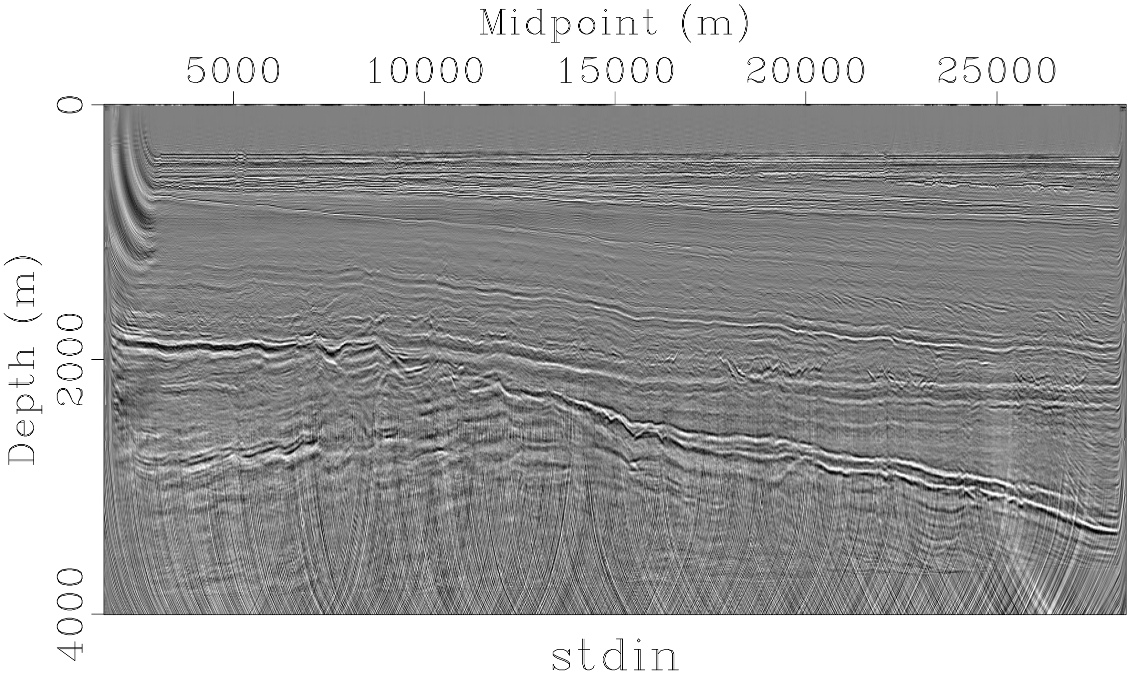
Midpoint (m)
5000
10000
15000
20000
25000
Depth (m)
0
2000
4000
stdin

Seismic Inversion
Part 1: Classic Imaging

William W. Symes

Contents

Chapter 1

Introduction

1.1 Goals

This book will explore some of the core concepts and methods of active-source seismic imaging and inversion from a mathematical point of view. The content arose from my efforts (and those of many others) to explain *precisely* the origin in the physics of waves of many clever and effective processing algorithms, widely used in industrial and scientific practice. Evidently, either such explanation is possible, or our understanding of seismic waves is lacking in important respects. To some extent this project has been successful: wave theory - linear elastodynamics - actually does explain *precisely* why some important seismic imaging and inversion methods work. Even better, the theoretical concepts required to understand modern imaging methods are themselves (relatively) modern - the necessary mathematics has developed in parallel with (and largely independent of) engineering/scientific methodology. Rich and beautiful connections have been drawn in both directions: imaging algorithms based on data-driven insight have inspired new mathematics, which has returned the favor by suggesting improved (or, in some cases, brand-new) imaging algorithms.

Active-source seismic surveys use various human-made and -controlled devices, or *sources*, to stimulate stress/displacement (sound, or more properly seismic) waves propagating through the earth, and records the resulting displacement or stress fields at positions on or near the earth's surface. The recordings encode the acoustic structure of the earth through which the waves has passed on their journey from sources to recording devices, or *receivers*. Seismic imaging aims to decode the earth's structure, or at least those features of it to which sound waves are sensitive, from the recorded data.

