Basics of Seismic Imaging

Gerard T. Schuster
KAUST

To be published by Press (ISBN 9780521871242).

May 9, 2010

Contents

1 Overview **1**

2 Practical Migration **23**

3 Time Migration and Migration Resolution **43**

4 Modeling and Green's Functions **59**

5 Reverse Time Migration **85**

6 Phase Shift Methods **105**

Preface

This book describes the theory and practice of seismic imaging, with an emphasis on migration of seismic data in exploration seismology. It is written at the level where it can be understood by physical scientists who have some familiarity with the principles of wave propagation, Fourier transforms, and numerical analysis. The book can be taught as a one-semester course for advanced seniors and graduate students in the physical sciences and engineering. Exercises are given at the end of each chapter, and many chapters come with MATLAB codes that illustrate important ideas.

Seismic waves are vibrations that propagate throughout the earth, excited by sources such as earthquakes or large vibrator trucks used by oil companies. The associated particle velocities (or displacements) are recorded as seismograms (i.e., seismic traces in time) by land-based geophones, or in the marine environment the pressure variations generated by air guns are detected by marine hydrophones. The goal of seismic imaging is to use these seismograms to reconstruct the earth model, such as the reflectivity distribution or the velocity distribution. These distributions can be used to understand the subsurface geology and locate deposits of oil and gas.

The key principles of seismic imaging are heuristically described in Chapter 2, and by its end the diligent reader will be using MATLAB code to generate synthetic traces, redatum these data by summed cross-correlations, and invert the redatumed traces for the earth's reflectivity structure. Later chapters reinforce these principles by deriving the rigorous mathematics of seismic imaging. In particular, the governing equation of imaging is known as the Lippmann-Schwinger equation. Many examples are presented that migrate both synthetic data and field data. The terminology and examples mostly come from the applied geophysics community, but there are examples and chapters devoted to earthquake applications. The non-geophysicist will benefit by reading the brief overview of exploration seismology in the first chapter.

Acknowledgments

The author wishes to thank the long term support provided by the sponsors of the Utah Tomography and Modeling/Migration consortium. Their continued financial support through both lean and bountiful years was necessary in bringing this book to fruition.

Finally, many thanks go to the following students or postdocs who generously donated their results or MATLAB codes to this book: Chaiwoot Boonyasiriwat, Weiping Cao, Shuqian Dong, Sherif Hanafy, Ruiqing He, Zhiyong Jiang, Jianming Sheng, Yibo Wang, Yanwei Xue, Xiang Xiao, and Ge Zhan. Their diligent efforts have resulted in many interesting results, some of which are contained in this book.

Chapter 1

Overview

This chapter presents an overview of the basic goals and procedures for seismic exploration, and defines the seismic migration operation sometimes known as linearized seismic inversion. It is the main imaging tool for hydrocarbon mapping, and is increasingly being used for medical, earthquake, and engineering applications. We also review some fundamental mathematics associated with seismic imaging, such as the Fourier transform and convolution.

1.1 Exploration Seismology

The principal goal of exploration seismology is to map out oil and gas reservoirs by seismically imaging the earth's reflectivity distribution. Exploration geophysicists perform seismic experiments ideally equivalent to that shown in Figure 1.1, where the source excites seismic waves, and the resulting primary reflections are recorded by a geophone located at the source position. For this ideal zero-offset (ZO) experiment we assume only primary reflections in the records and that waves only travel in the vertical direction.

After recording at one location, the source and receiver are laterally moved by about 1/2 source wavelength and the experiment is iteratively repeated at different ground positions. All recorded traces are lined up next to one another and the resulting section is defined as the zero-offset (ZO) or poststack seismic section, as shown by the traces in Figure 1.1. This section resembles the actual geology, where one side of the signal is colored black to help enhance visual identification of the interface. Note that the depth z of the first reflector can be calculated by multiplying the 2-way reflection time t by half the P-wave velocity v of the first layer, i.e. $z = tv/2$.

Seismic images of the subsurface are used to understand the geology of the earth. For example, Figure 1.2 shows both optical and seismic pictures of faults[1]. These images provide an understanding of the fault's characteristics and so aid geologists in deciphering the tectonic forces that shaped the earth. Faults also serve as impermeable traps for oil and gas deposits, waiting to be found by the explorationist with the most capable seismic camera.

[1]Faults are high angle planar breaks in the subsurface rock. Rocks on one side of the fault slide apart from the rocks on the other side. If this sliding is mostly along the vertical direction then this is a normal fault as shown in Figure 1.2; if the sliding is along the horizontal direction then this is a transverse fault.

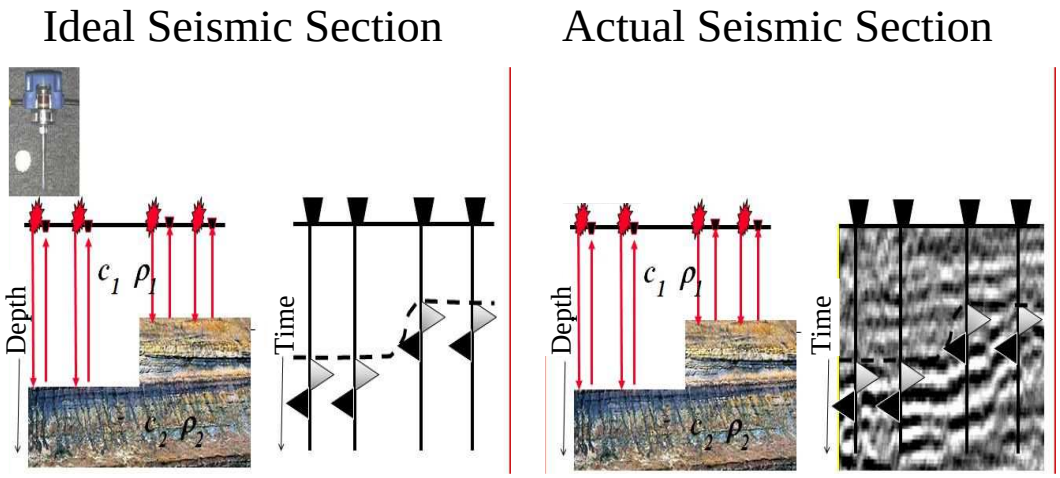


Figure 1.1: Earth model and idealized zero-offset (ZO) seismic section in time. Each trace is recorded by a geophone coincident with the source position; and the light (dark) colored amplitudes correspond to the particle velocity of the ground in the upward (downward) direction. The background image on the far right is an actual seismic section.

Other views of the earth using seismic data are shown in Figure 1.3.

1.1.1 Seismic sources

A land seismic source consists of a mechanical device or explosive located at $\mathbf{s}$ that thumps the earth (see Figure 1.4a) at time $t = 0$, and a geophone (see Figure 1.4b) at $\mathbf{g}$ records the time history of the earth's vertical particle velocity; the resulting seismogram is denoted as a seismic trace $d(\mathbf{g}, t|\mathbf{s}, 0)$. A marine source is usually an array of air guns. Larger amplitudes on the Figure 1.1 traces correspond to a faster particle velocity and the up-going (down-going) motion is denoted here by the unblackened (blackened) lobes. The lobe amplitude is roughly proportional to the reflectivity strength $m(\mathbf{x})$ of the corresponding reflector at $\mathbf{x} = (x, y, z)$. Assuming a constant density and a layered medium, the normal-incidence reflectivity model $m(\mathbf{x})$ is sometimes approximated as

$$m(\mathbf{x}) \quad = \quad \frac{v(z+dz) - v(z)}{v(z+dz) + v(z)} \approx \frac{1}{v(z)} \frac{dv(z)}{dz}, \tag{1.1}$$

where $v(z)$ is the P-wave propagation velocity at depth z and the depth interval $0.5dz$ is normalized to the value 1.

1.1.2 Non-zero offset seismic experiment

In practice, a ZO experiment cannot generate the ideal seismic section because the source also generates strong coherent noise and near-source scattering energy that drowns out the weak reflections. In addition, the waves are propagating in all directions and contain

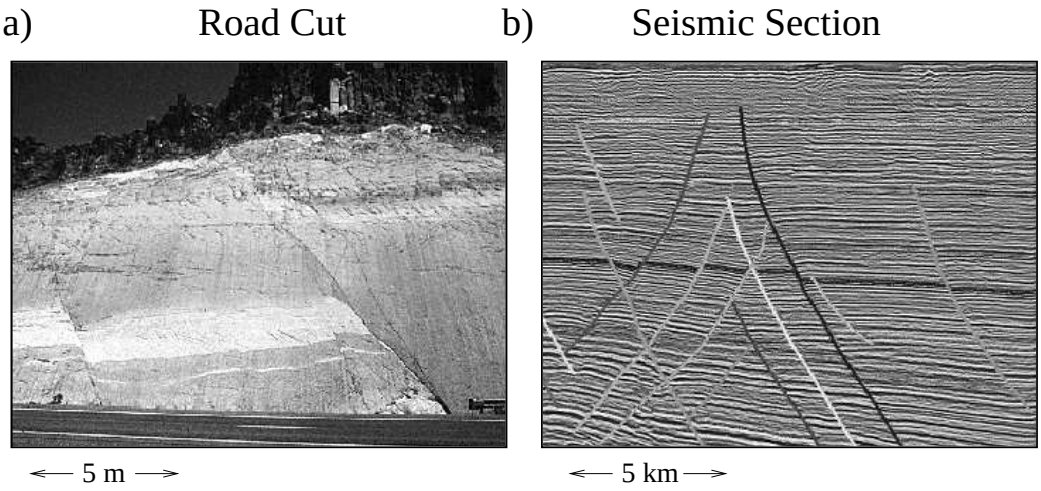


Figure 1.2: Geologic faults revealed by a) road cut and b) marine seismic section. The length scales above are roughly estimated.

distracting noise such as multiples, surface waves, scattered arrivals, out-of-the plane reflections, and converted waves. To account for these complexities, geophysicists perform non-zero offset experiments where the vibrations are recorded by many receivers as shown in Figure 1.5b. As before, each experiment consists of a shot at a different location except hundreds of active receivers are spread out over a long line for a 2D survey and a large area for a 3D survey[2].

1.1.3 Seismic processing

For surveys over a mostly layered medium, data processing consists of the following steps: 1). filtering of noise and near-surface statics corrections, 2). reassembly of common shot gather (CSG) traces in Figure 1.6a into common midpoint gathers (CMG) in Figure 1.6b where the source-receiver pair of each trace has the same midpoint location, 3). the traces in the CMG are time shifted to align the CMP reflections with the ZO reflection event in Figure 1.6c, 4) stack the traces in the time-shifted CMG to form a single trace (see Figure 1.6d) at the common midpoint position[3]. This stacked trace approximates a ZO trace with a high signal-to-noise at that position, and 5). repeat steps 3-4 for all midpoint gathers to give the seismic section shown in Figure 1.7. If the subsurface reflectivity is complex then steps 2-5 are skipped and instead the algorithm known as *seismic migration* is used, which is the subject of this book.

[2]Some surveys activate anywhere from 3,000 to 10,000 receivers per shot.

[3]The ensemble of traces from a surface seismic experiment is also known as surface seismic profile (SSP).

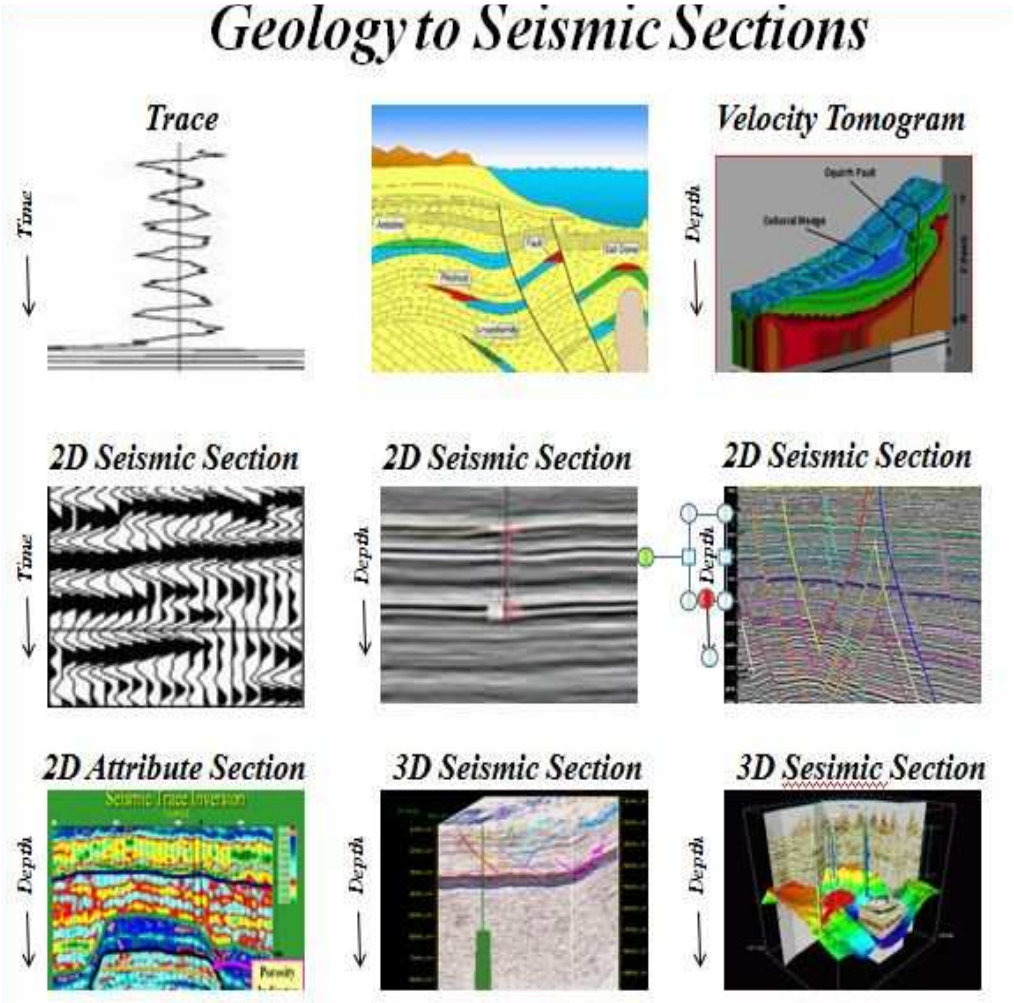


Figure 1.3: Different views of the earth extracted from seismic data. The single trace on the top left is the basic element of seismic recording, and after certain operations an ensemble of such traces can be transformed, e.g., into the 3D migration cube in the lower right or inverted to give the 3D velocity tomogram in the upper right.

a) Vibroseis Truck b) Geophones and Cables

Figure 1.4: a) Vibroseis truck and b) geophones attached to cables at a desert base camp. Inset is a particle velocity geophone about 12 cm long.

a) 2D SSP Land Survey b) 2D SSP Marine Survey

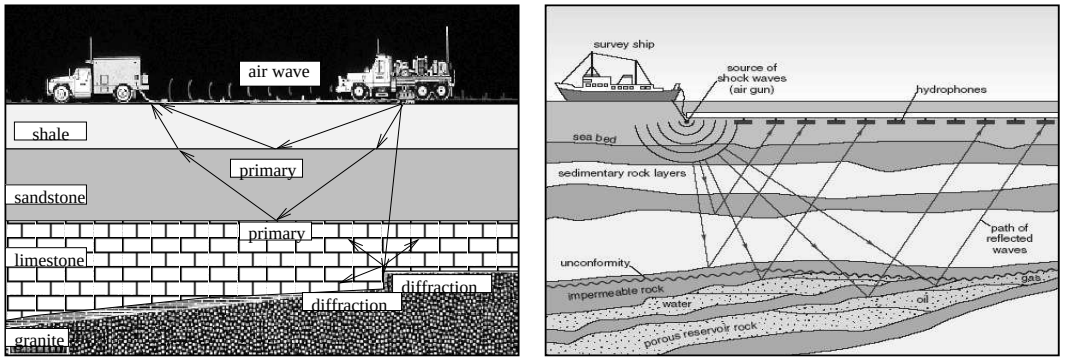


Figure 1.5: a) Land (courtesy of ConocoPhillips) and b) marine (courtesy of open-learn.open.ac.uk) survey geometries to record surface seismic profiles. The hydrophone cable for a marine survey can be as long as 12 km with a 30 m hydrophone spacing.

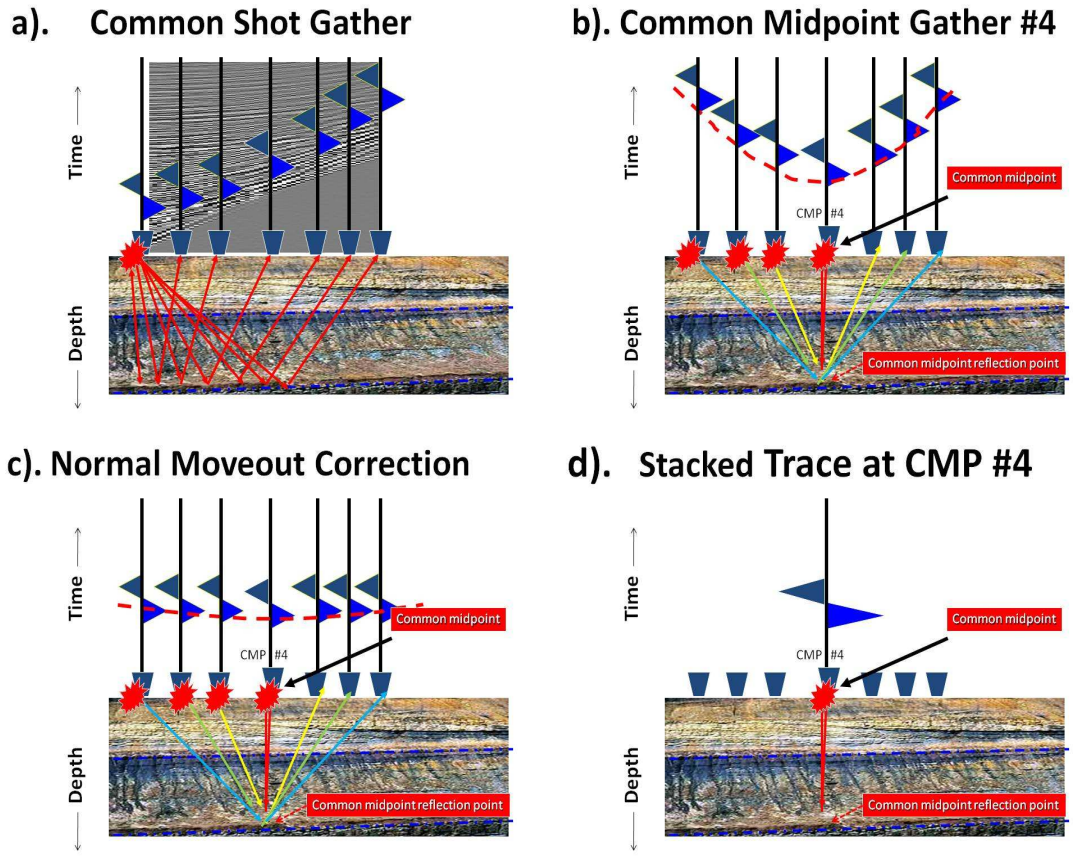


Figure 1.6: Different types of seismic trace formats. a). common shot gather where a single source is recorded by all of the geophones to form a CSG of traces. b). The traces from many CSGs can be reassembled to form a common midpoint gather (CMG) where a trace associated with each source-receiver pair has the same midpoint position on the surface. The common subsurface reflection point is known as the common reflection point; if the interface is dipping then rays associated with traces in the same CMG do not share a common reflection point. c). Traces after time shifting the CMP traces to align with the zero-offset trace; this time shifting is known as the normal moveout offset correction or NMO. d). The stacked trace is formed by adding the NMO corrected traces. In this case the fold is 7 and the signal/noise is enhanced by a factor of $\sqrt{7}$ if there is additive white noise in the CMG seismograms.

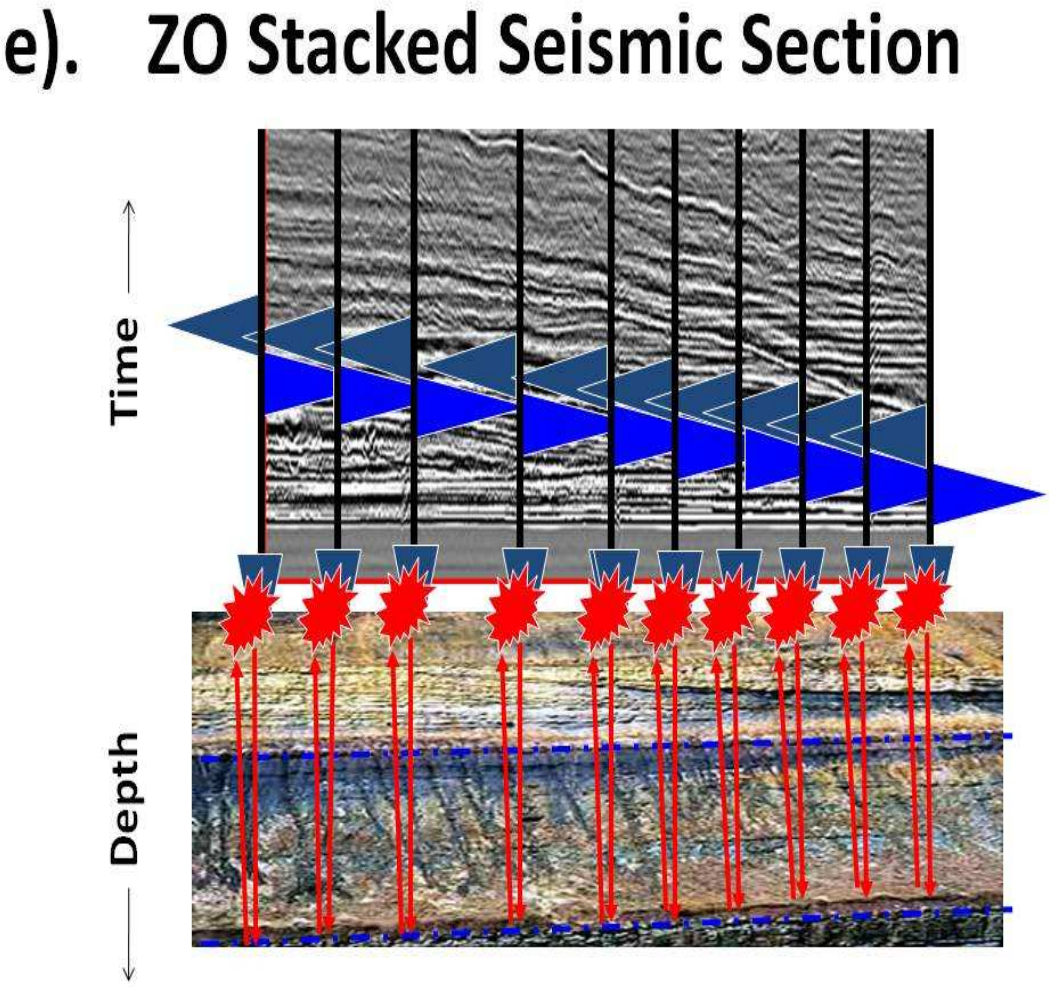


Figure 1.7: Diagram of stacked seismic section with background of an actual stacked seismic section.

1.2 Seismic Migration

For horizontal layers with homogeneous velocity, zero-offset seismic reflections will originate from reflection points directly beneath the geophone as shown in Figure 1.1. In this way the seismic section in time bears an accurate resemblance to the actual interface geometry in the subsurface If the subsurface model strongly violates the layer assumption, then interpretation of the seismic section will be inaccurate. The solution then is to apply migration to the data. Here, *migration is defined as the process which takes the seismic section* $d(x, y, t)$ *and moves the reflection events back to their origin at the interfaces.* Mathematically, migration maps the data $d(x, y, t)$ into the reflectivity distribution $m(x, y, z)$, denoted as the migration image.

Some departures from a layered model include the following examples.

1. **Dipping layers.** If the interfaces are dipping, as shown in Figure 1.8, then the dip of the seismic section can be noticeably different than the actual dip of the layer interface. Even if the time section were converted to depth z by a $z = v*t/2$ correction (where v is the homogeneous velocity of the medium and t is the 2-way time of the reflection event) the apparent dip in the seismic section would still be incorrect. The problem is that the reflection recorded in the trace is assumed to emanate from the interface point directly beneath the ZO geophone, which is an incorrect assumption if the reflector is dipping[4].

2. **Diffractions.** Another problem is that the subsurface interface might change its shape very rapidly with respect to the wavelength, and so strong diffraction energy can emanate from the distorted zone without having to honor Snell's law. An example is the faulted interface in Figure 1.9 that abruptly changes its tangent angle. This abrupt change in angle results in strong scattered energy that appears as a diffraction frown in the Figure 1.9 seismic section. This seismic section can then be incorrectly interpreted as indicating a false dome-like structure in the subsurface. *In this case migration moves the diffraction energy back to its origin at the diffractor point.* A synthetic example that clearly shows the uncollapsed diffraction frowns in the ZO section is shown in Figure 1.10, which are collapsed to points along the fault (dashed white line) in the migration image.

3. **Out-of-plane reflections.** Reflection events can originate outside the vertical plane coincident with the line of sources and receivers. These *out-of-plane reflections* will appear in the seismic section as coming from within the vertical plane, which can lead to falsely interpreted structures in the 2D seismic section. Hence, 3D migration must be applied to the data to map it into the correct reflectivity distribution.

4. **Conflicting dips.** If the geology is complicated then more than one type of reflection event may arrive at nearly the same time, but will be associated with different dip angles in $x - t$ space as shown in Figure 1.10. These dip angles differ from one

[4]According to the Snell's law, the transmission and reflection reflection rays must be parallel to one another for a ZO ray, and so they must be perpendicular to the interface at the specular reflection point. This means that, e.g. in a dipping layer model, the ZO green ray in Figure 1.8 is tilted to the left of the geophone rather than being vertical.

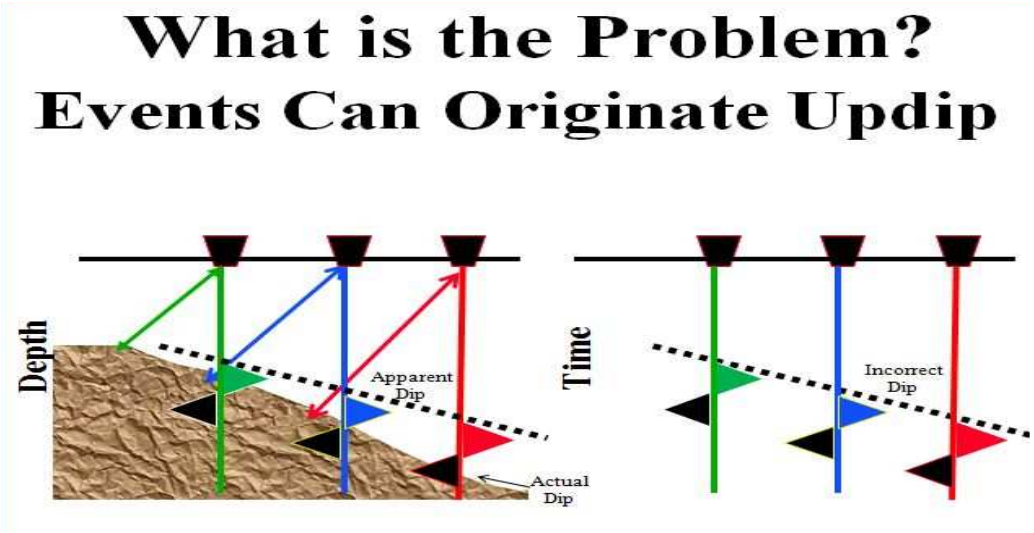


Figure 1.8: Zero-offset (ZO) reflections originate from reflection points updip from the recording geophone. Hence, the resulting seismic section appears to have a shallower dip than the actual seismic section. Two-sided arrows indicate raypaths for the ZO primary reflection events. Also, the migrated section will have a shorter reflector then the one interpreted in the seismic section.

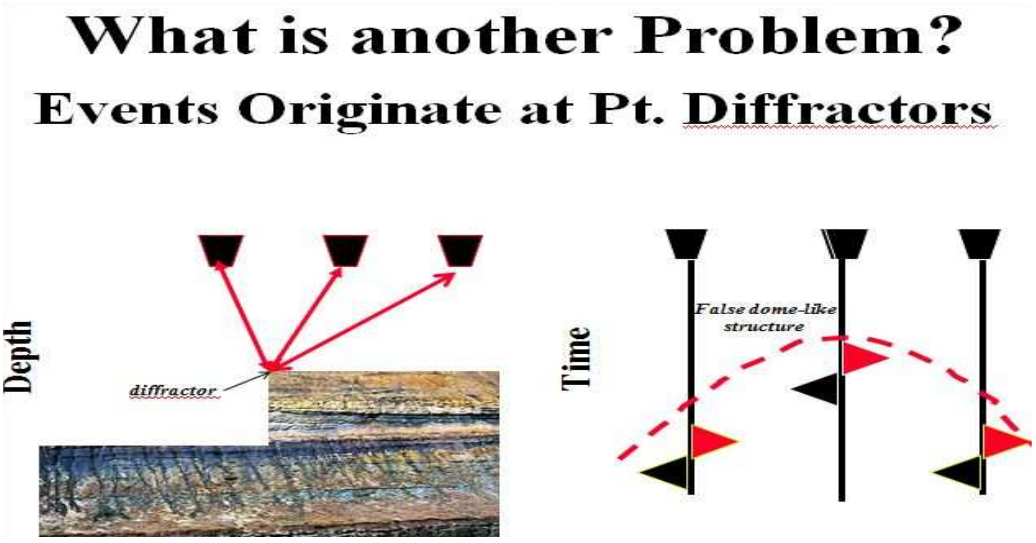


Figure 1.9: Reflectivity models with diffractors produce diffractions that appear as frowns in the seismic section, and can lead to falsely interpreted sections.

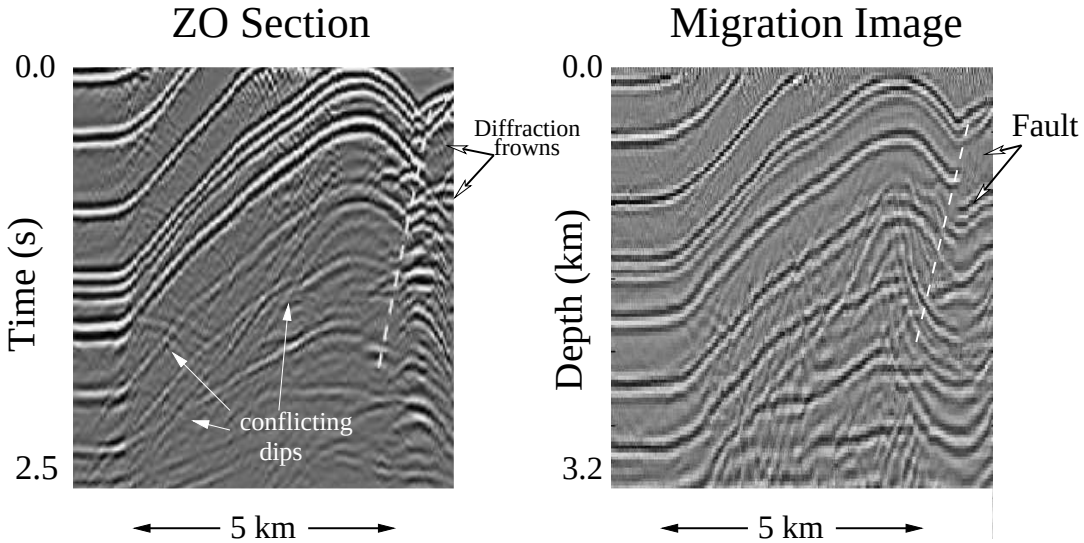


Figure 1.10: (Left) ZO section and (right) migration image of a faulted anticline model. Note, the fault denoted by dashed white lines is characterized by diffraction frowns in the ZO section; also, the dip of the right flank of the anticline is steeper in the migration image than in the ZO section. The conflicting dips seen in the ZO section are mostly eliminated in the migration image.

another because the associated reflections originate from different interfaces, each with a different dip angle in $x - y - z$ space.

Since stacking only assigns one moveout correction curve to each zero-offset time in a CMG, only type of the reflection events will be coherently stacked and the other will be inappropriately ignored. In this case, prestack migration should be used rather than conventional NMO and stacking.

1.2.1 Hand migration

A simple example of migrating traces by hand is described by the following three steps illustrated in Figures 1.11- 1.12.

1. In Figure 1.11a the lone reflection event in the single ZO trace is smeared back into the medium along a circle with radius equal to half the two-way reflectime time. Energy is also smeared within the annulus of width $T_0/2$ to account for the bandlimited source wavelet with period T_0. A point scatterer anywhere within this annulus could have accounted for some of the reflection energy in the single trace.

2. To determine more precisely the location of the reflector, a second trace is migrated (i.e., energy is smeared along the second fat circle in Figure 1.11b) and the intersection of these two annulus' pinpoints the scatterer location that can account for the events in both traces. Mathematically this intersection zone is delineated by summing the two traces and identifying the zone of strongest summed amplitudes.

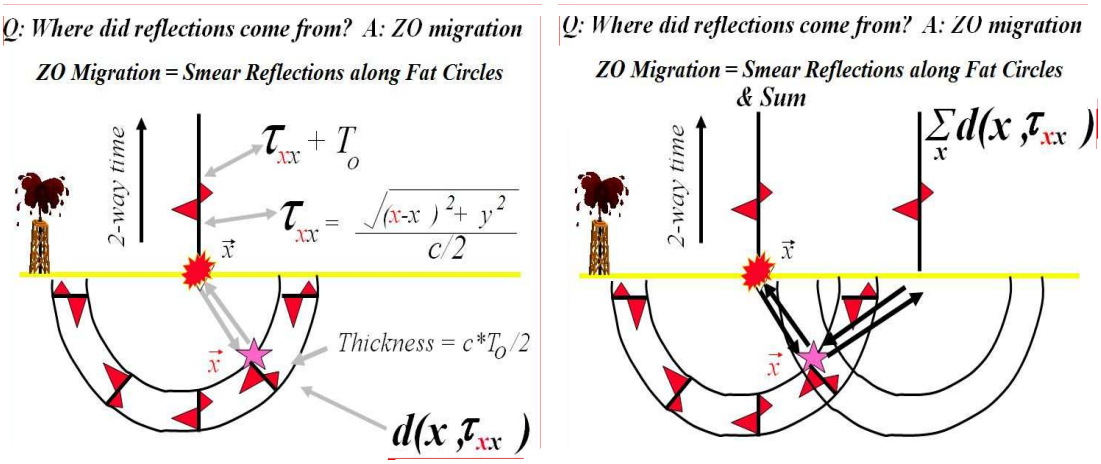


Figure 1.11: Migration images formed from (left) one and (right) two traces.

3. The previous step is repeated for many ZO traces to give the final poststack migration section shown in Figure 1.12.

1.2.2 Key problem with migration images

A *key difficulty in obtaining an accurate migration image is the estimation of a sufficiently accurate velocity model.* An inaccurate velocity model will lead to defocused and sometimes unusable migration images. In this case, the data must be be used to estimate not just the reflectivity model but also the migration velocity model. A common velocity model updating method is called traveltime tomography, which uses as input the traveltimes of either reflections or refractions in the data and inverts them for the velocity distribution as a function of (x, y, z). The upper right picture in Figure 1.3 depicts a velocity tomogram estimated by inverting refraction traveltimes[5]. A more expensive and complicated means for updating the velocity model is by full wavefield inversion, which is briefly discussed in the next section.

1.2.3 Algebraic description of migration

We are now ready to represent the seismic imaging operation in terms of matrix-vector algebra. The seismic trace (see upper left image in Figure 1.3) can be represented with the notation $d(\mathbf{g}, t|\mathbf{s}, 0)$, where $\mathbf{g}$ denotes the geophone location and $\mathbf{s}$ denotes the location of the source excited at time zero. Typically the trace might be about 10^3 samples long and so can be represented as a 10^3x1 vector. But many traces can be collected for thousands of different source and receiver positions. For example, an experiment with 10^4 shot positions, where 10^3 receivers record traces for each shot, gives rise to 10^7 traces, each with 10^3 samples. If these trace vectors are sequentially aligned on top of one other into a composite data vector then the input data can be denoted by the $10^{10}x1$ vector $\mathbf{d}$.

[5]Hotter colors in this case correspond to faster velocities.

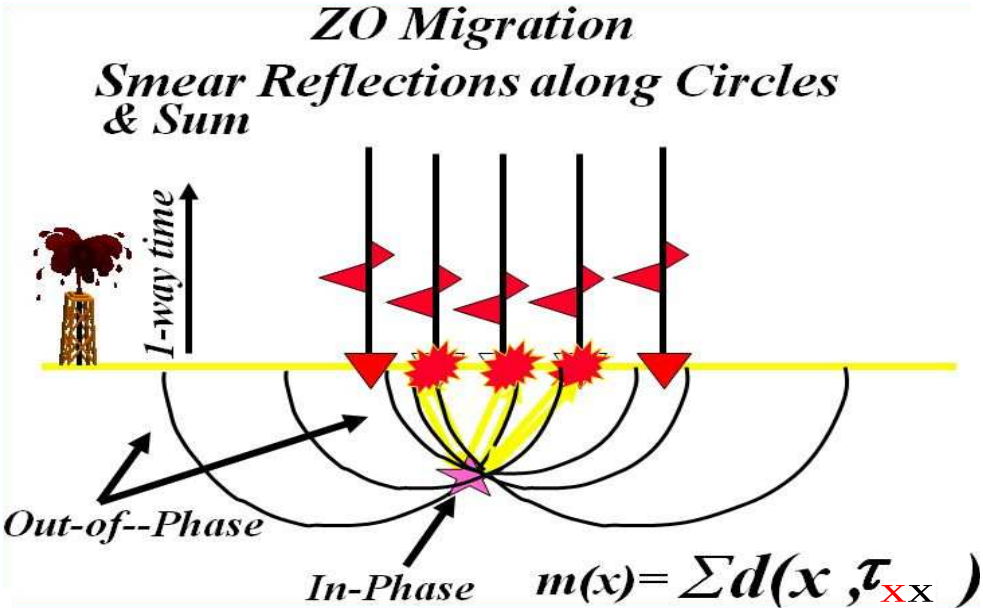


Figure 1.12: Migration image formed from many traces.

Similiarly, the reflectivity distribution can be approximated on a, for example, $10^3x10^3x10^3$ grid and reassembled as a 10^9x1 reflectivity vector $\mathbf{m}$. The mathematical operation $\mathbf{L}$ of forward modeling uses the input earth's reflectivity distribution as $\mathbf{m}$ to give the data $\mathbf{d}$, i.e.,

$$\mathbf{d} \quad = \quad \mathbf{L}(\mathbf{m}), \tag{1.2}$$

so in our example, $\mathbf{L}$ must be a $10^{10}x10^9$ matrix. Mother earth implicitly represents the forward modeling operation $\mathbf{L}$ as exact solutions to the anisotropic poro-viscoelastic wave equation, and so the actual seismic data can be very expensive to model in today's computers. To avoid this expense and only model the important primary reflections, the explorationist typically *assumes that the data are acoustic*, and so approximates $\mathbf{L}$ by solutions to the acoustic wave equation. This approximation was adequate for most processing in the 20th century, although taking into account anisotropic effects is becoming increasingly important if we wish to achieve more accurate images.

As we saw in the previous section, the seismic section $\mathbf{d}$ (e.g., ZO section in Figure 1.10) is not an accurate representation of the reflectivity distribution for complicated earth models. To estimate the actual reflectivity distribution, we migrate the reflection events to their place of origin along the interfaces. Mathematically, this is done by assuming a linearity approximation so equation 1.2 becomes $\mathbf{d} = \mathbf{Lm}$, and then solve it by finding the least squares solution:

$$\mathbf{m} \quad = \quad \overbrace{[\mathbf{L^T L}]^{-1}\mathbf{L^T d}}^{least\ squares\ migration}, \tag{1.3}$$

where $\mathbf{L^T}$ is the adjoint of the forward modeling operator. In this case we implicitly assume that we know how to compute the adjoint of the actual forward modeling operation of mother earth[6]. The above equation describes the *least squares migration* algorithm (Nemeth et al., 1998) and is carried out using a regularized conjugate gradient algorithm.

If the normal equation matrix $[\mathbf{L^T L}]$ is diagonally dominant, then the $[\mathbf{L^T L}]^{-1}$ can be approximated by a weighted diagonal matrix to get the *compensated illumination migration* equation:

$$\mathbf{m} \approx \overbrace{\mathbf{C L^T d}}^{migration}, \tag{1.4}$$

where the illumination compensation matrix $\mathbf{C}$ is a diagonal matrix with components $\mathbf{C}_{ij} = \delta_{ij}/[\mathbf{L^T L}]_{ii}$. Here $\mathbf{C}_{ij}$ compensates primarily for weakening of the signal due to geometric spreading in the reflected field and δ_{ij} is the Kronecker delta function that is equal to one if $i = j$, otherwise it is zero.

The above equations assume that $\mathbf{L}$ is independent of the reflectivity structure and only depends on the accurate smooth background velocity model; this means that the resulting forward modeling and inverse modeling equations are linear. If the background velocity is not accurately known and the reflectivity model contains large reflection coefficient values[7], then we should seek iterative updates to the velocity model. Such updates can be inexpensively obtained by traveltime tomography, or more accurately by the expensive means of waveform inversion. Due to the relentless increase in computational capabilities, it is now becoming possible to update the background velocity model in this way. The most common means for doing this is by an iterative non-linear gradient formula known as full wavefield inversion:

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \overbrace{[\mathbf{L}^T_{(k)}\mathbf{L}_{(k)}]^{-1}\mathbf{L}^T_{(k)}\mathbf{d}^{(k)}}^{full-wave\ inversion}, \tag{1.5}$$

where $\mathbf{d}^{(k)}$ represents the data residual[8] at the k^{th} iteration and $\mathbf{m}^{(k)}$ represents the velocity model at the k^{th} iteration. The modeling $\mathbf{L}_{(k)}$ and adjoint $\mathbf{L}^T_{(k)}$ operators depend on the iteration index because they are both updated with the new velocity model at each iteration. *Full wave inversion can be described as an iterative sequence of migrations, where the data residuals updated and migrated at each iteration to give the new model update. Moreover,* $\mathbf{L}$ *is updated at each iteration.*

Full wavefield inversion is a broad subject that will not be covered much in this book. Instead we we will mostly concentrate on showing the many ways of applying the adjoint approximation in equation 1.4 to seismic data, otherwise known as seismic migration. It is method that is used for the vast majority of seismic processing today, and so we will concentrate on its most popular implementations.

[6]We never know how to do this exactly because 1). it is too expensive and 2). we do not know the rock parameters with high precision.

[7]Reflection coefficients with absolute values much larger than 0.05 cause problems in the validity of the Born approximation, which will be discussed in later chapters.

[8]The data residual is computed by taking the difference between the observed seismic trace vector and the predicted trace vector. The predicted traces are obtained by forward modeling synthetic seismograms for the k^{th} velocity model.

1.3 Summary

This chapter describes the goals and key procedures of seismic exploration, where seismic data are collected and the goal is to invert it for the subsurface reflectivity and velocity distributions. In the early days of exploration (1960s) the ZO seismic section was mostly used to deduce the subsurface reflectivity structure, with the understanding that the interpreted geology became increasingly erroneous with increasing complexity in the subsurface geology. Some examples of complicated geology include dipping layers that produce conflicting dips, sharp changes in impedance that cause diffractions, and 3D geology that produces out-of-plane reflections. The remedy to these problems is to avoid NMO corrections and stacking, and directly map the primary reflections recorded in (x, y, t) data space onto their origin at the reflecting interfaces in (x, y, z) model space; this procedure is known today as migration and widely practiced since the 1960s. Migration should almost always be applied to prestack data. If the smooth background velocity model is well known, migration applies the adjoint operator $\mathbf{L}^T$ to the data $\mathbf{d}$ to get an estimate of the reflectivity distribution $\mathbf{m}$. This and accurate estimation of velocity models are the keystone tools for mapping oil and gas deposits by every oil company in the world today. The detailed study of seismic migration methods is the primary focus of this book.

1.4 Exercises

1. Appendix 1 defines the forward and inverse Fourier transforms as

$$F(\omega) \quad = \mathcal{F}[f(t)] = \quad \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)e^{i\omega t}dt, \tag{1.6}$$

$$f(t) \quad = \mathcal{F}^{-1}[F(\omega)] = \quad \int_{-\infty}^{\infty} F(\omega)e^{-i\omega t}d\omega. \tag{1.7}$$

   Show that $cos(\omega\tau) = [e^{i\omega\tau} + e^{-i\omega\tau}]/2$ has an inverse Fourier transform equal to $\pi\delta(t+\tau) + \pi\delta(t-\tau)$, where $\int e^{i\omega(t-\tau)}d\omega = 2\pi\delta(t-\tau)$.

2. The discrete convolution of the real N-point vectors $\mathbf{f} = [f[0]\ f[1]\ ...f[N-1]]$ and $\mathbf{g} = [g[0]\ g[1]...g[N-1]]$ is given by

$$h[i] \quad = \quad \sum_{i'=0}^{N-1} f[i-i']g[i'] = \sum_{i'=0}^{N-1} f[i']g[i-i']. \tag{1.8}$$

   The argument values refer to the time values of each element. Validate the above expression by computing the time series $h[t]$ for $\mathbf{f} = [1\ \ -2]$ and $\mathbf{g} = [-1\ 3]$.

3. The discrete correlation of two N-point vectors is given by

$$h[i] \quad = \quad \sum_{i'=0}^{N-1} f[i'-i]g[i'] = \sum_{i'=0}^{N-1} f[i']g[i+i'], \tag{1.9}$$

where the coefficients are assumed to be real. Compute the time series $h[t]$ for $\mathbf{f} = [1 \ -2]$ and $\mathbf{g} = [0 \ 1]$ and validate this equation (assume the elements in each vector are ordered from time zero for the 1st element and have increasing time indices for the other elements). Does the filter $\mathbf{g}$ shift $\mathbf{f}$ forward or backward in time?

4. For $f(t) = \delta(t-2) + \delta(t-4)$ and $g(t) = \delta(t-1)$, find $f(t) \otimes g(t)$ and $g(t) \otimes f(t)$. Is the cross-correlation operation commutative, i.e., is $f(t) \otimes g(t) = g(t) \otimes f(t)$?

5. Is convolution commutative? Mathematically prove your answer and test it in MATLAB.

```
f=[1 0 0 0 -2 0 0 0 0 4];
g=[0 0 -2 0 0 0 0 0 0 0];
h=conv(f,g);subplot(121);stem(h)
h=conv(g,f);subplot(122);stem(h)
```

Convolution of two M-length vectors produces a $2M-1$ length vector. Which is the zero-lag position in the MATLAB plot? The procedure for plotting the correct time axis labels is demonstrated in the next exercise.

6. The cross-correlation operation is equivalent to a reversed time convolution, i.e., $f(t) \otimes g(t) = f(-t) \star g(t)$ (see Appendix 1). The MATLAB program for cross-correlating two length M vectors $f(t)$ and $g(t)$ is

```
f=[1 0 0 0 -2 0 0 0 0 4];
g=[0 0 -2 0 0 0 0 0 0 0];M=length(g);
h=xcorr(f,g);
TMm1=length(h);
t=[1:TMm1];t=t-M;stem(t,h);
```

The result is a $2M-1$ length vector where the amplitude at lag zero is at the Mth element. Is this equal to $f(-t) \star g(t)$

```
h=conv(fliplr(f),g);subplot(121);stem(h)
```

or $g(-t) \star f(t)$?

```
h=conv(fliplr(g),f);subplot(122);stem(h)
```

Note that $fliplr(g)$ reverses the order of the vector g.

7. To create synthetic seismograms geophysicists often use a Ricker wavelet (Yilmaz, 2001) as their source wavelet. The formula in MATLAB script is given as

```
np=100;fr=20;dt=.001;
npt=np*dt;t=(-npt/2):dt:npt/2;
out=(1-t .*t * fr^2 *pi^2  ) .*exp(- t.^2 * pi^2 * fr^2 ) ;
plot(t,out);xlabel('Time (s)')
```

where (np, fr, dt) are equal to the number of samples, peak frequency (Hz), and time interval dt in seconds. The operation $t = (-npt/2) : dt : npt/2$ creates a vector of time units from time $-npt/2$ to $npt/2$, sampled at the time interval of dt. Plot out this wavelet using the above script except adjust the plotting code so the time units are in seconds, not samples. Repeat this exercise except choose (np, fr) so that a 5 Hz Ricker wavelet is plotted.

