be observed in figure 3.9b, where the hyperbolic approximation of the second interface reflection is plotted with a dashed line.

Applying NMO correction

Then, how do we apply this NMO correction? First we have to determine the stacking (i.e. root-mean-square) velocities for each zero offset time T_0 (see next section). Then, for each sample of the zero-offset trace will remain in its position. For a trace with offset x, we calculate the position of the reflection according to equation (3.19) and find the sample nearest to this time T. This sample is then time-shifted back with the time difference between T and T_0 (in fact it is mapped from time T to time T_0). In this simple scheme we have taken the sample nearest to the time T, but in general we can be much more accurate by using a better interpolation scheme. It is important to realize that with NMO we interpolate the data.

An artifact of the NMO correction is the NMO *stretch*. An example of this effect is shown in figure 3.10. How does this occur? We can see that the correction factor not only depends on the offset x and the velocity c_{rms}, but also on the time T_0. So given a certain stacking velocity and offset, the correction $T-T_0$ becomes smaller when T_0 becomes larger. This is visible in figure 3.11, where for the second event smaller time shifts need be applied compared to the first event. Thus, the correction is not constant along a trace, even if we have a constant offset and constant velocity. Also, we can see from this correction that the effect will become more prominent when the offset becomes larger as well. This effect is called NMO stretching. A measure for this stretch is the quantity $\Delta t_{NMO}/T_0$, with the applied NMO correction being defined as $\Delta t_{NMO} = T - T_0$. This can be seen by analyzing

the change in NMO correction as a function of T_0:

$$\frac{\partial \Delta t_{NMO}}{\partial T_0} = \frac{\partial \left((T_0^2 + x^2/c^2)^{1/2} - T_0 \right)}{\partial T_0} \tag{3.22}$$

$$= \frac{T_0}{(T_0^2 + x^2/c^2)^{1/2}} - 1 = \frac{T_0 - (T_0^2 + x^2/c^2)^{1/2}}{(T_0^2 + x^2/c^2)^{1/2}} \tag{3.23}$$

$$\approx -\frac{\Delta t_{NMO}}{T_0}. \tag{3.24}$$

This quantity relates to the frequency distortion by:

$$\frac{\Delta t_{\mathrm{NMO}}}{T_0} = \frac{\Delta f}{f_{dom}}, \tag{3.25}$$

where f_{dom} is the dominant frequency and Δf is the change in the dominant frequency due to the stretching, supposed we have a constant velocity. This frequency distortion is clearly visible if we take the amplitude spectrum of each NMO corrected trace in figure 3.10b, as displayed in figure 3.10c. The frequency content decreases dramatically towards the large offsets. In this example $T_0 = 0.44s$ and $\Delta t_{\mathrm{NMO}} = 0.5s$ at the largest offsets, giving a stretch of 1.14, meaning that the frequency reduces to less than half the original bandwidth, which can indeed be observed in figure 3.10c.

When we are processing the data, we do not want to have a too severe signal distortion, and therefore the data which is distorted more than a certain threshold (say 100%), is zeroed, called *the mute*. The measure for this distortion is as given in equation (3.25). Figure 3.12 shows an example of a CMP gather with two reflection in which NMO correction is applied with and without the mute applied. In figure 3.12c the signal is muted when more than 50% stretch is present.

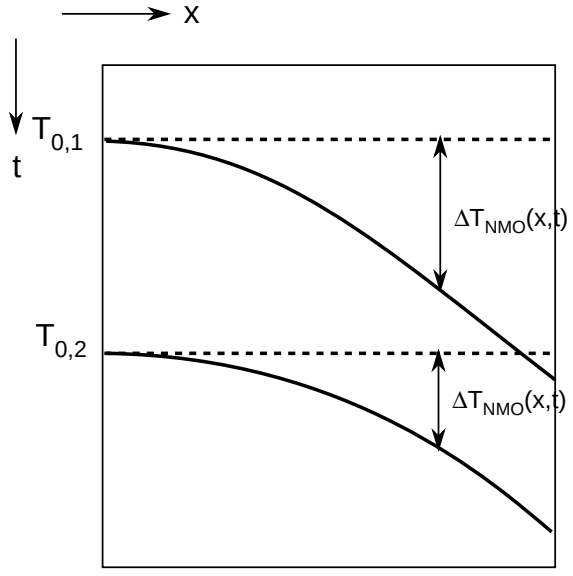


Figure 3.11: NMO correction is applied as a time and offset dependent time shift of the data.

Velocity estimation

In the application of the NMO correction, there is of course one big question: which velocity do we use? Indeed, we do not know the stacking velocity on beforehand. Actually, we use the alignment of a reflection in a CMP gather as a measure for the velocity. Since, if the velocity is right, the reflection will align perfectly. However, when the velocity is taken too small, the correction is too large and the reflection will not align well; in the same way, when the velocity is chosen too big, the correction is too small, and again the reflection will not align. An example of these cases is given in figure 3.13.

As the earth is consisting of more than one interface, we need to determine the velocities, although they may just be root-mean square velocities for each layer. The goal is the same as in the case of just one interface: we would like all the reflections to be horizontally aligned. A systematic way of determining these velocities is to make common-midpoint panels which are each NMO corrected for a constant velocity. Then we can see for those velocities the reflector will align or not; usually the deeper the interface the higher the (root-mean-square) velocity. An example of such an analysis is given for a four reflector median (see figure 3.5) in figure 3.14.

Another way of determining velocities is via $t^2 - x^2$ analysis. For this analysis we have to pick the traveltimes for a certain reflector and plot them as a function of x^2. As we have seen with multiple interfaces, the slope of this curve should be $1/c_{RMS}^2$, and thus we know the stacking velocity. This method can be quite accurate but depends on the quality of the data whether we are able to pick the reflection times from the data. An example is given in figure 3.15, where for a two-reflector situation the axis have been stretched such that the different velocities become visible as different dip lines.

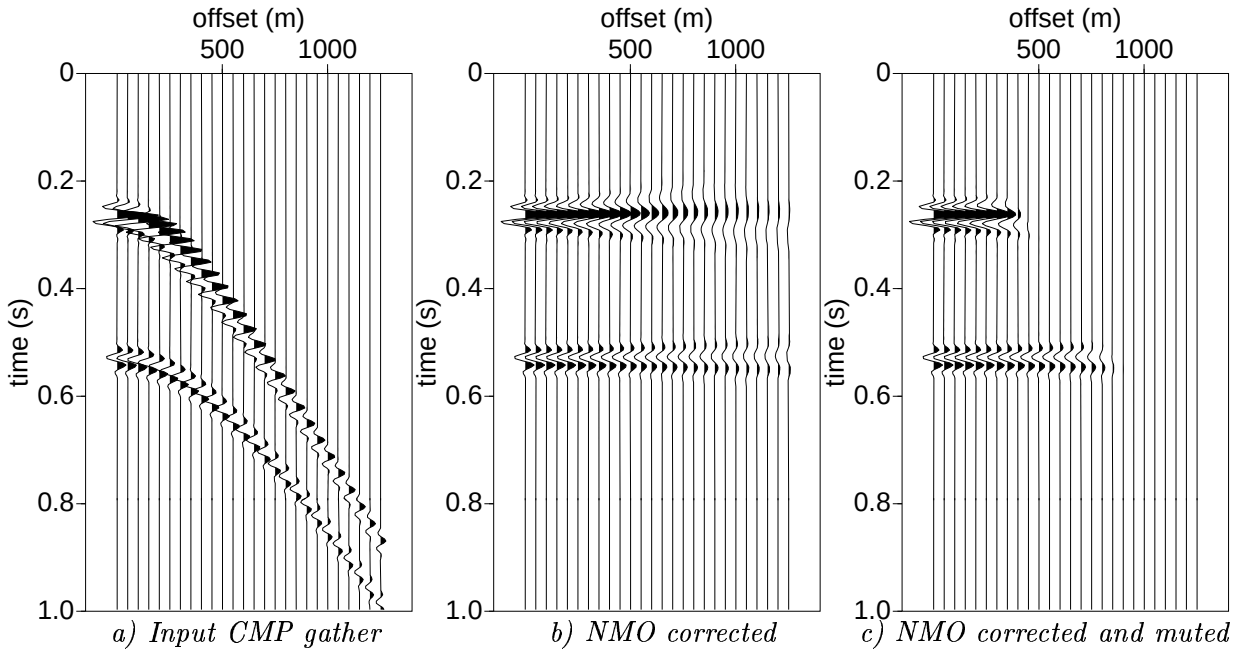


Figure 3.12: a) CMP gather with two reflections. b) CMP gather after NMO correction without stretch-mute. c) CMP gather after NMO correction and stretch-mute.

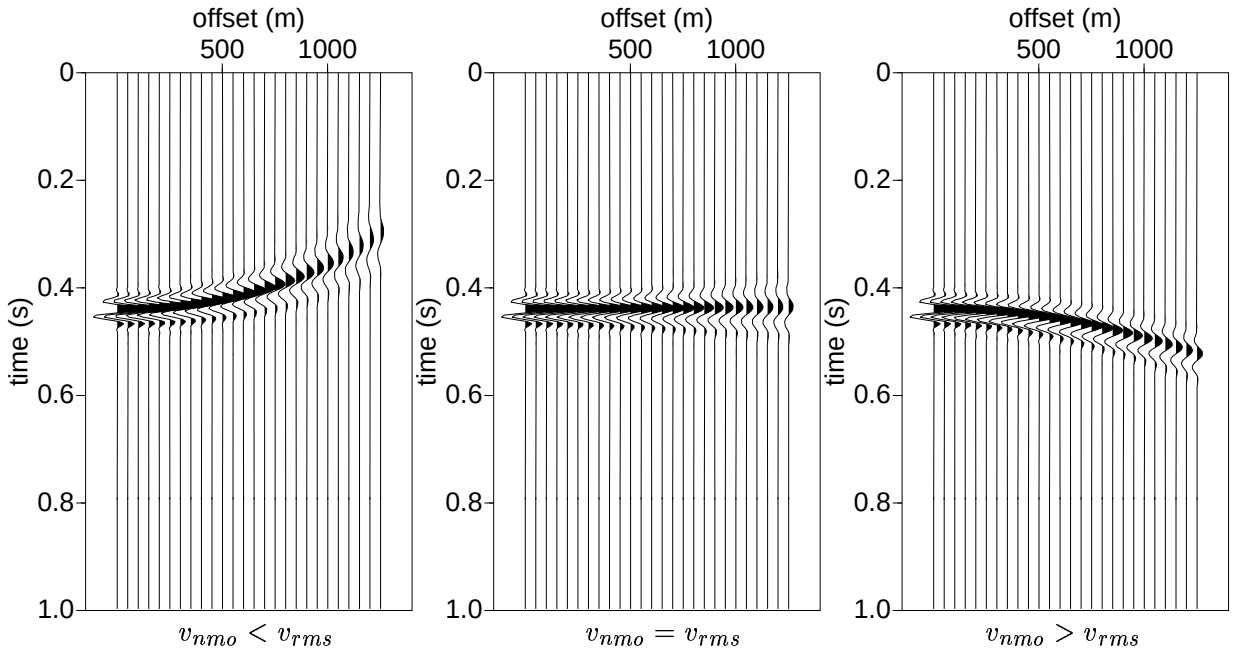


Figure 3.13: CMP gather with one reflection after NMO correction with too low, correct and too high stacking velocities.

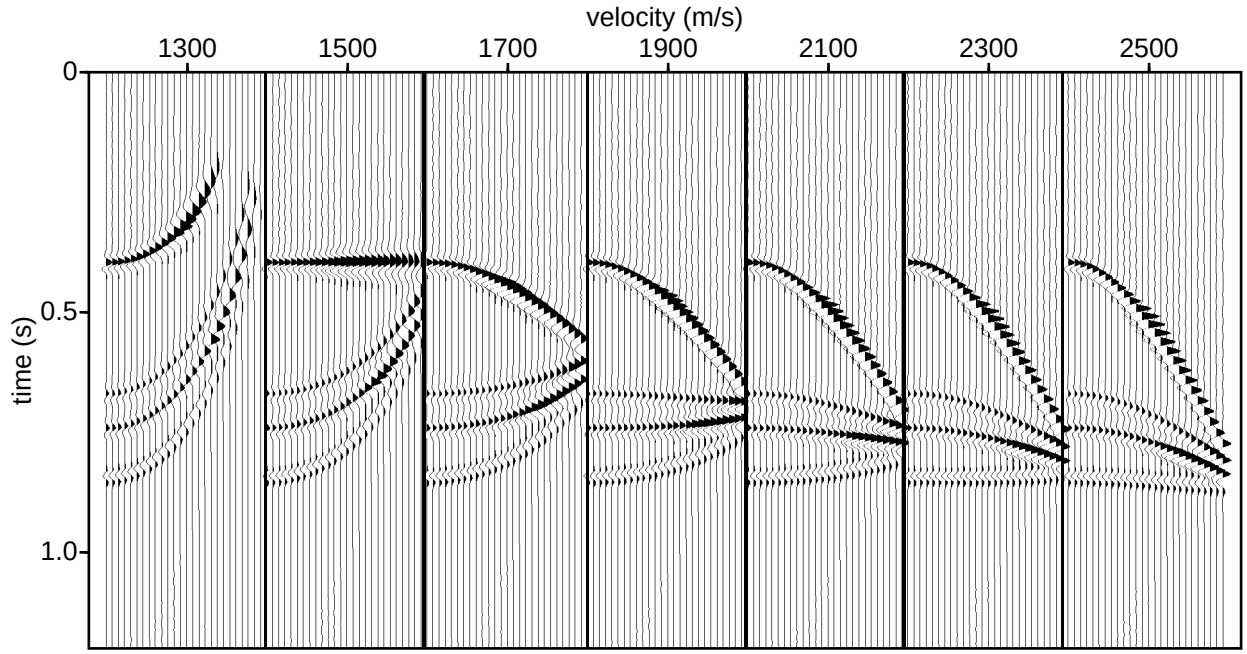


Figure 3.14: CMP gather NMO corrected with a range of constant NMO velocities from 1300 to 2700 m/s with steps of 200 m/s.

The most commonly used way of determining velocities is via the *velocity spectrum*, which can be seen as a hyperbolic transformation of the seismic reflections from the space-time domain to the velocity-time domain. In order to determine the velocity spectrum we correct the CMP gather for a certain stacking velocity and apply a coherency measure (i.e. weighted stack) to this data. This gives us one output trace. Then, for a next velocity, we do the same. For a complete set of velocities, we plot these results next to each other, the result being called the velocity spectrum. On the vertical axis we then have the time, while on the horizontal axis we have the velocity. As an example we consider again the synthetic CMP gather in the model of figure 3.5, for which we calculate the semblance for velocities between 1000 m/s and 3000 m/s with 50 m/s interval, see figure 3.16. The result we obtain is often displayed in contour mode or color mode.

As a coherency measure, the semblance is most often used. The semblance $S(t,c)$ at a time t for a velocity c is defined as:

$$S(t,c) = \frac{1}{M} \frac{\left(\sum_{m=1}^{M} A(x_m,t,c)\right)^2}{\sum_{m=1}^{M} A^2(x_m,t,c)}, \tag{3.26}$$

in which M is the number of traces in a CMP and A is the amplitude of the seismogram at offset x_m and time t after NMO correction with velocity c. For the definition of other coherency measures, the reader is referred to [Yilmaz, 1987] (page 169, 173). Note that if an event is perfectly aligned with constant amplitude for all offsets, the semblance has value 1. Therefore, the semblance has always values between 0 and 1.

Normally, the coherency measure is averaged over a few time samples, just to avoid too much computing time. Say, the k–th sample is in the center of the time gate, then

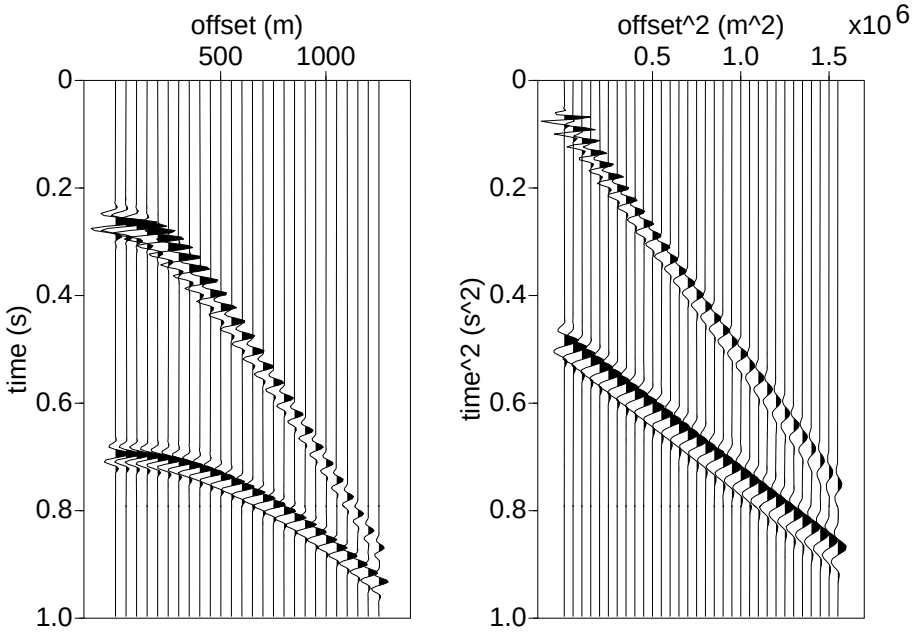


Figure 3.15: CMP gather with two reflections (left) after the t^2, x^2 axis stretching.

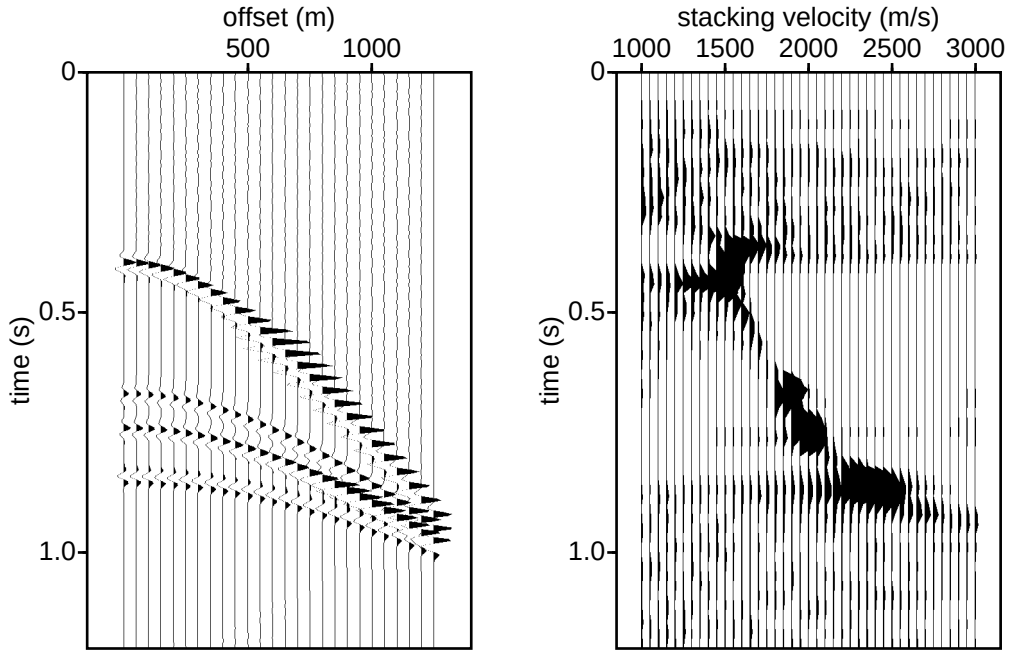


Figure 3.16: CMP gather with its velocity spectrum, using a semblance calculation with window length of 20 ms.

the semblance is:

$$S(t_k, c) = \frac{1}{M} \frac{\sum_{i=k-p}^{k+p} \left(\sum_{m=1}^{M} A(x_m, t_i, c) \right)^2}{\sum_{i=k-p}^{k+p} \sum_{m=1}^{M} A^2(x_m, t_i, c)}, \tag{3.27}$$

with $2p$ being the number of samples within the time gate. In practice, the time gate should be chosen quite small, at most the period of the dominant frequency of the signal ($1/f_{dom}$), usually between 20 and 40 ms, otherwise we loose too much temporal resolution. The effect of using different time gates is shown in figure 3.17.

For a more extensive discussion on the velocity analysis we would like to refer to [Yilmaz, 1987] (pp.166—182).

3.4 Stacking

A characteristic of seismic data as obtained for the exploration for oil and gas, is that they generally show a poor signal-to-noise ratio, not only due to coherent events such as surface waves, but also due to uncorrelated noise. Often, only the strong reflectors show up in raw seismic data. An important goal in seismic processing is to increase the signal-to-noise ratio, and the most important steps towards this goal, is CMP sorting and stacking. With stacking we add the NMO-corrected traces in a CMP gather to give one output trace. A better nomenclature is perhaps *horizontal* stacking because we stack in the horizontal direction. This is in contrast to *vertical* stacking, which is recording

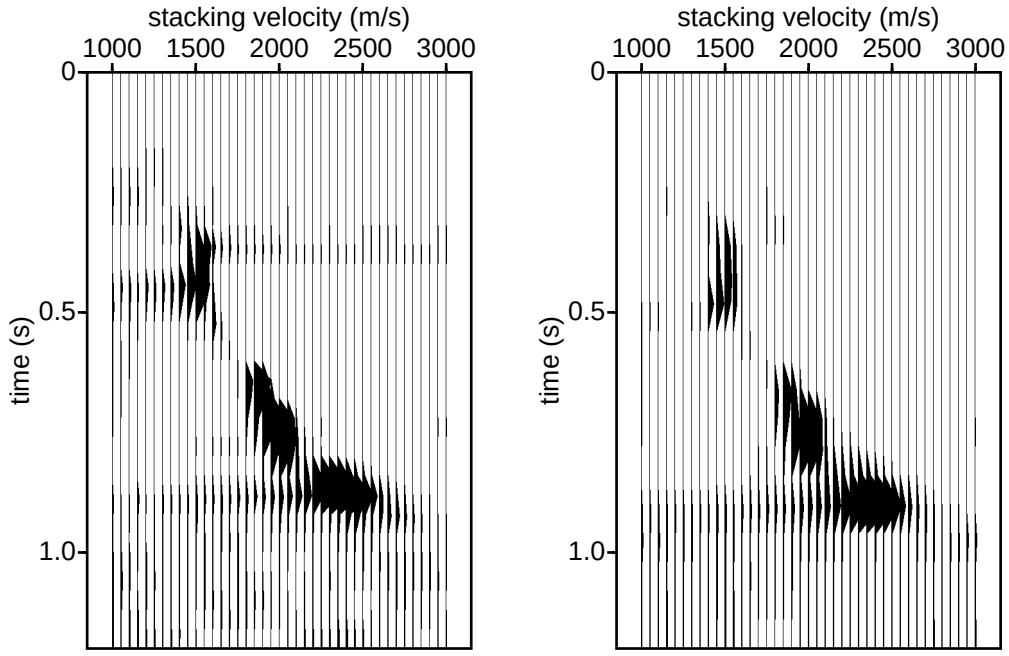


Figure 3.17: Semblance velocity spectrum of a CMP gather with a window length of 40 ms (left) and 60 ms (right).

the data at the same place from the same shot position several times and adding (i.e. averaging) these results. With stacking, we average over different angles of incidence of the waves, even in horizontally layered media. This means that we loose some information on the reflection coefficient since, as the reader may know, the reflection coefficient of an interface is angle-dependent. Therefore, the stacked section will contain the *average* angle dependent reflection information. It is important to realize that stacking the traces in the $x-t$ domain is equal to selecting the $k_x = 0$ line in the (f, k_x) domain. This is similar to the time domain situation: the summation of all time samples yields the d.c. component (i.e. $f = 0$ Hz) of the signal. In principal, all events in the NMO corrected CMP gather that are not present at $k_x = 0$ will disappear in the stack. Of course, an optimally aligned (i.e. NMO corrected) event will yield a large contribution at $k_x = 0$. This is visible in figure 3.18 where a CMP gather with two -primaries and one multiple is shown before and after NMO correction. The NMO correction maps all primary energy around the zero wavenumber axis, although the multiple energy is still spread out in the (f, k_x) domain. The resulting stack shows a reduced multiple energy, which is a desired feature of the stack.

Although the signal-to-noise ratio is increased by stacking, we will also have introduced some distortions. We have already discussed the NMO stretch and the approximation with the root-mean-square velocity. Therefore, when we add traces, we do not do a perfect job so we loose resolution. The effect of an erroneous velocity for the NMO is shown in figure 3.19, which shows a stacked section with the correct stacking velocities and with 7% too high stacking velocities for the data generated in the model of figure 3.2. One can

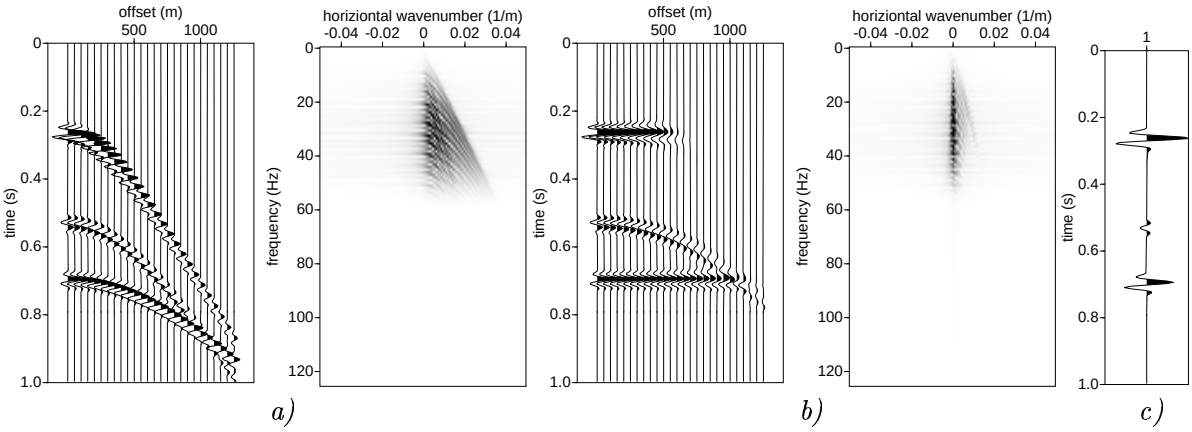


Figure 3.18: CMP gather with 2 primaries and 1 multiple in the (x,t) and (f,k_x) domain before (a) and after (b) NMO correction and after stacking (c).

see that the stacked trace is getting a lower frequency content and that the amplitudes are decreasing in some parts with the erroneous velocities. Note that a stacked section simulates a zero offset section, but with much better signal to noise ratio. Compare therefore the stacked result to the zero offset section of figure 3.6, which shows exactly the same region (1000 - 4000 m) in the model. Note the resemblance of the stack with the zero offset section. Note also that the stack is twice as dense sampled in the trace direction, due to the fact that there are twice as many CMP positions as there are shot positions.

Finally, it should be emphasized that, with stacking, we reduce the data volume. The amount of data reduction is the number of added traces in a CMP gather. There are certain algorithms which are expensive to compute and are therefore applied to stacked data rather than on pre-stack data. An example of this is migration as shall be discussed in the next section.

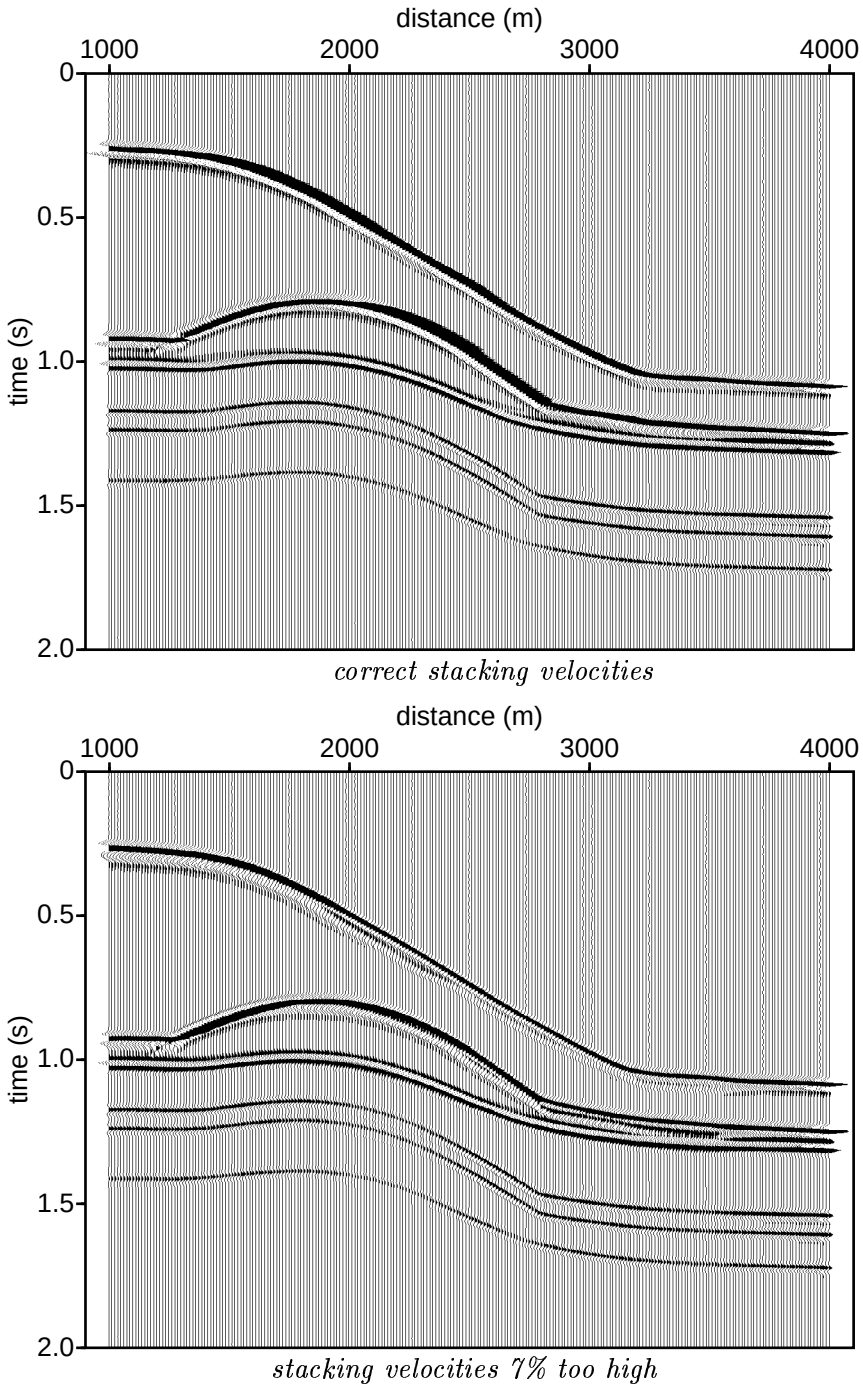


Figure 3.19: Stacked sections with correct and too high stacking velocities.

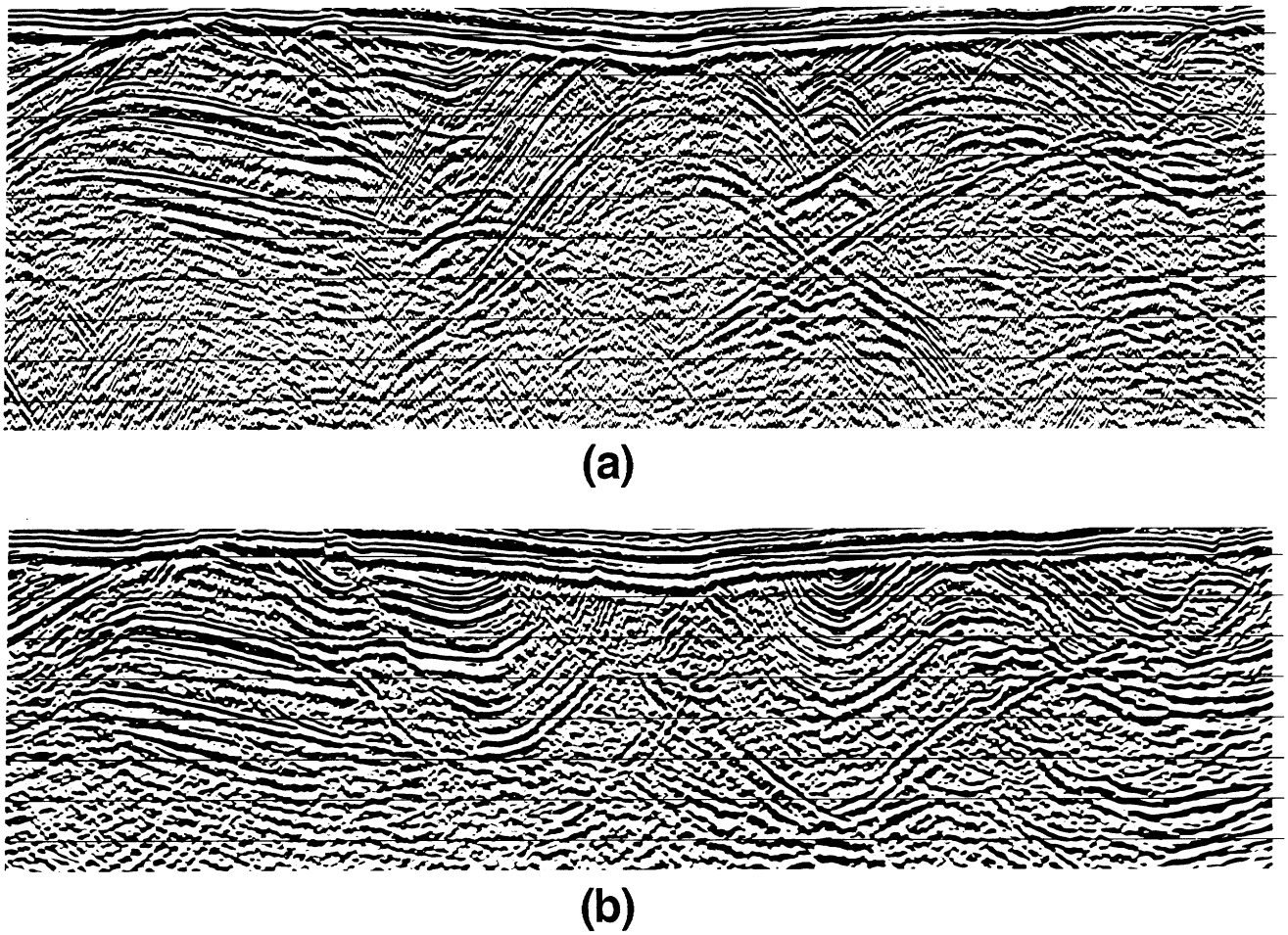


Figure 3.20: Stacked section (a) and its time migrated version (b) (from Yilmaz, 1987, fig. 4-20)

3.5 Zero-offset migration

Introduction

Although we have removed some timing effects with the NMO correction, this does not mean that we have removed the wave effects: it is just one of many. Migration deals with a further removal of wave phenomena in order to arrive at a section which is a better representation of the subsurface. After the NMO correction and stacking, we can still have diffractions in our stack, such as shown in figure 3.20. Also, dipping reflectors will not be positioned at their right place and will not show the true dip, as also shown in figure 3.20. In migration we are going to deal with this.

We have only summarized briefly the problem but in order to understand it better, we have to invoke some concepts and theory, such as the exploding reflector model and wave theory. But let us first summarize at what stage we have arrived. We sorted our

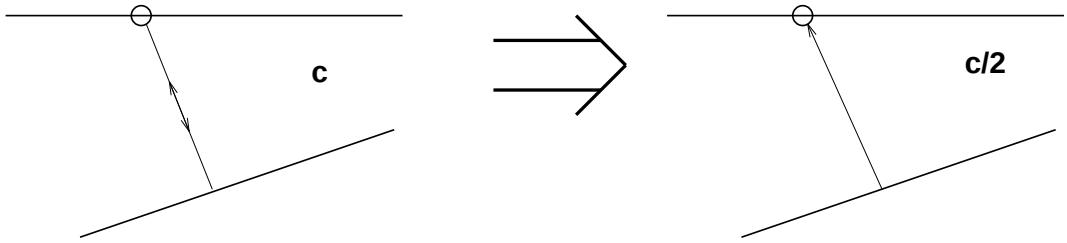


Figure 3.21: Exploding reflector model for zero offset data. A zero offset measurement can be considered as an exploding reflector response in a medium with half the velocity.

data in CMP gathers, determined the stacking velocities, applied the NMO correction and stacked the data. In the NMO correction we removed the offset dependence of the receiver position with respect to the source. In this step we made our data look as if it were obtained with coincident source-receiver position, thus zero-offset. For a reflection from a horizontal reflector, the reflection points on the reflector coincide for all traces in one CMP gather at the same lateral location as the midpoint at the surface (see figure 3.5). However, when we have a laterally varying structure, the reflection points on the reflector do not coincide for the traces in one CMP gather (see figure 3.4). Still, we assume we have done a good job and assume the reflection in the stacked section is coming from one point in the subsurface. The main purpose of the stacking process is to reduce the data volume and to improve the signal-to-noise ratio of the data.

Exploding reflector model

An important concept in migration is *the exploding reflector model*. Consider a simple model with one reflector in the subsurface. When we have a source which emits a signal at $t = 0$, the signal will propagate through the medium to the reflector, will be reflected and will arrive back at the receiver (= shot position for a zero offset experiment). This is shown in figure 3.21 at the left hand side. Say the wave takes a time T to do this. Apart from some amplitude differences, the data recorded in such a way would be the same if we could fire off the sources on the reflector at time 0 but assume half the velocity of the medium in between. Putting the sources on the reflector is called the exploding reflector model. This is shown in figure 3.21 at the right hand side. To get a good amplitude, each exploding source should have the strength of the reflection coefficient of the reflector. When we have recorded the data at time T, and would keep track of the time to get back from time T to the reflector, we would obtain the image at time $t = 0$, again assuming we have taken the right (i.e. half the original) velocity. The condition of $t = 0$ is called the *imaging condition*. In fact, from a kinematic point of view, any zero offset section can be considered as being a record of exploding reflectors in the subsurface (for primary wave paths).

Diffraction stack

Let us consider the simple example of a point diffractor in the subsurface. When we excite a source and let the wave travel to the diffractor and let it travel back to the receiver, we will obtain an image as shown in figure 3.22; it defines a so-called a diffraction hyperbola, according to the following formula, with the apex in (x_d, T_d):

$$T^2 = \left[\frac{2R}{c}\right]^2 = T_d^2 + \frac{4(x_r - x_d)^2}{c^2}, \tag{3.28}$$

where R being the distance in a homogeneous medium with velocity c from the diffractor at $(x_d, z_d = T_d/c_s)$ to the surface position x_r.
Before the computers were in use, people used to stack along the hyperbola in order to obtain the diffracted point. Of course, for this procedure to be effective we need to know the stacking velocity. When we have more than one diffractor, we can do the same procedure and stack all the hyperbolae. This is called a diffraction stack. In the early days of computers the diffraction stack was used to apply the migration by calculating the following formula (assumed to have a discrete number of x's, being the traces in a zero offset section):

$$p(x, T) = \sum_{x_s} p(x_s, T(x - x_s, c)), \tag{3.29}$$

with T defined as the hyperbolic move-out of equation (3.28) as a function of the offset $x - x_s$ and the stacking velocity c. This equation indicates that all seismic energy is added, or stacked, along a hyperbola with the apex in point (x, T). What we do when stacking along hyperbolae, is actually removing the wave propagation effect from the point diffractor to the receiver positions. A very nice feature about the diffraction stack is that it visualizes our intuitive idea of migration, and is very useful in a conceptual sense.

So far, we considered one point diffractor, but we can build up any rerflector by putting all point diffractors on that reflector. When the spacing between the point diffractors

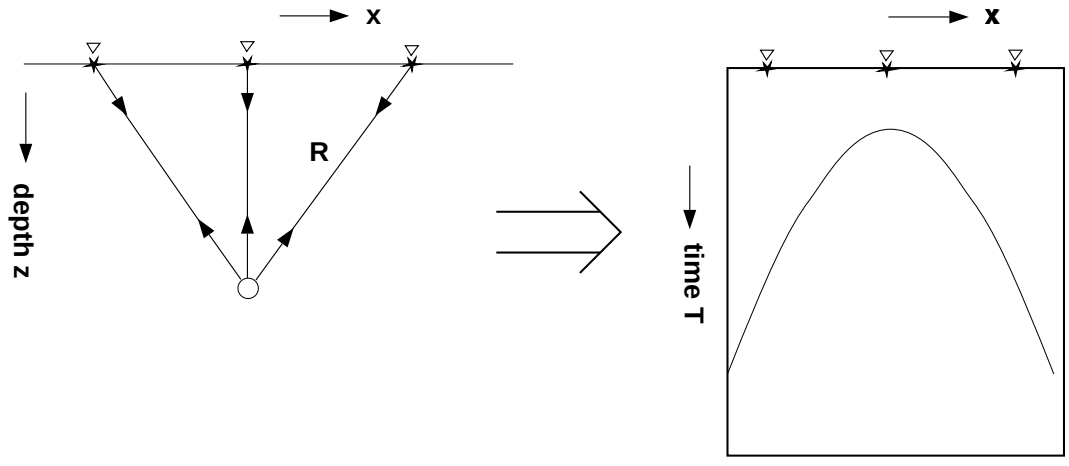


Figure 3.22: A diffractor (left) and its seismic response (right), a hyperbola in the zero offset section.

become infinitely small, the responses become identical. This concept agrees with Huygens' principle. As example, consider four point diffractors, as depicted on the left of figure (3.23). Each diffractor has the behaviour as discussed above, as can be seen on the right of figure (3.23), but the combination of the time responses shows an apparent dip. The actual dip goes, of course, through the apeces of the hyperbolae.

Let us now look at a full dipping reflector. Of course, it has some of the characteristics as we saw with the four point diffractors, only with a full reflector we no longer see the separate hyperbolae. Actually, we will only see the apparent dip. As we saw with the 4 point diffractors, we need to bring the reflection energy back to where they came from, namely the apex of each hyperbola. When connecting all the apeces of the hyperbolae, we get the real dip. This is depicted in figure (3.24).

The next figure (3.25) quantifies the effect of migrating the energy to its actual location. In particular, compare the figures in the middle and and on the right: the difference is a factor $\cos\theta$, where θ is the dip of the reflector with the horizontal. The zero offset traveltime at a certain x-value can be specified by $t_{ZO} = (2/c)x\sin\theta$, assuming that $x = 0$ corresponds to the point where the reflector hits the surface in figure 3.25a. The slope in the zero offset section is therefore $dt/dx = (2/c)\sin\theta$, see figure 3.25b. If this zero offset section is migrated and the result is displayed in vertical time $\tau = z/c$, the resulting slope of the reflector is $d\tau/dx = (2/c)\tan\theta$ (figure 3.25c). Thus, migration increases the time dip in the section by $\cos\theta$ and thus reflectors in the unmigrated section are increased in their up-dip direction in the migrated section. At the same time, migration decreases the apparent signal frequency by the factor $\cos\theta$. The reason that the dip is increased by $\cos\theta$ and the frequency decreased by $\cos\theta$ lies in the fact that the horizontal wavenumber is

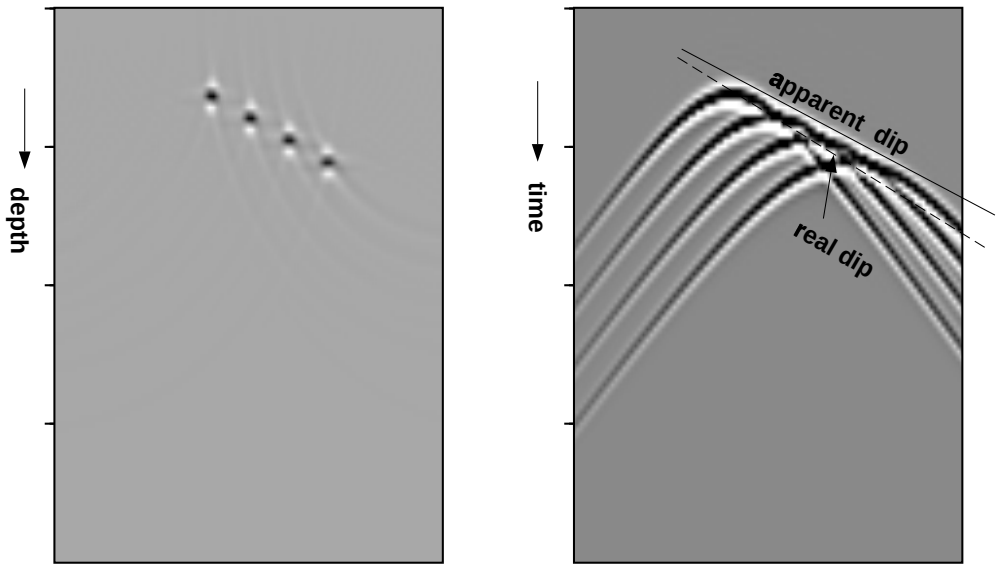


Figure 3.23: Four point diffractors (left) and their seismic responses (right). Note the apparent dip from the hyperbolae.

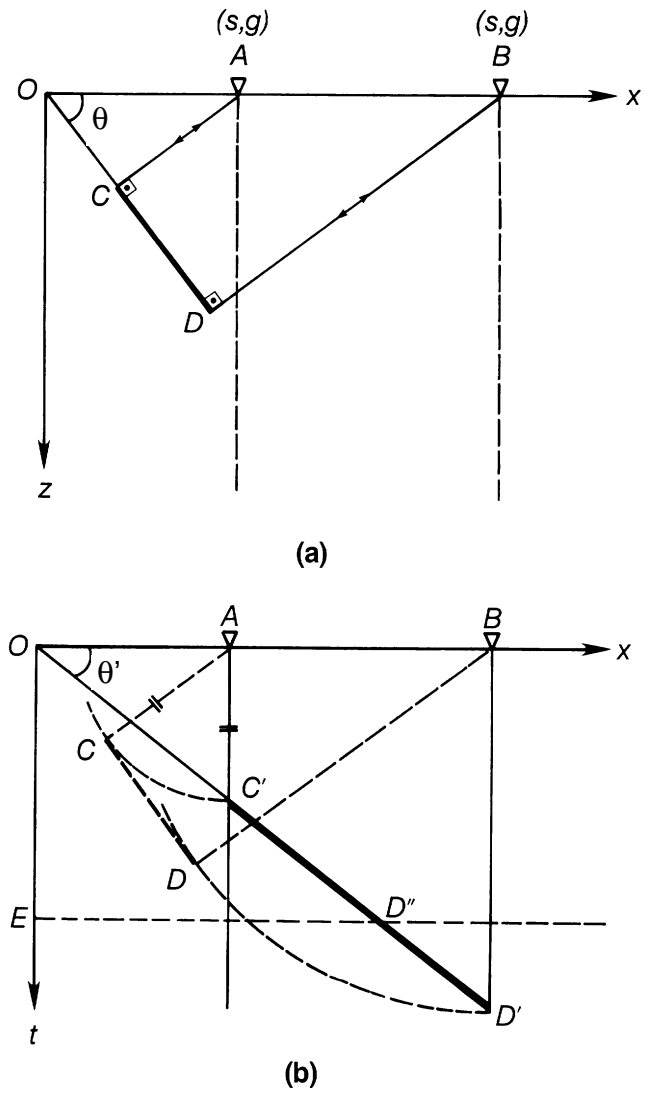
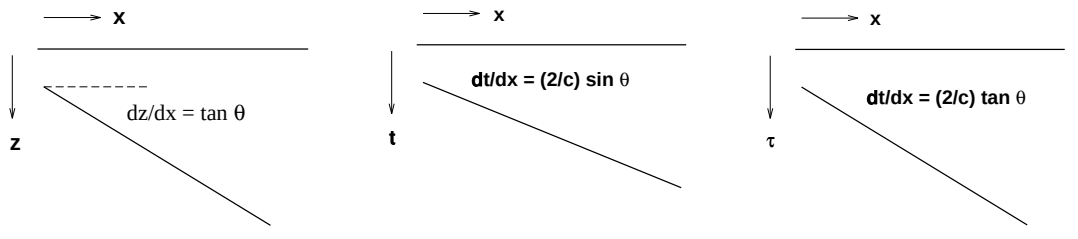


Figure 3.24: Relation between the reflection points in depth (a) and the traveltimes in the zero offset section (b) for a dipping reflector (from Yilmaz, 1987, fig. 4-14).



depth section *zero offset section* *migrated zero offset section*

Figure 3.25: Migration increases the dip in the zero offset section.

preserved.