This branch of earth science originated roughly one hundred years ago, in prospecting for ore bodies before the first World War, and in artillery ranging methods developed during it. Since the early successes in locating petroleum reservoirs along the Gulf of Mexico coast in the 1920's, active-source seismology has become the primary exploration tool of the oil industry, and has also supplied useful methodology to civil and environmental engineering. The topics that form the focus of this book, while at least implicit in the early development of the subject, flowered after the introduction of digital computing in the exploration industry, in the 50's and early 60's. Each subsequent twenty-year period saw the development of a distinct collection of topics:

- 1960's through 1970's: trace-by-trace processing: deconvolution, normal moveout, stack, post-stack migration ("classic imaging");
- 1980's through 1990's: pre-stack and reverse-time migration ("linearized inversion");
- last 20 years: full waveform inversion ("nonlinear inversion").

This book consists of three parts, each devoted to one of the topics listed above. The list above is full of jargon; one of the main goals of this book is to explain each term above (and many others used in this subject) in terms of the mathematics of waves. The story told here is not comprehensive by any means. However, in the pages to follow you will find a few complete accounts, starting in basic linear elastodynamics and arriving at some practical and important imaging and inversion algorithms.

These notes owe a great deal, in structure, content, and spirit, to Jon Claerbout. In preparing the text, I re-read the introduction to Jon's book *Image Estimation by Example*. In Jon's terms, you might say that this course is about "the theory of the similarity between theory and practice" (an emphasis that would clearly not appeal to Jon!). In particular, it is about both theory and practice, with practice represented by computational examples illustrating the theory, using open source software packages and public domain data, both synthetic and field.

These examples take the form of *reproducible computational research*, another pioneering concept of Claerbout's. The reader can rebuild the results from raw data and code by executing a few simple commands, included as auxiliary material with the book. Most importantly, the reader can also modify the examples to explore beyond the content of the text. This computational infrastructure is described in detail in Appendix A, and rests on three open source software packages (Seismic Unix, Madagascar, and TRIP), which in turn depend on Python, C/C++, and Unix. Madagascar provides the computational "glue" tying together tex, data, and code. The chapters form a directory tree of Madagascar papers, each one with examples and resulting figures produced from Python scripts in a project subdirectory. The participants can rebuild every one of these papers from scratch, and the scripts can (and should!) serve as departure points for additional projects.

1.2 Overview

- **Part 1: Basic Imaging.** Starts with an imaging exercise, using 2D marine survey data: RMS velocity estimation, NMO correction of CMP gathers, stack, and poststack time and depth migration. The object of this part of the course is to identify the origin in wave theory of each process used in this exercise. Two approximations to solution of the wave equation are central to this story: linearization, aka single scattering, aka the Born approximation, and geometric optics, aka high frequency asymptotics, aka ray theory. Applied to "locally layered" modeling, these lead to the convolutional model, basic time processing of seismic reflection surveys (deconvolution, statics, NMO, stack), and NMO based velocity analysis. Applied to models with mild lateral variation, it leads to poststack imaging. NMO based velocity analysis provides a starting point for prestack imaging (the topic of the second part of the course) and migration velocity analysis, and serves as an interesting and accessible test bed for inversion algorithms.

- **Part 2: Linear Inversion.** Prestack depth imaging via diffraction sums (Kirchhoff migration), Generalized Radon Transform inversion for acoustics and elasticity, Reverse Time Migration, Least Squares Migration (iterative and true amplitude migration). This time we start with the theory, as the computational examples will use Reverse Time and True Amplitude migration, but the justification for these is essentially the theory of the Generalized Radon Transform. We end by applying these techniques to the 2D marine survey of Part 1.

- **Part 3: Nonlinear Inversion.** Theory of reflection and transmission tomography, relation to migration velocity analysis in extended time and image domains. Inversion beyond single scattering - full waveform inversion. Application to the marine survey of segment 1.