8. The Ricker wavelet from the previous question is acausal if there are non-zero amplitudes prior to $t = 0$. Implement a time shift to make it causal and plot it. One can create a subroutine by the following command

```
function [rick]=ricker(np,dt,fr)
% Computes acausal\index{acausal} Ricker wavelet\index{Ricker wavelet} with peak frequency fr
% sampled at dt with a total of np points. Make
% sure you choose np to be longer than T/dt , where
% T=1/fr.
npt=np*dt;t=(-npt/2):dt:npt/2;
rick=(1-t .*t * fr^2 *pi^2  ) .*exp(- t.^2 * pi^2 * fr^2 ) ;
z=rick(np/2:np);
%rick=rick*0;rick(1:np/2+1)=z;% Causal\index{causal} 1/2 Ricker
```

and typing $np = 100; dt = .002; fr = 20; rick = ricker(np, dt, fr)$ to create a vector $rick$ that represents a Ricker wavelet with a peak frequency of 20 Hz.

9. A two-layer velocity model consists of a 500 m thick layer with velocity v_1 =1 km/s and an underlying layer of velocity v_2 =2 km/s; here density is assigned a unit value everywhere and the top interface is a free surface. For a zero-offset acquisition geometry with source and receiver at **A** just below the free surface, the synthetic seismogram that contains only a primary reflection is given by $s(t) = r\delta(t - \tau_{AyA})$, where the reflection coefficient is $r = (v_2 - v_1)/(v_2 + v_1)$ and τ_{AyA} is the two-way normal incidence reflection time for a source at **A** and reflection point at **y**. For a sampling interval of 0.001 s, plot the impulse $I(t)$ response that only consists of primaries and upgoing waves. Use MATLAB to plot the response of a Ricker wavelet that only consists of primaries. That is, if the impulse response and Ricker wavelet are respectively defined in MATLAB by the vectors I and R then the Ricker response is given by

```
s=conv(I,R);plot(s);
```

10. A more elaborate modeling program for a 2-layer model that generates the primary reflection, and the 1st- and 2nd-order multiples is given by

```
function [seismo,ntime]=forward(v1,v2,dx,nx,d,dt,np,rick,x)
% (v1,v2,d) -input- velocity of 1st & 2nd layer of thickness d
% (dx,nx,dt)-input- (phone interval, # of phones, time interval)
% (np,rick) -input- (# of samples Ricker, Ricker wavelet)
%  x        -input- Nx1 vector of x values of  phone at z=0
%seismo(i,j,k)-output- Shot gather at ith src, jth phone, kth time
r=(v2-v1)/(v2+v1);r1=r*r;r2=r1*r;
for ixs=1:nx  % Loop over sources
  xs=(ixs-1)*dx;
  t=round(sqrt((xs-x).^2+(2*d)^2)/v1/dt)+1;  %Primary Time
  t1=round(sqrt((xs-x).^2+(4*d)^2)/v1/dt)+1; %1st Multiple Time
  t2=round(sqrt((xs-x).^2+(6*d)^2)/v1/dt)+1; %2nd Multiple Time
if ixs==1; ntime=max(t2)+np; seismo=zeros(nx,nx,ntime); end;
  s=zeros(nx,ntime);
 for i=1:nx; % Loop over  receivers
   s(i,round(t(i)))=r/t(i);   %Primary
   s(i,round(t1(i)))=-r1/t1(i); % 1st-order Multiple
   s(i,round(t2(i)))=r2/t2(i); % 2nd-order Multiple
   ss=conv(s(i,:),rick); % Convolve Ricker & Trace
   s(i,:)=ss(1:ntime);   % Synthetic Seismograms
   seismo(ixs,i,1:ntime)=ss(1:ntime);
 end
```

```
  c=seismo(ixs,:,:);c=reshape(c,nx,ntime);
  imagesc([1:nx]*dx,[1:ntime]*dt,c');xlabel('X(km)');ylabel('Time (s)')
  title('Shot Gather for Two-Layer Model'); pause(.1)
end
```

Run this program to generate synthetic seismograms for a two-layer model. Adjust the frequency content of the Ricker wavelet. If the Nyquist sampling criterion says that you need more than two samples per period, how should you adjust dt as you increase the wavelet frequency fr? How should you adjust dx as you decrease the velocity of the 1st layer $v1$, where wavelength is $v1/fr$?

11. Generate seismograms s that only contain primary reflections; convolve s with s to generate the 1st-order free-surface multiple. Does it have the correct arrival time for a 1st-order multiple? Does it have the correct magnitude for the reflection coefficient?

12. This is the same question as the previous one, except generate the 2nd-order free-surface multiple by two sequential convolutions of the primary trace.

13. The free-surface reflection coefficient for an incident pressure field is -1. What adjustment should you make to the previous convolutions in order to correctly model the polarity of the reflections?

14. In MATLAB, use $forward.m$ in problem 10 to generate the synthetic seismograms that contain the direct wave, primary and free-surface multiples up to the 2nd order. Assume a 20-Hz Ricker wavelet for the source time history and a two-later model. Write down the mathematical expression for this seismogram in terms of delta functions.

Appendix 1: Fourier Identities

The forward and inverse Fourier transforms are respectively given by (Bracewell, 2000)

$$F(\omega) \quad = \mathcal{F}[f(t)] = \quad \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)e^{i\omega t}dt, \tag{1.10}$$

$$f(t) \quad = \mathcal{F}^{-1}[F(\omega)] = \quad \int_{-\infty}^{\infty} F(\omega)e^{-i\omega t}d\omega, \tag{1.11}$$

where we adopt the following convention throughout the book: a lower case letter indicates a time- or space-domain function and its capitalized version indicates the Fourier transform. The following are Fourier identities, where the double-sided arrows indicate the functions are Fourier pairs and $\mathcal{F}$ indicates the forward Fourier transform.

1. **Differentiation:** $\partial^n/\partial t^n \leftrightarrow (-i\omega)^n$. This property is proved by differentiating equation 4.65 w/r to t.

2. **Convolution Theorem:** $f(t) * g(t) = \int f(\tau)g(t-\tau)d\tau \leftrightarrow 2\pi F(\omega)G(\omega)$. This property is proved by applying the Fourier transform to the convolution equation

$$\begin{aligned} \mathcal{F}[f * g] &= \mathcal{F}[\int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau], \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t}[\int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau]dt. \end{aligned} \tag{1.12}$$

Interchanging the order of integration we get

$$\mathcal{F}[f * g] = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(\tau)[\int_{-\infty}^{\infty} e^{i\omega t}g(t-\tau)dt]d\tau, \tag{1.13}$$

and defining the integration variable as $t' = t - \tau$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(\tau)[\int_{-\infty}^{\infty} e^{i\omega(t'+\tau)}g(t')dt']d\tau, \tag{1.14}$$

and using the definitions of the Fourier transform of $g(t)$ and $f(t)$ we get

$$= G(\omega)\int_{-\infty}^{\infty} f(\tau)e^{i\omega\tau}d\tau, \tag{1.15}$$

$$= 2\pi F(\omega)G(\omega). \tag{1.16}$$

We will often denote the convolution of two functions $f(t) * g(t)$ by the $*$ symbol.

3. **For real $f(t)$: $f(-t) \leftrightarrow F(\omega)^*$.** This property is easily proven by taking the complex conjugate of equation 4.65 to get $f(t)^* = f(t) = \int F(\omega)^* e^{i\omega t}d\omega$ and then apply the transform $t = -t'$.

4. **Correlation. $f(-t) * g(t) = f(t) \otimes g(t)$:** By definition $f(-t) * g(t) = \int f(-\tau)g(t-\tau)d\tau$. By changing the dummy integration variable $\tau \rightarrow -\tau'$ we get $f(-t) * g(t) = \int f(\tau')g(t+\tau')d\tau' = f(t) \otimes g(t)$, where $\otimes$ represents correlation. By identities 2 and 3 we conclude $\mathcal{F}[f(t) \otimes g(t)] = \mathcal{F}[f(-t) \star g(t)] = 2\pi F(\omega)^* G(\omega)$.

Appendix 2: Glossary

The following is a glossary of acronyms and terms commonly used in this book. A more detailed description of such terms can be found in Yilmaz (2001).

- AGC - Automatic gain control. An amplitude gain procedure applied to the trace that equalizes the trace energy over a contiguous sequence of specified time windows. After application of AGC, attenuation and geometrical spreading effects can be roughly corrected for and reflection amplitudes are normalized to be about the same value.

- Autocorrelation - $\phi(\tau)_{gg} = g(t) \otimes g(t) = \int_{-\infty}^{\infty} g(t+\tau)g(t)dt$. If $g(t)$ is a vector then the autocorrelation function $\phi(\tau)$ can be interpreted as the dot product of $g(t)$ with shifted copies of itself. Large positive values of $\phi(\tau)$ indicate a high degree of positive similarity between $g(t)$ and $g(\tau + t)$, large negative values indicate a high degree of negative similarity, and zero values mean no similarity.

- Cross-correlation - $\phi(\tau)_{gf} = g(t) \otimes f(t) = \int_{-\infty}^{\infty} g(t-\tau) f(t) dt = \int_{-\infty}^{\infty} g(t) f(t+\tau) dt = g(-t) \star f(t)$.

- CMG - Common midpoint gather. A collection of traces all having the same midpoint location between the source and geophone.

- COG - Common offset gather. A collection of traces all having the same offset displacement between the source and geophone.

- CRG - Common receiver gather. A collection of traces all recorded with the same geophone but generated by different shots.

- CSG - Common shot gather. Vibrations from a shot (e.g., an explosion, air gun, or vibroseis truck) are recorded by a number of geophones, and the collection of these traces is known as a CSG.

- Fold - The number of traces that are summed together to enhance coherent signal. For example, a common midpoint gather of N traces is time shifted to align the common reflection events with one another and the traces are stacked to give a single trace with fold N.

- IVSP data - Inverse vertical seismic profile data, where the sources are in the well and the receivers are on the surface. This is the opposite to the VSP geometry where the sources are on the surface and the receivers are in the well (see Figure 1.13). An IVSP trace will sometimes be referred to as a VSP trace or reverse vertical seismic profile (RVSP) seismogram.

- OBS survey - Ocean bottom seismic survey. Recording devices are placed along an areal grid on the ocean floor and record the seismic response of the earth for marine sources, such as air guns towed behind a boat. The OBS trace will be classified as a VSP-like trace.

- Reflection coefficient. A flat acoustic layer interface that separates two homogeneous isotropic media with densities ρ_1 and ρ_2 and compressional velocities v_1 and v_2 has the pressure reflection coefficient $(\rho_2 v_2 - \rho_1 v_1)/(\rho_2 v_2 + \rho_1 v_1)$. This assumes that the source plane wave is normally incident on the interface from the medium indexed by the number 1.

- RTM - Reverse Time Migration. A migration method where the reflection traces are reversed in time as the source-time history at each geophone. These geophones now act as sources of seismic energy and the fields are backpropagated into the medium (Yilmaz, 2001).

- PDE - Partial differential equation.

- Stacking - Stacking traces together is equivalent to summation of traces. This is usually done with traces in a common midpoint gather after aligning events from a common reflection point.

- S/N - Signal-to-noise ratio. There are many practical ways to compute the S/N ratio. Gerstoft et al. (2006) estimates the S/N of seismic traces by taking the strongest amplitude of a coherent event and divides it by the standard deviation of a long noise segment in the trace.
- SSF - Split step Fourier migration. A migration method performed in the frequency, depth, and spatial wavenumber domains along the lateral coordinates (Yilmaz, 2001).
- SSP data - Surface seismic profile data. Data collected by locating both shots and receivers on or near the free surface (see Figure 1.13).
- SWD data - Seismic-while-drilling (SWD) data. Passive traces recorded by receivers on the free surface with the source as a moving drill bit. Drillers desire knowledge about the rock environment ahead of the bit, so they sometimes record the vibrations that are excited by the drill bit. These records can be used to estimate the subsurface properties, such as reflectivity (Poletto and Miranda, 2004).
- SWP data - Single well profile data with the shooting geometry shown in Figure 1.13. Data are collected by placing both shots and receivers along a well.
- VSP data - Vertical seismic profile data. Data collected by firing shots at or near the free surface and recorded by receivers in a nearby well. The well can be either vertical, deviated, or horizontal (see Figure 1.13).
- Xwell data - Crosswell data. Data collected by firing shots along one well and recording the resulting seismic vibrations by receivers along an adjacent well (see Figure 1.13).
- ZO data - Zero-offset data where the geophone is at the same location as the source.

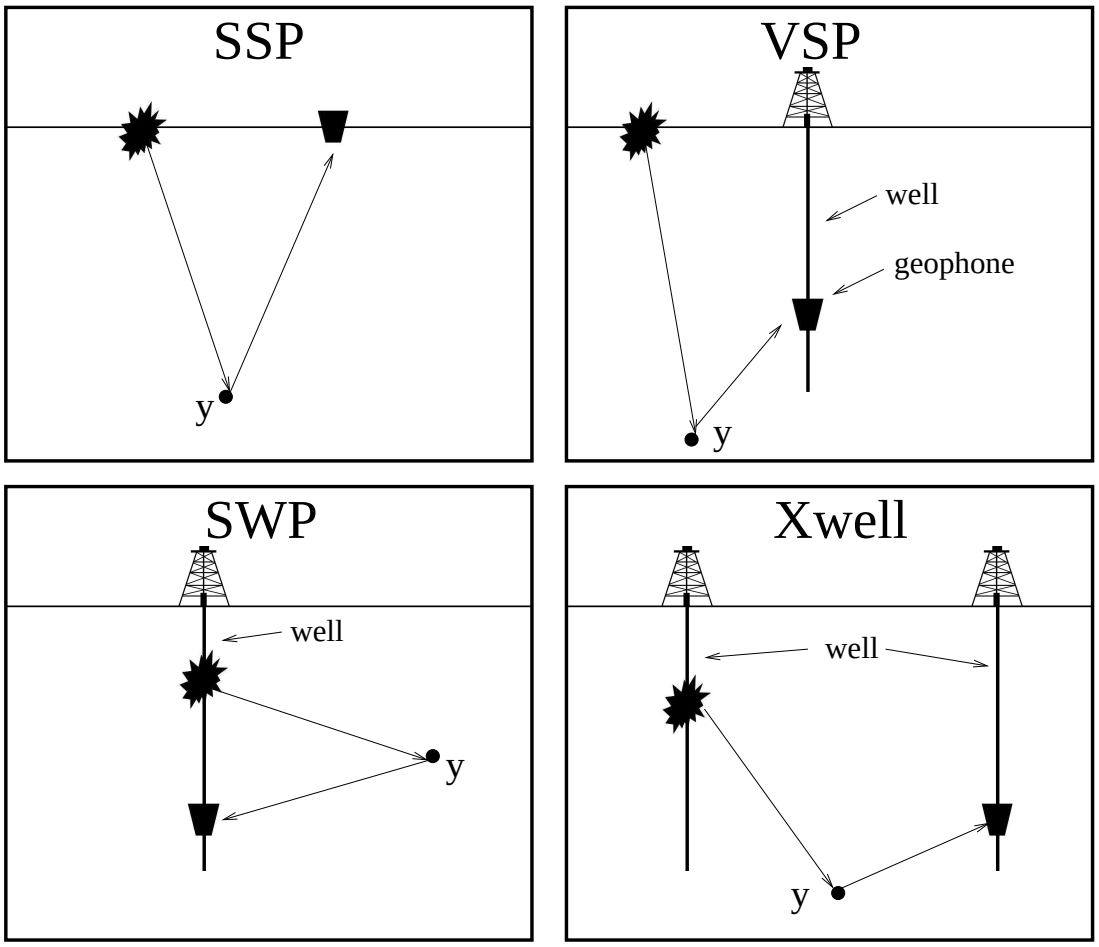


Figure 1.13: Source-receiver configurations for four different experiments: SSP=surface seismic profile, VSP=vertical seismic profile, SWP=single well profile, and Xwell=Crosswell. Each experiment can have many sources or receivers at the indicated boundaries (horizontal solid line is the free surface, vertical thick line is a well). The derrick indicates a surface well location, $\mathbf{y}$ denotes the reflection point, and the stars indicate sources.

Chapter 2

Diffraction Stack Migration

The basic assumption in resorting shot gathers into common midpoint gathers, applying NMO corrections, and stacking is that the reflectors are flat with no lateral velocity changes. If true, the origin of the stacked reflection energy originates somewhere directly beneath the trace, as shown in Figure 2.1. However, this will not be true for strong lateral velocity variations or layering that departs from the horizontal, such as the dipping layers shown in Figures 1.8, 2.2, 1.11, and 2.2, with the consequence that the reflection energy could have originated at a depth horizontally offset from the trace position. Thus we need to "migrate" the reflection to its point or origin, a process commonly known as migration. Not doing so will lead to mispositioning errors in estimating oil location and result in many dry holes.

In this section we describe the simplest type of migration known as diffraction stack migration, and is closely related to Kirchhoff migration. It can be formulated using intuitive reasoning only, but in a later chapter will be derived by a rigorous form of Green's theorem.

2.1 Poststack Migration

The goal in seismic imaging is to use seismic data to create reflectivity sections of the earth in either (x, z) or (x, y, z) space[1] If the reflection traces are all we have, how do we identify the (x, z) point as the origin point of reflections and migrate them to their reflection point of origin? The answer to this question for a point scatterer model is shown in Figure 1.11. The reflection energy at time t in trace C could have originated anywhere along a circle of radius vt centered at C, where v is the homogeneous velocity of the medium. Any reflector located along this circle could have contributed energy to trace C at time t. However, this ambiguity in location is not acceptable, so we improve our answer by asking the question:

[1]For 3D prestack experiments on a horizontal plane, the sources and geophones occupy areal positions (x_s, y_s) and (x_g, y_g), respectively, on the surface and we image the 3-dimensional reflectivity volume in model space coordinates of (x, y, z). Therefore, the prestack migration operation maps the 5-dimensional data space of (x_s, y_s, x_g, y_g, t) to the 3-dimensional model space of (x, y, z). For notational convenience we usually restrict our migration examples to 2D ZO migration where the mapping is from (x, t) space to (x, z) space.

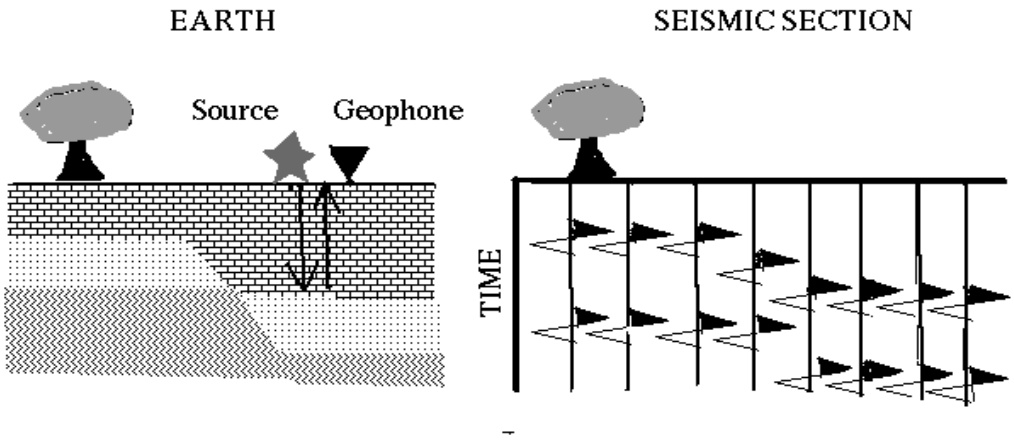


Figure 2.1: Earth model on top left and idealized zero-offset (ZO) seismic section on top right, where each trace was recorded by an experiment where the source has zero offset from the geophone. The above ZO seismic section represented by $data(x, z = 0, t)$ roughly resembles the earth's reflectivity model $m(x, z)$ because we unrealistically assume it contains only the primary reflections and the reflections in the data originate directly beneath the trace.

what kind of reflector could account for *all* the traces? The answer is found by drawing circles of radius vt centered at all trace positions, where t is the reflection time at that particular trace. The *common* tangent to these circles defines the reflector that could account for energy seen in all of the traces (see Figures 1.11 and 1.12).

2.1.1 Poststack Migration = Smear $d(x_g, t)$ along Semicircle(x, z)

Rather than drawing circles, the migration procedure can be automated by smearing reflection energy along fat circles where each doughnut is filled with the reflection wavelet $w(t)$, as shown in Figure 1.11. The width of the doughnut is $T_0/2$, where T_0 is the dominant period of the wavelet. Smearing for a single trace is carried out by plotting $d(x, t = 2r/v)$ in model space coordinates (x, z); here, $t = 2r/v = 2\sqrt{(x - x_g)^2 + z^2}/v$ is the two-way traveltime in a homogeneous model with velocity v. The smeared wavelets in the zone of common intersection will be in phase and add in a coherent manner, while outside this zone they will, on average, be out of phase and tend to cancel. Thus, *migration can be described as smearing and summing the reflections* along the appropriate doughnuts (Claerbout, 1992).

In matrix-vector notation, ZO migration is given by the formula:

$$\begin{aligned} m(x,z) &= \sum_{g=1}^{ntraces} \ddot{d}ata(x_g, t(x_g, x, z))/||(x - x_g)^2 + z^2||^2, \\ &\quad where \\ t(x_g, x, z) &= 2\sqrt{(x - x_g)^2 + z^2}/c, \end{aligned} \tag{2.1}$$

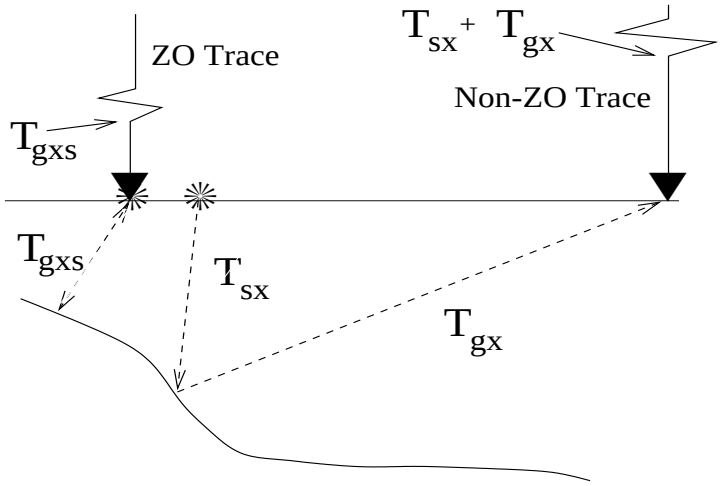


Figure 2.2: Raypaths and associated traces for a 2-layer medium. Note that there is no common reflection point associated with these rays, and so the 1D NMO formula incorrectly predicts the moveout of the reflection data. The subsequent stack will not produce the correct ZO trace illustrated on the left.

where $t(x_g, x, z)$ is the two-way time for energy to go from the surface point $(x_g, 0)$ to the model point (x, z) and back up to the receiver, and $data(x_g, t)$ is the reflection ZO trace in two-way time at position $(x_g, 0)$. The explanation for the geometric spreading factor will be deferred to a later chapter; and the double dot denotes second-order differentiation with respect to time, which is to undo the smoothing effects of integration in the data-space x' coordinate. Since phase changes mainly determine the constructive/destructive interference of energy, I will sometimes ignore the geometrical spreading factor in later formulations.

This MATLAB fragment for ZO migration can be written as

```
ZO Migration: Smear Impulse along Semi-circle in (x,z)

for ixtrace=1:ntrace;       Loop over ZO trace indices
for xs=istart:iend;         Loop over model space indices (xs,zs)
for zs=1:nz;
   r = sqrt((ixtrace*dx-xs*dx)^2+(zs*dx)^2);      compute radius of semi-circle
   time = round( 1 +  r/c/dt );                   compute 1-way time to circle
   m(xs,zs) = m(xs,zs) + data(ixtrace,2*time)/r;  sum data reflections into m(x,z)
 end;
 end;
end
```

For a single trace input with a single impulse value in time, the migration response of the above code is known as the *impulse response of the ZO migration operator*. Any of the doughnuts in Figure 2.3 represents a bandlimited impulse of the ZO migration operator. Impulse responses are used to test the fidelity of a migration algorithm, e.g., the actual ZO impulse response in a homogeneous medium should be a circle, and any departures from that is a cause for concern in a numerical migration operator.

Examples of migrating data associated with a simple scatterer model and a syncline model are shown in Figure 2.4. Note, the data in $data(ixtrace, time)$ are assumed to be in one-way time, otherwise the velocity v would have to be halved.

The above MATLAB code can have the outer loop exchanged with the two inner loops to give

```
ZO Migration: Sum Data along Hyperbolas and Dump into m(x,z)

for xs=istart:iend;         Loop over model space indices (xs,zs)
for zs=1:nz;
for ixtrace=1:ntrace;       Loop over ZO trace indices
   r = sqrt((ixtrace*dx-xs*dx)^2+(zs*dx)^2);      compute radius of semi-circle
   time = round( 1 +  r/c/dt );                   compute 1-way time to circle
   m(xs,zs) = m(xs,zs) + data(ixtrace,2*time)/r;  Sum data reflections into m(x,z)
 end;
 end;
end
```

For a fixed value of xs and zs, we can interpret the above ZO migration as *summing the trace amplitudes along a hyperbola in data space and dumping the result into $m(x, z)$; i.e., smear and dump.*

In 3D prestack migration, the data volume is 5 dimensions compared to the 3D model volume, and so 3D data typically must be stored on hard disk while the smaller 3D migration model can be stored in physical memory. This means that the smear-along-circle migration is most efficient because it visits each trace sample just once. In comparison, the "sum-along-hyperbola" migration, where the inner loop is along trace indices, revisits the same trace many times and therefore must inefficiently access slow hard disk many times.

2.1.2 Inhomogeneous Velocity

An implicit assumption in the above procedure is that the velocity medium is homogeneous so that possible reflector positions fall along a semi-circle. More realistically, the earth has a heterogeneous velocity distribution so that energy should fall along a irregular circle shown at the bottom of Figure 2.3. The traveltimes to the quasi-circles can be computed by using ray tracing to construct the traveltime table $t(x, z, ixtrace)$ for a source at $(ixtrace, 0)$; that is, find the 3D matrix $t(x, z, ixtrace)$ for all model coordinates (x, z) and trace offsets $(ixtrace, 0)$. Then, replace $time = round(1 + r/c/dt)$ in the above MATLAB code by $t(x, z, ixtrace)$.

2.1.3 3D Poststack Migration = Smear $d(x_g, t)$ along Hemisphere(x, z)

Instead of a line of ZO traces, a 3D section will contain an areal patch of ZO traces. The motivation for 3D recording (recording data over an areal patch of ground rather than a line of recordings) is that the reflection energy could have originated anywhere in the subsurface, such as out-of-the-plane reflections. To migrate these out-of-the-plane reflections to their origin point we should smear+sum reflection energy along semispheres rather than semicircles (see Figure 2.5). The change to the MATLAB code is to include extra loops over the model- and data-space y-axes.

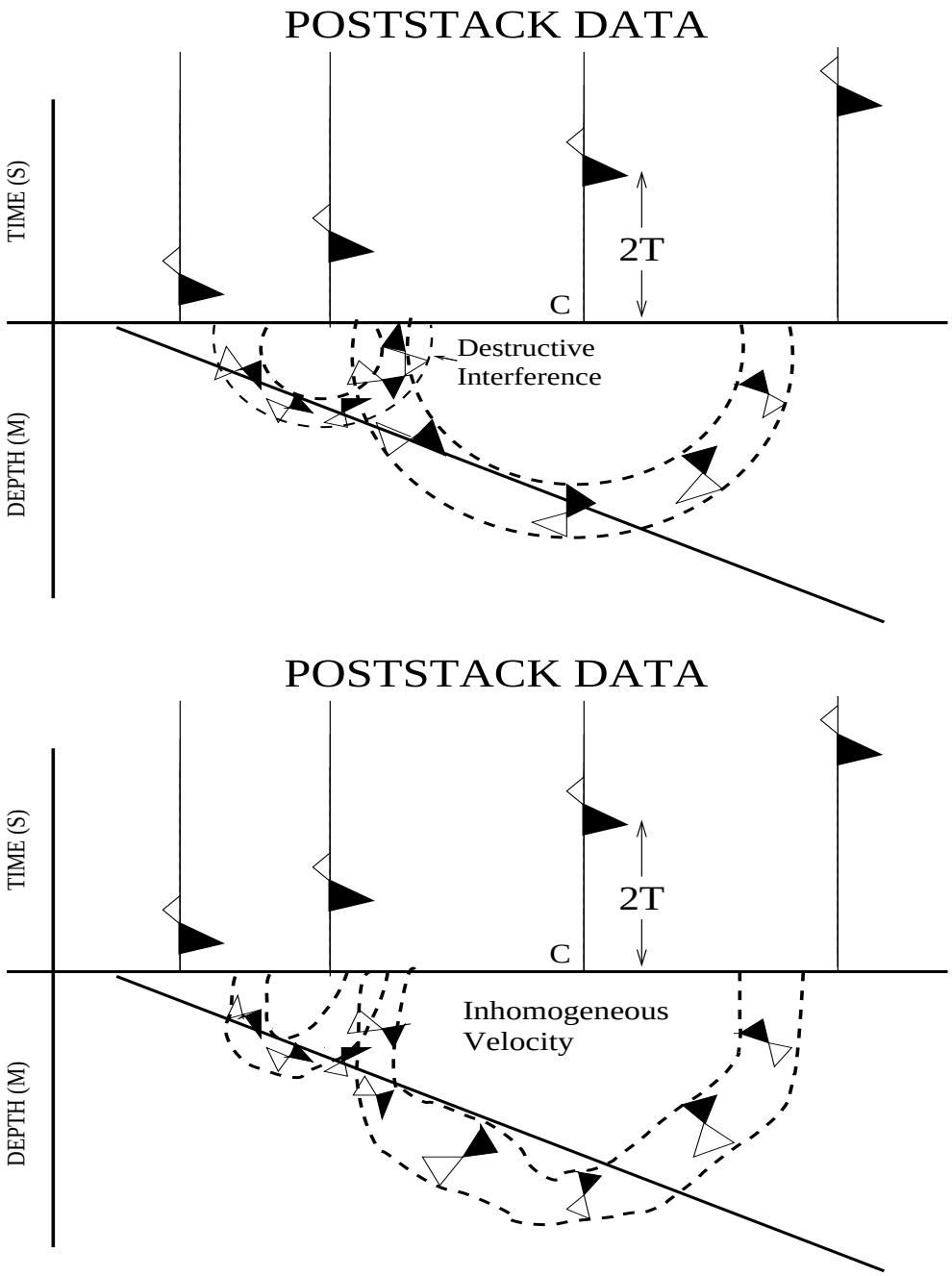


Figure 2.3: Fat migration "circles" for a (top) homogeneous and (bottom) an inhomogeneous velocity medium. The source wavelet fills in the fat "circles".

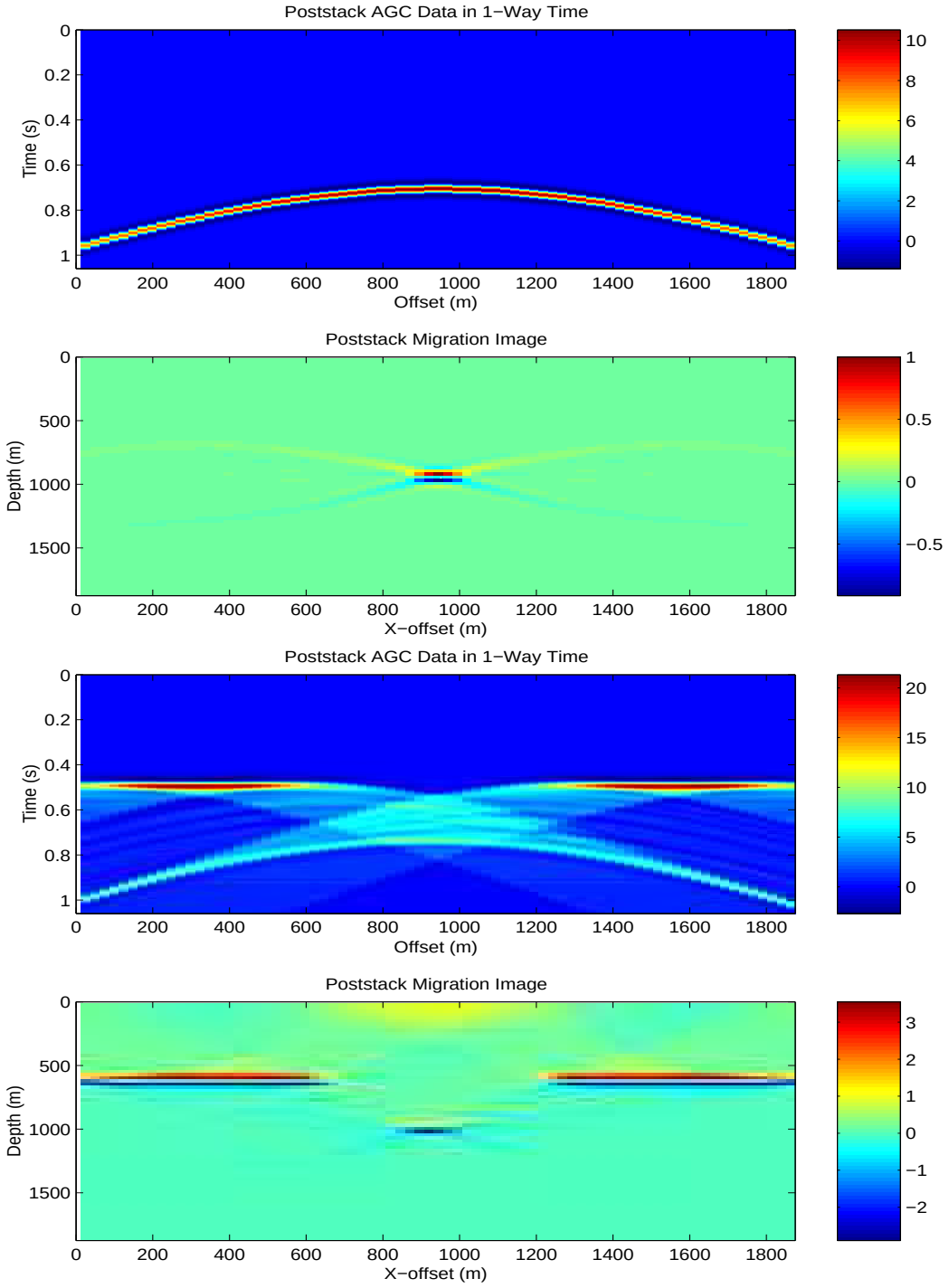


Figure 2.4: ZO traces and model generated by reflections from a (top 2 images) point scatterer and (bottom 2 images) synclinal model denoted by white dashed lines. Migration image in red and blue colors is shown as well.

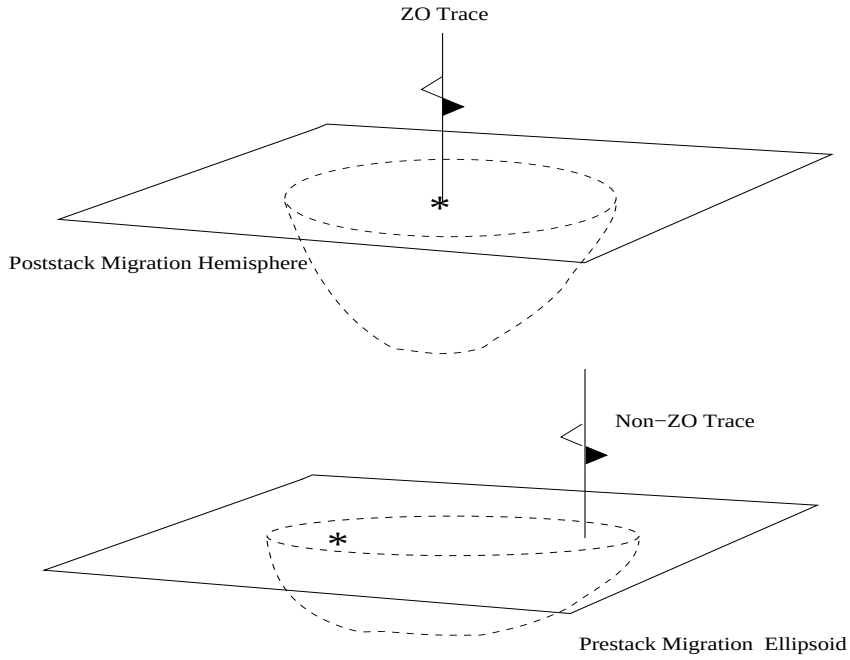


Figure 2.5: (Top) Poststack and (bottom) prestack migration impulse responses for 3D data.

3D MATLAB Poststack Migration Code

```
for ixtrace=1:ntrace;    % Loop over ZO trace indices
for IYTRACE=1:ntrace;
 for xs=istart:iend; % Loop over model space indices (xs,ys,zs)
 for zs=1:nz; for YS=1:ny;
  r = sqrt((ixtrace*dx-xs*dx)^2+(IYTRACE*dy-YS*dy)^2+(zs*dx)^2); % compute radius of
                                                                 % hemi-sphere
  time = round( 1 +  r/c/dt );                % compute 1-way time to circle
  m(xs,YS,zs) = m(xs,YS,zs) + data(ixtrace,IYTRACE,time)/r; % Smear and sum data
 end                                                        % reflections into (x,z)
 end;
 end;
end
end
```

Note that there are almost twice as many *end* statements in the 3D code compared to the 2D code, which means that 3D ZO migration is several orders of magnitude more expensive than 2D ZO migration.

2.1.4 Obliquity Factor

An improved poststack image can be obtained by including an obliquity factor in the diffraction stack equation:

$$m(x,z) = \sum_{g=1}^{ntraces} \ddot{d}ata(x_g, t(x_g,x,z))\hat{\mathbf{n}}\cdot\hat{\mathbf{r}}/||(x-x_g)^2+z^2||^2, \quad (2.2)$$

where $\hat{\mathbf{n}}$ is the unit normal at the surface and $\hat{\mathbf{r}}$ is the unit vector at the surface that is the incidence angle of the reflection ray between the trial image point and the geophone.

2.2 Prestack Migration

In a complex medium, such as beneath a salt dome, the NMO and stacking assumption is inappropriate so that the stacked section is too blurred, and so cannot be well imaged by poststack migration. The cure to this problem is to eliminate the NMO and stacking steps and perform prestack migration on the shot gather $d(x_g, 0, t_g|x_s, 0, t_s)$ recorded on a horizontal recording plane, where x_g and x_s denote the source and geophone x-coordinates for a shot gather. Here, t_g and t_s denote the listening time and source excitation time, respectively, and the source is typically assumed to be excited at time $t_s = 0$.

2.2.1 Prestack Migration = Smear $d(x_g, t)$ along Ellipse(x, z)

Which parts of the model could the reflection energy at time τ originate from? Similar to the poststack migration example, the answer is that the (x, z) parts of the model that satisfy the moveout equation for a fixed t, x_g and x_s. This equation defines an ellipse in model space with foci at $(x_g, 0)$ and $(x_s, 0)$.

Similar to the poststack migration formula, prestack migration smears a reflection sample into model space, but along the appropriate ellipse rather than a semi-circle. The formula for prestack migration is given by

$$m(x,z) = \sum_{x_g}\sum_{x_s} \ddot{d}(x_g, 0, \tau(x_g, x_s, x, z)|x_s, 0, t_s)/A(x, z, x_g, x_s), \quad (2.3)$$

Examples of stacked data, poststack migration and prestack migration are given in Figures 2.6-2.7.

2.2.2 Migration as a Pattern Matching Operation

The similiarity between two digital photos (each assumed to be a 100x100 pixelated image in the x-y plane) can be quantified by representing each photo by a $10,000x1$ vector and taking their dot product. If the photos are very similar then the dot product will yield a sum of mostly positive numbers to give a very high correlation coefficient. Conversely, the dot product between dissimilar photos will yield a sum of both positive and negative numbers to give a small correlation coefficient. Taking dot products of photos is a common *pattern matching* operation.

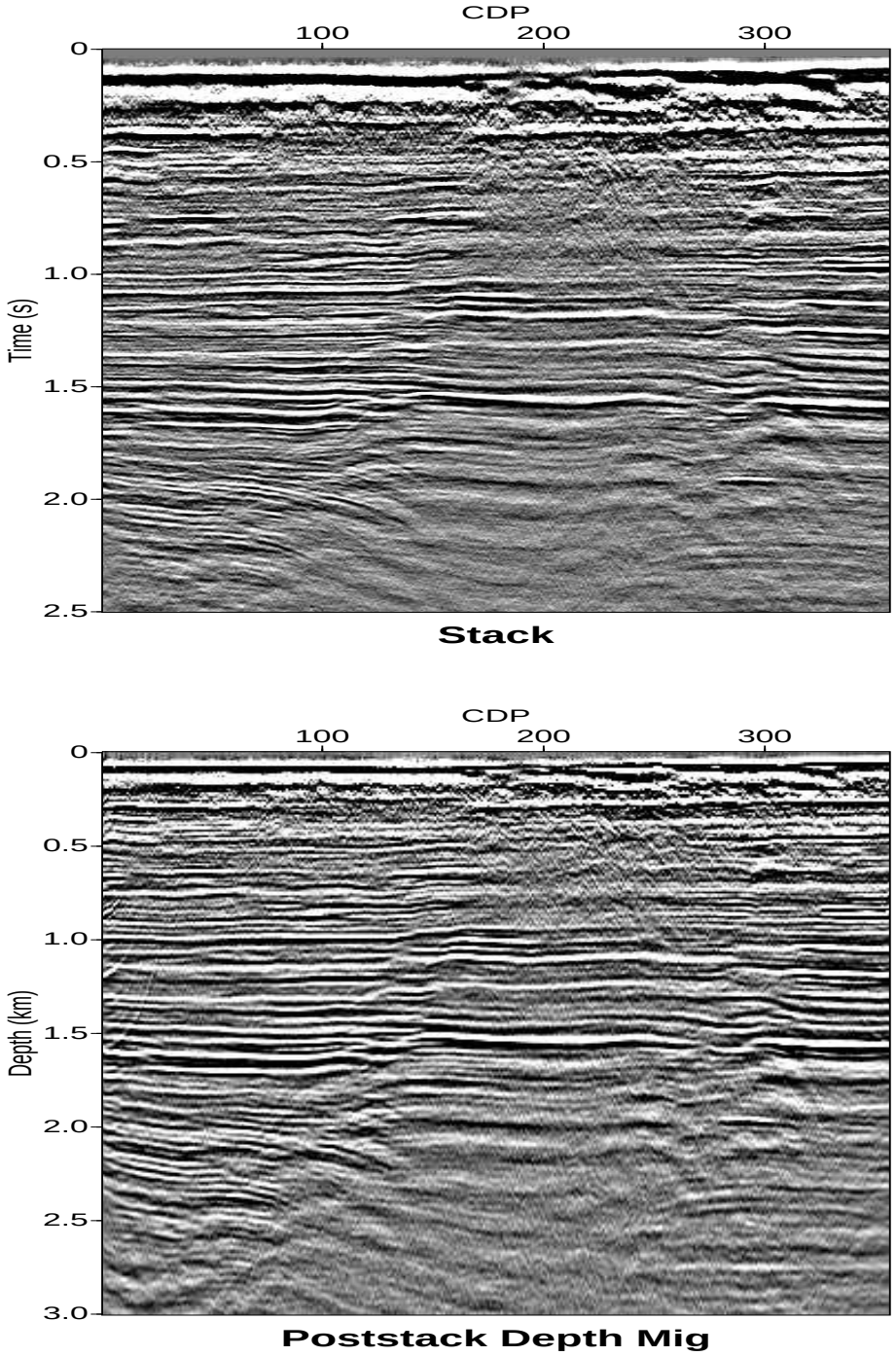


Figure 2.6: (Top) Stacked section and (bottom) poststack depth migration sections (courtesy of Jianhua Yu and Unocal). Notice how truncated reflectors have been shortened and diffractions frowns have been reduced by migration. Moreover, reflections with conflicting dips (events with the same arrival time but different dips) have been relocated to their places of origin along the interface.

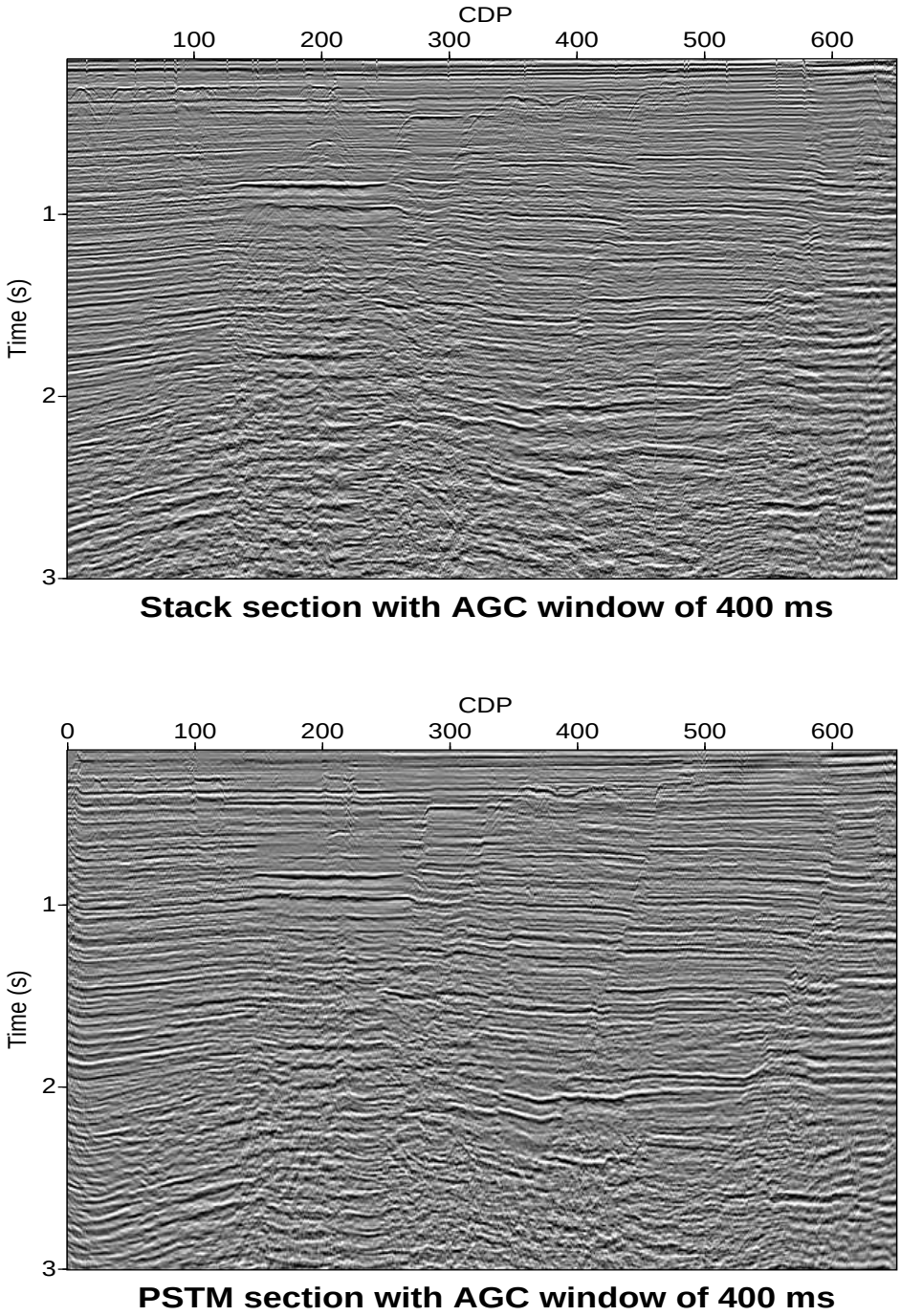


Figure 2.7: (Top) Stacked section and (bottom) prestack time migration (courtesy of Jianhua Yu and Unocal).

Summing the data over migration hyperbola in $x-t$ space can also be thought of as a pattern matching operation. The left image in Figure 2.8 depicts the migration curves in $x-t$ space as solid curves, where each colored curve corresponds to a different trial image point with the same color. Summing the data over a curve is equivalent to a 2-D dot product between the migration operator image and the data image. If the trial image point is near the actual scatterer, then the, e.g. black, migration operator in the left image of Figure 2.8 matches the data very well. Hence, the migration image at that trial image point has a high value. At other trial image points, the pattern of the migration operator correlates poorly with the data so the correlation is small to give a small value in the migration image.