Another commonly observed phenomena is the so-called "bow-tie" shaped zero offset response, due to synclinal structures in the earth. This is shown in figure 3.26, where it can be observed how in the middle above the syncline multi-valued arrivals are present. This behaviour can be predicted by considering small portions of the reflected signal, and increasing the dip of each portion of the reflected signal. Note that in figure 3.20 such structures are also visible.

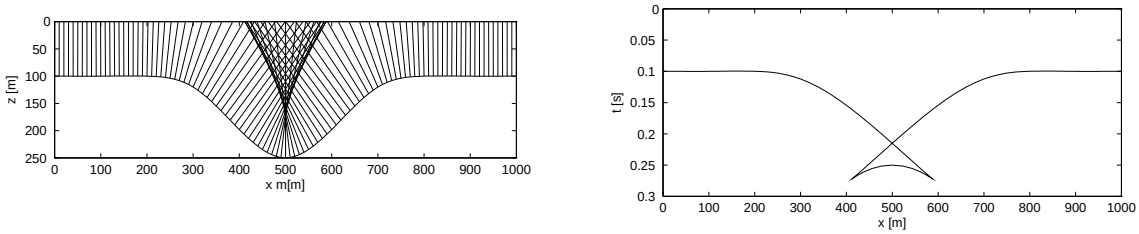
When we consider the above configurations, we can well understand the effect of migration of a real data set, as shown in figure 3.20. We can observe that all the diffractions in the stacked section are well collapsed after the migration. What is lacking in the approach of the diffraction stack is sound theory, and the final migrated result may be correct in position (if the diffraction responses can be assumed to have a hyperbolic shape, i.e. if the subsurface exhibits moderately velocity variations), but not in amplitude.

Migration using wave theory (zero offset)

The most elegant way to migrate seismic data in a way which takes wave theory into account, is Kirchhoff migration. Wave theory takes amplitdue effects into account, and is derived from fundamental physical laws in appendix B. Via the wave equation we can then derive a formulation for the Kirchhoff migration. The theoretical derivation is postponed until the next chapter (section 4.9). In this section, we are interested in the application of zero-offset (post-stack) migration. This is described by the Kirchhoff's migration formula [Schneider, 1978]:

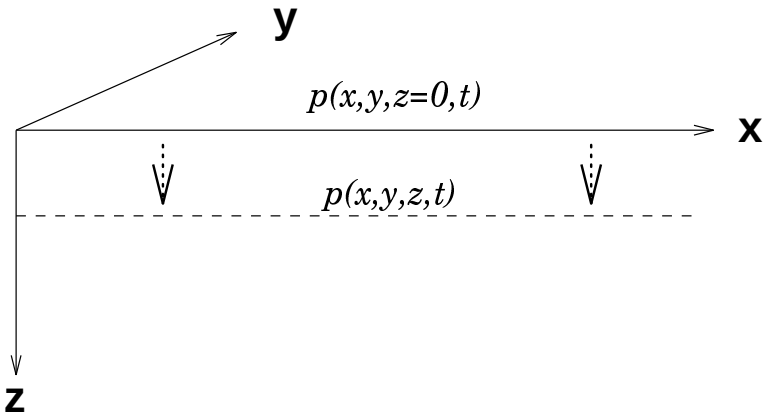
$$p(\mathbf{x},t) = \frac{-1}{2\pi}\partial_z \int_{z^s=0} \frac{p(\mathbf{x}^s, t+R/c)}{R}\, dA^s, \tag{3.30}$$

In the Kirchhoff migration formula (3.30), the $p(\mathbf{x}^s, t+R/c)$ in the integrand can be identified as the zero offset data or the stacked data (as a stack should simulate the zero offset recording). The data is recorded at the surface; Let us call our surface measurement p_{zo}, where the subsript zo stands for zero offset. Then we set $c \to c/2$ in order to correct for two-way traveltime. Then, we calculate the response for each point in the subsurface and put $t = 0$, the imaging condition, which images the exploding reflector which starts to act at $t = 0$. With other words, we start with our measurements at the surface and do



depth section with zero offset ray paths *zero offset section*

Figure 3.26: A syncline reflector (left) yields "bow-tie" shape in zero offset section (right).



Downward extrapolation: $p(x,y,z,t) = \frac{-1}{2\pi}\partial_z \int_{x^s,y^s} \frac{p(x^s,y^s,z^s=0,t+2R/c)}{R} dx^s dy^s$

Figure 3.27: Downward continuation step used in migration.

a downward continuation (inverse wave field extrapolation) to all depth levels, and pick the $t = 0$ point at each subsurface point. If there was a reflector at a certain point, it will be imaged with this procedure. If there is no reflector at a certain depth point, no contribution at $t = 0$ is expected for that point. So we can obtain a depth section by integrating over the surface to obtain:

$$p(\mathbf{x}) = \frac{-1}{2\pi}\partial_z \int_{z^s=0} \frac{p_{zo}(\mathbf{x}^s, 2R/c)}{R}\, dA^s. \tag{3.31}$$

Remember that R is the distance between the output point on the depth section and the particular receiver or trace location on the surface $z^s = 0$. So as we integrate along the surface $z^s = 0$ we are actually summing along diffraction hyperbolae (in the case of a constant velocity medium), defined by the time curve $t = 2R/c$, but then in a weighted fashion. Note indeed the large resemblance with the diffraction stack definition of equation (3.29). The extra $1/R$ factor takes the spherical divergence of the wave front into account and the factor ∂_z compensates for the frequency dependent and wave front angle dependent effects of the lateral summation process. Note that the integral over surface A^s will numerically be implemented as a summation over all (x_s, y_s) positions, i.e. a summation over all traces in the seismic section. Although the diffraction stack of equation (3.29) has been written as a summation over x_s only, the extension to 3D by adding a summation over the y_s coordinate is straightforward; in that situation the hyperbola is replaced by a hyperboloid: $T_s^2 = T^2 + 4[(x_s - x)^2 + (y_s - y)^2]/c^2$.

For inhomogeneous media, the diffraction responses are no longer hyperbolic, and the concept of diffraction stack is wrong. Here, we are doing the summation much better than the diffraction stack because we have included the wave equation.

The complete 3D zero offset migration procedure can now be as follows:

- Step 1 : Extract or simulate by stacking the zero offset dataset $p(x, y, z = 0, t)$.

Consider this to be measured in a half-velocity medium with exploding reflectors.

- Step 2 : Do a downward continuation (inverse extrapolation) step from the surface level to a level in the subsurface, according to:

$$p(x,y,z,t) = \frac{-1}{2\pi}\partial_z \int_{x^s,y^s} \frac{p(x^s,y^s,z^s=0,t+2R/c)}{R} dx^s dy^s. \tag{3.32}$$

  Note that for this extrapolation step we need the velocities in the subsurface. This extrapolation is visualized in figure 3.27.

- Select at each depth level the zero time component, which yields the migrated section:

$$p_{mig}(x,y,z) = p(x,y,z,t=0). \tag{3.33}$$

Our final result is a depth section, as we would obtain when we would make a geological cross-section through the subsurface (of course with a limited resolution). However, migration is not a simple process without any artifacts, and most importantly, we usually do not exactly know the velocity as a function of x, y and z. Therefore, we would like to be able to compare our original stacked section with the migrated section directly in order to see what the migration has done. Especially seismic interpreters need this type of comparison. To this aim, the depth coordinate z is mapped backed onto time τ via:

$$\tau = \frac{z}{c} \tag{3.34}$$

for a constant-velocity medium. ([Gazdag and Sguazerro, 1984], equation (43) etc.; [Schneider, 1978] p.56). For an inhomogeneous subsurface, this mapping is more complicated. For this purpose often ray-trace techniques are used to located the reflectors in time.

Time migration using the stacking velocities

To overcome the problem of not knowing the interval velocities in your medium, people have though of a work-around, using the stacking velocites. As we have done a stack in general, the stacking velocities are already known. In equation (3.32) we need to know the distance R from subsurface point to the surface (which depends on the velocities in the subsurface). It is often assumed that this path can be approximated by a straight line (as in a homogeneous medium) using the stacking velocity. Therefore, R is replaced by:

$$R/c \approx \tau' = (\tau^2 + \frac{4x^{s2} + 4y^{s2}}{c_{rms}^2})^{1/2}. \tag{3.35}$$

Furthermore, the extrapolated data is considered in migrated time τ and not in depth, which transforms equation (3.32) into:

$$p(x,y,\tau) \approx \frac{-1}{2\pi}\partial_z \int_{x^s,y^s} \frac{p(x^s,y^s,z^s=0,2\tau')}{c_{rms}\tau'} dx^s dy^s, \tag{3.36}$$

which again describes a diffraction stack (if we also neglect the derivative to depth). In these type of migrations, it is assumed that the structures in the subsurface are simple

enough to use the hyperbolic approximation of the response of an exploding reflector source.

Effects of wrong migration velocities

The only important parameter we can actually set is the velocity distribution. It is therefore important to know how a wrong velocity distribution will manifest itself in the final result. This is shown in figure 3.28 where we see a correctly and incorrectly migrated V-shaped reflector response. Note again the effect of migration: the increase of the slopes and the collapsing of the diffraction hyperbola into a point (i.e. the edge of the V-shape). When we put the velocity too low, the diffraction hyperbolae are not completely collapsed yet and we keep a hyperbola in our result. Such a section is undermigrated. In the same way, when the velocity is too high, then the diffraction hyperbolae are corrected too much, and an over-migrated section will arise. As such, migration can also be used to determine velocities: it is thát velocity that images the diffractor(s) in its original point with no diffraction effects visible anymore. A well-known effect of over-migrated sections is the creation of so-called "migration smiles" and crossing events, as visible in figure 3.29.

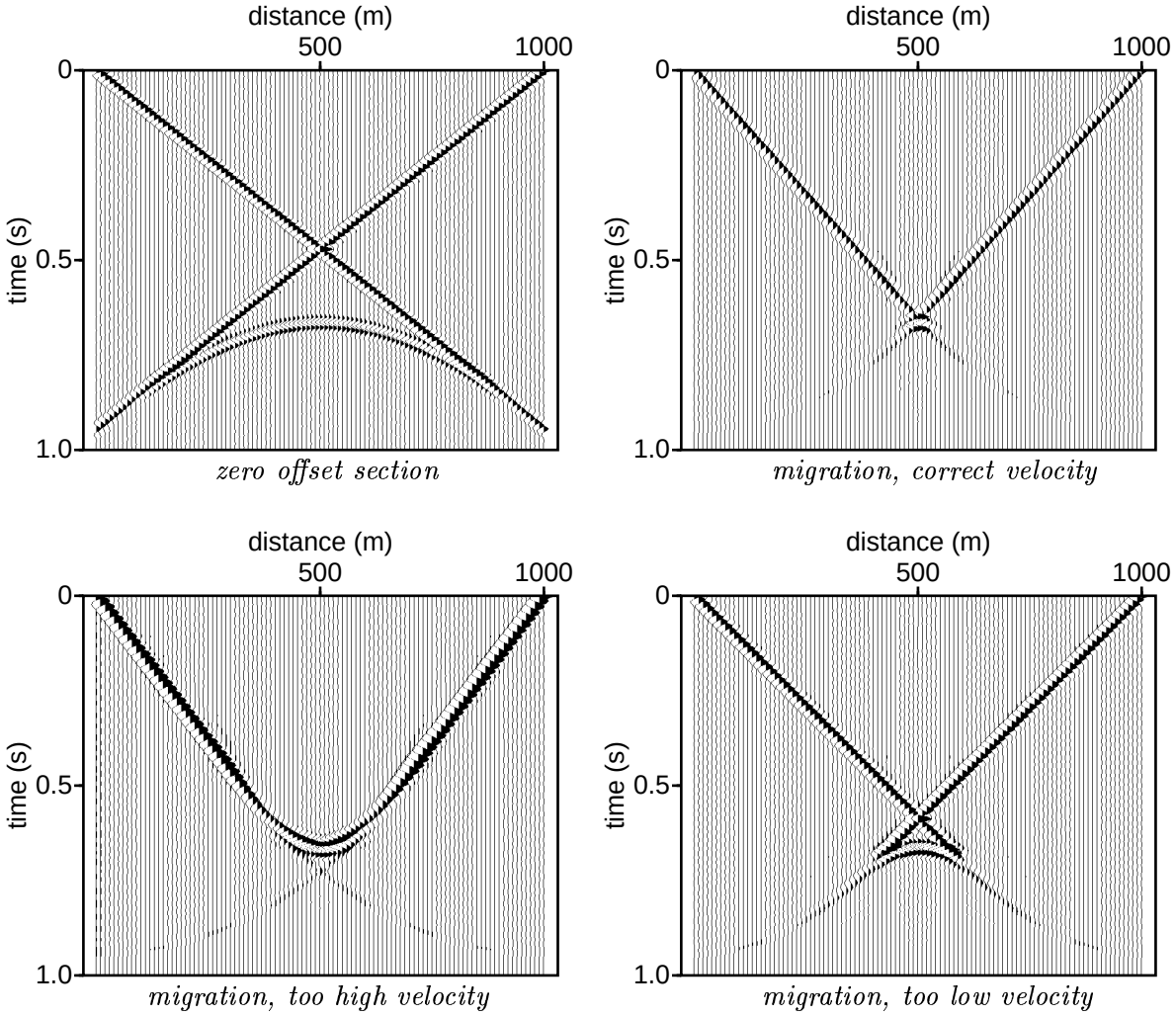


Figure 3.28: Stacked section and its time migrated version with the correct and wrong velocities.

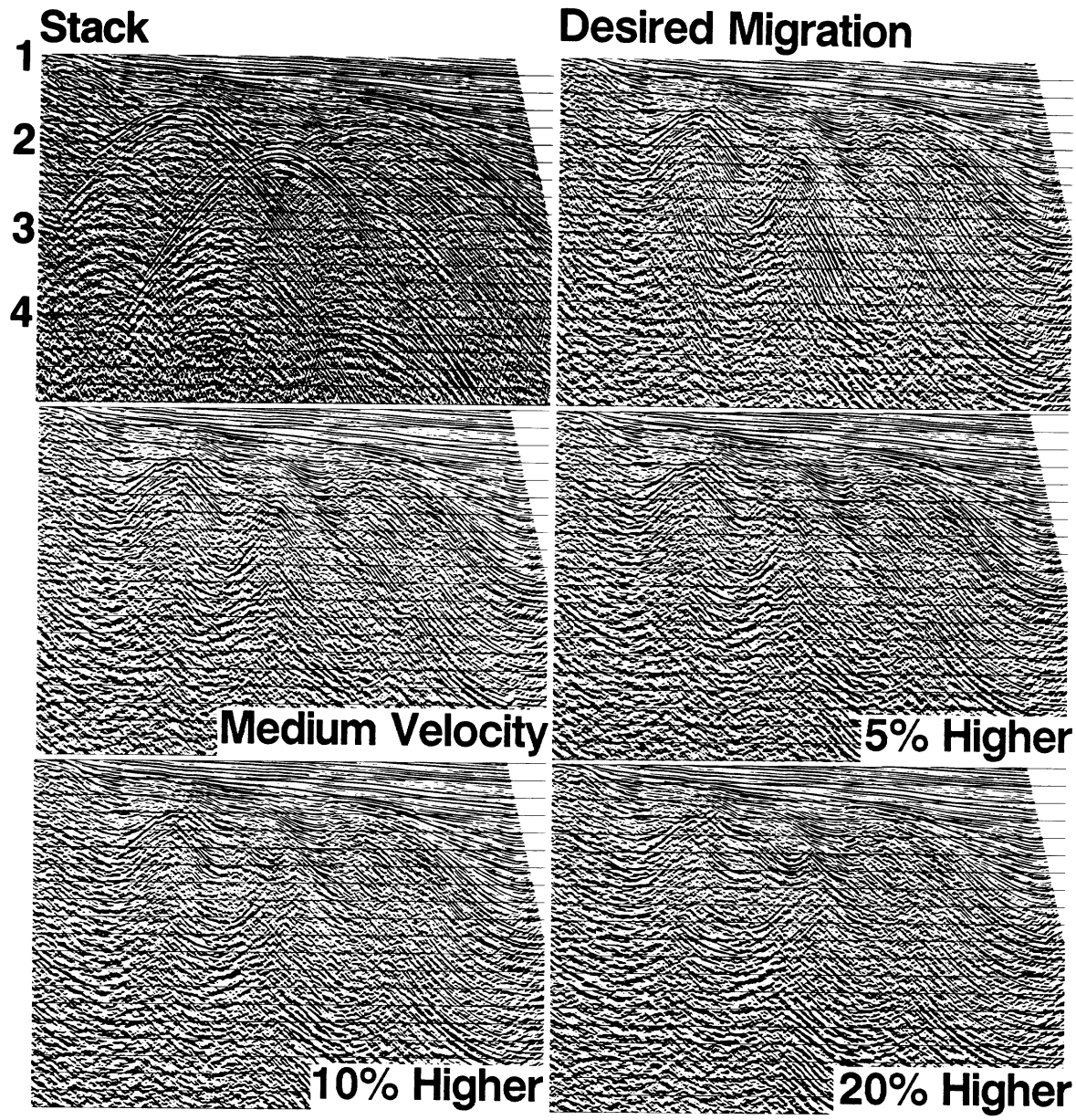


Figure 3.29: Stacked section and its time migrated version with the correct and wrong velocities (from Yilmaz, 1987, fig. 4-54).

Chapter 4

Extended Processing

4.1 Introduction

In this chapter we will deal with more specific subjects which are needed in processing. In chapter 3 we discussed a basic processing sequence which is common to data acquired on land as well as at sea. As was mentioned in that chapter, the processes which were described do not guarantee a good image of the subsurface in either case. On land, and at sea, there are specific phenomena which must be dealt with, and they shall be discussed in this chapter. However, these extra processing steps are common to either land and/or sea data in seismic processing production work. In certain circumstances they can still not be sufficient to get a good image of the subsurface. In that situation the data is given to people specialized in particular processing which is not commonly done, and often a close connection to research groups is kept.

There is perhaps one processing feature which is not dealt with explicitly but is very important nowadays in processing, namely 3-D. Many processes as described in this chapter apply to 3-D processing, and do not need any further explanation in their application to 3-D data sets, but some do. The most important one in that respect is 3-D migration. The largest improvement in data quality via 3-D shooting, has been due to migration of 3-D data in comparison with 2-D migration. We leave the discussion which is of particular interest to 3-D, to a later chapter on special processing.

In this chapter we will not discriminate between marine and land data, unless specifically mentioned. We will first discuss some data formats as currently being used in seismic exploration for oil and gas, before telling some things about surveying and field geometry-information. Then, in order to correct for amplitude variations, trace balancing is standardly applied to seismic data. A large nuisance in land data are static time shifts due to the influence of the shallow subsurface. Static corrections aim to correct for those local deviations and we will discuss them in a separate section. Another process which will be discussed is deconvolution, which can be used to remove short-time and long-time multiples, but also to shape the seismic wavelet. Another often applied process is filtering in the (f, k_x) domain, especially for the removal of surface waves in land data, and removal of multiples and noise in marine data. The last processes we will discuss in this chapter

are related to migration: first a separate section on Dip Move-Out (DMO), and then a more extensive discussion on migration techniques.

4.2 Data formats

Seismic data is processed on the computer, and therefore the data must be stored on a device. The data storage can be done in many ways, and in general there are quite a lot of ways to do that. Because the amount of data we are dealing with in seismics, it is important to do this in an efficient way. Already early in the use of computers seismic data is stored in a format specific to seismic data. And fortunately to the seismic industry, the formats introduced became standard. The standards in the seismic industry are made by a special standardization committee of the Society of Exploration Geophysicists (SEG), which is an American society. Quite a few standards came out of this committee, and two of them became very popular. These are the standards SEG-D and SEG-Y. There is a special booklet on the mostly used standards for data stored on tape devices, called Digital Tape Standards [SEG, 1980]. Since in the early days the disk capacity was rather small compared to the tape devices, hardly any format existed for seismic data stored on disk. However, times have changed and some standards have been defined, although not very much used (yet). Many processing packages have their own internal format, but lately the standard defined for tape devices, SEG-Y, is used for disk files as well. However, the data is unformatted written on disk, so it depends on what kind of internal word representation is used on the computer. SEG-Y is, by definition, IBM floating point. The other disk format for disk storage, is SEG-2, which is unfortunately only defined for a PC environment [Pullan, 1990].

SEG-D

Let us briefly tell something about the standard SEG-D. This is a standard which is mostly used in the field, and field tapes delivered to the processing centre are usually in this format. The first thing which is usually done when it arrives at the processing centre is that the tape is converted to a SEG-Y tape, and from then on only SEG-Y is used. SEG-D has some advantageous features for use in the field. One is that it can deal with multiplexed as well as demultiplexed data. Multiplexed data is that the data is not stored trace by trace, but that all the first samples of all traces (receivers) for one shot record are stored, then the second samples of all the traces, and so on. This way of storing very much links up to how the electronics deals with the data, i.e. how the data is sampled. Another advantage of the SEG-D format is that various word lengths are allowed which makes it possible to make use of less storage space than the conventional 4 bytes. As a final remark it can be said that the format is an awful standard to implement and understand, the description of it takes up more than 30 pages. This can be found in the publication by the SEG ([SEG, 1980]).

SEG-Y

The other tape standard mentioned is SEG-Y. This standard has been published [Barry et al., 1975] and has been enclosed in appendix C. A very nice feature of the standard is that it is not only a published standard, but it has also been adopted as

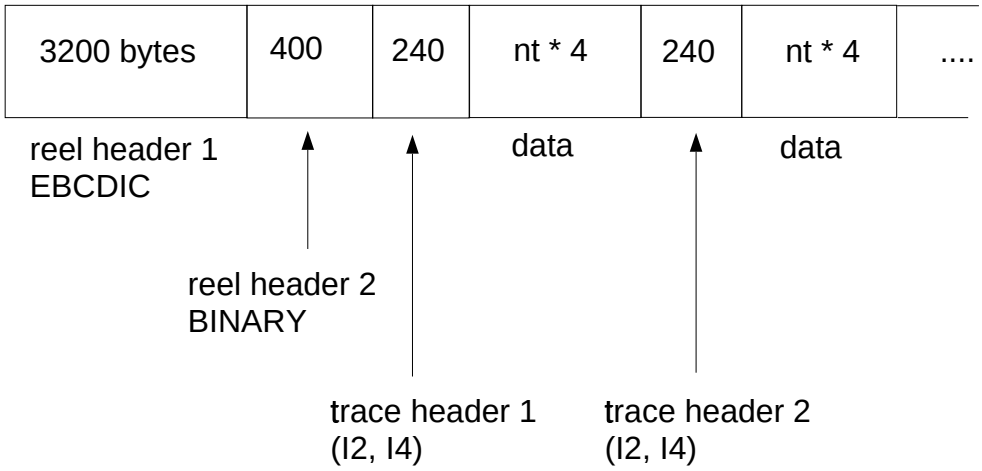


Figure 4.1: General set-up of a SEG-Y file

a standard for the oil industry. This has made data communication very simple, and most data transfer takes place via SEG-Y.

SEG-Y is a tape standard which was defined on top of the at-the-time standard of IBM-floating point, this means that floating-point numbers are stored as such. A SEG-Y tape consists of two so-called reel headers, so these are headers which occur only once on a tape. There are two reel headers, one being an EBCDIC header of 3200 bytes, and another being a binary header of 400 bytes. In the EBCDIC header one can put general information about the data on the tape, such as where it was shot and when, and by who, and so on. How this is filled is completely arbitrary, as long as it is in EBCDIC (which can easily be converted to ASCII). For the binary reel header this is different. In this header the way is built up is very precise, as can be seen in appendix C. In this header there are a few positions which are recommended to be filled (and thus MUST be filled), and are emphasized by an asterisk in the appendix. These are data such as line number, sampling rate, number of samples per trace, and so on.
The next blocks on a SEG-Y tape are trace blocks which consists of a trace header and the values of the trace itself. This trace header is by definition 240 bytes large, and information of the trace can be set in this header. Again, some positions are recommended to be filled and thus must be set and are emphasized by an asterisk in the table in appendix C. There are totally six fields which must be set, namely trace sequence number in the line, the original field record number, the trace number within the original field record number, a trace identification code, the number of samples, and the sampling interval; optionally, other information like source and receiver coordinates, water depth, elevation information, etc. After this header the values of the trace are given in one record, of course as many as given in the header. After these trace blocks comes the header of the next trace, and then the values of that next trace, and so on (see also figure (4.1)). For more details the reader is referred to again appendix C.

Summary

In this section the mostly used data formats for seismic reflection data are briefly discussed: SEG-D and SEG-Y.

4.3 Surveying information and field geometry

The next step after converting SEG-D to SEG-Y is incorporating the survey results and the field geometry. In the first step all the information from the surveyors has to be put into the SEG-Y reel and trace headers (i.e., coordinates, elevation information, etc.). It should be realized that the surveying of a seismic survey is an enormous amount of work, both at sea and on land. Sometimes it can take half a year to have processed the surveying data from a 3-D marine seismic survey. It should be remembered that surveying is a basic need of any seismic method: if one doesn't know where the measured data is located, the results are not of much use. Also, when the surveying is not done properly, one can choose the processing parameters as precise as possible, but the results will always be degraded by the surveying errors. Normally, positioning errors that are made vary between 1 and 5 meters.

When we set up the field geometry, we use the surveying data to specify how we have shot in the field. Setting the field geometry of a marine survey is rather simple compared to a land survey; on land we often have many obstacles which had to be avoided, or we had to follow a road which wasn't on a straight line. In setting up the field geometry, we must then also use the seismic observers log in order to avoid errors. Many types of processing problems arise from incorrectly setting the field geometry.

4.4 Trace editing and balancing

Trace editing

The next step is looking at the data itself. The data must be examined on their quality, and the seismic observers log must be worked through. An observers log is of report of all the measurements that have been done in the field and some comments about the acquisition circumstances and possible hardware problems. An example of part of such a log for a marine survey is shown in figure 4.2. In every survey there are bad shots and bad traces and they must be deleted or repaired from the file. Bad shots can be shots that triggered the instrument but not the shot itself, or the shots are not recorded at all. Bad traces can be dead traces (being completely zero), or traces that show polarity reversal due to bad soldering in the field, noisy traces due to 50 Hz power lines; many other reasons may exist why a certain shot or trace was bad. At sea, there is no way to repeat a shot at a certain position, as the boat will have to keep moving. On land, shots may be repeated at the same position such that bad shots can just be removed from the survey. When the seismic observer has done his or her job properly, it should be mentioned in the log why these features occurred. Trace editing is often a tedious and time-consuming job, but if this is not done properly, the resultant stack can sometimes be awful. In practice, when people start to process data for the first time, they often do a lot of other processing as well (e.g. make a raw stack) just to find out that there are many spurious features present in the stack. When the starting processor finds out what caused them, it often appears that they included a bad shot or a bad trace.

Trace balancing

The next step in a normal procedure is equalizing the amplitudes of samples, traces and shots. With regard to equalizing amplitudes on one trace, it may be obvious that reflections late in the section will be of a much smaller amplitude than the ones early in the section, simply due to energy losses in the subsurface. The most important energy losses are due to geometrical spreading, absorption of the rocks and transmission losses. Geometrical spreading is spreading of the wavefront since the energy of a wave is inversely proportional to $1/r^2$ (in a homogeneous medium). This means that the amplitude is decaying as $1/r$. Although the measured field is from a correct physical experiment, some aspects cannot be dealt with in standard processing. Therefore, corrections are made for the raw data to match with the assumptions in further data processing. The first assumption is that the data is processed as if it was measured in a two-dimensional world, which means a spreading which is proportional to $1/\sqrt{r}$ for the amplitude. A correction is made on the traces with a spreading function. This geometrical spreading

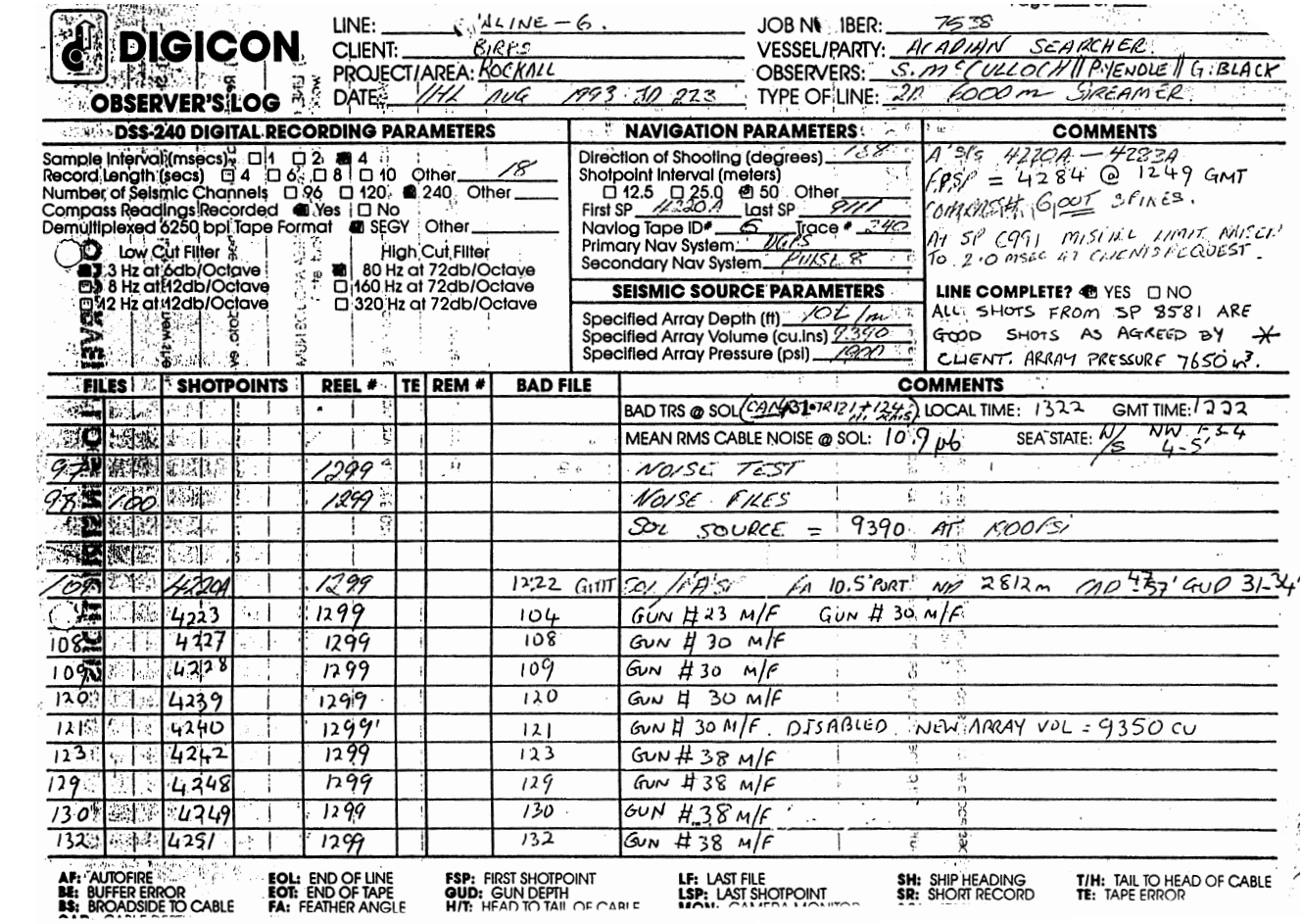


Figure 4.2: Example of an oberserver's log page for a marine dataset.

function is specified as a traveltime and a specific average velocity function. The other (gain) correction is applied to correct for absorption, and this one is corrected with an exponential gain function, where the factor in the exponential is related to the average absorption coefficient of the rocks. Then, the last correction is to confront the transmission losses at interfaces in the subsurface, although this correction is usually included in the exponential gain for the absorption losses.

The other amplitude equalization is within the traces of one shot. In one shot it can happen that a certain trace has over the whole trace a smaller amplitude than any of the neighboring traces, and this can be due to e.g. a different sensitivity of a geophone, or perhaps bad coupling of the geophone to the ground. This could be corrected for by calculating a total energy of a trace and making the energy of each trace in a shot the same, or at least smoothly varying as a function of offset.

Then the third equalization is equalization of different shots. Shots may generate different amounts of propagating energy due to different coupling conditions, and this is most severe for dynamite on land. This equalization becomes important when a different sorting than source-receiver is used, such as cmp-offset (which is mostly done). In one CMP, we combined traces from different shots, and if all the shots haven't generated the same signal energies, the traces in the CMP will show this. It may be possible to equalize those traces, but it is more reasonable, i.e. more related to the real cause, to make the energies of all the shots equal before any sorting is done.

A last equalization is purely for plotting purposes and does not serve to correct for a physical phenomenon, as discussed in the last three paragraphs. This equalization is Automatic Gain Control or AGC. With this type of scaling a trace sample is scaled to the sum of the absolute values of the neighbouring sample values. The number of time samples around this sample is the window length $W + 1$. In formula, AGC is applied by a scaling function $s_{AGC}[t]$ as follows:

$$x_{out}[t] = s_{AGC}[t]x_{in}[t] \tag{4.1}$$

where

$$s_{AGC}[t] = \frac{1}{\frac{1}{W+1}\sum_{\tau=t-W/2}^{t+W/2}|x_{in}[\tau]|} \tag{4.2}$$

With seismic data, it customary to plot the data with AGC because it makes the amplitudes along the section of comparable amplitude.

Summary

In this section some energy equalization methods are briefly discussed: geometrical spreading, receiver– and shot equalization, and Automatic Gain Control

4.5 Static corrections

Elevation and weathering statics

When we record our seismic data in the field, the ideal case would be that there would be no topography at all and that the shallow subsurface would not influence our resultant section at all. In the situation of marine seismics, this ideal situation is almost the case, but the case on land is much different. On land, we usually have topography, and we need to correct for that topography. In seismic terms, if one geophone would be on top of a small hill and the others not, the sound would need extra time to arrive at the geophone on the hill. This means that all the reflections of the subsurface would arrive later than in the neighboring traces. In order to get the timing right, we would have to apply some time shift to the whole trace. Applying a time shift is called a static correction; it is called static because it is one time correction for the complete trace, thus all samples will be shifted with the same constant amount. More specifically, when we correct for static shifts due to topography, we call it the elevation correction, the elevation referring to some datum level.

Besides the elevation statics, there is another type of statics: the weathering statics. This is caused by some low-velocity chunk of material right below the geophone. This kind of static shift is classified as a static shift due to the weathered layer, which is a general term for static shifts due to the top layers of the earth which are often weathered and irregular. Although in theory, the static shift for both types depends also on angle of incidence of the reflected waves, in practice a single time shift works often satisfactorily.

On land, these two types of static shifts, elevation and weathering, are often determined in a separate refraction survey. This is done because the data from the seismic reflection survey is not suitable for a good analysis of the shallow subsurface, and also because in many cases the static shifts can have such a detrimental effect on the stack that it is worth spending money on a separate survey in order to determine a model which explains these shifts as good as possible. The approach here is that we can determine the structure of the shallow subsurface in a deterministic fashion. The static corrections which are determined via field data, obtained from the refractions in the reflection data or from a separate refraction survey, are called *field* statics. The static corrections for surface topography are the elevation corrections, while corrections for the weathered layer are called weathering corrections.

The procedure of field statics is, in the order of being applied to the data:

- Weathered-layer correction
- Elevation correction

This is illustrated in figure (4.3). These corrections do a good job in the sense that reflectors in the subsurface align much better, but it may still not be good enough and often an extra static correction is applied which makes the alignment of reflectors even better. This aligning is not based on any physical measurement, but based on maximizing a correlation between adjacent traces. This is called the *residual* static correction. This

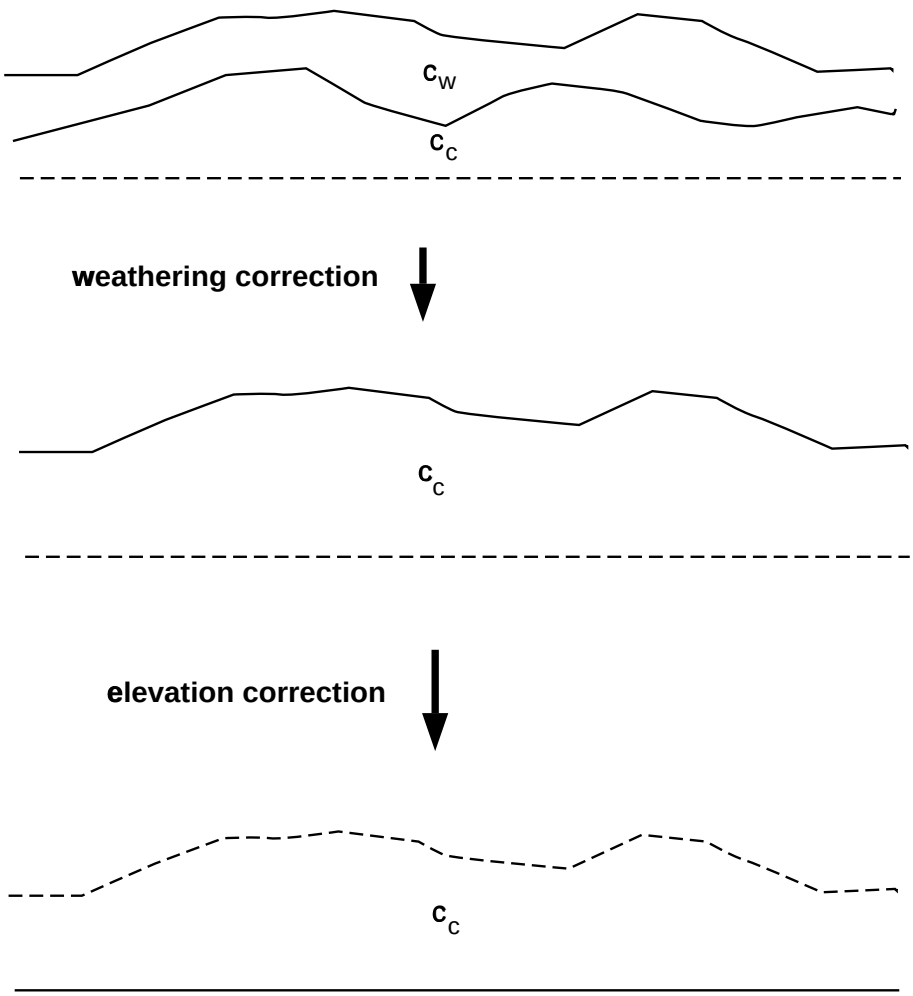


Figure 4.3: Static correction procedure: removing effects of weathered layer (from top to middle figure), and of surface topography (from middle to bottom figure).

correction may relate to some physical feature, perhaps due to unexplained features in the model determined from the refractions, but in essence it is a cosmetic correction. It should be realized that determining the residual static correction goes much better when the field statics have been applied and therefore determining the residual static without having applied field statics can be very difficult due to the irregular nature of the cross-correlation function.

The weathering corrections

Weathering corrections are applied to correct for differences in reflection time from trace to trace, due to variations in the transit time of the weathered layer along the seismic traverse. The weathered layer is a zone of less compacted young sediments with low seismic velocities. It is only present in the first few meters below the surface and may show considerable variations in thickness as well as in seismic velocity. Due to these variations in the weathered layer, irregular differences in arrival time between corresponding seismic

signals on different seismic traces of one shot record and between signals on corresponding traces from different shots may occur. As a result of these irregular "delay-time" variations, the reflection events on the records show an irregular (noisy) character.

The weathering correction is a traveltime correction; the delay time of the weathered layer is subtracted from the observed traveltime and replaced by a delay time of material with the velocity c_c of the first consolidated layer, i.e.:

$$\Delta T_w = -\frac{d_w}{c_w} + \frac{d_w}{c_c} \tag{4.3}$$

in which d_w denotes the thickness of the weathered layer and c_w denotes the seismic velocity of the weathered layer. In fact, the weathered layer is replaced by a layer with the velocity of the first consolidated layer. In the following we shall derive the necessary expressions to determine this expression, since we do not know d_w on beforehand.

The following derivation is one which allows irregular refracting boundaries between the weathered and first consolidated layer; it is also known as the plus-minus method. For simplification, we shall assume that the boundary is a dipping boundary; however, the analysis is also valid for irregular boundaries.

The computation of the weathering correction can be derived from separate refraction surveys. With a refraction survey, the distance between source and receiver is large compared to the depth of investigation: we must make sure the rays are critically refracted into the refractor and therefore we need large angles of incidence. Assuming the shot is fired near the surface, the wave first travels in the weathered (upper) layer, then travels in the first consolidated layer horizontally along the base of the weathered layer, and refracts critically through the weathered layer to the geophone.

With a refraction survey, it is standard procedure is to shoot from both sides in order to be able to determine possible dips in the refracting boundary. In appendix D, the traveltime T of the ray for one source and one geophone is derived:

$$T = \frac{1}{c_c \cos\alpha} x + \tau_S + \tau_G \tag{4.4}$$

where x is the distance between source and receiver, α is the angle of the refracting boundary with the horizontal, c_c is the wave speed in the refracting medium (the first consolidated layer), and τ_S and τ_G are some delay times at the source and receiver, respectively. What a delay time means, can be seen simplest by assuming a horizontal refractor ($\alpha = 0$); consider figure 4.4(a). The time of the part of the ray which goes from the refractor to the geophone G, is given by:

$$T_{AG} = T_{AD} + T_{DG}, \tag{4.5}$$

where A is the point from which the wave is diffracted from the base of the weathered layer. Now we can write the traveltime T_{AD} as

$$T_{AD} = \frac{AD}{c_w} = \frac{x_3}{c_w}\sin\theta_c = \frac{x_3}{c_w}\frac{c_w}{c_c} = \frac{x_3}{c_c} = T_{AG'} \tag{4.6}$$

in which c_w is the wave speed in the weathered layer, x_3 is the horizontal distance of the upgoing ray, and θ_c is the critical angle. Thus, the time from A to D is now converted in

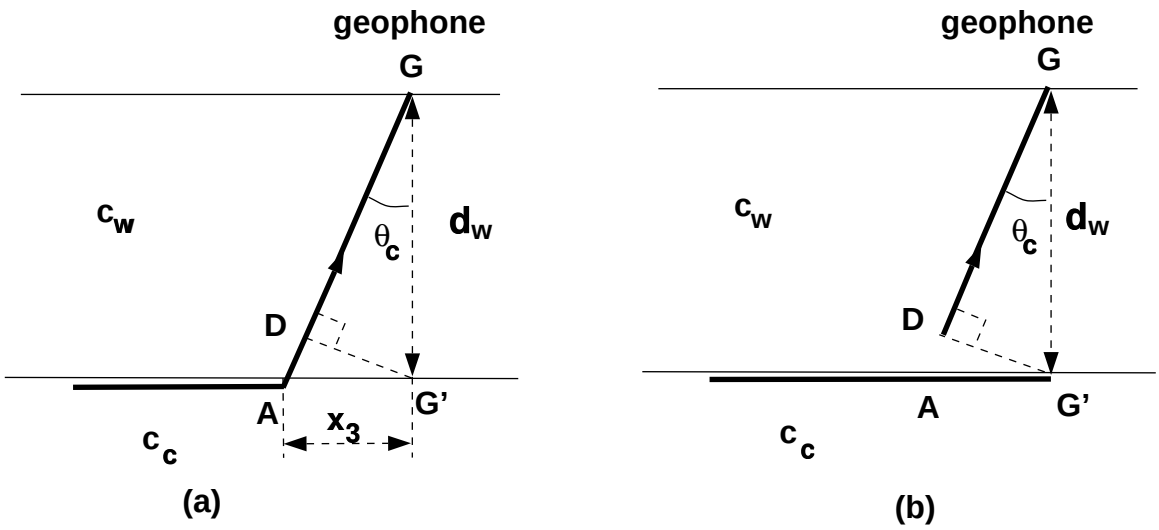


Figure 4.4: Horizontal reflector: upgoing ray with critical angle (a), with its equivalent path (b).

a time for the ray to travel the horizontal distance; this is shown in figure (4.4b). In the case of a dipping refractor, this term is included in the term $x/(c_2 \cos\alpha)$; the delay time is then taken with respect to the *normal* of the dipping boundary. The other term T_{DG} can be seen as a delay time, denoted by τ_G, so T_{AG} can be written as:

$$T_{AG} = \frac{x_3}{c_c} + \tau_G \tag{4.7}$$

Let us now consider the situation as drawn in figure (4.5); we have taken again only one dipping refracting boundary. In this situation we have a geophone placed between two shots S_1 and S_2. It is also assumed that a geophone is at position S_2 when shot S_1 is fired, and vice versa. Let us denote the total distance between shot S_1 and S_2 by L, and the distance from shot S_1 to the geophone G by x. Then the traveltimes from shot S_1 and S_2 are given by:

$$T_{S_1S_2} = \frac{1}{c_c \cos\alpha} L + \tau_{S_1} + \tau_{S_2} \tag{4.8}$$

$$T_{S_1G} = \frac{1}{c_c \cos\alpha} x + \tau_{S_1} + \tau_{G1} \tag{4.9}$$

$$T_{S_2G} = \frac{1}{c_c \cos\alpha} (L - x) + \tau_{S_2} + \tau_{G2} \tag{4.10}$$

where the numbers 1 and 2 in τ_{G1} and τ_{G2} denote that they are due to shot 1 and 2, respectively. When we add the last two equations, and substitute the first into it, the following equation is obtained:

$$T_{S_1G} + T_{S_2G} = T_{S_1S_2} + \tau_{G1} + \tau_{G2} \tag{4.11}$$

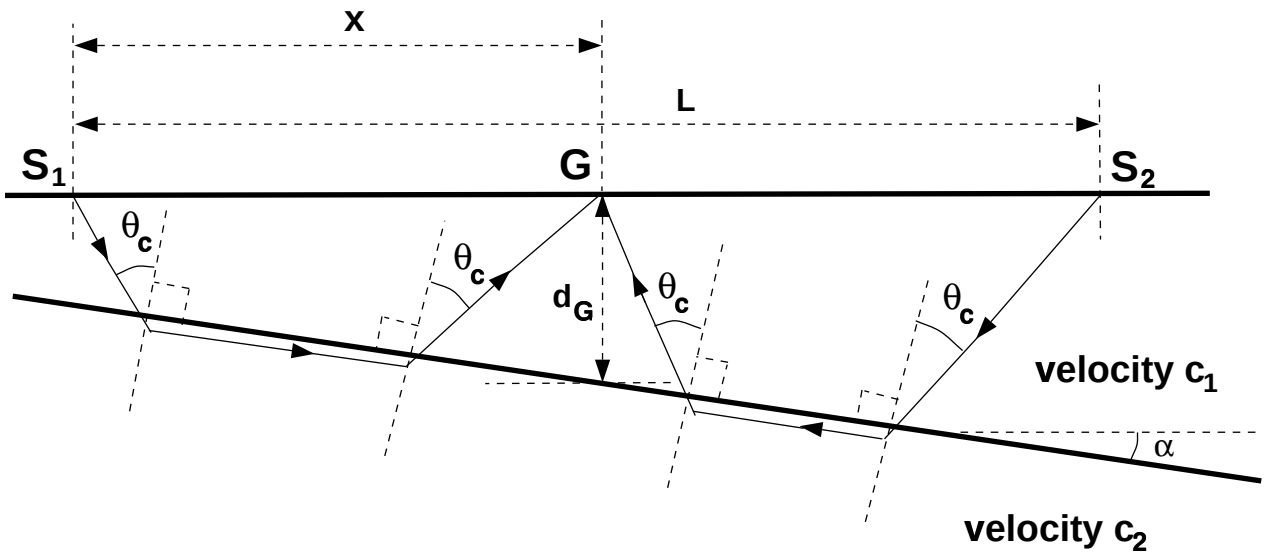


Figure 4.5: Configuration of dipping refractor for determining delay times of geophones, using shots from both ends.