Chapter 2

Basic Imaging of a 2D Marine Survey

2.1 Introduction

The purpose of this paper is to explain how to use simple scripts and public-domain software to create images of the subsurface from reasonably tame 2D seismic reflection data. By "reasonably tame", I mean having few amplitude anomalies and mostly primary reflection energy. These conditions pretty much imply high quality shallow water marine data from regions with low structural relief and well-defined dip trends, shot along dip lines, with successful multiple removal applied.

The Viking Graben (or Mobil AVO) data placed in the public domain by Mobil Research in 1994 appears to be reasonably tame, in the sense just described."Viking Graben" refers to a 25 km long marine line from the North Sea Norwegian sector. The data released by Mobil included both raw field tape (`tape1093`) and parabolic Radon demultiple (`paracdp`) version of the data, along with two logs of v_p, v_s, and ρ, and a far-field wavelet recording. Data are in SEGY format (IBM 4-byte floats, plus text and binary reel headers).

Robert Keys and Douglas Foster were the organizers of the SEG post-convention workshop at which the initial results of various groups' efforts to invert this data were presented. Besides the workshop report volume (Keys and Foster, 1998), Keys and Foster have recently published a web page on this data:

`s3.amazonaws.com/open.source.geoscience/open_data/`
`Mobil_Avo_Viking_Graben_Line_12/mobil_avo.html`

This section of the CAAM 641 course notes discusses so-called standard processing of the Viking Graben data, based on hyperbolic normal moveout (NMO). The connection between inversion and standard processing is the subject of the first part of this course.

An earlier treatment can be found in my MSRI lecture notes from 2013, available on the TRIP web page, downloadable materials, short course and summer school materials, especially part 2.4. I also show how to transform the by-products of standard processing into suitable input for prestack imaging via RTM.

My workflow is recorded in the `project/SConstruct` file. The reader should refer to this file for the detailed form of the commands used. The following sections outline the rationale for the choice of commands and their sequence. To reproduce my results, the reader will need to install SU, Madagascar, and TRIP packages. SU should be version 44 or later. The prestack part of the workflow requires parallel installation of TRIP, and at least 10 threads of execution, preferably more, to finish in reasonable time.

2.2 Preliminary Steps

Extract the data using `Fetch`, from the TRIP data repository. The result is `paracdp.segy`, a SEGY-format file. The SU command `segyread` converts this file to SU format (no reel headers, little-endian 4-byte floats, same number of traces). Besides parabolic radon demultiple, this data was delivered with a rather strict mute applied, with no signal before 1 s. No further mute is applied in the processing sequence described here.

Evidently there are some bad traces in some of the low-number CDPs, with floating point garbage in some of the samples. On average, random binary digits form numbers many orders of magnitude larger than the $O(10^8)$ samples encountered in "good" CDPs. So an adequate strategy is to use `sugain` to reject samples in the 99%ile, and this command is included in the conversion to SU format.

2.3 Extract Analysis CDPs

I chose 4 CDPs (Figure 2.1a, Figure 2.1b, Figure 2.2a, Figure 2.2b) roughly evenly spaced, including the end zones, separated by about 8 km. This is a preliminary guess as to the necessary lateral resolution for velocity analysis. It turned out to be reasonable.

Also, I truncated the traces at 3 s, as the logged zone indicated that only reflection arriving before that time would likely image the exploration targets. This is of course an ex post facto decision, but in fact it becomes clear quickly that most of the energy after perhaps 4 s must be other than primary reflection, either multiples or sideswipe.