2.2.3 Migration Operator for Multiple Arrivals

In theory, migrating many types of events with different arrival angles to their common reflector point leads to a better resolution at that point, and a cleaner migration image. This is similar to looking at a diamond from different view angles, each new view angle revealing a new facet of the gem.

To achieve this extra resolution with seismic images, one can tune a diffraction migration operator to migrate both primary reflections *and scattered multiples*[2] A representative multiple migration operator is illustrated on the right hand side of Figure 2.8, where, for a trial image point, the summation of energy is along the hyperbolic curves (e.g., the solid black curves) that represent the traveltimes for both primaries and multiples. For the correct trial image point at the black dot, a huge of amount of seismic energy gets placed at the scatterer's position by primary+multiple migration compared to that for primary migration.

From a pattern matching point of view, the complicated pattern of the primary+multiple migration operator correlates well with data only in a small neighborhood of the actual scatterer's point; thus the image resolution is very good. Compare this to matching the simple primary migration operator to the data; there is a relatively large neighborhood around the actual scatterer that gives a good match between the operator's pattern and the actual data. This means a migration image with worse resolution compared to the primary+multiple migration image.

The disadvantage of primary+multiple migration is that its migration image is especially sensitive to errors in the migration velocity model. Small migration velocity errors tend to give a noisier image compared to the primary migration image. In fact, the primary+multiple migration operator can be shown to be identical to that for reverse time migration, a subject to be discussed in a later chapter.

2.3 Spatial Sampling Issues

How does one choose the trace and shot intervals as well as the total recording aperture? The trace interval is selected so that the discrete representation of the data still contains unambiguous information about dips. Steeper data dips and higher source frequencies will

[2]A multiple migration algorithm can be constructed by ray tracing the traveltimes for *both primaries and multiples and including the extra summations in the diffraction stack migration formula.*

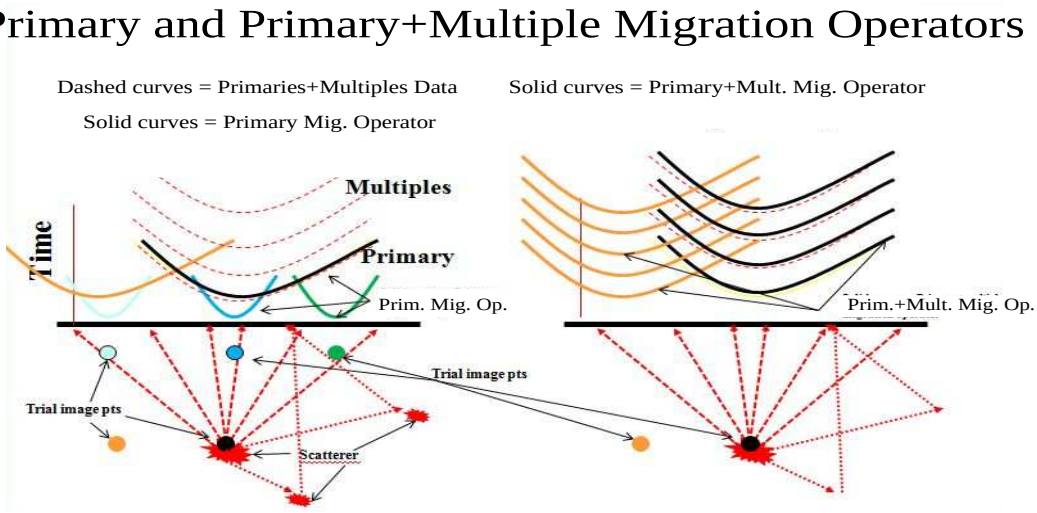


Figure 2.8: Migration of (left) primary and (right) primary+multiple reflection energy by diffraction stack migration of data (dashed curves). The solid curves represent the migration operators for different trial image points (represented by filled circles in model space). The primary+multiple migration operator associated with the black trial image point provides the best correlation between the "dashed-data" photo and the "black-migration-operator" photo.

demand finer trace spacing. Also, steeper dips in the model will give rise to ZO reflection rays that are almost horizontal, so will demand wider recording apertures.

If the trace spacing is too coarse, the "wings" of the hyperbolas will not completely cancel and so leave ugly smiles or frowns in the migration image, as shown in Figure 2.9. We say that the coarse trace spacing results in migration aliasing artifacts. Another point of view is that steep dips in the data appear more shallow if the trace spacing is too coarse; i.e., the data are aliased..

Aliasing can be cured by making sure that the shortest apparent wavelength in the data $(\lambda_x)_{min}$ is greater than 1/2 the trace spacing Δx. By definition $min(\lambda_x) = T_{min}(dx/dt)_{min}$ so that the anti-aliasing condition is

$$2\Delta x \quad < \quad min(\lambda_x) = (dx/dt)_{min} T_{min}, \tag{2.4}$$

where $min(\lambda_x)$ is the minimum apparent wavelength along the x-direction in the hyperbolas. This is known as the Nyquist sampling criterion, and its violation means that neighboring traces undergo more than 1/2 period of time shift. This assumes that the model grid is discretized finely enough to avoid problems in discretely sampling a continuous interface.

A visual example of aliased data is given in Figure 2.10. The bottom image is a snapshot of a horizontally propagating plane wave. The top and middle figures show the seismograms recorded on the horizontal surface, except the top section is aliased while the middle section is well sampled in space. The dip of the events in the middle section show the true propagation velocity (dx/dt=v) while the apparent velocity of the dip in the top section is slower because the data are aliased. Try running the MATLAB code below to see how aliasing affects recorded data.

```
colormap('gray'); subplot(311); plot(1:2,1:2); subplot(312);
plot(1:2,1:2); subplot(313); plot(1:3,1:3);
nangle=3;nt=50;nsnap=nt*nangle;itt=0;c=.2;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

rect=get(gcf,'Position');
rect(1:2)=[0 0];
M=moviein(nsnap,gcf,rect);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
for iangle=1:nangle; lam=6;angle=pi/4.2*(iangle-1);
kx=2*pi*cos(angle)/lam;ky=sin(angle)*2*pi/lam;
nx=100;k=sqrt(kx^2+ky^2);w=k*c; r=1:1:nx;i=ones(nx,1);
dx=i*r; dy=dx';tt=.001*[nt:-1:1];sei1=ones(nt,nx/4);
sei=ones(nt,nx);

for it=1:nt; itt=itt+1; pw=cos(dx*kx+dy*ky+w*it);pw=fliplr(pw);
for iii=1:25; sei1(it,iii)=pw(1,1+(iii-1)*4); end;
sei(it,:)=pw(1,:);
%%%%%%%%%%%%%%%
```

```
subplot(311);
imagesc(r,tt,sei1);
if iangle==1;title('Seismograms: Angle = 0 deg');end;
if iangle==2;title('Seismograms: Angle = 40 deg');end;
if iangle==3;title('Seismograms: Angle = 80 deg');end;
ylabel('Time (s)'); pause(.2)
%%%%%%%%%%%%%%%
subplot(312);
imagesc(r,tt,sei);
if iangle==1;title('Seismograms: Angle = 0 deg');end;
if iangle==2;title('Seismograms: Angle = 40 deg');end;
if iangle==3;title('Seismograms: Angle = 80 deg');end;
ylabel('Time (s)'); pause(.2)
%%%%%%%%%%%%%%%
subplot(313);
imagesc(pw);xlabel('Horizontal Offset (m)');
ylabel('Depth (m)');
if iangle==1;title('Snapshots: Angle = 0 deg');end;
if iangle==2;title('Snapshots: Angle = 40 deg');end;
if iangle==3;title('Snapshots: Angle = 80 deg');end;
pause(.2)
%%%%%%%%%%%%%%%
M(:,itt)=getframe(gcf,rect); end; end; N=1 ;
FPS=1 ;
movie(gcf,M,N,FPS,rect);
%mpgwrite(M, hot, 'filename', [1, 0, 1, 0, 10, 8, 10, 25]);
```

2.3.1 Aperture Limitation.

Wider apertures will lead to better horizontal resolution, and also allow for recording of events with nearly horizontal raypaths. The aperture width should be estimated by using the formula given in Figure 2.11.

2.4 Summary

The practical aspects of integral equation migration are reviewed. Migration can be viewed as either "smearing a time sample of data along the corresponding migration circle (or ellipse)", or equivalently it can be viewed as "summing energy along the appropriate hyperbola (for a fixed trial image point at $\mathbf{x}$ and source-receiver pair) and dumping it into the pixel centered at $\mathbf{x}$". The latter view is illustrated in Figure 2.12, where shallower trial image points lead to more steeply dipping wings of the hyperbola. Since we are summing amplitudes along these hyperbola we must ensure that the dip along the migration hyperbola is not as steep as the trace spacing allows; otherwise the data are aliased.

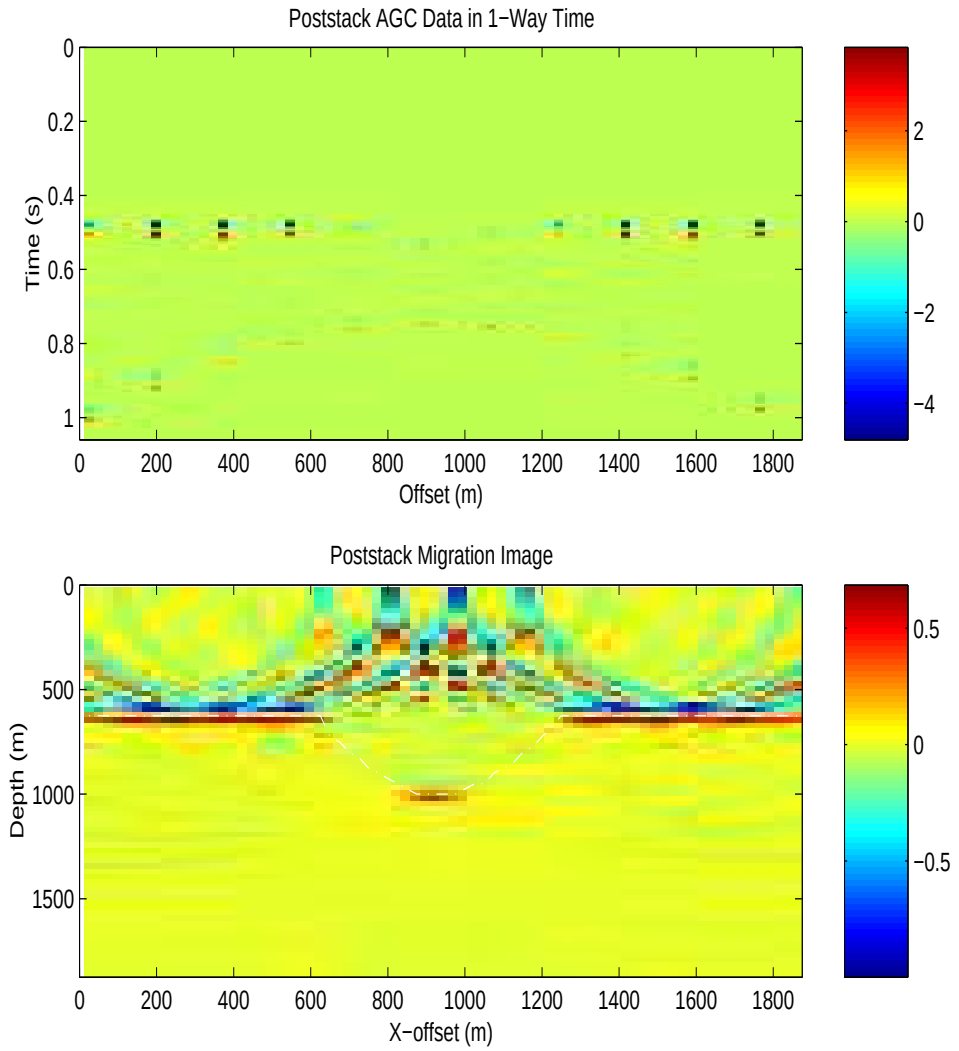


Figure 2.9: Same as synclinal model results except a coarse grid of data are the input to give rise to an aliased migration section with smiley artifacts. Note, the flat reflector data is not aliased but the steeply dipping part of data (syncline reflections) is aliased.

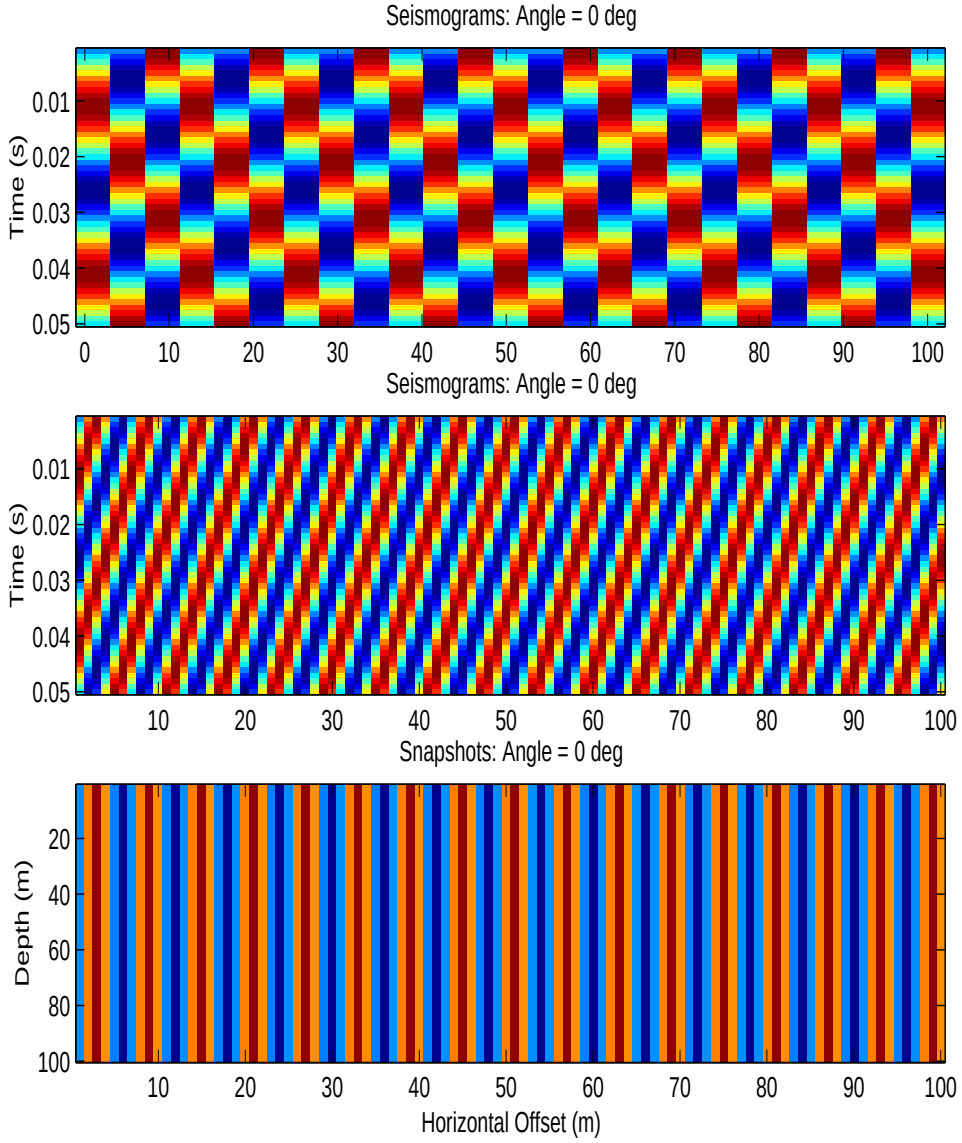


Figure 2.10: (Top) Aliased and (middle) well-sampled seismograms for the horizontally propagating wave seen in the bottom snapshot. Aliased data appears to have slower moving waves, just like a fan that appears to spin backward in a strobe light. Note, the top two sections represent seismograms in (x, t) space, while the bottom section is a snapshot of the propagating wave in model space coordinates (x, z).

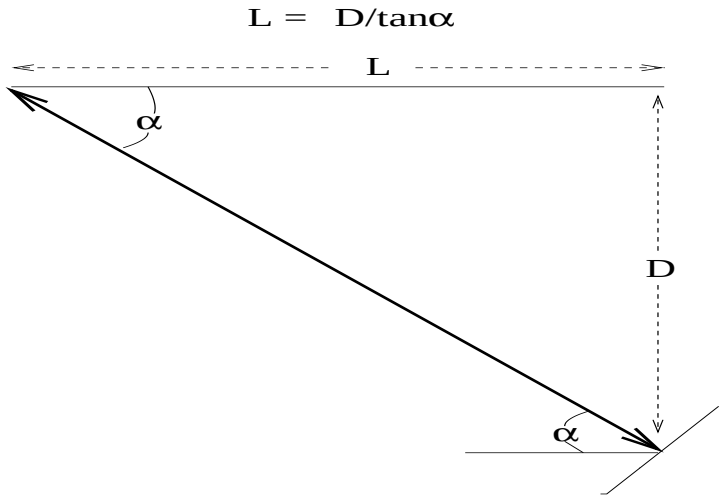


Figure 2.11: Aperture width L should be computed by estimating the maximum dip angle α of a reflector and its depth D, so that $L = D/tan\alpha$.

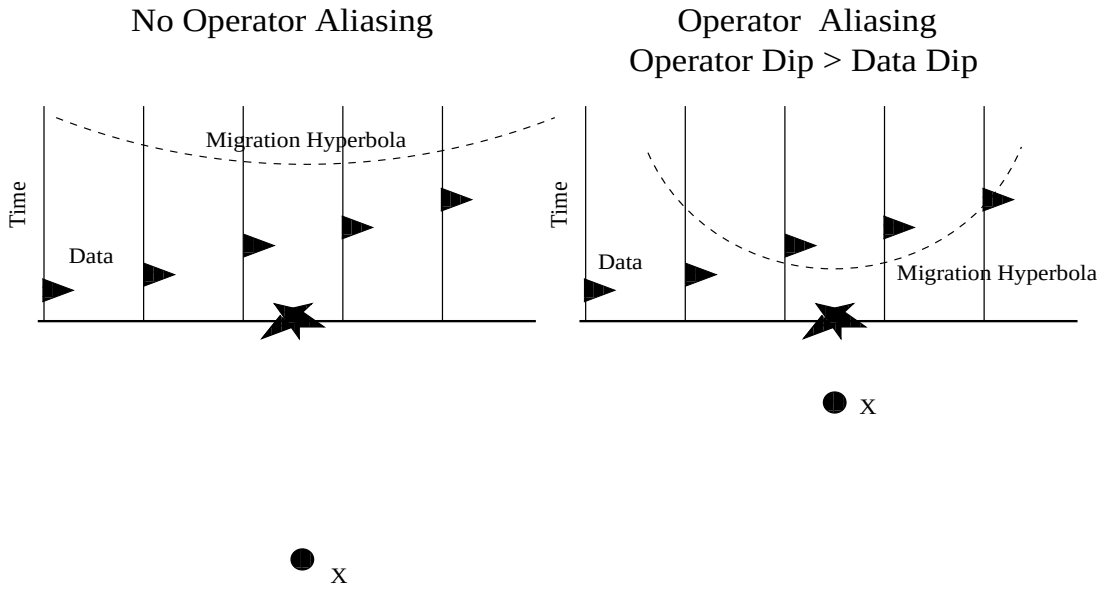


Figure 2.12: Dashed migration hyperbolas for a (left) deep and (right) shallow trial image point at **x**. The shallow migration hyperbola has steeper dips so requires a finer trace sampling than the one for a deep trial image point. Note, even though he data are not aliased the migration hyperbola can be.

A more economical migration can be carried out by reordering the loops so that the outer loops are over data space coordinates (i.e., traces), while the inner loops are over model space coordinates. This is because the data volume is five dimensional while the model volume is 3 dimensional; thus, IO cost is minimized if the trace sample is visited only once by having the outer loop over trace coordinates.

There are three critical parameters that should be selected prior to designing a seismic experiment or migrating seismic data:

- Spatial sampling interval Δx of geophones and sources. The spatial sampling interval is determined by the minimum horizontal wavelength λ_{min} of the seismic data such that $\lambda_{min} > 2\Delta x$. This minimum wavelength is estimated a priori from previous experiments in the area or from a test seismic experiment that spatially oversamples a shot gather and determines from this shot gather the minimum apparent wavelength in the data: $(dx/dt)_{min} T_{min} > 2\Delta x$. Normally, the surface waves are much slower than the body wave reflections (they travel along the surface at about the shear wave velocity at the near surface) so this would require too expensive of a geophone sampling interval. Instead companies will use an array of geophones (up to 70/group in Saudi Arabia!) for each group or channel to cancel out the slow moving surface wave yet retain the long wavelength body wave reflection.

- Aperture width. Wide angle reflections will allow for the capture of steeply dipping specular reflections. Estimation of the aperture width should use the formula given in Figure 2.11. An alternative estimation procedure, to be discussed later, is the Beylkin stretch formula where the user decides the acceptable resolution at a selected depth region and uses the stretch formulas to estimate aperture width. The best vertical resolution you can achieve is $\lambda_{min}/4$ but the best horizontal resolution is estimated by a complicated stretch formula that is a function of source and receiver coordinates.

- A digital antialiasing filter should be applied to the data to eliminate operator aliasing.

The diffraction stack formula is acceptable when there is plenty of redundancy in the data (i.e., high fold) but the obliquity factor should be used if fold is low. Experience shows that the obliquity factor suppresses aliasing-like artifacts in the final migration image. Unlike depth migration, time migration does not suffer from migration stretch but does suffer from mispositioning of events in complex geologic areas. Therefore it is rarely used today for subsalt imaging. However, the velocity model must be finely tuned in order to get depth imaging to show a coherent section. This compares to time migration which usually provides a good looking image because the stacking velocity is used to estimate the time-migration velocity. The stacking velocity is robustly estimated by efficient and automatic velocity scans while the velocity model for depth migration is typically estimated by a time-consuming and tedious process (e.g., reflection tomography or migration velocity analysis).

2.5 Exercises

1. Write a pseudo-MATLAB for 3D prestack migration.

2. Write a pseudo-MATLAB for primary and multiple migration.

3. Determine the maximum aperture for a seismic experiment in order to image $0 - 40$ degree dips at $z = 5$ km. Assume a homogeneous velocity of 5 km/s. What is the minimum geophone spacing in order to not spatially alias the data? Assume a minimum wavelet period of 0.01 s. Clearly show steps in your reasoning.

4. Estimate the horizontal wavelength of surface waves and body wave reflections in the Saudi shot gather (from a previous lab). What is a good array interval that would suppress surface waves but retain body waves ion the data. Test your estimate by applying an N-point spatial averaging filter to the data, where N is length of your estimated filter. Show results.

Chapter 3

Time Migration and Resolution

In this section time migration is defined and its pitfalls and benefits are compared to depth migration. In the far-field approximation the resolution formula for analyzing a migration image is derived and its application to practical problems is analyzed.

3.1 Time Migration vs Depth Migration

Depth migration in a variable velocity medium can be effected by replacing the $t = r/c$ statement in the MATLAB script by the traveltime obtained from ray tracing seismic energy from the trace position to the migrated image point. This is the correct procedure for downward continuing recorded energy to its place of origin, and such traveltimes can computed by a finite-difference solution to the eikonal equation.

Sometimes the velocity is not known very well so that the depth migration results do not look very good. In this case the geophysicists try to use a degraded form of migration known as time migration. Time migration is effected by replacing the $t = r/c$ statement in the MATLAB script by the NMO traveltime computed from the RMS velocity. The depth loop is replaced by a loop over the zero-offset traveltimes. This replacement automatically assumes an earth model with vertically varying velocity, negligible ray bending, and no lateral velocity variations. The reflection event is smeared along a circle, as shown in Figure 3.1. Note that this type of migration does not require ray tracing so it is faster than depth migration. The problem is that it still smears energy along circles, which is incorrect if there is significant ray bending across interfaces. This leads to the mispositioning of the lateral positions of irregular reflectors. As an example, the migration of the field data clearly images the dipping layer while the time migrated Kirchhoff migration smeared the reflector position. Below is a time-migration MATLAB code.

```
for ixtrace=1:ntrace;        Loop over ZO trace indicies
 for xs=istart:iend;         Loop over model space indicies (xs,t0)
 for t0=1:nt0;
  r = vnmo(t0)*t0           Compute geometrical-spreading radius
  time = sqrt( (t0/dt)^2 +  (x/vnmo(t0)/dt)^2 ); Compute 1-way time to circle
```

```
  m(xs,t0) = m(xs,t0) + data(ixtrace,time)/r; Smear/sum reflections into (x,t0)
 end;
 end;
end
```

One effect of time migration is that the time migration images sometimes looks better than depth migration images. This is because the images are plotted in offset vs two-way vertical time space, so there will be no wavelet stretch due to a faster velocity medium. That is, in depth migration the wavelet is smeared within the fat semicircle, but if the velocity of the semicircle is faster at depth (longer wavelength for a given period) then the upper part of donut is thin and lower part of donut is fat. Compare the stretching effect in Figure 2.6 to the no-stretch time migration section in Figure 2.7. Thankfully, time migration looks like it is becoming less common because of the inaccuracies in positioning events.

However, time migration is preferred if the velocity model is not very well known. Why? Because the NMO velocity is obtained by looking for the best hyperbola that most coherently stacks reflections. Thus, time migration focuses energy as well as any summation along hyperbolas can hope to achieve. Compare this to depth migration with a crummy velocity model. For a crummy velocity model we are summing energy along a correspondingly crummy quasi-hyperbola, resulting in a crummy focused image. Crumbs in, crumbs out.

3.2 Spatial Resolution

What is the minimum separation between two point scatterers that can be resolved in a migrated section? This minimum distance is known as the spatial resolution of the migrated section, where the horizontal resolution will differ from the vertical resolution. The spatial resolution will largely be a function of the wavelength λ, recording aperture L and/or depth of the scatterer.

The spatial resolution is quantitatively estimated by using the idea of a Fresnel zone (Elmore and Heald, 1969), where the first Fresnel zones makes the largest contribution to the recorded reflection signal, with a negligible contribution from the higher-order Fresnel zones. The definition of the 1st reflection Fresnel zone is the area on the reflector such that any reflection ray path (see Figure 3.5) from source-to-reflector-geophone differs in total path length by no more than a $\lambda/2$ (or arrival time difference of $T/2$). Therefore, two point scatterers located within each others first Fresnel zone will not be clearly distinguishable from one another on a migrated section because of interference effects. An example of imaging data for several neighboring point scatterers at different depths is given in Figure 3.3. Here, the lateral resolution becomes worse with depth because the migration circles become flatter with depth, so that the deepest pair of scatterers is not resolvable in the migration image. The migration ellipses interfere such that only one bump rather than two appear in the sideview image (Schuster, 1996; Chen and Schuster, 1999).

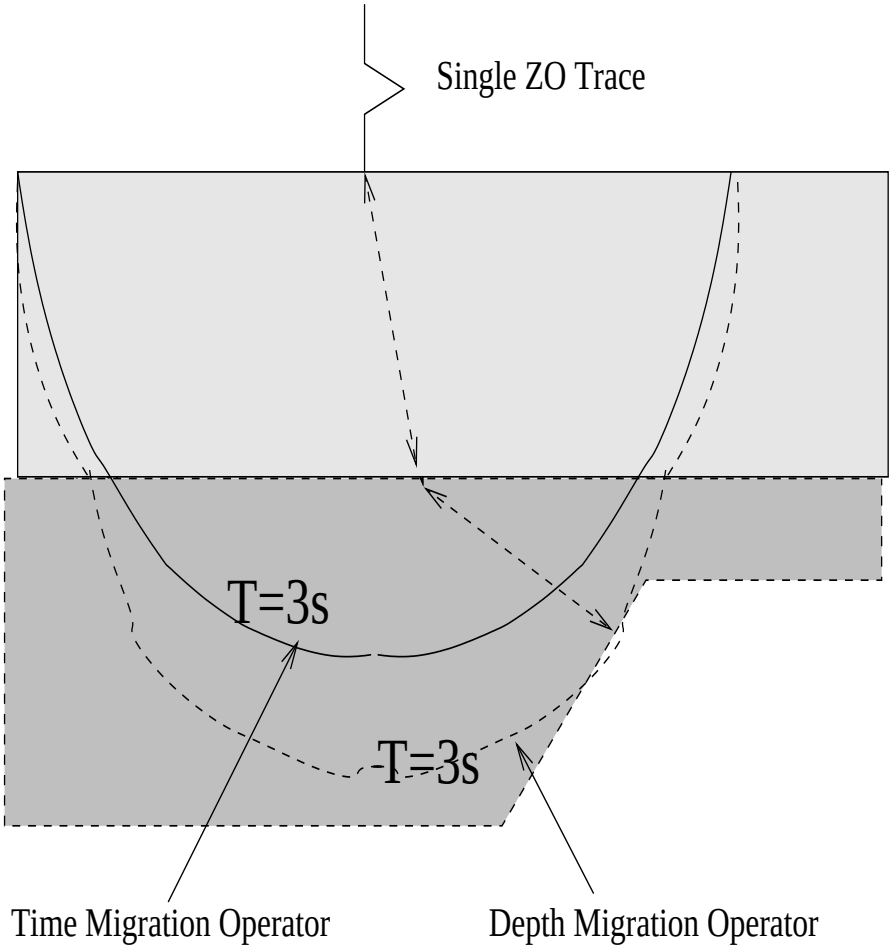


Figure 3.1: ZO time migration circles do not take into account raybending effects, so the result is that time migration laterally mispositions the events. Depth migration takes into account raybending across interfaces as well as lateral velocity variations, and so will correctly image the ZO reflection energy. The problem is that the ZO trace is formed by stacking CMP traces using a 1D NMO formula, which can be inappropriate for earth models with strong lateral velocity variations. The solution is prestack depth migration

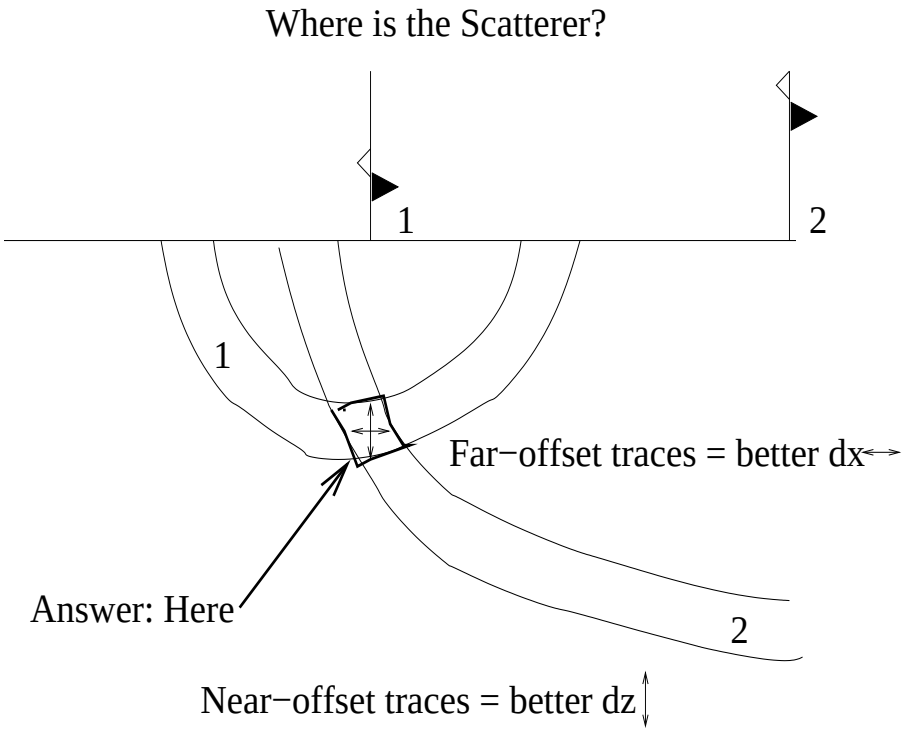


Figure 3.2: Where is the scatterer located that explains the reflections in the near- and far-offset ZO traces? Answer: The quadrilateral intersection zone. The width of the intersection zone is controlled by the far-offset (i.e., farthest from the scatterer) trace while its height is controlled by the near-offset (i.e., nearest to scatterer) trace.

An intuitive picture of resolution limits is given in Figure 3.2. Here, two traces are used to resolve the location of a scatterer, where the intersection of the migration donuts determines the approximate location of the point scatterer. Note, the width (i.e., horizontal resolution limit) of the intersection zone is controlled by the width of the far-trace donut at the scatterer point, while vertical height (i.e., vertical resolution limit) of the intersection zone is controlled by the thickness of the near-trace donut at the scatterer point. More generally, vertical resolution limits are controlled by the near-offset traces and the horizontal resolution limit is controlled by the far-offset traces. The next two sections quantify these resolution limits with analytic formulas.

Poststack Migration Resolution. As shown in Figure 3.5, a Fresnel zone encompasses

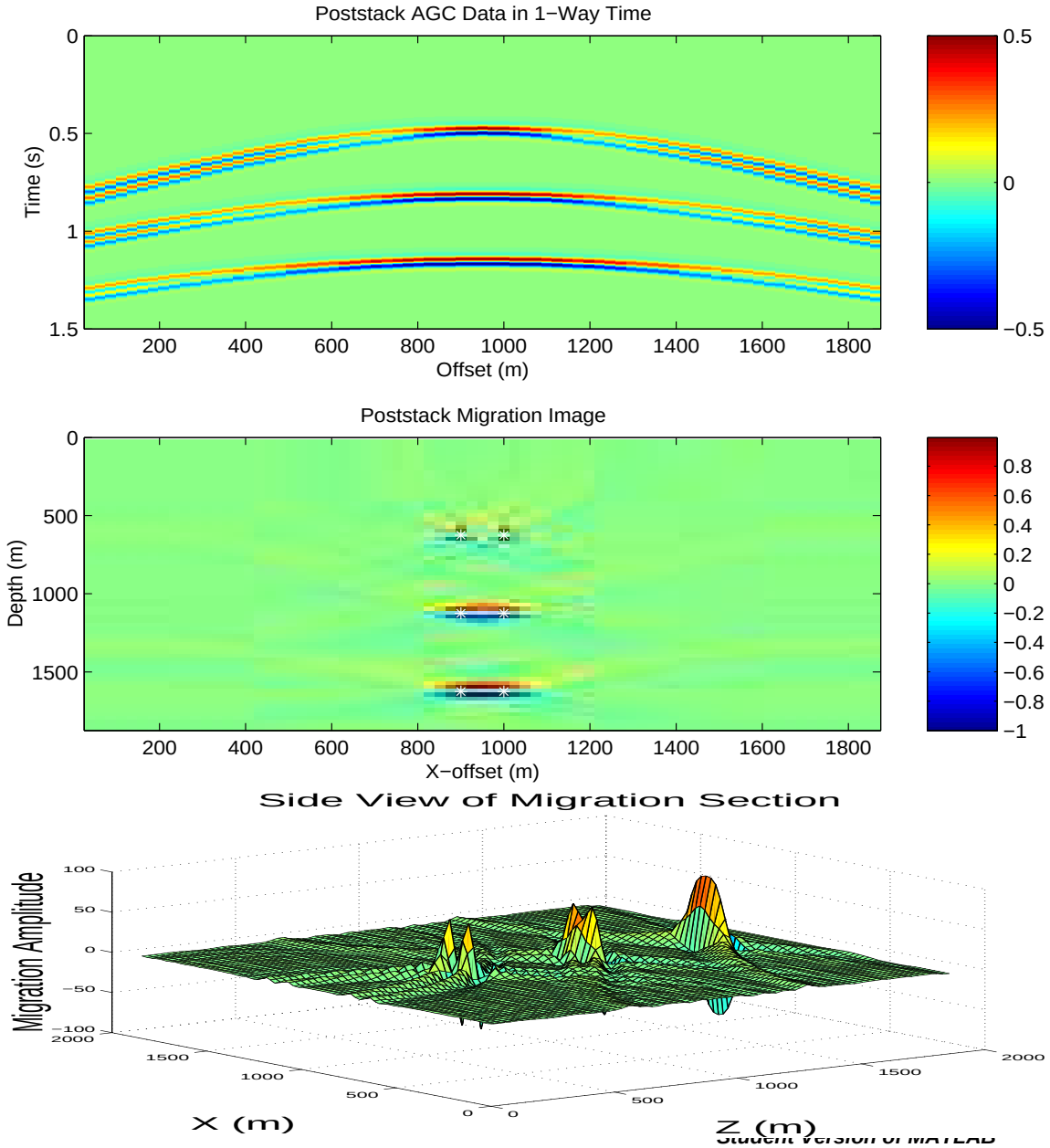


Figure 3.3: (Top) ZO data (middle and bottom) migration images for a model with 3 pairs of neighboring scatterers (white stars). Lateral resolution becomes worse with depth because the migration circles become less tilted and Fresnel zones become wider. The pair of scatterers separated by 125 m at $z = 1700$ m are not resolvable, where the central frequency of wavelet is 15 Hz, velocity is 1500 m/s, and the aperture width is 1875 m.

a spatial region in which the length difference between the shortest and longest reflection ray is $1/2\lambda$. For the ZO trace in Figure 3.5, the 1st horizontal Fresnel zone has a radius of

$$\begin{aligned} T/2 &= T_{ACA} - T_{ABA}, \\ &= 2AB\sqrt{1 + BC^2/AB^2}/c - 2 * AB/c, \\ &\approx 2AB(1 + 0.5BC^2/AB^2)/c - 2 * AB/c, \\ &= BC^2/(AB * c), \end{aligned} \tag{3.1}$$

and rearranging, noting that $cT = \lambda$, setting $AB = z_{depth}$ and solving for BC gives

$$BC \quad = dx = \quad \sqrt{\lambda * z_{depth}/2}. \tag{3.2}$$

where $AB = z_{depth}$ is the depth of the scatterer directly beneath the ZO trace. The length BC is proportional to the horizontal resolution directly below a ZO trace, and so horizontal resolution becomes better for smaller wavelengths and shallower scatterers.

But what is the horizontal resolution of a ZO trace for a scatterer laterally offset from the ZO trace? We can cook up an analytic formula similar to the above, except now we use a short cut. That is, differentiate the traveltime equation 2.1 w/r to x to get

$$dt/dx \quad = \quad 2 * (x - x_g)/(\sqrt{(x - x_g)^2 + z^2} * c), \tag{3.3}$$

where the factor 2 comes about because we are using the two-way traveltime equation. Setting $dt = T/2$ and solving for dx gives

$$\begin{aligned} dx &= (cT\sqrt{(x - x_g)^2 + z^2})/(4(x - x_g)), \\ &= \lambda\sqrt{(x - x_g)^2 + z^2}/(4 * (x - x_g)). \end{aligned} \tag{3.4}$$

We seek the parameters that give the minimum resolution value dx and denote this as $\Delta x = min\ dx$. If the scatterer depth z is much larger than the aperture $L = max(x - x_g)$, then equation 3.4 becomes

$$\begin{aligned} \Delta x \ = min(dx) = \quad & .25T\sqrt{L^2 + z^2} * c/L, \\ = \quad & .25\lambda z\sqrt{1 + L^2/z^2}/L, \\ \approx \quad & .25\lambda z/L, \end{aligned} \tag{3.5}$$

for $z >> L$. Thus larger apertures, shallower scatterers, and smaller wavelengths lead to better horizontal resolution.

Similar considerations show that the vertical resolution can be obtained by subtracting $T_{ABA} - T_{ACA}$ for the vertical raypath shown in the right plot of Figure 3.5:

$$\begin{aligned} T/2 &= T_{ACA} - T_{ABA}, \\ &= 2 * ABA/c - 2 * ACA/c, \\ &= 2 * BC/c. \end{aligned} \tag{3.6}$$

Solving for BC gives

$$BC = \quad \Delta z \quad = \lambda/4. \tag{3.7}$$

Figure 3.4 shows the depth migrated traces for reflections from a thinning bed, and suggests that we can distinguish there are two reflectors from the migration section if their thickness is greater than or equal to $1/4\lambda$.

Equation 3.7 says that vertical resolution does not depend on the recording aperture. In designing a recording array, the horizontal and vertical resolution limits can be estimated (see Figure 3.5) for ZO migrated sections by the formulas 3.5 and 3.7. In fact, equation 3.5 can be used to validate that the deepest pair of scatterers in Figure 3.3 are not laterally resolvable.

Prestack Migration Resolution. Similar considerations define the resolution limits for migrating a prestack gather of CMP traces, as shown in Figure 3.6. Here, the minimum vertical resolution of the migrated gather is governed by the migrated ZO trace, which is $\Delta z = \lambda/4$. On the other hand, the far-offset trace will govern the minimum horizontal resolution, which is $dx = \lambda * z/(4L)$, where L is the recording aperture.

A formula for Δx can be derived by a procedure similar to that of the ZO trace resolution; or by differentiating the traveltime equation

$$\begin{aligned} \tau &= \tau_{sx} + \tau_{xg}, \\ &= \overbrace{\sqrt{(x_s - x)^2 + (y_s - y)^2 + (z_s - z)^2}/c}^{downgoing\ time} + \overbrace{\sqrt{(x_g - x)^2 + (y_g - y)^2 + (z_g - z)^2}/c}^{upgoing\ time}, \end{aligned} \tag{3.8}$$

with respect to the x-coordinate x of the trial image point, $\partial\tau/\partial x$, setting $d\tau \to T/2$, and solving for dx. These are known as migration stretch formula and give both stretch and resolution estimates along the x and z directions.

How does one find the minimum dz and dx at image point (x_0, y_0, z_0) for an entire ensemble of prestack traces? Simply find the source receiver pairs that minimize these resolution estimates.

The above formulas are restricted to homogeneous velocity media, but the real earth is inhomogeneous in velocity. Resolution estimates can be obtained for inhomogeneous media by using the simple idea that the sum of the unit vectors of the downgoing and negative upgoing rays is proportional to the wavenumber $\mathbf{k_{model}}$ of the model that can be reconstructed For example, the ZO vertical rays suggest that the wavenumber of the model that can be reconstructed is proportional to $\vec{k}_{model} = (0, 0, 2\pi/\lambda)$. Note, the k_x and k_y components are zero because the ray is perfectly vertical and has no x-y component (Gesbert, 2003). Formally, the model wavenumber that can be reconstructed (see Figure 3.7) is given by

$$\vec{k} \quad = \quad \omega\nabla(\tau_{down}(x) + \tau_{up}(x)). \tag{3.9}$$

Similar considerations can be used to estimate resolution for tomographic images (Sheng and Schuster, 2003).

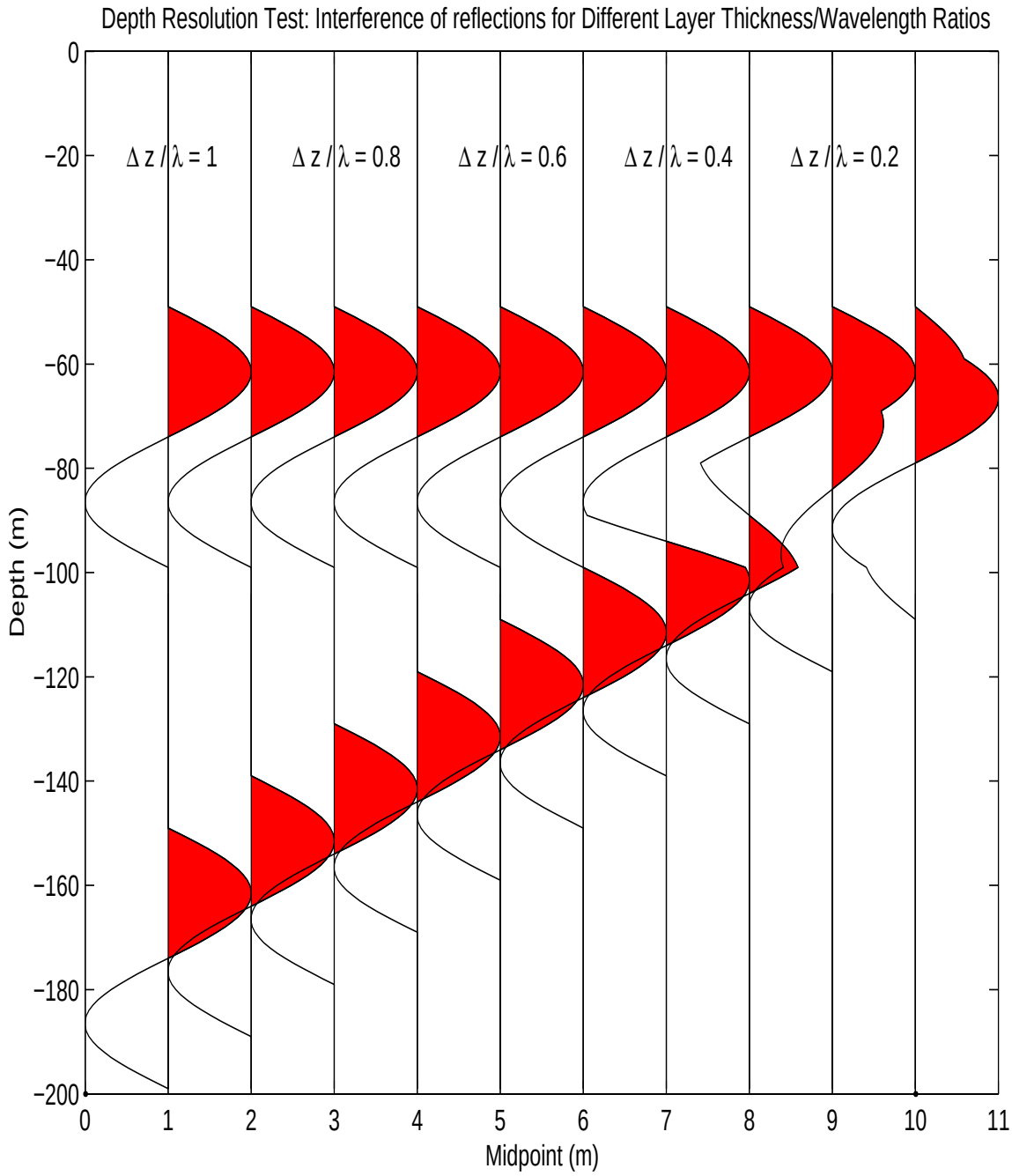


Figure 3.4: Vertical resolution limit is reached when the thickness between two neighboring reflectors $\Delta z = \lambda/4$.

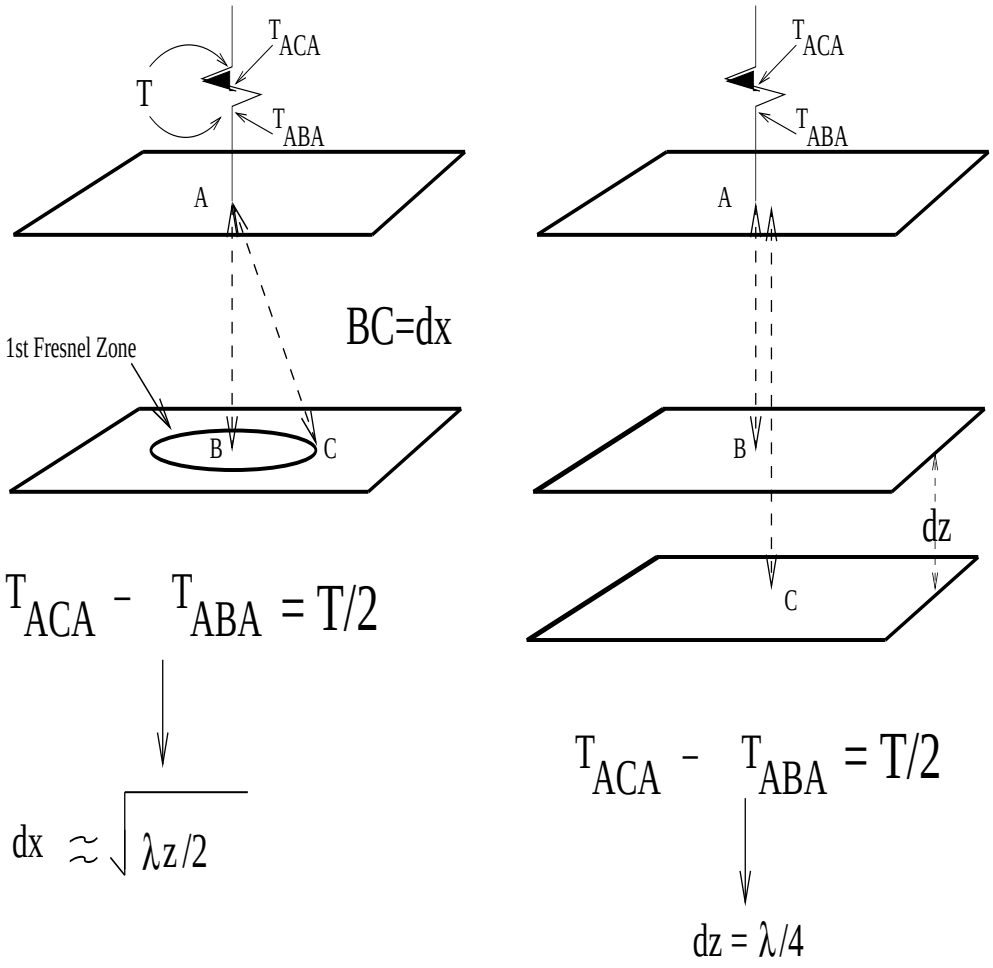


Figure 3.5: Extent of horizontal (left) and vertical (right) Fresnel zones for a ZO trace, where the 1st Fresnel zone defines the area where the difference in the shortest and longest raypath is equal to half the wavelength. The approximation formula $0.25\lambda z/L$ for horizontal resolution is valid when the B point is far to the right of the square and $z >> L$ so points A and A' appear to be coincident, resembling a ZO configuration. In contrast, if B is laterally offset $\approx L/2$ from the center at depth $z >> L$ then $\Delta x \approx 0.5\lambda z/L$ for the prestack configuration. If B is centered below the square then the horizontal resolution is $\approx \sqrt{z\lambda}$.