Hence,

$$\tau_{G1} + \tau_{G2} = T_{S_1G} + T_{S_2G} - T_{S_1S_2} \tag{4.12}$$

What can now be seen in this equation, is that the delay times of the geophone can be determined from the times which are observed in the refraction seismograms.

The most important result from this equation, is when we look at the ray paths of the different shots. What can be seen, is that the delay times are partly obtained from *subtracting the paths which are similar as indicated by the numbered ray segments* ① *to* ④ This means, that the delay time which is obtained here, is a *local* traveltime. In the case of a dipping refractor, the analysis is exact; however, when the refractor is irregular, the above is an approximation but still applicable as long as the boundaries are not too irregular.

In the weathering correction which needs to be applied, the depth of the weathered layer is used; however, we do not know the depth on beforehand. To that end, consider the delay time $\tau_{G1} + \tau_{G2}$, i.e., (see also appendix D):

$$\begin{aligned} \tau_{G1} + \tau_{G2} &= d_G \left(\frac{\cos(\theta_c - \alpha)}{c_w} + \frac{\cos(\theta_c + \alpha)}{c_w} \right) \\ &= d_G \frac{2\cos\theta_c \cos\alpha}{c_w} \end{aligned} \tag{4.13}$$

Using this equation, the depth at the geophone d_G can be expressed in the delay time:

$$d_G = \frac{c_w(\tau_{G1} + \tau_{G2})}{2\cos\alpha\cos\theta_c} \tag{4.14}$$

Let us now turn to the weathering correction, as mentioned in the introduction of the section. The weathering correction is defined as the difference between the actual travel time to a point at the surface and the traveltime it would have taken to reach the same surface point if the weathered layer were replaced by material having the velocity of the first consolidated layer. We thus obtain (see figure (4.4)):

$$\begin{aligned}
\Delta T_w &= \frac{d_G}{c_c} - \frac{d_G}{c_w} = d_G \frac{c_w - c_c}{c_c c_w} \\
&= \frac{(\tau_{G1} + \tau_{G2})}{2 \cos\alpha \cos\theta_c} \frac{c_w - c_c}{c_c} = \frac{(\tau_{G1} + \tau_{G2})}{2 \cos\alpha} \frac{1}{(1 - c_w^2/c_c^2)^{1/2}} \frac{c_w - c_c}{c_c} \\
&= -\frac{(\tau_{G1} + \tau_{G2})}{2 \cos\alpha} \left(\frac{c_c - c_w}{c_c + c_w} \right)^{1/2} \qquad (4.15)
\end{aligned}$$

Now using the delay time as derived earlier (eq.(4.12)), the final expression becomes:

$$\Delta T_w = -\frac{(T_{S_1 G} + T_{S_2 G} - T_{S_1 S_2})}{2 \cos\alpha} \left(\frac{c_c - c_w}{c_c + c_w} \right)^{1/2} \qquad (4.16)$$

The factor of the square root $\sqrt{(c_c - c_w)/(c_c + c_w)}$ is of the order 0.6 to 0.9 and can often be assumed to vary slowly along the seismic traverse or to be constant in a particular region. For the calculation of this factor the weathering velocity c_w must be known; this may be determined from the short distance geophones of the refraction spreads.

When the seismic source is an explosive, buried under the weathered layer, it is also possible to determine the weathering corrections at the shot-points from the so-called uphole time: a geophone is planted close to the shot hole in order to determine the vertical travel time from the source to the surface. The weathering correction is found by subtraction of the "uphole time" T_{up} from the source depth Z_{shot} divided by the velocity of the first consolidated layer c_c:

$$\Delta T_w = \frac{Z_{shot}}{c_c} - T_{up} \qquad (4.17)$$

At the shot holes, the factor with the square root in equation (4.15) can now be checked by comparing the weathering correction from the uphole times with the weathering corrections derived from geophones. Increased density of shooting along the seismic traverse, as is the case in multiple-coverage shooting, increases the number of checkpoints along the seismic line accordingly.

Elevation corrections

Once the weathering corrections have been applied, elevation corrections have to be applied for sources and detectors. These elevations are determined via surveyors who have determined the coordinates, so also the elevation, with their instruments. In the calculations of the time correction it is assumed that the weathered layer has been replaced by material having the velocity c_c of the first consolidated layer. The elevation correction for the shot is:

$$\Delta T_{e,s} = \frac{Z_S - d_S}{c_c} \qquad (4.18)$$

in which Z_S is the elevation of the shot and d_S is the shot depth. The elevation correction for the geophones is:

$$\Delta T_{e,g} = \frac{Z_G}{c_c} \tag{4.19}$$

in which Z_G is the geophone elevation. These corrections have to be subtracted or added to the weathering corrections, depending on the chosen datum level (station above or below datum).

Residual statics

In the former discussion of refraction spreads, it should be realized that refraction surveys have their limitations, such as non-detection of low-velocity zones (no critical refractions!) and blind zones, and thus refraction statics are not perfect. Therefore, when these corrections have not cleaned up the data sufficiently, one can apply the so-called residual static corrections. Normally, the residual statics are in the order of the dominant seismic wavelength. They can be determined via calculations of time shifts using cross-correlations between pairs of seismic traces belonging to large sets of multiple-coverage seismic reflection data, being CMP gathers, shot gathers, offset gathers and/or geophone gathers. These statistically determined static corrections are often applied in a "surface-consistent" way: the same time shift at the same location. So if the multiplicity (or fold) of the data is N, and there are M surface positions, one has to solve $M \times N$ equations with M unknowns. These residual static corrections can be determined more efficiently when the field statics have been applied, with less iterations and with a more accurate result.

Summary

The static errors due to a weathered layer and surface topography are derived. The weathering correction is usually determined from refraction spreads. In this section this correction is derived, where the final result only contains traveltimes, the velocities of the layers and a dip, if present; no depths are needed. To seismic reflection data, the weathering correction is applied first. Next, the elevation correction is applied to correct for surface topography. Then finally, residual statics are applied to line reflections up even better to correct for errors not determined from the refractions spreads.

4.6 Deconvolution

Deconvolution concerns itself with removing a part of the data which is convolutional. For instance, we know that a total seismic response consists of the convolution of the seismic source wavelet with the earth response, convolved with the response from the seismic detector, convolved with the seismic response from the recording system. If we consider only the seismic source signature $s(t)$ and the impulse response of the earth $g(t)$, then the seismic signal can be written as:

$$x(t) = g(t) * s(t). \tag{4.20}$$

Usually, we are not interested in the responses from the seismic source, detector or recording system, so we want to remove them. The most critical response in this list is usually the seismic source. Removing the seismic source from a seismic recording is called signature deconvolution. We can distinguish two types of deconvolution: *deterministic* and *statistical* deconvolution: the first type describes the deconvolution of a known signature $s(t)$ from the data, whereas the second type describes the removal of signature or other effects from the data based on some statistical assumptions (e.g. "whiteness" of reflectivity series, minimum phase assumption of wavelet).

4.6.1 Deterministic deconvolution in the frequency domain

Let us assume we have a wavelet with a known spectrum, $S(f)$. Neglecting receiver and recording-system responses, the response is seen as a convolution of only the earth response and the seismic source in the time domain. Then the convolution becomes a multiplication in the frequency domain:

$$X(f) = S(f)G(f), \tag{4.21}$$

in which $X(f)$ is the spectrum of the seismic recording, and $G(f)$ is the spectrum of the earth response. Now if we want to remove the source signature, then we have to divide each side by $S(f)$, or equivalently apply the inverse operator $F(f) = 1/S(f)$ to each side, obtaining:

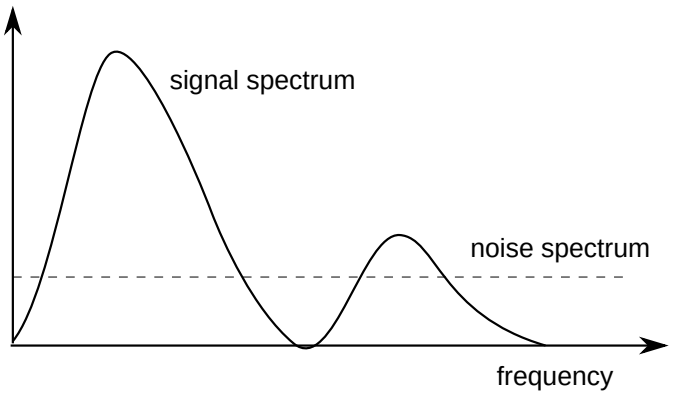
$$\frac{X(f)}{S(f)} = G(f). \tag{4.22}$$

Of course, this states the problem too simple. For instance, a seismic recording always contains noise. When the seismic recording is taken as the earth response together with some noise term, i.e., $X(f) = S(f)G(f) + N(f)$ in which $N(f)$ denotes the noise term, then the deconvolution in the frequency domain becomes:

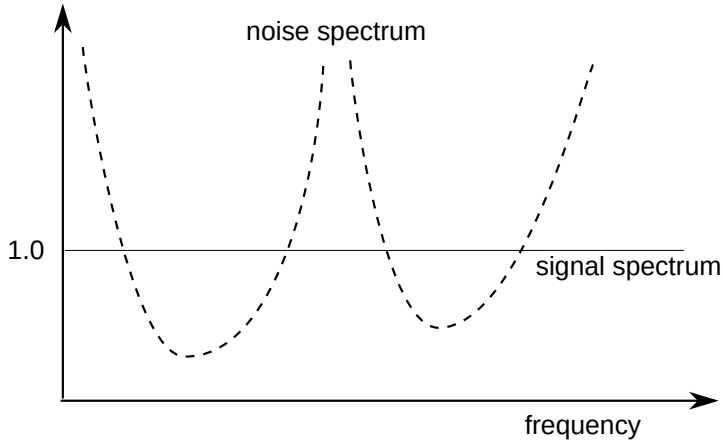
$$\frac{X(f)}{S(f)} = G(f) + \frac{N(f)}{S(f)}. \tag{4.23}$$

The next problem is that due to this division, the noise is blown up outside the bandwidth of signal $S(f)$. This effect is shown in figure (4.6).

There are two ways to tackle this problem. The first one is that we stabilize the division. This is done by not applying a filter $F(f) = 1/S(f)$ but first multiplying both



a) Before deconvolution



b) After deconvolution

Figure 4.6: The effect of deconvolution in the frequency domain in the presence of noise.

the numerator and the denominator by the complex conjugate of the source spectrum, $S^*(f)$, and since the denominator is now real we can add a small (real) constant ϵ to it. Thus instead of $1/S(f)$, we apply the filter:

$$F(f) = \frac{S^*(f)}{S(f)S^*(f) + \epsilon^2}. \tag{4.24}$$

Often we take ϵ as a fraction of the maximum value in $|S(f)|$, e.g. $\epsilon = \alpha MAX(|S(f)|)$ with α in the order of 0.01 - 0.1. In this way we have controlled the noise, but it can still be large outside the bandwidth of $S(f)$ (see figure (4.6)). As an example, figure (4.7) shows the result for deconvolution

The other way of dealing with the blowing up of the noise is only doing the division in a certain bandwidth which is equivalent to shaping the wavelet $s(t)$ into a shorter one, which we call $d(t)$. In fact the signal $d(t)$ is our desired output wavelet on the seismic

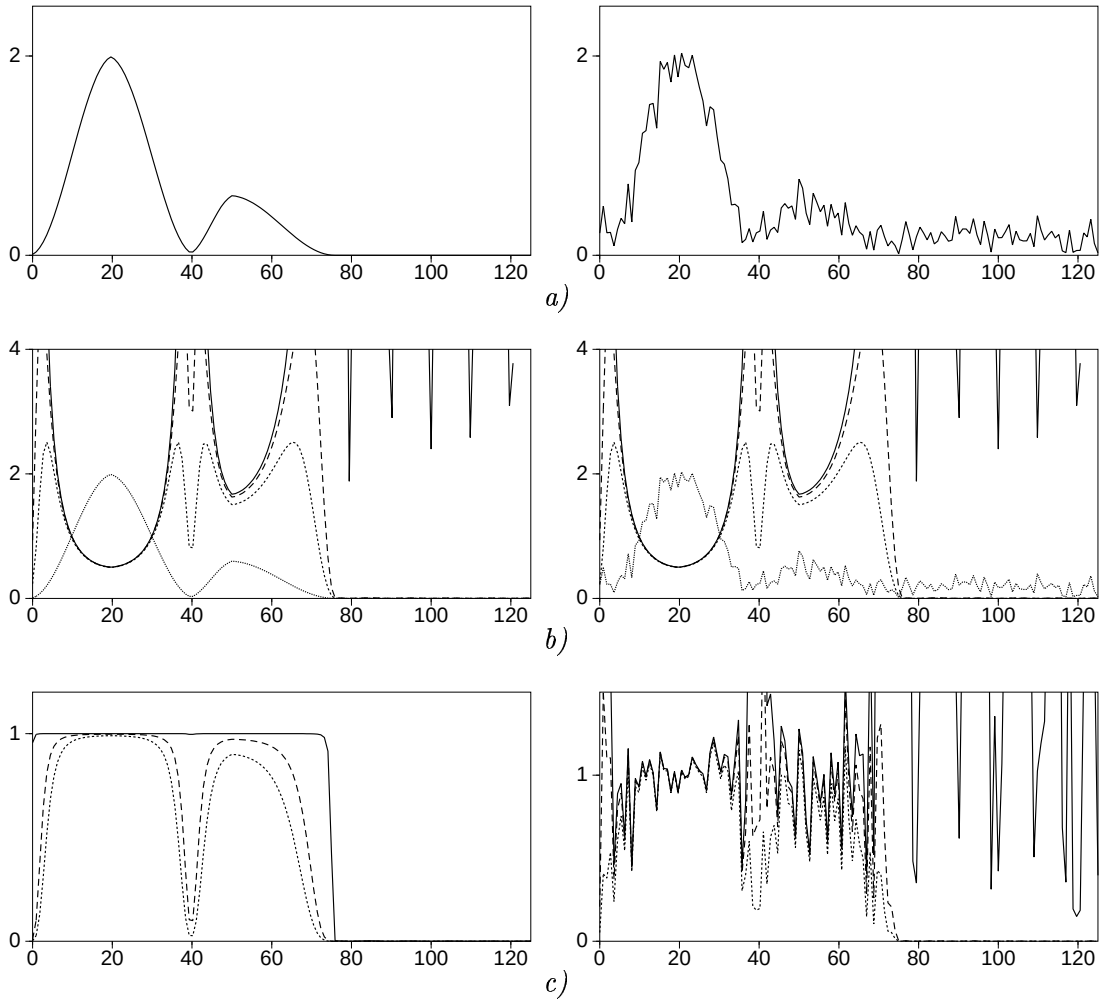


Figure 4.7: Applying stabilized inversion in the frequency domain, left for a noise free wavelet, right for a noisy wavelet. a) Spectrum of signal to be inverted. b) Spectra of inverse operators with 3 stabilization constants (e = 0, 0.05, 0.1). c) Multiplication of inverse filters with original spectrum of a), i.e. the deconvolution results.

data, which is often a nice-looking signal in the time domain, such that we have optimal resolution in our seismic sections. In this case we do not apply the filter $1/S(f)$ but instead we use $D(f)/S(f)$. Then the deconvolution amounts to:

$$\frac{X(f)D(f)}{S(f)} = G(f)D(f) + \frac{N(f)D(f)}{S(f)}, \tag{4.25}$$

where $|D(f)|$ is approximately equal to $|S(f)|$, i.e.:

$$a < \frac{|D(f)|}{|S(f)|} < b, \tag{4.26}$$

in which b/a is less than 10, say. As mentioned, we would like to end up with a signal that is short in the time domain. This means that the spectrum of $D(f)$ must be smooth

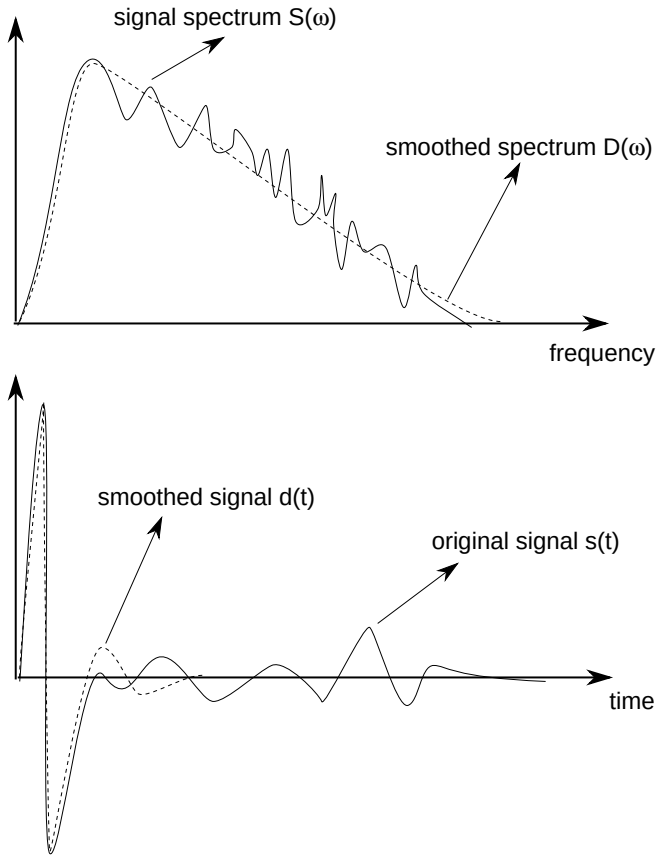


Figure 4.8: Designing a desired wavelet via smoothing in the frequency domain.

compared to the true wavelet spectrum $S(f)$. Note that a short signal in time corresponds with a smooth (i.e. oversampled) signal in frequency, as the major part of the time signal will be zero. Practically this means when we know the spectrum we can design some smooth envelope around the source spectrum $S(f)$, or we can just pick a few significant points in the spectrum and let a smooth interpolator go through these picked points. An example of designing such a window is given in figure (4.8).

As a last remark of deconvolution in the frequency domain it can be said that in practice both ways of control over the division by $S(f)$ are used. That is that we apply a filter :

$$F(f) = \frac{D(f)S^*(f)}{S(f)S^*(f) + \epsilon^2} \tag{4.27}$$

to the data, resulting in:

$$\frac{X(f)D(f)S^*(f)}{S(f)S^*(f) + \epsilon^2} = \frac{G(f)D(f)S(f)S^*(f)}{S(f)S^*(f) + \epsilon^2} + \frac{N(f)D(f)S^*(f)}{S(f)S^*(f) + \epsilon^2}. \tag{4.28}$$

This is about the best we can do given the constraints of bandwidth and signal-to-noise ratio.

4.6.2 Deterministic deconvolution in the time domain : Wiener filters

So far we have dealt with deconvolution in the frequency domain which was just a modification of a division of spectra. Let us look specifically at the filter $F(f)$ for the case that the stabilized division and shaping is applied, i.e.:

$$(S(f)S^*(f) + \epsilon^2)F(f) = D(f)S^*(f), \tag{4.29}$$

in which we have brought the denominator of the division in equation (4.27) to the other side. When we transform this equation back to the time domain, we obtain:

$$\{\phi_{ss}(t) + \epsilon^2\delta(t)\} * f(t) = \phi_{ds}(t) \tag{4.30}$$

These equations are called the *normal equations*. In words, by calculating the autocorrelation of the signal $s(t)$ and the cross-correlation of the desired signal $d(t)$ and the actual signal $s(t)$, the corresponding filter can be determined via the above equations. The length of the filter can be chosen by the user. Note that the definitions of auto-correlation and cross-correlation can be found in Chapter 2. In Appendix E the concepts of correlation are further explained and some examples are shown.

These normal equations can also be derived starting from a different point of view, namely minimizing the squared error between a desired output and an actual output, as is common for Wiener filters. In that situation we want to minimize the error between a desired signal $d(t)$ and the convolution of the in put signal $s(t)$ and a filter $f(t)$, and stabilizing this by also minimizing the energy of the coefficients of the filter:

$$E = \sum_t [d(t) - f(t) * s(t)]^2 = minimum. \tag{4.31}$$

In this way the length of the filter $f(t)$ can be chosen according to the user's definition. This may be desired because we do not always need an *exact* deconvolution filter, but a *short* filter that will do 95% of the job. To avoid that the filter coefficients $f(t)$ will take large values, often the energy of the coefficients of the filter are minimized as well:

$$E = \sum_t [d(t) - f(t) * s(t)]^2 + \epsilon^2 \sum_t f^2(t) = minimum. \tag{4.32}$$

The minimum energy can befound by taking the derivative of the above equation to the filter coefficients and equlaizing them to zero. We shall not do this derivation here but refer the interested reader to Appendix F. The result of this derivation is the set of *normal equations* as described above. There exist a fast scheme of calculating the filter coefficients for such a system, namely the Levinson recursion scheme (e.g. see [Robinson and Treitel, 1980]), which makes it possible to determine the coefficients efficiently on a computer. But in order to use this scheme we must rewrite the system to a slightly different form such that a Toeplitz matrix arises. To this effect, write the normal equations as a discrete convolution:

$$\sum_{n=0}^{N} \phi_{ss}[i-n]f[n] + \epsilon^2 f[i] = \phi_{ds}[i] \qquad \text{for} \quad i = 0, 1, 2, ..., N, \tag{4.33}$$

where we use square brackets to denote that we deal with discrete data. Note that the filter $f[i]$ is a causal one here. Writing this in matrix form, using the fact that the autocorrelation is symmetric around zero correlation time, i.e. $\phi_{ss}[n] = \phi_{ss}[-n]$, we obtain:

$$\begin{pmatrix} \phi_{ss}[0]+\epsilon^2 & \phi_{ss}[1] & \phi_{ss}[2] & \cdots & \phi_{ss}[N] \\ \phi_{ss}[1] & \phi_{ss}[0]+\epsilon^2 & \phi_{ss}[1] & \cdots & \phi_{ss}[N-1] \\ \phi_{ss}[2] & \phi_{ss}[1] & \phi_{ss}[0]+\epsilon^2 & \cdots & \phi_{ss}[N-2] \\ \vdots & & \ddots & & \vdots \\ \phi_{ss}[N] & \phi_{ss}[N-1] & \phi_{ss}[N-2] & \cdots & \phi_{ss}[0]+\epsilon^2 \end{pmatrix} \begin{pmatrix} f[0] \\ f[1] \\ f[2] \\ \vdots \\ f[N] \end{pmatrix} = \begin{pmatrix} \phi_{ds}[0] \\ \phi_{ds}[1] \\ \phi_{ds}[2] \\ \vdots \\ \phi_{ds}[N] \end{pmatrix} \tag{4.34}$$

The matrix with the autocorrelation coefficients has the Toeplitz form, which is exploited by Levinson's algorithm. This algorithm reduces computation from N^3 operations to N^2, where N is the number of rows or columns in the matrix. Thus, this can be computed efficiently on a computer.

What is now the advantage of determining the filter in the time domain? An advantage is that the formulation makes it easy to take any number of coefficients for the filter $f[n]$. Also, we can take a few negative filter coefficients, and many positive ones, or vice versa; this can be advantageous in many circumstances. When doing the deconvolution in the frequency domain there is no control on the length of the filter in the time domain. Moreover, it should be realized that in order to do the deconvolution in the frequency domain, we have to transform the data to the frequency domain first, which takes up some computing time. This becomes only worthwhile for a certain number of filter samples, say about 100. For a small number of filter coefficients, it is more efficient to calculate the filter coefficients via the Levinson scheme.

Via the time domain, we have a lot of flexibility in the choice of the filter coefficients, such as length, desired wavelet and the amount of negative and positive samples. But how can we optimize these parameters? In order to address this problem, consider the wavelet as given in figure (4.9). Is there a good (short) causal filter that converts the wavelet into a spike? In general, yes, but we should not require the spike to be at a zero delay. The optimal choice is usually desiring a spike that has a delay that corresponds to the point where the wavelet $s[k]$ has most of its energy.

This can be inspected via defining the partial energy of a signal, i.e.:

$$p[k] = \sum_{n=-\infty}^{k} s^2[n] \tag{4.35}$$

and plotting the partial energy as a function of k. Note that most signals in seismics are causal so the summation will start at $n = 0$. Using the partial energies we can classify certain wavelets. For instance, the *minimum-phase* or *minimum delay* wavelet has the fastest energy build-up, thus most of the energy is in the beginning. In the same way, a *maximum phase* or *maximum delay* wavelet has the slowest energy build-up, so most of the energy is at the end. Any case in between these extremes is called a mixed-phase wavelet. The wavelet of figure (4.9) is a mixed-phase wavelet and its cumulative energy is shown in figure (4.10). It appears that the main peak is around 12 ms (fastest energy build-up). It is now clear that the optimum delay of a spike for a minimum-phase wavelet,

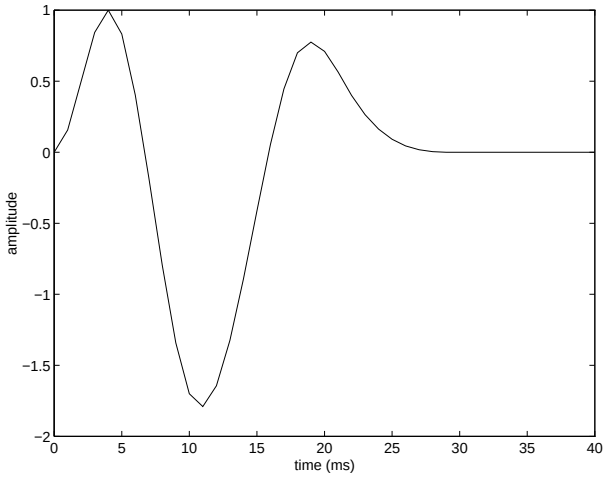


Figure 4.9: A mixed-phase wavelet to find the inverse for.

is no delay. For any mixed phase wavelet, we must delay the spike of the desired output signal. This delay time os often called the *lag*.

For Wiener filters, there is some measure of how good the filter does its work, and that is of course via the error which is calculated via the square of the difference between the actual output, and the desired output. The performance P of the deconvolution is defined as:

$$P = 1 - \frac{\sum_n [d[n] - f[n] * s[n]]^2}{\sum_n d^2[n]}, \tag{4.36}$$

which is a number between 0 and 1, $P = 1$ meaning that we have a perfect filter. Let us consider again the wavelet of figure (4.9) and design deconvolution filters for a fixed filter length of 20 ms, but with different lag times in the desired output signal. Figure (4.11) shows the result for the deconvolution for a number of lag times, with the dotted lines showing the desired output signal, being the delayed pulse. The best result is achieved for

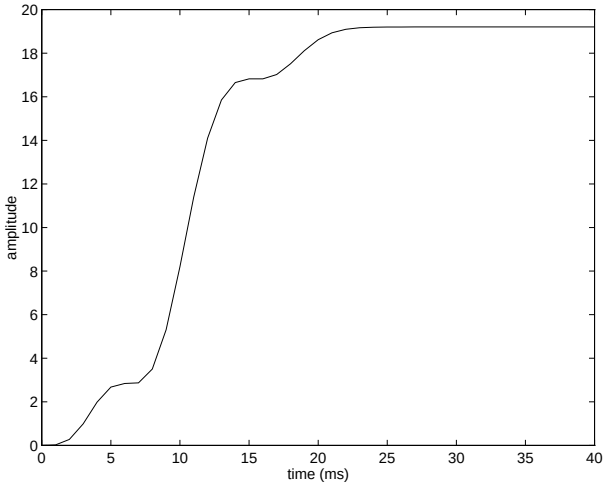


Figure 4.10: Partial energy function for the wavelet of figure (4.9).

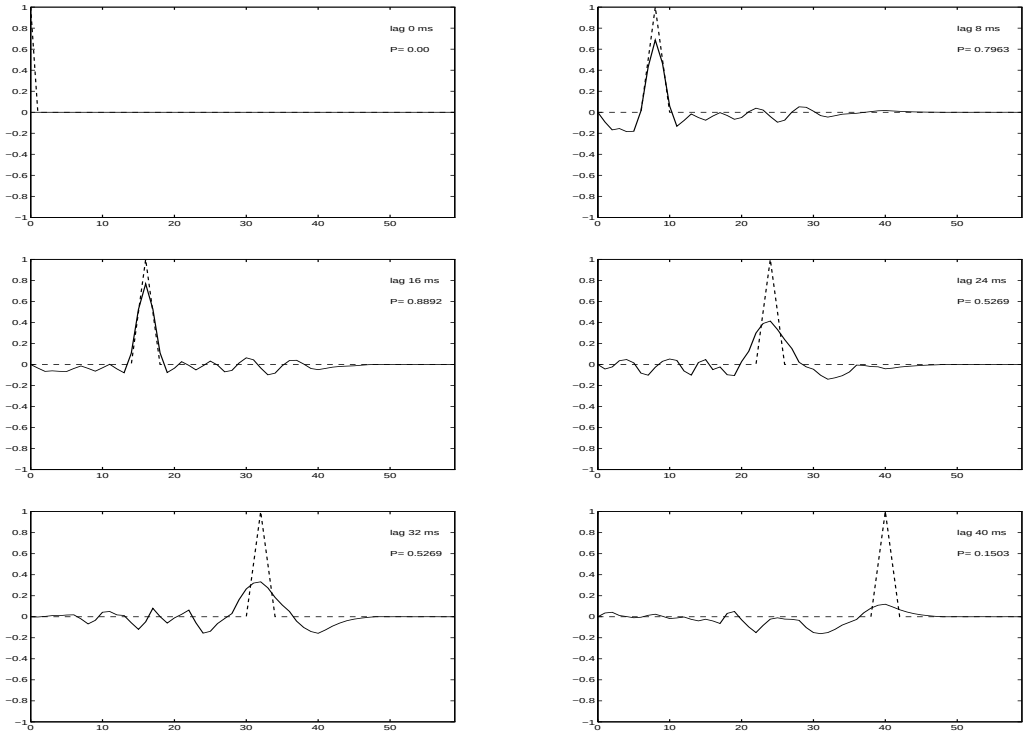


Figure 4.11: The actual output for spiking deconvolution for the wavelet given in figure (4.9), for different delays. The dashed line defines the desired wavelet, i.e. a shifted spike. The length of the filter is 20 ms

lag 16 ms. If we repeat this experiment for a range of lag times and plot the performance as a function of lag, figure (4.12) is the result. There we can see that the optimum lag of this wavelet is 14 ms.

As is obvious, when we increase the filter length, the performance will be better; in the extreme of infinite filter length, we will find an (almost) perfect inverse filter. However, with a finite delay of the spike and filter coefficients which start at t=0 (i.e. causal filter as considered in these examples), we cannot obtain the exact inverse filter. We can also make a curve which gives the performance as a function of the filter length. Such an example is given in figure (4.13) for our wavelet under consideration. The lag has now be kept constant at 14 ms. As can be seen in this figure, is that the asymptotic value does not go up to 1, but a lower constant value. This is the best we can do with a finite delay, a causal filter with an infinite number of samples.

So the two important parameters in optimizing the filter design, are the *lag* (i.e. delay) in the desired output, and the *filter length*. We can always improve the performance by increasing the filter length and changing the lag time. Therefore, for design purposes, it is useful to plot values for P as a function of lag and filter length as is given in figure (4.14). Note that the previous two figures showed cross-sections of this 2 dimensional function.

There is some rule of thumb when designing filters. That is that if we inspect the

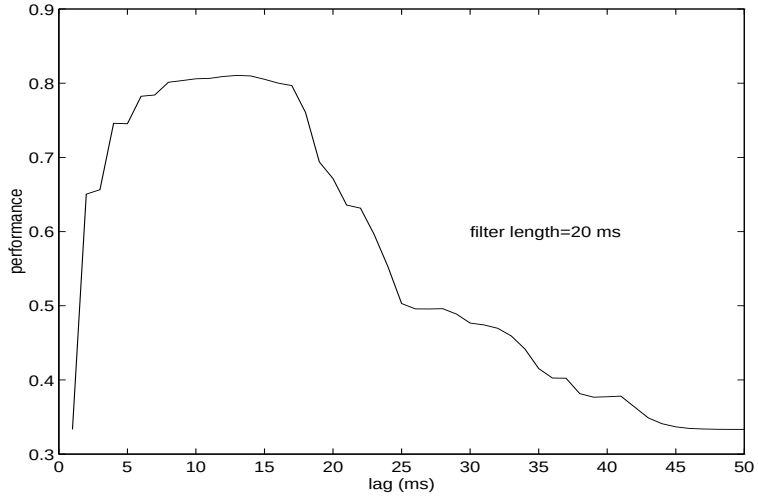


Figure 4.12: The performance curve as function of delay for spiking deconvolution for the wavelet given in figure (4.9) for a filter length of 20 ms.

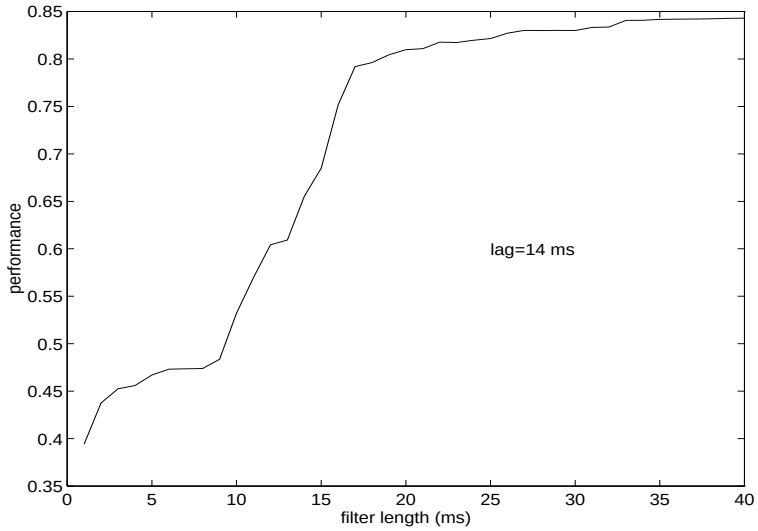


Figure 4.13: The performance of the deconvolution filter as a function of filter length for spiking deconvolution of the wavelet as given in figure (4.9) with a lag of 14 ms.

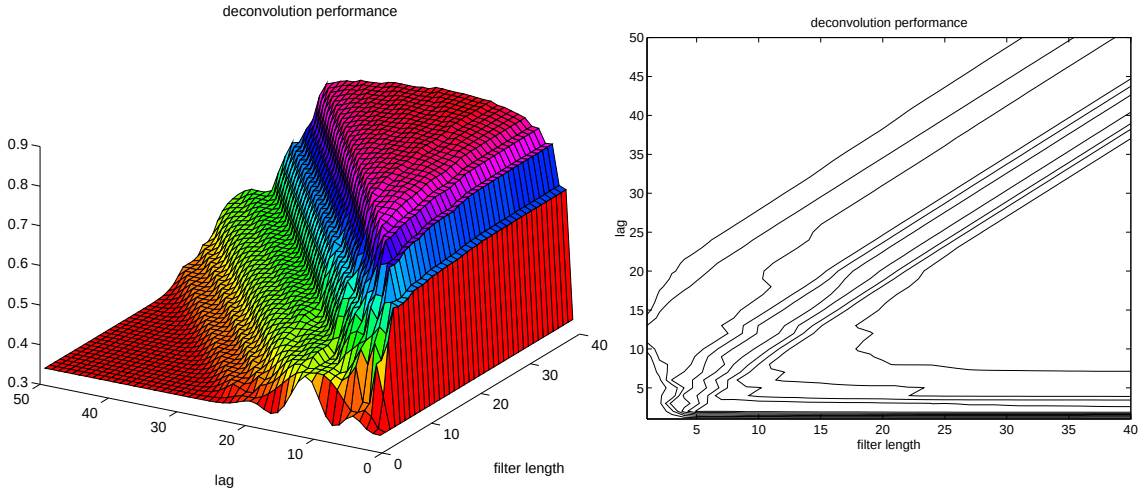


Figure 4.14: 3d and contour plot of the spiking deconvolution filter performance as a function of filter length and lag for the wavelet as given in figure (4.9).

wavelet and look where most of the energy is (via partial energies) then the optimum lag distance will be half the distance of the filter length and the peaking of the energy of the wavelet. So, for instance, if the wavelet seems to put its energy mostly around 12 ms, and we choose a filter length of 20 ms, the optimum lag will be around 0.5*(12+20) = 16 ms. This rule of thumb is found to be quite useful in practice.

There is perhaps one parameter which we have left out in the discussion so far and that is the stabilization constant ϵ^2. This parameter is not only needed in the deconvolution via the frequency domain, but also for the deconvolution in the time domain. In order to solve the matrix system for the filter in a stable fashion, we need to add some noise to the diagonal of that matrix. Common values for this stabilization factor are 1% to 5% of the maximum of the power of the wavelet ($\phi_{ss}[0]$).

4.6.3 Statistical deconvolution: minimum phase deconvolution

Up to now we have assumed that the wavelet where we want to deconvolve for is known. However, in practice this is most of the time not the case. Therefore, common practice is to make some assumptions on the statistical properties of the wavelet and of the earth response. One of these approaches leads to the method of *minimum phase deconvolution*. For this we go back to the convolution model of the earth response, as given in equation (4.6):

$$x(t) = g(t) * s(t). \tag{4.37}$$

Suppose we now define our desired deconvolved earth response as $d(t)$ by applying filter $f(t)$ to the measured response $x(t)$ and we want that $d(t) = g(t)$. This can be achieved

by solving the following energy equation:

$$E = \sum_t [g(t) - f(t) * x(t)]^2 = minimum. \tag{4.38}$$

Note that at this stage $g(t)$ is still unknown, but we assume it for the moment to be known. Similar as described in the previous sub-sections, solving this will lead to the following normal equations (omitting the stabilization constant):

$$\phi_{xx}(\tau) * f(\tau) = \phi_{gx}(\tau), \tag{4.39}$$

where $\phi_{xx}(\tau)$ is the autocorrelation $x(t)*x(-t)$ and $\phi_{gx}(\tau)$ is the cross-correlation $g(t)x(-t)$. The following expressions can now be formulated:

$$\begin{aligned} \phi_{xx}(\tau) &= x(\tau) * x(-\tau) \\ &= s(\tau) * g(\tau) * s(-\tau) * g(-\tau) \\ &= g(\tau) * g(-\tau) * s(\tau) * s(-\tau) \\ &= \phi_{gg}(\tau) * \phi_{ss}(\tau), \end{aligned} \tag{4.40}$$

for the autocorrelation function and

$$\begin{aligned} \phi_{gx}(\tau) &= g(\tau) * x(-\tau) \\ &= g(\tau) * s(-\tau) * g(-\tau) \\ &= \phi_{gg}(\tau) * s(-\tau) \end{aligned} \tag{4.41}$$

for the cross-correlation function. With these expressions the normal equation can be written as:

$$\phi_{gg}(\tau) * \phi_{ss}(\tau) * f(\tau) = \phi_{gg}(\tau) * s(-\tau) \tag{4.42}$$

i.e. fully in terms of the autocorrelation of the true earth response, the autocorrelation of the wavelet and the wavelet itself.

Now, the assumptions come into play. First let us assume that the actual earth response $g(t)$ is *statistically "white"* :

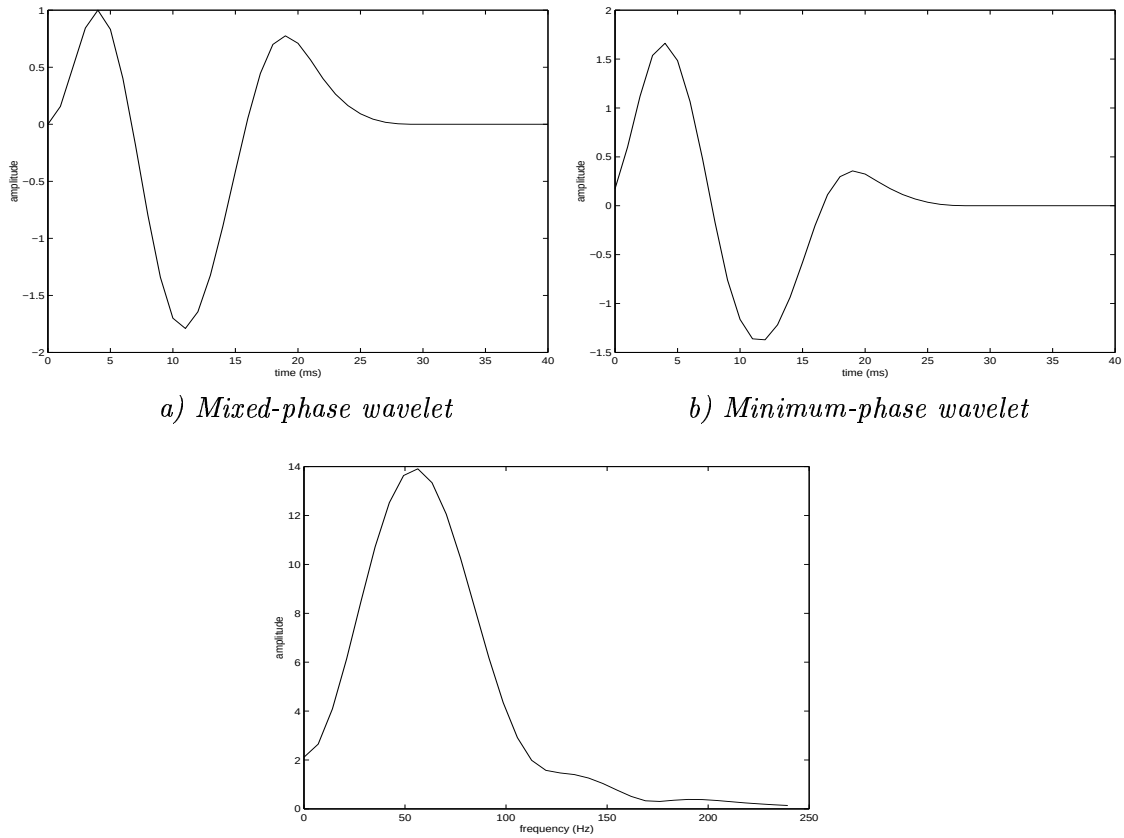
$$\phi_{gg}(\tau) = G\delta(\tau), \tag{4.43}$$

which means that it consists of a series of uncorrelated reflections. In practice we know that this is not the case in general, although to some extend this assumption is acceptable (see also the examples in Appendix E). With this assumption, the normal equation can be rewritten as:

$$\phi_{ss}(\tau) * f(\tau) = s(-\tau). \tag{4.44}$$

The influence of the earth has been fully removed from this equation. In fact we assume that by looking at the auto-correlation of the seismic signal, the earth reflectivity only gives a scaling factor, but that the auto-correlation can be directly interpreted as the auto-correlation of the source signal (besides a scaling factor).

The second assumption is that the wavelet is considered to be *minimum phase*. This means that given the power spectrum of the wavelet ($\phi_{ss}(\tau)$), the time domain response $s(\tau)$ can be constructed by defining that the wavelet is as short as possible, i.e. it has the fastest possible energy build-up. See also [Robinson and Treitel, 1980] for more details on



a) Mixed-phase wavelet

b) Minimum-phase wavelet

c) Amplitude spectrum

Figure 4.15: A mixed phase wavelet (a) and its minimum phase equivalent (b). The amplitude spectrum (c) is the same for both signals.

the minimum phase concept. At this stage it is important to know that given the power spectrum of a signal, its minimum phase time domain representation can be constructed. In Figure 4.15 an example of the minimum phase equivalent of the mixed-phase wavelet of Figure 4.9 is given. Note that the amplitude spectrum is the same of both signals (by definition) and that the minimum phase equivalent of the input signal has most of its energy concentrated around zero time. In that case the signal $s(-\tau)$ in equation (4.6.3) can be determined from the power spectrum and the filter $f(t)$ can be calculated. As the factor G in equation (4.6.3) is not known, the (assumed minimum phase) source wavelet can only be determined up to an absolute scale factor. This means that the filter ($f(t)$ will remove this wavelet only up to this absolute scale factor.

In this subsection, we have seen that under two assumptions: the whiteness of the reflection series and the minimum phase behavior of the source signal the effect of the source can be removed from the seismic traces.

4.6.4 Statistical deconvolution: predictive deconvolution

We can now assume that the wavelet has been shortened by minumum phase deconvolution or the wavelet has already a short duration in time (i.e. dynamite wavelet). However, there are other effects in the seismic response that are undesired and can be removed using the Wiener filter deconvolution method. One of these effects is reverberations of the seismic wavefield between two reflectors, for example reverberations in the water layer. Figure (4.16) shows the effect of a water layer on the seismic response. Each up-going reflection will be followed by a train of reverberations, alternating in sign because the surface has a reflection of approximately -1. These reverberations have a nice property, i.e. they repeat themselves after the reverberation period $\tau = 2\Delta z/c$, with Δz the thickness of the water layer and c the water velocity, and each reverberation is scaled by $-R_1$, the reflectivity of the water bottom. Deconvolution via the time domain has specific applications in the removal of these multiple reflections, and the Wiener filter designed for this is called *predictive deconvolution.*

The main idea of predictive deconvolution is that if we delay the signal with α seconds, which is chosen in the order of $\tau = 2\Delta z/c$ seconds, and adapt the amplitudes (by convolving with filter f_t), it will fit again in itself. This means that a primary reflection will fit with the first order reverberation, the first order reverberation will match the second order reverberation, etc. So we want to minimize the error E for the filter coefficients f_t according to the following formulation (see also appendix F):

$$E = \sum_t (x_t - f_t * x_{t-\alpha})^2, \tag{4.45}$$

or, equivalently,

$$E = \sum_t (x_{t+\alpha} - f_t * x_t)^2. \tag{4.46}$$

The crucial factor in this type of deconvolution is choosing the desired wavelet: we choose it as $d_t = x_{t+\alpha}$, thus the original signal, but then delayed with α. The solution can again be defined as a set normal equations as follows (neglecting the stabilization):

$$\phi_{xx}(t) * f(t) = \phi_{xx}(t + \alpha). \tag{4.47}$$

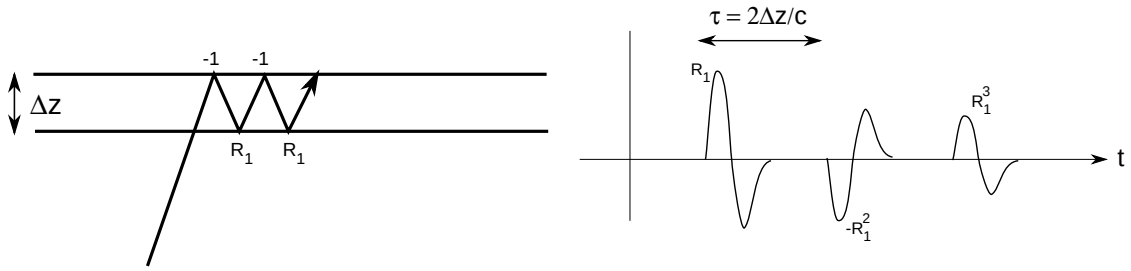


Figure 4.16: Reverberations in a water layer in depth (left) and its effect on the seismic trace (right).