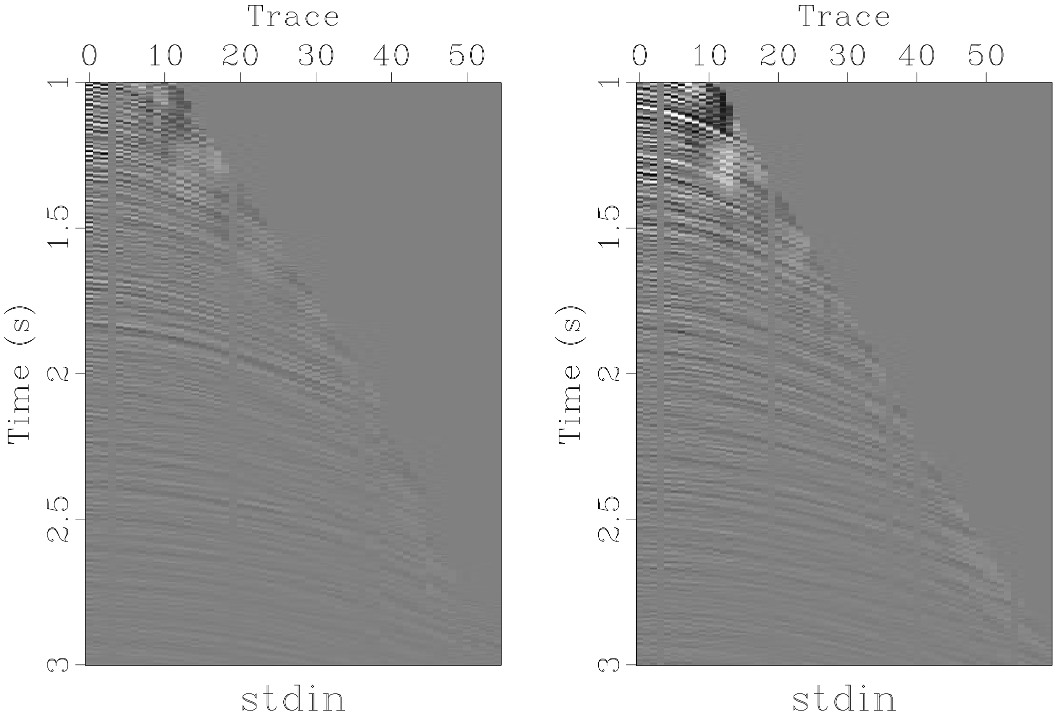

Figure 2.1: Left: CDP 200; Right: CDP 700

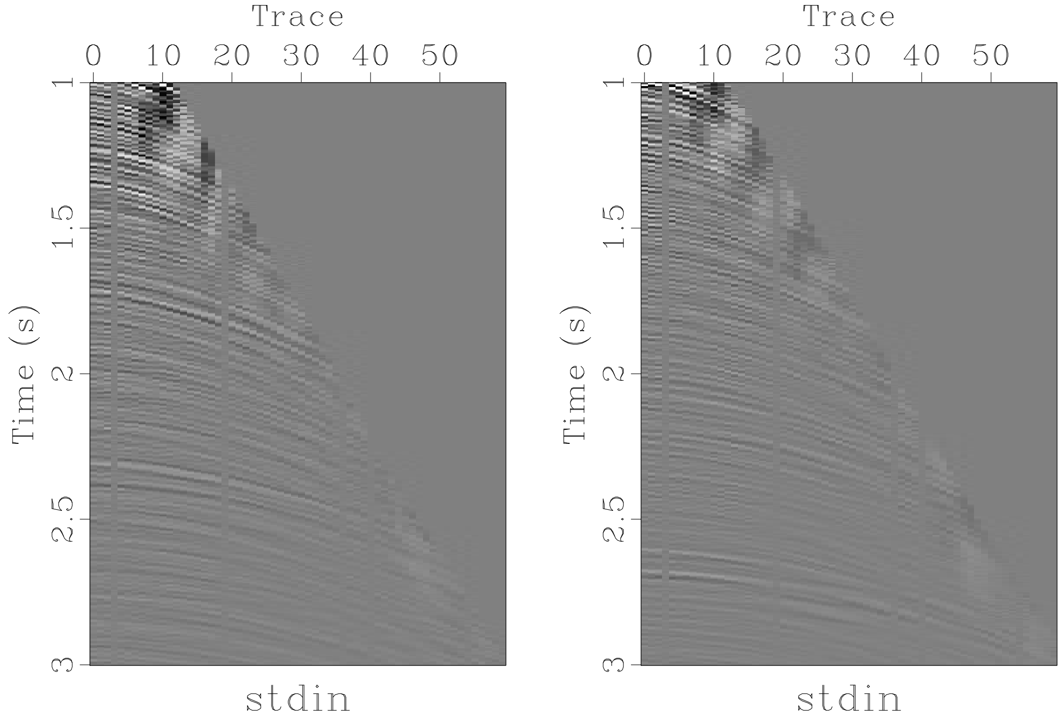

Figure 2.2: Left: CDP 1300; Right: CDP 2000

2.4 Velocity Analysis

Create velan panels, one for each CDP: Figure 2.3a, Figure 2.3b, Figure 2.4a, and Figure 2.4b. Note that in each case there is a trend of strong bullets or streaks at lower velocities than some nearby higher velocity peaks. The higher velocity peaks are weaker, but pick them anyway. The stronger