Is it really important to be concerned about the subtle nuances of improving lateral resolution? Jianhua Yu shows in Figure 3.8 that modest lateral resolution leaves doubt about the existence of fault, but slightly improving this lateral resolution via migration deconvolution leaves no doubt about the existence of a fault.

Beylkin Resolution → Migration Stretch. A simple interpretation of equation 3.9 is that it is equivalent to the formula for migration stretch. Specifically, the migration stretch formulas in the different coordinate directions are given by differentiating the traveltime formula 3.8 with respect to the coordinates of the trial image point (x, y, z):

$$\begin{aligned} \partial\tau/\partial x &= \partial\tau_{sx}/\partial x + \partial\tau_{xg}/\partial x, \\ \partial\tau/\partial y &= \partial\tau_{sx}/\partial y + \partial\tau_{xg}/\partial y, \\ \partial\tau/\partial z &= \partial\tau_{sx}/\partial z + \partial\tau_{xg}/\partial z. \end{aligned} \tag{3.10}$$

Using a Fresnel-zone argument, we set $\partial\tau \approx T/2$ and $\partial x \approx \Delta x$, $\partial y \approx \Delta y$ and $\partial z \approx \Delta z$ on the left hand side of the above equation, and rearrange terms to get analytical expressions for the resolution limits $(\Delta x, \Delta y, \Delta z)$, or migration stretch in the three coordinates similar to the expression in Figure 3.6. Here, T is the dominant period of the wavelet. Multiplying the three stretch formula in equation 3.10 by ω yields the components proportional to the Beylkin wavenumber formula in equation 3.9 (e.g., recall $\omega\partial\tau/\partial x = k_x$). This establishes the equivalency between Beylkin's fancy resolution formula and the well known migration stretch formulas.

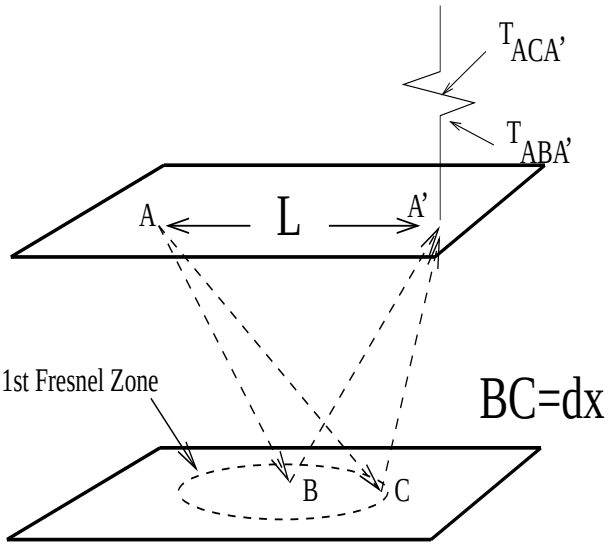
3.3 Summary

Both time migration and spatial resolution are defined. The Beylkin stretch formula is used so that the user decides the acceptable horizontal resolution at a selected depth region and uses the stretch formulas to estimate aperture width. The best vertical resolution you can achieve is $\lambda_{min}/4$ but the best horizontal resolution is estimated by a complicated stretch formula that is a function of source and receiver coordinates.

Unlike depth migration, time migration does not suffer from migration stretch but does suffer from mispositioning of events in complex geologic areas. Therefore it is rarely used today for subsalt imaging. However, the velocity model must be finely tuned in order to get depth imaging to show a coherent section. This compares to time migration which usually provides a good looking image because the stacking velocity is used to estimate the time-migration velocity. The stacking velocity is robustly estimated by efficient and automatic velocity scans while the velocity model for depth migration is typically estimated by a time-consumming and tedious process (e.g., reflection tomography or migration velocity analysis).

3.4 Exercises

1. The CDP interval in Figure 3.8 is 50 m, and assume the total depth of 4 km in the image. The velocity at $z = 1$ km is 2 km/s and at $z = 2.5$ km it is 4 km/s, and the recording aperture is 350 CDP's wide for any shot point. Calculate the



$$T_{ACA'} - T_{ABA'} = T/2 \longrightarrow dx \approx \frac{\lambda z}{4L}$$

Figure 3.6: Extent of horizontal Fresnel zone for a trace with non-zero offset between the source and receiver. The 1st horizontal Fresnel zone for a reflection defines the area where the difference in the shortest and longest reflection raypaths is equal to half the wavelength.

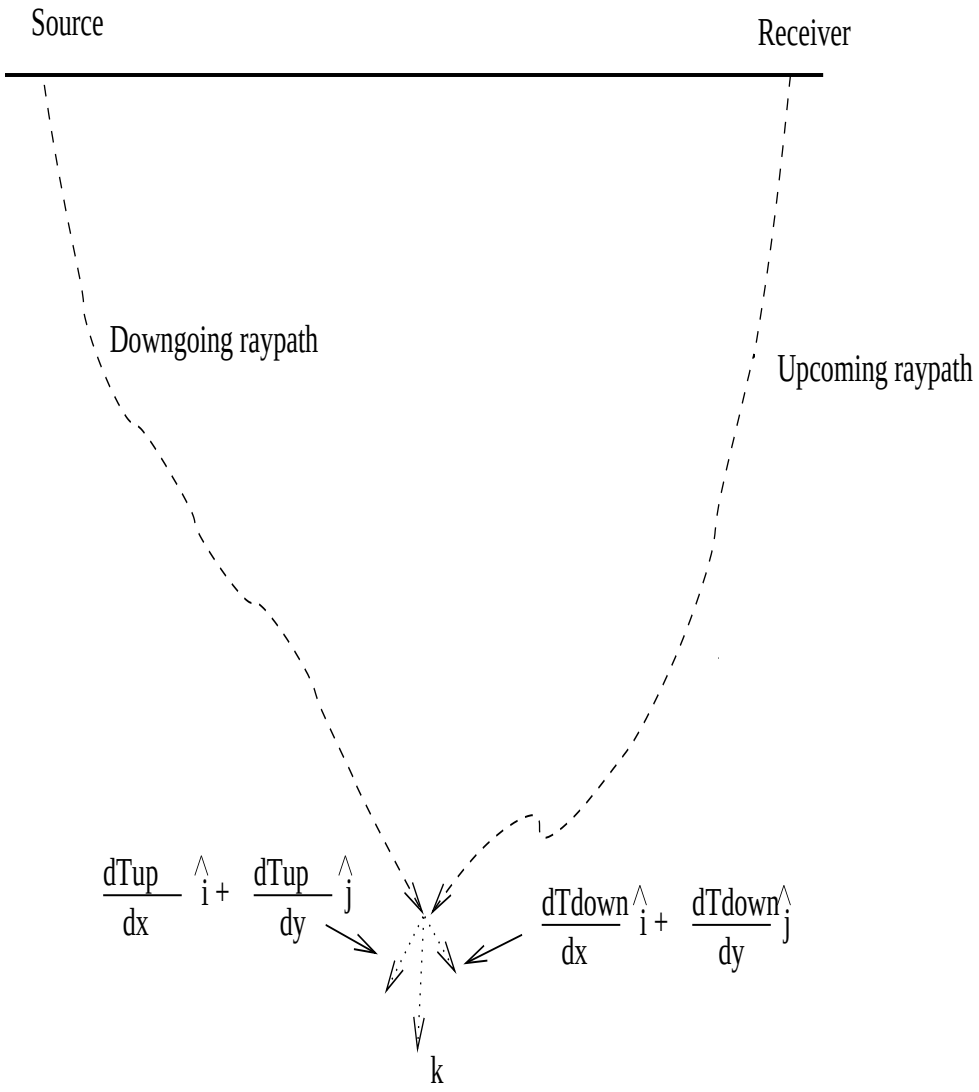


Figure 3.7: Wavenumber $\mathbf{k}$ of model that can be reconstructed by the downgoing and upgoing rays of the reflection. Here $\mathbf{k} = \omega(\nabla\mathbf{T_{down}} + \nabla\mathbf{T_{up}})$, so that wavenumbers that point sideways (down) indicate good lateral (vertical) resolution.

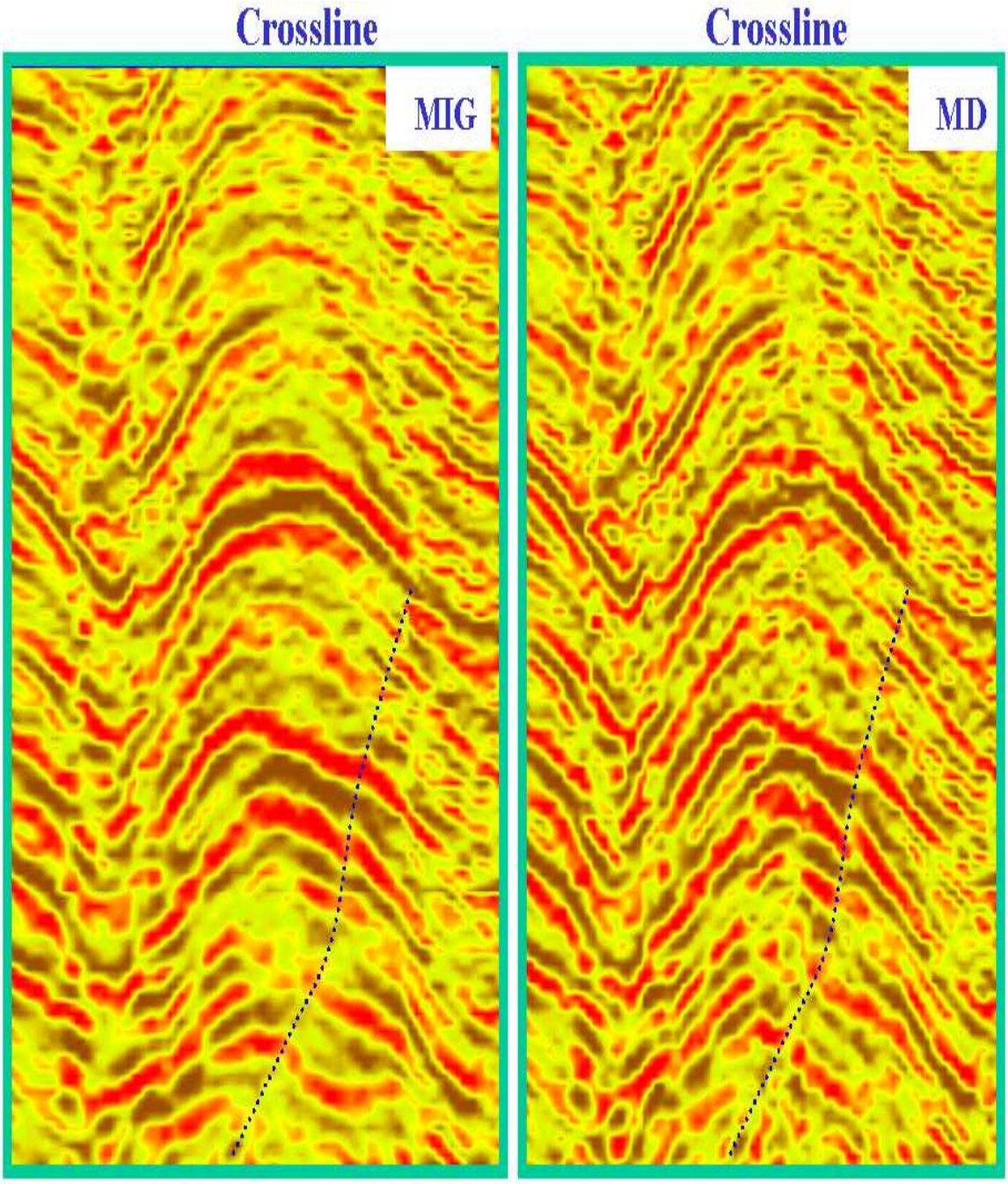


Figure 3.8: 3D prestack migration images (left) before and (right) after migration deconvolution. Lateral resolution improves by about 20 percent so that the right image leaves no doubt about the existence of a fault.

minimum horizontal and vertical resolutions for a prestack migration image at the points $(x, z) = (0\ km, 1\ km)$, $(x, z) = (1\ km, 1\ km)$, $(x, z) = (0\ km, 2.5\ km)$, and $(x, z) = (1\ km, 2.5\ km)$. Same question as before, except calculate minimum horizontal and vertical resolutions for a poststack migration image. Show work. In the above example why would it be better to use $.25\lambda z\sqrt{1 + (L/z)^2/L}$ rather than $.25\lambda z/L$ for the horizontal resolution formula?

2. Using fat circles and fat ellipses, compare the vertical and horizontal resolution limits for poststack migration and prestack migration for a point at $(x, z) = (1\ km, 2.5\ km)$ in Figure 3.8. Assume a homogeneous velocity.

3. Convert your poststack depth migration code into a prestack time migration code. The output should be in the offset-time domain. Show poststack depth and time migration images.

4. Determine the maximum aperture for a seismic experiment in order to image $0 - 40$ degree dips at $z = 5$ km. Assume a homogeneous velocity of 5 km/s. What is the minimum geophone spacing in order to not spatially alias the data? Assume a minimum wavelet period of 0.01 s. Clearly show steps in your reasoning.

5. Derive and use the Beylkin stretch formulas (starting point for derivation is equation 3.10) to estimate the best horizontal and vertical resolutions for a 12 km wide poststack depth section at the following coordinates: (0,12), (12,12), (6,12), (6,6), (0,6), (3,3), (0,12). Assume the origin of coordinate system is at upper left portion of migration section and positive z is pointing downward. Assume a homogeneous velocity of $v = 3\ km/s$ from 0-3 km, $v = 5\ km/s$ from 3-7 km, and any deeper than 7 km we have a velocity of 6 km/s. Assume a maximum useful frequency of 80 Hz. Assume straight rays everywhere. Why is there better horizontal resolution at shallow depths?

6. Time migrate the radar data from your lab exercise. Choose a suitable time migration velocity by trial and error. Show results. Which results look more coherent, the time migration or depth migration results? Why does depth migration suffer from migration stretch which is avoided by time migration. Wgat are benefits and liabilities of time migration vs depth migration?

7. Estimate the horizontal wavelength of surface waves and body wave reflections in the Saudi shot gather (from a previous lab). What is a good array interval that would suppress surface waves but retain body waves ion the data. Test your estimate by applying an N-point spatial averaging filter to the data, where N is length of your estimated filter. Show results.

8. Using Figure 3.3, estimate the maximum trace sampling interval that will avoid aliasing the earliest reflections seen in the seismograms. Same question, except the latest-arriving reflections. Show work.

9. Same as previous question, except find the maximum trace sampling intervals that avoid aliasing of the ZO migration operators? Show work.

10. Assume that Figure 3.4 represents a ZO seismic migration section and that it agrees with the exact earth model. Assume a homogeneous velocity of 500 m/s. What trace spacing in the original stacked data would have been required to just avoid data aliasing for the a). dipping reflector, b). flat reflector? Show work.

11. Same as previous question, except what is the trace spacing that would avoid aliasing for the ZO migration operator? Show work.

12. Same as previous question except what is trace spacing that would just avoid aliasing the prestack migration operator?

Chapter 4

Forward and Adjoint Modeling using Green's Functions

Forward modeling of the wave equation is defined as computing the seismograms for a given source-receiver distribution and a velocity model. This chapter shows how to compute these seismograms using Green's theorem, which casts the solution of the wave equation in terms of a boundary integral and a source-related volume integral. These integral solutions are too costly to compute by the standard boundary integral equation method, so a Lippmann-Schwinger (LS) equation is used under the Born approximation. The LS equation computes the seismograms by a model-coordinate summation of weighted Green's functions, where the weights are the reflection-like coefficients in the model. Another name for this is diffraction-stack modeling. If the adjoint of the Green's function is used (complex conjugate of the Green's function), then the migration image is obtained by a data-coordinate summation of the weighted adjoint Green's functions, where the weights are the seismograms. Another name for this is diffraction-stack migration. These methods are used in later chapters that describe least squares migration and waveform tomography.

The mathematics for the forward modeling equations can be somewhat complicated so that it is easy to lose the physical meaning and motivation for forward modeling. Thus, the next section intuitively derives the equations for convolutional forward modeling of a 1D earth model. Later sections use Green's theorem to generalize these equations to arbitrary media.

4.1 1D Convolutional Modeling

The motivation for modeling waves, i.e., solving the wave equation, is to make a direct connection between geology and the wiggles seen in the recorded seismograms. For example, the sonic and density logs can be used to estimate the 1D layered model of the earth and from these the associated zero-offset (or stacked) seismograms are computed by a forward modeling method. The reflections from a, say, sand-shale interface can be identified in the synthetic trace and correlated to the corresponding event in the recorded stacked sections. Thus, the lithology of each wiggle in the recorded data can be geologically, in principle, identified. This identification will tremendously increase the chances for a correct geological

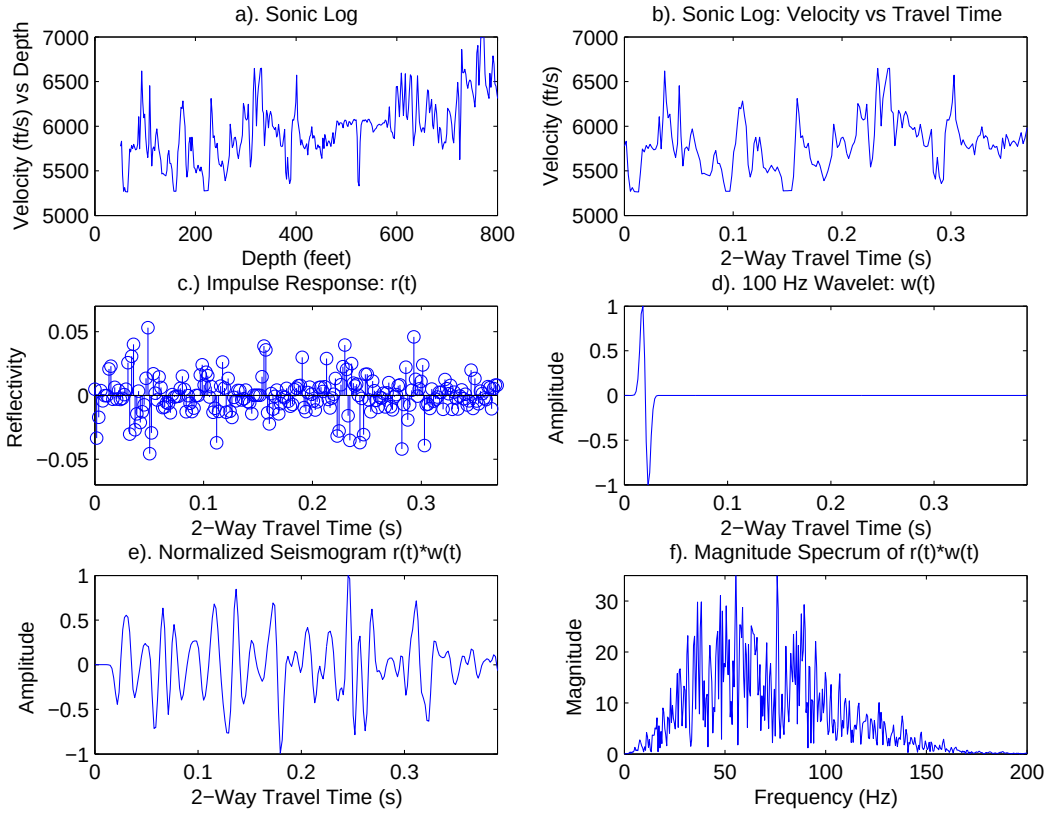


Figure 4.1: a). Sonic log vs depth from Texas, b). sonic log vs 2-way traveltime, c). reflectivity vs time, d). source wavelet vs time, e). normalized seismogram $s(t) = r(t) \star w(t)$, f). magnitude spectrum of seismogram.

interpretation of the seismic traces.

The simplest forward modeling procedure is known as the 1D convolutional model of the earth. In this procedure, a 1D layered model is assumed and parameterized by the density $\rho(z)'$ and velocity $v(z)'$ as a function of depth, as shown in Figure 4.1a. These two functions undergo a depth to 2-way time transformation using the following coordinate transformation:

$$t(z) \quad = \quad 2 \int_0^z dz'/v(z')', \tag{4.1}$$

where $t(z)$ is the 2-way propagation time for energy to go vertically downward from the surface to the horizontal reflector at depth z and back up to the surface in the 1D layered model. A MATLAB script for this mapping from depth to time is given as

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Finds t(z) from v(z). Assumes
% v(z) starts at free surface.
% v(z)   - input- sonic log as function of z
% dz     - input- depth sampling interval of sonic log
```

```
% t(z)   - input- 2-way time as function of z
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function [time]=depth2time(v,dz)
nz=length(v);time=zeros(nz,1); time(1)=dz/v(1);
for i=2:nz; time(i)=time(i-1)+dz/v(i); end
time=time*2; plot(dz*[1:nz],time);
xlabel('Depth (ft)'); ylabel('Time (s)')
title('Depth vs 2-way Time');figure
plot(time,v);xlabel('Time (s)');ylabel('Velocity (ft/s)')
```

The velocity model as a function of time $v(t)$ (see Figure 4.1b) is usually unevenly sampled. To perform convolutional forward modeling, we must convert to an evenly sampled function in time $v(t) = v(t(z))'$; the MATLAB code for getting an even sampled function sampled at the sampling interval dt from an unevenly sampled function is given in Appendix D.

Assuming that the velocity function $v(t)$ is now an evenly sampled function, the evenly sampled zero-offset (ZO) reflection coefficients as a function of time can be estimated by $r(t) = (\rho(t)v(t) - \rho(t-dt)v(t-dt))/((\rho(t)v(t) + \rho(t-dt)v(t-dt))$, which in MATLAB script becomes for constant density:

```
y=diff(vpp);nl=length(y);dt=diff(time);add=vpp(1:nl)+vpp(2:nl+1);
rc=y(1:nl)./add(1:nl);stem(time(1:nl),rc(1:nl));
```

If the density profile is known then the density can be put into the above reflection coefficient formula.

The bandlimited response of the medium for a plane wave input (with source wavelet $w(t)$ as shown in Figure 4.1d) into the surface is a combination of arrivals, including primary and multiple reflections. If attenuation, transmission losses and multiples are excluded then the 1D convolutional model of the seismogram $s(t)$ is given by

$$s(t) = \int_{-\infty}^{\infty} r(\tau)w(t-\tau)d\tau, \tag{4.2}$$

which is the definition of convolution of $r(t)$ with $w(t)$, often abbreviated as $s(t) = r(t) \star w(t)$.

The above formula can be derived by taking the special case of the impulse response where the source wavelet is a Dirac delta function that is excited at time equal to zero: $w(t) = \delta(t)$, where the delta function is defined as $\delta(t) = 0$ if $t \neq 0$, otherwise $\delta(t) = 1$ in the sense $\int \delta(t)dt = 1$. Plugging this impulse wavelet into the above equation yields:

$$\begin{aligned} s(t) &= \int_{-\infty}^{\infty} r(\tau)\delta(t-\tau)d\tau, \\ &= r(t), \end{aligned} \tag{4.3}$$

which describes the reflection coefficient series shown in Figure 4.1c. Thus, the 1D impulse response of the earth under the above assumptions perfectly describes the reflection coefficient series as a function of 2-way traveltime. If the source wavelet were weighted by the scalar weight $w(\tau_i)$ and delayed by time τ_i then $w(t) = w(\tau_i)\delta(t-\tau_i)$ then the delayed impulse response of the earth would be

$$\begin{aligned} s(t)' &= \int_{-\infty}^{\infty} r(\tau)w(\tau_i)\delta(t-\tau_i-\tau)d\tau, \\ &= w(\tau_i)r(t-\tau_i), \end{aligned} \tag{4.4}$$

which is a weighted delayed version of the original impulse response in equation 4.2. If we were to sum these two seismograms we would get, by linearity of integration,

$$s(t)' + s(t) \quad = \quad w(\tau_0)r(t-\tau_0) + w(\tau_i)r(t-\tau_i), \tag{4.5}$$

where $\tau_0 = 0$ and $w(\tau_0) = 1$. By the superposition property of waves (i.e., interfering wave motions add together), we could have performed these two seismic experiments at the same time and the resulting seismograms would be identical mathematically to equation 4.5. More generally, the earth's resposne to an arbitrary wavelet $w(\tau)$ is given by

$$\begin{aligned} s(t) &= \sum_i w(\tau_i)r(t-\tau_i), \\ &\approx \int_{-\infty}^{\infty} w(\tau)r(t-\tau)d\tau, \end{aligned} \tag{4.6}$$

and in the limit of vanishing sampling interval dt the approximation becomes an equality (see Figure 4.1e). Under the transformation of variables $\tau' = t - \tau$ equation 4.6 becomes

$$s(t) \quad = \quad \int_{-\infty}^{\infty} w(t-\tau')r(\tau')d\tau', \tag{4.7}$$

which is precisely the convolution equation shown in equation 4.2. The equality of equations 4.6 and 4.7 also shows that convolution commutes, i.e., $s(t) = r(t) \star w(t) = w(t) \star r(t)$. The convolutional modeling equation was practically used by many oil companies starting in the 1950's, and is still in use today for correlation of well logs to surface seismic data.

4.1.1 Multiples

Multiples associated with a strong reflector and the free surface can be accounted for in the 1D modeling equations. For a sea-bottom with depth d and two-way ZO traveltime τ_w, the sea-bed impulse response for ZO downgoing pressure waves a source and pressure receiver just below the sea surface is given by

$$m(t) \quad = \quad w(t) + \sum_{i=1}^{\infty} (-R)^i w(t - i\tau_w) \tag{4.8}$$

where R is the ZO reflection coefficient of the sea floor; the -1 accounts for the free surface reversal of polarity and $w(t)$ is the source wavelet of the airgun modified by the interaction with the sea-surface reflectivity. We assume that the propagation time between the surface and hydrophone streamer is negligible compared to the propagation time from the surface to the sea floor. as illustated in Figure 4.2a.

The upgoing multiples each act as a secondary source on the sea surface, so we can consider the "generalized" source wavelet to be $m(t)$. Thus the response of the medium is given by

$$s(t) \quad = \quad r(t) \star m(t). \tag{4.9}$$

These multiples tend to blur the reflectivity response so we should try to deconvolve the multiples.

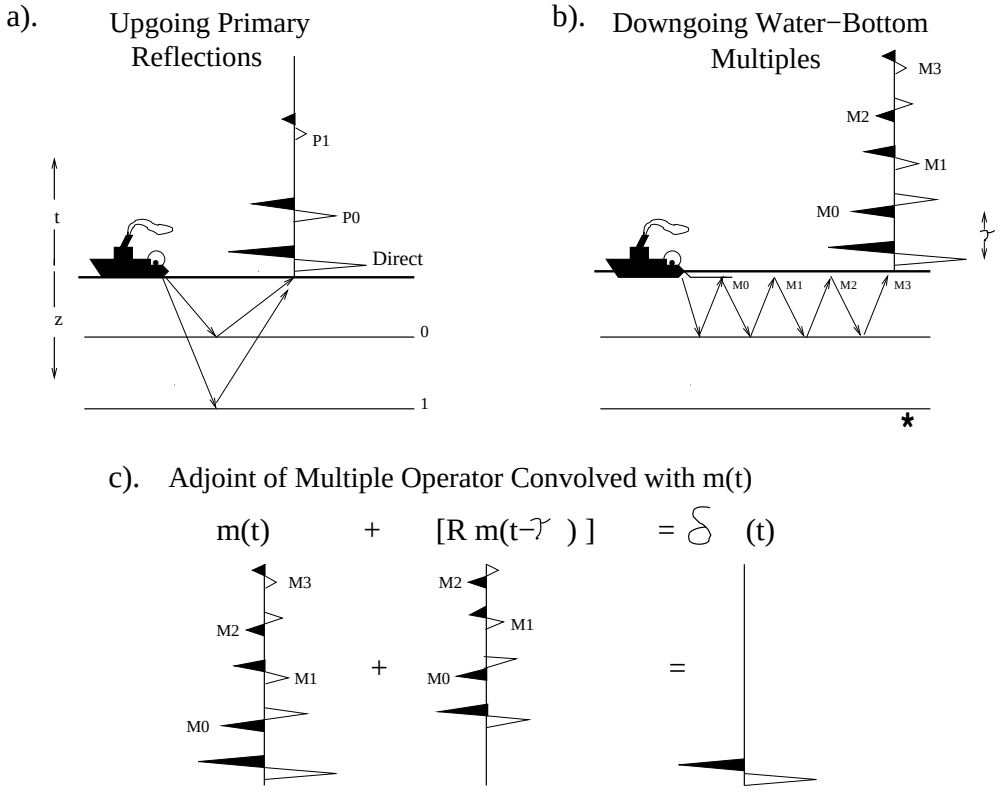


Figure 4.2: a). Upgoing primary reflections, b). downgoing sea-floor multiples only and c). adjoint of multiple operator applied to multiples to give a quasi-spike at $t = 0$. Here, vertical incidence angles are assumed in a layer-cake model and the wavelet in c). is a delta function.

The "deblurring" or deconvolution operator can be found in the frequency domain by taking the Fourier transform of equation 4.8:

$$\begin{aligned} M(\mathbf{Z}) &= W(\mathbf{Z})(1+\sum_{i=1}^{\infty}(-R)^i\mathbf{Z}^{i\tau_w}), \\ &= \frac{W(\mathbf{Z})}{1+R\mathbf{Z}^{\tau_w}}, \end{aligned} \tag{4.10}$$

where $\mathbf{Z} = e^{-j\omega}$. Thus, the free-surface demultiple operator for downgoing waves in equation 4.9 is given in the frequency domain by

$$\begin{aligned} M(\mathbf{Z})^{-1} &= (1+R\mathbf{Z}^{\tau_w})/W(\mathbf{Z}), \\ &= (1+R\mathbf{Z}^{\tau_w})W(\mathbf{Z})^*/|W(\mathbf{Z})|^2, \end{aligned} \tag{4.11}$$

or in the time domain as

$$m(t)^{-1} = (1+R\delta(t-\tau_w)) * w(t)^{-1}. \tag{4.12}$$

as illustrated in Figure **??**c. The deblurring operator says that the seismogram must be added to a weighted-delayed copy of itself, where the weighting is R and the delay is the two-way propagation time through the water layer. In addition, the wavelet inverse must be convolved with this result.

4.2 1D Adjoint Convolutional Modeling

The adjoint operator is sometimes a good approximation to the inverse operator, as explained in the previous chapter. For equation 4.7, the inverse can be approximated by applying the adjoint operator $\int_{-\infty}^{\infty} w(t+t')dt$ to $s(t)$, i.e.,

$$r(\tau) \approx \int_{-\infty}^{\infty} w(\tau+t)s(t)dt. \tag{4.13}$$

This can be seen to be a good approximation by substituting $s(t) = \int_{-\infty}^{\infty} r(\tau')w(t-\tau')d\tau'$ to give

$$\begin{aligned} r(\tau) &\approx \int_{-\infty}^{\infty}[\int_{-\infty}^{\infty} w(t-\tau')w(\tau+t)dt]r(\tau')d\tau', \\ &= \int_{-\infty}^{\infty}\phi_{ww}(\tau-\tau')r(\tau')d\tau', \\ &\approx r(\tau). \end{aligned} \tag{4.14}$$

where the approximation follows if the wavelet is a vibroseis chirp over a wideband of frequencies so that $\phi_{ww}(\tau-\tau') \approx \delta(\tau-\tau')$. In this case we say that $\int_{-\infty}^{\infty} w(t-t')dt$ is an approximate inverse operator or we can say it is the adjoint modeling operator.

4.2.1 1D Adjoint for Multiples

The inverse to the free-surface multiple operator in equation 4.10 can be approximated by its adjoint operator

$$\begin{aligned} M(\mathbf{Z})^{-1} &\approx \frac{W(\mathbf{Z})^*}{1+R\mathbf{Z}^{-\tau_w}}, \\ &= \frac{W(\mathbf{Z})^*(1+R\mathbf{Z}^{\tau_w})}{|1+R\mathbf{Z}^{-\tau_w}|^2}, \end{aligned} \tag{4.15}$$

which has exactly the same phase spectrum as the actual inverse in equation 4.11. Their magnitude spectrums only differ by the absence of the $|W(\mathbf{Z})|^2$ in the denominator of equation **??** and the inclusion of the $|1+R\mathbf{Z}^{-\tau_w}|^2$ term. These terms can be harmlessly ignored if they have a constant value over the frequency band of the signal.

.

4.3 3D Integral Equation Forward Modeling

The previous section assumed no geomteric spreading losses, no multiples, a layered media and no transmission losses. These assumptions are unrealistic for many purposes, so we must learn how to solve the wave equation for arbitrary acoustic media using Green's theorem. Extensions to anelastic media are straightforward.

The 3D Helmholtz equation is given by

$$(\nabla^2_{\mathbf{x}'} + \omega^2 s(\mathbf{x}')^2)P(\mathbf{x}') = F(\mathbf{x}'), \tag{4.16}$$

where the wavenumber $k = \omega/c(\mathbf{x})$ is for an inhomogeneous medium with velocity $c(\mathbf{x})$, $F(\mathbf{x})$ is the source term associated with a harmonically oscillating source at $\mathbf{x}$, and $P(\mathbf{x}')$ is the associated pressure field.

The goal of forward modeling is, given the source-receiver coordinates and the velocity model, use Green's theorem to find the pressure field in the form of an integral equation. Towards this goal, we first define a Green's function and then use it to derive Green's theorem.

4.3.1 Green's Functions

The Green's function associated with the 3-D Helmholtz equation for an arbitrary medium (Morse and Feshback, 1953) solves

$$(\nabla^2_{\mathbf{x}'} + k^2)G(\mathbf{x}'|\mathbf{x}) = \delta(\mathbf{x}'-\mathbf{x}), \tag{4.17}$$

where $k = \omega/c(\mathbf{x})$ and $\delta(\mathbf{x}'-\mathbf{x}) = \delta(x-x')\delta(y-y')\delta(z-z')$.

There are two independent solutions to this 2nd-order PDE: an outgoing Greens function $G(\mathbf{x}|\mathbf{x}')$ and its complex conjugate the incoming Green's function $G(\mathbf{x}|\mathbf{x}')^*$. The outgoing Green's function (see Appendices A and B) for a homogeneous medium with velocity c_0 is given by

$$G(\mathbf{x}|\mathbf{x}') = -\frac{1}{4\pi}e^{-ik_0|\mathbf{x}'-\mathbf{x}|}/|\mathbf{x}-\mathbf{x}'|, \tag{4.18}$$

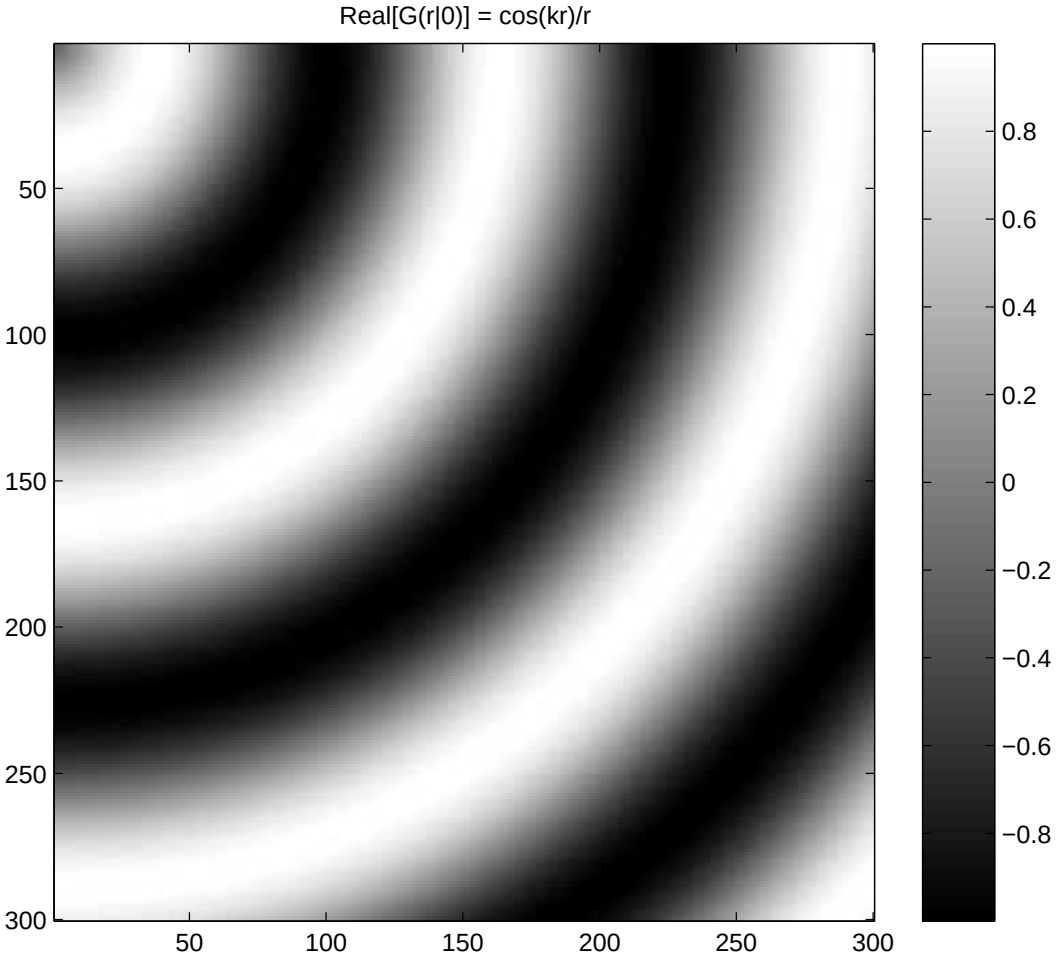


Figure 4.3: Snapshot of the real part of the harmonic Green's function, where the geometrical spreading factor has been ignored.

where the wavenumber is given by $k_0 = \omega/c_0$, the denominator represents the geometrical spreading factor while the exponential phase factor is proportional to the with distance between the observer $\mathbf{x}$ and the source point $\mathbf{x}'$. The interpretation of the Green's function is that it is the acoustic response measured at $\mathbf{x}$ for a harmonically oscillating point source located at $\mathbf{x}'$.

For a homogeneous medium the Green's function plots out as a series of concentric circles centered about $\mathbf{x}'$ with wavelength $\lambda_o = 2\pi c_o/\omega$ (see Figure 4.3). Note, the source and receiver locations can be interchanged in equation 4.18 so that $G(\mathbf{x}|\mathbf{x}') = G(\mathbf{x}'|\mathbf{x})$, which is the reciprocity principle. This says that a trace recorded at position A excited by a source at B will be the same as the trace located at B for a source excitation at A.

4.3.2 $(\nabla^2_{\mathbf{x}'} + k^2)^{-1}$ by Green's theorem

Green's theorem will be used to invert equation 4.16. Multiplying the Helmholtz equation by the Green's function gives

$$G(\mathbf{x}|\mathbf{x}')(\nabla^2_{\mathbf{x}'} + k^2)P(\mathbf{x}') \quad = \quad G(\mathbf{x}|\mathbf{x}')F(\mathbf{x}'), \tag{4.19}$$

where the differentiation in $\nabla_{\mathbf{x}'}$ is with respect to the primed coordinates. Inserting the identity $G\nabla^2_{\mathbf{x}'}P = P\nabla^2_{\mathbf{x}'}G + \nabla_{\mathbf{x}'} \cdot (G\nabla_{\mathbf{x}'}P - P\nabla_{\mathbf{x}'}G)$ (which can be proved using differentiation by parts) into the above equation gives

$$P(\mathbf{x}')(\nabla^2_{\mathbf{x}'} + k^2)G(\mathbf{x}|\mathbf{x}') + \nabla_{\mathbf{x}'} \cdot (G\nabla_{\mathbf{x}'}P - P\nabla_{\mathbf{x}'}G) = G(\mathbf{x}|\mathbf{x}')F(\mathbf{x}'). \tag{4.20}$$

Integrating equation 4.20 over the entire model volume yields

$$\int_{vol}[P(\mathbf{x}')(\nabla^2_{\mathbf{x}'} + k^2)G(\mathbf{x}|\mathbf{x}') + \nabla_{\mathbf{x}'} \cdot (G\nabla_{\mathbf{x}'}P - P\nabla_{\mathbf{x}'}G)]dV' = \int_{vol} G(\mathbf{x}|\mathbf{x}')F(\mathbf{x}')dV' \tag{4.21}$$

and inserting equation 4.17 yields

$$\int_{vol} P(\mathbf{x}')\delta(\mathbf{x}' - \mathbf{x})dV' + \int_{vol} \nabla_{\mathbf{x}'} \cdot (G\nabla_{\mathbf{x}'}P - P\nabla_{\mathbf{x}'}G)dV' = \int_{vol} G(\mathbf{x}|\mathbf{x}')F(\mathbf{x}')dV', \tag{4.22}$$

or

$$P(\mathbf{x}) + \int_{surf} (G\nabla_{\mathbf{x}'}P - P\nabla_{\mathbf{x}'}G) \cdot d\vec{A}', = \int_{vol} G(\mathbf{x}|\mathbf{x}')F(\mathbf{x}')dV', \tag{4.23}$$

where the unit normal is pointing outward from the body. The conversion from a volume integral to a surface integral follows by the divergence theorem (Morse and Feshbach, 1953; Barton, 1989). Rearranging gives the integral equation solution to the Helmholtz equation:

$$P(\mathbf{x}) \quad = \quad -\int_{surf} (G\nabla_{\mathbf{x}'}P - P\nabla_{\mathbf{x}'}G) \cdot d\vec{A}' + \int_{vol} G(\mathbf{x}|\mathbf{x}')F(\mathbf{x}')dV'. \tag{4.24}$$

The above integral is the starting point for the Boundary Integral Equation modeling method (Brebbia, 1978; Schuster, 1985 and many others).

If only outgoing waves are considered and the surface integral is at infinity then the surface integration is zero in equation 4.25 to give

$$P(\mathbf{x}) \quad = \quad \int_{vol} G(\mathbf{x}|\mathbf{x}')F(\mathbf{x}')dV'. \tag{4.25}$$

The physical interpretation of this term is that it represents the direct wave. Therefore, the boundary integral in equation 4.24 represents the boundary reflections and echoes that contribute to the field inside the volume.

This solution of the Helmholtz equation given by equation 4.25 is a weighted sum of the Green's functions, where the weights are the source amplitudes at each point in the medium. This representation is valid for an arbitrary acoustic medium, and shows that $\int dV' G(\mathbf{x}|\mathbf{x}')\cdot = (\nabla^2_{\mathbf{x}'} + k^2)^{-1}$. In other words, applying the *appropriate* Green's function to the Helmholtz equation and integrating over the volume is the inverse operator to the Helmholtz equation.

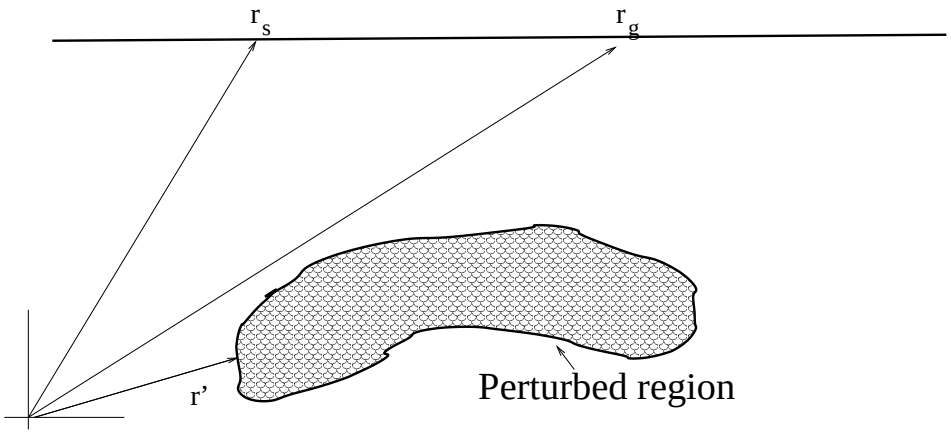


Figure 4.4: Background medium with an embedded potato-like perturbation of slightly different velcoity.

4.3.3 Lippmann-Schwinger Solution

Equation 4.24 is expensive to solve for the unknown field values $P(\mathbf{x})$ because they are on both the left and right hand sides, which means an expensive matrix inverse (Schuster, 1985) must be computed. To avoid this expense we use the Lippmann-Schwinger equation under the Born approximation, as described below.

Assume a medium composed of the background slowness and a perturbed slowness given by $s(\mathbf{x}) + \delta s(\mathbf{x})$, where $\delta s(\mathbf{x})$ is sufficiently small. For example, Figure 4.4 shows the background medium as homogeneous and the perturbed region as potato-shaped with a slightly different velocity.

The pressure field P in the perturbed medium is governed by

$$(\nabla^2_{\mathbf{x}'} + \omega^2[s(\mathbf{x}') + \delta s(\mathbf{x}')]^2)P(\mathbf{x}') \quad = \quad F(\mathbf{x}'), \tag{4.26}$$

or by rearranging we get

$$(\nabla^2_{\mathbf{x}'} + \omega^2 s(\mathbf{x}')^2)P(\mathbf{x}') \quad = \quad -2\omega^2 s(\mathbf{x}')\delta s(\mathbf{x}')P(\mathbf{x}') + F(\mathbf{x}'), \tag{4.27}$$

where second-order terms in the perturbation parameter are neglected. The above Helmholtz equation can be inverted by applying $\int dV' G(\mathbf{x}|\mathbf{x}')\cdot$ to both sides of the above equation to

give the Lippmann-Schwinger equation:

$$P(\mathbf{x}) = -2\omega^2 \int_{vol} \overbrace{G(\mathbf{x}|\mathbf{x}')}^{Upgoing\ field} \overbrace{s(\mathbf{x}')\delta s(\mathbf{x}')}^{Reflectivity} \overbrace{P(\mathbf{x}')}^{Downgoing\ field} dV' + \int_{vol} G(\mathbf{x}|\mathbf{x}')F(\mathbf{x}')dV', \tag{4.28}$$

where the Green's function is that for the unperturbed medium and satisfies equation 4.17 for $k = \omega s(x)$. The term $P(\mathbf{x}')$ represents the total field on the potato exited from an overlying source distribution so we call this the *downgoing field*. Similarly, the Green's function $G(\mathbf{x}|\mathbf{x}')$ takes the incident field on the potato and extrapolates it to the surface as an *upgoing field*; the strength of this upgoing field is proportional to the magnitude of the $reflectivity - like$ term $s(\mathbf{x}')\delta s(\mathbf{x}')$.

Note, the unknown field value $P(\mathbf{x}')$ is on both the left and right hand side of equation 4.28; this means that this unknown can be determined by discretizing the field values, the slowness distribution and replacing the integral by a numerical quadrature to give a large system of integral equations, similar to the boundary integral equation method (Brebbia, 1978).

4.3.4 Neumann Series Solution

An alternative to the direct inversion of this system of equations is an iterative solution, such as the Neummann series solution to equation 4.28. A Neumann series solution is similar to the scalar geometric series

$$(1-x)^{-1} = 1 + x + x^2 + x^3 + ... \tag{4.29}$$

that converges for $||x|| < 1$, except x represents an operator. Rearranging equation 4.28 so that the terms in the unknown $P(\mathbf{x}')$ equation 4.28 are on the lefthand side gives

$$[\int_{vol} \delta(\mathbf{x}' - \mathbf{x})dV' + 2\omega^2 \int_{vol} G(\mathbf{x}|\mathbf{x}')s(\mathbf{x}')\delta s(\mathbf{x}')dV']P(\mathbf{x}') = \int_{vol} G(\mathbf{x}|\mathbf{x}')F(\mathbf{x}')dV', \tag{4.30}$$

so that the operator x in the Neumann series here is identified as

$$x \rightarrow -2\omega^2 \int_{vol} G(\mathbf{x}|\mathbf{x}')s(\mathbf{x}')\delta s(\mathbf{x}')dV'. \tag{4.31}$$

Therefore, the Neumann series solution to equation 4.28 is

$$P(\mathbf{x}) = \int_{vol} \overbrace{G(\mathbf{x}|\mathbf{x}')F}^{\psi_0}(\mathbf{x}')dV' - 2\omega^2 \int_{vol} \overbrace{G(\mathbf{x}|\mathbf{x}')s(\mathbf{x}')\delta s(\mathbf{x}') \int_{vol} G(\mathbf{x}'|\mathbf{x}'')F(\mathbf{x}'')}^{\mathcal{G}_0 V \psi_0} dV'dV'' +$$

$$4\omega^4 \int_{vol} \overbrace{G(\mathbf{x}|\mathbf{x}')s(\mathbf{x}')\delta s(\mathbf{x}') \int_{vol} G(\mathbf{x}'|\mathbf{x}'')s(\mathbf{x}'')\delta s(\mathbf{x}'') \int_{vol} G(\mathbf{x}''|\mathbf{x}''')F(\mathbf{x}''')}^{\mathcal{G}_0 V \mathcal{G}_0 V \psi_0} dV' dV'' dV''' + ...$$

$$= \int_{vol} G(\mathbf{x}|\mathbf{x}')F(\mathbf{x}')dV' + \sum_{i=1}^{\infty}[-2\omega^2 \int_{vol} G(\mathbf{x}|\mathbf{x}')s(\mathbf{x}')\delta s(\mathbf{x}')dV']^i \int_{vol} G(\mathbf{x}'|\mathbf{x}'')F(\mathbf{x}'')dV'';$$

or in more compact notation we can represent the above equation by

$$P = \sum_{i=0}^{\infty}[\mathcal{G}_0 V]^i \psi_0 = P^{(0)} + P^{(1)} + P^{(2)} + ... , \tag{4.32}$$

where the *nth* term represents the *nth* order response of the perturbed medium. For example, the zeroth-order term $P^{(0)} = \int_{vol} G(\mathbf{x}|\mathbf{x}')F(\mathbf{x}')dV'$ is the direct-wave response of the background medium, the first-order term $P^{(1)}$ is the primary $reflection$ response between the *potato* and the background medium, the second-order term $P^{(2)}$ is its 1st-order multiple response, and so on. The first-order term $P^{(1)}$ is also known as the Born approximation, as discussed in the next section.

4.3.5 Born Approximation

The Born approximation is nothing more than approximating the scattered field $P - P_0$ by the first-order term of the Born series in equation 4.36:

$$P - P^{(0)} \approx P^{(1)}. \tag{4.33}$$

The physical meaning of this approximation is now gievn.

If the perturbed field is "close" to that of the unperturbed field and if the actual source is a point source located at $\mathbf{x}_s$, then $P(\mathbf{x}')$ on the RHS of equation 4.28 can be replaced by $G(\mathbf{x}'|\mathbf{x}_s)W(\omega)$ (where $W(\omega)$ is the source spectrum) to give the Born approximation to the Lippmann-Schwinger equation:

$$\begin{aligned} P(\mathbf{x}|\mathbf{x}_s) &= -2\omega^2 \int_{vol} G(\mathbf{x}|\mathbf{x}')dV's(\mathbf{x}')\delta s(\mathbf{x}')W(\omega)G(\mathbf{x}'|\mathbf{x}_s) \\ &\quad + \int_{vol} G(\mathbf{x}|\mathbf{x}')F(\mathbf{x}')dV', \end{aligned} \tag{4.34}$$

where the first term on the RHS accounts for the primary reflection from the perturbation and the second term accounts for the direct wave in the unperturbed medium. The approximation is valid if multiples inspired by δs can be neglected.

Example of Non-Zero Offset Modeling: We will now derive the diffraction stack modeling equation for a smoothly varying velocity model at high source frequencies, and for a non-zero offset (ZO) configuration where each source shoots into receivers with finite offset.

Assume high frequencies so that the asymptotic Green's function (Bleistein, 1984) can be used:

$$G(\mathbf{x}|\mathbf{x}_s) = \frac{e^{-i\omega\tau_{sr}}}{r}, \tag{4.35}$$

where τ_{sr} is the time for energy to propagate from the source at $\mathbf{x}_s$ to the interrogation point at $\mathbf{x}$, and $1/r_{sr} = 1/|\mathbf{x} - \mathbf{x}_s|$ is the geometrical spreading term that approximates the solution to the transport equation. For convenience we will ignore the $-1/4\pi$ factor.

Substituting equation 4.35 into equation 4.34 we get the diffraction stack equation of forward modeling:

$$P(\mathbf{x}|\mathbf{x}_s)^{scatt} = -2\omega^2 \int_{vol} s(\mathbf{x}')\delta s(\mathbf{x}')W(\omega)e^{-i\omega(\tau_{sr'}+\tau_{rr'})}/(|\mathbf{x}' - \mathbf{x}_s||\mathbf{x} - \mathbf{x}'|)dV', \tag{4.36}$$

where the direct wave is neglected to give the scattered field $P(\mathbf{x}|\mathbf{x}_s)^{scatt}$. Applying an inverse Fourier transform $\int d\omega e^{i\omega t}$ to the above equation yields the time-domain diffraction stack modeling formula:

$$p(\mathbf{x}, t|\mathbf{x}_s, 0)^{scatt} = -2 \int_{vol} s(\mathbf{x}')\delta s(\mathbf{x}')\ddot{w}(t - \tau_{sr'} - \tau_{rr'})/(|\mathbf{x}' - \mathbf{x}_s||\mathbf{x} - \mathbf{x}'|)dV'. \tag{4.37}$$

This formula can be used to generate synthetic seismograms of primary reflections by setting $\delta s(\mathbf{x}) = 1$ along reflector boundaries, otherwise $\delta s(\mathbf{x}) = 0$.

A special case is for a point scatterer $\delta s(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}_0)$ and a geophone at $\mathbf{x} \to \mathbf{x}_g$ so that equation 4.37 becomes

$$p(\mathbf{x}_g, t|\mathbf{x}_s, 0)^{scatt} = -2s(\mathbf{x}_0)\delta s(\mathbf{x}_0)\ddot{w}(t - \tau_{sr_0} - \tau_{gr_0})/(|\mathbf{x}_0 - \mathbf{x}_s||\mathbf{x}_g - \mathbf{x}_0|). \tag{4.38}$$

Here, $w(t)$ describes the time history of the source wavelet, and $w(t-\tau_{sr_0}-\tau_{gr_0})$ is the source wavelet delayed by the amount of time it takes to go from the source down to the scatterer and up to the receiver. The support of $w(t - \tau_{sr_0} - \tau_{gr_0})$ exists along the hyperbola shown in Figure 4.6 because $t = \tau_{sr_0} + \tau_{gr_0} = [\sqrt{(x_0 - x_s)^2 + {z_0}^2} + \sqrt{(x_g - x_0)^2 + {z_0}^2}]/c$ traces out a hyperbola in data-space coordinates $(x_g, 0, t)$. Thus the point scatterer response is obtained by smearing the reflectivity $\delta s(\mathbf{x})$ amplitude along the appropriate hyperbola in data space.

For a string of contiguous point scatterers the seismograms are obtained by *smearing* and *summing* the perturbation amplitudes along the appropriate hyperbolas, one for each scatterer as shown in Figure 4.5.

Example of Zero-Offset Modeling: We will now derive the diffraction stack modeling equation for a zero-offset (ZO) configuration, where each source shoots into only the receiver located at the source point. Recall, after nomral moveout correction of common midpoint gathers and stacking, the resulting section is somewhat equivalent to a ZO section.

When the source is coincident with the geophone then $\tau_{s\mathbf{x}} = \tau_{g\mathbf{x}}$ and equation 4.38 becomes

$$p(\mathbf{x}_s, t|\mathbf{x}_s, 0)^{scatt} = -2s(\mathbf{x}_0)\delta s(\mathbf{x}_0)\ddot{w}(t - 2\tau_{sr_0})/|\mathbf{x}_0 - \mathbf{x}_s|^2. \tag{4.39}$$

In this case the data are smeared along the hyperbola with an apex at the source coordinate.

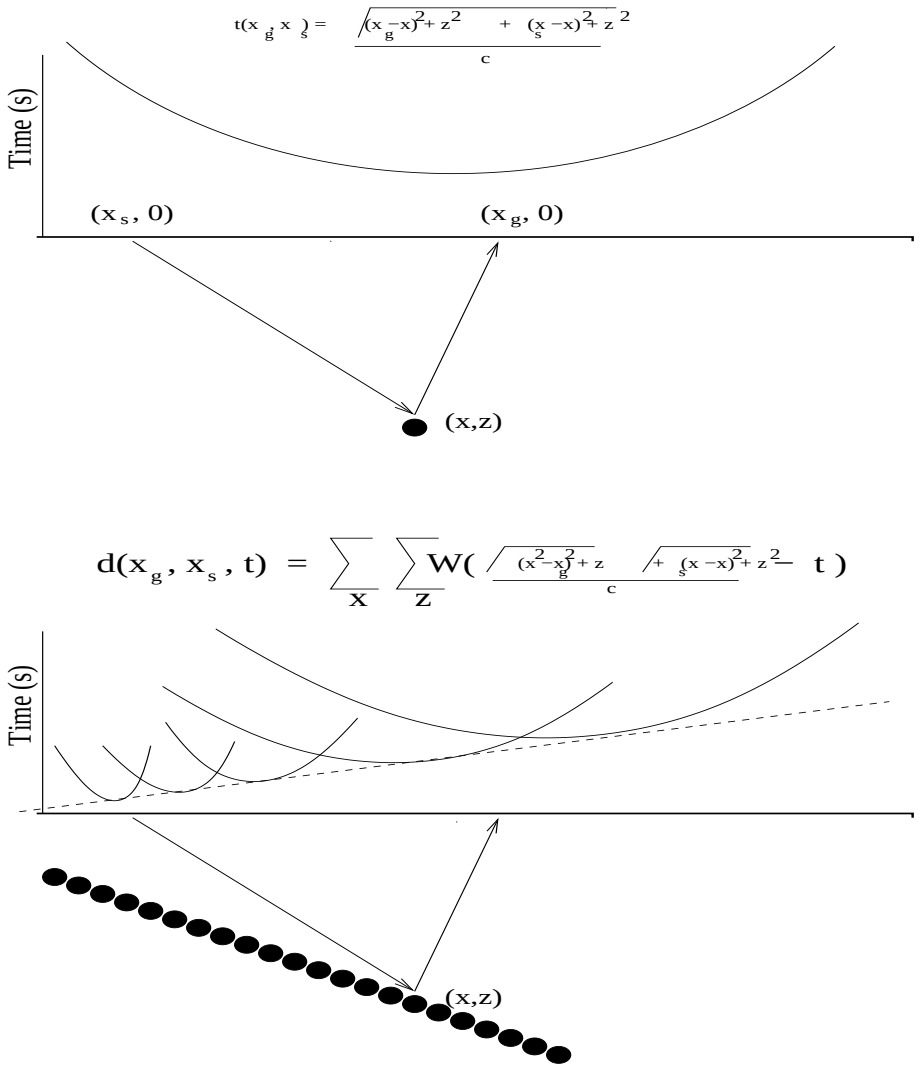


Figure 4.5: Shot gather of traces for a (top) point scatterer and a (bottom) dipping reflector. Bottom formula is that for Born forward modeling, except geometric spreading effects have been neglected. Each reflectivity point is mapped (i.e., smeared) from the model space point (x, z) into the data space as a shifted weighted wavelet, and these events falls along a hyperbola and summed to give the shot gather traces.

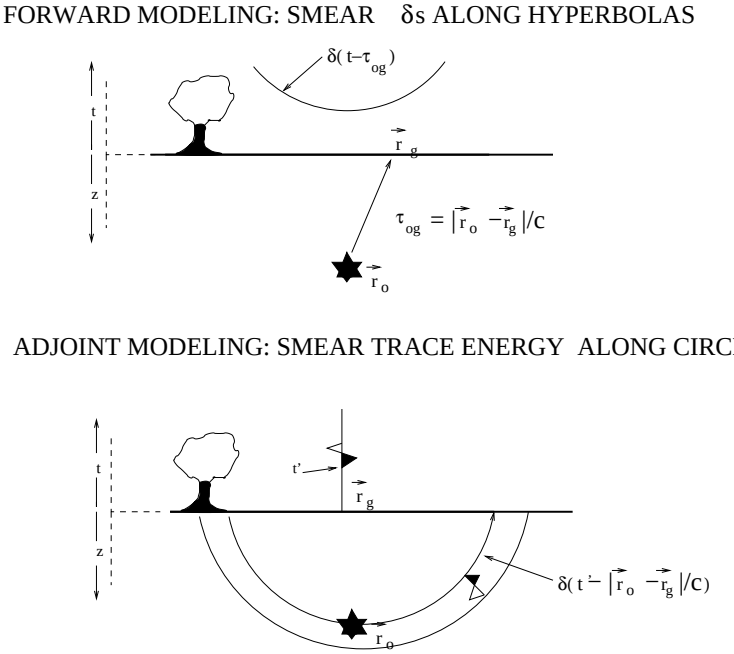


Figure 4.6: (Top) Forward modeling of a point scatterer consists of smearing reflectivity energy along the appropriate hyperbola in data space, and (bottom) adjoint modeling consists of smearing reflection energy in an impulsive trace along the appropriate circle in model space. Zero-offset traces are assumed where the source is coincident with the receiver.

More generally, a model with an interface can be approximated by a sequence of contiguous point scatterers, so that the data are given by the summed contribution of all the point scatterers, i.e.,

$$\begin{aligned} p(\mathbf{x}_s, t|\mathbf{x}_s, 0)^{scatt} &= \sum_x \sum_z r(x,z)\ddot{w}(t = 2\sqrt{(x-x_s)^2 + z^2}/c)/A \\ & \textit{where} \\ A(x,z,x_s,x_s) &= ||\mathbf{x_s} - \mathbf{x}||^2. \end{aligned} \tag{4.40}$$

This formula includes primary reflections, but excludes multiples and angle-dependent reflection coefficients. It is known as the Born modeling formula, and can be described as smearing+summing the reflectivity values $r(x,z)$ along the appropriate hyperbolas in data space. For an inhomogeneous medium, ray tracing can be used to compute a traveltime table that replaces the square root expression in equation 4.40.

The MATLAB code for zero-offset (i.e., the source position is the same as the receiver position) diffraction stack modeling is given below. The top of Figure 4.7 depicts a zero-offset seismic section generated by this code.