If we consider a discrete version of this equation, introduce a short length filter $f[n]$ and also take stabilization into account, the normal equations are as follows:

$$\sum_{n=0}^{N} \phi_{xx}[i-n]f[n] + \epsilon^2 f[i] = \phi_{xx}[i+N_\alpha] \qquad \text{for} \quad i = 0, 1, 2, ..., N, \tag{4.48}$$

With N_α being the discrete equivalent of the time lag α. As in the subsection on Wiener filtering, this set of normal equations can be written as a matrix-vector equation as follows:

$$\begin{pmatrix} \phi_{xx}[0]+\epsilon^2 & \phi_{xx}[1] & \phi_{xx}[2] & \cdots & \phi_{xx}[N] \\ \phi_{xx}[1] & \phi_{xx}[0]+\epsilon^2 & \phi_{xx}[1] & \cdots & \phi_{xx}[N-1] \\ \phi_{xx}[2] & \phi_{xx}[1] & \phi_{xx}[0]+\epsilon^2 & \cdots & \phi_{xx}[N-2] \\ \vdots & & \ddots & & \vdots \\ \phi_{xx}[N] & \phi_{xx}[N-1] & \phi_{xx}[N-2] & \cdots & \phi_{xx}[0]+\epsilon^2 \end{pmatrix} \begin{pmatrix} f[0] \\ f[1] \\ f[2] \\ \vdots \\ f[N] \end{pmatrix} = \begin{pmatrix} \phi_{xx}[N_\alpha] \\ \phi_{xx}[N_\alpha+1] \\ \phi_{xx}[N_\alpha+2] \\ \vdots \\ \phi_{xx}[N_\alpha+N] \end{pmatrix} \tag{4.49}$$

In these equations only the autocorrelation of the seismic trace, $\phi_{xx}[n]$, and the shifted autocorrelation, $\phi_{xx}[n+N_\alpha]$, are used to define the filter samples $f[n]$.
Note that if we have found the optimum solution $f[n]$ we have to apply it on the data as the predictive deconvolution filter $\delta[n] - f[n-N_\alpha]$ to remove the reverberations from the input.

Figure (4.17) shows the result of a predictive deconvolution filtering procedure to a shot record from a field dataset. In the figure (4.18) the autocorrelations of the two shots have been displayed, showing clearly that predictive deconvolution removes the ringing in the autocorrelation function of the data.

Analytical example of predictive-error filtering

We will analyze the response:

$$G(\omega) = \frac{R_1 \exp(i\omega\tau_1)}{1 + R_1 \exp(i\omega\tau_1)}. \tag{4.50}$$

This is the normal-incidence reflection response from a layer above a half space. The upper boundary has reflection coefficient —1 (like the sea surface) and the lower boundary has a reflection coefficient R_1. The two-way traveltime is given by τ_1 (see also figure (4.19)). When we expand the denominator, we obtain:

$$G(\omega) = R_1 \exp(i\omega\tau_1)[1 - R_1 \exp(i\omega\tau_1) + R_1^2 \exp(2i\omega\tau_1) - R_1^3 \exp(3i\omega\tau_1) + \cdots]. \tag{4.51}$$

In each of these terms we recognize a multiple. In the time domain this gives the response:

$$g_t = 0, 0, \ldots, 0, R_1, 0, 0, \ldots, 0, -R_1^2, 0, 0, \ldots, 0, R_1^3, 0, 0 \ldots. \tag{4.52}$$

where we have $N_1 - 1$ zeroes at the beginning, with $N_1 = \tau_1/\Delta t$ and Δt being the discrete time sampling. Also between R_1 and $-R_1^2$ we have $N_1 - 1$ zeroes, between $-R_1^2$ and R_1^3 we have again $N_1 - 1$ zeroes, etc. We now assume we have measured this seismogram (with infinite bandwidth) so for the moment leave all the earlier considerations on source,

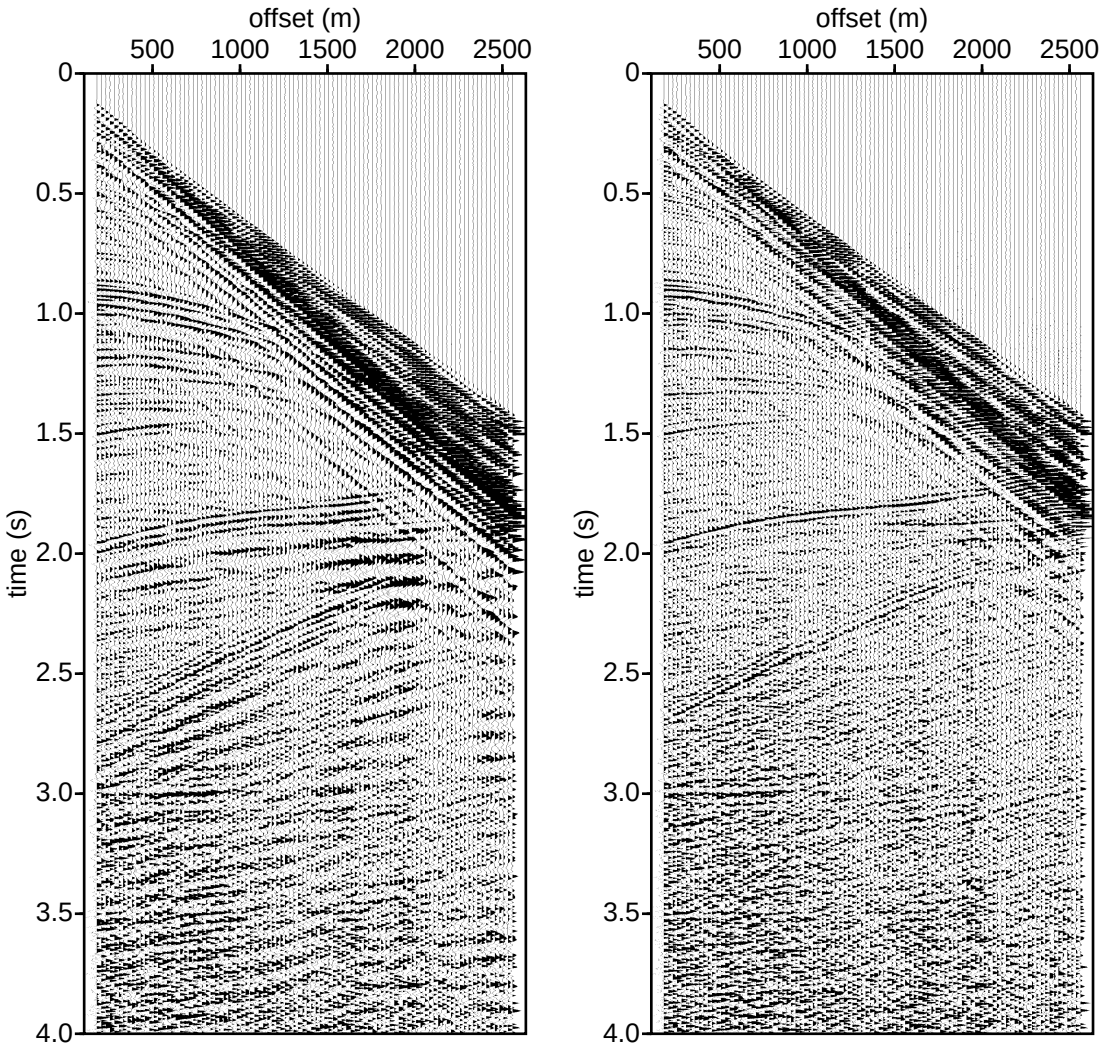


Figure 4.17: Shot record with reverberations (left) and after predictive deconvolution (right).

receivers, etc. out. We wish to convert this seismogram to a seismogram without multiples. For application of the Wiener-Levinson technique we require the auto-correlation function $\phi_{gg}(\tau)$ of g_t:

$$\begin{aligned} \phi_{gg}(\tau) &= R_1^2(1 + R_1^2 + R_1^4 + \cdots) \quad \text{for } \tau = 0 \\ &= 0 \qquad\qquad\qquad \text{for } 0 < \tau < \tau_1 \\ &= -R_1^3(1 + R_1^2 + R_1^4 + \cdots) = -R_1\phi_{gg}(0) \quad \text{for } \tau = \tau_1. \end{aligned} \tag{4.53}$$

Thus the discrete version of the autocorrelation of $g(\tau)$ can be written as:

$$\phi_{gg}, [n] = E_g, 0, 0, \ldots, 0, -R_1 E_g, \ldots. \tag{4.54}$$

where E_g denotes the energy of g_t, and we have $N_1 - 1$ zeroes. Let the filter length N_f be less than N_1, and let the prediction distance be $\alpha = \tau_1$, or $N_\alpha = N_1$. The normal

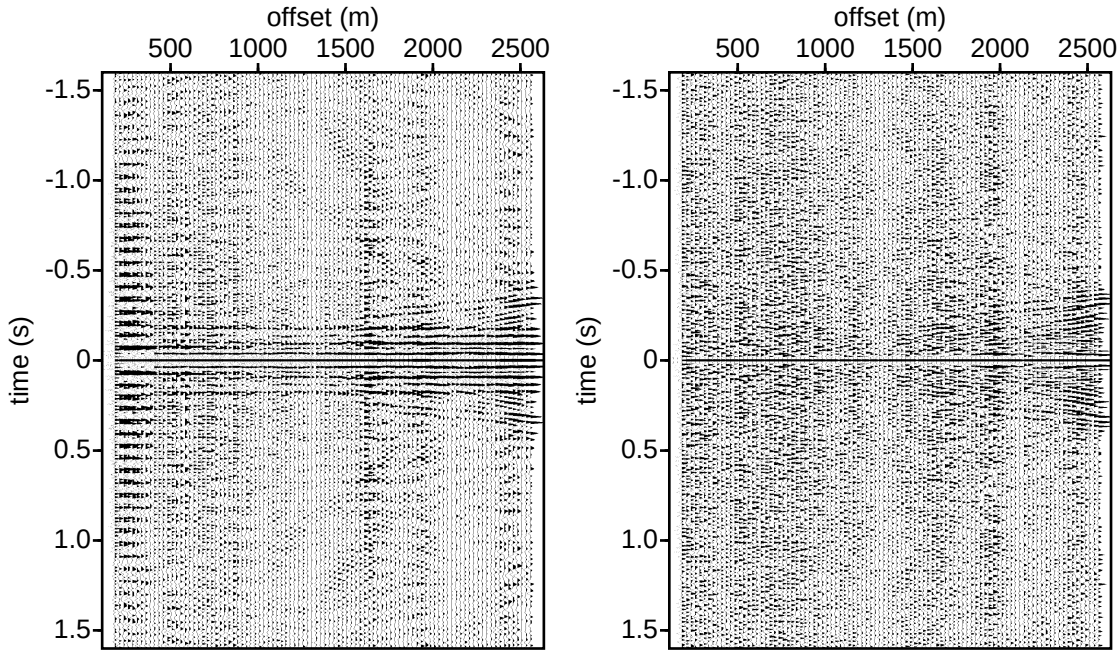


Figure 4.18: Autocorrelation functions of the shot record with reverberations (top left) and after predictive deconvolution (top right)

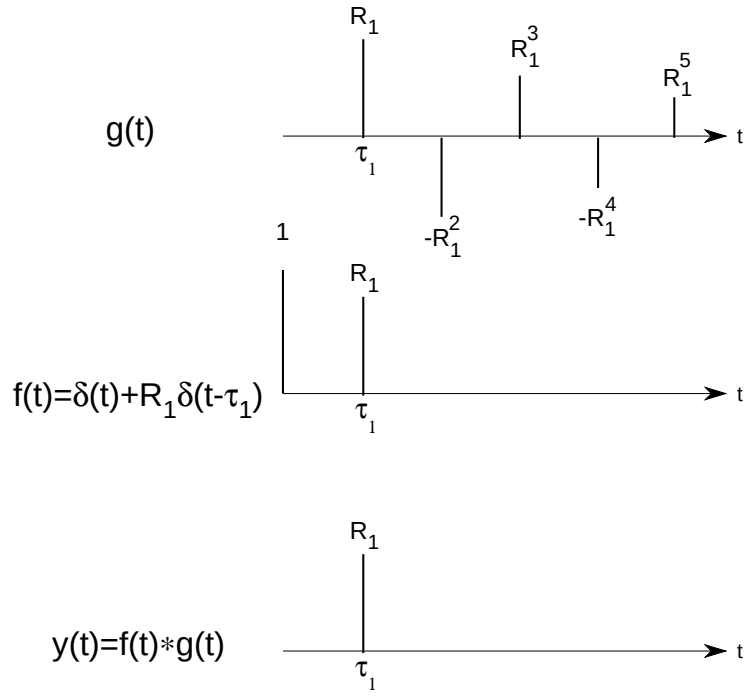


Figure 4.19: The input, operator and output for the analytical example of multiples within one layer.

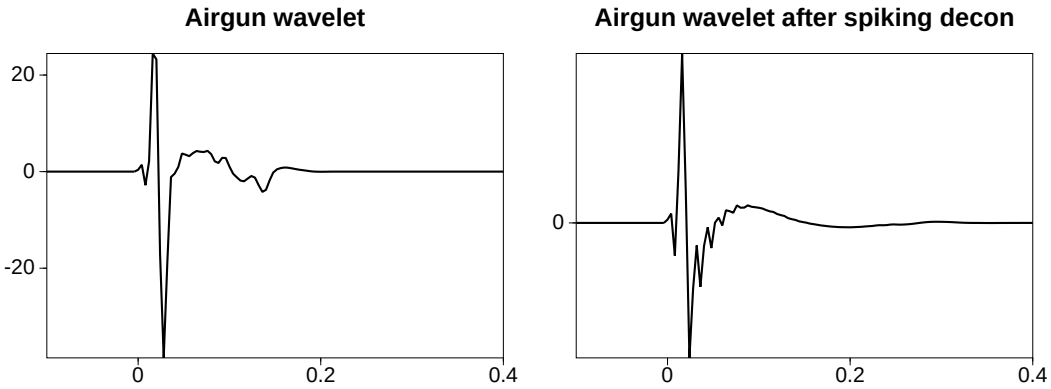


Figure 4.20: Measured airgun wavelet and the result after spiking deconvolution.

equations are then:

$$\begin{pmatrix} \phi_{gg}[0] & 0 & \cdots & 0 \\ 0 & \phi_{gg}[0] & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & \phi_{gg}[0] \end{pmatrix} \begin{pmatrix} f[0] \\ f[1] \\ \vdots \\ f[N-1] \end{pmatrix} = \begin{pmatrix} \phi_{gg}(N_1) \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{4.55}$$

The only member of this system whose right-hand side does not vanish is:

$$\phi_{gg}[0]f[0] = \phi_{gg}[N_1] \tag{4.56}$$

and thus

$$f[0] = \frac{\phi_{gg}[N_1]}{\phi_{gg}[0]} = \frac{-R_1 E_g}{E_g} = -R_1. \tag{4.57}$$

The total prediction-error operator becomes:

$$1, 0, 0, \ldots, R_1, \tag{4.58}$$

(note the minus sign disappearing). In practice it is not necessary to set the prediction distance exactly to N_1. The model here allows us to choose N_α to take any value as long as it is less than or equal to N_1. Also, we must have $N_\alpha + N_f > N_1$. So we can now apply our filter to the seismogram, as depicted in figure (4.19).

It is interesting to note that the original seismogram was periodic in time. This is only the case for normal-incidence data when we consider a plane-layered earth. If we do not have normal-incidence data, it is not periodic (in the (t, x) domain). In that case, the predictive deconvolution procedure can only partly solve the reverberation problem, but taking a filter length with $N_f > 1$ it can still handle the situation to some extend. Another solution of this is to go to the so-called linear Radon domain, or $\tau - p$ domain, where data is mapped into plane waves. There, the periodicity for each angle of incidence is again constant, and predictive deconvolution can be applied more accurately. A discussion of the linear Radon transform is beyond the scope of this course notes.

4.6.5 Spiking deconvolution

One of the most used deconvolution procedures in practice is the *spiking deconvolution.* In fact the spiking deconvolution is a special case of predictive deconvolution: by taking the lag to be one sample in predictive deconvolution, a spiking deconvolution is obtained! Examples of spiking deconvolution are shown in figures (4.20) and (4.21), in which spiking deconvolution results on a single airgun wavelet and on a synthetic shot record with this airgun wavelet are shown. Note that although the spiking is not perfect, clearly the temporal resolution is increased after the spiking deconvolution procedure: events have sharpened and each event starts with a large positive peak, which is desired in interpretation. It can be shown that spiking deconvolution performs best when the wavelet is minimum phase: then the spiking deconvolution filter is stable and appears to be a perfect inverse filter for the minumum phase signal.

Summary of deconvolution

With a deconvolution process an undesired distortion on the seismic signals is removed, which is either caused by the acquisition tools (e.g. airgun bubble effect) or by (the shallow) part of the earth (e.g. reverberations in a water layer). If this effect can be considered as a convolution effect on the desired seismic signals, by convolution with a filter with the inverse effect, this distortion can be removed from the data. In most cases the objective of deconvolution is to increase the time resolution of the data. Two main types of deconvolution are considered: 1) deterministic and 2) statistical deconvolution.

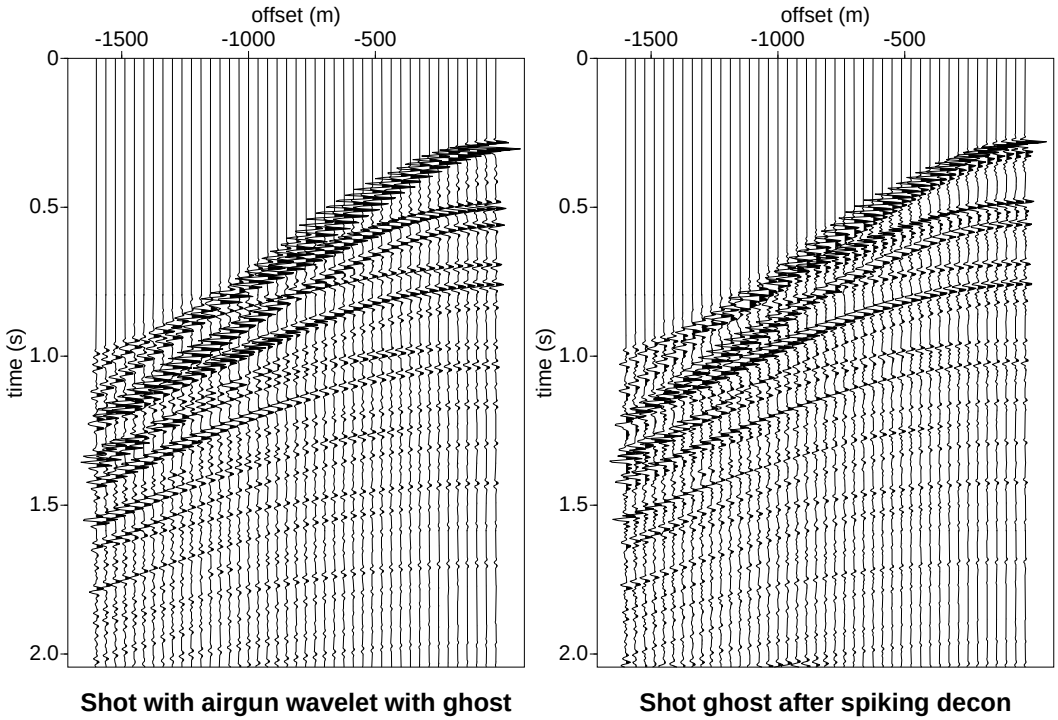


Figure 4.21: Synthetic shot record with measured airgun wavelet and the result after spiking deconvolution.

In the deterministic deconvolution, the convolutional distortion is known and need be removed in an optimal way. For the statistical deconvolution the undesired effect is not precisely known, but based on some statistical assumptions on the data this effect is reduced. Assumptions that are often used is uncorrelated reflectors, minimum phase behavior of the source signal and repeating patterns within the signal.

The deterministic deconvolution can be performed in either the frequency domain or in the time domain. Statistical deconvolution, with the aid of prediction error filtering, is always performed in the time domain. In the time domain domain the deconvolution problems can often be written as a set of so-called normal equations in which auto- and cross-correlation functions of the input signals are related. This type of equations can be written as a (Toeplitz) matrix-vector equation and can be solved efficiently with a Levinson recursion scheme. The advantage of applying the deconvolution in the time domain is that the user has a good control on the time length and possible instability effects. Often the deconvolution with a short filter in the time domain gives an optimal stable result, but with a slight loss in accuracy.

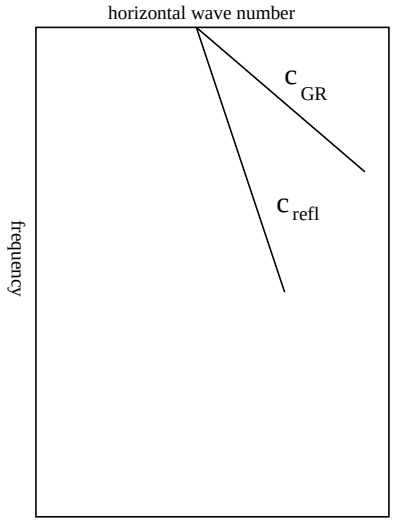


Figure 4.22: Ground roll and a reflection in the (f, k_x) domain.

4.7 Filtering in the (f, k_x) domain

We have seen in the chapter on basic signal analysis (chapter 2) that a dipping broadband event in the (t, x) domain gives a dipping event in the (f, k_x) domain. The dip is not the reciprocal of the velocity as in the (t, x) domain, but the velocity itself. In this section, we shall discuss two applications which make use of these characteristics of different slopes in the (f, k_x) domain, together with some other properties in this domain. These applications are the removal of ground roll in land data, and removal of multiples in marine data.

Ground roll filtering

Let us first look at ground roll in land data. Ground roll is the general term for surface waves which travel along the surface, and only if the earth would be a homogeneous half space, the ground roll would be a perfect Rayleigh wave, giving one event in the measurements. Since this is not the case for the real earth, it is a combination of resonances which show a very dispersive character due to interferences of the different modes.

At the same times that we receive the ground roll, we get the reflections back from the subsurface, so they overlap. A difference between these two types of arrivals, i.e. the ground roll and the reflections, is that they travel with a different speed along the surface. The ground roll travels along the surface with a relative low speed, while the reflections from below arrive nearly at the same time for a close group of receivers with a slight hyperbolic move-out because in reflection seismology we are aiming at reflections from small angles of incidence. In the (t, x) domain, we obtain the response as derived in chapter 2, i.e.:

$$p_{GR}(t, x) = s_{GR}(t - \frac{x}{c_{GR}}) \tag{4.59}$$

where s_{GR} is the wavelet for the ground roll, where the subscript GR stands for ground

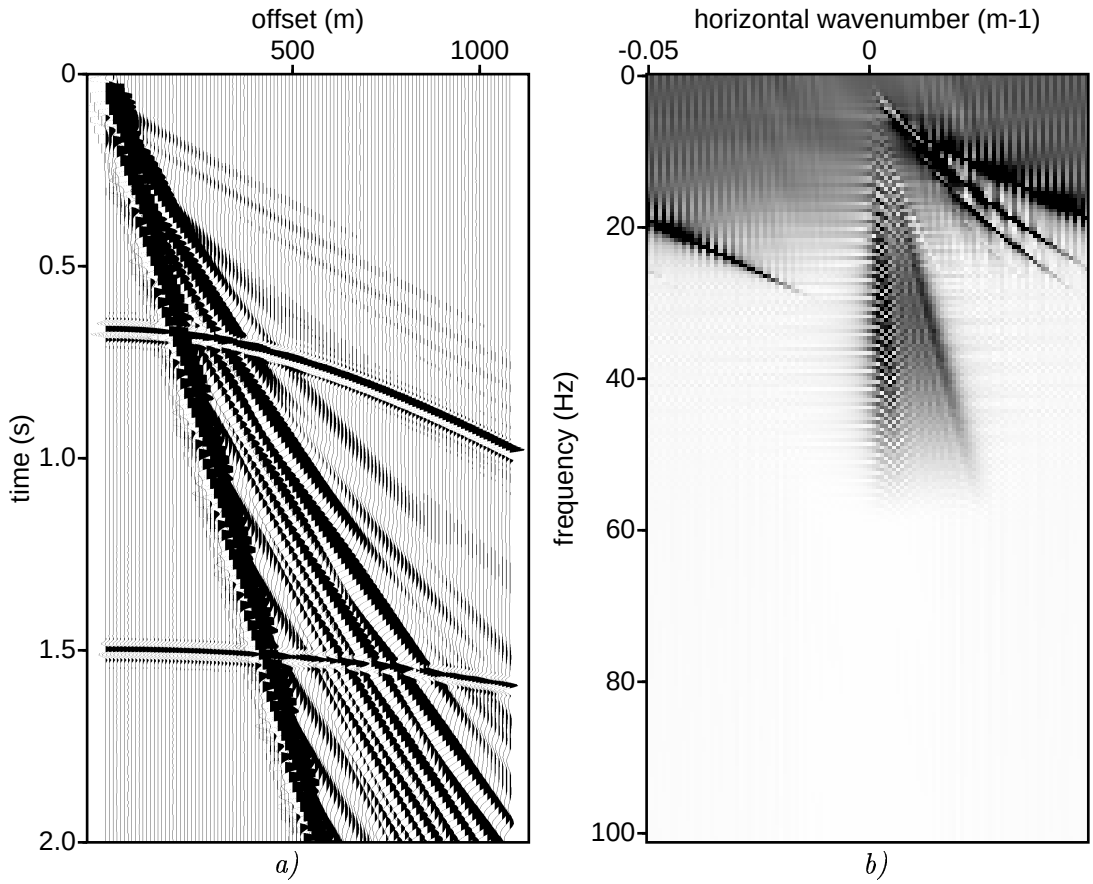


Figure 4.23: Simulated shot record with dispersive ground roll. a) Shot record with ground roll and two reflection in (t, x) domain. b) Shot record of a) in (f, k_x) domain.

roll. For the reflected events, we have the response in the (t, x) domain:

$$p_{Refl}(t, x) = p_{Refl}(t - \frac{x \sin(\theta)}{c_{Refl}}) \tag{4.60}$$

where we made use of the fact that the apparent speed along the surface for the reflected events is $c_{Refl}/\sin(\theta)$, where θ denotes the angle the normal on the wave front makes with the vertical. Since θ is small, the apparent velocity is large. In principle the reflection events are located within a pie-slice in the (f, k_x) domain for apparent velocities between c_{Refl} and ∞. We can now make a sketch of these two arrivals in the (f, k_x) domain, as is given in figure 4.22. A nice feature of this diagram is that the ground roll and reflections are now well separated.

A shot record with dispersive ground roll, simulated in an earth model with a few thin layers on top, together with two reflections is displayed in figure 4.23. Note the separation of the ground roll and reflections in the (f, k_x) domain.

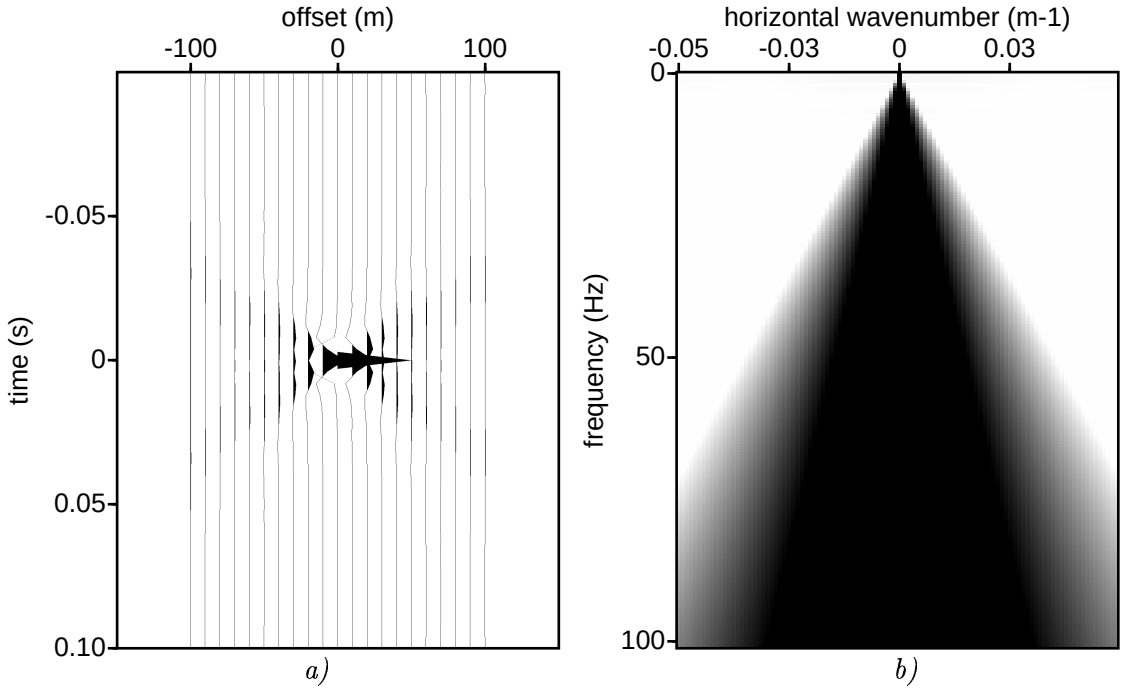


Figure 4.24: "Pie-slice" filters in the (t, x) and (f, k_x) domain.

The use of the two-dimensional Fourier transform enables us to carry out filtering in the (f, k_x) domain. In the case of one-dimensional filtering we are used to perform multiplication of the complex spectrum of an input signal with the complex spectrum of a filter function, being equivalent to convolution in the time domain:

$$\mathcal{F}_t[h(t) * f(t)] = H(f)F(f) \tag{4.61}$$

in which $h(t)$ is the input time response, and $f(t)$ the filter response. Similarly, in the (f, k_x) domain, filtering may be carried out through multiplication with some window function or a more general function of k_x or f only.
However, the two-dimensionality of the (f, k_x) domain provides the extra freedom to design filters of which the transfer characteristics are functions of both k_x and f. In particular, filters with boundaries being linear functions of k_x and f, i.e. filters with sector-shaped pass-band or rejection-band boundaries (see figure 4.24). These filters, called "pie-slice" filters of which the (linear) sector boundaries are specified by apparent velocities, may be designed in such a way that, say, reflection energy is passed in a sector covering the f–axis (apparent velocity is ∞) and a range of relatively high apparent velocities on both sides of the f–axis, while suppressing energy from the regions further away from the f–axis, e.g. in the region between the $+k_x$–axis and a line through the origin with velocity parameter $c_{GR} = f/k_{x,GR}$ somewhat larger than the highest apparent velocity of the low-velocity ground roll events.

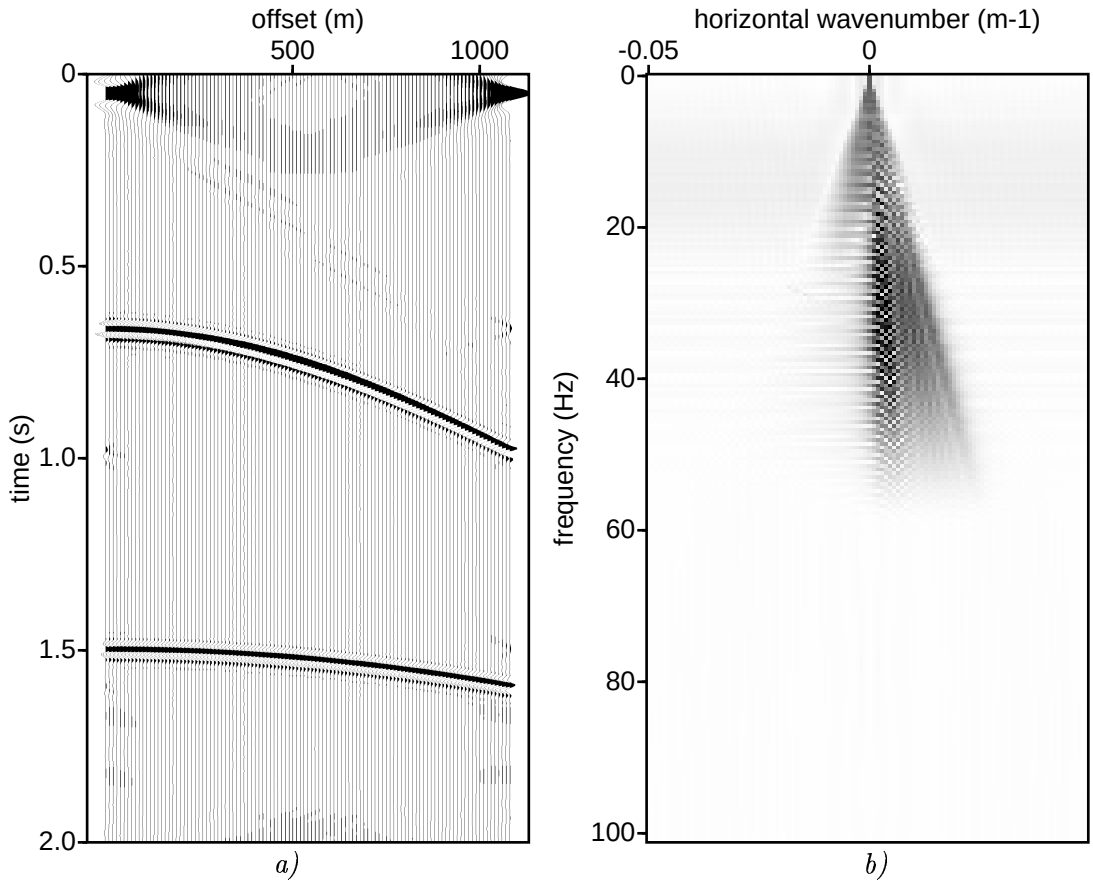


Figure 4.25: Simulated shot record with dispersive ground roll after (f, k_x) domain pie-slice filtering. a) Filtered shot record in (t, x) domain. b) Filtered shot record of a) in (f, k_x) domain.

This is shown in figure 4.25 in which the example of figure 4.23 has been filtered such that the ground roll is removed (except for some edge effects).
A field example of a section with ground roll and without ground roll, removed by (f, k_x) filtering, is shown in figure 4.26. Note that the reflection events (e.g. at 0.8 and 1.5 seconds) become better visible after the filtering. However, in the lower part of the data, some of the ground-roll appears to be aliased, and the (f, k_x) filtering procedure will smear this aliased ground roll. It will be difficult to make a distinction between smeared ground roll and true reflection events in that region.

So far, we established a separation of the arrivals in the (f, k_x) domain by means of different apparent velocities, but the process is even more accentuated by the fact that ground roll usually contains much lower frequencies than the reflection events; it can sometimes even happen that this separation can be achieved only on frequency considerations, but even in that case (f, k_x) "pie-slice" filtering has the preference.

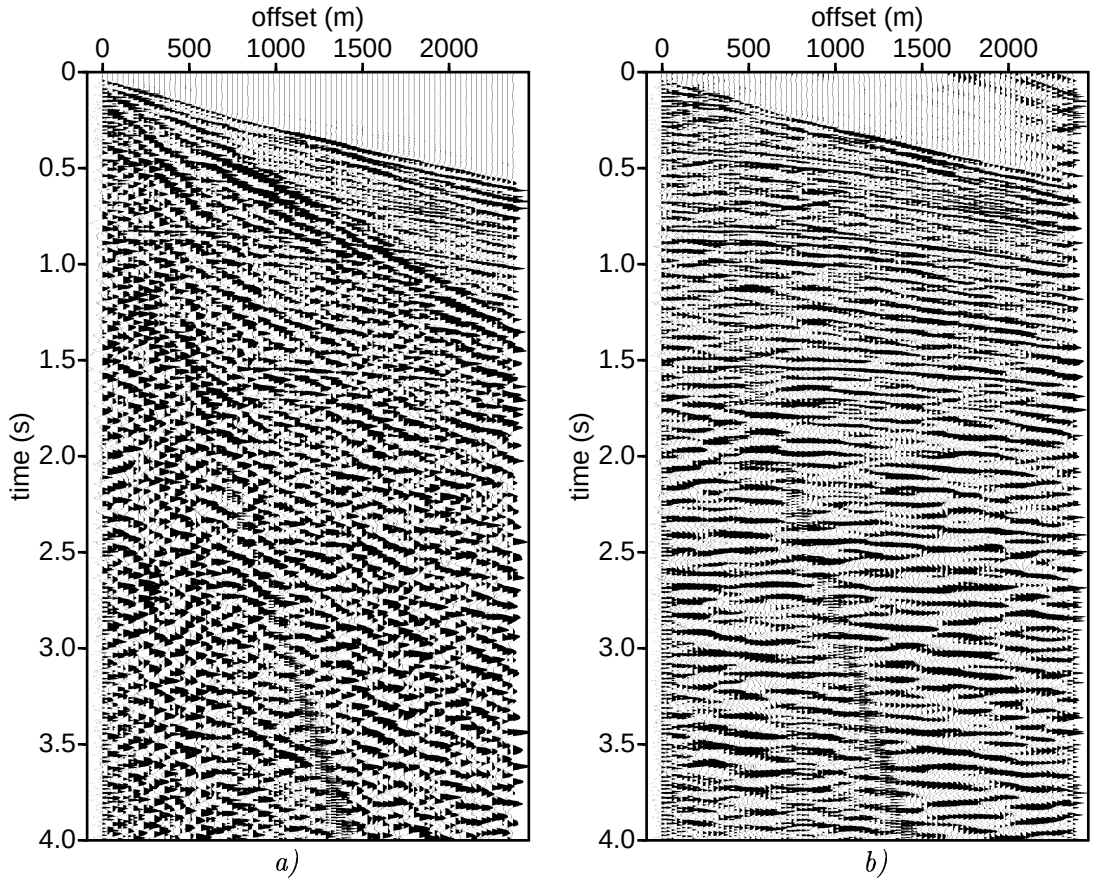


Figure 4.26: Field example of a shot gather (after statics correction) with ground roll and after removal of ground roll by dip-filtering.

Multiple removal using (f, k_x) domain filtering

Another application of (f, k_x) domain filtering is removing multiples on marine data. We already discussed multiple-removal in the context of predictive deconvolution, but there we quite depend on the statistical nature of the process and on the plane layering of the reflectors causing the multiples. Also, for long-period multiples, this filtering procedure will not yield good results. With other words, the assumption that multiples appear as a strict periodic sequence does not always apply in pratice, especially for non-zero offset data. For these situations, (f, k_x) filtering can help. Here again, we make use of the fact that there is a different apparent velocity of the multiple, compared to the reflections coming from deeper down. The reflections from deeper down have generally encountered a higher wave speed, and so these arrivals will arrive more vertically at the surface (see also figure 4.27), compared to multiples that arrive at the same traveltime (and therefore have propagated longer in shallower layers). However, due to structure in the subsurface, this observation may not be true anymore. Therefore, for multiple elimination, the data is sorted into CMP gathers (see also chapter 3) in order to make the hyperbolic assumption of the reflection events better valid.

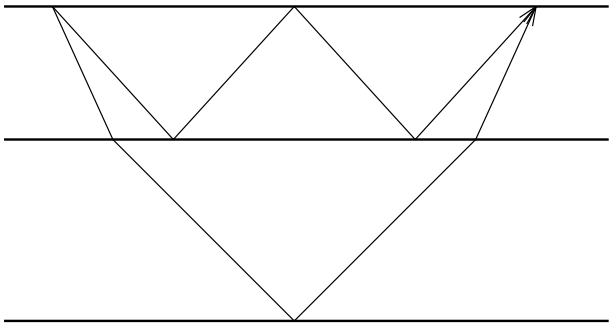


Figure 4.27: The ray path of a multiple and a reflection from deeper down.

The way the multiple elimination filtering is applied is as follows:

- Sort data into CMP gathers.
- Apply NMO correction with velocity in between primary and multiple events, which will map the upward curved primaries to the negative k_x values and the downward curving multiples to the positive k_x plane.
- Apply (f, k_x) filtering by removing one half of the k_x plane, containing the multiple events.
- Resort data into shot gathers or continue processing in CMP gathers.

A field example for a CMP gather from a marine line in the North Sea is given in figure 4.28. Here we see an enormous amount of multiple reflections, which can be identified after NMO correction, with velocities in between primary and multiple velocities, as shown in figure 4.28b. The upward curving events are identified as primaries, whereas the downward curving events are the multiples. In the (f, k_x) domain, the positive k_x plane is zeroed, and after inverse NMO correction, figure 4.28d is the result. Note that due to the mute in the NMO correction, some reflection information is lost. Note also that in the lower part of the region, we are not sure that all the remaining events are primaries: generally the difference in move-out velocity is not present for all multiples, especially in the deeper part of the data.

Summary of F-K filtering

When disturbing events need be removed from the seismic data, and there is a clear difference in slope between desired and undesired events, a filtering procedure in the (f, k_x) domain can be applied. In this domain, certain slopes can be rejected from the seismic data, resulting in removing all events related to these slopes. Applications can be found in removal of groundroll and other low velocity events (especially in land data) and the removal of multiples (especially in marine data).

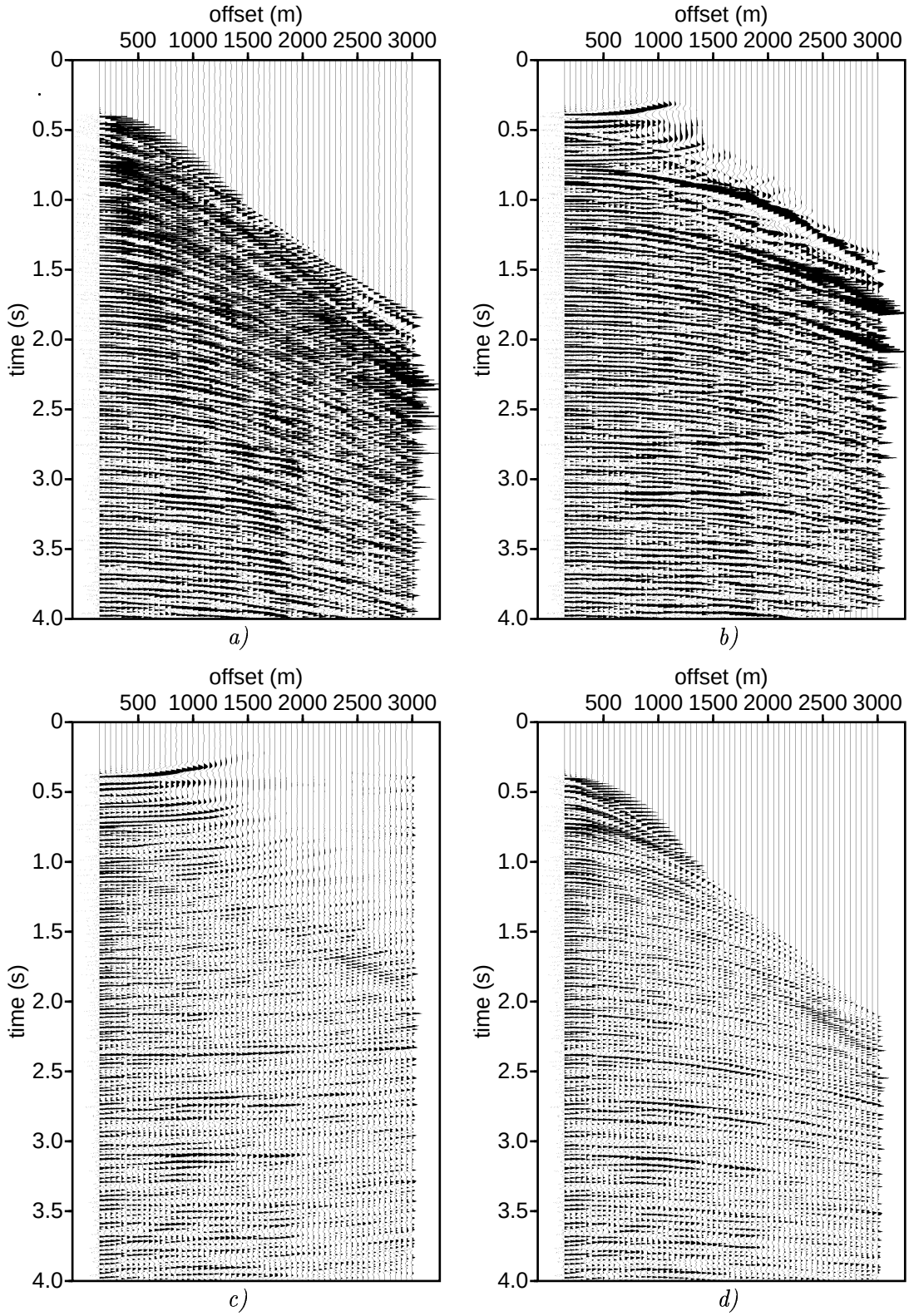


Figure 4.28: Field example of multiple elimination on a marine CMP gather, using NMO correction. a) Input CMP gather. b) NMO corrected CMP gather with NMO velocities in between primary and multiple velocities. c) Result of b) after removing positive dipping events. d) Filtered CMP gather after inverse NMO correction.

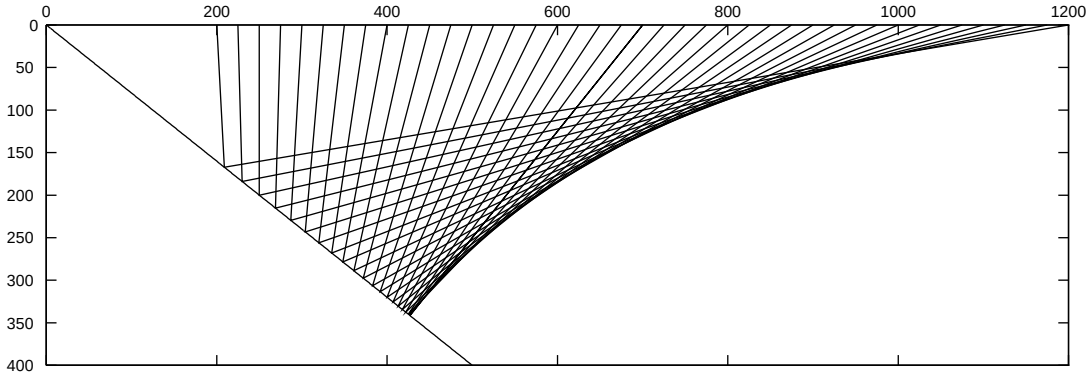


Figure 4.29: The effect of a dipping reflector on the ray pattern.