lower velocity features represent multiple reflections - the suppression of multiple energy notwithstanding. (If you carry out this exercise vvwith the field tapes, you will see nothing but water and near-water velocity features).

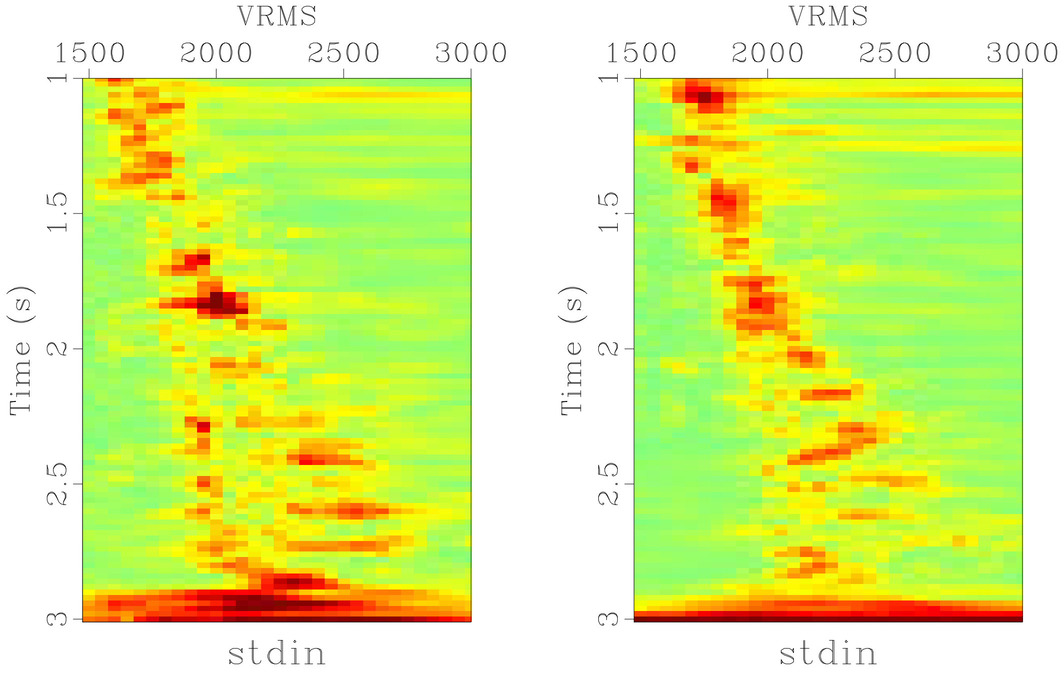

Figure 2.3: Left: Velocity spectrum for CDP 200; Right: Velocity spectrum for CDP 700

To QC your picks, plot the NMO correction of the corresponding CDPs as Figure 2.5a, Figure 2.5b, Figure 2.6a, and Figure 2.6b.