```
% ZERO-OFFSET DIFFRACTION STACK FORWARD MODELING
```

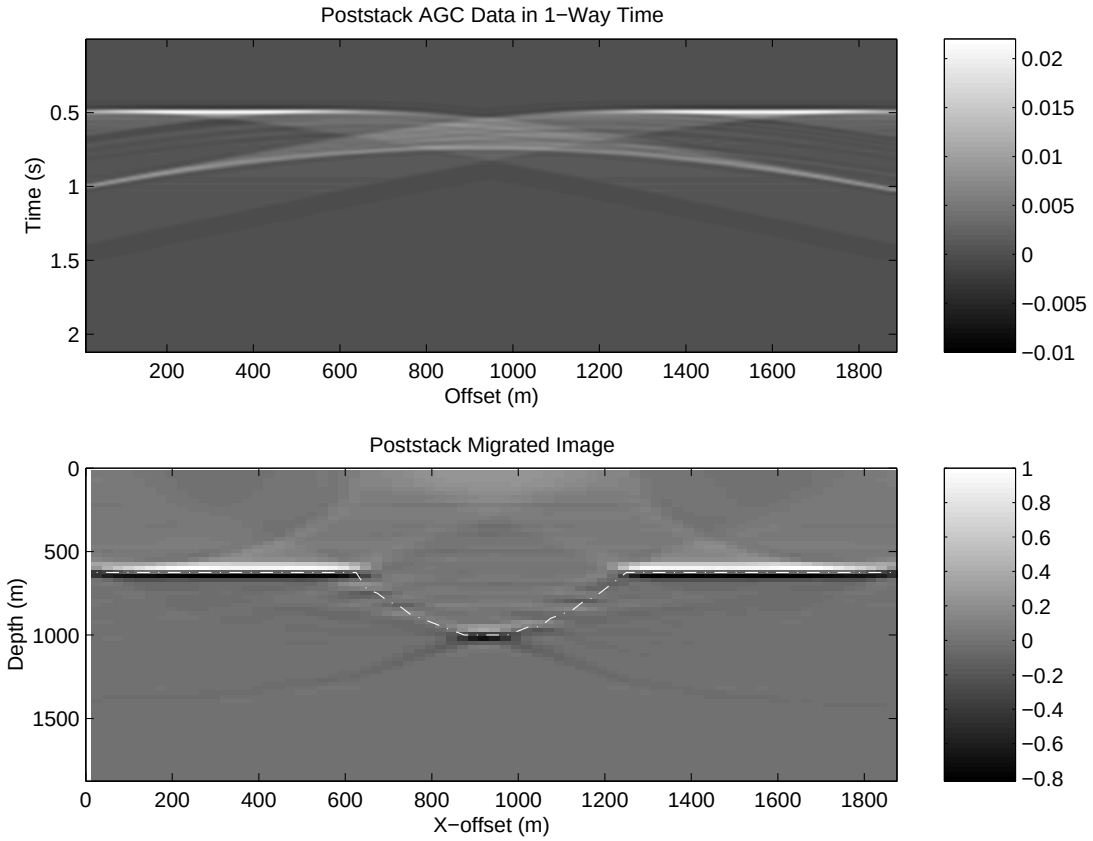


Figure 4.7: (Top) Poststack seismograms and (bottom) migration section generated by diffraction stack forward modeling and adjoint modeling codes, respectively. The model consisted of 75 contiguous point scatterers that defined the synclinal reflector denoted by white dashes in the bottom figure.

```
%
for ixtrace=1:ntrace; %LOOP OVER ALL TRACES
    for ix=1:nx;     %LOOP OVER ALL X-Coord IN MODEL
    for iz=2:nz;     %LOOP OVER ALL Z-Coord IN MODEL
      r = sqrt((ixtrace*dx-ix*dx)^2+(iz*dz)^2); DISTANCE
      time = 1 + round( r/c/dt ); PROPAGATION TIME
      data(ixtrace,time) = mig(ix,iz)/r + data(ixtrace,time);
    end;
    end;
 data1(ixtrace,:)=conv2(data(ixtrace,:),rick)% CONVOLVE
                              % TRACES WITH BANDED WAVELET
                              % rick
end;
```

4.3.6 Matrix Operator Notation

Equation 4.37 gives the integral representation for Born forward modeling. We will now discretize the scattered pressure field, slowness model and modeling operator.

If the slowness field is discretized onto a mesh of N square cells with a constant slowness perturbation Δs_i in the ith cell, equation 4.37 reduces to a linear system of equations for the high-frequency approximation:

$$p_i^{scatt} = \sum_{j=1}^{N} l_{ij} \Delta s_j, \tag{4.41}$$

for $i = 1, .., M$ and

$$l_{ij} \leftrightarrow l_{rst,j} \approx 2\delta(t - \tau_{jr} - \tau_{js}) \frac{dx^2}{|\mathbf{x}_j - \mathbf{x}_r||\mathbf{x}_j - \mathbf{x}_s|}, \tag{4.42}$$

$$p_i^{scatt} \leftrightarrow p_{rst} = p(\mathbf{x}, t|\mathbf{x}_s, 0) \ . \tag{4.43}$$

Here dx is the width of a cell; $i \leftrightarrow (r, s, t)$ denotes the 1-1 mapping between the single data index i and the 3-tuple of data indices (r, s, t); and r, s, and t correspond to the integer indices associated with discretizing the receiver locations $\mathbf{x}$, source locations $\mathbf{x}_s$ and time values t, respectively. Here, the model indices are denoted by j. More compactly, the above system of equations can be represented by

$$\mathbf{L}\ \Delta\mathbf{s} \quad = \quad \mathbf{p^{scatt}}, \tag{4.44}$$

where $\mathbf{L}$ represents the MxN matrix with components l_{ij}, $\mathbf{p^{scatt}}$ is the $Mx1$ vector of pressure data, and $\Delta\mathbf{s}$ is the $Nx1$ vector of slowness perturbations. Note, in the above MATLAB script the $\mathbf{L}$ matrix is not explicitly formed, rather, each trace is computed by the summation indicated in equation 4.41.

4.4 3D Integral Equation Adjoint Modeling

In a later chapter, we will see that the least squares formalism solves for $\mathbf{m}$ in equation 4.44 by using the inverse to the normal equations:

$$\mathbf{m} \quad = \quad [\mathbf{L}^{*\mathbf{T}}\mathbf{L}]^{-1}\mathbf{L}^{*\mathbf{T}}\mathbf{p} \approx \mathbf{L}^{*\mathbf{T}}\mathbf{p}. \tag{4.45}$$

The estimate for the reflectivity model $\mathbf{m}$ approximated by $\mathbf{L}^{*\mathbf{T}}\mathbf{p}$ is known as the migration image, where $\mathbf{L}^{*\mathbf{T}}$ is the transposed conjugate of $\mathbf{L}$ othrewise known as the adjoint of $\mathbf{L}$.

Adjoint Matrix and Kernel. If the elements of $\mathbf{L}$ are defined as $[\mathbf{L}]_{ij} = l_{ij}$, then the elements of its adjoint are given by $[\mathbf{L}^{*T}]_{ij} = l^*_{ji}$. Simply put, the element of the adjoint matrix $\mathbf{L}^{*T}$ is obtained by exchanging the order of the element indices in l_{ij} and taking its complex conjugate.

In a similar way we can define the adjoint kernel (Morse and Feshbach, 1954) of the integral equation in equation 4.36. Interchange the order of the data and model variables in the kernel $-2\omega^2 W(\omega)G(\mathbf{x}_g|\mathbf{x}')G(\mathbf{x}'|\mathbf{x}_s)$ and take its complex conjugate.

Specifically, multiply the adjoint kernel is $-2\omega^2 W(\omega)^*\, G(\mathbf{x}_g|\mathbf{x}')^* G(\mathbf{x}'|\mathbf{x}_s)^* = -2\omega^2 W(\omega)^* e^{ik(|\mathbf{x}'-\mathbf{x}_s|+|\mathbf{x}'-\mathbf{x}_g|)}/\mathbf{x}_g||\mathbf{x}'-\mathbf{x}_s|$, by the scattered data $P(\mathbf{x}_g|\mathbf{x}_s)^{scatt}$ and integrate over all data coordinates (including frequency) to give diffraction stack equation of adjoint modeling:

$$\mathbf{m} \approx \mathbf{L}^{*T}\mathbf{p} \quad = \quad \delta\mathbf{s}(\mathbf{x})^{mig},$$

or

$$\delta\mathbf{s}(\mathbf{x})^{mig} \quad = \quad \int_{data} \frac{2}{|\mathbf{x}-\mathbf{x}_s||\mathbf{x}-\mathbf{x}_g|}\{\int_{-\infty}^{\infty}[(i\omega^2)P(\mathbf{x}_g|\mathbf{x}_s)^{scatt}W(\omega)^*]e^{-i\omega(\tau_{sr}+\tau_{rg})}d\omega\}dx_g dx_s. \tag{4.46}$$

Appendix C presents the differentiation theorem which says that $(i\omega^2)P(\mathbf{x}_g|\mathbf{x}_s)^{scatt}$ Fourier transforms as the second-time derivative of $p(\mathbf{x}_g, t|\mathbf{x}_s, 0)^{scatt}$, and the convolution theorem says that the product of two spectrums $(i\omega^2)P(\mathbf{x}_g|\mathbf{x}_s)^{scatt}$ and $W(\omega)^*$ Fourier transforms as their convolution $2\pi\ddot{p}(t) * w(-t)$. Thus, equation 4.46 becomes

$$\delta s(\mathbf{x})^{mig} \quad = \quad 4\pi \int_{data} \frac{[\ddot{p}(\mathbf{x}_g, t|\mathbf{x}_s, 0)^{scatt} * w(-t)]_{t=\tau_{sr}+\tau_{rg}}}{(|\mathbf{x}-\mathbf{x}_s||\mathbf{x}-\mathbf{x}_g|)} dx_g dx_s. \tag{4.47}$$

Here *mig* denotes the *migration* image of the slowness perturbation, and the integration is over the planar free surface where the source and geophones are restricted to $z = 0$.

Physical Meaning. To interpret the physical meaning of equation 4.47, assume a single source at $\mathbf{x}_s$, a single receiver at $\mathbf{x}_g$, and a wideband source so that $w(t) \rightarrow \delta(t)$ and $p(t) * w(-t) \rightarrow p(t) * \delta(-t) = p(t)$. In this case, equation 4.47 becomes

$$\delta s(\mathbf{x})^{mig} = 4\pi \frac{\ddot{p}(\mathbf{x}, \tau_{sr}+\tau_{rg}|\mathbf{x}_s, 0)^{scatt}}{|\mathbf{x}-\mathbf{x}_s||\mathbf{x}_g-\mathbf{x}|}. \tag{4.48}$$

For 2D, the term $\tau_{sr} + \tau_{rg} = constant$ describes an ellipse in model space coordinates $\mathbf{x} = (x, z)$ with foci at the source $\mathbf{x}_s$ and receiver $\mathbf{x}_g$ points. Thus, the migration image is formed by smearing the reflection energy at time $\tau_{sr} + \tau_{rg}$ along the appropriate ellipse in $\mathbf{x}$ coordinates. For numerous traces, the migration image is formed by smearing and summing reflection energy along the appropriate ellipses in model space. The bottom illustration in Figure 4.6 shows the migration of a single trace for a zero-offset trace.

A MATLAB code for zero-offset diffraction stack migration is given below.

```
% ZERO-OFFSET ADJOINT MODELING OR
% ZERO-OFFSET DIFFRACTION STACK MIGRATION
%
for ixtrace=1:ntrace; %LOOP OVER ALL TRACES
    for ix=1:nx      %LOOP OVER ROWS OF MODEL PIXELS
    for iz=1:nz;     %LOOP OVER COLUMNS OF MODEL PIXELS
      r = sqrt((ixtrace*dx-ix*dx)^2+(iz*dx)^2);
      time = round( 1 +  r/c/dt );
      migi(ix,iz) = migi(ix,iz) + cdp3(ixtrace,time)/r;
    end;
    end;
end;
```

4.5 Summary

The Green's function is the point source impulse response of the model. In our case the model is any inhomogeneous acoustic medium where the wave motion honors the acoustic wave equation. The wave motion $P(\mathbf{x})$ (or corresponding inhomogeneous Green's function) in an arbitrary acoustic model is impossible to determine analytically, but $P(\mathbf{x})$ can be found numerically by inserting the Green's function for a homogeneous medium into Green's theorem (equation 4.24). This says that the field values within the volume are a superposition of the direct wave in the background medium and the reflections and scattering from the boundaries. These boundaries define the parts of the medium where the actual medium is not the same as the background medium. Unfortunately, the surface integral term requires knowledge of the field values that we are trying to calculate. To eliminate this stumbling block, we use the Born approximation to the Lippmann-Schwinger equation, which gives the field values as a superposition of the primary reflections from the slowness perturbations. Here, primary reflections are propagating with the background velocity, and present a good approximation to the actual seismograms if the slowness perturbations are weak, i.e., multiple scattering events are negligible. The adjoint kernel applied to the data gives the migration, the topic for the next chapter. These forward and adjoint equations under the Born approximation are the primary modeling tools and imaging in exploration geophysics.

4.6 Exercises

1. Prove the identity $G\nabla^2_{\mathbf{x}'}P = P\nabla^2_{\mathbf{x}'}G + \nabla_{\mathbf{x}'} \cdot (G\nabla_{\mathbf{x}'}P - P\nabla_{\mathbf{x}'}G)$. Show work.

2. Show that if the Green's function for the Laplace's equation is given by

$$\nabla^2_{\mathbf{x}'} G(\mathbf{x}|\mathbf{x}') = \delta(\mathbf{x} - \mathbf{x}'), \tag{4.49}$$

then the Green's theorem for Laplace's equation

$$\nabla^2_{\mathbf{x}'} P(\mathbf{x}|\mathbf{x}') = -F(\mathbf{x}), \tag{4.50}$$

is given by equation 4.24, except now the Green function is given by equation 4.49. Show derivation in a step by step fashion, just like the Green theorem for the Helmholtz equation.

3. Compute movie of propagating harmonic wave, as described by MATLAB code below. Copy MATLAB script and run it in MATLAB. I have deliberately omitted geometrical spreading term to avoid imbalanced amplitude.

```
x=[1:300];y=[1:300];[X,Y] = meshgrid(x,y);k=.05; % Set up
                            %arrays of X and Y cooridnates
R=sqrt(X.^2+Y.^2);G=cos(k*R);imagesc(G) % Compute Green
                                        % Function for k=1;
polarity=1; % Causal Green function if polarity=1; Acausal
                            %Green Function if polarity=-1
for j=1:100;
G=cos(k*R-polarity*j*.1);imagesc(G) % Compute Movie of
                          %Propagating Wave as time increases
title([' Green Function t = ',num2str(j*.01)]); colorbar
xlabel('X (m)');
ylabel('Y (m)');
pause(.01)
end
```

Look at movie and script, and try to understand script. For example, type help meshgrid to understand how meshgrid works. What is the period of this harmonic wave? What is the wavelength? What is propagation velocity? What is apparent velocity in x direction? What is apparent velocity in y direction. Change polarity to polarity=-1. Which way is wave propagating, outgoing or incoming to the origin?

4. The Helmholtz equation for a spherically symmetric Green's function is given by

$$\nabla^2_{\mathbf{x}} G + k^2 G \;\; = \;\; \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 \frac{\partial G}{\partial r}) + k^2 G = 0 \quad (for\ r \neq 0). \tag{4.51}$$

Note that if the source point $\mathbf{x}' = 0$ is at the origin then $G(\mathbf{x}|\mathbf{x}') = e^{ikr}/r$ where $r = |\mathbf{x} - \mathbf{x}'| = |\mathbf{x}|$. Show that the Green's function for either an outgoing (equation 4.18) or incoming wave satisfies this equation as long as $r \neq 0$.

5. Recall the volume integral in spherical coordinates over a sphere of radius R is given by

$$\int_0^R \int_0^\pi \int_0^{2\pi} r^2 dr\ sin\theta d\theta d\phi \quad = \quad 4/3\pi R^3. \tag{4.52}$$

Prove this. Also prove that the surface integral area is equal to the following:

$$\int_0^\pi \int_0^{2\pi} R^2\ sin\theta d\theta d\phi \quad = \quad 4\pi R^2. \tag{4.53}$$

Here $d\Omega = sin\theta\ d\phi d\theta$ is the differential of the solid angle, with the identity $\int d\Omega = 4\pi$.

6. A previous problem asked to show that the Green's function satisfied the Helmholtz equation when $r \neq 0$. Now you will show (Morse and Feshbach, 1953) that integrating equation 4.17 over a small sphere that surrounds the source point $\mathbf{x}' = 0$ centered at the origin yields

$$\int_{volume} (\nabla^2_{\mathbf{x}'} + k^2) G(\mathbf{x}'|\mathbf{x})\ r^2 dr\ sin\theta d\theta d\phi \quad = \quad \int_{volume} \delta(\mathbf{x} - \mathbf{x}') dx^3 = 1. \tag{4.54}$$

The integration is in the observer space. Proving that the RHS becomes 1 is obvious, by definition of the delta function. The LHS can be shown to be equal to 1 by using the following argument. Assume the integration is about a small spherical surface with radius ϵ. In this case, the volume integration over the k^2-like term in equation 4.55 becomes

$$\begin{aligned} |k^2 \int_{volume} G r^2 dr d\Omega| &= |k^2/4\pi \int_{volume} e^{ikr} r dr d\Omega|, \\ &\leq k^2/4\pi \int_{volume} r dr d\Omega, \\ &= k^2\epsilon^2/2, \end{aligned} \tag{4.55}$$

which goes to zero as the volume radius $\epsilon \rightarrow 0$. The inequality follows from the Schwartz's theorem, which says $\int f(x)dx \leq max(|f(x)|) \int dx$.

The only remaining thing to do is to show that the integration of the Laplacian $\nabla^2_{\mathbf{x}'} G$ goes to 1. This can be shown by using Gauss's theorem: $\int_{volume} \nabla^2_{\mathbf{x}'} G\ r^2 dr\ d\Omega = \int_{surface} \partial G/\partial r r^2 d\Omega$, plugging in the Green's function $G = -(4\pi)^{-1} e^{ikr}$ and letting ϵ go to zero. Do this. (Hint: $\partial(1/r)/\partial r = -1/r^2$).

7. Convert the zero-offset forward modeling code into a prestack modeling code where the source and receiver positions are on the free surface but are not coincident. Demonstrate the effectiveness of this code by creating shot gathers for two earth models: a point scatterer model and a single reflector interface model. A shot gather is a collection of traces where the events are excited by a common source.

Appendix A: Causal and Acausal Green's Functions

Causal Green's Functions. Applying the inverse Fourier transform

$$\mathcal{F}\cdot \;=\; \int_{-\infty}^{\infty} d\omega\; e^{i\omega(t-t_s)}\cdot, \tag{4.56}$$

to equation 4.18 gives the causal Green's function

$$\begin{aligned} g_c(\mathbf{x}, t-t_s|\mathbf{x}', 0) \;&= \mathcal{F}(G(\mathbf{x}|\mathbf{x}')) = \; -\frac{1}{4\pi}\int_{-\infty}^{\infty} e^{-i\omega(|\mathbf{x}'-\mathbf{x}|/c-(t-t_s))}/|\mathbf{x}-\mathbf{x}'|d\omega, \\ &= \qquad\qquad -\frac{1}{2}\delta(t-t_s-|\mathbf{x}'-\mathbf{x}|/c)/|\mathbf{x}-\mathbf{x}'|, \end{aligned} \tag{4.57}$$

where

$$\delta(t-t_s-|\mathbf{x}'-\mathbf{x}|/c) = \begin{cases} 1 & \text{if } t-t_s = |\mathbf{x}'-\mathbf{x}|/c, \\ 0 & \text{if } t-t_s \neq |\mathbf{x}'-\mathbf{x}|/c. \end{cases} \tag{4.58}$$

More precisely, $\delta(t)$ is infinite at $t=0$ and is a generalized *delta functional* which can be only defined in terms of an inner product with a sufficiently regular function (Zemanian, 1965). With this understanding, we symbolically denote its value to be 1 for $t=0$.

Equation 4.58 describes an expanding circle centered about the source point; and the initiation time of the impulsive source is at time t_s and the observation time is denoted by t. The radius of the expanding circle is equal to $c(t-t_s)$ as long as the $t > t_s$. Equation 4.57 is a causal Green's function because the wave is not observed until it has propagated from the source point to the observation point in the *retarded time* given by $|\mathbf{x}'-\mathbf{x}|/c$. Figure 4.8a shows the causal Green's function expanding as a "light cone" in $x-y-t$ space.

It is easy to see that the the Green's function for a homogeneous medium is stationary so that $g_c(\mathbf{x}, t|\mathbf{x}', t_s) = g_c(\mathbf{x}, t-t_s|\mathbf{x}', 0)$. This says that the observed field depends on the time difference between the source initiation time and the observation time, no matter when the source was initiated. Thus, the shot gather obtained at noon should look like the shot gather obtained at midnight, except for a delay time of 12 hours.

The stationarity property is also true for any acoustic medium (Morse and Feshbach, 1953). A simple proof is given by multiplying equation 4.17 by a phase shift term $e^{i\omega\tau_0}$. The RHS term now is $\delta(\mathbf{x}'-\mathbf{x})e^{i\omega\tau_0}$, which is equivalent to delaying the source initiation time by the delay time τ_0. But the linearity of the LHS PDE says that the solution $G(\mathbf{x}'|\mathbf{x})e^{i\omega\tau_0}$ is the same Green's function as before, except for the time delay of τ_0. Note $G(\mathbf{x}'|\mathbf{x})e^{i\omega\tau_0}$ would not be a solution if the PDE was non-linear with terms, e.g., such as $G(\mathbf{x}'|\mathbf{x})^2$.

Acausal Green's Function. Equation 4.57 describes a causal Green's in a homogeneous medium where the wavefield is *only* observed after the source initiation time τ_s. In contrast, the *acausal* Green's function is denoted by the adjoint kernel $G(\mathbf{x}|\mathbf{x}')^*$, which for a homogeneous medium is given as $-\frac{1}{4\pi}e^{ik|\mathbf{x}-\mathbf{x}'|/c}/|\mathbf{x}-\mathbf{x}'|$.

Applying the inverse Fourier transform

$$\mathcal{F}\cdot \;=\; \int_{-\infty}^{\infty} d\omega\; e^{i\omega(t-t_s)}\cdot, \tag{4.59}$$

a) Causal Green's Function b). Acausal Green's Function

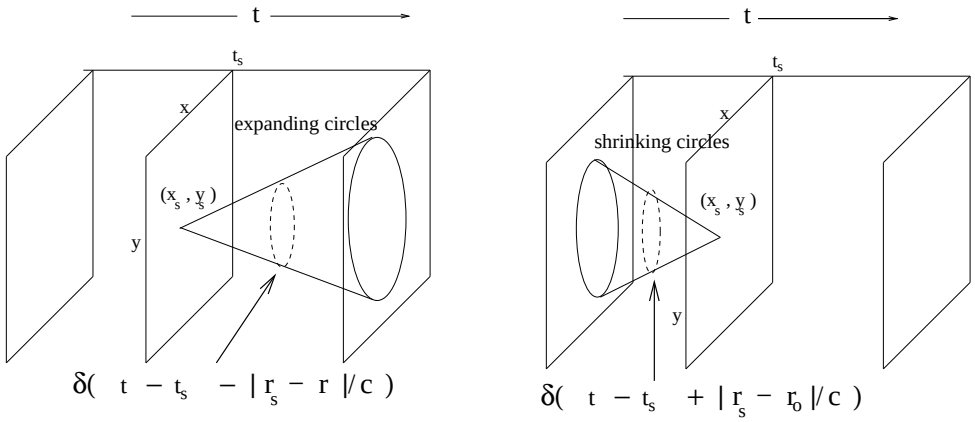


Figure 4.8: (a). Causal and (b). acausal Green's function in $x - y - t$ space.

to $G(\mathbf{x}|\mathbf{x}')^*$ yields the time domain acausal Green's function:

$$\begin{aligned} g_a(\mathbf{x}, t-t_s|\mathbf{x}', 0) &= \mathcal{F}(G(\mathbf{x}|\mathbf{x}')^*) = -\frac{1}{4\pi}\int_{-\infty}^{\infty} e^{i\omega(|\mathbf{x}'-\mathbf{x}|/c+(t-t_s))}/|\mathbf{x}-\mathbf{x}'|d\omega, \\ &= -\frac{1}{2}\delta(t-t_s+|\mathbf{x}'-\mathbf{x}|/c)/|\mathbf{x}-\mathbf{x}'|, \end{aligned} \tag{4.60}$$

where

$$\delta(t-t_s+|\mathbf{x}'-\mathbf{x}|/c) = \begin{cases} 1 & \text{if } t_s - t = |\mathbf{x}'-\mathbf{x}|/c, \\ 0 & \text{if } t_s - t \neq |\mathbf{x}'-\mathbf{x}|/c. \end{cases} \tag{4.61}$$

The acausal Green's function describes a contracting circular wavefront centered at the source point and is extinguished at time t_s and later. It is acausal because it is alive and contracting (se Figure 4.8) prior to the source initiation time t_s, and turns off after the source turns on! This Green's function is important because it is used for seismic migration which focuses wavefronts to their place of origin, compared to the causal Green's function that is used for modeling of exploding reflectors.

Appendix B: Sommerfeld Radiation Conditions

We will show that the outgoing Green's function in equation 4.17 satisfies the Sommerfeld outgoing boundary condition at infinity (Butko, 1972):

$$\lim_{r\to\infty}(r\frac{\partial G}{\partial r} + ikrG) = 0. \tag{4.62}$$

Differentiating $G = e^{-ikr}/r$ with respect to r and substituting into equation 4.62 we get:

$$\lim_{r\to\infty}(ikre^{-ikr}/r - ikrG - G) = \lim_{r\to\infty}(ike^{-ikr} - ike^{-ikr} - e^{-ikr}/r) \to 0. \tag{4.63}$$

Notice that the acausal or incoming Green's function $G = e^{ikr}/r$ will not satisfy this radiation condition. The Sommerfeld boundary condition is also needed to show that the integrand of the surface integral in equation 4.24 goes to zero at infinity.

Appendix C: Fourier Identities

The foward and inverse Fourier transforms are respectively given by

$$F(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt, \tag{4.64}$$

$$f(t) = \int_{-\infty}^{\infty} F(\omega)e^{i\omega t}d\omega, \tag{4.65}$$

The following are Fourier identities, where the double-sided arrows indicate the functions are Fourier pairs and $\mathcal{F}$ indicates the forward Fourier transform.

1. **Differentiation:** $\partial^n/\partial t^n \longleftrightarrow (i\omega)^n$. This property is proved by differentiating equation 4.65 w/r to t.

2. **Convolution Theorem:** $f(t)*g(t) = \int f(\tau)g(t-\tau)d\tau \longleftrightarrow F(\omega)G(\omega)$. This property is proved by applying the Fourier transform to the convolutional equations

$$\begin{aligned}\mathcal{F}[f*g] &= \mathcal{F}[\int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau], \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega t}[\int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau]dt. \end{aligned} \tag{4.66}$$

   Interchanging the order of integration we get

$$\mathcal{F}[f*g] = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(\tau)[\int_{-\infty}^{\infty} e^{-i\omega t}g(t-\tau)dt]d\tau, \tag{4.67}$$

   and defining the integration variable as $t' = t - \tau$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(\tau)[\int_{-\infty}^{\infty} e^{-i\omega(t'+\tau)}g(t')dt']d\tau, \tag{4.68}$$

   and using the definitions of the Fourier transform of $g(t)$ and $f(t)$ we get

$$= G(\omega)\int_{-\infty}^{\infty} f(\tau)e^{-i\omega\tau}d\tau, \tag{4.69}$$

$$= 2\pi F(\omega)G(\omega). \tag{4.70}$$

   We will often denote the convolution of two functions $f(t)*g(t)$ by the $*$ symbol.

3. **For real $f(t)$: $f(-t) \longleftrightarrow F(\omega)^*$:** This property is easily proven by taking the complex conjugate of equation 4.65 to get $f(t)^* = f(t) = \int F(\omega)^* e^{-i\omega t}d\omega$ and then apply the transform $t = -t'$.

4. **Correlation: $f(t)*g(-t) = f(t)\otimes g(t)$:** By definition $f(t)*g(-t) = \int f(t-\tau)g(-\tau)d\tau$. By changing the dummy integration variable $\tau \rightarrow -\tau'$ we get $f(t)*g(-t) = \int f(t+\tau')g(\tau')d\tau' = f(t)\otimes g(t)$, where $\otimes$ represents correlation.

Appendix D: Uneven sampling to Even Sampling MATLAB Code

The MATLAB code for getting an even sampled function sampled at dt from an unevenly sampled function is

```
function [even,deceven]=resmp(uneven,time,dt)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Resamples input series uneven(t) with argument t in seconds
% so that output series deceven(t1) is evenly sampled with dt sampling interval
% Assumes an even sampled time interval in t1;
```