4.8 Dip Move-Out (DMO) / Pre-stack Partial Migration

The (t, x)–curve for a dipping reflector

When we applied the NMO correction to the CMP gather, we assumed we were dealing with horizontal layers, giving rise to quasi-hyperbolic events. When we have dipping reflectors, the NMO correction still corrects for the hyperbolic move out, but the velocities we use are not the true velocities any more, since they include the dip of the reflector. In order to obtain the true velocity, an extra term needs to be added, and the extra correction for the dip is called Dip Move-out, abbreviated to DMO.

Let us consider figure (4.29). We see that when we take a line perpendicular to the reflector at subsurface reflection point for a finite offset in the subsurface and take the intersection of this line with the surface ($z = 0$), that this point does not lie at the midpoint between source and receiver. This would not be so troublesome if the subsurface reflection point would be the same for the neighboring source-receiver pair in the CMP. But, as can be seen in the figure, the reflection points are smeared out over the reflector.

We will now derive the extra term due to the reflection-point smear. To this purpose, consider figure (4.30). We have a source S with a receiver R; the distance between these two is called $2x_h$, where the subscript h stands for half-offset. The depth of the reflector, measured perpendicular to the interface at the receiver location, is called d_R; the depth of the reflector at point H half between S and R is called d_H; the angle of the reflector with the horizontal is called α. When we take the image of the receiver, we can apply the "cosine-rule" to determine the distance r of the ray path from source to receiver, i.e.:

$$\begin{aligned} r^2 &= (2x_h)^2 + (2d_R)^2 - 2(2x_h)(2d_R)\cos(\frac{\pi}{2} + \alpha) \\ &= 4x_h^2 + 4d_R^2 + 8x_h d_R \sin(\alpha) \end{aligned} \tag{4.62}$$

so we see an extra term arising in the distance, and thus also in the traveltime. But before writing down the traveltime, we should consider that we want to get the same common

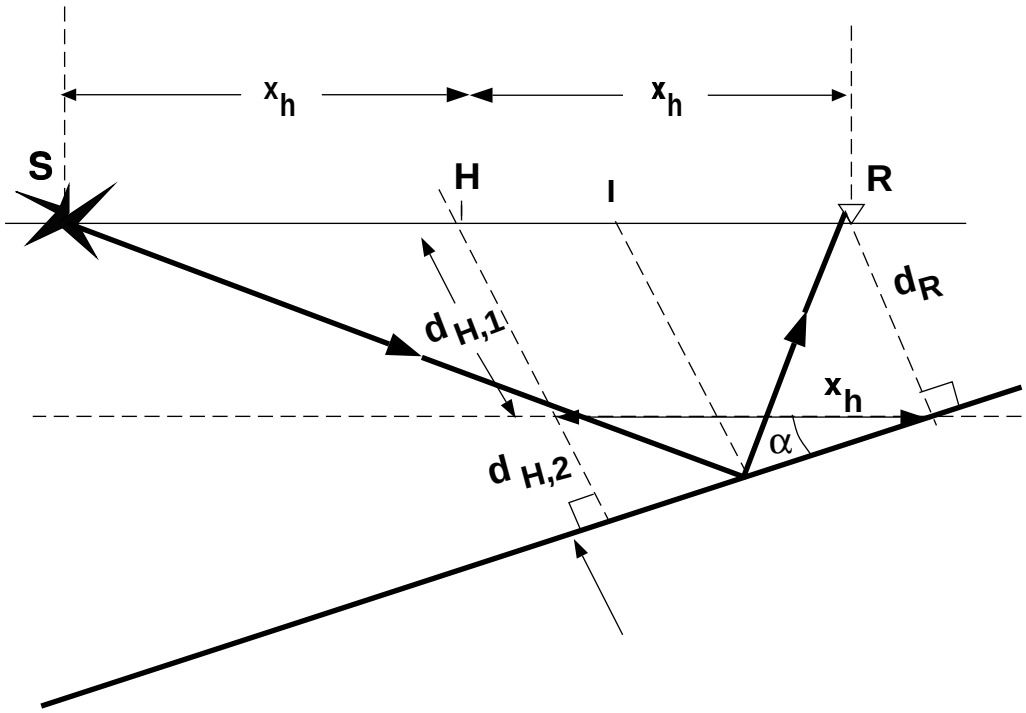


Figure 4.30: A model to derive the DMO term.

reflection point for a CMP gather, so we do not want d_R in the equation but d_H. To this effect, consider the extra lines drawn in figure (4.30) to determine the relation between d_H and d_R. Hence,

$$\begin{aligned} d_H &= d_{H,1} + d_{H,2} \\ &= d_R + x_h \sin(\alpha) \end{aligned} \tag{4.63}$$

Now substituting d_H for d_R, we obtain:

$$r^2 = 4x_h^2 + 4(d_H - x_h \sin\alpha)^2 + 8x_h \sin(\alpha)(d_H - x_h \sin\alpha) \tag{4.64}$$

$$= 4d_H^2 + 4x_h^2 \cos^2\alpha \tag{4.65}$$

This is the distance travelled by the ray, so the traveltime becomes:

$$t = \frac{r}{c} = \sqrt{t_H^2 + \frac{4x_h^2 \cos^2(\alpha)}{c^2}} \tag{4.66}$$

in which t_H is given by:

$$t_H = \frac{2d_H}{c} \tag{4.67}$$

We see that we have a dip-dependent velocity c_{dip} which is related to the true velocity c by $c_{dip} = c/\cos(\alpha)$.

Here we see that the NMO correction with a velocity of $c = c_{dip} = c/\cos(\alpha)$ will do a good job, only the velocity used for the correction is not the true: it includes the dip.

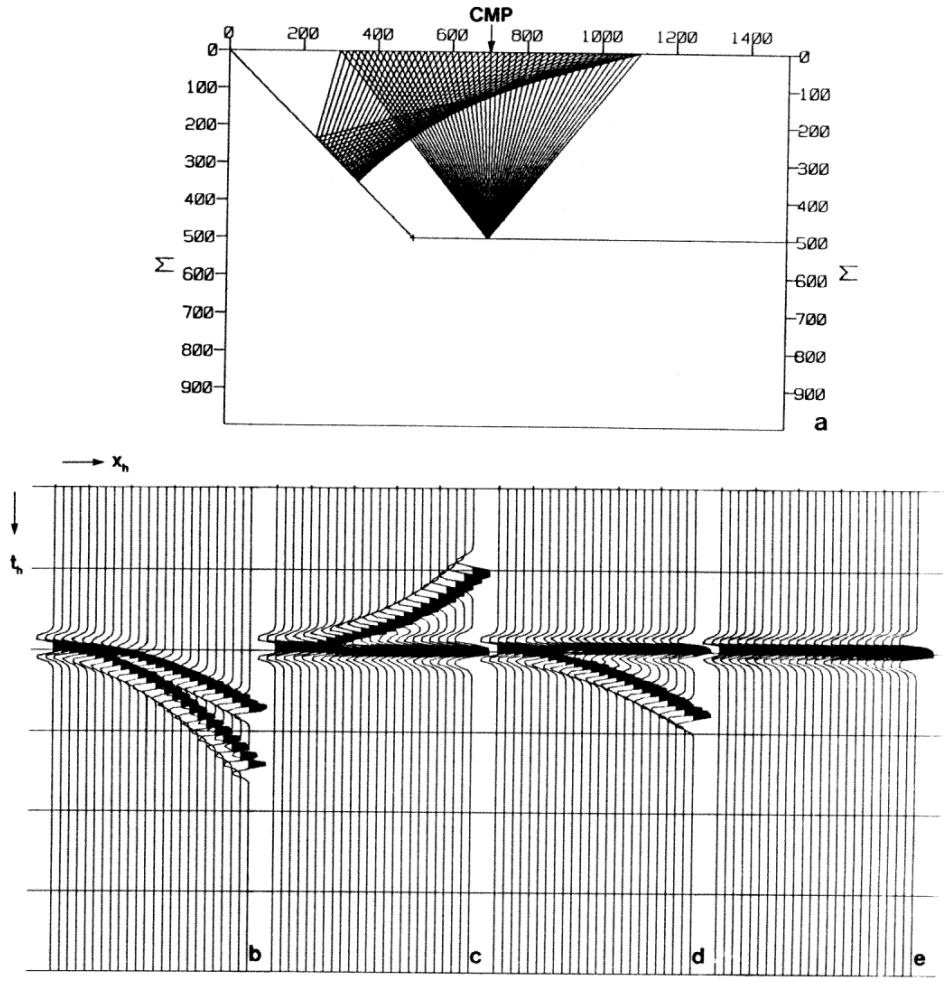


Figure 4.31: A model of two reflectors, showing the problem of conflicting dips.

However, when we have two dips which arrive at the same time, there is a problem: which velocity do we take ? This is the problem of conflicting dips and is illustrated in figure (4.31). We have two reflectors, one horizontal and one dipping. The CMP gather looks as given in figure (b). In figures (c) and (d) we have corrected with the right NMO velocity for the horizontal and dipping reflector, respectively. As can be seen, if we have the right velocity for the one, it is wrong for the other. After correcting for DMO, we take the effect of dip into account and correct for the dip in the right manner, as shown in figure (e).

The DMO correction

In this subsection we will derive the term needed for correcting for the extra time effect. To that purpose, we apply the correction for velocity and dip in two steps, via

writing $\cos^2(\alpha) = 1 - \sin^2(\alpha)$ and splitting the above equation (4.66) as:

$$t = \sqrt{t_{DMO}^2 + \frac{4x_h^2}{c^2}} \tag{4.68}$$

in which t_{DMO} is defined as:

$$t_{DMO} = \sqrt{t_H^2 - \frac{4x_h^2 \sin^2(\alpha)}{c^2}} \tag{4.69}$$

The first equation can be seen as the NMO correction. The other equation has been termed dip move-out or DMO. It can be seen that the time t_{DMO} is equal to the time t_H when the offset between source and receiever is zero; when there is some offset, the time t_{DMO} will be *smaller* than t_H (and remember that the NMO correction has already taken place).

Let us now return to the figure given in the beginning, figure (4.29). In that figure, we saw the reflection-point smear along the subsurface reflector. We need to put each source-receiver pair in the right CMP gather. So, this means that for each offset the data has to be shifted to another CMP position. For a single dipping reflector with a constant velocity layer in between, we can derive what the reflection-point smear is. This is derived in appendix G. The derivation involves some quite elaborate algebra, so the full derivation is left out here. From the reflection-point smear together and some extra relations, we can derive the equation which describes the time effect due to dip, as is also shown in appendix G. The result is:

$$t_I = t_{DMO} \left(1 - \frac{(x_I - x_h)^2}{x_h^2} \right)^{1/2} \tag{4.70}$$

where t_I is the true time from the reflection point upwards, perpendicular to the reflector. This equation is the equation of an ellipse. An example of such an ellipse is given in figure (4.32).

What is very striking and very nice of the expression above, is that the operation does not depend on the velocity of subsurface! For this configuration, DMO is a velocity-independent process, and can thus be robustly included in any processing scheme. However, when the model consists of more than one layer, the DMO correction in general still depends on the velocities of the layers but is not so sensitive to it. When we would add the traces only with an NMO correction, then the resulting image will be of a lower quality. So, we must somehow correct for this, and that is DMO. *The function of DMO is to migrate to a true zero-offset section.*

The DMO operator has to be applied to *common-offset* sections. So therefore the usual procedure in applying NMO and DMO is:

- Apply NMO
- Sort the data to common-offset gathers
- Apply DMO

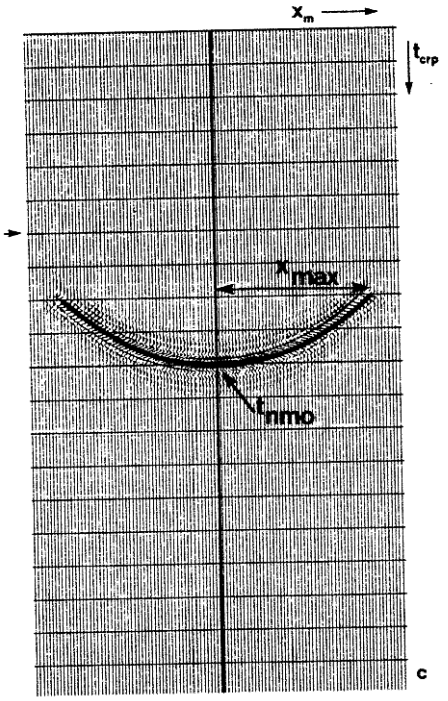


Figure 4.32: The DMO operator : an ellipse in the common offset domain. (Note that t_{NMO} in figure $= t_{DMO}$ in text; x_m in figure $= x_I$ in text)

- Sort the data back to CMP gathers
- Inverse NMO with the velocities from the first NMO
- Apply NMO with the true velocities

An example is given in figure (4.33). These days, this is a standard procedure in data processing. Following DMO, it is often possible to re-pick velocities on the DMO corrected data and re-apply NMO thus giving both a better quality stack and better velocities for input to post-stack migration, interpretation and other velocity-dependent techniques.

In the discussion so far we discussed DMO as a process via the space-time domain as operators. Also, we only discussed it via kinematic effects (ray theory), not yet putting any wave theoretical aspects in it. DMO can be given a wave-theoretical basis by way of the Kirchhoff integral. There are still problems with amplitude and phase distortions in this approach, but despite these problems, integral DMO methods are extremely popular today due to their speed and their adaptability to irregular surface sampling as is common in 3-D seismics. As with many migration algorithms, DMO can be applied via different routes, such as via the (f, k_x) domain. This latter has been done by [Hale, 1984].

When should we use which method? We follow the recommendations given by [Deregowski, 1986]. If you have regular spaced 2D data and amplitudes are an important feature then it is best to use an (f, k_x)-domain method such as the log-stretch technique. This is especially true if the data contains a range of different dips. If amplitudes are not of prime concern,

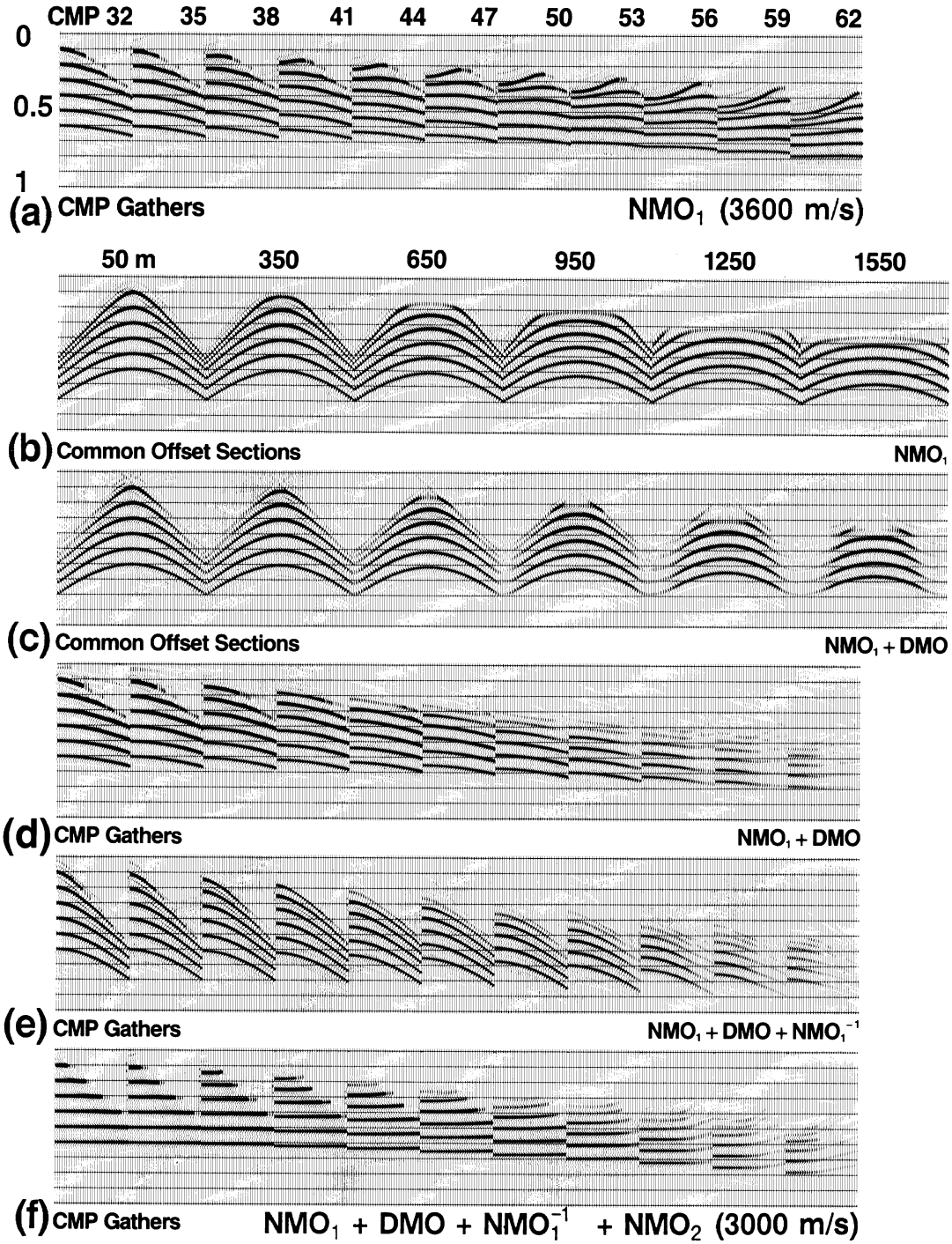


Figure 4.33: Figure from [Yilmaz, 1987], figure 4-125.

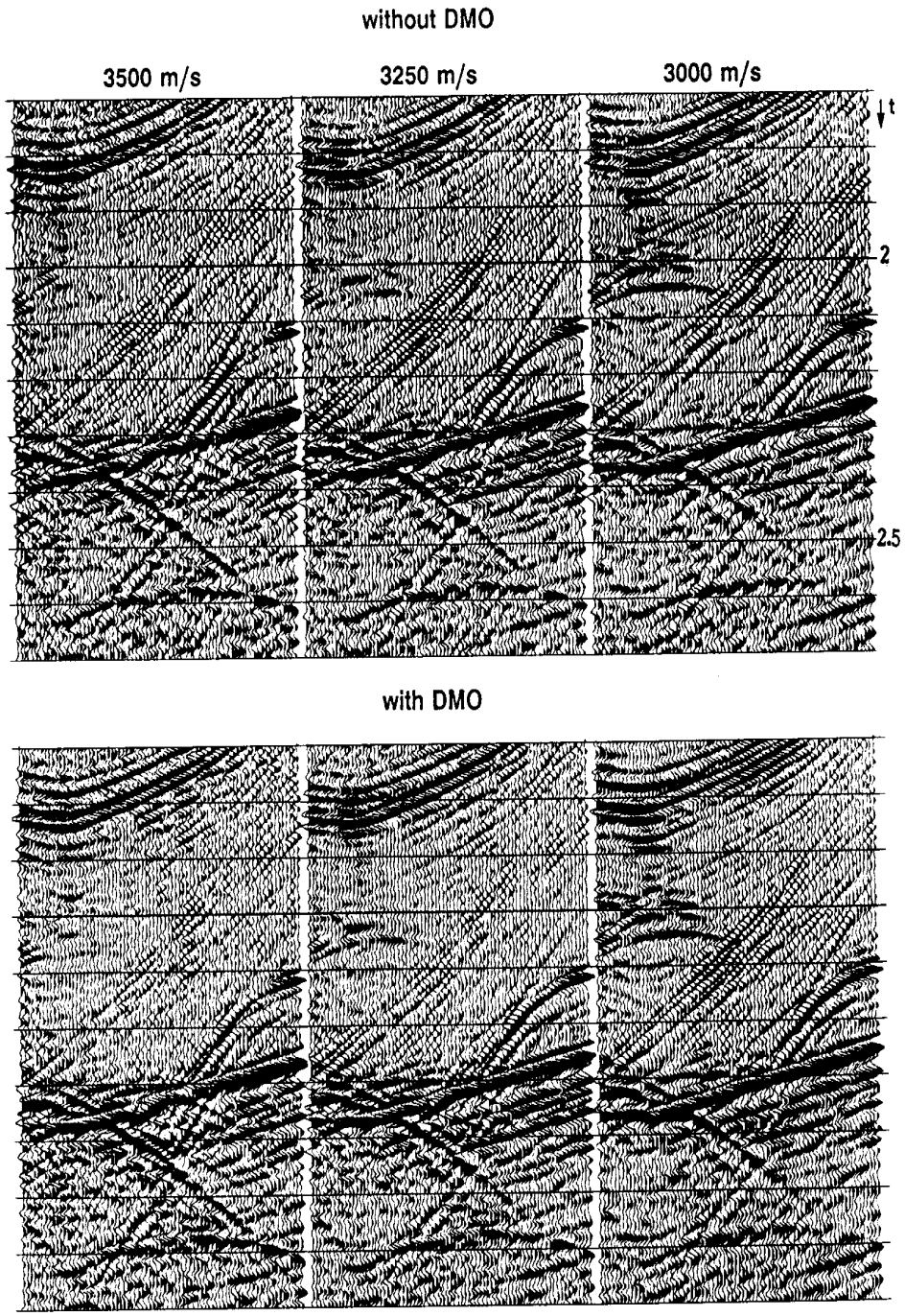


Figure 4.34: The effect of conflicting dips (Schoot,1989).

then (t, x)-domain methods are adequate provided some care has been taken to pass the steeper dips and that the operator is anti-aliased.

If a complete proper image of the subsurface is the goal, pre-stack migration is needed. However, this technique suffers from a large computational cost and the requirement of a

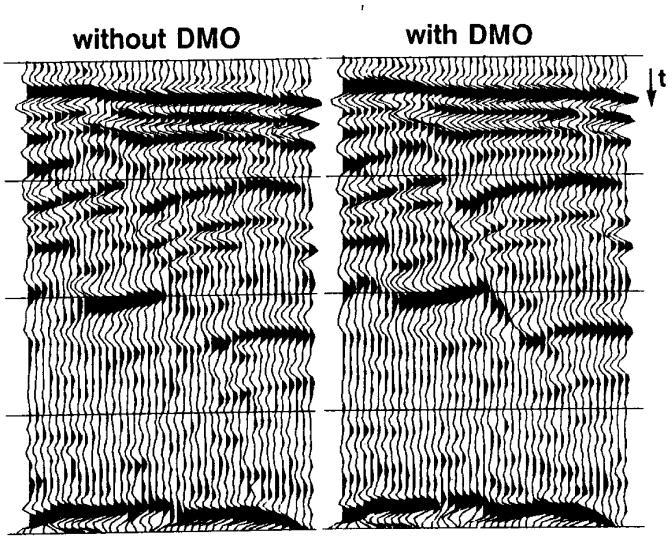


Figure 4.35: The better-resolved dipping fault using DMO (Schoot,1989).

detailed velocity model, but for sufficiently complicated structures it is the only method to correctly image the data. DMO is a method which bridges these extremes. DMO increases the costs via computational effort, but is still fast enough to be part of a standard processing sequence.

Finally, we would like to give a field example as shown in figures (4.34) and (4.35), taken from [van der Schoot, 1989].

Summary of DMO

In this section, the correction due to the dip of a reflector is given. The most important feature is that for a single layer the correction does not depend on the velocity model. When more layers are present, DMO slightly depends on the velocity model. Mostly, DMO is applied in a separate processing step, using common-offset gathers.

4.9 Zero offset (poststack) migration algorithms

In our basic processing sequence in chapter 3, we discussed Kirchhoff migration because it corresponded with our intuitive notion of collapsing diffraction hyperbolae to their apexes, as was common in the early days with the diffraction stack. In Kirchhoff migration we not only get the timing right, but we also take account of amplitude variations along the diffraction hyperbolae. In practice, Kirchhoff migration is hardly used in post-stack migration; instead finite-difference methods, $f - k_x$ and $f - x$ algorithms are employed. These techniques will be the topic in this section. Again, we will derive these techniques but will not go into all its fine mathematical details. We hope to give the most important properties of these techniques.

With $f - k_x$, $f - x$ and finite-difference techniques we will focus on both time and depth migration. The basic difference between time migration and depth migration is that in time migration the migration result is expressed in "vertical time", i.e. the time domain variant of depth. After application of time migration, a time-to-depth conversion is required in order to obtain the final depth image (see section 4.10). The application of time migration versus depth migration implies that there is less dependence on the velocity variations expected in the subsurface; depth migration is used in regions where large lateral velocity variations exist. We will point out some differences between time and depth migration. For each method we will discuss what kind of parameters are needed, and what the influence is of these parameters. Finally, we give a summary of when to use which technique.

In general migration consists of two steps:

1. Inverse extrapolation of the input data (i.e. the stacked section representing a zero offset section) from surface level to a certain depth (or vertical time) level.
2. Select the t=0 component, which is the migrated result for that depth (or vertical time) level.

Step 1 involves the wave equation. Therefore, we will derive the Kirchhoff integral, and show how this is used for migration.

The Kirchhoff integral

These days the theory has developed such that we can describe the migration process much better via wave theory. We will discuss some concepts in wave theory and discuss migration in these terms. The migration methods based on wave theory are referred to as *wave-equation migration*. Mathematically, there are two ways to solve the wave equation, one is via integral methods, and the other via differential methods. We choose here to find a solution via integral methods, because of two reasons. First, it adheres to the intuitive idea of the diffraction stack (which in fact is also an integration along hyperbolic paths). Secondly, the integral method gives the general exact solution, while the differential methods involve some approximations. The classical paper explaining the integral method as applied to seismic migration, i.e. Kirchhoff migration, is the one by [Schneider, 1978]. In appendix H the derivation from the Kirchhof integral can be found.

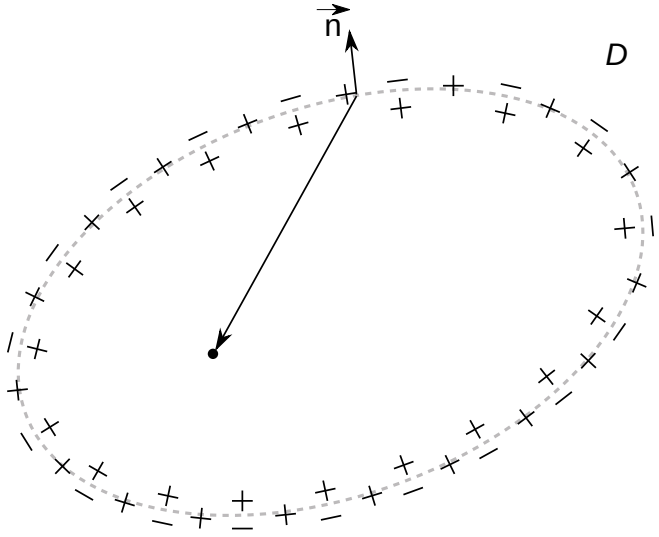


Figure 4.36: A pressure field can be synthesized from the wave fields of a monopole and dipole distribution on a closed surface, using respectively the particle velocity and pressure of the actual wave field at this boundary as their source strengths (after [Berkhout, 1984], figure 5.1)

The Kirchhoff integral looks as follows:

$$\chi(\mathbf{x})p(\mathbf{x},t) = \frac{1}{4\pi}\int_{-\infty}^{\infty}\int_{\partial D}(-\rho G\partial_t\mathbf{v} - p\nabla G)\cdot\mathbf{n}dA^s dt^s, \tag{4.71}$$

where $p(\mathbf{x},t)$ and $\mathbf{v}(\mathbf{x},t)$ describe the pressure field and the particle velocity vector field of a wave field. $G(\mathbf{x},t)$ descirbes a so-called *Green's function*, which is the solution of a point source in the same medium where the actual wave field is present.

This equation expresses that if we know the pressure p and the time derivative of the normal component of the particle velocity on a closed surface, the pressure can be computed in every point inside D. Also, we recognize that the pressure at a certain position is synthesized by means of a monopole (i.e. G) and dipole (i.e. $\nabla G\cdot\mathbf{n}$) distribution on a closed surface ∂D. The propagation of the secondary sources at the boundary ∂D to the observation point $(\mathbf{x},t)$ is described by the Green's function G. The same kind of expression can be derived for the particle velocity, see [Berkhout, 1984], chapter 5, from which figure 4.36 has been drawn.

Using the Kirchhoff integral for migration

The configuration of figure 4.36 is not suited to do zero offset migration directly. The seismic measurements are done on the surface of the earth, and only one type of wave field (either pressure or vertical velocity component) is measured. Therefore, the Kirchhoff integral can be rewritten for the special case of a flat surface of the earth into a more convenient shape.

The derivation of this migration integral can also be found in appendix H. Besides choosing a special situation of the boundary δD, being a flat surface combined with a

semi-hemisphere that extends to infinity, also a special type of Green's function is used. As we only have to define that the Green's function is the solution to the wave equation, any combination of two Green's function to a new one is also valid. In the derivation of the migration formula, the Green's functions are chosen to be two monopoles that are chosen just above and below the surface, with opposite signs.

With these two choices, one term in the Kirchhof integral will disappear and finally the following equation remains:

$$p(\mathbf{x},t) = \frac{-1}{2\pi}\partial_z \int_{z^s=0} (\frac{p(\mathbf{x}^s, t+r/c)}{r}) dA^s. \tag{4.72}$$

This is Kirchhoff's migration formula, given by [Schneider, 1978]. In this equation A^s represents the surface and R the distance between the subsurface location point $\mathbf{x}$ and a point at the surface, and c the propagation velocity. We would like to stress that this result does not involve any approximations, the result is only dependent on the knowledge of the velocity distribution (i.e. vertical derivative of the pressure field) at the surface.
In fact it states that the wave field in any point in the subsurface $p(\mathbf{x},t)$ can be calculated from the wave field $p(\mathbf{x}^s,t)$ recorded at a plane reference level z_s, assumed that we have a recording from $-\infty$ until $+\infty$ at the surface.
Note that if the term $t+r/c$ is replaced by $t-r/c$ the *inverse* propagation from the surface to point $(\mathbf{x},t)$ becomes a *forward* extrapolation.

A general zero offset migration procedure

With the Kirchhoff's migration formula the wavefield at a certain depth level can be constructed from the wave field measured at the surface. This can be used in a zero offset migration procedure as follows:

- Consider a stacked section (eventually after DMO) as a zero offset section. This will be correct in traveltime but not in amplitudes.
- Consider a zero offset section to be an exploding reflector measurement, in which all reflectors in the subsurface are considered to be sources that explode at t=0 and travel to the surface with half the medium velocity (see also Chapter 3 of this lecture notes). This again is good for explaining the traveltimes in a zero offset section, but not perfect for amplitudes.
- Inverse extrapolate the wave field measured at the surface (i.e. the zero offset section) to a depth level in the earth (by means of the migration formula equation (4.72)).
- select the $t = 0$ component of this extrapolated wave field, which will contain the exploding reflector contributions of the depth level under consideration. Save this $t = 0$ component in the migrated output section.
- Repeat the last two steps for all depth levels, such that the complete migration result is constructed for all depth levels.

The procedure is illustrated in figure 4.37 for a wedge-shaped reflector model. The zero offset response (figure 4.37a) contains the dipping contributions from the flanks and a

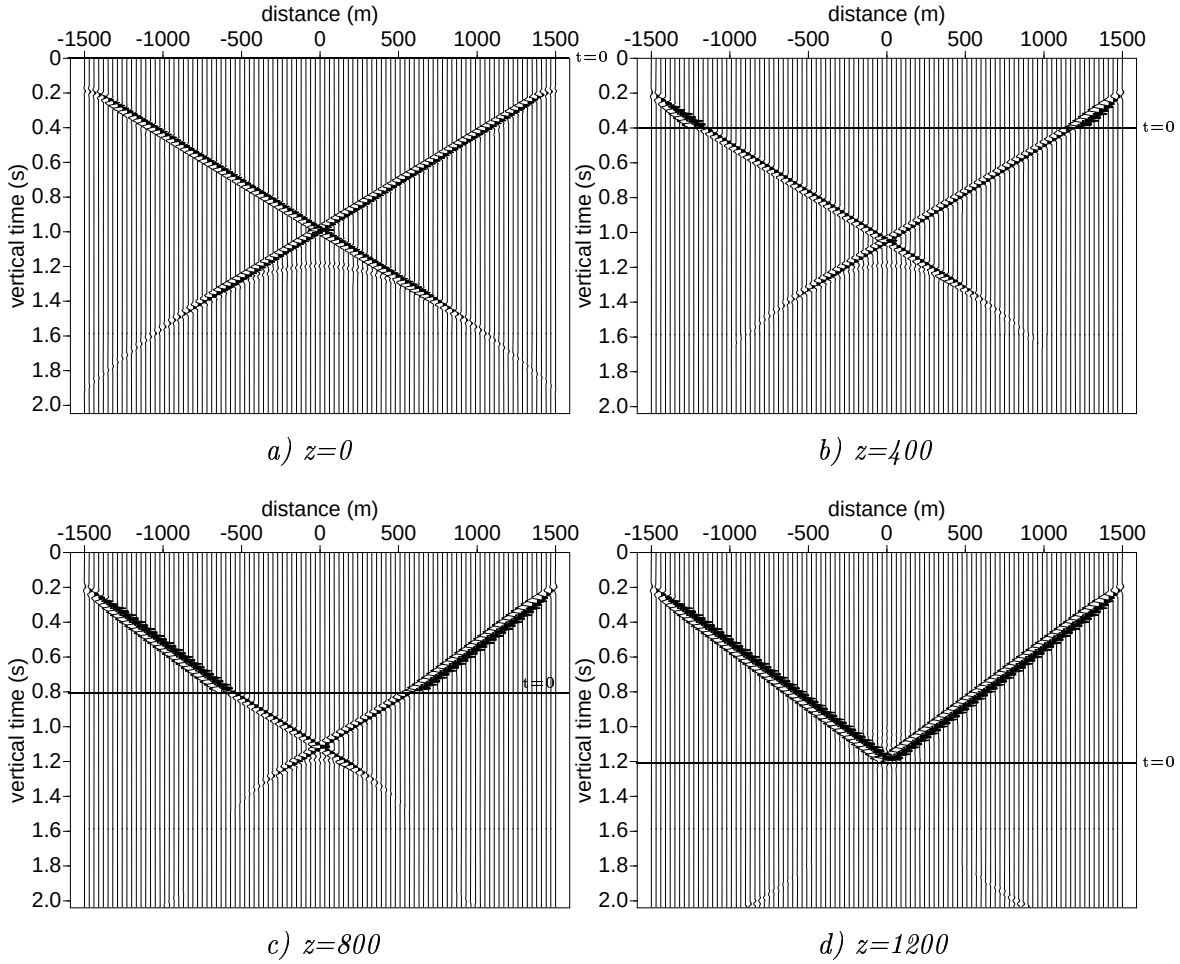


Figure 4.37: Zero offset migration in action. Above the indicated t=0 line the migrated image is visible, below the t=0 line the inverse extrapolated result at the current depth level. At t=0 the actual imaging takes place for that depth. a) Zero offset data of wedge shaped reflector. b) Wave field after inverse propagation to 400 m depth. c) Result at 800 m depth. d) Result at 1200 m depth.

diffraction from the tip of the wedge. For three different depth levels a combined picture of the migrated image (above the t=0 line) and the inverse extrapolated data (below the t=0 line) is shown. Note that at the indicated $t = 0$ line the imaging of the current depth level takes place. Note also that the diffraction of the tip of the wedge is collapsing towards a point at 1200 m depth and that the dip of the sides of the wedge are correctly positioned after imaging.

Overview of the discussed migration procedures

In the following a number of migration techniques are discussed. Depending on the required accuracy and of the constraints given by the subsurface model, different techniques will be selected in practice.

In chapter 3 already the so-called **Kirchhoff migration** procedure is discussed: by modeling the operators that describe propagation from each point in the subsurface to the surface level, the wave field (zero offset section) can be extrapolated to each subsurface point. By selecting the $t = 0$ component of each extrapolated wave field the subsurface image is built up. The quality of the modeled operators will define the quality of the image. In its simplest form, the local stacking velocities (i.e. hyperbolic approximations) are used to describe these propagation operators and the result is often called the diffraction stack. Such an approximation is only valid if the medium velocities vary very smoothly in both lateral and vertical direction. For more complex models often raytracing methods are used to calculate the traveltimes from the surface to each image point.

The Kirchhoff migration procedure is a *non-recursive* method: for the result at each image point, the extrapolation is done via one operator from the original wave field (i.e. zero offset section). Another approach is *recursive* migration: the wave field is extrapolated in small depth steps the output for one depth level being the input for the next extrapolation step. In the following, some well-known recursive migration procedures are discussed in more detail.

The first discussed method is the **Gazdag phase shift migration**, which is a relatively efficient procedure for a velocity model where the velocity only varies with depth. This means that each extrapolation step (from one to the next depth level) can be achieved by a simple multiplication in the wavenumber-frequency domain. So extrapolation will consist of three steps: (i) Forward Fourier transform of the wave field from $x - y - t$ to $k_x - k_y - \omega$, (ii) multiplication with the so-called phase shift operator, (iii) inverse Fourier transform to $x - y - t$ and selection of the $t = 0$ component.

If the velocity medium is completely homogeneous, an even more efficient method can be used:the **Stolt migration**. By a forward Fourier transform to the frequency-wavenumber domain, an interpolation procedure (axis transformation) and an inverse Fourier transform, the complete image for all vertical times is immediately obtained (no recursion procedure needed). This means that a direct conversion from the zero offset data (as shown in figure 4.37a) to the final image (figure 4.37d) is achieved.

In practice, the subsurface hardly ever can be assumed by a homogenous velocity model, and other techniques need be used that can handle more or less these velocity variations. One of the procedures is the migration using a **finite-difference** operator, where the extrapolation and imaging is done in small steps, each step considering the local velocity. Depending whether this migration procedure is carried out in the vertical time domain or in the true depth domain, less or more accurate results can be achieved.

For laterally more complex media, the so-called **recursive space-frequency depth migration** is the good solution, which also takes small steps in the lateral and depth direction, each time using an exact wave field extrapolation operator belonging to the local velocity.

As can be expected, everything is a matter of efficiency: the cheapest method is the Stolt migration, but it can handle the least model complexities. The more complex the subsurface model, the more expensive the migration method will be that must be selected for an acceptable image.

Time migration via the $f - k_x - k_y$ domain (Gazdag phase shift migration)

We can use the result from Kirchhoff migration as given in equation (4.72). This is written in a compact form. Because of the form of r, we can write this formula symbolically as a three-dimensional convolution, i.e.:

$$p(x, y, \Delta z, t) = p(x, y, 0, t) * \frac{1}{2\pi} \partial_z \frac{\delta(t + \bar{r}/c)}{\bar{r}}, \tag{4.73}$$

in which $\bar{r}$ is given by:

$$\bar{r}^2 = x^2 + y^2 + (\Delta z)^2. \tag{4.74}$$

In this expression we can see that we have a three-dimensional convolution over x, y and t of the data with a term describing the *inverse* propagation through the medium from the surface to depth Δz. It is assumed that we can consider a homogenous layer of thickness Δz, otherwise equation (4.74) would be a more complicated function of the coordinates. For this situation a convolution in t, x and y means that in the $f - k_x - k_y$ domain this is a multiplication of the transformed data with the $f - k_x - k_y$–transform of the last term in equation(4.73). Note that the transformation is only valid for models for which the velocity c is a function of the depth z only.

Let us determine what the $f - k_x - k_y$ transform of this propagation term is. If we call this term W, then we can write:

$$\tilde{W}(k_x, k_y, \Delta z, f) = -\frac{\partial}{\partial z} \frac{1}{2\pi} \int \int \frac{\exp(+2\pi i f \bar{r}/c)}{\bar{r}} \exp(2\pi i k_x x + 2\pi i k_y y) dx dy. \tag{4.75}$$

Now transforming this equation to polar coordinates for x and y, via $x = r\cos(\phi)$ and $y = r\sin(\phi)$, and also for k_x and k_y via $k_x = k_r \cos(\theta)$ and $k_y = k_r \sin(\theta)$ we arrive at:

$$\begin{aligned} dxdy &= r dr d\phi \\ k_x x + k_y y &= k_r r \cos(\theta - \phi) \\ \bar{r} &= \sqrt{r^2 + \Delta z^2}. \end{aligned} \tag{4.76}$$

Substituting this in the above, we get:

$$\begin{aligned} \tilde{W}(k_r, \Delta z, f) &= -\frac{\partial}{\partial z} \int_0^\infty \frac{\exp(+2\pi i f \bar{r}/c)}{\bar{r}} r dr \left[\frac{1}{2\pi} \int_0^{2\pi} \exp(2\pi i k_r r \cos(\theta - \phi)) d\phi \right] \\ &= -\frac{\partial}{\partial z} \int_0^\infty \frac{\exp(+2\pi i f r/c)}{r} r dr J_0(2\pi k_r r), \end{aligned} \tag{4.77}$$

in which we have used the Bessel function J_0. The last integral is a standard Fourier-Bessel transform, so we get:

$$\begin{aligned} \tilde{W}(k_r, \Delta z, f) &= -\frac{\partial}{\partial z} \frac{\exp(+2\pi i k_z \Delta z)}{2\pi i k_z} \\ &= \exp(+2\pi i k_z \Delta z) \end{aligned} \tag{4.78}$$

in which k_z is defined as:

$$k_z = \left(\frac{f^2}{c^2} - k_r^2 \right)^{1/2}. \tag{4.79}$$

So this is a relatively simple function in the $f - k_x - k_y$–domain. Note that for *forward* extrapolation, the phase shift is given as:

$$\tilde{W}(k_r, \Delta z, f) = \exp(-2\pi i k_z \Delta z). \tag{4.80}$$

Figure 4.38 shows a picture of the amplitude of the (complex valued) W function in the k_x, k_y domain. This operator is normally referred to as the *phase shift* operator. For a homogeneous medium a depth step Δz can be taken by a simple multiplication in the wavenumber-frequency domain:

$$\tilde{P}(k_x, k_y, \Delta z, f) = \tilde{W}(k_x, k_y, \Delta z, f)\tilde{P}(k_x, k_y, 0, f), \tag{4.81}$$

in which the sign of the exponent in the phase shift operator will define whether it is a forward ($-$ sign) or inverse ($+$ sign) extrapolation. It appears that in a 2-D medium a similar expression can be found:

$$\tilde{P}(k_x, \Delta z, f) = \tilde{W}(k_x, \Delta z, f)\tilde{P}(k_x, 0, f), \tag{4.82}$$

with the phase shift operator defined as:

$$\tilde{W}(k_x, \Delta z, f) = \exp(\pm 2\pi i k_z \Delta z) \tag{4.83}$$

and the value of k_z defined as:

$$k_z = \left(\frac{f^2}{c^2} - k_x^2 \right)^{1/2}. \tag{4.84}$$

In equation (4.83) the sign of the exponent determines again forward or inverse propagation.

So far, we considered a general extrapolation scheme for any wave field at the surface that we want to inverse propagate into a medium with velocity c. For the case of a zero offset section (i.e. our stacked section), we have shown in chapter 3 that we can consider this as a response of exploding reflectors in the subsurface, considering a medium with half velocity $c/2$. Therefore, from now on we have to use $c/2$ in our expressions. Then equation (4.79) becomes:

$$k_z = \left(\frac{4f^2}{c^2} - k_x^2 - k_y^2 \right)^{1/2}. \tag{4.85}$$

As mentioned in the beginning of this section, we are usually interested in results which is mapped into vertical time and in this expression, we have a step Δz in depth instead of in time. To convert to time steps, we rewrite the inverse propagation operator as:

$$\tilde{W}(k_x, k_y, \Delta z, f) = \exp(+2\pi i \Omega \Delta \tau), \tag{4.86}$$

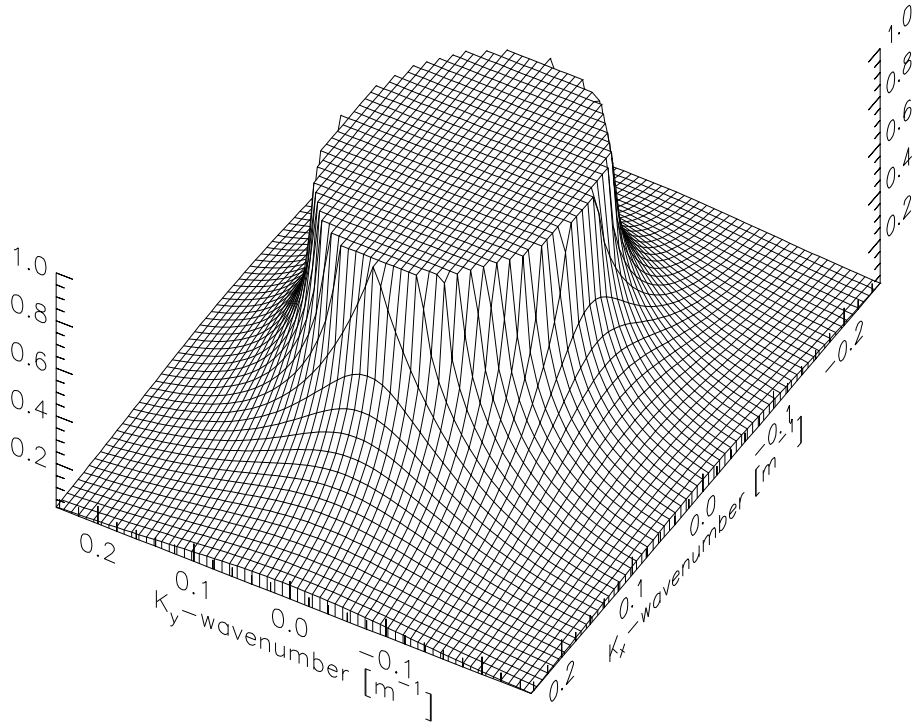


Figure 4.38: Amplitude of the homogeneous phase shift operator in the wavenumber domain.

in which Ω and $\Delta\tau$ are given by:

$$\Omega = \left(f^2 - \frac{k_x^2 c^2}{4} - \frac{k_y^2 c^2}{4} \right)^{1/2} \tag{4.87}$$

and

$$\Delta\tau = \frac{2\Delta z}{c}. \tag{4.88}$$

In this expression, we consider a zero offset section along the x and y coordinate, so the horizontal wavenumbers are k_x and k_y. Note that Ω is related to k_z via $k_z = 2\Omega/c$. $\tilde{W}$ represents a simple phase shift which is a wavefield extrapolation over the time difference $\Delta\tau$. In a sense, we do not get the data in z, but in τ so it would be better to introduce a $\tilde{W}'$, which is related to $\tilde{W}$ via:

$$\tilde{W}'(k_x, k_y, \Delta\tau, f) = \tilde{W}(k_x, k_y, \Delta z, f). \tag{4.89}$$

For migration, we forward $f - k_x - k_y$ transform the data, apply the phase shift as above, inverse $f - k_x - k_y$ transform the result and gather the data at $t = 0$, to obtain the migrated data. This method is called the *phase-shift method*, described by [Gazdag, 1978].