These CDPs are pretty easy to pick from the velocity spectra; however it is instructive to see what might go wrong. For CDP 2000, suppose you were to pick velocities of 1800 m/s at t=1.2, 1.6. and 1.8 s, and 2200 m/s at t=2.1 s - each of these coordinates with a feature on the velocity panel - instead of 1950, 2050, and 2250 m/s at 1.25, 1.6, and 2.25 s, as was done to create Figure 2.6b. Then you would generate Figure 2.7 instead.

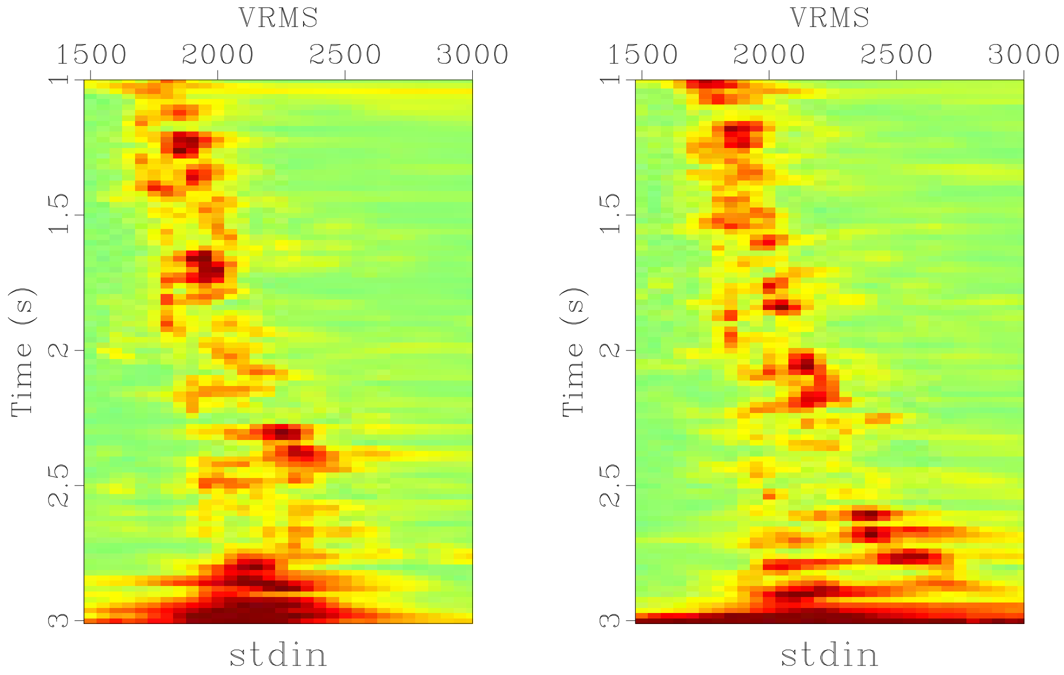

Figure 2.4: Left: Velocity spectrum for CDP 1300; Right: Velocity spectrum for CDP 2000