```
% Crude nearest neighbor interpolation..you might try use fancier interpolator.
% uneven  - input- Uneven sampled function
% time    - input- 2-way time of each uneven value
% dt      - input- sample interval in time for resampled data
% even    -output- Even sampled function
% deceven -output- Decimated data at dt sampling rate
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
nun=length(uneven);dtprime=dt;
tmin=min(time); tmax=max(time);
dt=max(diff(time))*3;
dtt=tmax-tmin;
nt=round(dtt/dt+1);
even=zeros(nt,1);istart=1;
for i=1:nun-1;   % Loop over uneven samples in time
   x1 = uneven(i); x2 = uneven(i+1);x=[x1 x2];
   t1 = time(i); t2 = time(i+1);t12= round((t2-t1)/dt+1);
   y=resample(x,t12,1);iend=round(t12/2)+1;
   even(istart:istart+iend-1)=y(1:iend);
   istart=istart+iend+1;
end
nl=length(even);ttime=[1:nl]*dt+time(1);
for i=1:nl-1;    % Kill 0 values of even
if even(i)<.1;even(i)=even(i-1);end
end
dt=dtprime;
deceven=even;
nttt=round(dtt/dt)+1;  % Decimate data if too many
if nttt<nl;
r=round(nl/nttt);
deceven=decimate(even,r);
end
subplot(311);plot(time,uneven);title('Original Velocity vs Time')
n=length(even);
subplot(312);plot(even);title('Original Velocity vs Sample Number')
n=length(deceven);t=[1:n]*dt+time(1);
subplot(313);plot(t,deceven);title('Decimated Even Sampled dt  Velocity vs Time')
```

Chapter 5

Reverse Time Migration

This chapter describes the reverse time migration (RTM) method and its equivalent interpretation known as generalized diffraction stack migration (GDM). GDM is a diffraction-stack interpretation of RTM that can be used to design filtering strategies for RTM. The GDM migration image is formed by taking the dot product between the shot gather and the migration operator. Compared to other migration methods, RTM and GDM accounts for all of the arrivals in the wavefield, including both primaries and multiples. This can lead to much better resolution in the image, but the image blurring is more sensitive to accuracy in the velocity model. As we will see later, full wavefield inversion is nothing more than an iterative sequence of migrating data residuals by RTM.

5.1 Introduction

An earlier chapter introduced the concept of Kirchhoff migration (KM), which forms the migration image by summing the amplitudes of the CSG traces along the associated "moveout curve". If the medium is homogeneous then the moveout curve in offset-time coordinates forms an hyperbola, otherwise it describes an irregular hyperbola-like curve. The problem with KM is that it only accounts for single arrival imaging and does not account for the fact that multiple arrivals may have reflected off of the same reflection point at $\mathbf{x}_0$. In fact, the single arrival associated with the "moveout curve" may be quite weak relative to other reflections at $\mathbf{x}_0$ so the diffraction stack image here might be of low quality. This problem is particularly common beneath salt bodies that tend to defocus the incident and reflection energy beneath the salt.

To overcome the defocusing problem caused by large velocity contrasts, Whitmore (1982) and McMechan (1983) introduced the concept of reverse time migration (RTM). Instead of using ray tracing to compute the Green's functions, a finite-difference solution to the wave equation is used to estimate $G(\mathbf{x}|\mathbf{x}')$. If the velocity model is accurately known then all arrivals are accounted for that might reflect off any trial image point, including multiples, diffractions, and converted waves. The benefit can be, if the velocity model is accurately known, a great increase in signal-to-noise ratio in the migration image and a greater ability to see beneath bodies with high velocity contrast. However, the danger is that the velocity model is not exactly known so small errors in the model estimate can lead to strong noise

in the migration image. Thus, the more accurate the velocity model the better the RTM image.

To facilitate an understanding of RTM, we start with a description of wavefield extrapolation in a layered medium.

5.2 Extrapolation by Phase Shifts

Wavefield extrapolation is defined as using the boundary values of a propagating wave to predict the wavefield distant from the boundary. If the wavefield prediction is at an earlier time we call this backpropagation, and if the predicted wavefield is at a later time we call this forward propagation. Typically, extrapolation is accomplished by applying weighted phase shifts to the boundary data in the frequency domain. For monofrequency plane wave data, this operation can simply be described by multiplying the boundary data by $e^{\pm ik_z \Delta z}$ for plane waves with wavenumber k_z; and for data on an arbitrary data extrapolation can be generalized using the Fresnel-Huygen's principle or Green's theorem.

For example, assume a downgoing harmonic plane wave $\tilde{D}(x,z) = W(\omega)e^{i(k_x x - k_z z - \omega t)}$ that reflects across the lower interface in Figure 5.1, and the resulting upgoing wave at z is recorded as

$$\tilde{U}(x,z) \quad = \quad \tilde{D}(x,z_0)Re^{i(k_x x + k_z z - \omega t)}, \tag{5.1}$$

where the upgoing wavenumber vector is denoted as $\mathbf{k} = (k_x, k_z)$, R is the reflection coefficient, and $W(\omega)$ is the source wavelet spectrum at angular frequency ω. For an upgoing wave, the phase $\mathbf{k}\cdot\mathbf{x} = k_x x + k_z z$ must increase with increasing time to keep the exponential argument a constant, as expected when following an upgoing wavefront.

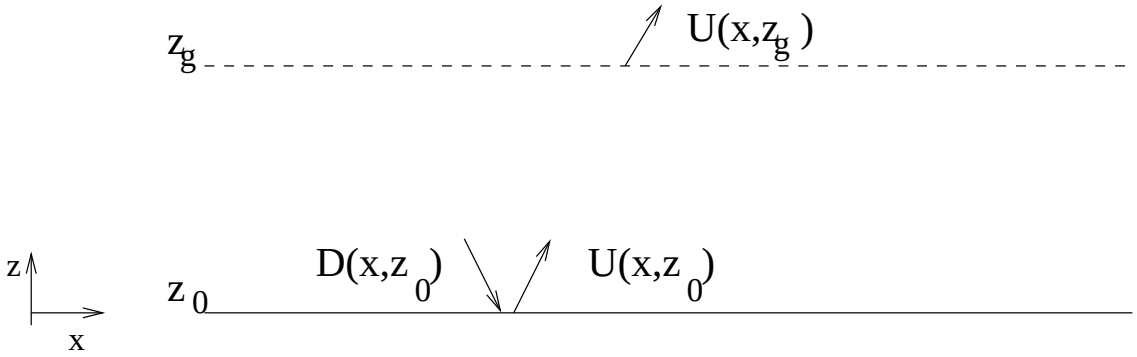


Figure 5.1: Downgoing and upgoing waves reflected at an interface at depth z_0.

To *forward extrapolate* the wavefield $\tilde{U}(x,z_0)$ at z_0 to an overlying interface at $z_0 + |z - z_0| = z_0 + \Delta z$, multiply the upcoming wavefield at $z = z_0$ by the forward extrapolation operator $e^{ik_z \Delta z}$ to get

$$\tilde{U}(x,z) \quad = \quad \tilde{U}(x,z_0)e^{ik_z \Delta z}. \tag{5.2}$$

We call $e^{ik_z \Delta z}$ the forward extrapolation operator for upgoing waves because it predicts the upgoing field forward in time by adding the positive[1] phase term $k_z \Delta z$ to the exponential

[1]This assumes k_z and ω are positive.

argument of $\tilde{U}(x, z_0)$. For example, if upcoming waves were measured along the $z = 0$ plane in Figure 5.2 then the formula could be used to estimate the data at $z = z_g$.

Conversely, predicting the deeper (or prior) upgoing waves at $z = z_0$ from upgoing fields along the upper boundary $z = z_g$ is obtained by subtracting phase from the exponential argument in $\tilde{U}(x, z_g)$. That is, simply rearrange equation 5.2 to get

$$\tilde{U}(x, z_0) \quad = \quad \tilde{U}(x, z) e^{-ik_z \Delta z}|_{z=z_g}, \tag{5.3}$$

where $e^{-ik_z \Delta z}$ is the backward extrapolation operator for upgoing waves[2]. Similar to the acausal Green's function, this operator subtracts phase from the boundary data at z_g to predict upcoming events earlier in time below the measurement boundary. For example, the above data at $z = z_g$ could be used to predict the data measured at the deeper depth $z = 0$ in Figure 5.2.

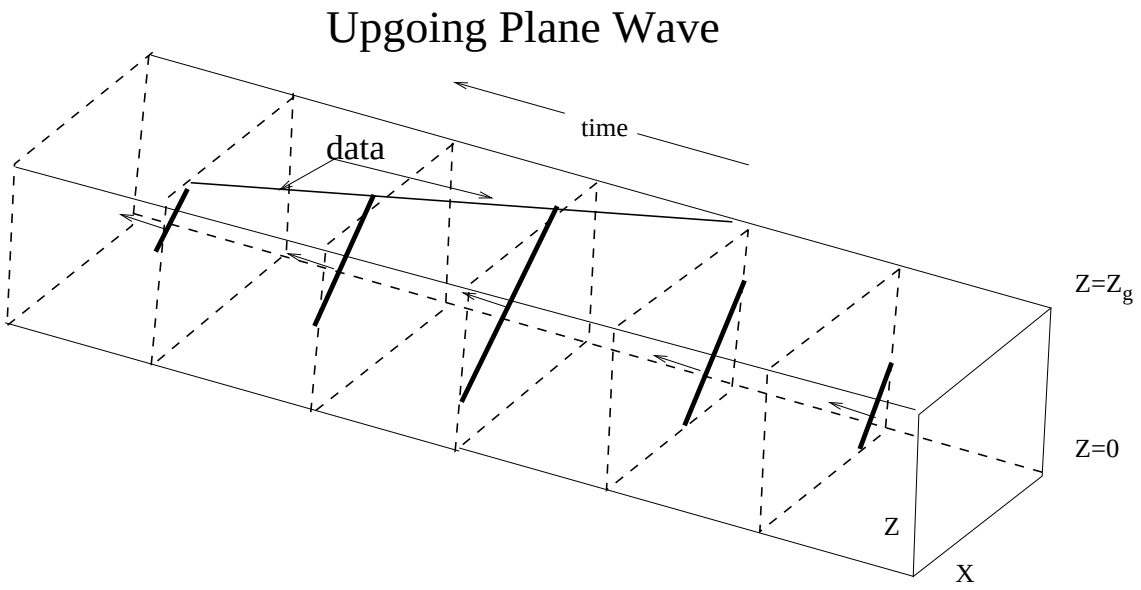


Figure 5.2: Upgoing plane waves n $x - z - t$ data cube; the measured data can be along the $z = 0$ or $z = z_g$ planes.

5.2.1 Estimating the Reflection Coefficient

Migration can be defined as estimating the earth's reflectivity distribution from recorded reflection data. For example, the simplest reflector distribution is that of a horizontal layer with a reflectivity distribution $R(x, z) = R_0 \delta(z - z_0)$ that separates two homogeneous half spaces at the depth z_0. Claerbout (1975) suggested that $R(x, z)$ could be estimated by taking the ratio of the downgoing field and the estimated upgoing field in equation 5.3 and evaluating at the interface at $z = z_0$:

$$R(x, z)|_{z=z_0} = \frac{\tilde{U}(x, z)}{\tilde{D}(x, z_0)}|_{z_0} = \frac{\tilde{U}(x, z)\tilde{D}(x, z)^*}{|\tilde{D}(x, z)|^2 + \epsilon}|_{z=z_0}, \tag{5.4}$$

[2]Conversely, $e^{-ik_z \Delta z}$ is the forward extrapolation operator for downgoing waves.

where ϵ is a small positive constant to avoid division by zero. The composite estimate of the reflectivity distribution over all frequencies is given by

$$r = \sum_{\omega} R(x,z), \tag{5.5}$$

where the summation is over the usable frequency band in the signal. At the reflection interface $z = z_0$, the phase of the predicted upgoing wave $\tilde{U}(x,z_0)$ will be canceled by the phase of the downgoing field $\tilde{D}(x,z_0)^*$ for any frequency so there will always be constructive interference at the interface for $\sum_\omega R(x,z)$. In contrast, if $z \neq z_0$, the numerator $\tilde{U}(x,z)\tilde{D}(x,z)^*$ in equation 5.4 will differ in phase for different frequencies so there will be destructive interference away from the interface.

5.3 Reverse Time Migration

The plane wave source can thought of as being initiated by a distant point source at $\mathbf{s}$, so we can re-express it in Green's function notation:

$$\tilde{D}(\mathbf{x}) \rightarrow W(\omega)G(\mathbf{x}|\mathbf{s}). \tag{5.6}$$

Similarly, the extrapolation operator $e^{-ik_z|\Delta z|}$ uses virtual sources at the measurement boundary (x_g, z_g) to predict upcoming waves below it at $\mathbf{x}$, i.e.,

$$\tilde{U}(x,z) = \tilde{U}(x,z_g)e^{-ik_z|\Delta z|}e^{ik_x x_g} \rightarrow \tilde{U}(\mathbf{x}) = \tilde{U}(\mathbf{x}_g)G(\mathbf{x}|\mathbf{g})^*, \tag{5.7}$$

where the phase shift term $e^{ik_x x_g}$ is used to conveniently account for the horizontal geophone coordinate x_g not being equal to the horizontal listening coordinate x. Plugging equations 5.6-5.7 into equation 5.4 and summing over all frequencies gives the migration formula for a multifrequency plane wave:

$$\begin{aligned} r(\mathbf{x}) = \sum_{\omega} R(\mathbf{x}) &= \sum_{\omega} \overbrace{e^{i\omega 0}}^{lagtime=0} \frac{\overbrace{\tilde{U}(\mathbf{x}_g)G(\mathbf{x}|\mathbf{x}_g)^*}^{backpropagated\ upgoing\ refl.} \overbrace{W(\omega)^*G(\mathbf{x}|\mathbf{s})^*}^{downgoing\ source}}{|W(\omega)G(\mathbf{x}|\mathbf{s})|^2 + \epsilon}, \\ &= u(x,t) \otimes d(x,t)|_{t=0}. \end{aligned} \tag{5.8}$$

This follows from the property that the inverse Fourier transform of the product of a spectrum $B(\omega)$ by the conjugated spectrum $A(\omega)^*$ yields the correlation of $a(t) \otimes b(t)$. If the time variable $t \rightarrow 0$ then this is zero-lag correlation. If the Green's functions are computed by numerical solutions to the wave equation then equation 5.8 is the formula for wave equation migration.

If there is more than one plane wave then an extra summation over k_x is used to estimate the composite migration image. The above formula is derived for a homogeneous medium, but can easily be extended to a layered medium as well, which becomes the phase shift migration method (Yilmaz, 2001). Similar to the interpretation of equation 5.8, the reverse time migration image at $\mathbf{x}$ is computed by taking the weighted dot product between the backpropagated data trace at $\mathbf{x}$ with the source field trace at $\mathbf{x}$.

5.3.1 Reverse Time Migration in Frequency Domain

The reverse time migration formula can be intuitively derived from the above considerations. Assume a medium with arbitrary velocity and define the associated Green's function as $G(\mathbf{x}|\mathbf{s})$. We will assume the Green's function can be decomposed into upgoing $U(\mathbf{x}_g|\mathbf{s})$ and downgoing fields, and that $U(\mathbf{x}_g|\mathbf{s})$ is measured at $\mathbf{x}_g$ on a level measurement surface above the source position. The Huygen's-Fresnel principle (Gu, 2000) says that each point $\mathbf{x}_g$ on the measurement plane acts as a secondary point source with the source spectrum of $U(\mathbf{g}|\mathbf{s})$, and excites an expanding quasi-spherical wavelet[3] The wavefield amplitude a small distance away is the superposition of these secondary wavefronts. Therefore, the upgoing field below the listening plane (i.e., backpropagated wavefield) can be reconstructed by weighting the recorded data $U(\mathbf{x}_g|\mathbf{s})$ with the acausal Green's functions $G(\mathbf{x}|\mathbf{g})^*$ and integrating over the geophone $\mathbf{g}$ indices to get the reconstruct the upgoing field[4]

$$\tilde{U}(\mathbf{x}) \quad \rightarrow \quad \omega^2 \int_g G(\mathbf{x}|\mathbf{g})^* U(\mathbf{g}|\mathbf{s}) dg. \tag{5.9}$$

Plugging equations 5.9 into equation 5.8 yields the RTM formula:

$$\mathbf{m}(\mathbf{x}) \quad = \quad \frac{\int_\omega \omega^2 \int_{data\ space} \overbrace{[G(\mathbf{g}|\mathbf{x})^* D(\mathbf{g}|\mathbf{s})]}^{backpropagated\ upgoing\ refl.} \overbrace{[W(\omega) G(\mathbf{x}|\mathbf{s})]^*}^{downgoing\ source} dg d\omega}{|I(\mathbf{x},\mathbf{s})|^2}, \tag{5.10}$$

where only one shot gather is assumed and

$$|I(\mathbf{x},\mathbf{s})|^2 \quad = \quad |W(\omega) G(\mathbf{x}|\mathbf{s})|^2, \tag{5.11}$$

can be interpreted as the source illumination compensation term, or a preconditioning factor. This factor backs out geometric spreading effects in the field and deconvolves the source wavelet as well. If many shot gathers are available then there is an extra integration over the source variables.

5.3.2 Reverse Time Migration in Time Domain

Assuming that $|I(\mathbf{x},\mathbf{s})| = 1$ and integrating equation 5.10 over all frequencies gives the traditional formula for reverse time migration:

$$\mathbf{m}(\mathbf{x}) \quad = \quad \int_{data\ space} \overbrace{[g(\mathbf{g}, -t|\mathbf{x}, 0) \star d(\mathbf{g}, t|\mathbf{s}, 0)]}^{backpropagated\ upgoing\ refl.} \otimes \overbrace{[w(t) \star g(\mathbf{x}, t|\mathbf{s}, 0)]}^{downgoing\ source} |_{t=0} dg. \tag{5.12}$$

where the the data function is assumed to have a double time derivative. The backpropagation term says that the acausal Green's function (seen as backward pointing light cones in Figure 5.3) is convolved with each trace at the surface to estimate earlier upcoming

[3] Exactly spherical in a homogeneous medium.

[4] The upgoing field above the recording plane can be *predicted* by replacing the acausal Green's function with a causal Green's function.

waves at depth. In this case, the resulting wavefield converges to a subsurface point with decreasing time. In contrast, the forward propagation of the source field is accomplished by temporally convolving the source wavelet at the source point with the causal Green's function. As illustrated with the forward pointing light cones on the rightside of Figure 5.3, the forward propagation of the field from surface traces gives rise to an expanding wavefield with increasing time.

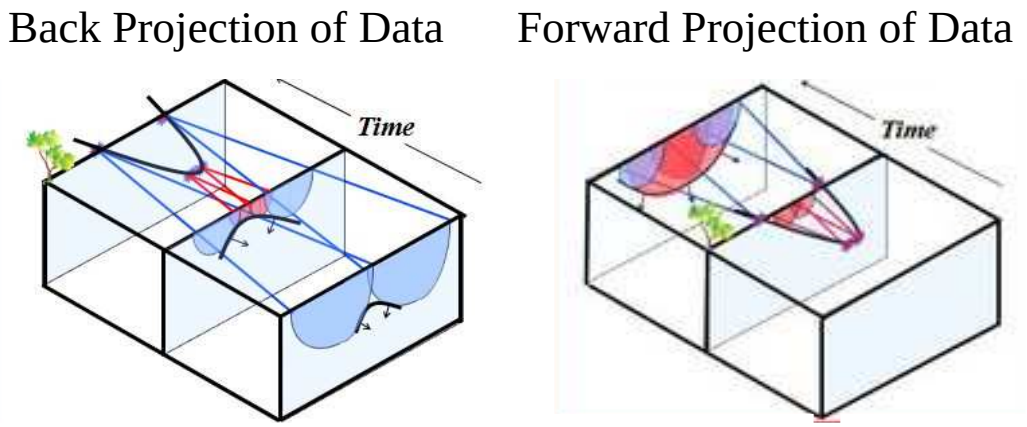


Figure 5.3: (left) Backward and (right) forward propagation of data measured along the surface.

Figure 5.4 illustrates the (bottom) backpropagated scattered field denoted by $b(\mathbf{x}, \mathbf{x}_s, t) = \int g(\mathbf{g}, -t|\mathbf{x}, 0) \star d(\mathbf{g}, t|\mathbf{x}_s, 0) dg$ and the (middle) forward source field denoted as $f(\mathbf{x}, \mathbf{x}_s, t) = w(t) \star g(\mathbf{x}, t|\mathbf{x}_s, 0)$ for a single source at $\mathbf{x}_s$. The temporal dot product[5] is zero everywhere, except at the scattering point where the both the source and backpropagated scattered fields are simultaneously alive. In MATLAB, the element product of an NxM mask matrix $\mathbf{A}$ with the NxM matrix $\mathbf{B}$ is denoted as $\mathbf{A}.*\mathbf{B}$; we shall call this a *spatial mask product*. Therefore, for a single shot gather, RTM is a spatial mask product in (x, y) and temporal dot product between the 3-dimensional matrices $f(\mathbf{x}, \mathbf{x}_s, t)$ and $b(\mathbf{x}, \mathbf{x}_s, t)$ in Figure 5.3.

The following MATLAB code implements equation 5.12 as reverse time migration of the common shot gather defined as $CSG(x, t)$.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Reverse time migration of single shot gather CSG(1:NX,1:it)
%
% (NX,NZ,NT)    - input- (Horiz,Vert) gridpt dimens. of vel model
%                         & # Time Steps
% FR            - input- Peak frequency of Ricker wavelet
% c             - input- NXxNZ matrix of velocity model
% (dx,dt)       - input- (space, time) sample intervals
% CSG(NX,NT)    - input- NXxNT matrix of data seismograms at z=2.
```

[5]The "zero-lag correlation of two time series" at $\mathbf{x}$ is a fancy way of saying "the dot product of two vectors", where one vector is the discrete time series for the forward modeled field $f(\mathbf{x}, \mathbf{x}_s, t)$ and the other is that for the backpropagated field $b(\mathbf{x}, \mathbf{x}_s, t)$, both at the spatial position (x, y, z).

```
%                          The direct waves are muted out and only upgoing
% FORW(NX,NZ,NT)- input- NXxNT matrix of source wavefield
% For(NX,NT)    - input- NXxNT matrix of forward modeled field
% IH(NX,NZ)     - input- NXxNZ preconditioner matrix of approx. inverse Hessian
% (p2,p1,p0)    -calcul- (future,pres,past) NXxNZ matrices of modeled
%                          pressure field after 2nd-order time derivative
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
p0=zeros(NX,NZ);p1=p0;p2=p0; cns=(dt/dx*c).^2;
for it=NT-1:-1:2
    p2 = 2*p1 - p0 +  cns.*del2(p1);
%
% Add Source Terms CSG(NX,NT) at NX points along z=2 below the surface
%
    p2(:,2)= p2(:,2) + CSG(:,it);
    m(:,:)=mig(:,:)+p2(:,:).*FORW(:,:,it);% Accumulate migration image
    p0=p1;p1=p2;
  end;
    mig=mig.*IH;  % Apply preconditioner inverse Hessian
    mig=del2(mig);% Apply Laplacian
```

No absorbing boundary conditions have been included in the above code, but if they are included then any first-order derivatives in time must take into account the sign change due to *backward* differencing in time[6]. Some procedures that lessen the artifacts in RTM images are described in Appendix A.

5.3.3 Generalized Diffraction Stack Migration

Equation 5.10 can be rearranged to give the formula for generalized diffraction stack migration (Schuster, 2002):

$$\mathbf{m}(\mathbf{x}) \;=\; \frac{\omega^2 \int_{data\ space} \overbrace{[W(\omega)G(\mathbf{g}|\mathbf{x})G(\mathbf{x}|\mathbf{s})]^*}^{diffraction-stack\ focusing\ kernel} \overbrace{[D(\mathbf{g}|\mathbf{s})}^{data}\, dg}{|I(\mathbf{x},\mathbf{s})|^2}, \tag{5.13}$$

or in the time domain

$$\mathbf{m}(\mathbf{x}) \;=\; \int_{data\ space} \overbrace{[w(t) \star g(\mathbf{g},t|\mathbf{x},0) \star g(\mathbf{x},t|\mathbf{s},0)]}^{diffraction-stack\ focusing\ kernel} \otimes \overbrace{[d(\mathbf{g},t|\mathbf{s},0)]}^{data} |_{t=0} dg. \tag{5.14}$$

This equation yields an alternative interpretation and implementation of RTM. It says that the *RTM migration image* $\mathbf{m}(\mathbf{x})$ *at* $\mathbf{x}$ *can be interpreted as the spatial (in* $\mathbf{g}$*) and temporal dot product of the migration Green's function* $w(t) \star g(\mathbf{g},t|\mathbf{x},0) \star g(\mathbf{x},t|\mathbf{s},0)$ *with the shot gather traces* $d(\mathbf{g},t|\mathbf{s},0)$. This migration Green's function (MGF) is the*point scatterer response of the migration operator (Schuster and Hu, 2000) and describes a diffraction hyperbola in data space for a single scatterer in a homogeneous medium*, as shown in Figure 5.5a. The apex of the hyperbola is centered above the point scatterer and the strength of the migration image

[6]It is too costly to store and retrieve the multidimensional matrix $FORW([1:nx],[1:nz],[1:nt])$; see exercise 2 for a parsimonious strategy for computing this forward field.

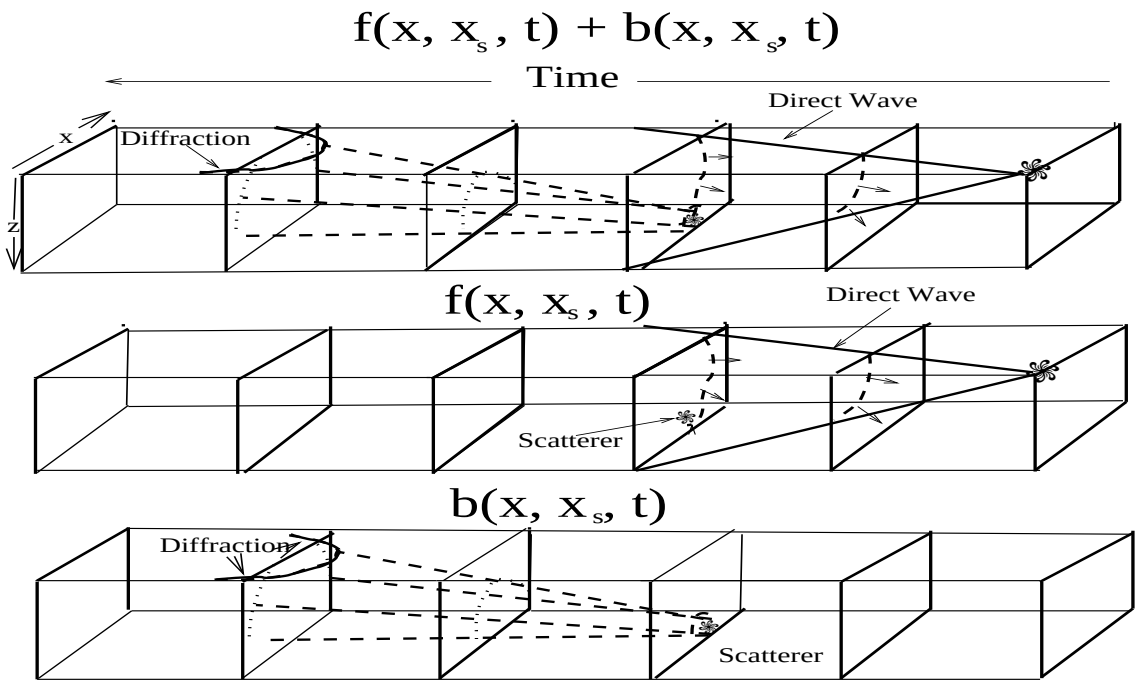


Figure 5.4: Snapshots of wavefronts for a point source located at the origin, and a buried scatterer denoted by *. Top figure depicts the total wavefield snapshots, middle figure depicts the forward propagated wavefield in a homogeneous medium, and bottom figure depicts the backpropagated scattered field.

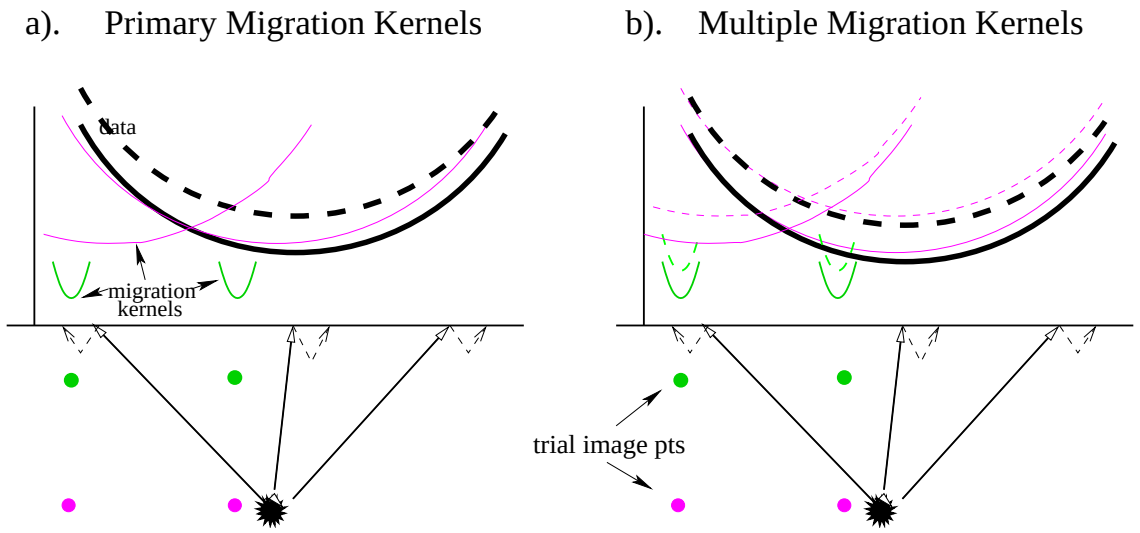


Figure 5.5: Migration kernels plotted in data space for a) primary and b) primary+multiple events associated with shallow (green) and (deep (pink) trial image points. The best match between the data (black hyperbolas) and migration curves (pink and green) is when the trial image point is near the actual scatterer's position; the highest correlation (i.e., resolution) for the multiple migration kernel on the right.

at $\mathbf{x}$ depends how well the data patterns, i.e. hyperbolas, match those of the migration kernel[7] with trial image point $\mathbf{x}$. The thick black line denotes the primary reflection data from the source (star) and the dashed black line denotes the secondary multiple scattering events associated with the near surface layer.

With GDM, all of the multiple scattering events are taken into account compared to the restriction to primary reflection events for standard diffraction stack migration. This is illustrated in Figure 5.5b where the pink and green hyperbolas describe the GDM migration kernel for primary and multiple scattering events associated with different trial image points; each filled circle belongs to a different set of migration hyperbolas. Here, the data (denoted by black hyperbolas) consist of both primary and multiple scattering events emanating from the buried black star. The migration image $m(\mathbf{x})$, or dot product (i.e., correlation) between the migration hyperbolas and data hyperbolas, is poor unless the trial image point $\mathbf{x}$ is in close proximity to the actual scatterer point. This means that the migration image will be able to distinguish, or highly resolve, two closely spaced scatterers from one another. In contrast, the migration kernel that describes a single scattering hyperbola (see Figure 5.5b) correlates somewhat well with the data hyperbolas in a relatively larger neighborhood of the actual scatterer's position. Therefore, images obtained by single scatterer migration have less resolution than those obtained by multiple scattering migration.

Unfortunately, the hyperbolas associated with the RTM migration Green's function can undergo large changes when there are small errors in the velocity model. For example, a small error in the thickness of a reverberating layer will be magnified N times in the arrival time of an Nth-order multiple. Such large changes in the MGF will lead to more blurring and artifacts in the migration image compared to the migration Green's function of standard diffraction stack migration, which only describes a single hyperbola. We therefore conclude that image quality of RTM is much more sensitive to migration velocity errors than standard diffraction stack migration or the one-way wave equation migration (OWEM) methods.

5.4 Numerical Examples of RTM

The chief benefit of RTM compared to conventional one wave equation migration (OWEM) or Kirchhoff migration (KM) methods is that all of the arrivals, including multiples and turning waves are accounted for by the finite-difference solution to the wave equation. If the migration model is accurate, then these events get relocated back to their places of origin along reflecting interfaces. This is illustrated by the diagram in Figure 5.5 and the synthetic data example in Figure 5.6; here, marine seismic reflections are migrated by the KM and RTM methods. Since the KM method only migrates primary reflections, then "prismatic" multiple reflections are ignored. Such events might help illuminate the "root" of the salt as seen in the RTM image on the right; on the other hand, accurate migration of the prismatic multiples requires the difficult task of accurately locating the root's interface! A comparison of KM, OWEM, and RTM applied to field data is shown in Figure 5.7, where the sides of the salt flank are best imaged by RTM.

It is often the case that the OWEM and KM methods produce migration images with less noise than RTM images. This is because the $U(\mathbf{x})$ and $D(\mathbf{x})$ wavefields extrapolated

[7]Appendix B describes the similarity between pattern matching and migration.

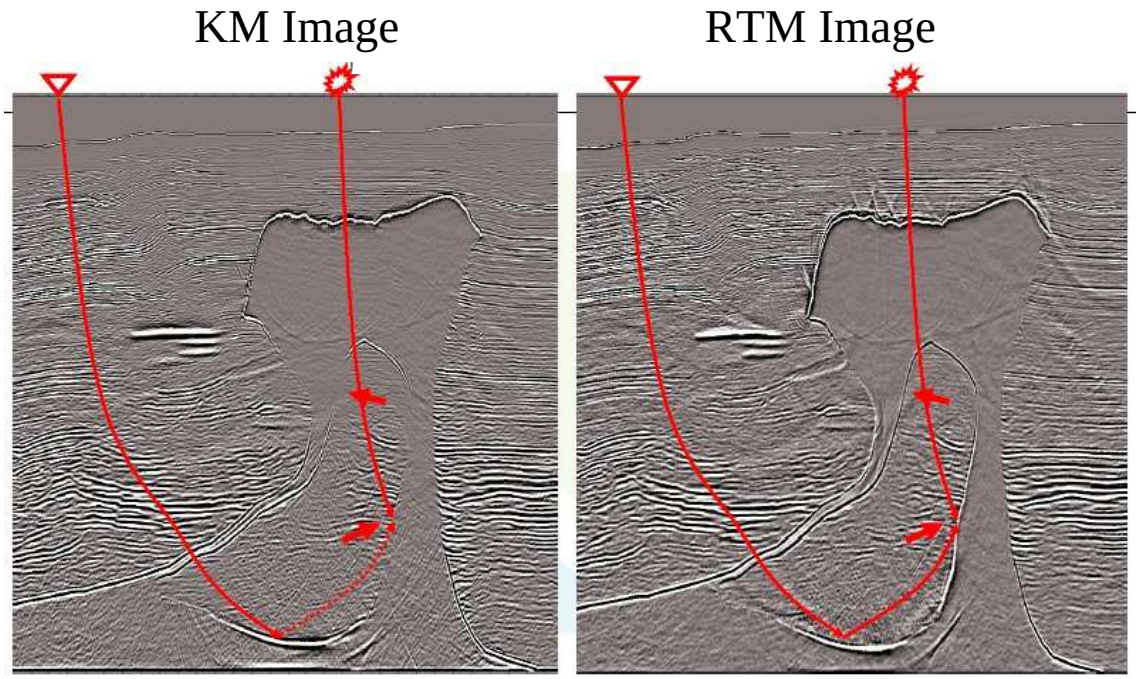


Figure 5.6: Images obtained by (left) KM and (right) RTM methods applied to synthetic acoustic data (courtesy of Yu Zhang of CGGVeritas).

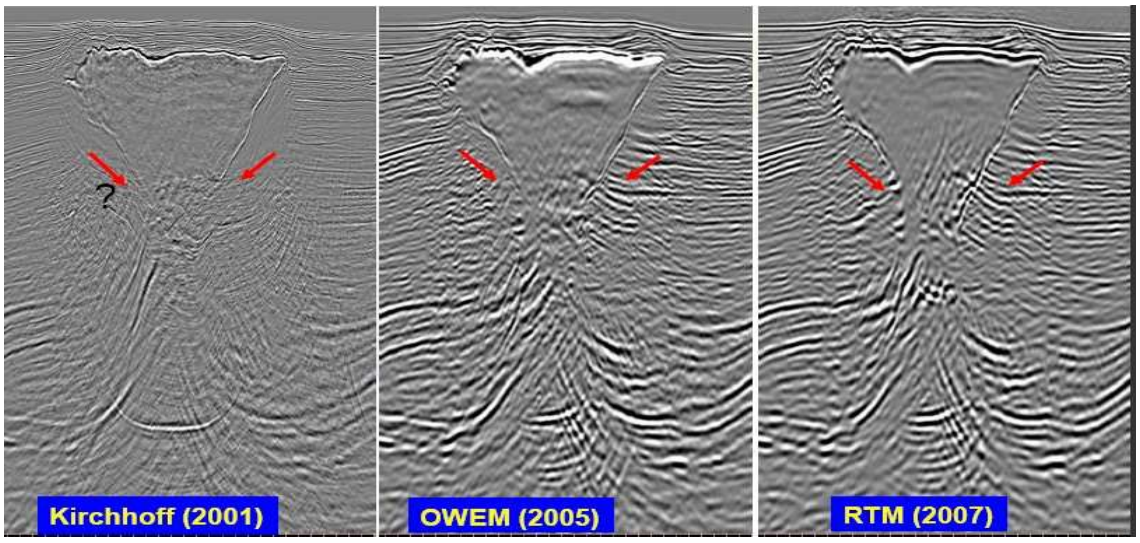


Figure 5.7: Images obtained by (left) KM, (middle) OWEM, and right) RTM methods applied to marine data (courtesy of Yu Zhang of CGGVeritas).

by OWEM methods are, respectively, purely upward and downward extrapolated waves; therefore they only have a non-zero dot product for $\mathbf{x}$ at the reflecting interfaces. This is not the case with RTM because the finite-difference solutions to the wave equation can inadverdantly generate both upgoing and downgoing waves in the extrapolation process. For example, multiple reverberations can get excited by extrapolating a plane wave through a strongly reflecting slab of limestone. Another example is that a strong diving wave can be recorded in the data. RTM backpropagation of this diving wave will show strong correlations with the forward propagated diving wave everywhere along the wavepath, not just along the bottom part of the diving ray. The result is unwanted migration energy away from the reflecting interfaces, as illustrated by the low-wavenumber artifact seen in Figure 5.8b. To suppress this artifact a high boost filter can be applied to the data, such as spatially convolving a finite-difference approximation to the Laplacian operator to the raw RTM image (see Figure 5.8c). Another remedy, like all migrations, is to eliminate the direct wave in the data by modeling it and adaptively subtracting it from the data. A summary of the

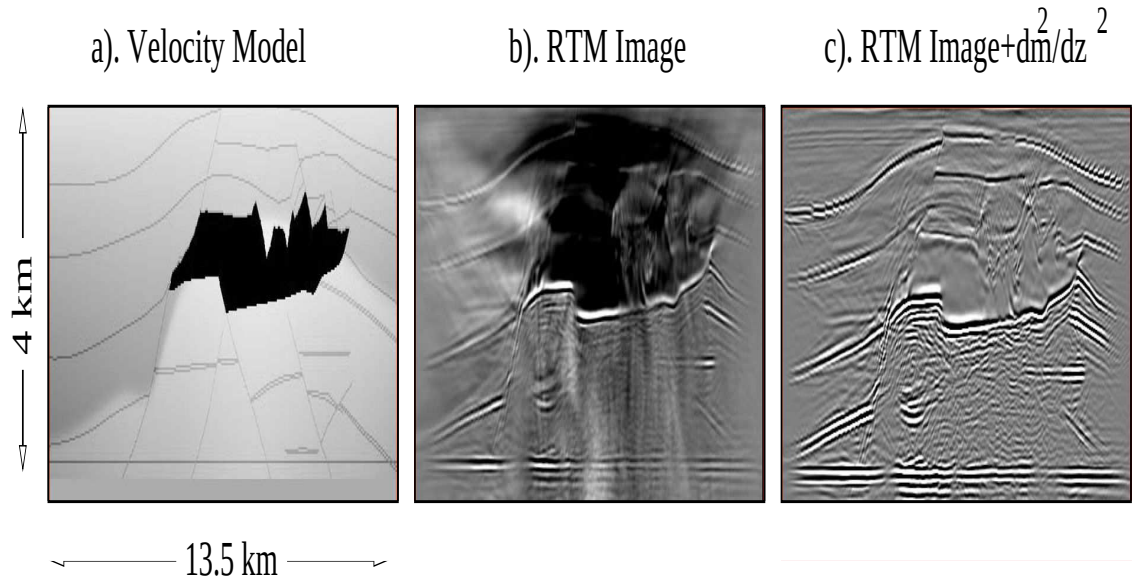


Figure 5.8: Images after b) RTM and c). RTM followed by a second derivative in the depth direction. Applying a Laplacian to the RTM image did not significantly improve the image compared to the c) image (courtesy of Wei Dai).

steps for practical implementation of RTM to field data is summarized in Appendix A.

5.5 Decomposition of Migration Kernel

The noise seen in the RTM image in Figure 5.8 can be attributed to the undesirable correlation of the source and backpropagated data events at places other than reflection interfaces[8].

[8] An early paper by Schuster (1993) showed that the waveform gradient (i.e., the RTM migration operator) could be decomposed into both short and long wavelength features, the long wavelength features associated with the transmitted portion of the reflection rays and the short wavelength portion associated with the rays

To clarify, assume a single reflector model so that the scattered wave Green's function is represented by $G(\mathbf{x}|\mathbf{s})$. It can be decomposed into reflected $G(\mathbf{x}|\mathbf{s})^{refl.}$ and direct $G(\mathbf{x}|\mathbf{s})^{dir.}$ wave components in the first layer as illustrated in Figure 5.9:

$$\begin{aligned} \mathbf{x}\epsilon\ interface + interbed; \quad G(\mathbf{x}|\mathbf{s}) &= G(\mathbf{x}|\mathbf{s})^{refl.} + G(\mathbf{x}|\mathbf{s})^{dir.}, \\ \mathbf{x}\epsilon\ interface + interbed; \quad G(\mathbf{x}|\mathbf{g}) &= G(\mathbf{x}|\mathbf{g})^{refl.} + G(\mathbf{x}|\mathbf{g})^{dir.}, \end{aligned} \tag{5.15}$$

and $\mathbf{x}$ is the trial image anywhere along the specular portion of the reflection ray. The

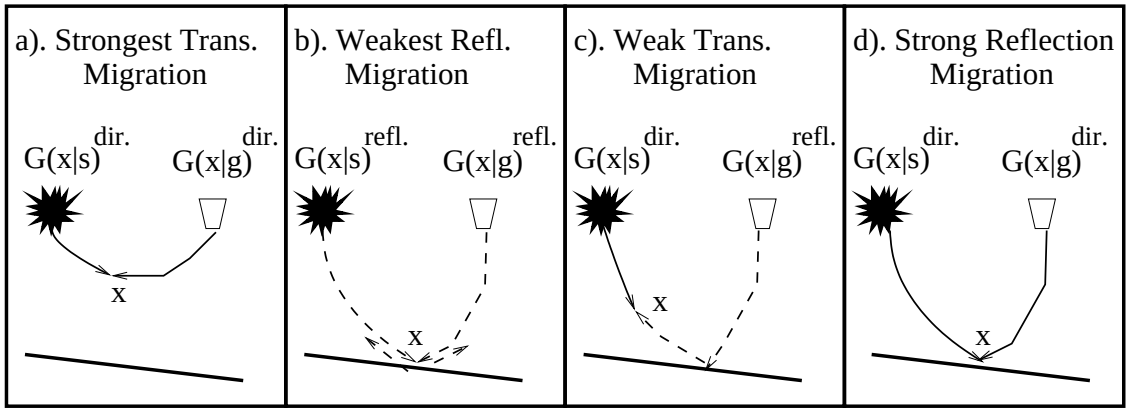


Figure 5.9: Raypaths associated with product terms of $[G(\mathbf{x}|\mathbf{g})^{dir.} + G(\mathbf{x}|\mathbf{g})^{refl.}][G(\mathbf{x}|\mathbf{s})^{dir.} + G(\mathbf{x}|\mathbf{s})^{refl.}]$. The phase of these products at certain trial image points $\mathbf{x}$ will annihilate the phase of either $D(\mathbf{g}|\mathbf{s})^{dir.}$ or $D(\mathbf{g}|\mathbf{s})^{refl.}$, but not both. The diving wavepath in a). can lead to strong artifacts in the RTM image.

contributions to the migration image $m(\mathbf{x})$ will be along the raypaths where the phase of $G(\mathbf{x}|\mathbf{s})^* G(\mathbf{x}|\mathbf{g})^*$ cancels that of $D(\mathbf{g}|\mathbf{s})$. This statement is quantified by inserting equation 5.15 into the RTM equation 5.14 to give

$$\begin{aligned} m_{mig}(\mathbf{x}) \;=\; & \omega^2 \int_{data\ space} dgds D(\mathbf{x}_g|\mathbf{x}_s)[\\ & \overbrace{G(\mathbf{x}|\mathbf{g})^{dir.} G(\mathbf{x}|\mathbf{s})^{dir.}}^{strongest\ direct\ trans.\ mig.} \quad x\ \epsilon\ direct\ raypath \\ & + \overbrace{G(\mathbf{x}|\mathbf{g})^{refl.} G(\mathbf{x}|\mathbf{s})^{refl.}}^{weakest\ refl.\ mig.} \quad x\ \epsilon\ interface \\ & + \overbrace{G(\mathbf{x}|\mathbf{g})^{refl.} G(\mathbf{x}|\mathbf{s})^{dir.}}^{weak\ src-side\ trans.\ mig.} \quad x\ \epsilon\ src-side\ interbed\ raypath \end{aligned}$$

intersecting at the reflection interface. This analysis was done by decomposing the Green's functions into transmitted, upgoing, and downgoing rays. These different components are weighted by different powers of the reflection coefficient and so gave rise to an imbalance in the weighting of different components of the gradient. Balancing these terms by reweighting them was suggested as a means for accelerating convergence and reducing artifacts. Later, a paper by Liu et al. (2007) also decomposed the Green's functions into one-way components and deduced that applying the imaging condition to appropriate components could lead to less noise in the RTM image. Their numerical tests validated this conjecture.

$$+ \overbrace{G(\mathbf{x}|\mathbf{g})^{dir.} G(\mathbf{x}|\mathbf{s})^{refl.}}^{weak\ geo-side\ trans.\ mig.} \quad x \in geo-side\ interbed\ raypath$$

$$+ \overbrace{G(\mathbf{x}|\mathbf{g})^{dir.} G(\mathbf{x}|\mathbf{s})^{dir.}]^*}^{strong\ refl.\ mig.} \quad x \in interface, \tag{5.16}$$

where $\mathbf{x}$ is the trial image point. Each of the terms in the above integrand has a unique physical interpretation as illustrated in Figure 5.9.

- The first term $[G(\mathbf{x}|\mathbf{g})^{dir.}\ G(\mathbf{x}|\mathbf{s})^{dir.}]^*$ in equation 5.16 annihilates the phase of $D(\mathbf{g}|\mathbf{s})^{dir.}$ everywhere along the direct raypath. This represents the strongest-amplitude update to $m(\mathbf{x})$ along the direct wavepath in Figure 5.9a and is valuable for updating large areas of velocity in waveform inversion. However, it is an undesirable noise feature in RTM (see low-wavenumber artifact in Figure 5.8)b and should be avoided.

- The "weakest-amplitude reflection migration" term $G(\mathbf{x}|\mathbf{g})^{refl.} G(\mathbf{x}|\mathbf{s})^{refl.}$ in Figure 5.9b is characteristic of a weak-amplitude kernel for Kirchhoff migration with the trial image point at the interface. It is denoted as "weakest" because each Green's function has a strength of $O(r)$ so when multiplied by the reflection data has strength $O(r^3)$.

- The terms $G^*(\mathbf{x}|\mathbf{g})^{dir.} G^*(\mathbf{x}|\mathbf{s})^{refl.}$ and $G^*(\mathbf{x}|\mathbf{s})^{dir.} G^*(\mathbf{x}|\mathbf{g})^{refl.}$ for $\mathbf{x} \in interbed\ raypaths$ represent weak-amplitude transmission migration kernels that cancel the phase of reflection arrivals as they transmit between the interfaces. They have a strength of $O(r^2)$ because the , e.g. Figure 5.9c, reflection events in $D(\mathbf{g}|\mathbf{s})$ and $G^*(\mathbf{x}|\mathbf{g})^{refl.}$ each have a strength of $O(r)$ while the direct wave in $G^*(\mathbf{x}|\mathbf{s})^{dir.}$ has the strength $O(1)$. This is a desirable velocity update mechanism for iterative full waveform inversion which updates the velocity distribution between reflector interfaces. It also is a source of bothersome artifacts[9] in RTM, which can be eliminated by excluding this term. The OWEM and KM methods do not produce $G^*(\mathbf{x}|\mathbf{s})^{refl.}$ so their images are not polluted with such artifacts.

- The rays in Figure 5.9d depict the kernel $G^*(\mathbf{x}|\mathbf{g})^{dir.}\ G^*(\mathbf{x}|\mathbf{s})^{dir.}$ for $\mathbf{x} \in interface$, which represents a $strong-amplitude$ Kirchhoff reflection migration kernel with strength $O(r)$. The annihilation of the phase in $D(\mathbf{g}|\mathbf{s})$ for all ω only takes place along the reflecting boundary at the specular reflection point. This is the only term desired for *true* RTM, but without filtering the migration image contains effects from all of the other terms.

The decomposition of the migration kernel in equation 5.16 can lead to filtered RTM images that are less noisy than standard RTM. An example is given in Figure 5.10 that only uses early arrivals in calculating the Green's functions, and ignores later events such as reflections. It can be seen that this strategy leads to cleaner migration images.

5.6 Summary

The equations of reverse time migration are derived, and the standard interpretation is that the migration image at $\mathbf{x}$ is formed by zero-lag correlation between the forward propagated

[9] RTM is designed to exclusively image reflector boundaries and not update velocities between interfaces.

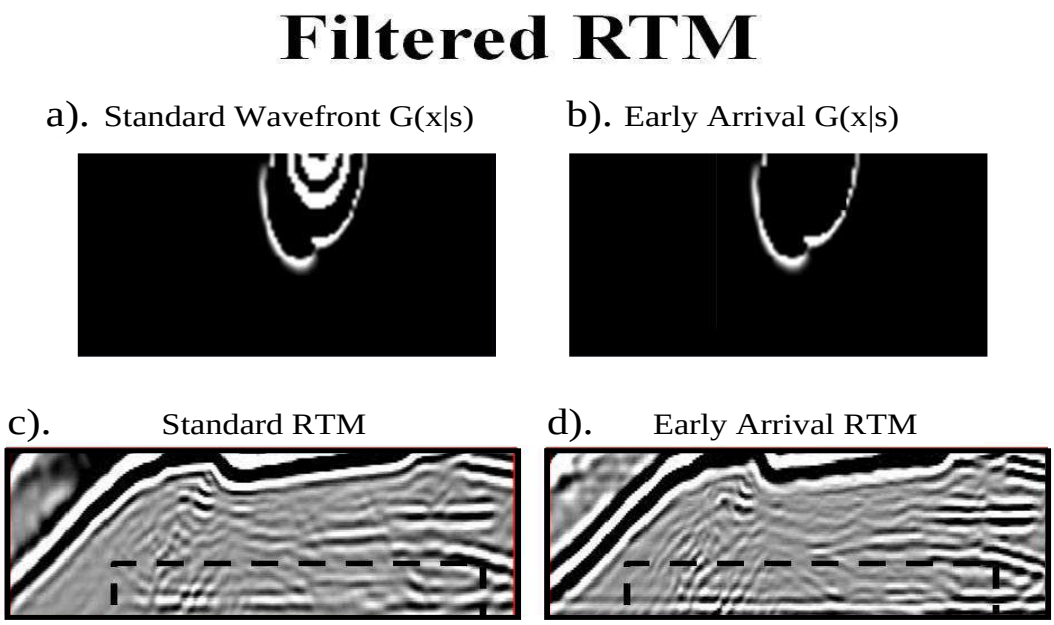


Figure 5.10: Snapshots of propagating wave for a). standard finite-difference solution to the wave equation and b). early-arrival solution to the wave equation for the SEG/EAGE salt model. The associated RTM images are in c). and d).

source field and the backpropagated scattered field. The advantage of RTM is that, unlike OWEM and KM methods, RTM accounts for all scattered events in the data if the migration velocity model is accurate enough. This can lead to illumination of salt flanks normally invisible to KM or OWEM methods, and can lead to proper focusing beneath bodies with large velocity contrasts, such as salt lenses. Another potential advantage is higher spatial resolution, which can only be achieved if the velocity model is accurate enough. The main disadvantages of RTM are that it is typically an order of magnitude more expensive than OWEM and KM methods, and its migration image is much more sensitive to errors in the velocity model.

An alternative interpretation of RTM is that it is a generalized diffraction stack migration method. The generalized migration image at a point $\mathbf{x}$ is obtained by taking the dot product of the appropriate migration operator with the data. This is a generalization of simple diffraction stack migration which sums the data over the appropriate hyperbola. Using the GDM approach, the migration kernel for RTM can be filtered to eliminate artifacts in the migration image. The drawback is that the storage requirements of the GDM kernels are enormous; e.g., a 7-dimensional matrix of GDM kernels $\Gamma(\mathbf{x}_g, \mathbf{x}_s, \mathbf{x})$ per frequency is required for a 3D problem.