Time migration by Fourier transform (Stolt migration)

Let us consider the special case that the wave speed is constant. In this situation, the phase shift migration procedure can be very efficiently rewritten in a one-step migration procedure, which is commonly known as *Stolt migration*. The wavefield extrapolation process in this homogeneous velocity case can be written as:

$$p(x,y,\Delta\tau,t) = \int\int\int \tilde{P}(k_x,k_y,\tau=0,f)\exp(2\pi i\Omega\Delta\tau)\exp(-2\pi i(k_x x+k_y y)+2\pi i f t)dk_x dk_y df. \tag{4.90}$$

For migration we need the $t=0$ component of the extrapolated wave field, which results into:

$$p(x,y,\Delta\tau,t=0) = \int\int\int \tilde{P}(k_x,k_y,\tau=0,f)\exp(-2\pi i(k_x x+k_y y)+2\pi i\Omega\Delta\tau)dk_x dk_y df. \tag{4.91}$$

If we can change the integration over f into an integration of Ω, it describes a simple inverse Fourier transform. Therefore we are going to rewrite Ω as given in equation (4.87) as:

$$f = \left(\Omega^2 + \frac{k_x^2c^2}{4} + \frac{k_y^2c^2}{4}\right)^{1/2} \tag{4.92}$$

Changing the integration over f into one over Ω using requires the followng relation:

$$df = \frac{\Omega}{(\Omega^2 + k_x^2c^2/4 + k_y^2c^2/4)^{1/2}}d\Omega \tag{4.93}$$

and using this in the migration formula (4.91), we get the result:

$$\begin{aligned} p(x,y,\Delta\tau,t=0) \quad = \quad & \int\int\left[\frac{\Omega}{(\Omega^2 + k_x^2c^2/4 + k_y^2c^2/4)^{1/2}}\right] && (4.94)\\ & \tilde{P}\left(k_x,k_y,\Delta\tau=0,(\Omega^2 + k_x^2c^2/4 + k_y^2c^2/4)^{1/2}\right) \\ & \exp(-2\pi i(k_x x+k_y y)+2\pi i\Omega\Delta\tau)dk_x dk_y d\Omega. && (4.95)\end{aligned}$$

The nice feature of this formulation is that it describes an inverse Fourier transform over the coordinates $k_x - k_y - \Omega$, yielding the complete migration for all $\Delta\tau$ values. The result as given above is the constant-velocity Stolt (time) migration. To resume this procedure, the following steps have to be taken:

- Forward Fourier transform from $x-y-t$ to k_x-k_y-f of the zero offset data
- Interpolate from $\tilde{P}(k_x,k_y,f)$ to $\tilde{P}(k_x,k_y,\Omega)$
- Scale this result with the Jacobian factor of equation (4.93)
- Inverse Fourier transform from $k_x-k_y-\Omega$ to $x-y-t$

The second step describes a mapping from the energy in the $k_x - k_y - f$ domain to another position. Therefore this method is often called k_z *mapping*.

Let us now discuss some practical aspects of the $f - k$ migrations, first the Stolt migration. Stolt migration is not much applied in practice simply because it assumes that the velocity is constant in the earth. Stolt made an extension of his original scheme, and this one should be considered as a separate algorithm. In this extended algorithm he introduced a so-called stretch-factor in which the term "stretch" is used because it stretches the time-axis. This stretch factor is, theoretically, a complicated function of velocity and stretch-coordinate variables; in practice it is often set to a scalar. The stretch factor usually varies between 0 and 2, where for a constant velocity medium, the stretch factor is exactly 1. A too small stretch factor gets undermigrated data, while a too large stretch factor overmigrates the data. For a further discussion the reader is referred to [Yilmaz, 1987], page 298 and page 514.

The other migration algorithm we derived, is the Gazdag phase-shift method which is valid for velocity functions that vary with depth only, so more general than in the Stolt migration. A parameter in the phase-shift method is the 'depth' step size $\Delta\tau$. This step size must be set smaller when the dips are becoming larger. In practice, the 'depth' step size is typically taken between the half and full-dominant period of the wave field (which is dependent on the steepness of the dips in the section). Again, for a good discussion on results with this method, the reader is referred to [Yilmaz, 1987], page 301.

Recursive depth migration in the space-frequency domain

A restriction of the above mentioned migration methods is that they operate in the wavenumber domain, under the assumption that the velocity field is laterally invariant. This can be a serious hurdle in practice, as the earth in general does not behave like that. In such situation, the so-called *recursive* depth migration in the $x - \omega$ domain can be a solution. For this we go back to the phase shift operator for homogeneous media, as given by equation (4.80). With this operator, we can describe the propagation of a wave field from depth level z over a distance Δz (in the 2D domain):

$$\tilde{P}(k_x, z + \Delta z, f) = \tilde{W}(k_x, \Delta z, f)\tilde{P}(k_x, z, f), \tag{4.96}$$

which describes a forward propagation. The inverse propagation, as used in migration, is given by:

$$\tilde{P}(k_x, z + \Delta z, f) = \tilde{W}^*(k_x, \Delta z, f)\tilde{P}(k_x, z, f). \tag{4.97}$$

This equation can also be rewritten in the space domain as a convolution:

$$P(x, z + \Delta z, f) = W^*(x, \Delta z, f) * P(x, z, f), \tag{4.98}$$

With $W^*(x, \Delta z, f)$ a convolution operator, which is the inverse Fourier transform of the (complex conjugate of the) phase shift operator.

Normally, we are in a medium with varying velocity as a function of x and z. However, if we can assume a *local* homogeneous medium, i.e. that the medium is homogenous within an area with the horizontal length of the spatial convolution operator $W^*(x, \Delta z, f)$ and

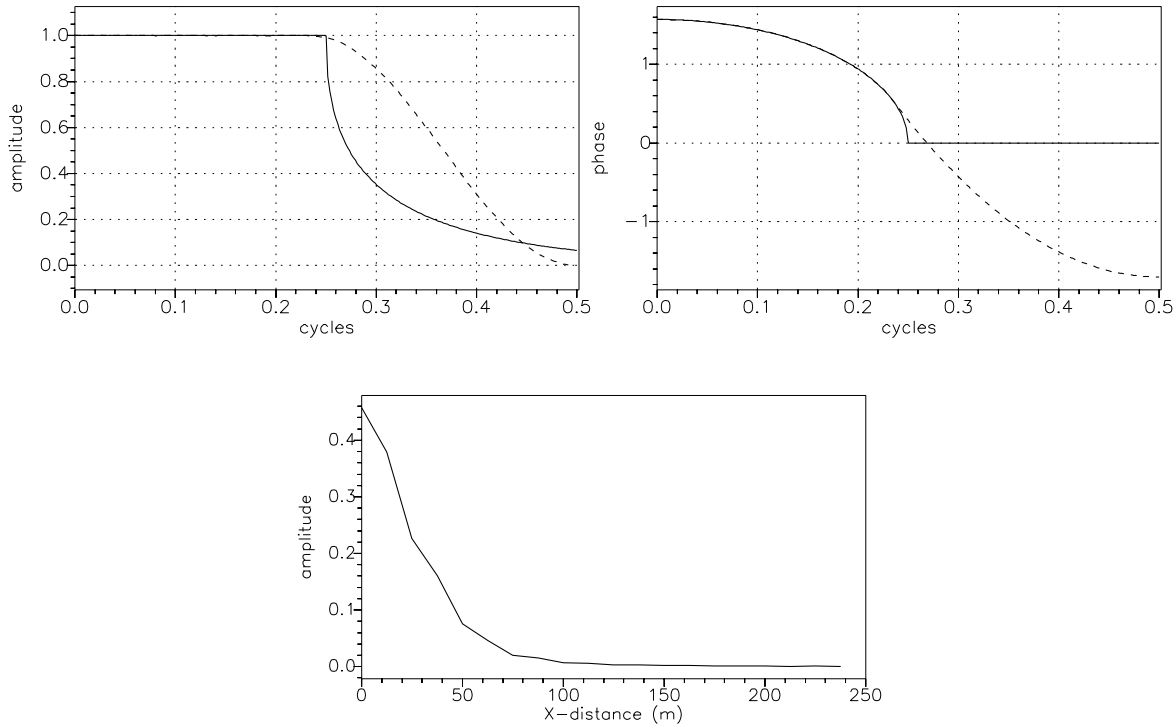


Figure 4.39: Amplitude and phase of the 2-dimensional homogeneous phase shift operator in the wavenumber domain. The solid line is the true operator, the dashed line the optimized short version. Below, the inverse Fourier transform of the optimized operator to the space domain (the amplitude is shown).

depth Δz. Then, we can for each part in the medium use a different convolution operator $W^*(x, \Delta z, f)$, based on the local velocity, which is in fact a slice from the 3D operator as shown in figure 4.38. Figure 4.39 shows the phase shift operator in the wave number domain. Often, the objective for these type of migrations is to create operators in the space domain that are as short as possible. The dashed line in figure 4.39 is such an optimized operator, which is identical to the true operator up to a certain wavenumber (i.e. propagation angle), and is short in the space domain (see figure 4.39 below). We see that for the given example the medium should be homogeneous within an area of approximately 200 m.

The final migration algorithm consists of the following steps:

- Create short operators for different velocities and frequencies and store them in a table.
- Start with the zero offset data at surface in the frequency domain $P(x, z = 0, f)$.
- Apply an extrapolation to depth level Δz for all frequencies, using space dependent operators:
  $P(x, z + \Delta z, f) = W^*(x, \Delta z, f) * P(x, z, f)$.

- Select the $t = 0$ component of the inverse extrapolated result, which is the migrated result at that depth level:
  $P_{mig}(x, z) = \int P(x, z, f)df$.
- Repeat the recursive extrapolation and imaging for each depth level.

Note that this process is easily extended to the 3D case, using short convolutional operators in the (x, y, f) domain.

Finite-difference time migration

The other type of migration we discuss here and which is often applied in seismic processing, is finite-difference migration. It is based on a *differential* approach, rather than an *integral* approach (as was the case with the previous methods). It involves the so-called one-way wave equation, and some approximation to the vertical wavenumber.

Let us return to the homogeneous wave equation which was used as a starting point for Kirchhoff migration, but now in the 2D situation, i.e.:

$$\nabla^2 p(x, z, t) - \frac{1}{c^2}\partial_t^2 p(x, z, t) = 0. \tag{4.99}$$

For models for which the velocities only vary in the vertical direction, we can easily apply our two-dimensional Fourier transform to this equation to obtain:

$$\frac{d^2}{dz^2}\tilde{P}(k_x, z, f) + 4\pi^2\left(\frac{f^2}{c^2} - k_x^2\right)\tilde{P}(k_x, z, f) = 0. \tag{4.100}$$

This is an equation which is easy to solve:

$$\tilde{P}(k_x, z, f) = \tilde{P}_0 \exp(\pm 2\pi i \int_0^z k_z dz), \tag{4.101}$$

where k_z is as given earlier (4.79) and the minus sign in the exponent is chosen for *forward* extrapolation, or for *inverse* extrapolation we have to use the plus sign. This solution is very simple, and actually a solution of a simpler differential equation than we started off with, namely the so-called *one-way wave equation*:

$$\frac{d}{dz}\tilde{P}(k_x, z, f) = -2\pi i k_z \tilde{P}(k_x, z, f). \tag{4.102}$$

This assumes that we are considering a wave field propagating in one direction, without interaction with inhomogeneities (reflection etc). Again, the phase shift operator as given by equation (4.80) is a solution of this one-way wave equation. As for the $f - k_x$ migration, we are interested in time migration so we convert z to τ via:

$$\tau = 2\int_0^z \frac{1}{c} dz, \tag{4.103}$$

so that $d\tau = 2dz/c$. Also invoking the property that we deal with zero offset data in a half velocity medium, we obtain:

$$\frac{d}{d\tau}\tilde{P}'(k_x, \tau, f) = -2\pi i \Omega \tilde{P}'(k_x, \tau, f) \tag{4.104}$$

where we have used Ω as before (equation (4.87)), and we have used P' to denote that it does depend on τ rather than z.

In finite-difference migration, an approximation is made of Ω, by making a Taylor expansion of the square root of Ω :

$$\begin{aligned} \Omega &= \left(f^2 - \frac{k_x^2 c^2}{4}\right)^{1/2} = f\left(1 - \frac{k_x^2 c^2}{4f^2}\right)^{1/2} \\ &\simeq f\left(1 - \frac{k_x^2 c^2}{8f^2}\right) = f - \frac{k_x^2 c^2}{8f}. \end{aligned} \tag{4.105}$$

With this approximation the one-way wave equation is written as:

$$\frac{d}{d\tau}\tilde{P}(k_x,\tau,f) \approx -2\pi i[f - \frac{k_x^2 c^2}{8f}]\tilde{P}(k_x,\tau,f) \tag{4.106}$$

$$= -2\pi i f\tilde{P}(k_x,\tau,f) + 2\pi i\frac{k_x^2 c^2}{8f}\tilde{P}(k_x,\tau,f). \tag{4.107}$$

For a simpler solution, we assume that we have a wavefield $\tilde{P'}$ that is related to the wavefield-extrapolated wavefield $\tilde{Q}$ by:

$$\tilde{P'} = \tilde{Q}\exp(-2\pi i f\tau), \tag{4.108}$$

in which τ is again the vertical traveltime. This can be seen as the time shift for vertical propagation, i.e. the solution for $k_x = 0$. The extrapolation process (going from one to another depth level) has mainly two effects on the wave field: reducing the overall time by a vertical time shift and a contraction of the wave field (e.g. collapsing of diffractors). The general time shift with τ compensates this first effect. The remaining effect will be described in the function $\tilde{Q}$. With other words, the function $\tilde{Q}$ describes a correction for other than vertical angles on this solution. Using this formulation for $\tilde{P'}$, working out $d\tilde{P'}/d\tau$ and using the approximation for Ω, the approximated one-way wave equation becomes:

$$\frac{d\tilde{Q}}{d\tau} = 2\pi i\frac{k_x^2 c^2}{8f}\tilde{Q}. \tag{4.109}$$

The factor of the right-hand side can be written as $(c^2/8)\cdot(-2\pi i k_x)^2/(2\pi i f)$ and we can recognize $(2\pi i f)$ as a differentiation to t in time and $(-2\pi i k_x)$ as a differentiation with respect to x. The inverse Fourier transform of this equation then becomes:

$$\frac{\partial^2 Q}{\partial\tau\partial t} = \frac{c^2}{8}\frac{\partial^2 Q}{\partial x^2}. \tag{4.110}$$

This is the equation used for finite-difference time migration. The most important approximation to derive this, is the expansion of Ω.

For implementation in a computer algorithms, the derivatives are written as finite differences, e.g.:

$$\frac{\partial Q}{\partial t} \approx \frac{Q(t+\Delta t) - Q(t-\Delta t)}{2\Delta t}. \tag{4.111}$$

For spatial derivatives a similar expression is used. With these approximations, it means that each derivative involves a small operator in time and space. The output is calculated as a recursive application of these finite difference operators to calculate the result for small steps in τ.

Note that the 3D extension is rather straightforward, as the spatial derivative $\frac{\partial^2 Q}{\partial x^2}$ is replaced by $\frac{\partial^2 Q}{\partial x^2} + \frac{\partial^2 Q}{\partial y^2}$.

This expression is theoretically valid for velocities which vary only in the vertical direction, but practically are also used for (smooth) lateral velocity variations. Because of the expansion of Ω, the scheme is only valid for certain ranges of Ω, which can be easily related to the structural dip. Theoretical studies have pointed out that dips up to 15 degrees can be handled accurately enough, and that's why this is called the 15-degree finite-difference scheme. However, in practice, dips up to 35 degrees can be handled. Higher-order approximations to Ω can be used, and this is done in the so-called steep-dip or 45 degrees finite-difference algorithm [Claerbout, 1985]. As the name suggests, the method can handle dips up to 45 degrees to a sufficient degree.

Apart from the velocity being input to the finite-difference migration scheme, there is another important parameter and that is the depth step size, $\Delta\tau$. Since we deal with a difference scheme, the scheme can be expensive when the depth step size is taken very small. However, too large a step in $\Delta\tau$ causes the algorithm to undermigrate, so not migrate enough. Often, the undermigration is accompanied by some dispersive "noise" which is an effect of approximating differential operators with difference operators. For a more extensive discussion on practical aspects, the reader is recommended to read the appropriate sections in [Yilmaz, 1987], page 277.

Finite difference depth migration

Before discussing depth migration in more detail, we would like to point out the sometimes badly used word depth in it. With depth migration we mean that we map onto depth instead of onto time. Sometimes depth migration is used for a migration which includes an extra term which corrects for lateral velocity variations ([Yilmaz, 1987]). However, in this terminology it suggests that the extra term only exist when we map onto depth, and not in time; this is not true. Depth migration is important when strong lateral velocity variations exist and then the picture in time can suggest some structure while if the image would be made in depth, we would see that there is no structure at all. An example of this has already been given in the previous chapter in figure (3.26). Of course, the section can always be better in depth because then it corresponds more to a geological cross-section, but this is often a difficult task because of the sensitivity to the velocity. Here also another advantage of migration to a time section becomes important, namely that the algorithm is not so sensitive to the velocity. The corrections due to a migration to vertical traveltime are more corrections due to dip, while with a migration to depth we at the same time correct for velocities as well. This sensitivity can be compared to determining interval velocities on a time-migrated section which is also a very sensitive process. A seismic interpreter should be well aware of the velocity effect when interpreting a time section. We will adopt the approach that by depth migration, we mean that we map onto depth,

not necessarily including the extra term in the migration procedure.

In the above, we derived the migration in terms of the vertical traveltime τ by converting the depth z to τ by means of equation (4.103). We will not do that in this section. On top of that, we will specifically consider strong lateral velocity variations because then it is strongly recommended to map onto depth rather in time in order to prevent misinterpretation.

Let us first introduce a velocity c_{lat} which depends on the horizontal as well as vertical coordinates, $c_{lat} = c_{lat}(x, z)$. Using this velocity we define a velocity $\bar{c}$ which is a horizontal average of the velocity c_{lat} so:

$$\bar{c}(z) = \frac{1}{x_{max} - x_{min}} \int_{x_{min}}^{x_{max}} c_{lat} dx. \tag{4.112}$$

This velocity $\bar{c}$ is used as being constant locally. Now let us define the vertical traveltime τ by:

$$\tau = 2 \int_0^z \frac{dz}{\bar{c}} = \frac{2z}{\bar{c}}, \tag{4.113}$$

where we have now used the laterally invariant $\bar{c}$.

We can now follow the same procedure as before when we employed the one-way wave equation. Since we here deal with depth migration, we will use the one-way wave equation in z rather than in τ. When we go through the same derivation as for the finite-difference time migration, we now have that $d\tau/dz = 2/\bar{c}$, while for k_z we use the same expansion as for Ω (equation(4.105)) so:

$$k_z \simeq \frac{2f}{c_{lat}} - \frac{c_{lat} k_x^2}{4f}, \tag{4.114}$$

in which we have kept c_{lat}. Again we assume a solution in the form of:

$$\tilde{P} = \tilde{Q} \exp(-2\pi i f \tau) = \tilde{Q} \exp(-4\pi i f z/\bar{c}). \tag{4.115}$$

Using these in the one-way wave equation of equation (4.102) yields:

$$\frac{d\tilde{Q}}{dz} \approx 2\pi i \frac{c_{lat} k_x^2}{4f} \tilde{Q} + 2\pi i f \left(\frac{1}{\bar{c}} - \frac{1}{c_{lat}} \right) \tilde{Q}. \tag{4.116}$$

We see that if the velocity c_{lat} is only a function of the depth z, then the extra term on the right-hand side vanishes and the equation is equivalent to the one we derived before for time migration (equation (4.109)). The first term on the right-hand side in this equation is called the *diffraction* term, while the second term is called the *thin-lens* term.

Rewriting using the formulation for spatial and temporal derivation, we obtain in the $t - x$ domain:

$$\frac{\partial^2 Q}{\partial t \partial z} \approx \frac{c_{lat}}{4} \frac{\partial^2 Q}{\partial x^2} + 2 \left(\frac{1}{\bar{c}} - \frac{1}{c_{lat}} \right) \frac{\partial^2 Q}{\partial t^2}. \tag{4.117}$$

Because of the dependence on z, this is called depth migration. The extra term accounts for lateral velocity variations, and is only important when strong lateral velocity variations exist. For slow lateral velocity variations, the approximation as discussed in time migration, is sufficiently accurate.

Note that we could have included lateral velocity variations in the time migration as well. Then we obtain in the (k_x, f) domain, following the same procedure as before,

$$\frac{d\tilde{Q}}{d\tau} \approx 2\pi i \frac{\bar{c} c_{lat} k_x^2}{8f} \tilde{Q} + 2\pi i f \left(1 - \frac{\bar{c}}{c_{lat}}\right) \tilde{Q}. \tag{4.118}$$

Note the extra term on the right.

When to use which technique

We have so far discussed the most popular techniques to perform the migration, but all of them have some advantages and some disadvantages. In practice, the seismic processor will decide which algorithm to use, based on the situation that is faced. As some (depth) migration algorithms require interval velocities, which may not be known yet, the availble information can limit the number of possibilities. The various options are listed below:

Kirchhoff time migration

Advantages

- Simplicity, based on NMO velocities
- Can handle steep dips
- Adaptable to unusual source-receiver geometries

Disadvantages

- Cannot handle low Signal-to-Noise ratio's
- Cannot handle lateral velocity variations
- Improper amplitudes
- Possible aliasing

Finite-difference time migration

Advantages

- Can handle slow lateral velocity variations
- Can handle low Signal-to-Noise ratio's

Disadvantages

- Slow in computational speed
- Cannot handle dips above 45 degrees
- Cannot handle complex media

Migration by Fourier transform (Gazdag, Stolt)

Advantages

- Fast in computational speed

- Able to handle steep dips
- Can handle low Signal-to-Noise ratio's

Disadvantages

- Cannot handle lateral velocity variations

Depth migration (in comparison to time migration)

Advantages

- Can handle vertical and lateral velocity variations
- Can handle steep dips

Disadvantages

- Requires accurate interval velocity model
- Slow in computational speed

Another description of when to use which type of migration, is given in [Yilmaz, 1987], table 4-1, page 246.

The usual goal of a seismic processor is to obtain a section which represents as much as possible a geological cross-section (where we only have an impedance map!). So some conversion to depth has to take place. Also here in different circumstances, it is known when to use which technique. This is given in [Yilmaz, 1987], figure 5-8, p. 361. (A note must be made about Yilmaz here. He suggests that only a migration mapped onto depth includes the extra thin-lens term; this is not true. Only when a depth section is desired, then the depth migration can directly map onto depth instead of the vertical intercept-time (τ).)

Summary of zero offset migration

After stacking the resulting section is considered as a zero offset experiment. This zero offset experiment can again be considered as the result of a so-called exploding reflector *experiment in the half velocity medium: each reflector point is considered as a source; all these source fire at the same time. By inverse extrapolation of this exploding reflector experiment through this half-velocity medium and imaging at t=0 at each depth level, the resulting migrated image is retrieved.*

For this inverse extrapolation the Kirchhoff integral is used. This integral has been derived form the wave equation and states that the wavefield of each point within a volume can be determined from the wavefield values at the boundary of this volume and a number of Green's functions that describe the propagation from the desired point to all boundary locations.

In migration practice, the boundary is chosen as the earth's surface, and all points below the surface can be determined. Depending on the complexity of the medium several migration approaches can be followed. In the simplest case: the medium is homogeneous, a mapping procedure (Stolt migration) can be used, that achieves a perfect migration result by a forward and inverse temporal and spatial Fourier transform of the data, and a coordinate

mapping in between. If the medium is inhomogeneous, other procedures need be followed. If the velocity distribution only varies as a function of z then a recursive migration via the wavenumber-frequency domain can be applied (Gazdag phase shift migration) and in the case of (local) lateral and depth variations, Kirchhoff migration (based on traveltime functions), finite difference methods (in time or depth) or recursive $f-x$ migration can be used. Each method has its pros and cons, and the seismic processor makes his/her decision based on the complexity of the model and the available resources (computer power or financial).

4.10 Conversion from time to depth

In the previous section we have spoken of time and depth migration, referring to whether the output section is in time or depth, respectively. In time, we do not need to know the velocities that well, stacking velocities will often do. In depth migration we need to know the velocities very well, which is often a difficult task. Still, our goal is to obtain a section which is as close as possible to a geological cross-section; to that effect we want to have our section in depth. In this section we will briefly discuss the conversion from time to depth, especially in which circumstances you can use certain techniques.

Dix formula

Let us first consider a model with plane horizontal layers. We showed in Chapter 3 that we could determine the root-mean-square velocities from the interval velocities via:

$$c^2_{rms,N} = \frac{1}{T_{tot,N}(0)} \sum_{i=1}^{N} c_i^2 T_i(0), \tag{4.119}$$

where we have included an extra N in the notation of $c_{rms,N}$ and $T_{tot,N}$. We can invert this formula, which means that we can determine the interval velocities from the root-mean-square velocities. When we consider the root-mean-square velocities for $N=2$ and $N=3$, we have:

$$c^2_{rms,2} = \frac{c_1^2 T_1(0) + c_2^2 T_2(0)}{T_1(0) + T_2(0)} \tag{4.120}$$

$$c^2_{rms,3} = \frac{c_1^2 T_1(0) + c_2^2 T_2(0) + c_3^2 T_3(0)}{T_1(0) + T_2(0) + T_3(0)} \tag{4.121}$$

We bring the denominator on the right-hand side to the left-hand side, subtract the first equation from the second, and obtain:

$$c^2_{rms,3}(T_1(0) + T_2(0) + T_3(0)) - c^2_{rms,2}(T_1(0) + T_2(0)) = c_3^2 T_3(0) \tag{4.122}$$

in which we recall that $T_3(0)$ is the zero-offset traveltime through layer 3, so in fact the difference between the total time up to the time at level 3 minus the time at level 2, so $T_3(0) = T_{tot,3}(0) - T_{tot,2}(0)$. So then the interval velocity c_3 becomes:

$$c_3 = \sqrt{\frac{c^2_{rms,3} T_{tot,3}(0) - c^2_{rms,2} T_{tot,2}(0)}{T_{tot,3}(0) - T_{tot,2}(0)}} \tag{4.123}$$

In general the interval velocity for the i^{th} layer is given by:

$$c_i = \sqrt{\frac{c_{rms,i}^2 T_{tot,i}(0) - c_{rms,i-1}^2 T_{tot,i-1}(0)}{T_{tot,i}(0) - T_{tot,i-1}(0)}} \tag{4.124}$$

The values for $c_{rms,n}$ and $T_{tot,n}$ can directly be obtained from the velocity file as used for stacking the data. This is Dix formula [Dix, 1955]. Dix' formula converts RMS-velocities to interval velocities.

In our procedure to get a depth section for a model with horizontal plane layers, we convert the time axis on our (zero-offset) stacked section to a depth axis using this formula.

Although we derived Dix formula for horizontal layers, the formula will still be good when we have mild lateral velocity variations. It has been shown that even in the case of dipping events, the formula will still be good. In that case however, in order to obtain a good depth section, we must first time-migrate the data (without the thin-lens term) before we can convert the time axis to a depth axis.

Image rays

The above formula breaks down when the lateral velocity variations become larger. However, the concept of image rays has helped us still to be able to convert the time section to a depth section. What is an image ray? An image ray is a ray which goes down vertically from the surface; this is in contrast to normal-incidence rays for which the ray is perpendicular to the reflection interface. An example is given in figure (4.40), taken from [Yilmaz, 1987].

Let us now look at a simple model in which we have a point diffractor buried below a dipping interface between two media which show quite a strong contrast in velocity (4.41). When we look at the zero-offset section we see a shape which does not look like a hyperbola any more; it is skewed. It has the lowest time, its apex, at a receiver location that is not right above the diffractor, but laterally shifted. In the figure, **B** is the point above the diffractor, and **A** is the point of the apex of the traveltime curve. However, the ray picture shows a very interesting feature, namely that at point **A** the ray path is perpendicular to the surface. This was first recognized by [Hubral, 1977]. We can use the image ray to perform a lateral shift, which is equivalent to applying the thin-lens term in the finite-difference migration scheme. We should realize that with the image ray, we apply first the diffraction term in the (time) migration, and only later perform the lateral shift as predicted by the image ray.

In order to convert the time section to a depth section (note: section instead of axis) for models for which the lateral velocity variations are not too large, we can use the image-ray concept rather than full depth migration. Time migration will always position the reflection information at the apex of the migration operator. The procedure to obtain the depth section is then to do a time migration with only the diffraction term, and then convert to depth *along the image rays.* In this way, that reflection information is positioned at the correct depth and lateral location.

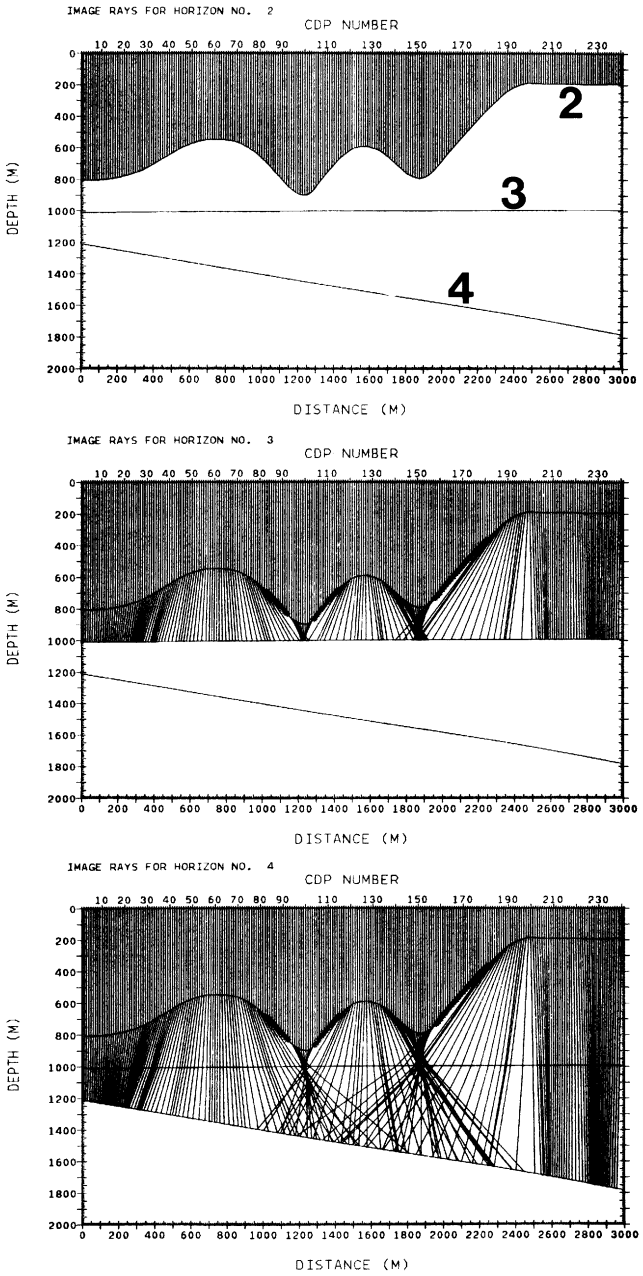


Figure 4.40: The image rays for a model with 3 reflectors. From Yilmaz (1987), fig 5-7

Conversion to depth via depth migration

There is some point when even the image-ray approach breaks down. It was already said that the image rays are used *after* a time migration has been done, but sometimes the migration cannot be split into these two parts, and one has to apply the thin-lens term alternately with the diffraction term. Whether the image-ray approach breaks down can be inspected via a plot of the image rays. When more than one image ray is associated with a subsurface point, we should apply the diffraction and thin-lens term alternately and use a full depth migration in order to get the appropriate depth section. Such a case

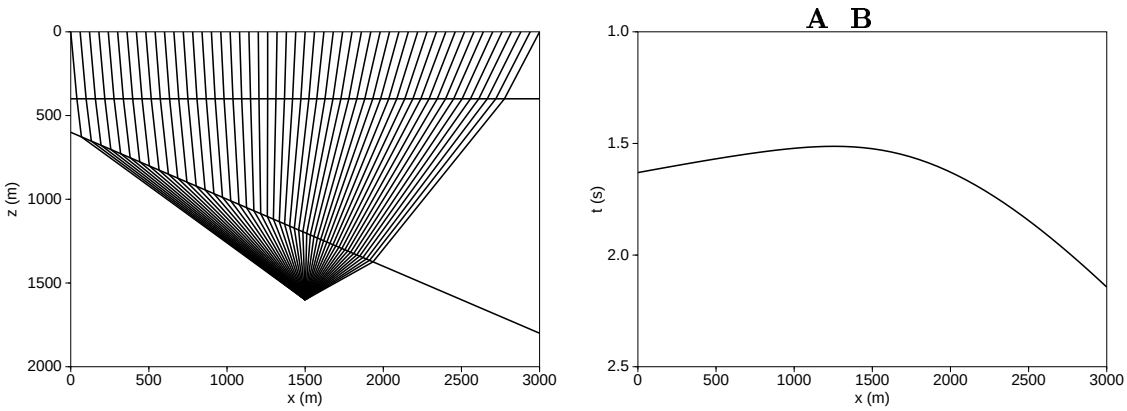


Figure 4.41: The ray paths for a point diffractor in a zero-offset section (left), with its time section (right). (Adapted from Yilmaz (1987))

in shown in figure (4.40), where for the lowest reflector the many parts are illuminated by more than one image ray.

All the considerations of this section are summarized in the figure as can be found in [Yilmaz, 1987], fig5-8, p.361, which is included here for completeness in figure (4.42).

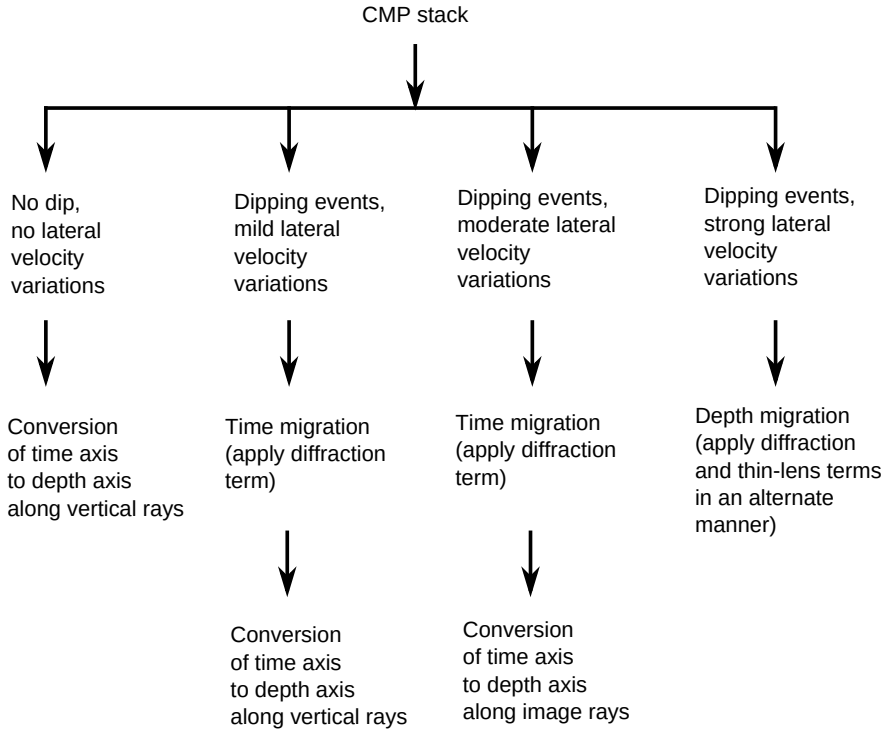


Figure 4.42: Strategy for obtaining a depth section. (Adapted from Yilmaz (1987))

4.11 Prestack migration

In the previous chapters and sections, a general seismic processing procedure has been discussed that consists of the following major steps:

1 CMP sorting

2 NMO correction

3 DMO correction

4 Stack

5 Poststack time/depth migration

In a number of these steps some assumptions have been made that are not valid for general inhomogeneous earth models, such as:

1 CMP sorting: if the earth is inhomogeneous and reflectors have complex shapes the reflection events within a CMP gather do not belong to one subsurface reflection point (see section 4.8).

2 NMO correction: in complex subsurface media the moveout in a CMP gather is not hyperbolic, so a perfect moveout correction cannot be achieved.

3 DMO: if strong lateral velocity and/or reflector geometry variations are present, the DMO procedure still will not resolve the reflection point smear within a CMP gather.

4 Stack: as the events within a CMP gather do not belong to the same subsurface reflection point, stacking of these events will mix subsurface information.

5 Poststack time/depth migration: given the approximations in the previous steps, a stacked section does not represent a true zero offset section and migration of this stack will therefore not result in an exact image. Furthermore, some poststack migration algorithms have limitations due to the assumed simple velocity field or limitation in the maximum dip that can be handled.

In fact this procedure is still used so often as it is robust and very efficient. Furthermore, as a first indication it still serves a prupose to do some fast and robust CMP-oriented processing. However, if the results are not satisfactory, a true *prestack migration* procedure need be followed: the original field shot records need be directly migrated into a subsurface image. In such a procedure, the exploding reflector model does not work anymore: the prestack dataset is a *two-way wavefield*, with a mix of up and downgoing propagation effects. A prestack migration procedure is a more complex procedure, but could be explained as the following sequence:

1 Put a point source in the shot position of the shot record under consideration and apply a forward extrapolation to each subsurface point. For this step still the Kirchhoff extrapolation procedure can be used, but now with the *full* velocity model (not half velocity).

2 The wave field of the shot record (i.e. the measured response at the receivers) is *inverse* extrapolated to each subsurface point, using again the Kirchhoff formula.

3 At each depth level the two wave fields (source and receiver field) are *correlated*, which means that the downward extrapolated receiver wave field is corrected in time by the corresponding source wave field. The contribution that appears at $t = 0$ is the prestack depth migration contribution for this subsurface point.

4 By repeating this procedure for all shot records in the dataset, and adding all resulting images, the complete subsurface image is obtained.

Note that this procedure is much more expensive, as the stacking procedure is now applied *after* migration. However, complex subsurface models with large lateral velocity variations (e.g. in a salt dome environment) require such an elaborate method in order to correctly image structures below or near the edges of such a salt dome structure.

Stacking after migration has also another advantage: as neighboring shot records illuminate largely the same subsurface area, it means that (if the correct velocity model has been used) after migration at a certain reflector location, all migrated contributions from each experiment should have the same depth. Based on this knowledge, prestack-related velocity analysis can be done: if the reflections do not align from the different seismic experiments after migration, the velocity model need be adapted and migration is recalculated. although this procedure is much more computational intensive than the velocity analysis based on NMO correction, the main advantage again is the accuracy in complex models, where CMP gathers do not show any hyperbolic moveout behavior, and the NMO based velocity analysis does not apply at all. A similar list of migration procedures also exist for prestack depth migration, as still the wave field extrapolation process takes a central role. Therefore, Kirchhoff techniques, Gazdag, Stolt, finite difference and recursive depth migration algorithms are available to the user. As the prestack migration procedure was not the main objective of this course, we will not go into these details.

Chapter 5

3D seismic processing

5.1 Introduction

So far we considered most of the seismic processing only in a 2D mode: assuming sources and receivers were positioned along one coordinate, and also the earth was assumed to be invariant in the direction perpendicular to the acquisition line. In reality this is definitely not the case. Up to the seventies, seismic processing was indeed fully 2D oriented. Acquisition was done along lines. But by combining the results from many parallel lines, a 3D view of the earth could still be obtained (although the sampling in the cross-line direction was very coarse). In the eighties, full 3D seismic came into development. For each shot several parallel lines of receivers were recorded simultaneously, thus creating a much denser coverage of the earth. Nowadays, true 2D seismic acquisition belongs to the exception, and is sometimes only done for certain research purposes. Together with the change in the acquisition method, the processing algorithms had to follow the transition from 2D to 3D. In this chapter, an overview of the major consequences of this extra dimension are described.

5.2 Midpoint oriented processing

Even nowadays, the vaste amount of data is processed still in a midpoint-oriented way. This means that prestack data is sorted into common midpoint (CMP) gathers and stacked after application of the proper moveout correction. The midpoints are now defined on a two-dimensional grid at the surface, instead of along one coordinate only. Typically, the midpoint spacing for a 3D survey is in the order of 25 m in both directions. Sometimes the in-line spacing is smaller than the cross-line spacing. The major difference with 2D in this processing approach is just that the amount of data has increased enormously. A small 3D seismic survey of 10x10 km, with 50 traces at every 25x25 m grid cell (which is called a *bin*) contains thus $8 \text{x} 10^6$ traces, which results in 0.1 Tbyte of data. Typical 3D surveys can easily consist of several Tbytes.

Figure 5.1: 3D marine seismic acquisition, with multiple streamers towed behind a vessel. *(Picture taken from Veritas DGC website)*

A 3D midpoint-oriented processing scheme will look typically like:

1. Geometry assign
2. Preprocessing: groundroll removal, statics correction, deconvolution
3. CMP sorting
4. NMO velocity analysis and correction
5. 3D DMO
6. Stack
7. 3D Poststack time/depth migration

In the following some of these steps are further described.

3D geometry assignment

An important step in 3D seismic processing is the assignment of the correct geometry to the seismic measurements. At first this sounds like a trivial taks, but in practice this needs careful labour work. Before 3D seismic acquisition takes place, the survey is designed, based on a compromise between costs of acquisition and the required accuracy in the image. For this, the expected presence of strong noise in the seismic records or special requirements in the imaging of the subsurface structures can play an important role.

For 3D marine seismic the design of the survey is limited by the practical possibilities: sources and receivers need be towed behind one (or more) vessels (see Figure 5.1). Typically, a 3D marine survey consists of a dual source, alternatively shooting into 7 streamers. Each streamer has a length from 3 km up to sometimes 8 km. The streamers have a typical cross-line spacing of 100 m, such that with the two sources the cross-line midpoint spacing becomes 25 m. However, during shooting these long cables cannot be kept in a straight line behind the ship. Cross-line currents will move the cables such that so-called feathering and snaking effects can occur. This means that the receivers can be easily shifted 100 m away from its supposed position. Using GPS systems mounted on the cables, the exact position of each hydrophone can be determined. All this position information is separately stored. After all data has been shot, the geometry information needs be merged with the seismic measurements.

For 3D land data, the design of an seismic acquisition is not bounded by the fact that sources and receiver need be in-line with each other. But there other aspects play a role, e.g., the cost of planting geophones in the ground, the conditions of the surface, obstacles in the area (water, buildings, mountains etc.). Typically, a 3D acquisition on land has a cross-spread type geometry: sources are densely sampled in one direction and coarsely sampled in the perpendicular direction, and the receiver spacing is also dense in one and coarse in the other direction, but in opposite directions compared to the sources. In this way the coverage of the subsurface becomes well distributed over the area. Except in desert areas, the intended acquisition geometry can never be exactly achieved in practice, and in the end an irregular geometry is obtained. Like in the marine case, the position of each source and receiver is determined and this information need be transferred to the seismic traces after the acquisition has finished.

All traces will get the correct header information (source and receiver x and y coordinates). Besides that, each trace will get an in-line and cross-line CMP (i.e. bin) number. Once this geometry information is assigned to the seismic data, all kinds of quality control plots can be made.

An example of such a plot is a coverage (or fold) plot, like shown in Figure 5.2. For a 3D marine dataset the number of traces that have a midpoint in a certain bin is displayed in color-code. If the sources and receiver positions would be stricitly regular, this coverage plot would have one color only, as the geometry was designed that way. However, due to the above mentioned effects, this fold plot can display a variable coverage. Moreover, also in marine acquisition obstacles can be present, like an oil platform. This will be observed in the coverage plot as a low-fold area, as the ship has to manouver around the obstacle.

Besides the fact that in an irregular acquisition geometry not all bins are covered by the same amount of traces, also the exact midpoint positions do not fall in the center of each bin. In fact, the binning process assign a bin number to each traces, meaning that the geometric midpoint of such a trace falls within the boundaries of a grid cell. But the actual midpoint can also be plotted on a map in order to quantify the midpoint smearing within each grid cell. Figure 5.3 shows such a plot for a 3D marine dataset, that has encountered feathering effects. Therefore, midpoints are spread out over the grid cells (even moving to a neighbouring cells). Another effect that one is faced with in 3D acquisition is that the distribution of offsets for one CMP gather can vary from bin to bin. Figure 5.4 shows for

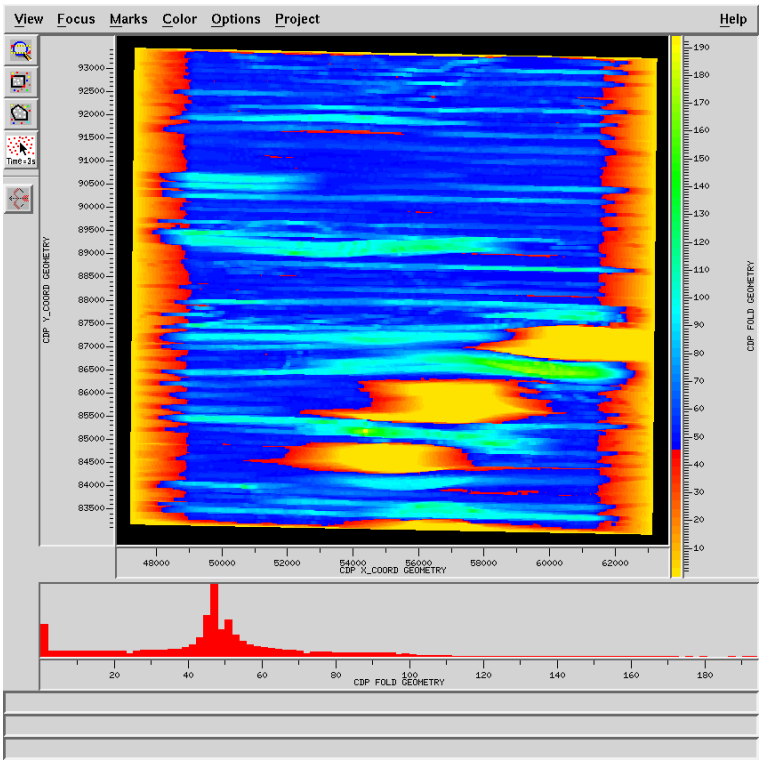


Figure 5.2: An example of a coverage plot for a 3D marine survey. Note the holes due to oil platforms in this area.

one fixed cross-line location in Figure 5.3 the distribution of offsets as a function of the in-line midpoint coordinate. The feathering effects can be clearly observed here. Note that for some midpoint locations (like the left box at CMP 2100) a very regular distribution of offsets is visible, whereas other midpoints (e.g. the box around CMP 2115) have a very irregular offset distribution. An irregular offset distribution has the effect that correlated noise events - like multiples or gorundroll - will not be suppressed that well. Also due to AVO effects in the desired reflection data (AVO = Amplitude versus Offset) the stacked amplitude will vary with varying offset distribution. This may result into false amplitude anomalies on the stacked section.

CMP sorting and velocity analysis

After the geometry has been assigned to the traces, and also all midpoints have been defined (i.e. the binning process), the seismic data can be sorted into CMP gathers. For the two CMP locations in Figure 5.4 the seismic trace are plotted at their actual offset location in Figure 5.5. An NMO correction has been applied to facilitate the interpretation of the events. The evens that show a residual curvature in these plots are identified as multiples or converted waves. Note the irregular offset distribution for CMP 2115 (right hand side of Figure 5.5).