The liabilities of RTM compared to the KM and OWEM methods include much slower computation speeds and a greater sensitivity to errors in the velocity model. Another liability is that there is generally much more noise in the RTM image because the backpropagated reflections correlate with the source field away from reflector boundaries, which is useful

for waveform inversion but not for migration images. Filtering methods are discussed that alleviate this problem, and a more detailed analysis is given in a later chapter.

Exercises

1. Make a movie of waves emanating from a buried point source in a 300x200 grid model with a Ricker wavelet time history for the point source. Let c=5000 ft/s and choose the dx and the peak frequency of the Ricker source wavelet so that there are about 15 points/wavelength, where the minimum wavelength is twice the peak frequency of the Ricker wavelet. The code for the zero-phase Ricker wavelet is given below, and the wavelet is delayed in time to insure causality of the source wavelet.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% NT     - input- # points Ricker wavelet
% FR     - input- Peak frequency of Ricker wavelet
% dt     - input- Temporal sampling interval
% RICKER -output- Time delayed Ricker wavelet
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
t=[0:1:NT-1]*dt-0.95/FR;RICKER=zeros(length(t));
RICKER= (1-t .*t * FR^2 *pi^2  ) .*exp(- t.^2 * pi^2 * FR^2 ) ;
```

2. Write a parsimonious MATLAB code to sequentially compute the forward field $FORW(x, z, t)$ at decreasing time steps starting from final conditions; the parsimony should be for minimizing storage of arrays. It is too costly to store and retrieve the multidimensional matrix $FORW([1 : nx], [1 : nz], [1 : nt])$; instead, a finite-difference simulation is used to compute the forward field all along the exterior spatial boundaries of the medium. These boundary values are saved for all time steps and the last two $x - z$ panels of pressure are also saved. With these saved traces, the finite-difference program can be run backwards in time, starting at the 2nd-to-last time step, to sequentially recreate the forward field $FORW([1 : nx], [1 : nz], t)$ everywhere in space and at decreasing time t.

3. Write a RTM code, using your code from a previous chapter that solves the acoustic wave equation by a finite difference method. Use RTM to migrate the reflection data generated in the previous exercise. For the forward field use the analytic Green's function in a homogeneous medium. Show movies.

4. Repeat previous exercise except use an illumination compensation method to apply to the migration image.

5. The reflection coefficient distribution is estimated by taking the reflection events measured at the surface and backpropagating them to depth where they arrived at earlier times. But what happens if the backward extrapolation operator is accidentally applied to a downgoing wave measured at $z = z_g$? The measured downgoing wave $e^{-ik_z z_g}$ will be extrapolated as $e^{-ik_z(z+|\Delta z|)}$, which predicts the *future* behavior of the downgoing field in depth from events measured at the surface. Unfortunately, such predicted events have nothing to do with estimating the reflectivity distribution $r(\mathbf{x})$

and instead will lead to artifacts in the migration image. This suggests that the measured data should undergo upgoing and downgoing separation prior to migration, and only the upgoing data migrated. Propose a reflector model, and use the FD code to generate pressure fields that are a sum of upgoing and downgoing reflections (include a free surface); also generate upgoing only data (exclude free surface). Separately migrate the two types of data, and compare the two results.

6. Prove the statement "$e^{-ik_z \Delta z}$ is the forward extrapolation operator for downgoing waves."

Appendix A: Practical Implementation of RTM

Several improvements can be useful for creating RTM images with few artifacts: direct wave muting, illumination compensation (i.e., preconditioning), smoothing of the velocity model, tapering of data at edges of aperture, and up- and downgoing wave separation.

1. Preconditioning. An ill-posed problem is characterized by slow convergence with a gradient optimization method. Instead of a bullseye minimum, the global minimum might have resided at the bottom of a long narrow valley. Preconditioning is recommended to deform the long valleys into a bullseye so that steepest descent method can quickly converge to the bullseye. As will be shown in a following chapter, one pass of RTM is equivalent to the 1st iteration of a steepest descent minimization of the data misfit function. Therefore, a preconditioning operator should be applied to the RTM algorithm, also known as illumination compensation.

   The simplest preconditioning operator is that of Beydoun and Mendes (1989), who uses the reciprocal of the diagonal elements of the Hessian matrix $[\mathbf{L}^T\mathbf{L}]$

$$[\mathbf{L}^T\mathbf{L}]_{ij}^{-1} \approx \frac{\delta(i-j)}{[\mathbf{L}^T\mathbf{L}]_{ii}}. \tag{5.17}$$

   Either a straight ray approximation or ray tracing method can be used to estimate the Hessian.

   The formula for the diagonal term in the Hessian can be derived from the Lippmann-Schwinger equation (i.e., $\mathbf{D} = \mathbf{Lm}$) as:

$$D(\mathbf{x}_d) = 2\omega^2 \int_{V_m} \Gamma(\mathbf{x}_d, \mathbf{x}_m) m(\mathbf{x}_m) d\mathbf{x}_m \tag{5.18}$$

   where $\mathbf{x}_d$ is the position vector composed of $(\mathbf{x}_g, \mathbf{x}_s)$, $\mathbf{x}_m$ denotes the position vector in the model space, and $\Gamma(\mathbf{x}_d, \mathbf{x}_m) = G(\mathbf{x}_g|\mathbf{x})G(\mathbf{x}|\mathbf{x}_s)$. The integration volume V_m is over that of the model. Multiplying equation 5.18 by $\Gamma(\mathbf{x}_d, \mathbf{x}'_m)^*$ and integrating over the data volume denoted by V_d, yields:

$$\begin{aligned} \hat{m}(\mathbf{x}'_m) &= \int_{V_d} \Gamma(\mathbf{x}_d, \mathbf{x}'_m)^* D(\mathbf{x}_d) d\mathbf{x}_d \\ &= 2\omega^2 \int_{V_m} [\int_{V_d} \Gamma(\mathbf{x}_d, \mathbf{x}'_m)^* \Gamma(\mathbf{x}_d, \mathbf{x}_m) d\mathbf{x}_d] \; m(\mathbf{x}_m) d\mathbf{x}_m. \end{aligned} \tag{5.19}$$

The integration term in square brackets can be considered as the kernel for the $\mathbf{L}^T\mathbf{L}$ operator, where $\mathbf{D} = \mathbf{Ls}$ and $\mathbf{L}$ is the forward modeling operator. The diagonal component of $[\mathbf{L}^T\mathbf{L}]$ is given by by setting $\mathbf{x}'_m = \mathbf{x}_m = \mathbf{x}_i$:

$$[\mathbf{L}^T\mathbf{L}]_{ii} \quad = \quad \int_{V_d} |\Gamma(\mathbf{x}_d|\mathbf{x}_i)|^2 d\mathbf{x}_d. \tag{5.20}$$

Thus, the *ith* diagonal component of the Hessian is the integration of the squared amplitudes of the wavepaths that visit the *ith* slowness cell for all source-receiver pairs. This quantity can be numerically computed and the inverse to the Hessian can be approximated as:

$$[\mathbf{L}^T\mathbf{L}]_{ij}^{-1} \quad \approx \quad \frac{\delta_{ij}}{[\mathbf{L}^T\mathbf{L}]_{ii}}. \tag{5.21}$$

Ray tracing can be used to approximate this preconditioning operator because in the high frequency approximation $|G(\mathbf{x}|\mathbf{x}')| = 1/|\mathbf{x}_{xx'}|$ where $|\mathbf{x}_{xx'}|$ is the raypath length between $\mathbf{x}$ and $\mathbf{x}'$. Therefore, for a single source and receiver $[\mathbf{L}^T\mathbf{L}]_{xx}^{-1} \approx 1/(|\mathbf{x}_{xs}|^2|\mathbf{x}_{xg}|^2)$.

2. If a wave equation method is used, then the computed forward field can be used to find the illumination compensation term $I(\mathbf{x})$ in equation 5.11 for each shot gather; this is denoted as source-side compensation. An alternative strategy is to use receiver-side compensation[10] where the backpropagated reflected data is used to find the redatumed field at each model point; and the total energy of this backpropagated field at $\mathbf{x}$ is used to get the illumination compensation term $I(\mathbf{x})$ for a migrated shot gather. See Figure 5.11 for an example of applying $I(x)$ to each point of an RTM image obtained from one shot gather.

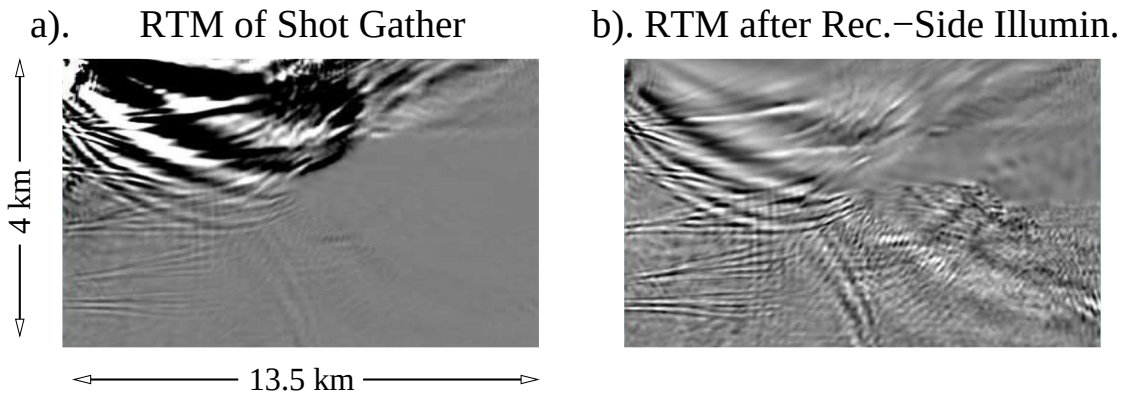


Figure 5.11: RTM image of a shot gather a). before and b). after receiver-side illumination compensation (courtesy of Wei Dai).

3. Velocity model smoothing. A smooth velocity model should be used for RTM, otherwise reflections from impedance discontinuities will be generated and create artifacts

[10] Wei Dai claims that this idea originated from a SEP report at Standford University.

in the migration image. On the other hand, if the reflection boundaries are accurately known then the density can be adjusted so that the impedance ρv is the same across a velocity discontinuity. This leads to a zero reflection coefficient at a normal incidence angle.

4. Taper the edges of the shot gather so that the traces are weighted by zero at the end and smoothly weighted by (e.g., use a cosine taper) to the trace about one wavelength from the edge. Tapering reduces migration artifacts from a sharply truncated data aperture of traces.

5. Mute the nearly horizontal traveling parts of the direct arrival. If the direct arrival is not muted prior to migration, then there will be strong artifacts along the diving wavepath (see Figure 5.9a). One strategy is to model the direct waves in a smooth velocity model and adaptively subtract them from the data.

6. Apply a high-pass filter to the migration image to eliminate the low-wavenumber noise. Sometimes a simple derivative in the depth direction in the migration image will be sufficient, but a Laplacian filter or 2nd-order vertical derivative filter can be used (see Figure 5.8c). An alternative strategy is to numerically detect dips along the interfaces and apply directional derivatives perpendicular to the interface boundary to eliminate long wavelength interbed artifacts.

Appendix B: Migration as a Pattern Matching Operation

The similarity between two digital photos (each assumed to be a 100x100 pixelated image in the x-y plane) can be quantified by representing each photo by a $10,000x1$ vector and taking their dot product. If the photos are very similar then the dot product will yield a sum of mostly positive numbers to give a very high correlation coefficient. Conversely, the dot product between dissimilar photos will yield a sum of both positive and negative numbers to give a small correlation coefficient. Taking dot products of photos is a common *pattern matching* operation.

Summing the data over migration hyperbola in $x-t$ space can also be thought of as a pattern matching operation. Figure 5.5a depicts the primary migration curves in $x-t$ space as pink and green curves, where each colored curve corresponds to a different trial image point with the same color. Summing the data over a curve is equivalent to a 2-D dot product between the migration operator image and the data image. If the trial image point is near the actual scatterer, then the, e.g. pink, migration operator in Figure 5.5a matches the data very well. Hence, the migration image at that trial image point has a high value. At other trial image points, the pattern of the migration operator correlates poorly with the data so the correlation is small to give a small value in the migration image.

Migration operator for multiple arrivals. In theory, migrating many types of events with different arrival angles to their common reflector point leads to a better resolution at that point, and a cleaner migration image. This is similar to looking at a diamond from different view angles, each new view angle revealing a new facet of the gem. To achieve this extra resolution with seismic images, one can tune a diffraction migration

operator to migrate both primary reflections *and scattered multiples*[11] A representative primary+multiple migration operator is illustrated in Figure 5.5b, where, for a trial image point, the summation of energy is along the hyperbolic curves (e.g., the solid and dashed pink curves or the solid and dashed green curves) that represent the traveltimes for both primaries and multiples. For the correct trial image point at the black star, a huge of amount of seismic energy gets placed at the scatterer's position by primary+multiple migration compared to that for primary migration.

From a pattern matching point of view, the complicated pattern of the primary+multiple migration operator correlates well with data only in a small neighborhood of the actual scatterer's point; thus the image resolution is very good. Compare this to matching the simple primary migration operator to the data in Figure 5.5a; there is a relatively large neighborhood around the actual scatterer that gives a good match between the operator's pattern and the actual data. This means a migration image with worse resolution compared to the primary+multiple migration image.

The disadvantage of primary+multiple migration is that its migration image is especially sensitive to errors in the migration velocity model. Small migration velocity errors tend to give a noisier image compared to the primary migration image. In fact, the primary+multiple migration operator can be shown to be identical to that for reverse time migration, a subject to be discussed in a later chapter.

[11] A multiple migration algorithm can be constructed by ray tracing the traveltimes for *both primaries and multiples and including the extra summations in the diffraction stack migration formula.* An alternative method is by using a finite-difference method to solve the wave equation to get appropriate Green's functions.

Chapter 6

Phase Shift Migration Methods

6.1 Introduction

Prestack migration techniques for subsurface imaging are widely used in oil exploration and play an important role in imaging the complex subsurface structure image. The ray-based diffraction-stack and Kirchhoff migration (KM) methods are presently considered the most popular and flexible technique for 3D migration with generally good image quality. The main reason is that the KM method has the capability of target-oriented processing, steep dip imaging and somewhat efficient computation. The ray-based migration methods usually use single path ray tracing to get traveltimes, but sometimes multi-path arrivals are needed for proper imaging in complex areas. Therefore KM typically generates a poor migration in areas with complex geological structure unless multiarrivals are properly accounted for. In order to remedy such drawbacks and increase the image quality, new ray migration methods such as Gaussian Beam migration were developed to consider the multi-path arrivals and caustics.

Although the ray-based migration methods, such as Kirchhoff migration (KM) and Born migration/inversion, are considered the most popular imaging tools for 3D migration with generally good image quality, the more expensive wave-equation migration methods can produce more accurate images in complex area. It has drawn keen attention from oil industry (Huang et al., 2000; Lee et al, 1991; Bonomi and Cazzola, 1999; Ristow and Ruhl, 1994; Sun et al., 2001; Stoffa et al., 1990; Wu and Jin, 1997; Claerbout, 1974). As an example, reverse-time migration solves the two-way wave equation for imaging, which is accurate but at the cost of an increase in computation time. For shot migration, one needs to both forward propagate the source and backward propagate the receiver wavefield. In order to increase computational efficiency, the forward propagation can be implemented by a ray tracing technique More widely-used wave equation methods are based on more efficiently solving the one-way acoustic wave equation.

The one-way wave equation imaging methods computed in the frequency domain are known as phase-shift migration methods. Here we describe the basic phase-shift migration algorithm, which is strictly valid for vertically layered media, but are still useful for media with mild lateral variations in velocity. Stronger lateral variations in velocity can be handled by a later generation of phase-shift methods, such as the Split-Step migration and Finite Difference Fourier migration.

6.2 Phase Shift Migration

Zero-offset seismic traces do not provide an accurate picture of the subsurface layers when there is a great deal of lithological complexity. For example, dips in the seismic section are not the true dips of dipping reflectors, grabens look like bowties in the seismic record, and point scatterers appear as diffraction hyperbolas. To correct for this distortion we apply migration to the zero-offset seismic data. A family of wave equation-based migration methods is known as phase shift migration, which applies phase shifts to the data in the Fourier domain. They all assume upcoming reflections only and only account for one-way wave propagation. For $v(z)$ media, we have standard phase shift migration (Gazdag, 1978; Gazdag and Sguazzero, 1984), for mild velocity contrasts we have split-step Fourier migration, and for moderate lateral velocity contrasts we have Finite Difference Fourier migration.

The starting PDE is the Helmholtz equation given by

$$\nabla^2 P(x,z,\omega) + \omega^2/v^2 P(x,z,\omega) \quad = \quad 0. \tag{6.1}$$

where $v = v(x,z)$. This equation can be rearranged and factorized as a concatenation of two square root operators:

$$\frac{\partial^2 P(x,z,\omega)}{\partial z^2} \quad = \quad \overbrace{-\sqrt{k^2+\partial^2/\partial x^2}}^{upgoing\ waves} \ \overbrace{\sqrt{k^2+\partial^2/\partial x^2}}^{downgoing\ waves} P(x,z,\omega). \tag{6.2}$$

The above equations admit two independent solutions, the downgoing $P^+ = Be^{i\kappa_{z_o} z}$ and upgoing $P^- = Ae^{-i\kappa_{z_o} z}$ solutions, where

$$\kappa_{z_o} \quad = \quad \sqrt{(\omega/v_o)^2 + \partial^2/\partial x^2}. \tag{6.3}$$

Differential operators in the square root are not operationally meaningful unless the square root is expanded in some series approximation (Claerbout, 1985). Thus they are to be interpreted in the sense that they symbolize some expansion.

The upgoing and downgoing portions of the coefficient in equation 6.2 honor are associated with either the upgoing or downgoing wave equations:

$$\frac{\partial P^-}{\partial z} \quad = \quad -i\sqrt{k_o^2 + \partial^2/\partial x^2}\ P^-, \tag{6.4}$$

$$\frac{\partial P^+}{\partial z} \quad = \quad i\sqrt{k_o^2 + \partial^2/\partial x^2}\ P^+. \tag{6.5}$$

The solution to equation 6.4 is

$$P^-(x,z,\omega) \quad = \quad Ae^{-i\kappa_z z}, \tag{6.6}$$

where A is an arbitrary constant and $e^{-i\kappa_z z}$ is the phase-shift term. The upcoming wave equation 6.4 and the corresponding phase shift operator will be used to derive the phase shift, SSF and FFD migration algorithms.

6.2.1 Phase Shift Migration

Acoustic data $p(x, z = 0, t)$ are measured along the plane at $z = 0$, and Fourier transformed in time to give $P(x, z, \omega)$. it is assumed that the earth is a layered $v(z)$ medium and only upcoming primary reflections $P(x, z, \omega) \rightarrow P^+(x, z, \omega)$ are recorded so that equation 6.4 is the governing equation. The direct waves have been muted and the surfaces waves+multiples+converted waves are filtered out. In other words, the exploding reflector model is valid. The z-axis increases upward and $e^{-i\kappa_z z}$ corresponds to upward traveling waves.

The goal is to use phase-shift migration and downward continue 2-D data $p(x, z = 0, t)$ to get $p(x, z, t)$ for $z < 0$. The reflectivity is estimated by invoking the $t = 0$ imaging condition $r(x, z) \approx p(x, z, t = 0)$. Accordingly, the phase shift migration algorithm consists of 3 steps.

1. Assume a $v(z)$ medium discretized into N homogeneous layers with thickness dz, each with velocity given by v_n. Within the nth layer, the pressure field $P(x, z, \omega)$ satisfies the homogeneous wave equation with constant velocity v_n. Under a Fourier transform in the x coordinate, the upcoming wave equation 6.4 becomes:

$$\frac{d\tilde{P}^-}{dz} = -i\sqrt{k_n^2 - k_x^2}\ \tilde{P}^-, \tag{6.7}$$

because $\partial^2/\partial x^2$ transforms as $-k_x^2$; and $k_n = \omega/v_n$.

2. Defining the vertical wavenumber $k_{z_n} = \sqrt{(\frac{\omega}{v_n})^2 - k_x^2}$, the solution to the above ODE is given as $\tilde{P}^- = Ae^{-ik_{z_n}z}$ with unknown A. The boundary condition is that the solution must match the data at $z = 0$ so that $A = \tilde{P}(k_x, 0, \omega)$. Therefore the solution to equation 6.7 is

$$\tilde{P}(k_x, z, \omega)^- = \tilde{P}(k_x, 0, \omega)e^{-ik_{z_n}z} \tag{6.8}$$

where $\tilde{P}(k_x, 0, \omega)$ is the measured data at the $z = 0$ plane. The above equation is the *downward continuation* step because the phase shift term $e^{-ik_{z_n}z}$ shifts the phase of the surface data by $k_{z_n}z$ to give data at a deeper depth z for $z < 0$.

3. Since the exploding reflectors started exploding at $t = 0$ then the location of the reflectors is given by inverse transforming the solution 6.8 in k_x and ω and evaluating at $t = 0$:

$$r(x, z) \approx p(x, z, t = 0) = \sum_{k_x}\sum_{\omega} \tilde{P}(k_x, 0, \omega)e^{-i(k_x\ x\ +\ k_z z)}. \tag{6.9}$$

The condition $r(x, z) = p(x, z, t = 0)$ is known as the zero-offset *migration imaging condition.*

Question: Why do solutions to the wave equation in the form $e^{i(\omega t - k_z z)}$ correspond to upward propagating waves?

Answer: Jump on a wavefront moving upward and you notice that the amplitude or phase remains the same under your feet. Therefore, as time increases and z increases, the phase must stay the same (which it does for $e^{i(\omega t - k_z z)}$, but the phase changes for $e^{i(\omega t + k_z z)}$). In other words, as time t and z both increase, the signs of $k_z z$ and ωt must be opposite to each other if waves are to propagate in the positive z direction.

Question: Why do solutions to the wave equation in the form $e^{i(\omega t + k_z z)}$ correspond to downward propagating waves?

Answer: Similar to before, if you jump on a downward propagating wavefront t increases but z decreases so the phase/amplitude of $e^{i(\omega t + k_z z)}$ stays the same. In other words, as time t increases and z decreases, the signs of $k_z z$ and ωt must be be the same if the waves are to propagate in the negative z direction.

Remarks:

- We will often specify the sign of k_z as:

$$k_{z_n} \quad = \quad sgn(\omega)\sqrt{\omega^2/v_n^2 - {k_x}^2}, \tag{6.10}$$

  where $sgn(\omega) = 1$ if the sign of ω is positive, otherwise it is negative. The upward traveling waves will demand a negative sign in the exponent of the continuation operator $e^{-isgn(\omega)|k_{z_n}|z}$ because we assume an IFFT kernel of $e^{i\omega t}$. And the $sgn(\omega)$ is to insure that the continuation operator will still be for upward moving waves when $\omega < 0$ or $\omega > 0$ (recall integration limits for a Fourier transform is from ∞ to $+\infty$).

- The $k_z z$ term in the exponent of the downward continuation equation 4 will become imaginary when $k_x > \omega/c$; these wavenumbers are called evanescent wavenumbers and correspond to exponentially decaying or growing waves. Evanescent waves do not propagate, they evanesce (i.e., disappear) in certain directions! This will cause the continuation operation to be unstable if you are continuing the data towards the scattering point. However, it will be ok if you continue data away from source point (i.e., upward continuation) because now the sign of the exponent will be negative and the evanescent waves will decay. Therefore you must restrict the summation in equation 5 to exclude evanescent waves for downward continuation. This approximation does not damage your migrated section too much!

- Upward continuation is when we extrapolate the seismic data at $z = 0$ to above the measurement plane. Upward continuation is a way of getting rid of local topographic variations and obtaining the field on a flat measuring plane above the topography $z > 0$. It also is used as a way to eliminate data problems due to irregular surface topography. That is, you downward continue to the bottom of the shallow irregular

subsurface weathering layer. Then you replace the the weathering layer velocity by a fast subweathering velocity and upward continue to a flat measuring plane. This procedures helps to eliminate diffractions and time shifts due to waves passing through the weathering interface.

- A fragment of a pseudo-MATLAB migration code looks like the followings script.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% ZO Phase Shift Migration of Data
%
%(nt,dt) - input- No. samples and interval along time  axis
%(nx,dx) - input- No. samples and interval along offset  axis
%(nz,dz) - input- No. samples and interval along depth  axis
% c      - input- Migration velocity
% Datax  - input- Freq-Kx Fourier transform of data
% M(z,x) -output- ZO Migration image in (x,z) domain
%
% Author: Ruqing He
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
F=[0:(nt+1)/2-1,-(nt+1)/2+1:-1]/nt;   % Define freq. range
Kx=[0:(nx+1)/2-1,-(nx+1)/2+1:-1]/nx;  % Define Kx range
[gkx,gf]=meshgrid(Kx,F);              % Define range matrices Kx and f
FKz=(gf/c/dt).^2-(gkx/dx).^2; S=FKz>0;% Exclude decaying waves
FKz=2*pi*sign(gf).*sqrt(FKz.*S); % Argument of phase shift exponential
nz=nt; dz=dt*c;                  % migration model spacing
M=zeros(nz,nx);                  % Zero reflectivity model
Datax=Datax.*(FKz~=0); dps=exp(i*FKz*dz);
for iz=2:nz                      % Loop over depth
    Datax=Datax.*dps;            % Phase shift Data to dz deeper
    M(iz,:)=sum(real(ifft(Datax,[],2))); % IFFT in kx and sum over freq.
end
```

6.3 Prestack Phase Shift Migration

For prestack data, the upcoming $P(x,0,\omega)^-$ must be downward continued to give $P(x,z,\omega)^-$, and the downgoing source field at the surface $S(x,0,\omega)^+$ must be downward continued to give $S(x,z,\omega)^+$. The extrapolation operator for downward continuing the downgoing source field is $e^{ik_z z}$, that is $S(k_x,z,\omega)^+ = S(k_x,0,\omega)^+ e^{ik_z z}$; while the downward continuation operator for the data is $e^{-ik_z z}$, that is $P(k_x,z,\omega)^- = P(k_x,0,\omega)^- e^{-ik_z z}$,

At the reflector, the upgoing reflection event was created at the same *time* as the downgoing source field struck the reflector; thus $S(k_x,z,\omega)^+$ and $P(k_x,z,\omega)^-$ should have the same phase at $(x,z)\epsilon\ Reflector$. We would like their phases to cancel at the reflector so upon summation over all frequencies in (x,z,ω) will lead to constructive interference at all frequencies. This can be achieved by taking the product of $P(k_x,z,\omega)^-$ with the conjugate

of $S(k_x, z, \omega)^+$ to give the prestack migration imaging condition:

$$R(x, z, \omega) = S(x, z, \omega)^* P(x, z, \omega). \tag{6.11}$$

Conjugating $S(x, z, \omega)$ insures that its phase will be equal and opposite to the phase of the continued data $P(x, z, \omega)^-$ at the reflector point. Summing the reflectivity estimate $R(x, z, \omega)$ for all ω we have

$$\begin{aligned} r(x, z) &= \sum_{\omega} R(x, z, \omega), \\ &= \sum_{\omega} S(x, z, \omega)^* P(x, z, \omega). \end{aligned} \tag{6.12}$$

In summary, prestack phase shift migration is the same as poststack migration in that the data are continued downward, but in addition the source field is also downward continued. The product $S(x, z, \omega)^* P(x, z, \omega)$ is then summed over all frequencies to give the migration image.

6.4 Phase Shift Migration for Lateral Velocity Variations

The standard phase-shift migration method assumed no lateral velocity variations. To account for mild lateral velocity variations, the Split-Step Fourier method (SSF) was developed.

The starting point is to approximate the square root in equation 6.4 by splitting κ_z into a homogeneous term $\kappa_{z_o} = \sqrt{(\omega/v_0)^2 - \partial^2/\partial x^2}$ and the perturbation term $\Delta\kappa_z = \kappa_z - \kappa_{z_o}$:

$$\begin{aligned} \kappa_z &= \kappa_{z_o} + \Delta\kappa_z, \\ &\approx \overbrace{\kappa_{z_0}}^{Phase-Shift} + \overbrace{\omega(1/v - 1/v_o)}^{Split-Step} + \overbrace{other\ term}^{FFD}, \end{aligned} \tag{6.13}$$

where the *other term* is important for propagation angles that "significantly" depart from the vertical. The SSF (Stoffa, 1990) method corresponds to dropping the "other term" so that equation 6.14 becomes

$$\kappa_z \approx \kappa_{z_o} + \overbrace{\omega(1/v - 1/v_o)}^{Thin-Lens}, \tag{6.14}$$

where the extrapolation operation in equation 6.6 becomes

$$P(x, dz, \omega)^- = e^{i\omega(1/v - 1/v_o)dz} \mathcal{F}^{-1}[e^{i\kappa_{z_o} dz} \tilde{P}(k_x, 0, \omega)^-]. \tag{6.15}$$

The exponential outside the square brackets represents the phase shift correction for vertically propagating plane waves in parts of the medium that depart from the v_0 velocity. The term in the square brackets can be computed in the (k_x, z, ω) domain while the $e^{i\omega(1/v - 1/v_o)dz}$ can be applied in the (x, z, ω) domain using the actual velocity $v(x, z)$ at each (x, z) point. The correction term is also called the thin lens term (Claerbout, 1985) because this is the phase shift photons undergo in passing through a thin lens when propagating along the optic axis.

In summary, Split Step Fourier migration consists of three steps:

$$Step\ 1:\ Extrapolation\ in\ (k_x, dz, \omega):\ \tilde{P}(k_x, z, \omega)^- \ = \ e^{ik_{z_o}dz}\tilde{P}(k_x, 0, \omega)^-,$$

$$Step\ 2:\ Inverse\ Fourier\ Transform:\ \ \tilde{P}(k_x, dz, \omega)^- \ \rightarrow \ P(x, dz, \omega)^-,$$

$$Step\ 3:\ Phase\ Correction\ in\ (x, z, \omega):\ P(x, dz, \omega)^- \ = \ P(x, dz, \omega)^- e^{i\omega(1/v - 1/v_o)dz}.$$

The above steps demonstrate that the implementation of SSF migration includes two parts: Perform wavefield extrapolation of the data in the frequency-wavenumber domain, then transfer the wavefield into the frequency-space domain and apply the phase-shift or phase correction which accounts for the lateral velocity variations. For strong velocity lateral variations, this single perturbation is not enough for imaging and more than one reference velocity is required as the strategy used in PSPI to get a more accurate result. However, the penalty is an increase in the computational cost (Kessings, 1992; Huang et al, 1999).

Phase Shift Plus Interpolation Method

PSPI method is a phase-shift-like method for dealing with strong lateral velocity variations. (Gadgaz, 1984). The basic idea of PSPI is to introduce several reference velocities to account for the lateral velocity variation in each extrapolation step and obtain the multi-reference wavefields in the frequency-wavenumber domain. Based on the relationship of the local velocity and reference velocity, the final migration result is obtained by interpolating the reference wavefields in the frequency-space domain. The basic formulas are:

$$P_0(x, y, z, \omega) = P(x, y, z, \omega).e^{i\frac{\omega}{v(x,y,z)}dz}, \tag{6.16}$$

and

$$P'(k_x, k_y, z + dz, \omega) = P_0(k_x, k_y, z, \omega).e^{i(k'_z - \frac{\omega}{v'_{ref}})dz}, \tag{6.17}$$

where k'_z is obtained using the reference velocity. After the reference wavefield is Fourier-transformed back to the frequency-space domain, the final migration result is obtained by linear interpolation.

Obviously, the choice of the reference velocities is a crucial task for PSPI migration mainly because the cost of PSPI is proportional to the number of reference velocity values used in each extrapolation step. In order to decrease the cost, the adaptive strategy of (Bagaini, 1995) is adopted for the numerical results shown in this chapter. This adaptive strategy of selecting reference velocities not only reduces the cost of PSPI, but it also computes the reference velocities according to the distribution of velocities. More reference velocities will be used when the lateral velocity variation is strong and fewer velocity values will be used when the velocity contrast is small.

Fourier Finite Difference Method

Even though SSF and PSPI can handle lateral velocity variations, they will gives less accurate results when the lateral velocity variation is strong. The correction term is only a zero-order approximation to the one-way wave equation and propagates accurately only at small angles. For this problem, we need a more accurate approximation for the dispersion equation by adding additional terms, such as the extended local Born-Fourier migration and pseudo-screen propagator methods (Huang et al., 1999). In 1994, Ristow and Ruhl proposed a Fourier finite-difference method which is the combination of the phase-shift method in the frequency-wavenumber domain and the FD method in the frequency-space domain.

The starting point for FFD is to retain the *other term* in equation 6.14 so that extrapolation from one level to the next is given by

$$P(x,dz,\omega)^- \;=\; e^{i(other\ term)dz}e^{i\omega(1/v-1/v_o)dz}e^{i\kappa_{z_o}dz}\mathcal{F}^{-1}[\tilde{P}(k_x,0,\omega)^-]. \qquad (6.18)$$

Applying $e^{i(other\ term)dz}$ is a finite-difference calculation in the (x,z,ω) domain (see Appendix). Thus, FFD migration consists of four steps:

$$\begin{aligned}
&Step\ 1:\ Extrapolation\ in\ (k_x,z,\omega):\ \tilde{P}(k_x,dz,\omega)^- \;=\; e^{ik_{z_o}dz}\tilde{P}(k_x,0,\omega)^-,\\
&Step\ 2:\ Inverse\ Fourier\ Transform:\quad \tilde{P}(k_x,dz,\omega)^- \;\rightarrow\; P(x,dz,\omega)^-,\\
&Step\ 3:\ Thin-Lens\ Correction\ in\ (x,z,\omega):\ P(x,dz,\omega)^- \;=\; P(x,dz,\omega)^-e^{i\omega(1/v-1/v_o)dz}.\\
&Step\ 4:\ FFD\ Phase\ Correction:\ P(x,dz,\omega)^- \;=\; P(x,dz,\omega)^-e^{i(other\ term)dz}.
\end{aligned} \qquad (6.19)$$

Encoding Strategy for Wave Equation Migration

To drastically reduce the cost of prestack migration, a possible solution is to randomly compress the source and recording data in the frequency-space domain. The one pseudo-source term is generated by a linear combination of all single source terms. In a similar way the super common shot gather is also obtained by superimposing all shot gathers. The resulting seismic data are considered as the new input shot gather for migration.

$$\mathbf{S_s}(\mathbf{x},\omega) \;=\; \sum_n \alpha_n.S_n(\mathbf{x},\omega), \qquad (6.20)$$

$$\mathbf{D_s}(\mathbf{x},\omega) \;=\; \sum_n \alpha_n.D_n(\mathbf{x},\omega), \qquad (6.21)$$

where $\mathbf{S_n}$ and $\mathbf{D_n}$ denotes the *nth* source term and shot gather respectively; $\mathbf{S_s}$ and $\mathbf{D_n}$ are the encoded source term and shot gather; the coefficients α_n denotes a complex number

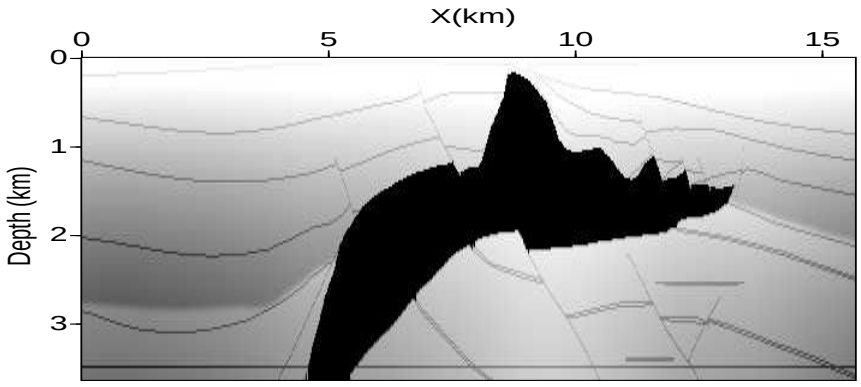


Figure 6.1: SEG/EAGE 2D velocity model.

which can be obtained in several ways (Romero, 2000; Bonomi, 1999). The following three encoding strategies were used to generate α_n. First randomly set one of the ± 1 values to every a_n with probability of 50 percent. The second way is linear encoding, which generates $a_n = e^{(i\theta)}$, where θ is uniformly distributed over the range $[0, 2\pi]$. In the third algorithm, a_n is obtained by randomly sampling from a Gaussian distribution with mean zero and variance of 1.

NUMERICAL RESULTS

All the migration methods were tested on data for the SEG/EAGE salt model. The data set consists of 325 shots, each shot contains 176 records with a recoding length of 5 s and a sampling interval of 8 ms. Shot and receiver intervals are 160 ft and 80 ft, respectively. The velocity model contains 645×150 grids with a gridpoint spacing of 80 ft. The velocity model is shown in Figure 6.1.

First we tested the SSF, PSPI and SSF+FD and KM method with two types of data: one free from multiples and another that includes multiples as seen in the Figure 6.2 shot gathers.

The migration results are shown in Figure 6.3 which correspond to the migrations images obtained from the salt model data without multiples. All results were achieved using one PII $450MHz$ processor with 512 Mbytes of memory. The comparison of computational times for these methods is shown in Figure 6.4. As we noted previously, the cost of KM is the lowest. The wave equation migration cost is several times that of the KM. However, the SSF+FD migration provides superior image quality, especially in the subsalt part where geologists are most interested. SSF and PSPI give comparable results to that of the SSF+FD method. In the PSPI method, its cost is dependent on the number of references velocities used in migration. The computational time of the PSPI method is about twice that of SSF, whereas SSF+FD is slightly more costly than SSF.

In order to drastically reduce the computational time of the wave equation method,

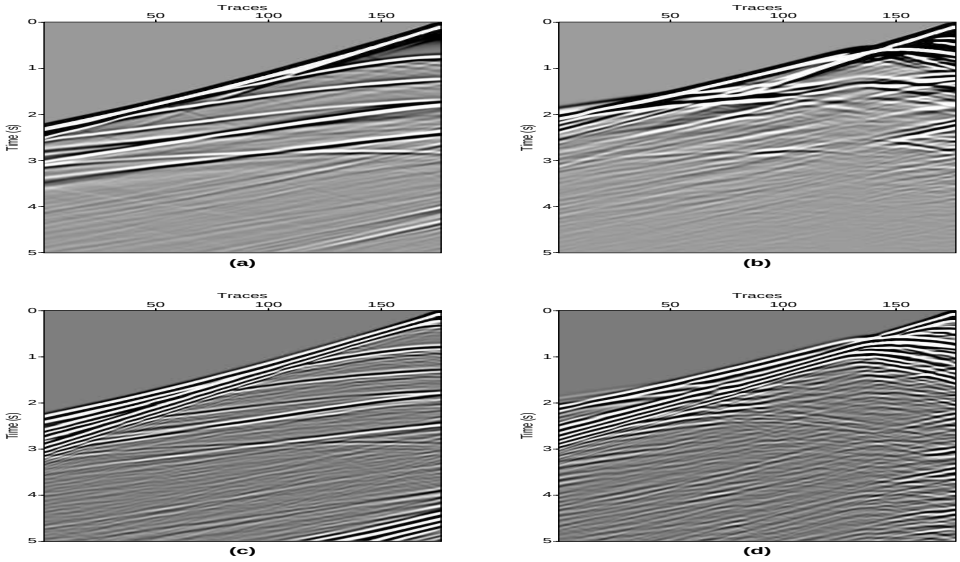


Figure 6.2: Common shot gathers No. 100 and 200: (a). shot 100 without multiples; (b). shot 200 without multiples; (c). shot 100 with multiples; (d). shot 200 with multiples.

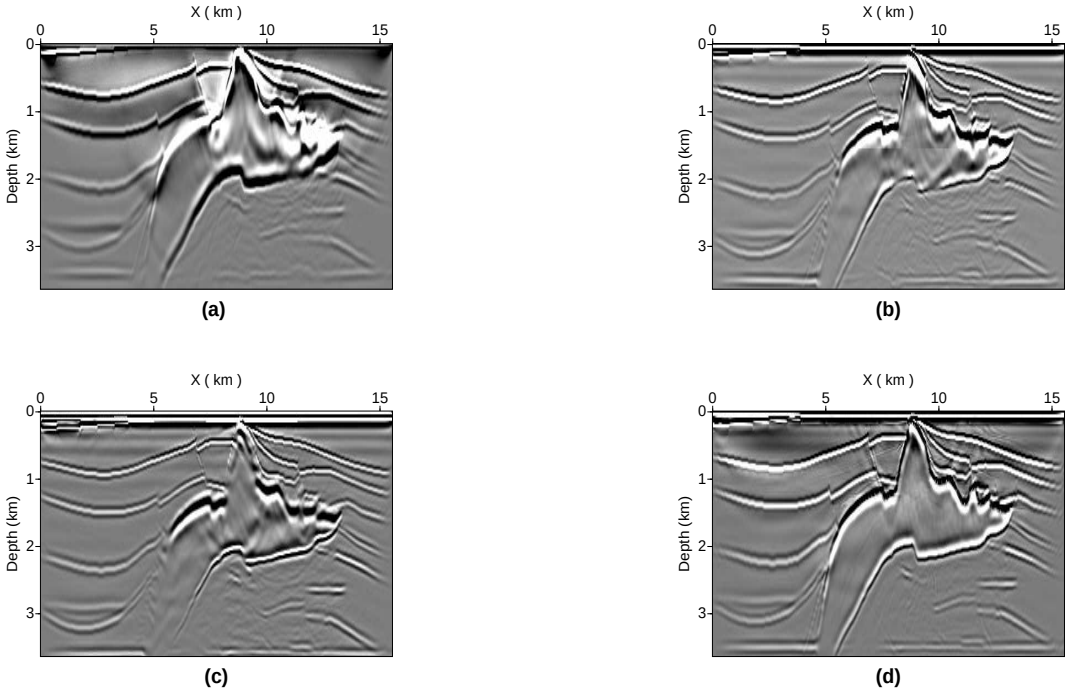


Figure 6.3: Comparison of four migration methods using data from the SEG/EAGE salt model without multiples. Images are from: (a). Kirchhoff migration, (b). SSF migration, (c). PSPI migration, and (d). SSF+FD migration.

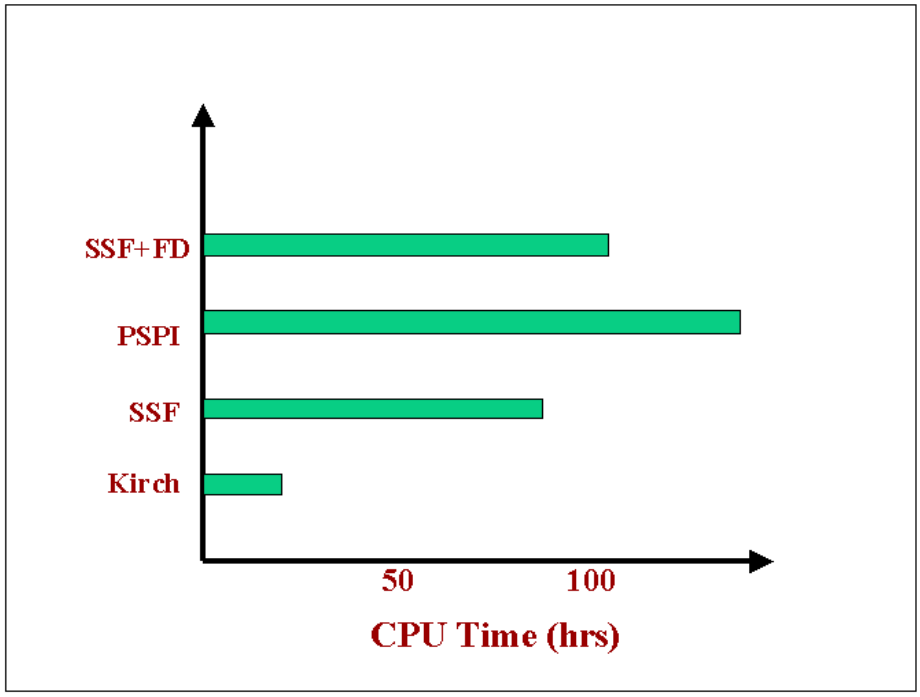


Figure 6.4: Computational time of migration algorithms tested on SEG/EAGE salt model. The Kirchhoff method is faster at the cost of worst equality. The SSF+FD method gives a most accurate migration image, and the CPU time of the PSPI method is dependent on the number of reference velocities used.

an encoding strategy was applied (Morton et al., 1998). Figure 6.5 shows the results of SSF+FD migration with three encoding algorithms. All three encoding algorithms yield almost the same results. In the following tests, the second phase encoding algorithm was adopted. Note, the input data for these migration comparisons includes data with multiple reflections.

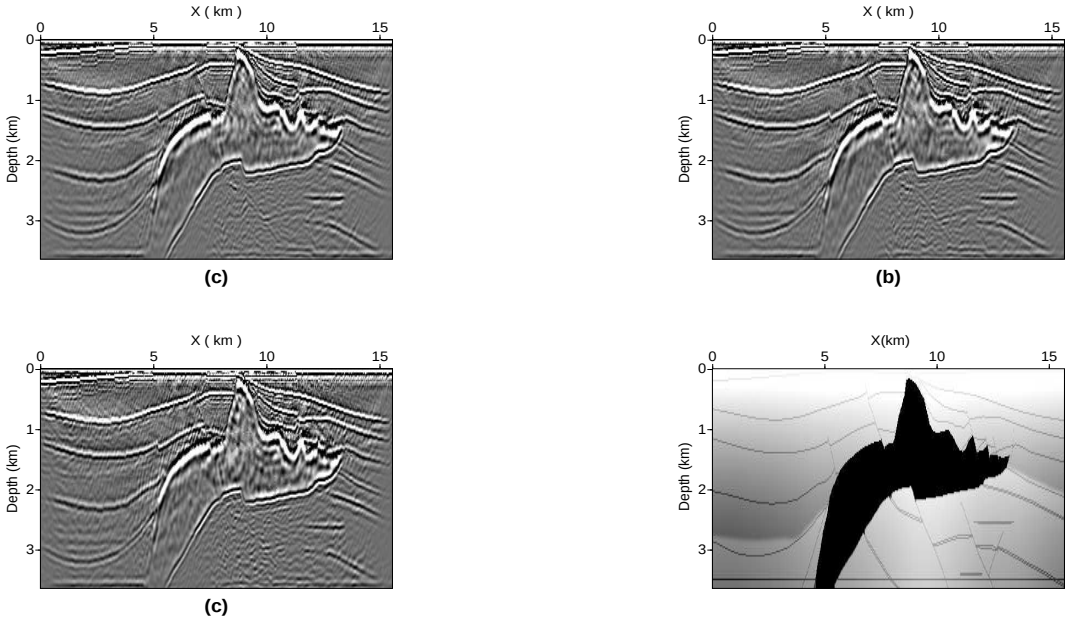


Figure 6.5: SSF+FD migration images with different encoding strategies: (a). uniquely random distribution; (b). linear encoding; (c). Gaussian distribution; (d) SEG/EAGE Salt model. Here the migration results were obtained by migrating a supergather stack a total of 60 times.

In the following tests of wave equation migration, two strategies for implementing phase encoding migration were used: in the first, we encode several adjacent shot gathers into one pseudo shot gather. After that, the total number of shot gathers is reduced and the migration method was applied to these new pseudo shot gathers. Another strategy is to encode a total of 320 shot gathers into one supergather, then apply the migration method to this unique supergather. After one migration, repeat the encoding and the migration steps with different random phase encoding until a satisfactory migration image is obtained.

Figures 6.6 and 6.7 show the SSF+FD migration results using the two phase encoding strategies. The CPU times of the SSF+FD migration method using two phase encoding strategies are shown in Figure 6.9. From this comparison, phase encoding of the SSF+FD migration still generates good migration result, especially in the part of subsalt region, with a running time reduced by a factor two or three times. The SSF migration tests also give the same conclusion as shown in Figures 6.10, 6.11 and 6.12 except that the subsalt image is not as clear as that obtained by SSF+FD migration. Parallel encoding of the wave equation migration codes were developed with MPI Fortran and accomplished on ICEBOX cluster. Figure 6.13 shows the CPU time for the SSF+FD encoding with 10 processors. From these

tests, encoded wave-equation migration can reduce the computational time by a factor of around $5-7$ and achieve a better image than Kirchhoff migration.

The 3D migration algorithm was implemented and tested on the 3D SEG/EAGE salt model. We hope to soon develop an efficient 3D wave equation based method for providing an accurate image of the subsurface geology.