The processing per CMP gather is similar to the 2D case. A stacking velocity analysis is carried out in the CMP domain. The result is a stacking velocity field as a function of

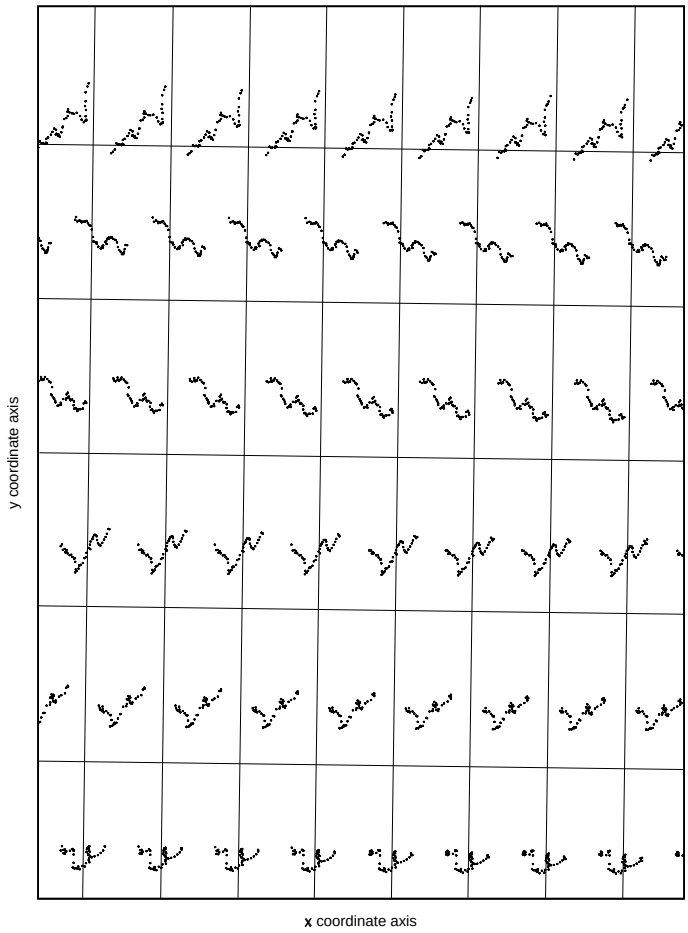


Figure 5.3: An example of a midpoint plot for a 3D marine survey.

the in-line and cross-line coordinate (i.e. bin number). Then each seismic trace is NMO corrected with its velocity function and these traces are stacked to simulate the zero offset trace - with high SNR - at that particular CMP position.

DMO correction

Like in the 2D case, the effect of dipping structures in the subsurface can be corrected for in a separate DMO step. However, all calculations are now carried out in a 3D sense, i.e. along the x- and y-coordinate. As DMO acts as a lateral summation process (i.e. a partial migration) and it is often applied to 3D seismic data even if there is no complex strucutural information. It has a tendency to smear out all seismic amplitudes, and thus reduces the so-called *acquisition footprint*. This means that due to the acquisition, not all CMP gathers are evenly filled (e.g. varying coverage or offset distribution within each bin). This will generate an amplitude pattern on the stack. Applying 3D DMO will reduce this effect.

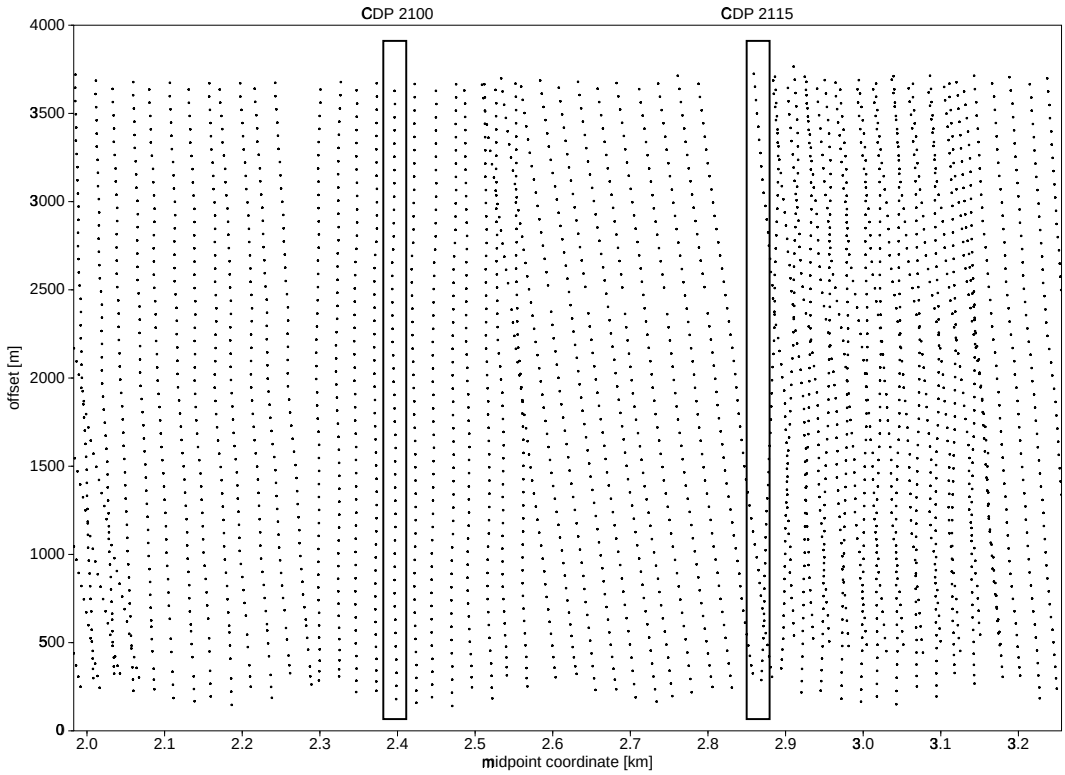


Figure 5.4: Plot of the distribution of offsets as a function of the in-line midpoint location for one fixed cross-line position for the data of Figure 5.3. Note the irregular coverage of offsets for the various midpoints. The two rectangles indicate the CMP locations for which the seismic data is shown in Figure 5.5.

Stacking

After the NMO correction, and optionally a DMO correction, all traces from one CMP gather are stacked such that for each bin location one trace is produced, simulating the zero offset trace at this position. Thus, after stacking a volume of stacked traces is obtained. This stack is now densely sampled in the in-line and cross-line direction.

5.3 3D Poststack migration

To obtain the final subsurface image after the stacking process, a 3D poststack migration needs be carried out. In fact, all migration algorithms as discussed in the previous chapter have their implementation for both 2D and 3D data. Moreover, most migration formulas have been derived in the 3D sense there. If the earth has variations in all directions, a full 3D migration process is the only way to accurately reveal its structures (see Figure 5.6). For a homogenous earth this means that each stacked trace is smeared out along a sphere. Or viewed from another way, all seismic amplitudes along a hyperboloid will be added together to form the reflection for one subsurface point. Of course, the earth is non-homogenous, and depending on the complexity of the earth a choice between various

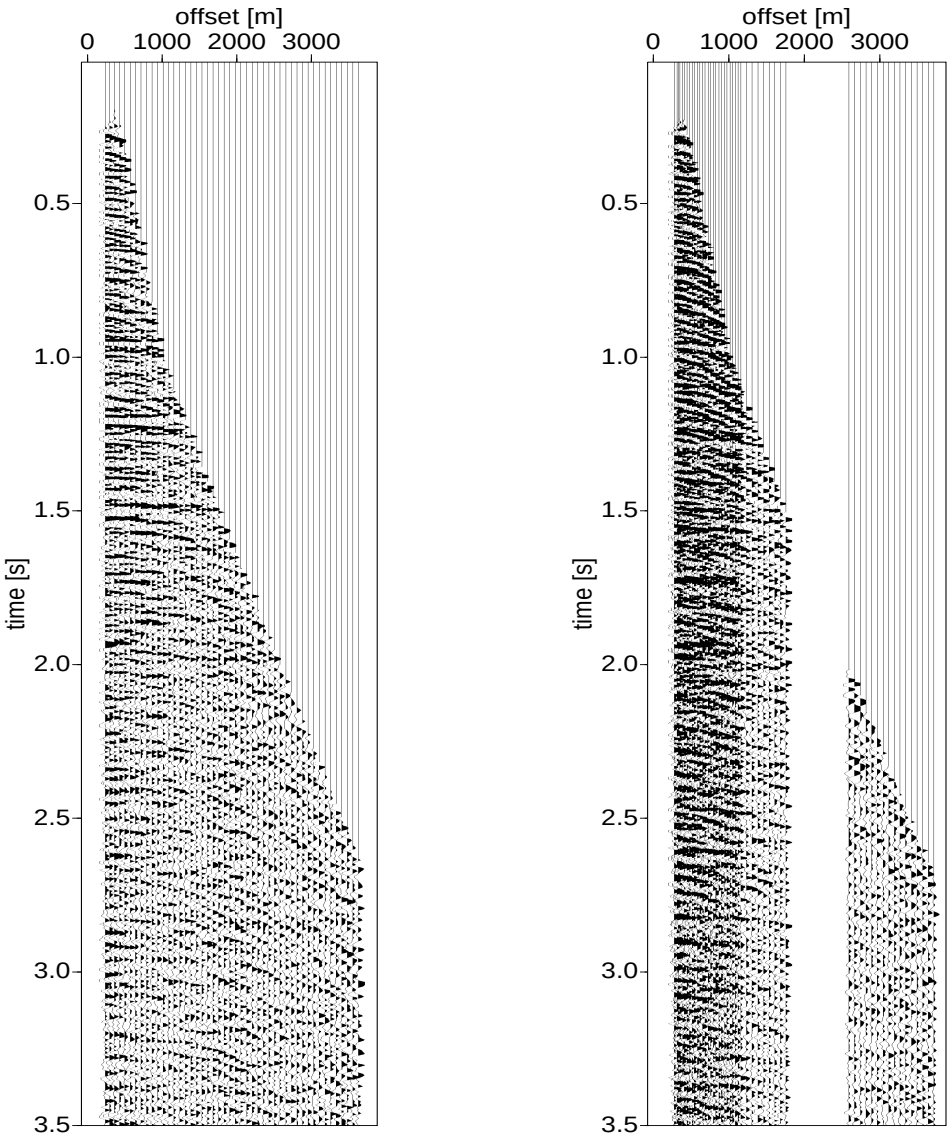


Figure 5.5: For the two CMP locations indicated in Figure 5.4 the seismic traces are plotted at the true offset locations. Note that the left hand side CMP has a very regular and the right hand side CMP a very irregular distribution of offsets.

3D migration algorithms need be made.

Besides the extra dimension involded in the processing steps, also in the interpretation of the seimic images a new dimension has been added. Small events that look like artifacts in a 2D cross-section of the earth's subsurface appear to be genuine structures if you look at the image in a 3D sense.

Examples of such structures can be prehistoric riverbeds that appear as channels throughout a seismic section. In Figure 5.7 one slice at constant depth through a 3D migrated seismic dataset is displayed. To obtain a better view of the fine details, the

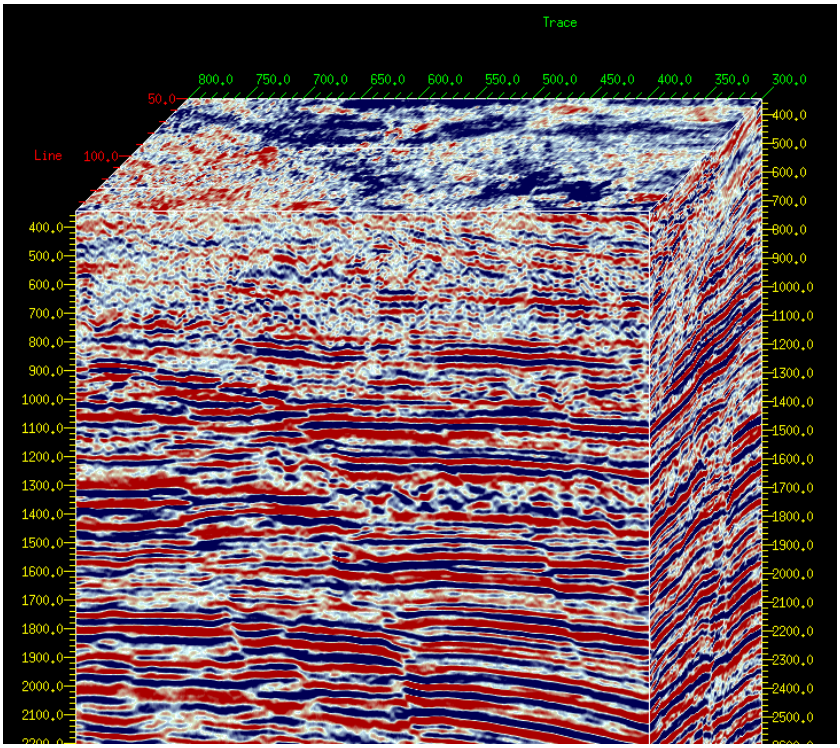


Figure 5.6: After 3D processing and migration, a volume image of the earth is obtained. (*picture taken from ExxonMobil website*)

original data (left hand side) has been filtered such that small edges become better visible (right hand side). The ancient rivers that are present at this depth become well visible. Often, river sands are good oil reservoirs. Therefore, the detection of these structures is of high importance to the oil and gas exploration industry. Such channels are hardly recognizable on a 2D cross-section of the earth. Note that the circular shapes in the top of the left side image in Figure 5.7 are due to a salt dome; these circular events show the boundaries of this dome (like cutting throug a mountain). Inside this dome, the imaging shows a very complicated pattern (see also the same area after filtering). This is a well-known feature of salt bodies: the salt has a very heterogenous character.

Furthermore, several interfaces can be picked or tracked in a volumetric way, resulting in a 3D map of the structures of interest, that can be viewed from different angles. An example of an integrated display of a seismic vertical cross-section, a few well-logs and an interpreted horizon is shown in Figure 5.8.

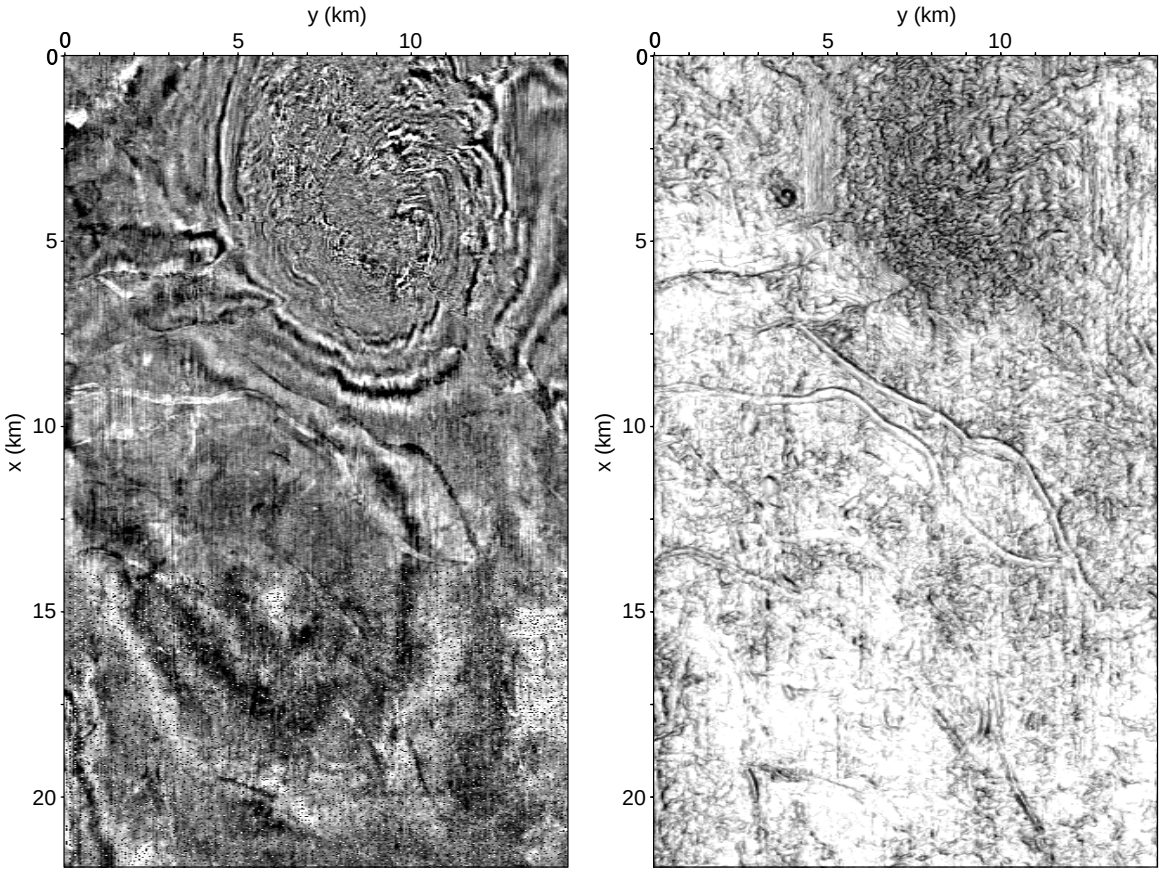


Figure 5.7: Depth slice through 3D migrated section before (left) and after (right) edge-detection filtering.

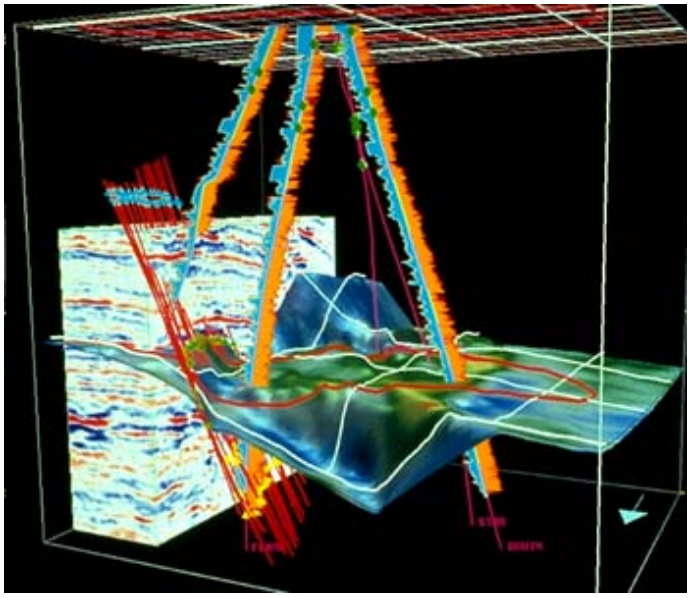


Figure 5.8: One vertical slice of a 3D migration combined with an interpreted horizon through the full survey. (*picture found on the internet*)

Appendix A

Discretisation of Fourier transform

The continuous Fourier integrals are nearly always used for deriving mathematical results, but, in performing transforms on numerical data, the integrals are always replaced by summations. The continuous signal $a(t)$ becomes the discrete signal, or time series, a_k, in which k is an integer, and the sampling has taken place at regular intervals $k\Delta t$. Thus the discrete signal corresponds exactly to the continuous signal at times

$$t = k\Delta t \tag{A.1}$$

Consider the evaluation of the Fourier transform (2.2) at the discrete times $k\Delta t$:

$$a_k = \int_{-\infty}^{\infty} A(f) \exp(2\pi i f k\Delta t) df \qquad k = ..., -2, -1, 0, 1, 2, ... \tag{A.2}$$

where a_k stands for the fact that time is now discrete so:

$$a_k = a(t), \text{when } t = k\Delta t \qquad k = ..., -2, -1, 0, 1, 2, ... \tag{A.3}$$

This integral may be replaced by an infinite sum of pieces of the integral:

$$a_k = \sum_{m=-\infty}^{\infty} \int_{\frac{m}{\Delta t}-\frac{1}{2\Delta t}}^{\frac{m}{\Delta t}+\frac{1}{2\Delta t}} A(f) \exp(2\pi i f k\Delta t) df \qquad k = ..., -2, -1, 0, 1, 2, ... \tag{A.4}$$

In order to get the bounds of the integral from $-1/(2\Delta t)$ to $+1/(2\Delta t)$, we change to the variable $f' = f - m/\Delta t$ to yield:

$$a_k = \sum_{m=-\infty}^{\infty} \int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} A(f' + \frac{m}{\Delta t}) \exp(2\pi i \{f' + \frac{m}{\Delta t}\} k\Delta t) df' \qquad k = ..., -2, -1, 0, 1, 2, ... \tag{A.5}$$

Changing the order of the integration and summation, and noting that the exponential becomes periodic (so $\exp(2\pi i m k) = 1$), this becomes

$$a_k = \int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} \left[\sum_{m=-\infty}^{\infty} A(f' + \frac{m}{\Delta t}) \right] \exp(2\pi i f' k\Delta t) df' \qquad k = ..., -2, -1, 0, 1, 2, ... \tag{A.6}$$

The Fourier transform of the discrete time series is thus

$$a_k = \int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} A_D(f) \exp(2\pi i f k \Delta t) df \qquad k = ..., -2, -1, 0, 1, 2, ... \tag{A.7}$$

provided

$$A_D(f') = \sum_{m=-\infty}^{\infty} A(f' + \frac{m}{\Delta t}) \tag{A.8}$$

So this is an infinite series of shifted spectra as shown in figure (A.1b) The discretisation of the time signal forces the Fourier transform to become periodic. In the discrete case we get the same spectrum as the continuous case if we only take the period from $-1/(2\Delta t)$ to $+1/(2\Delta t)$, and else be zero; the signal must be *band-limited*. So this means means that the discrete signal must be zero for frequencies $|f| \geq f_N = 1/(2\Delta t)$. The frequency f_N is known as the Nyquist frequency.

Let us now look at the other integral of the continuous Fourier- transform pair, i.e. (2.1). We evaluate the integral by discretisation , so then we obtain for $A_D(f)$:

$$A_D(f) = \Delta t \sum_{k=-\infty}^{\infty} a_k \exp(-2\pi i f k \Delta t) \tag{A.9}$$

In practice the number of samples is always finite since we measure only for a certain time. Say we have N samples. Then we have to determine the function $B_D(f)$ which resembles closely the one with infinitely many samples. Say, we define $B_D(f)$ as:

$$B_D(f) = \Delta t \sum_{k=0}^{N-1} b_k \exp(-2\pi i f k \Delta t) \tag{A.10}$$

and we have to determine the coefficients b_k. For that purpose, we use the least-squares method, that means that we determine the coefficients b_k such that the average quadratic error:

$$E_N = \int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} (A_D(f) - B_D(f))^2 df \tag{A.11}$$

is minimized for a fixed value of N. E_N can be viewed as a function of $\{b_n\}$, thus:

$$E_N = E_N(b_0, b_1, b_2, \ldots, b_{N-2}, b_{N-1}) \tag{A.12}$$

In order to minimize E_N we must put the partial derivatives with respect to b_n to zero, i.e.,:

$$\partial E_N / \partial b_n = 0, \qquad \text{for } n = 1, 2, \ldots, N-1 \tag{A.13}$$

Working this out gives us:

$$\begin{aligned} \partial E_N / \partial b_n &= 2 \int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} (A_D(f) - B_D(f)) \{ -\frac{\partial B_D(f)}{\partial b_n} \} df \\ &= -2 \int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} (A_D(f) - B_D(f)) \exp(-2\pi i f n \Delta t) \Delta t df \end{aligned} \tag{A.14}$$

$$= -2\Delta t \int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} \Delta t \sum_{k=-\infty}^{\infty} a_k \exp(-2\pi i f k \Delta t) \exp(-2\pi i f n \Delta t) df$$

$$-2\Delta t \int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} (-\Delta t) \sum_{k=0}^{N-1} b_k \exp(-2\pi i f k \Delta t) \exp(-2\pi i f n \Delta t) df$$

Since the integration variable does not depend on the summation variable, we can interchange the summation and the integration:

$$\partial E_N/\partial b_n = -2(\Delta t)^2 \sum_{k=-\infty}^{\infty} a_k \int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} \exp(-2\pi i f k \Delta t) \exp(-2\pi i f n \Delta t) df$$

$$+2(\Delta t)^2 \sum_{k=0}^{N-1} b_k \int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} \exp(-2\pi i f k \Delta t) \exp(-2\pi i f n \Delta t) df \qquad \text{(A.15)}$$

Because of the orthogonality relations, the integrals are zero, unless k equals n; in that case the integral becomes $(1/\Delta t)$. We thus obtain:

$$\partial E_N/\partial b_n = -2(\Delta t)^2 a_n \frac{1}{\Delta t} + 2(\Delta t)^2 b_n \frac{1}{\Delta t} \qquad n = 0, 1, 2, \ldots, N-1 \qquad \text{(A.16)}$$

Since each of these partial derivatives are set equal to zero, we obtain the nice result that the coefficients b_k are equal to the coefficients a_k which are the coefficients for the infinite sum.

Combining all this information, we obtain the Fourier pair:

$$A_D(f) = \Delta t \sum_{k=0}^{N-1} a_k \exp(-2\pi i f k \Delta t) \qquad \text{(A.17)}$$

$$a_k = \int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} A_D(f) \exp(2\pi i f k \Delta t) df \qquad k = 0, 1, 2, ..., N-1 \qquad \text{(A.18)}$$

This is the transform pair for continuous frequency and discrete time. Notice that the integral runs from $-1/2\Delta t$ to $+1/2\Delta t$, i.e. one period where one spectrum of $A_D(f)$ is present.

So far, we considered the fact that the values for frequencies above the Nyquist frequency must be set to zero. This way of looking at it, is a frequency-domain consideration. We can translate this to the time domain by saying that if there is no information in the continuous time signal $a(t)$ at frequencies above f_N, the maximum sampling interval Δt is

$$\Delta t_{max} = \frac{1}{2 f_N} \qquad \text{(A.19)}$$

This is the sampling theorem.

If we choose Δt too large , we under-sample the signal and we get aliasing as shown in Figure (A.2). The original signal appears to have a lower frequency.

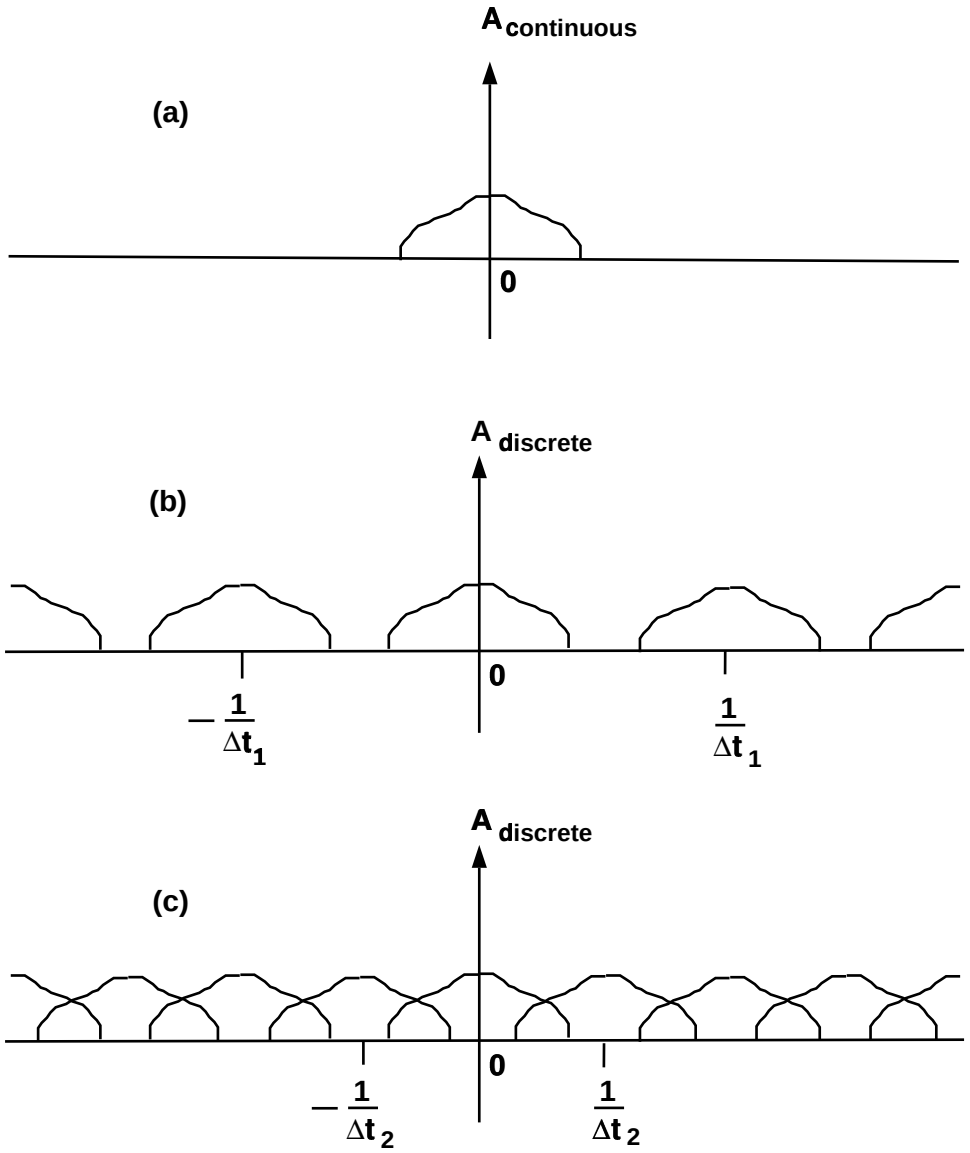


Figure A.1: Effect of time-discretisation in frequency domain: (a) continuous spectrum; (b) properly time-sampled spectra giving rise to periodicity (period $1/\Delta t_1$; (c) too coarse time sampling Δt_2 such that spectra overlap (= aliasing in time domain).

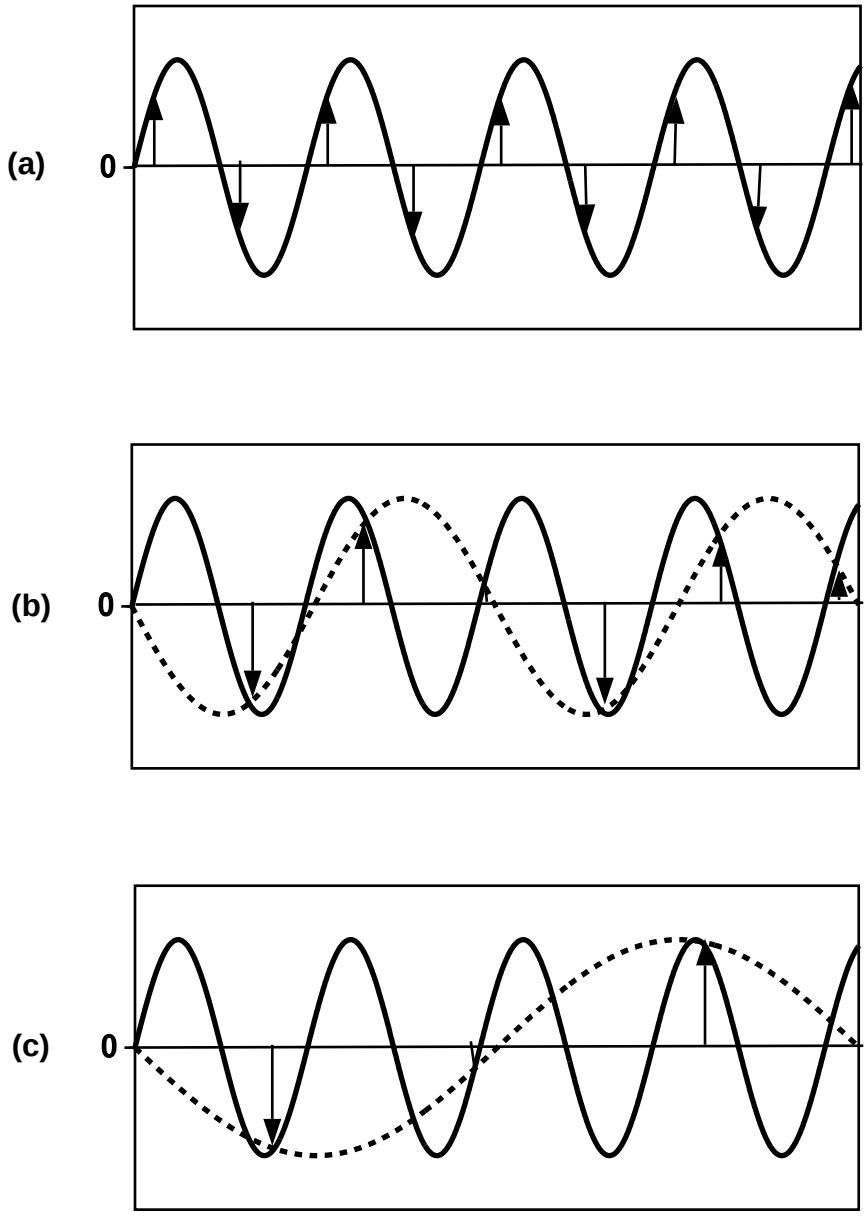


Figure A.2: Effect of discretisation in time: (a) properly sampled signal; (b) just under-sampled signal; (c) fully under-sampled signal.

In practice the number samples in a time series is always finite. We wish to find the discrete Fourier transform of a finite length sequence. We approach the problem by dividing the definite integral (A.7) into the sum of N pieces of equal frequency interval Δf. Because $A_D(f)$ is periodic, with period $1/\Delta t$, we may first rewrite the integral with different limits, but with the same frequency interval:

$$a_k = \int_0^{\frac{1}{\Delta t}} A_D(f) \exp(2\pi i f k \Delta t) df \qquad k = 0, 1, 2, ..., N-1 \tag{A.20}$$

Writing the integral as a summation, we obtain

$$a_k = \Delta f \sum_{n=0}^{N-1} A_n \exp(2\pi i n \Delta f k \Delta t) \qquad k = 0, 1, 2, ..., N-1 \tag{A.21}$$

where

$$A_n = A_D(f), \text{when} f = n\Delta f. \tag{A.22}$$

We now notice that the series a_k is periodic with period N:

$$a_{k+N} = \Delta f \sum_{n=0}^{N-1} A_n \exp(2\pi i n \Delta f \{k+N\} \Delta t) \tag{A.23}$$

$$= \Delta f \sum_{n=0}^{N-1} A_n \exp(2\pi i n \Delta f k \Delta t + 2\pi i n \Delta f N \Delta t) \tag{A.24}$$

$$= \Delta f \sum_{n=0}^{N-1} A_n \exp(2\pi i n \Delta f k \Delta t) \tag{A.25}$$

$$= a_k \tag{A.26}$$

since $N\Delta f = 1/\Delta t$ and so $\exp(2\pi i n) = 1$. Thus we arrive at the following discrete Fourier transform pair for a finite-length time series

$$A_n = \Delta t \sum_{k=0}^{N-1} a_k \exp(-2\pi i n k/N) \qquad n = 0, 1, 2, ..., N-1 \tag{A.27}$$

$$a_k = \Delta f \sum_{n=0}^{N-1} A_n \exp(2\pi i n k/N) \qquad k = 0, 1, 2, ..., N-1 \tag{A.28}$$

These two equations are the final discrete-time and discrete- frequency Fourier transform pair.

Appendix B

Derivation of the wave equation

In this appendix we will derive the wave equation for homogeneous media, using the conservation of momentum (Newton's second law) and the conservation of mass. In this derivation, we will follow [Berkhout, 1984] (appendix C), where we consider a single cube of mass when it is subdued to a seismic disturbance (see figure (B.1)). Such a cube has a volume ΔV with sides $\Delta x, \Delta y$ and Δz.

Conservation of mass gives us:

$$\Delta m(t_0) = \Delta m(t_0 + dt) \tag{B.1}$$

where Δm is the mass of the volume ΔV, and t denotes time. Using the density ρ, the conservation of mass can be written as:

$$\rho(t_0)\Delta V(t_0) = \rho(t_0 + dt)\Delta V(t_0 + dt) \tag{B.2}$$

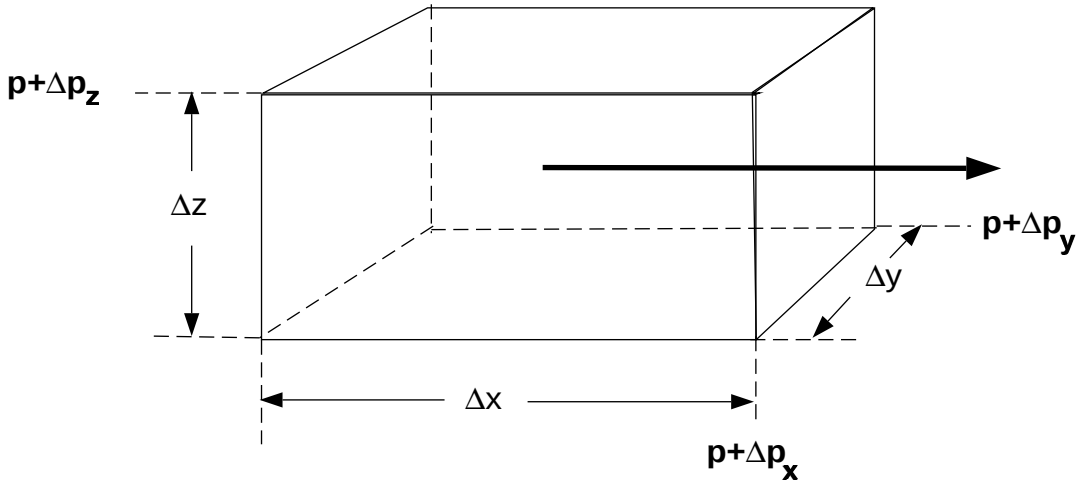


Figure B.1: A cube of mass, used for derivation of the wave equation.

Making this explicit:

$$\begin{aligned} \rho_0 \Delta V &= (\rho_0 + d\rho)(\Delta V + dV) \\ &= \rho_0 \Delta V + \rho_0 dV + \Delta V d\rho + d\rho dV \end{aligned} \tag{B.3}$$

Ignoring lower-order terms, i.e., $d\rho dV$, it follows that

$$\frac{d\rho}{\rho_0} = -\frac{dV}{\Delta V} \tag{B.4}$$

We want to derive an equation with the pressure in it so we assume there is a linear relation between the pressure p and the density:

$$dp = \frac{K}{\rho_0} d\rho \tag{B.5}$$

where K is called the bulk modulus. Then, we can rewrite the above equation as:

$$dp = -K \frac{dV}{\Delta V} \tag{B.6}$$

which formulates Hooke's law. It shows that for a constant mass the pressure is linearly related to the relative volume change. Now we can also derive that:

$$\begin{aligned} \frac{dV}{\Delta V} &= \frac{(\Delta x + dx)(\Delta y + dy)(\Delta z + dz)}{\Delta x \Delta y \Delta z} - \frac{\Delta x \Delta y \Delta z}{\Delta x \Delta y \Delta z} \\ &\simeq \frac{dx}{\Delta x} + \frac{dy}{\Delta y} + \frac{dz}{\Delta z} + O(dxdy) + O(dxdz) + O(dydz) \end{aligned} \tag{B.7}$$

For dx we can write:

$$\begin{aligned} dx &= (v_x dt)_{x+\Delta x} - (v_x dt) \\ &= \frac{\partial (v_x dt)}{\partial x} \Delta x \end{aligned} \tag{B.8}$$

where v_x denotes the particle velocity in the x–direction. We can do the same for the y and z–component and obtain:

$$\begin{aligned} \frac{dV}{\Delta V} &\simeq \left[\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} \right] dt \\ &= (\nabla \cdot \mathbf{v}) dt \end{aligned} \tag{B.9}$$

Substitute this in Hooke's law (equation B.6):

$$dp = -K(\nabla \cdot \mathbf{v}) dt \tag{B.10}$$

or

$$\frac{1}{K} \frac{dp}{dt} = -\nabla \cdot \mathbf{v} \tag{B.11}$$

The term on the left-hand side can be written as :

$$\frac{1}{K}\frac{dp}{dt} = \frac{1}{K}\left[\frac{\partial p}{\partial t} + \mathbf{v}\cdot\nabla p\right] \tag{B.12}$$

Ignoring the second term in brackets (low-velocity approximation), we obtain for equation (B.11):

$$\frac{1}{K}\frac{\partial p}{\partial t} = -\nabla\cdot\mathbf{v} \tag{B.13}$$

This is one basic relation needed for the derivation of the wave equation.

The other relation is obtained via Newton's law applied to the volume ΔV :

$$\Delta\mathbf{F} = \Delta m\frac{d\mathbf{v}}{dt} \tag{B.14}$$

where $\mathbf{F}$ is the (vectorial) force working on the element ΔV. Consider the force in the x–direction:

$$\begin{aligned}\Delta F_x &= -\Delta p_x \Delta S_x \\ &= -\left(\frac{\partial p}{\partial x}\Delta x + \frac{\partial p}{\partial t}\Delta t\right)\Delta S_x \\ &\simeq -\frac{\partial p}{\partial x}\Delta V \end{aligned} \tag{B.15}$$

ignoring the term with Δt, and ΔS_x is the surface in the x–direction, thus $\Delta y\Delta z$. So we can write:

$$\begin{aligned}\Delta\mathbf{F} &= -\left(\frac{\partial p}{\partial x}, \frac{\partial p}{\partial y}, \frac{\partial p}{\partial z}\right)^T \Delta V \\ &= -\Delta V(\nabla p)\end{aligned} \tag{B.16}$$

Substituting in Newton's law (equation B.14), we obtain:

$$\begin{aligned}-\Delta V(\nabla p) &= \Delta m\frac{d\mathbf{v}}{dt} \\ &= \rho\Delta V\frac{d\mathbf{v}}{dt}\end{aligned} \tag{B.17}$$

We can write $d\mathbf{v}/dt$ as $\partial\mathbf{v}/\partial t$; for this we have used a low-velocity approximation:

$$\frac{d\mathbf{v}}{dt} = \frac{\partial\mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v} \approx \frac{\partial\mathbf{v}}{\partial t} \tag{B.18}$$

We divide by ΔV to give:

$$-\nabla p = \rho_0\frac{\partial\mathbf{v}}{\partial t} \tag{B.19}$$

This equation is called the equation of motion.

We are now going to combine the conservation of mass and the equation of motion. Therefore we let the operator $(\nabla\cdot)$ work on the equation of motion:

$$\begin{aligned} -\nabla\cdot(\nabla p) &= \nabla\cdot\left(\rho\frac{\partial \mathbf{v}}{\partial t}\right) \\ &= \rho\frac{\partial}{\partial t}(\nabla\cdot\mathbf{v}) \end{aligned} \tag{B.20}$$

for constant ρ. Substituting the result of the conservation of mass gives:

$$-\nabla^2 p = \rho_0\frac{\partial}{\partial t}\left(-\frac{1}{K}\frac{\partial p}{\partial t}\right) \tag{B.21}$$

Rewriting gives us the wave equation:

$$\nabla^2 p - \frac{\rho_0}{K}\frac{\partial^2 p}{\partial t^2} = 0 \tag{B.22}$$

or

$$\nabla^2 p - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = 0 \tag{B.23}$$

in which c can be seen as the velocity of sound, for which we have: $c = \sqrt{K/\rho}$.

Appendix C

The definition of SEG-Y

In the following pages you will find a reprint of the article from [Barry et al., 1975] in which the SEG-Y seismic data format is defined.

.

Appendix D

Traveltime equation for a dipping refracting boundary

In this appendix, the traveltime for a refraction on a dipping boundary is derived. In order to obtain the desired expression, the dipping boundary is defined mathematically, and Snell's law is rewritten in a suitable form, before the total traveltime is determined.

Consider the configuration given in figure (D.1). The boundary between the first and second layer has an angle α with the horizontal. The equation for the dipping boundary is given by:

$$z = z_0 + x \tan\alpha \tag{D.1}$$

or, using the depth at the end point z_3:

$$z = z_3 - x \tan\alpha \tag{D.2}$$

A ray goes to the refractor (downgoing ray), is refracted critically in the second layer (refracted ray) and goes back to the surface (upgoing ray). Since the ray is critically refracted in the second layer, we have the following relations:

$$\theta_U = \theta_c + \alpha \tag{D.3}$$

$$\theta_D = \theta_c - \alpha \tag{D.4}$$

in which θ_c is the critical angle, θ_D is the angle with the vertical of the downgoing ray and θ_U of the upgoing ray. Since Snell's law holds for all rays, we have:

$$\frac{\sin(\theta_D + \alpha)}{c_1} = \frac{\sin(\theta_U - \alpha)}{c_1} \tag{D.5}$$

$$\frac{\sin(\theta_D + \alpha)}{c_1} = \frac{1}{c_2} \tag{D.6}$$

$$\frac{\sin(\theta_U - \alpha)}{c_1} = \frac{1}{c_2} \tag{D.7}$$

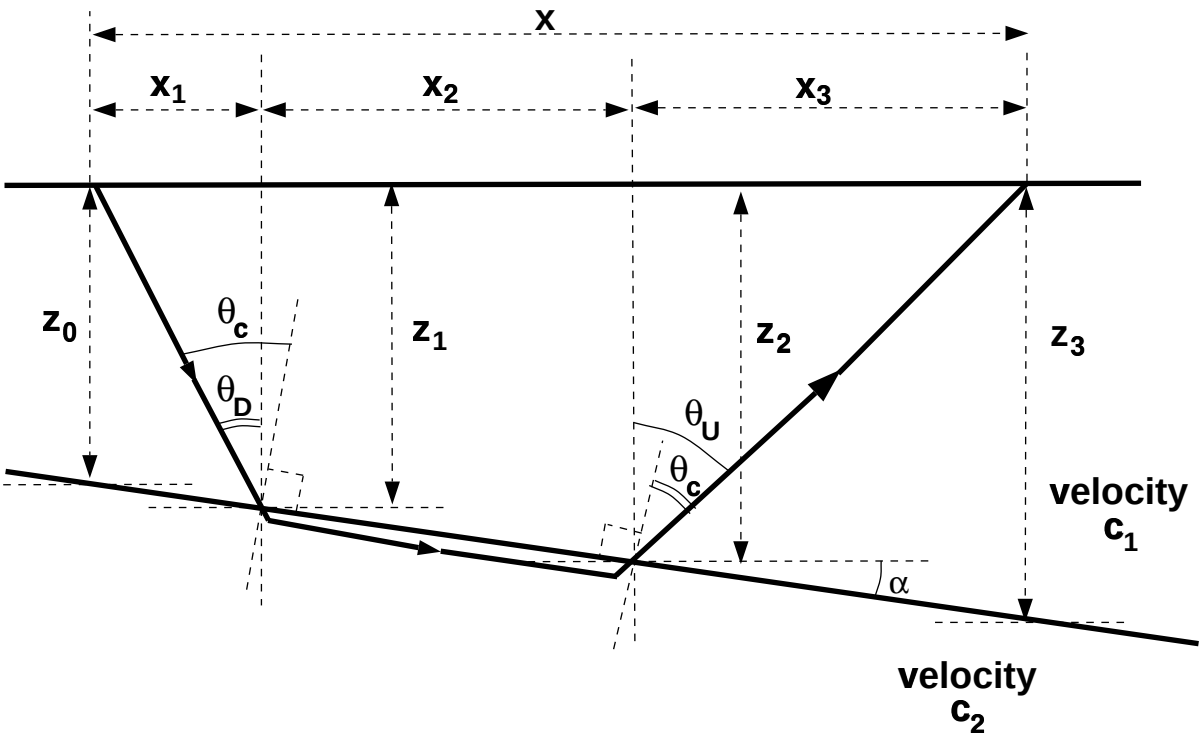


Figure D.1: Configuration for refraction on dipping plane boundary.

For convenience, the horizontal slowness p and vertical slowness q are introduced:

$$p_D = \frac{\sin\theta_D}{c_1} \tag{D.8}$$

$$p_U = \frac{\sin\theta_U}{c_1} \tag{D.9}$$

$$q_D = \frac{\cos\theta_D}{c_1} \tag{D.10}$$

$$q_U = \frac{\cos\theta_U}{c_1} \tag{D.11}$$

Writing the sine's out in equations (D.5)–(D.7), the introduction of the p_D, p_U, q_D and q_U gives the relations:

$$p_D + q_D \tan\alpha = p_U - q_U \tan\alpha \tag{D.12}$$

$$p_D + q_D \tan\alpha = \frac{1}{c_2 \cos\alpha} \tag{D.13}$$

$$p_U - q_U \tan\alpha = \frac{1}{c_2 \cos\alpha} \tag{D.14}$$

We have now all the necessary formulae to derive the traveltime for the refracted ray as drawn in figure (D.1). The total distance between source and receiver is denoted by x; the horizontal distances x_1, x_2 and x_3 are the distances of each part of the ray as drawn in the figure. The total traveltime T becomes:

$$\begin{aligned} T &= p_D x_1 + q_D z_1 + \frac{1}{c_2 \cos\alpha} x_2 + p_U x_3 + q_U z_2 \\ &= p_D x_1 + q_D (z_0 + x_1 \tan\alpha) + \frac{1}{c_2 \cos\alpha} x_2 + p_U (x - x_1 - x_2) + \\ &\quad q_U \left[z_3 - (x - x_1 - x_2) \tan\alpha \right] \end{aligned} \tag{D.15}$$

where we used $x = x_1 + x_2 + x_3$ and the equations for the interface (eqs.(D.1) and (D.2)), where the z_0–equation is put with the downgoing ray and the z_3–equation with the upgoing ray. Now grouping together the terms with x_1, x_2 and x, T reads:

$$\begin{aligned} T &= x_1 (p_D + q_D \tan\alpha - p_U + q_U \tan\alpha) + x_2 (-p_U + q_U \tan\alpha + \frac{1}{c_2 \cos\alpha}) + \\ &\quad x (p_U - q_U \tan\alpha) + q_D z_0 + q_U z_3 \\ &= \frac{1}{c_2 \cos\alpha} x + q_D z_0 + q_U z_3 \end{aligned} \tag{D.16}$$

where we used Snell's law as derived above (eqs.(D.12) and (D.14)). Notice that the second and third terms represent times associated with the begin- and end-point of the ray at the surface. Therefore it can also be written as:

$$T = \frac{1}{c_2 \cos\alpha} x + \tau_0 + \tau_3 \tag{D.17}$$

These last two equations ((D.16) and (D.17)) are the desired expressions.

Appendix E

Correlation of signals

In this appendix we will give some more background information on the correlation process. In Chapter 2 the correlation ϕ_{ab} of two signals $a(t)$ and $b(t)$ was defined as follows:

$$\phi_{ab}(\tau) = \int_{-\infty}^{\infty} a(t)b^*(t-\tau)dt. \tag{E.1}$$

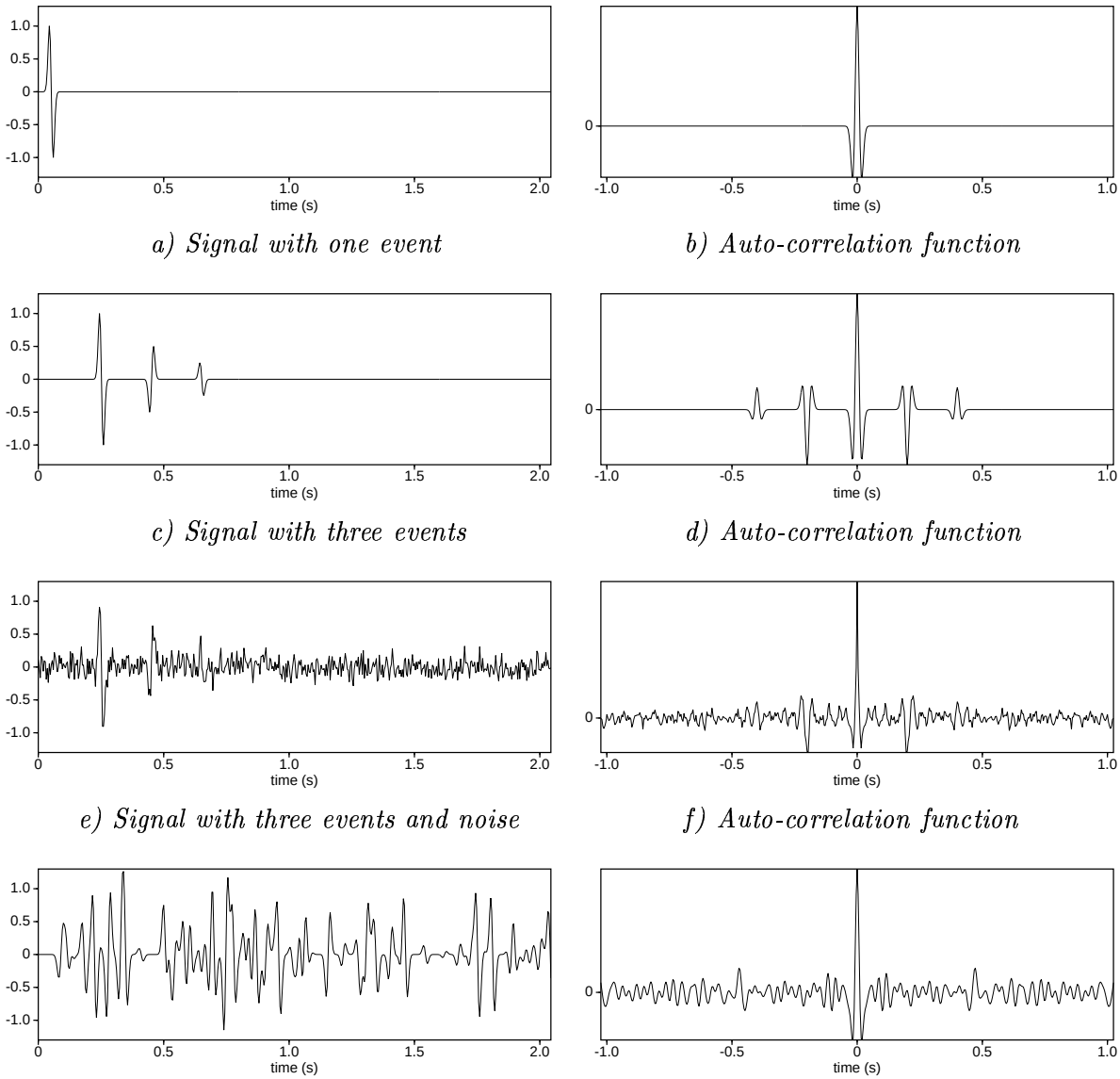
In the frequency domain this becomes:

$$\Phi_{ab}(f) = A(f)B^*(f). \tag{E.2}$$

If A and B are the same signal, then ϕ_{ab} represents the auto-correlation function. Due to the fact that in the freqeuncy domain the auto-correlation spectrum becomes $A(f)A^*(f)$ this is a real valued function and the auto-correlation function in time is symmetric around zero time.

The purpose of a correlation function is to find resemblences between two signals (or within signals). Looking at equation (E.1) we see that the correlation is calculated by shifting one signal across the other and do a sample-by-sample multiplication. This summation result will be larger if the two signals will look similar to each other (such that positive values are multiplied by positive and negative by negative, all contributing to a large correlation value).

Examples of auto-correlation functions for a few signals are displayed in Figure E.1. The first signal is a first derivative of a Gauss wavelet. Note that the autocorrelation of each signal is indeed symmetric around zero time (i.e. zero phase behaviour). Note also that the noisy signal in Figure E.1e has a very spiky autocorrelation. This can be easily understood by considering that a noise signal will only correlate with itself if there is not shift. As soon the signal is shifted by one or more time samples, there is only accidental local correlation. That is why we can assume in practice that a random signal as an autocorrelation function that is a scaled delta pulse. In Figure E.1e the signal with the three shifted Gauss wavelets $y(t)$ is added with noise, resulting in $x(t) = y(t) + n(t)$. The autocorrelation of this is $\phi_{xx} = \phi_{yy} + \phi_{nn} + \phi_{yn} + \phi_{ny}$. Neglecting the cross terms, we see indeed that Figure E.1f seems to be the summation of Figure E.1d and a delta function. In Figure E.1g we convolved a random reflection series with the Gauss wavelet of Figure



a) Signal with one event *b) Auto-correlation function*

c) Signal with three events *d) Auto-correlation function*

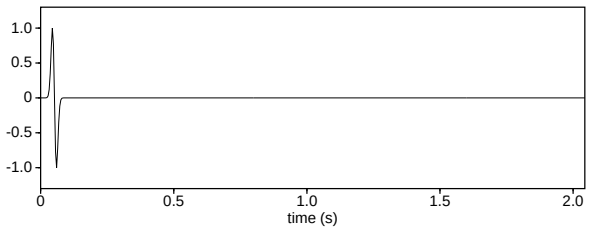
e) Signal with three events and noise *f) Auto-correlation function*

g) Signal with random events *h) Auto-correlation function*

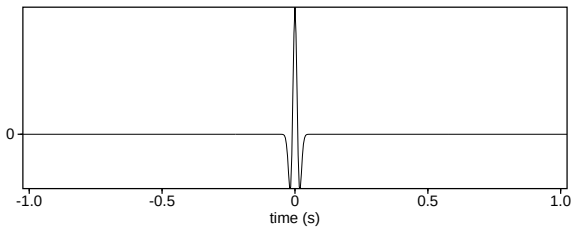
Figure E.1: Four signals, based on the first derivative of Gauss wavelets, and their auto-correlation function.

E.1a. The autocorrelation of this signal $x(t) = g(t) * s(t)$ with $g(t)$ being the random reflection series is given by $\phi_{xx} = \phi_{gg} * \phi_{ss} \approx G\phi_{ss}$. This means that Figure E.1h is a scaled version of Figure E.1b, which appears to be true to some extend.

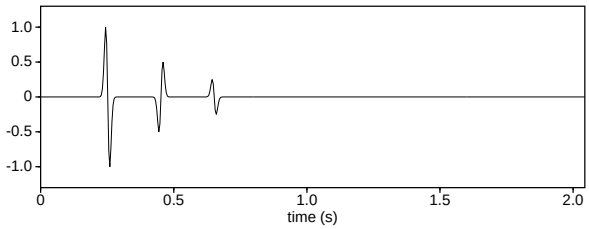
In Figure E.2 the same four signals are shown, but now accompanied by the cross-correlation between each signal and the original Gauss wavelet. It can be clearly observed that these cross-correlation signals are look like the original signal, but everywhere the Gauss wavelet is replaced by its autocorrelation.



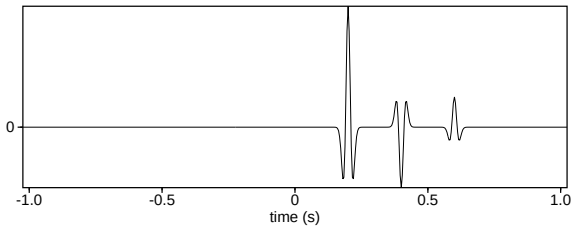
a) Signal with one event



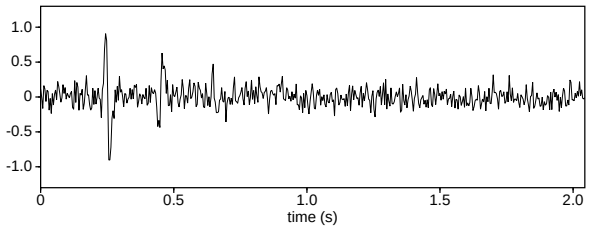
b) Auto-correlation function



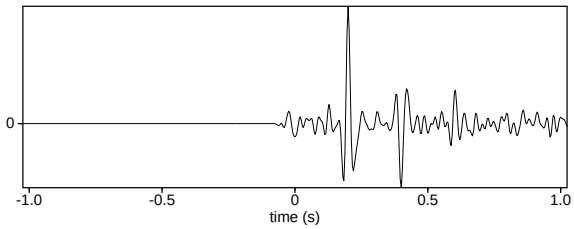
c) Signal with three events



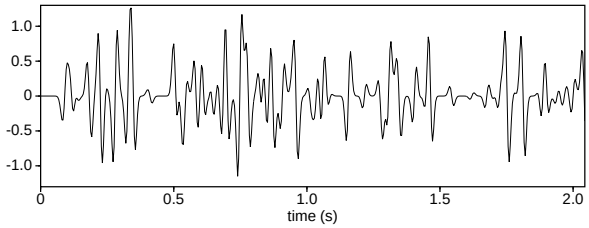
d) Cross-correlation function of c and a



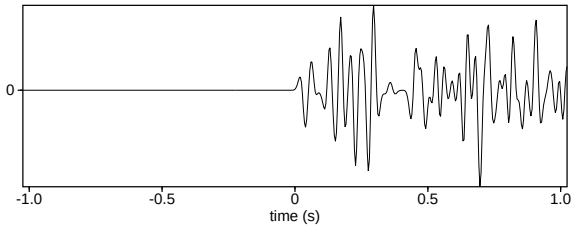
e) Signal with three events and noise



f) Cross-correlation function of e and a



g) Signal with random events



h) Cross-correlation function of g and a

Figure E.2: Four signals, based on the first derivative of Gauss wavelets, and some cross-correlation functions.

Appendix F

Wiener filters

We assume stationary time series: input x_t and desired output d_t are stationary. A stationary time series is a time series with statistical properties that do not change with time. We want to find the best filter such that the actual output $a_t = f_t * x_t$ is close to our desired signal d_t in which the asterisk stands for convolution. In a Wiener filter via the time-domain we minimize the error energy, which is defined as:

$$E = \sum_t (d_t - a_t)^2 = \sum_t (d_t - f_t * x_t)^2 = minimum \tag{F.1}$$

Least-squares solution

We assume the filter f_t has a finite length of $N+1$ points. So the goal is to minimize this error of the energy of the output. To find the solution for filter f_t such that the energy of the error is minimum is called the *least-squares solution* of this problem. Minimizing the error can be achieved by requiring the first derivative with respect to the filter coefficients f_i to be zero, i.e.:

$$\frac{\partial E}{\partial f_i} = 0 \qquad \text{for } i = 0, 1, 2, \ldots, N. \tag{F.2}$$

Working this out gives:

$$\frac{\partial E}{\partial f_i} = 2 \sum_t \left(d_t - \sum_{n=0}^{N} f_n x_{t-n} \right) x_{t-i} = 0 \tag{F.3}$$

or

$$\sum_t d_t x_{t-i} - \sum_t \sum_{n=0}^{N} f_n x_{t-n} x_{t-i} = 0. \tag{F.4}$$

Bringing one term to the other side

$$\sum_t \sum_{n=0}^{N} f_n x_{t-n} x_{t-i} = \sum_t d_t x_{t-i} \tag{F.5}$$

and interchanging the order of summation on the left-hand side gives:

$$\sum_{n=0}^{N} f_n \sum_t x_{t-n} x_{t-i} = \sum_t d_t x_{t-i}. \tag{F.6}$$

Now substituting $s = t - i$, we obtain:

$$\sum_{n=0}^{N} f_n \sum_{s} x_{s+(i-n)} x_s = \sum_{s} d_{s+i} x_s \tag{F.7}$$

and we recognize the auto- and cross-correlation function on the left- and right-hand side, respectively:

$$\sum_{n=0}^{N} f_n \phi_{xx}[i-n] = \phi_{dx}[i] \qquad \text{for } i = 0, 1, 2, \ldots, N, \tag{F.8}$$

where the correlation is denoted by ϕ (see also Chapter 2). Using the fact that the signal is real, so $\phi_{xx}[i] = \phi_{xx}[-i]$, we obtain the matrix system:

$$\begin{pmatrix} \phi_{xx}[0] & \phi_{xx}[1] & \phi_{xx}[2] & \cdots & \phi_{xx}[N] \\ \phi_{xx}[1] & \phi_{xx}[0] & \phi_{xx}[1] & \cdots & \phi_{xx}[N-1] \\ \phi_{xx}[2] & \phi_{xx}[1] & \phi_{xx}[0] & \cdots & \phi_{xx}[N-2] \\ \vdots & & & & \vdots \\ \phi_{xx}[N] & \phi_{xx}[N-1] & \phi_{xx}[N-2] & \cdots & \phi_{xx}[0] \end{pmatrix} \begin{pmatrix} f[0] \\ f[1] \\ f[2] \\ \vdots \\ f[N] \end{pmatrix} = \begin{pmatrix} \phi_{dx}[0] \\ \phi_{dx}[1] \\ \phi_{dx}[2] \\ \vdots \\ \phi_{dx}[N] \end{pmatrix}. \tag{F.9}$$

This is completely equivalent to equation (4.34) in Chapter 4: the matrix has the well-known Toeplitz structure.

Analytic example

Let us take a simple example, namely the wavelet (x_0, x_1) as input signal. We wish to get the wavelet (d_0, d_1, d_2). We want to get a filter with length 2, i.e. (f_0, f_1). The actual output is:

$$\begin{aligned} (a_0, a_1, a_2) &= (f_0, f_1) * (x_0, x_1) \\ &= (f_0 x_0, f_0 x_1 + f_1 x_0, f_1 x_1) \end{aligned} \tag{F.10}$$

We want to determine the filter coefficients by minimizing its error energy:

$$\begin{aligned} E &= \sum_{t=0}^{2} (d_t - a_t)^2 \\ &= (d_0 - f_0 x_0)^2 + (d_1 - f_0 x_1 - f_1 x_0)^2 + (d_2 - f_1 x_1)^2 \end{aligned} \tag{F.11}$$

Next we set each partial derivative with respect to the filter coefficients zero, i.e. for filter coefficient f_0:

$$\frac{\partial E}{\partial f_0} = 0 : 2(d_0 - f_0 x_0) \cdot (-x_0) + 2(d_1 - f_0 x_1 - f_1 x_0) \cdot (-x_1) = 0 \tag{F.12}$$

or

$$f_0(x_0^2 + x_1^2) + f_1 x_0 x_1 = d_0 x_0 + d_1 x_1. \tag{F.13}$$

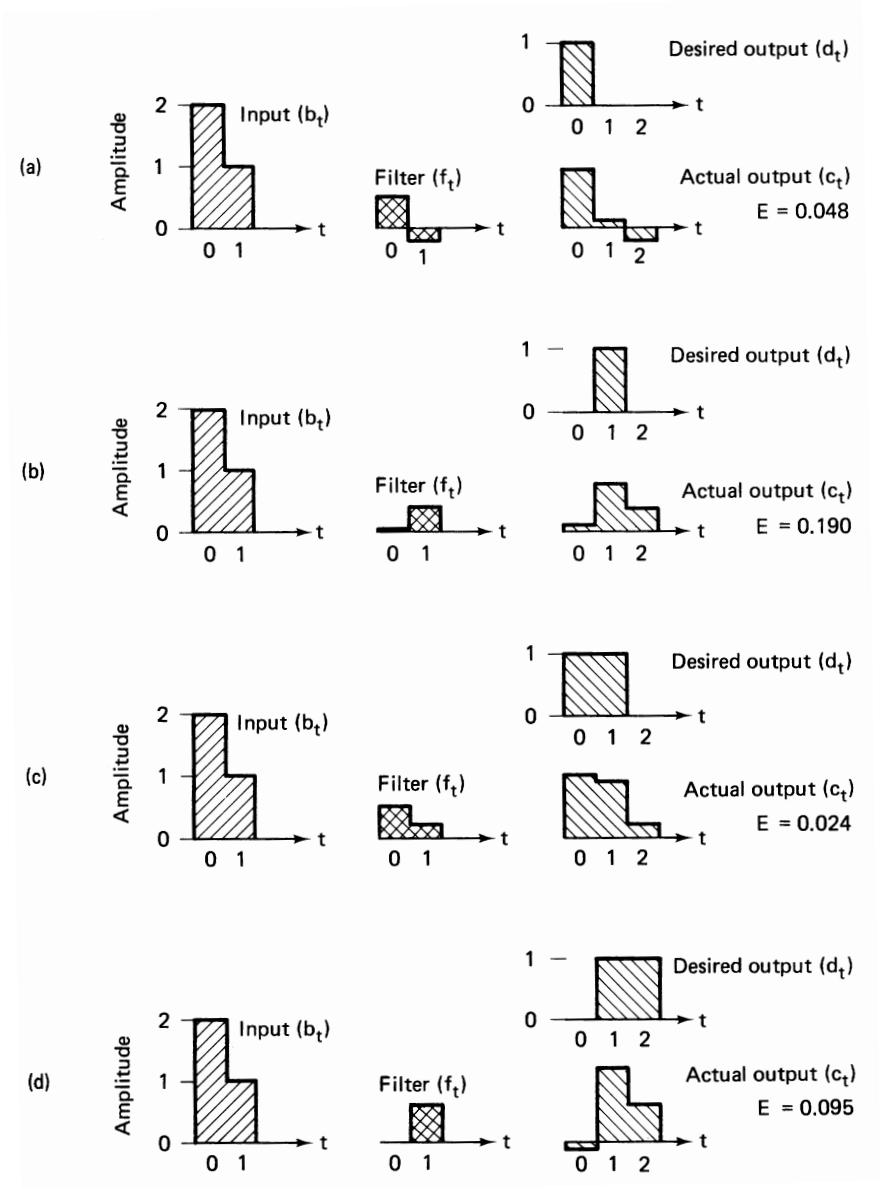


Figure F.1: Some simple examples of Wiener filters (from Robinson and Treitel 1980).

The same for filter coefficient f_1 :

$$\frac{\partial E}{\partial f_1} = 0 : 2(d_1 - f_0 x_1 - f_1 x_0) \cdot (-x_0) + 2(d_2 - f_1 x_1) \cdot (-x_1) = 0 \qquad \text{(F.14)}$$

or

$$f_0 x_0 x_1 + f_1 (x_0^2 + x_1^2) = d_1 x_0 + d_2 x_1. \qquad \text{(F.15)}$$

Combining the two equations yields:

$$\phi_{xx}[0] f_0 + \phi_{xx}[1] f_1 = \phi_{dx}[0]$$

$$\phi_{xx}[1]f_0 + \phi_{xx}[0]f_1 = \phi_{dx}[1]. \tag{F.16}$$

Now let x be (2,1), and d be (1,0,0). Then we can solve for f :

$$5f_0 + 2f_1 = 2$$

$$2f_0 + 5f_1 = 2. \tag{F.17}$$

It follows then that $f = (10/21, -4/21)$. For this solution the actual output is:

$$a_t = (2,1) * (10/21, -4/21) = (20/21, 2/21, -4/21). \tag{F.18}$$

The error energy between the desired output $d = (1,0,0)$ and the actual output is:

$$E_{min} = \left(1 - \frac{20}{21}\right)^2 + \left(0 - \frac{2}{21}\right)^2 + \left(0 - \left(-\frac{4}{21}\right)\right)^2 = \frac{1}{21} \approx 0.048. \tag{F.19}$$

Of course, if we would allow f_t to be 3 filter points long, a zero energy could be achieved: $f_t = (1/2, -1/2, 1/2)$. It is convenient to normalize E such that it lies between 0 and 1. We can achieve this by normalizing with the power of the signal ($\phi_{dd}[0]$):

$$\bar{E}_{min} = \frac{E_{min}}{\phi_{dd}[0]} = \frac{\frac{1}{21}}{1} = \frac{1}{21}. \tag{F.20}$$

Some more examples with desired outputs for this wavelet and filter lengths are found in figure (F.1).

Damped least-squares solution

Sometimes we need an extra constraint in order to prevent the filter coefficients f_i to become unstable. This is then achieved by also minimizing the energy of the filter coefficients. Thus, the total energy to be minimized is:

$$E = \sum_t (d_t - f_t * x_t)^2 + \epsilon \sum_t f_i^2 = minimum, \tag{F.21}$$

in which the stabilization parameter ϵ defines to what extend the energy of the filter coefficients need be involved. In a similar way to the unstabilized problem, the derivatives of the energy with respect to the filter coefficients is put to zero, which yields the following set of normal equations:

$$\sum_{n=0}^{N} f_n \phi_{xx}[i-n] + \epsilon f_i = \phi_{dx}[i] \qquad \text{for } i = 0, 1, 2, \ldots, N. \tag{F.22}$$

Again this can be written as a matrix-vector system:

$$\begin{pmatrix} \phi_{xx}[0]+\epsilon & \phi_{xx}[1] & \phi_{xx}[2] & \cdots & \phi_{xx}[N] \\ \phi_{xx}[1] & \phi_{xx}[0]+\epsilon & \phi_{xx}[1] & \cdots & \phi_{xx}[N-1] \\ \phi_{xx}[2] & \phi_{xx}[1] & \phi_{xx}[0]+\epsilon & \cdots & \phi_{xx}[N-2] \\ \vdots & & & & \vdots \\ \phi_{xx}[N] & \phi_{xx}[N-1] & \phi_{xx}[N-2] & \cdots & \phi_{xx}[0]+\epsilon \end{pmatrix} \begin{pmatrix} f[0] \\ f[1] \\ f[2] \\ \vdots \\ f[N] \end{pmatrix} = \begin{pmatrix} \phi_{dx}[0] \\ \phi_{dx}[1] \\ \phi_{dx}[2] \\ \vdots \\ \phi_{dx}[N] \end{pmatrix}. \tag{F.23}$$

As such, stabilization of a linear system of equations is obtained by adding a stabilization factor to the main diagonal of the autocorrelation matrix. Note that if the stabilization factor ϵ is taken very large, the estimated filter is just a scaled version of the cross-correlation ϕ_{dx} of the input and the desired signal.

Appendix G

Derivation of the DMO-ellipse

In this appendix we will derive the DMO-ellipse, which is a the common-reflection point time t_I as a function of the coordinate x_I of point I. The derivation is quite elaborate in the sense that many geometrical quantities have to be determined.

Let us start with figure (G.1). When a ray reflects at a dipping reflector, it does not reflect at the point H' of the line perpendicular to the interface going through the point H, half-way between the source and geophone. Because of the dip α of the reflector, the reflection point is slightly shifted, as we also showed in the main text as the reflection smear. We are first going to determine this shift.

In figure (G.1) we have drawn a ray which is reflected at an interface which has a dip α. The source S is situated at the origin (for convenience), the point H is at the surface at half-offset x_h and the receiver R is at the surface at distance $2x_h$. The equation for the interface, and the line perpendicular to the interface going through H are respectively given by:

$$z = -x\tan\alpha + \frac{d_H}{\cos\alpha} + x_h\tan\alpha \tag{G.1}$$

$$z = \frac{x}{\tan\alpha} - \frac{x_h}{\tan\alpha} \tag{G.2}$$

in which d_H is the distance from H to H'. The intersection point of these two lines gives the coordinates of H':

$$x_{H'} = x_h + d_H\sin\alpha$$

$$z_{H'} = d_H\cos\alpha$$

In order to determine the equation of the ray, we put an image source at S^* to obtain:

$$z = \frac{\cos\alpha(d_H + x_h\sin\alpha)}{d_H\sin\alpha - x_h\cos^2\alpha}x - \frac{2x_h(d_H + x_h\sin\alpha)\cos\alpha}{d_H\sin\alpha - x_h\cos^2\alpha} \tag{G.3}$$

The intersection point I' between the ray and the interface is given by:

$$x_{I'} = x_h + d_H\sin\alpha + \frac{x_h^2}{d_H}\sin\alpha\cos^2\alpha \tag{G.4}$$

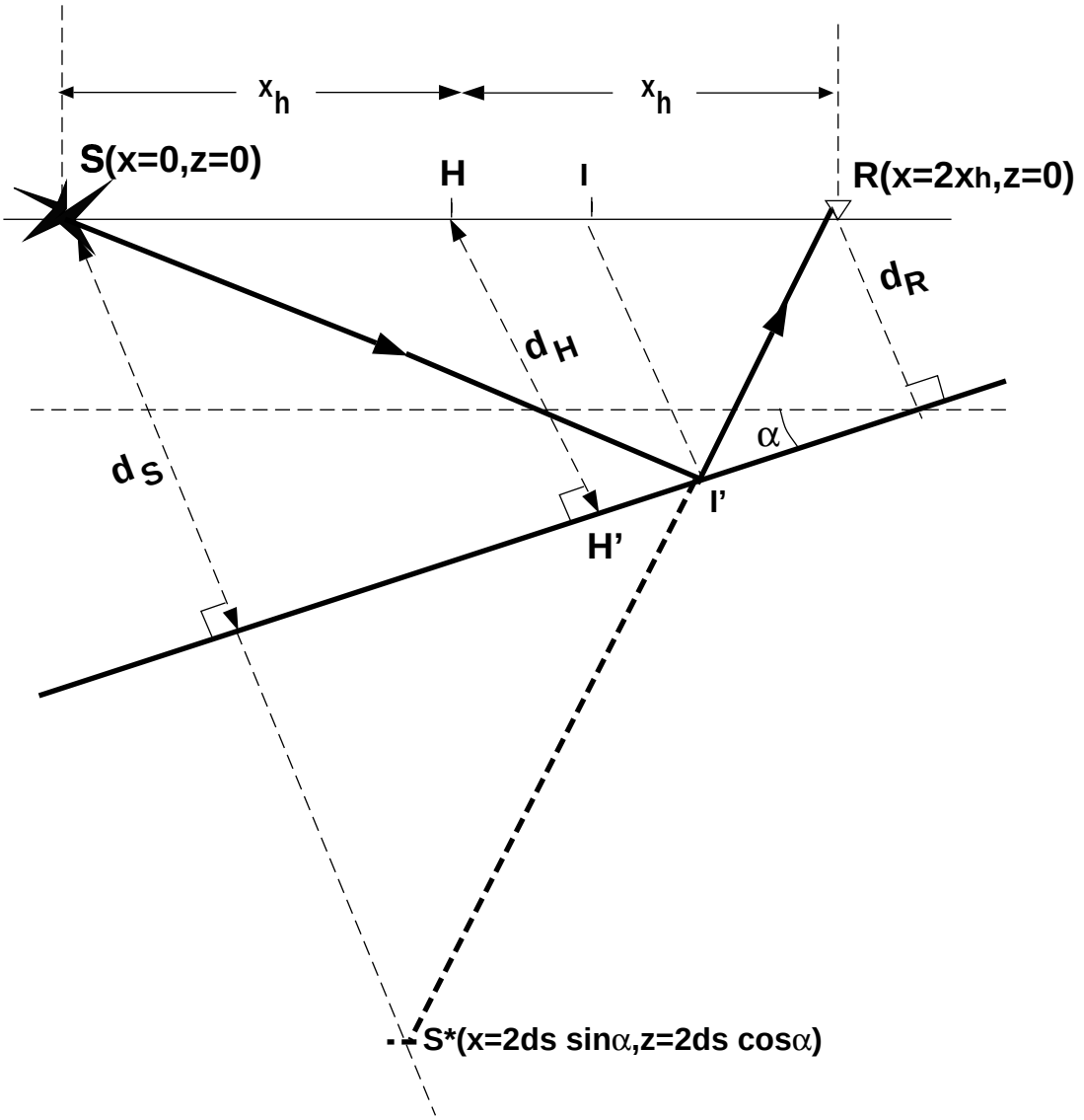


Figure G.1: The model for deriving the DMO ellipse.

$$z_{I'} \quad = \quad d_H \cos\alpha - \frac{x_h^2}{d_H} \sin^2\alpha \cos\alpha \tag{G.5}$$

The reflection-point smear can now be evaluated by the distance between the points I' and H', being $(x_h^2/d_H)\sin\alpha\cos\alpha$. Using the traveltime t_H rather than d_H ($t_H = 2d_H/c$), gives as final result:

$$x_I - x_H = \frac{x_{I'} - x_{H'}}{\cos\alpha} = \frac{x_h^2}{\frac{c}{2}t_H} \sin\alpha \tag{G.6}$$

where $x_H = x_h$.

The next equation follows straightforwardly from figure (G.1) by looking at the difference in distance travelled perpendicular to the interface, i.e., $t_I = t_{II'I}$ and $t_H = t_{HH'H}$, which is $(x_I - x_h)\sin\alpha$:

$$t_I = t_H - \frac{x_I - x_h}{c/2} \sin\alpha \tag{G.7}$$

Finally we need the equation as derived in the main text:

$$t_{DMO}^2 = t_H^2 - \frac{4x_h^2}{c^2}\sin^2\alpha \tag{G.8}$$

We are now going to combine the last three equations in order to obtain an expression of t_I as a function of x_I. To that purpose we write equation (G.6) as

$$\frac{\sin\alpha}{c/2} = \frac{x_I - x_h}{x_h^2}t_H \tag{G.9}$$

and substitute this in the other two equations ((G.7) and (G.8)):

$$t_I = t_H\left(1 - \frac{(x_I - x_h)^2}{x_h^2}\right) \tag{G.10}$$

$$t_H = t_{DMO}\left(1 - \frac{(x_I - x_h)^2}{x_h^2}\right)^{-1/2} \tag{G.11}$$

The last step is now to combine these two equations in order to get an expression in which t_I is a function of t_{DMO} rather than t_H. We then obtain the final result:

$$t_I = t_{DMO}\left(1 - \frac{(x_I - x_h)^2}{x_h^2}\right)^{1/2} \tag{G.12}$$

This is the equation of an ellipse in the (t_I, x_I) domain.

Appendix H

Derivation of the Kirchhoff integral

The classical paper explaining the integral method as applied to seismic migration, i.e. Kirchhoff migration, is the one by [Schneider, 1978].

From wave equation to Kircchoff integral

Let us start with the wave equation for the pressure $p = p(\mathbf{x}) = p(x, y, z, t)$:

$$\nabla^2 p - \frac{1}{c^2}\partial_t^2 p = 0, \tag{H.1}$$

which is the wave equation for a source-free 3 dimensional homogeneous medium with propagation velocity c. This equation combines Newton's second law, and Hooke's law, i. e. the conservation of mass for a homogeneous, isotropic medium with wave velocity c. This derivation is given in appendix B, which is taken from [Berkhout, 1984]. When we want to invoke Huygens' principle, any wave field can be thought to be the cumulative effect of an infinite number of point sources. The solution of the wave equation for a point source is called a Green's function. Thus a Green's function G is the solution of the equation:

$$\nabla^2 G - \frac{1}{c^2}\partial_t^2 G = -\delta(\mathbf{x} - \mathbf{x}^s)\delta(t - t^s), \tag{H.2}$$

in which the superscript s denotes that it pertains to the source, i.e. the point source at position $\mathbf{x}^s$ "explodes" at time t^s.

In the following, we are going to combine the latter two equations, thereby invoking Huygens' principle: the Green's function will act as Huygens source that will bring the measurements from one depth level to the other. When we multiply equation (H.2) with the pressure p, multiply equation (4.99) by G, and subtract the two equations from each other, we obtain:

$$p\delta(\mathbf{x} - \mathbf{x}^s)\delta(t - t^s) = G\nabla^2 p - p\nabla^2 G - \frac{1}{c^2}(G\partial_t^2 p - p\partial_t^2 G). \tag{H.3}$$

We can rewrite this as:

$$p\delta(\mathbf{x} - \mathbf{x}^s)\delta(t - t^s) = \nabla \cdot (G\nabla p - p\nabla G) - \frac{1}{c^2}\partial_t(G\partial_t p - p\partial_t G). \tag{H.4}$$

Now we integrate over some volume D and over all time to obtain:

$$\chi(\mathbf{x})p(\mathbf{x},t) = \int_{-\infty}^{\infty}\int_{D} \nabla\cdot(G\nabla p - p\nabla G)dV^s dt^s - \frac{1}{c^2}\int_{-\infty}^{\infty}\int_{D}\partial_t(G\partial_t p - p\partial_t G)dV^s dt^s, \tag{H.5}$$

in which the function $\chi(\mathbf{x})$ is defined as:

$$\chi(\mathbf{x}) = \{1, \frac{1}{2}, 0\} \qquad \text{when } x\epsilon\{D, \partial D, D'\}. \tag{H.6}$$

The term on the left-hand side occurs because of integrating out the δ–function, and it takes the values 1, 1/2 and 0 depending on whether the argument of the δ–function becomes zero within D, at the boundary of D, called ∂D, or outside D, called D'. This integration means physically that we assume there are Green's functions at all positions within volume D and also at all different times t^s. So this integration describes the total effect of all these Green's functions.
Let us look at the second integral in equation (H.5). This integral can be integrated by parts with respect to time giving a term of the form:

$$\left[\int_D (G\partial_t p - p\partial_t G)dV^s\right]_{t^s=-\infty}^{\infty}. \tag{H.7}$$

Since the pressure and its time derivative are zero before the Green's sources are fired, the contribution of the integral at $-\infty$ is zero. At $+\infty$ we assume the radiation condition (which states that the wave fields die out rapid enough such that integration over the volume D is still negligible), so there the contribution of the integral is zero too. We are now only left with the first integral in equation (H.5). Applying Gauss' theorem to the result, we obtain the integral:

$$\chi(\mathbf{x})p(\mathbf{x},t) = \frac{1}{4\pi}\int_{-\infty}^{\infty}\int_{\partial D}(G\nabla p - p\nabla G)\cdot\mathbf{n}dA^s dt^s, \tag{H.8}$$

where $\mathbf{n}$ is the outward pointing unit normal on ∂D. Finally, to obtain a form more connected to the seismic imaging, we substitute the particle velocity $\mathbf{v}$ for the gradient of the pressure via the equation of motion (equation (B.19) in appendix B). Then, we obtain the Kirchhoff integral:

$$\chi(\mathbf{x})p(\mathbf{x},t) = \frac{1}{4\pi}\int_{-\infty}^{\infty}\int_{\partial D}(-\rho G\partial_t\mathbf{v} - p\nabla G)\cdot\mathbf{n}dA^s dt^s. \tag{H.9}$$

This equation expresses that if we know the pressure p and the time derivative of the normal component of the particle velocity on a closed surface, the pressure can be computed in every point inside D. Also, we recognize that the pressure at a certain position is synthesized by means of a monopole (i.e. G) and dipole (i.e. $\nabla G\cdot\mathbf{n}$) distribution on a closed surface ∂D. The propagation of the secondary sources at the boundary ∂D to the observation point $(\mathbf{x},t)$ is described by the Green's function G. The same kind of expression can be derived for the particle velocity, see [Berkhout, 1984], chapter 5, from which figure H.1 has been drawn.

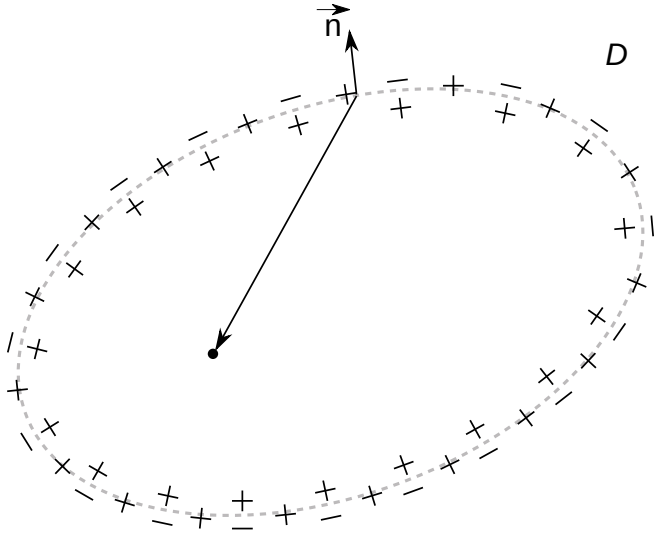


Figure H.1: A pressure field can be synthesized from the wave fields of a monopole and dipole distribution on a closed surface, using respectively the particle velocity and pressure of the actual wave field at this boundary as their source strengths (after [Berkhout, 1984], figure 5.1)

Using the Kirchhoff integral for migration

We will apply our results to seismic migration. To that end we consider data recorded at $z = 0$, which means that we take ∂D to be this surface, and assume that the contribution from an arbitrarily large hemisphere to the integral is zero. As a simple example, let us look at the case of a constant velocity medium, which means that there are no boundaries in the problem. We can first specify the Green's function, something we have left out so far in our discussion. Finding the solution for the free-space Green's function is very standard, and we leave the derivation as an exercise, or we find it in the literature (e.g. [Aki and Richards, 1980]):

$$G(\mathbf{x}, t; \mathbf{x}^s, t^s) = \frac{\delta(t - t^s \pm r/c)}{4\pi r}, \tag{H.10}$$

in which r is defined as:

$$r = \sqrt{(x - x^s)^2 + (y - y^s)^2 + (z - z^s)^2}. \tag{H.11}$$

For a constant velocity earth, we want to obtain two Green's functions such that they vanish on the surface $z = 0$. This is because we want to get rid of the first term in equation (H.9), such the expression becomes more simple. And we are free to choose our Green's function as long as they are solution of the wave equation. Therefore, we consider also an image source, with the surface $z_s = 0$ as the mirroring surface and with opposite amplitude:

$$G(\mathbf{x}, t; \mathbf{x}^s, t^s) = \frac{\delta(t - t^s \pm r/c)}{4\pi r} - \frac{\delta(t - t^s \pm r'/c)}{4\pi r'}, \tag{H.12}$$

in which r' is defined as:

$$r' = \sqrt{(x - x^s)^2 + (y - y^s)^2 + (z + z^s)^2}. \tag{H.13}$$

Note that there are two solutions each time, one with the $-$ and one with the $+$ sign in equation (H). The first is for *forward* and the second for *backward* (i.e. inverse) propagation. Since we aim to propagate the diffraction hyperbolae back to their origin in migration, we only use the anti-causal solution for the Green's function, i.e. the $+$ signs. We can plug this Green's function into Kirchhoff's integral. Making use of the fact that the Green's function is zero on the boundary ∂D the first term in brackets in Kirchhoff's integral (H.9) vanishes. Next, we have to apply the gradient of the Green's function which in our case is just the vertical derivative since we consider the surface $z^s = 0$. Then, ∇G becomes (notat that the positive z-axis is defined as pointing down):

$$\begin{aligned} \nabla G \cdot \mathbf{n} &= -\partial_{z^s} G \\ &= -2\partial_{z^s} \frac{\delta(t - t^s + r/c)}{4\pi r}, \end{aligned} \tag{H.14}$$

using the fact that $r'(z^s) = r(-z^s)$.
Since z appears only in the combination $z - z^s$, we can take $\partial_{z^s} = -\partial_z$ and take the derivative outside the integral:

$$p(\mathbf{x}, t) = \frac{1}{2\pi}\partial_z \int_{-\infty}^{\infty} \int_{z^s=0} (-p(\mathbf{x}^s, t^s) \frac{\delta(t - t^s + r/c)}{r}) dA^s dt^s. \tag{H.15}$$

Evaluating the integral with the δ–function, we obtain the result:

$$p(\mathbf{x}, t) = \frac{-1}{2\pi}\partial_z \int_{z^s=0} (\frac{p(\mathbf{x}^s, t + r/c)}{r}) dA^s. \tag{H.16}$$

This is Kirchhoff's migration formula, given by [Schneider, 1978]. We would like to stress that this result does not involve any approximations, the result is only dependent on the knowledge of the velocity distribution (i.e. vertical derivative of the pressure field) at the surface.
In fact it states that the wave field in any point in the subsurface $p(\mathbf{x}, t)$ can be calculated from the wave field $p(\mathbf{x}^s, t)$ recorded at a plane reference level z_s, assumed that we have a recording from $-\infty$ until $+\infty$ at the surface.
Note that if the term $t + r/c$ is replaced by $t - r/c$ the *inverse* propagation from the surface to point $(\mathbf{x}, t)$ becomes a *forward* extrapolation.

Bibliography

[Aki and Richards, 1980] Aki, K. and Richards, P. G. (1980). *Quantitative Seismology.* W. H. Freeman and Co.

[Barry et al., 1975] Barry, K. M., Cavers, D. A., and Kneale, C. W. (1975). Report on recommended standards for digital tape formats. *Geophysics*, 40(2):344–352.

[Berkhout, 1984] Berkhout, A. J. (1984). *Seismic migration, imaging of acoustic energy by wave field extrapolation, B: practical aspects.* Elsevier.

[Bracewell, 1978] Bracewell, R. N. (1978). *The Fourier transform and its applications.* McGraw-Hill Book Company.

[Claerbout, 1985] Claerbout, J. F. (1985). *Imaging the earth's interior.* Blackwell Scientific Publications.

[Deregowski, 1986] Deregowski, S. M. (1986). What is dmo? *First Break*, 4(7):7–24.

[Dix, 1955] Dix, C. H. (1955). Seismic velocities from surface measurements. *Geophysics*, 20(1):68–86.

[Gazdag, 1978] Gazdag, J. (1978). Wave equation migration with the phase shift method. *Geophysics*, 43:1342–1351.

[Gazdag and Sguazerro, 1984] Gazdag, J. and Sguazerro, P. (1984). Migration of seismic data by phase shift plus interpolation. *Geophysics*, 49(2):124–131.

[Hale, 1984] Hale, D. (1984). Dip-moveout by fourier transform. *Geophysics*, 49(6):741–757.

[Hubral, 1977] Hubral, P. (1977). Time migration - some ray theoretical aspects. *Geophys. Prosp.*, 25(4):738–745.

[Pieuchot, 1984] Pieuchot, M. (1984). *Seismic instrumentation (Handbook of Geophysical Exploration, section I:Seismic Exploration).* Geophysical Press, Amsterdam. ISBN 0-946631-02-6.

[Pullan, 1990] Pullan, S. E. (1990). Recommended standard for seismic (radar) files in the personal computer environment. *Geophysics*, 55(9):1260–1271.

[Robinson and Treitel, 1980] Robinson, E. A. and Treitel, S. (1980). *Geophysical signal analysis*. Prentice Hall, Inc.

[Schneider, 1978] Schneider, W. A. (1978). Integral formulation for migration in two and three dimensions. *Geophysics*, 43(1):49–76.

[SEG, 1980] SEG (1980). *Digital Tape Standards*. Society of Exploration Geophysicists.

[van der Schoot, 1989] van der Schoot, A. (1989). *Common reflection point stacking, a model driven approach to dip moveout*. PhD thesis, Delft University of Technology.

[Yilmaz, 1987] Yilmaz, O. (1987). *Seismic data processing*. Society of Exploration Geophysicists.

Contents

1 Introduction **1**

1.1 Exploration seismics 1

1.2 Structure of lecture notes 4

2 Basic signal analysis **5**

2.1 Introduction 5

2.2 The Fourier transform 5

2.3 The discrete Fourier transform 11

2.4 The spatial Fourier transform 12

2.5 The two-dimensional Fourier transform (in relation to waves) 13

3 A basic seismic processing sequence **18**

3.1 Seismic processing and imaging 18

3.2 Sorting of seismic data 20

3.3 Normal move-out and velocity analysis 25

3.4 Stacking 34

3.5 Zero-offset migration 38

4 Extended Processing **49**

4.1 Introduction 49

4.2 Data formats 50

4.3 Surveying information and field geometry 52

4.4 Trace editing and balancing 52

4.5 Static corrections 55

4.6 Deconvolution 62

4.6.1 Deterministic deconvolution in the frequency domain 62
4.6.2 Deterministic deconvolution in the time domain : Wiener filters . . . 66
4.6.3 Statistical deconvolution: minimum phase deconvolution 71
4.6.4 Statistical deconvolution: predictive deconvolution 74
4.6.5 Spiking deconvolution 79
4.7 Filtering in the (f, k_x) domain 81
4.8 Dip Move-Out (DMO) / Pre-stack Partial Migration 88
4.9 Zero offset (poststack) migration algorithms 96
4.10 Conversion from time to depth 113
4.11 Prestack migration 117

5 3D seismic processing **119**
5.1 Introduction 119
5.2 Midpoint oriented processing 119
5.3 3D Poststack migration 124

A Discretisation of Fourier transform **129**

B Derivation of the wave equation **135**

C The definition of SEG-Y **139**

D Traveltime equation for a dipping refracting boundary **141**

E Correlation of signals **144**

F Wiener filters **147**

G Derivation of the DMO-ellipse **152**

H Derivation of the Kirchhoff integral **155**