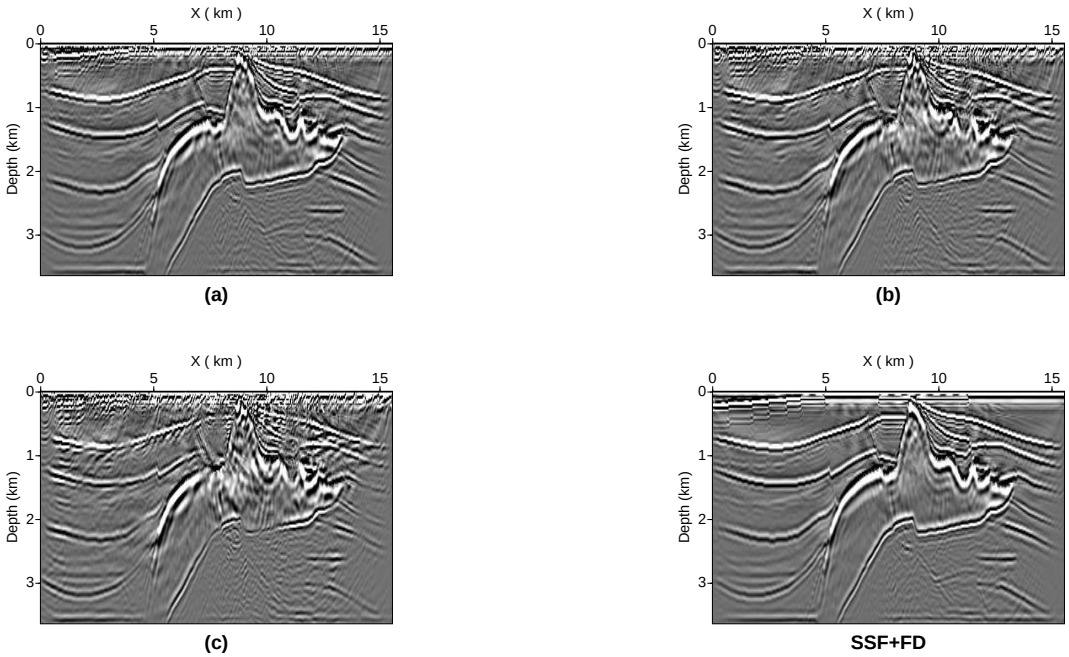


Figure 6.6: Migration comparison of SSF+FD method using different numbers of encoded shot gathers: (a). 2; (b); 4; (c). 10; (d). no encoding.

6.5 Conclusions

Phase-shift migration is a wave equation imaging method that overcomes the single-arrival limitation of diffraction-stack migration. Basic phase-shift migration is strictly valid for $v(z)$ geology, but extensions for lateral velocity variations are available in the form of SSF, PSPI, and SSF+FD migration. An advantage of phase-shift migration is that the power spectrum of the seismic source is band-limited with the cutoff frequency far below the temporal Nyquist frequency. This means we don't need all the frequencies to the Nyquist frequencies. Therefore, mapping the seismic data into the space-frequency domain allows for a significant compression of data and decreases the computational effort. In fact, full volume migration by the phase-shift method is considered by some to be faster and more accurate (except for really steep dips) than full-volume Kirchhoff migration.

Several encoded phase-shift migration methods were tested on the SEG/EAGE salt model data. For the prestack wave-equation migration tests, SSF is the most computationally efficient compared with other wave-equation based methods for our codes. The

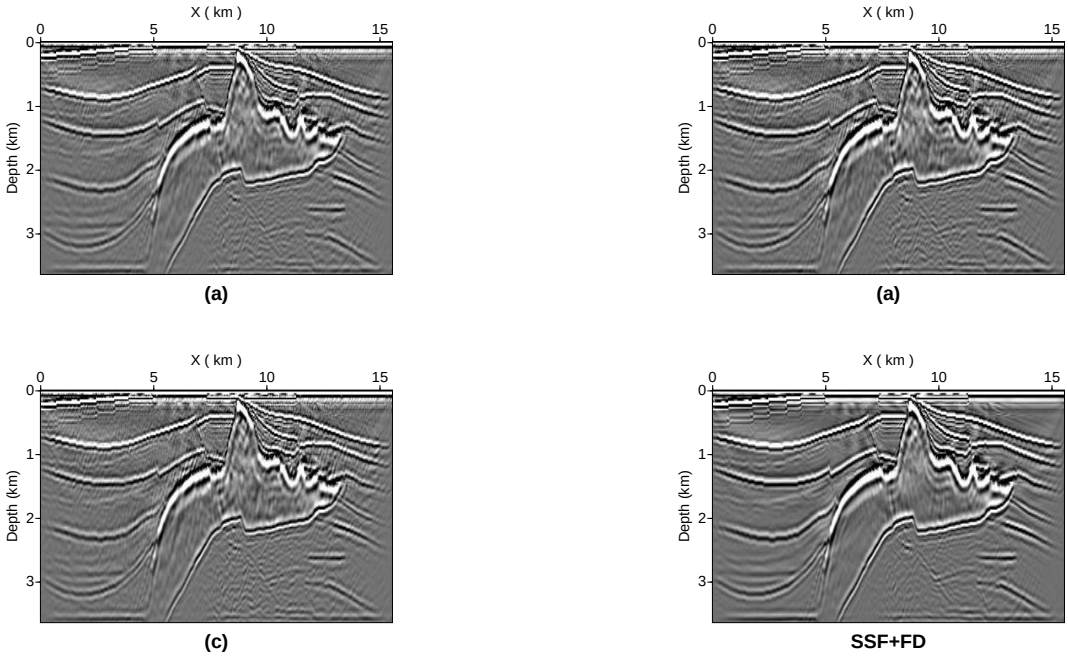


Figure 6.7: Migration comparison of SSF+FD method using different number for migrating an encoded super shot gather: (a). 100; (b); 80; (c). 60 and (d). no encoding.

PSPI method yields a migration image of comparable quality but at twice the cost of SSF migration. The SSF+FD migration generates the most accurate image when the velocity model has strong lateral variations but requires more CPU time than SSF migration and less CPU time than PSPI migration.

We also implemented these algorithms in parallel and with phase-encoding. Results demonstrate that phase-encoding will reduce the cost of wave-equation migration by a factor of 2 or 3. In some cases, the encoded wave-equation migration is competitive with the cost of UTAM's Kirchhoff migration code. We also note that the encoding technique has more influence on the shallow part of the migration image. The extension of SSF and PSPI migration algorithms to 3D is easily implemented by incorporating an additional Fourier transform in the cross-line direction.

6.6 Exercises

1. Show that $P = Ae^{-ik_z z}$ satisfies equation 4.

2. What is the ODE that $P = Ae^{ik_z z}$ satisfies?

3. Note the definition $k_z = \sqrt{k^2 - k_x^2}$. If k_x is less than k, then the k_z is real and $P = Ae^{-ik_z z}$ is a propagating plane wave. Which direction does it propagate for a positive frequency with the Fourier kernel $e^{i\omega t}$; make sure you define which direction is increasing z. If k_x is greater than k, then k_z is purely imaginary and $P = e^{isgn|kz|k_z z}$

is an evanescent wave. What should the exponent sign be (i.e., positive or negative) so that this evanescent wave decreases with increasing depth z? In practice such waves are muted out in the migration.

4. List the disadvantages and advantages of Phase Shift migration relative to Kirchhoff migration and RTM.

5. Write down the MATLAB code for prestack phase shift migration.

6. Write the Split Step Phase Shift migration code.

7. Use MATLAB to plot $\sqrt{1-x^2}$ vs x and Pade approximations for various values of a and b in equation 6.22. Discuss which pair of a and b values leads to the best approximation.

8. Derive equation 6.26 and repeat previous exercise except use equation 6.26 rather than 6.22.

6.7 Appendix: Pade Approximation of Extrapolation Operator

The term $\kappa_z = \omega/v\sqrt{1-(v/\omega)^2\partial^2/\partial x^2}$ can be recast as $\kappa_z = \omega/v\sqrt{1-X^2}$, where $X^2 = (v(x,z)/\omega)^2\partial^2/\partial x^2$. Using Pade polynomials, a ratio of two polynomials, we can approximate the square root by

$$\sqrt{1-X^2} \approx 1 - \frac{aX^2}{1-bX^2}, \tag{6.22}$$

where a and b are constants that are determined by some optimization procedure such as least squares fitting. Note that the numerator and denominator are polynomials of degree two, and for $a = 1/2;\ b = 4$ the vertical wavenumber is that for the 45 degree continued fraction approximation to the wave equation (Claerbout, 1985). For $a = .4$ and $b = .5$ the approximation plots out as Figure 6.8.

In the general case, a and b are constants that are determined to give a best fit to the wave equation in the least squares sense. The continued fraction expression for $a = b = 1$ yields $\sqrt{1-X^2} \approx 1 - \frac{X^2}{1-X^2}$ can be derived by setting $R_0 = 1$ in

$$R_{n+1} = 1 - \frac{X^2}{1+R_n}. \tag{6.23}$$

and iterating to get R_2.

From equation 6.3 we get for $\Delta\kappa_z$:

$$\Delta\kappa_z = \omega/v\sqrt{1-X^2} - \omega/v_o\sqrt{1-X_o^2}, \tag{6.24}$$

where $X_o^2 = (v_o/\omega)^2\partial^2/\partial x^2$. Inserting equation 6.22 into equation 6.24 we get

$$\Delta\kappa_z \approx [\omega/v - \omega/v_o] - \omega/v\frac{aX^2}{1-bX^2} + \omega/v_o\frac{aX_o^2}{1-bX_o^2}, \tag{6.25}$$

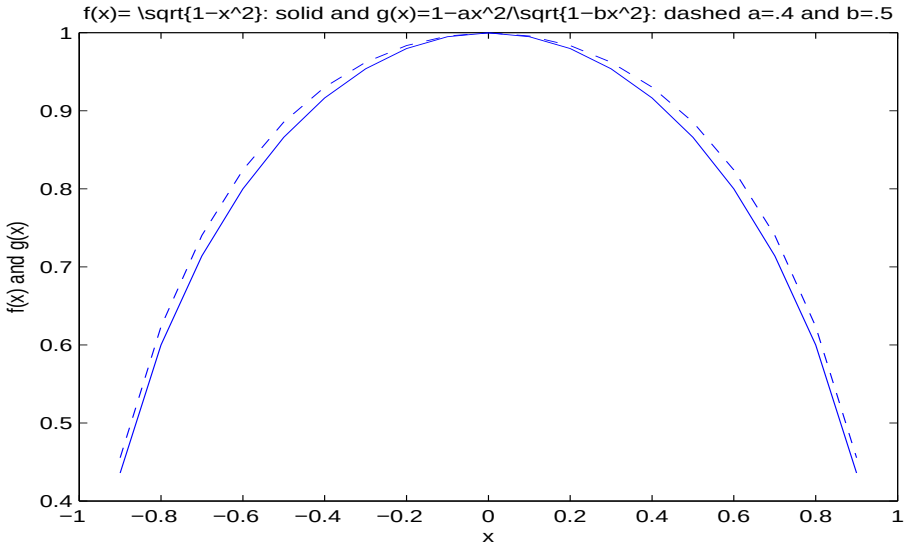


Figure 6.8: Plot of $\sqrt{1-x^2}$ vs x (solid lines) and $1-ax^2/(1-bx^2)$ vs x (dashed lines).

Setting $m = v/v_o$, we get $X_o^2 = -X^2/m^2$ and equation 6.25 becomes:

$$\begin{aligned}
\Delta\kappa_z &= [\omega/v - \omega/v_o] - \omega/v\frac{am^2X^2}{1-bm^2X^2} + \omega/v_o\frac{aX_o^2}{1-bX_o^2}, \\
&= [\omega/v - \omega/v_o] - \omega/v_o[\frac{amX_o^2}{1-bm^2X_o^2} - \frac{aX_o^2}{1-bX_o^2}], \\
&= [\omega/v - \omega/v_o] - \omega/v_o\frac{amX_o^2 - abmX_0^4 - aX_0^2 + abm^2X_o^4}{1-bm^2X_o^2 - bX_o^2 + b^2m^2X_o^4}, \\
&\approx [\omega/v - \omega/v_o] - \omega/v_o\frac{a(m-1)X_o^2}{1-b(m^2+1)X_o^2}. \qquad (6.26)
\end{aligned}$$

where the last term assumes 4th-order values of X_0 are dropped. This is true if the field propagates close to vertical so that the k_x components are small.

6.7.1 Finite Difference Approximation

The last term in equation 6.26 represents the *other term* in equations 6.13 and 6.18. Thus, the PDE associated with the *other term* solution is

$$\frac{\partial P(x,z,\omega)}{\partial z} = i\frac{\alpha\frac{\partial^2}{\partial x^2}}{1+\beta\frac{\partial^2}{\partial x^2}}P(x,z,\omega) \qquad (6.27)$$

where $\alpha = a(m-1)(v_o/\omega)^2$ and $\beta = b(m^2+1)(v_o/\omega)^2$. Rearranging the above equation yields

$$(1+\beta\frac{\partial^2}{\partial x^2})\frac{\partial P(x,z,\omega)}{\partial z} = i\alpha\frac{\partial^2}{\partial x^2}P(x,z,\omega). \qquad (6.28)$$

Replacing the 1st-order differential by the 1st-order correct finite difference approximation yields

$$\frac{\partial P(x,z,\omega)}{\partial z} \quad = \quad (P^x_{z+1} - P^x_z)/\Delta z, \tag{6.29}$$

and the second order differential by the *averaged* second-order correct finite-difference approximation yields

$$\frac{\partial^2 P(x,z,\omega)}{\partial x^2} \quad = \quad 0.5[(P^{x+1}_{z+1} - 2P^x_{z+1} + P^{x-1}_{z+1}) + (P^{x+1}_z - 2P^x_z + P^{x-1}_z)]/\Delta x^2. \tag{6.30}$$

Substituting equations 6.28-6.30 into equation 6.27 gives

$$(1 + \beta\frac{\partial^2}{\partial x^2})(P^x_{z+1} - P^x_z)/\Delta z \quad =$$

$$\frac{i\alpha}{2\Delta x^2}[(P^{x+1}_{z+1} - 2P^x_{z+1} + P^{x-1}_{z+1}) \quad + \quad (P^{x+1}_z - 2P^x_z + P^{x-1}_z)]. \tag{6.31}$$

Replacing the second-order derivative by a second-order correct finite-difference approximation and isolating the $z+1$ components to be on the left hand side yields

$$P^x_{z+1} + (\frac{\beta}{\Delta x^2} - \frac{i\alpha\Delta z}{2\Delta x^2})(P^{x+1}_{z+1} - 2P^x_{z+1} + P^{x-1}_{z+1}) \quad =$$

$$P^x_z \quad + \quad (\frac{\beta}{\Delta x^2} + \frac{i\alpha\Delta z}{2\Delta x^2})(P^{x+1}_z - 2P^x_z + P^{x-1}_z). \tag{6.32}$$

Letting $\alpha' = \frac{\beta}{\Delta x^2} - \frac{i\alpha\Delta z}{2\Delta x^2}$ and $\beta' = \frac{\beta}{\Delta x^2} + \frac{i\alpha\Delta z}{2\Delta x^2}$ we get the final equation

$$\alpha' P^{x+1}_{z+1} + (1 - 2\alpha')P^x_{z+1} + \alpha' P^{x-1}_{z+1} \quad = \quad \beta' P^{x+1}_z + (1 + 2\beta')P^x_z + \beta' P^{x-1}_z. \tag{6.33}$$

This is an implicit finite-difference equation with unconditional stability and can be efficiently solved by a tridiagonal matrix inversion scheme (Claerbout, 1985).

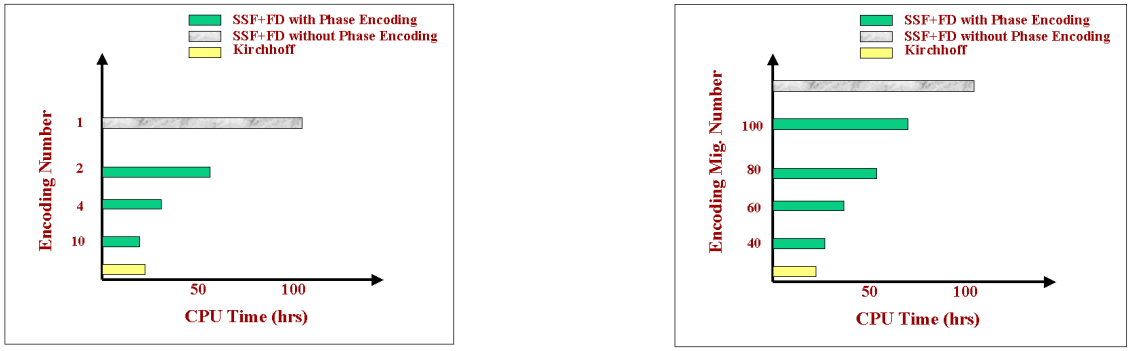


Figure 6.9: (Top). CPU time comparison for encoded SSF+FD migration with different number of shot gathers; *Encoding Number* denotes the number of adjacent shot gathers to be phase encoded. (Bottom) CPU time comparison for encoding SSF+FD migration with the second strategy. Note, the CPU time of no-encoding SSF+FD migration and Kirchhoff migration are also presented for comparison.

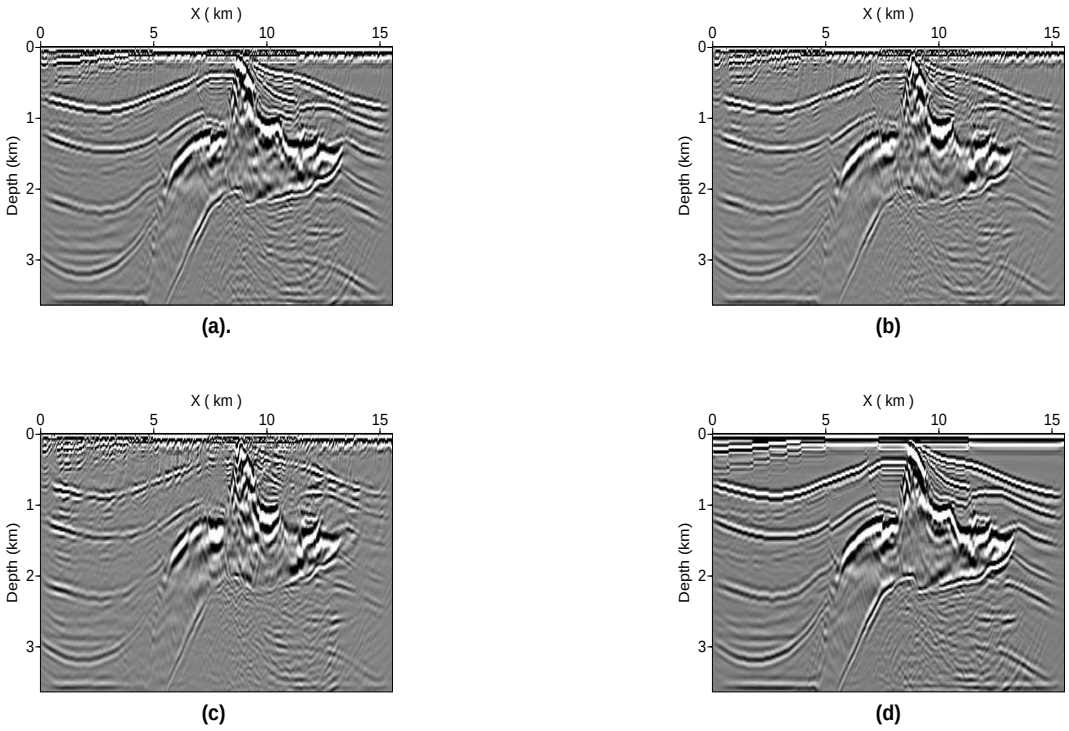


Figure 6.10: Migration comparison of SSF method with different encoded shot numbers: (a). 2; (b); 4; (c). 10 and (d). standard migration result.

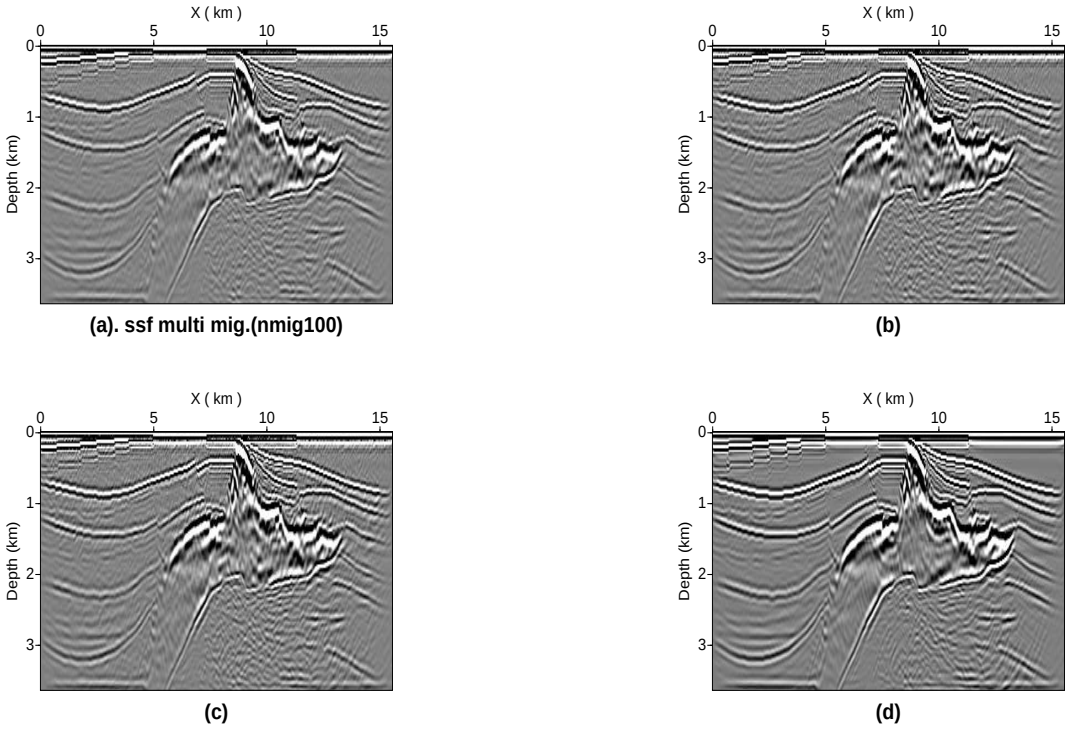


Figure 6.11: Migration comparison of encoded SSF method using different number of migration of a super encoded shot gather: (a). 80; (b); 60; (c). 40 and (d). standard migration result.

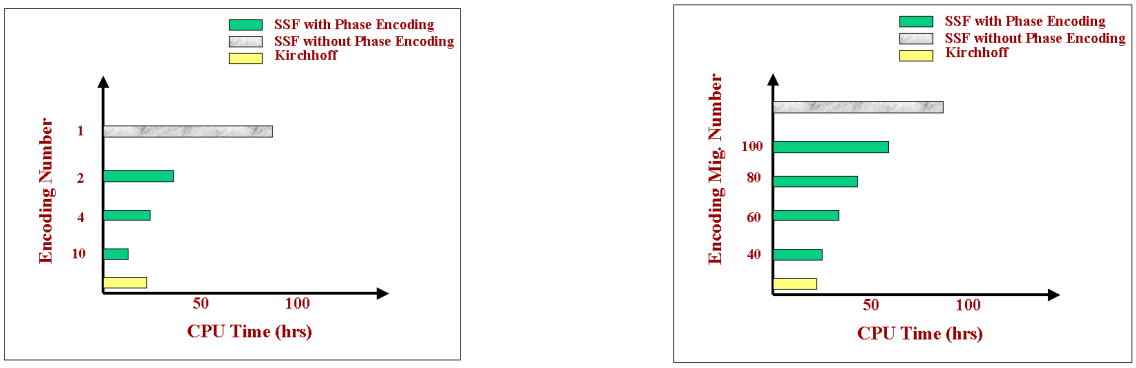


Figure 6.12: (Top). CPU time comparison for encoded SSF migration with different number of shot gathers; *Encoding Number* denotes the number of adjacent shot gathers that were phase encoded. (Bottom) CPU time comparison for encoding SSF migration with the second strategy. Note, the CPU time of no-encoding SSF+FD migration and Kirchhoff migration are also presented for comparison.

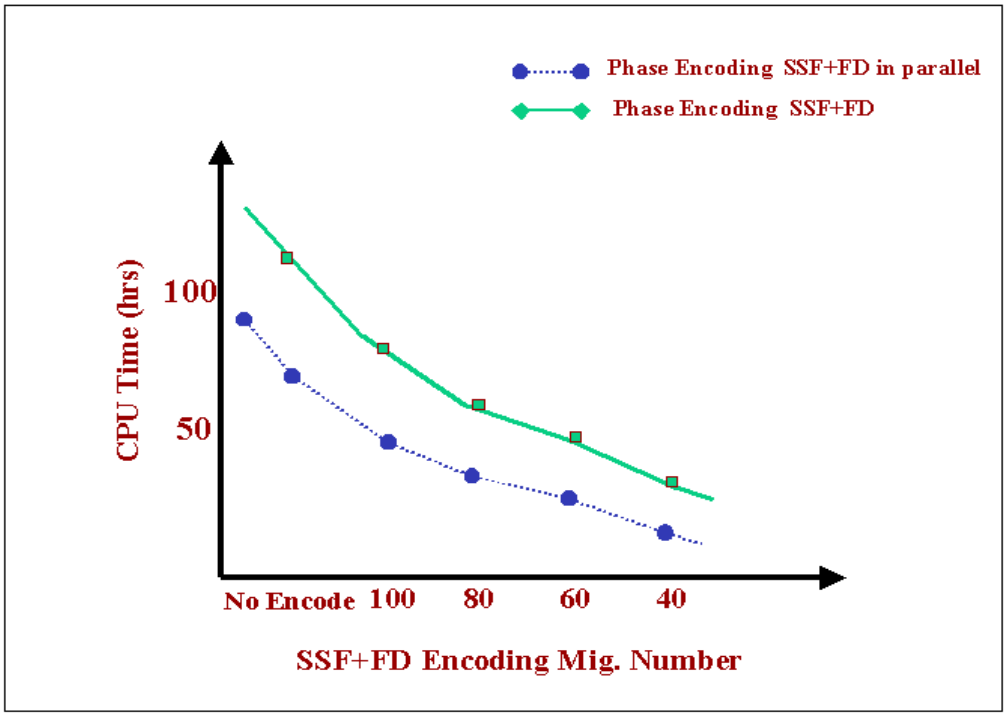


Figure 6.13: CPU time for SSF+FD migration method applied to the SEG/EAGE salt dome data. The solid line denotes the CPU time SSF+FD migration using one node. The dashed line is the encoded SSF+FD migration method executed on 10 nodes.

Bibliography

[1] Aarts, E., and Korst, J., 1991, Simulated annealing and Boltzmann machines: J. Wiley and Sons, NY, NY.

[2] Aki, K., Christoffersson, A., and Husebye, 1978, Determination of the three-dimensional seismic structures of the lithosphere: J. Geophys. Res., **82**, 277-296.

[3] Aki, K., and Richards, P., 1980, Quantitative Seismology: W.H. Freeman and Co., NY, NY.

[4] Aminzadeh, F., Brac, J., and Kunz, T., 1997, 3-D Salt and Overt hrust Models, SEG/EAGE 3-D M odeling Series No.1, Society of Exploration Geophysicists. Tulsa, OK.

[5] Alford, R., Kelly, K., and Boore, D., 1974, Accuracy of finite-difference modeling of the acoustic wave equation: Geophysics, **39**, 834-842.

[6] Anton, H., 1977, Elementary Linear Algebra: J. Wiley Co.

[7] Bagani, C., Bonomi, E., and Pieroni, E., 1995, Data parallel implementation of 3D PSPI: 65th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 188-191.

[8] Baker, G., Steeples, D., and Drake, D., 1998, Muting the noise cone in near-surface reflection data: An example from southeastern Kansas, *Geophysics*, **63**, 1332-1338.

[9] Barton, G., 1989, Elements of Green's functions and propagation: Oxford Science Publ., pp. 465.

[10] Baysal, E., D. Kosloff, and J. Sherwood, 1983, Reverse time migration: Geophysics, **48**, 1514–1524.

[11] Bednar, B., 2005, A brief history of migration: Geophysics, **70**, 3MJ–20MJ.

[12] Bednar, J. B., C. J. Bednar, and C. S. Shin, 2006, Two-way vs. one-way: A case study style comparison: 76th Annual International Meeting, SEG Expanded Abstracts, 2343–2347.

[13] Berryman, J., 1991, Lecture notes on nonlinear inversion and tomography; Law. Liv. Nat. Lab Publ. UCRL-LR-105358-rev.1

[14] Beylkin, G., 1994, Inversion and applications of the GRT: in "75 Years of Radon Transform", Conference and Proceedings and Lecture Notes in Mathematical Physics: V. IV, edited by S. Gindikin and P. Michor, Int. Press, Boston.

[15] Biondi, B., and G. Shan, 2002, Prestack imaging of overturned reflections by reverse time migration: 72nd Annual International Meeting, SEG, Expanded Abstracts, 1284–1287.

[16] Bishop, T., Bube, K., Cutler, R., Langan, R., Love, P., Resnick, J., Shuey, R., Spindler, D., and Wyld, H., 1985, Tomographic determination of velocity and depth in laterally varying media: Geophysics, **50**, p. 903-923.

[17] Bleistein, N., 1984, Mathematical methods for wave phenomena: Academic Press, NY, NY.

[18] Brebbia, C.A., 1978, The boundary element method for engineers: J. Wiley and Co., NY, NY, 189 p.

[19] Bregman, N. D., Bailey, R. C., and Chapman, C. H., 1989, Ghosts in tomography: The effect of poor angular coverage in 2-D seismic traveltime inversion: Canadian J. Expl. Geophysics, 25, 7-27.

[20] Bonomi E. and Cazzola L., 1999, Prestack imaging of compressed seismic data: a Monte Carlo approach, 69th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts.

[21] Bube, K., and Langan, R., 1994, A Hybrid l^1/l^2 minimization with applications to tomography: Geophysics, 62, 1183-1195.

[22] Bube, K., and Meadows, M., 1998, Characterization of the null space of a generally anisotropic medium in linearized cross-well tomography: J. Geophys. Res., 100, 12,4949-12,458.

[23] Bube, K., Jovanovich, D., Langan, R., Resnick, J., Shuey, R., and Spindler, D., 1985, Well-determined and poorly determined features in seismic reflection tomography: Part II; Expanded Abstracts of 55th Annual SEG Meeting, 608-610.

[24] Buck, B., and Macauly, V., 1994, Maximum Entropy in Action: Clarendon Press, Oxford, pp. 220.

[25] Cadzow, J., 1987, Foundations of Digital Signal Processing and Data Analysis: Mcmillian Publ. Co., NY, NY.

[26] Calnan, C., 1989, Crosswell tomography with reflection and transmission data; MS thesis, University of Utah.

[27] Calnan, C. and Schuster, G.T., 1989. Reflection + transmission cross-well tomography, Expanded Abstracts of the 1989 Technical Program of the Society of Exploration Geophysicists with Biographies, p. 908-910.

[28] Calvetti, D., Lewis, B., and Reichel, L., 1998, Smooth or abrupt: a comparison of regularization methods: Advanced Signal Processing Algorithms, Architectures and Implementations VIII, ed. F.T. Luk, Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE), vol. 3461, The International Society for Optical Engineering, Bellingham, WA, 286–295.

[29] Calvetti, D., Reichel, L., and Zhang, Q., 1999, Iterative solution methods for large linear discrete ill-posed problems: Applied and Computational Control, Signals and Circuits, **1**, 313–367.

[30] Cary, P., 1998, The simplest discrete radon transform: Expanded Abstracts of 68th Int. SEG meeting, New Orleans, 1999-2002.

[31] Chen, S.T., Zimmerman, L.J., and Tugnait, J.K., 1990, Subsurface imaging using reversed vertical seismic profiling and crosshole tomographic methods: Geophysics, 55, 1478-1487.

[32] Claerbout, J. F., 1985, Imaging the Earth's interior, Blackwell Scientific Publications.

[33] Claerbout, J, 1992, Earth soundings analysis: Processing vs inversion: Blackwell Scientific Publ., SF, CA.

[34] Claerbout, J, 2001, Geophysical estimation by example: Environmental soundings image construction: Multidimensional autoregression, (in progress). (sepwww.stanford.edu/sep/prof/toc_html/gee/toc_html/index.html)

[35] Constable, S., Parker, R., and Constable, C., 1987, Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data: Geophysics, **52**, 289-300.

[36] Davies, S., Packer, K., Baruya, A., and Grant, A., 1994, Enhanced information recovery in spectroscopy using the maximum entropy meth od: in "Maximum Entropy in Action", edited by Buck, B., and Macauly, V., Clarendon Press, Oxford, pp. 220.

[37] deGroot-Hedlin, C., and Constable, S. C., 1990, Occam s inversion to generate smooth, twodimensional models from magnetotelluric data: Geophysics, 55, 1613-1624.

[38] Dyer, B., and Worthington, M. H., 1988, Some sources of distortion in tomographic velocity images: Geophys. Prosp., 36, 209-222.

[39] Duquet B., Marfurt K., and Dellinger J., 2000, Kirchhoff modeling, inversion for reflectivity, and subsurface illumination: Geophysics, **65** (4), pp.1195-1209.

[40] Fletcher, R., l987, Practical Methods of Optimization: J. Wiley and Sons, NY, NY.

[41] Fomel, S., 2000, 3-D data regularization: PhD dissertation, Stanford University. (sepwww.stanford.edu/public/docs/sep107/paper_html/).

[42] Fomel, S., 2003, Asymptotic inverse operators: Geophysics, Geophysics, 1032-1042.

[43] Gazdag, J., 1978, Wave equation migration with phase-shift method. Geophysics,43, 1342-1351.

[44] Gazdag, J., and Sguazzero, P., 1984, Migration of seismic data by phase shift plus interpolation: Geophysics, 49, 124-131.

[45] Gill, P., Murray, W., and Wright, M., 1981, Practical Optimization: Academic Press, NY, NY.

[46] Golub, G., Heath, M., and Wahba, G., 1979, Generalized cross-validation as a method for choosing a good ridge parameter: Technometrics, **21**, 215-223.

[47] Golub, G., and von Matt, U., 1997, Generalized cross validation for large-scale problems: J. Comput. Graph. Stat. **6**, 1-34.

[48] Groetsch, C., 1993, Inverse Problems in Mathematical Science: Vieweg Publ., Wiesbaden, Germany.

[49] Gonzales, R., and Woods, R., 1992, Digital Image Processing, Addison-Wesley Publ. Co., Menlo Park, Calif., pp 716.

[50] Han, B., 1998, A comparison of four depth-migration methods, 68th Ann. Internat. Mtg., Soc. Expl. Geophys..

[51] Hansen, P., 1992, Analysis of discrete ill-posed problems with the L-curve: SIAM Review, **32**, 561-580.

[52] Hansen, P., and O'Leary, 1993, The use of the L-curve: SIAM, J. Sci. Statist. Comput., **14**, 1487-1503.

[53] Hanke, M., 1996, Limitations of the L-curve method in ill-posed problems: BIT 36, 287-301.

[54] Hu, J., Schuster, G.T., and Valasek, P., 2001, Migration deconvolution: Poststack data: Geophysics, **66**, 939-952.

[55] Huang, L., Fehler C. F., Wu, R., 1999, Extended local Born Fourier migration method: Geophysics, 64,1524-1534.

[56] Huang, L., Fehler C. F., Robert, M. P., and Burch, C. C., 1999, Extended local Rytov Fourier migration method: Geophysics, 64,1535-1545.

[57] Ivansson, S., 1985, A study of methods for tomographic velocity estimation in the presence of low-velocity zones: Geophysics, **50**, 969-988 .

[58] Jackson, D. D., 1972, Interpretation of inaccurate, insufficient, and inconsistent data: Geophys. J. Roy. Astr. Soc., **28**, 97-109.

[59] Jasson, P., 1997, Modern constrained nonlinear methods: in "Deconvolution of Images and Spectra", ed. by Petter A. Jasson, Academic Press, 107-181.

[60] Jupp, D. L., and Vozoff, K., 1975, Stable iterative methods for the inversion of geophysical data: Geophys. J. Roy. Astr. Soc., **42**, 67-72.

[61] Kapur, J., and Kesavan, H., 1992, Entropy Optimization Principles with Applications: Academic Press Inc., pp. 408.

[62] Kennett, B.L., Sambridge, M.S., and Williamson, P.R., 1988, Subspace methods for large inversion problems with multiple parameter classes: Geophysical Journal, **94**, 237-247.

[63] Kreyszig, E., 1978, Introductory Functional Analysis with Applications: J. Wiley and Sons, NY, NY.

[64] Lailly, P., 1983, The seismic inverse problem as a sequence of before-stack migrations, in J. Bednar, ed., Conference on inverse scattering: Theory and Applications: SIAM, 206–220.

[65] Lailly, P., 1984, Migration methods: partial but efficient solutions to the seismic inverse problem: in Inverse problems of acoustic and elastic waves, edited by F. Santosa, Y.H. Pao, Symes, W., and Ch. Holland, SIAM, Philadelphia.

[66] Lailly, P., 1983, The seismic inverse problem as a sequence of before-stack migrations, in J. Bednar, ed., Conference on inverse scattering: Theory and Applications: SIAM, 206–220.

[67] Lawson, L., and Hanson, R., 1974, Solving least squares problems: Prentice-Hall, Englewood Cliffs, NJ.

[68] Lee, M., Mason, L. M., and Jackson, G. M., 1991, Split-step Fourier shot-record migration with deconvolution imaging. Geophysics, 56(11), 1786-1793.

[69] Levander, A., 1988, Fourth-order finite-difference P-SV seismograms: Geophysics, **53**, 1425-1436.

[70] Lines, L.R., and Treitel, S., 1984, Tutorial, review of least-squares inversion and its application to geophysical problem: Geophy. Prosp., 32, 159-186.

[71] Liu, F., Zhang, G., Morton, S., and Leville, J., 2007, Reversetime migration using oneway wavefield imaging condition: SEG, Expanded Abstracts, 26 , no. 1, 2170-2174.

[72] Luo, Y., Q. Liu, Y. E. Wang, and M. N. AlFaraj, 2006, Imaging reflection-blind areas using transmitted PS-waves: Geophysics, **71**, S241–S250.

[73] Ma, X., 2001, Constrained global inversion method using an overdetermined scheme: Application to poststack data: Geophysics, **66**, 613-626.

[74] Mackie, R., and Madden, T., 1993, 3-D MT inversion using CG: Geophys. J. Internat., 115, 215-229.

[75] Menke, W., 1984, Geophysical data analysis: Discrete inverse theory: Academic Press, NY, NY.

[76] Mora, P., 1987, Elastic Wavefield Inversion: PhD Dissertation, Stanford University, pp. 143.

[77] Morey, D., and Schuster, G.T., 1999, Paleoseismicity of the Oquirrh fault, Utah from shallow seismic tomography, *Geophys. J. Int.*, **138**, 25-35.

[78] Morse, P., and Feshbach, H., 1953, Methods of theoretical physics: McGraw-Hill book Co., NY, NY.

[79] Morton S. A. and Ober C. C., 1998, Faster shot-record depth migration using phase encoding, 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts.

[80] Miller, D., Oristaglio, M., and Beylkin, G., 1987, A new slant on seismic imaging: Migration and integral geometry: *Geophysics* , **52**, 943-964.

[81] Mulder, W., and R.-E. Plessix, 2004, A comparison between one-way and two-way wave equation migration: Geophysics, 69, 1491–1504.

[82] Natterer, F., 1986, The mathematics of computerized tomography; J. Wiley and Sons, NY, NY.

[83] Nemeth, T., Qin, F., and Normark, E, 1997, Dynamic smoothing in crosswell traveltime tomography: Geophysics, 62, 168-176.

[84] Nemeth, T., 1996, Imaging and filtering by least squares migration: PhD Dissertation, U of Utah, pp. 121.

[85] Nemeth, T., Sun, H., and Schuster, G.T., 2000, Separation of signal and coherent noise by least squares filtering: Geophysics, **65**, 574-583.

[86] Nemeth, T., Wu, C., and Schuster, G.T., 1999, Least squares migration of incomplete data: Geophysics, **64**, 208-221.

[87] Nolet, G., l987, Seismic Tomography: D. Reidel Publishing, Boston, MA.

[88] Nolet, G., l987, Seismic wave propagation and seismic tomography: in "Seismic Tomography" edited by G. Nolet, D. Reidel Publishing, Boston, p. 1.-24.

[89] Nocedal, J. and Wright, S., 1999, Numerical Optimization: Springer Verlag Co., NY, NY.

[90] Pan, G., 1993, Measure for the resolving power and bias effect in the nonlinear inverse analysis of geophysical potential fields; Geophysics, v. 58, 626-636.

[91] Peterson, J. E., Paulsson, B. N. P., and McEvilly, T. V., 1985, Applications of algebraic reconstruction techniques to crosshole seismic data: Geophysics, 50, 1566-1580.

[92] Phillips, W. S., and Fehler, M. C., 1991, Traveltime tomography: A comparison of popular methods: Geophysics, 56, 1639-1649.

[93] Press, W., Teukolsky, S., Vetterling, W., and Flannery, B., 1992, Numerical Recipes in Fortran: Cambridge University Press, NY, NY.

[94] Rector III, J. W., and Washbourne, J. K., 1994, Characterization of resolution and uniqueness in crosswell direct-arrival traveltime tomography using the Fourier projection slice theorem: Geophysics, 59, 1642-1649.

[95] Ristow, D., and Ruhl, T., 1994, Fourier finite-difference migration. Geophysics, 59(12), 1882-1893.

[96] Qin, F., Luo, Y., Olsen, K., Cai, W., and Schuster, G.T., 1992, Finite-difference solution of the eikonal equation along expanding wavefronts: *Geophysics*, 57, p. 478-487.

[97] Oldenburg, D., McGillivray, P., and Ellis, R., 1993, Generalized subspace methods for large scale inverse problems: geophys. J. Internat., 114, 12-20.

[98] Rodi, W., and Mackie, R., 2001, Nonlinear conjugate gradient algorithm for 2-D MT inversion: Geophysics, 66, 174-187.

[99] Salo, E, and Schuster, G.T., 1989, Traveltime inversion of both direct and reflected arrivals in VSP data; Geophysics, v. 54, 49-56.

[100] Saito, H., 1990, Ray tracing based on reciprocity: 60th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1028-1031.

[101] Sava, P., and Fomel, S., 2003, Angle-domain common-image gathers by wavefield continuation methods: Geophysics, **68**, 1065-1074.

[102] Scales, L., 1985, Introduction to Non-linear Optimization: Springer-Verlag New York Inc.

[103] Scales, J., 1987, Tomographic inversion via the conjugate gradient method: *Geophysics*, 52, p. 179-185.

[104] Schonewille, M. and Duijndam, A., 2001, Parabolic Radon transform, sampling and efficiency: Geophysics, **66**, 314-333.

[105] Schuster, G. T., 1985. A hybrid BIE and Born series modeling scheme: Generalized Born series. J. Acoust. Soc. Am., v. 77 (3), p. 865-879.

[106] Schuster, G.T., 1989a, Analytic generalized inverse for transmission+reflection tomography; Geophysics, v. 54, 1046-1049.

[107] Schuster, G.T., 1989b, Derivation of analytic generalized inverses for transmission+reflection tomography; Inverse Problems, v. 5, 1117-1129.

[108] Schuster, G.T., 1996, Resolution limits for crosswell migration and traveltime tomography: Geophysical J. Int., 427-440.

[109] Schuster, G.T., 1993, Decomposition of the waveform gradient into short and long wavelength components: 1993 annual UTAM report, 79-88.

[110] Schuster, G.T., Johnson, D., and Trentman, D., 1988, Numerical verification and extension of an analytic inverse for CDP and VSP traveltime equations; Geophysics, Vol. 53, 326-333.

[111] Sen, M., and Stoffa, P., 1991, Nonlinear one-dimensional seismic waveform inversion using simulated annealing: Geophysics, 56, 1624-1638.

[112] Sevink, A., and Herman, G., 1994, Fast iterative solution of sparsely sampled seismic inverse problems: Inverse Problem, **10**, 937-948.

[113] Sheley, D., Crosby, T., Zhou, C., Giacoma, J., Yu, J., He, R., and Schuster, G., 2003, 2-D seismic trenching of colluvial wedges and faults: in Contributions to Neotectonics and Seismic Hazard from Shallow Geophysical Imaging, Tectonophysics, (in press).

[114] Sheng, J., and Schuster, G.T., 2003, Finite-frequency resolution limits of wave path traveltime tomography for smoothly varying models: Geophys. J. Int., **152**, 669-676.

[115] Stoffa, P. L., Fokkemam J. T, Freir, R. M., and Kessinger, W. P., 1990, Split-step Fourier migration: Geophysics, 55, 410-421.

[116] Stoffa, P., and Sen, 1991, Nonlinear multiparameter optimization using genetic algorithms: Inversion of plane-wave seismograms: Geophysics, **56**, 1974-1810.

[117] Stolt, R., and Benson, A., 1986, Seismic Migration: Theory and Practice: in Handbook of Geophysical Exploration, Volume 5, Geophysical Press, London, UK.

[118] Stork, C., and Clayton, R., 1986, Analysis of the resolution between ambiguous velocity and reflector position for traveltime tomography: 56th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 545-550.

[119] Stork, C., 1991, High resolution SVD analysis of the coupled velocity determination between ambiguous velocity and reflector position for traveltime tomography: 61st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 981-985.

[120] Squires, L., and Cambois, G., 1992, A linear filter approach to designing off-diagonal damping matrices for least-squares inverse problems: Geophysics, No. 7, 948-951.

[121] Sun, Y., 1992, Tomographic smoothing, cooperative inversion, and constraints by a Lagrange Multiplier method: 1992 UTAM annual report, 78-90.

[122] Tarantola, A., 1986, Linearized inversion of seismic reflection data: Geophys. Prosp., 32, 998-1015.

[123] Tarantola, A., 1987, Inverse problem theory methods for data fitting and model parameter estimation: Elsevier Science Publ. Co..

[124] Thorson, J., and Claerbout, 1985, Velocity-stack and slant-stack stochastic inversion: Geophysics, **50**, 2727-2741.

[125] Trad, D., Ulrych, T., and Sacchi, M., 2002, A hybrid linear-hyperbolic Radon transform: J. Seis. Expl., **9**, 303-318.

[126] Trad, D., Ulrych, T., and Sacchi, M., 2003, Latest views of the sparse Radon transform: Geophysics, **68**, 386-399.

[127] Vidale, J., 1988, Finite-difference calculation of traveltimes: *Bull. Seis. Soc. Am.*, 78, 2062-2076.

[128] Whiting, P., 1998, Reflection tomography without picking: 68th SEG Ann. Mtg., Expanded Abstracts, 1 226-1230.

[129] Whitmore, N. D., 1983, Iterative depth migration by backwards time propagation: 53rd Annual International Meeting, SEG, Expanded Abstracts, 382–385.

[130] Williamson, P. R., 1991, A guide to the limits of resolution imposed by scattering in ray tomography: Geophysics, 56, 202-207.

[131] Williamson, P. R., and Worthington, M. H., 1993, Resolution limits in ray tomography due to wave behavior: Numerical experiments: Geophysics, 58, 727-735.

[132] Woodward, M., 1992, Wave equation tomography: Geophysics, **57**, 15-26.

[133] Wu, R. and Jin, S., 1997, Windowed GSP (Generalized Screen Propagators) migration applied to SEG-EAGE salt model data: 67th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1746-1749.

[134] Yi, M., 2003, Enhancing the resolving power of least-squares inversion with Active Constraint Balancing: Geophysics (in press).

[135] Yoon, K., K. Marfurt, and E. W. Starr, 2004, Challenges in reverse time migration: 74th Annual International Meeting, SEG, Expanded Abstracts, 1057–1060.

[136] Yoon, K., C. Shin, S. Suh, L. Lines, and S. Hong, 2003, 3D reverse-time migration using the acoustic wave equation: An experience with the SEG/EAGE data set: The Leading Edge, 22, 38–41.

[137] Zemanian, A., 1965, Distribution theory and transform analysis: Dover Publications, NY, NY.

[138] Zhang, J. and Toksoz, N., 1998, Nonlinear refraction traveltime tomography:

[139] Zhou, C., Cai, W., Luo, Y., Schuster, G.T., and Hassanzadeh, S., 1995, Acoustic wave equation traveltime and waveform inversion of crosswell data: Geophysics, 60, 765-773.

[140] Zhou, C., Schuster, G.T., and Hassanzadeh, S., and Harris, J., 1997, Elastic wave equation traveltime and waveform inversion of crosswell data: Geophysics, 853-868.