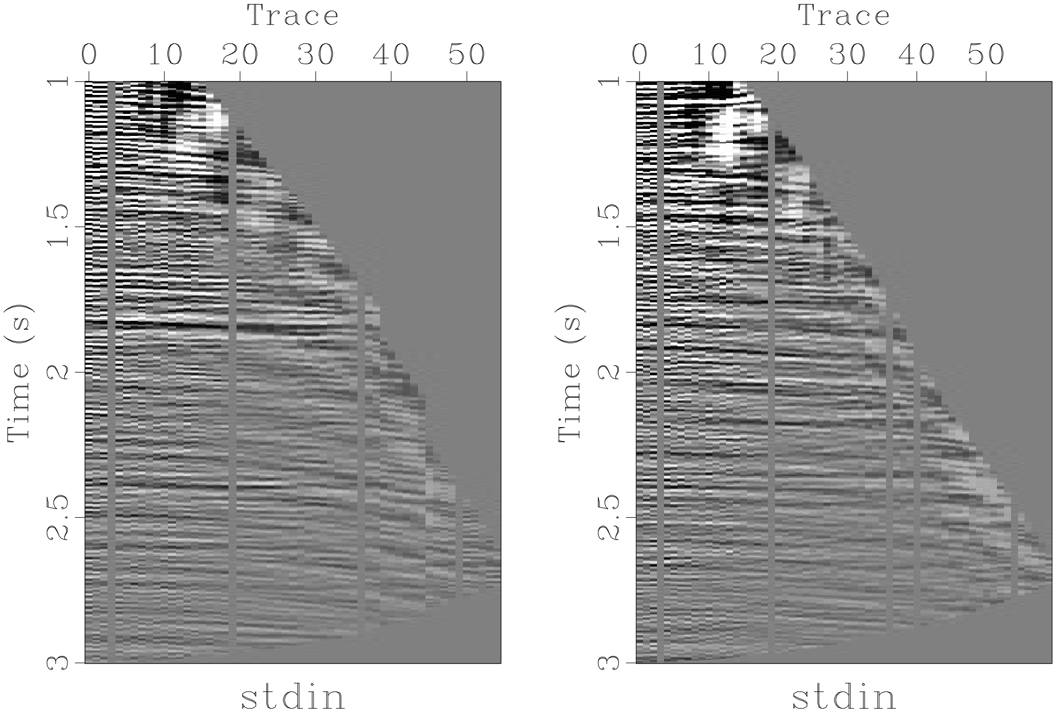

Figure 2.5: Left: NMO corrected CDP 200; Right: NMO corrected CDP 700

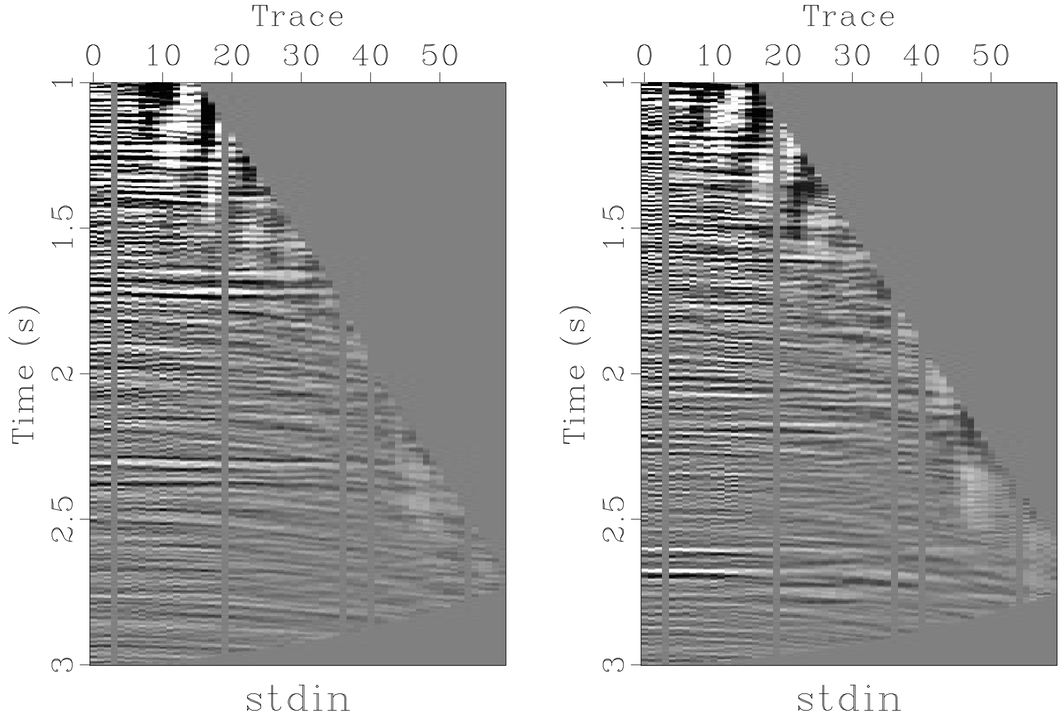

Figure 2.6: Left: NMO corrected CDP 1300; Right: NMO corrected CDP 2000

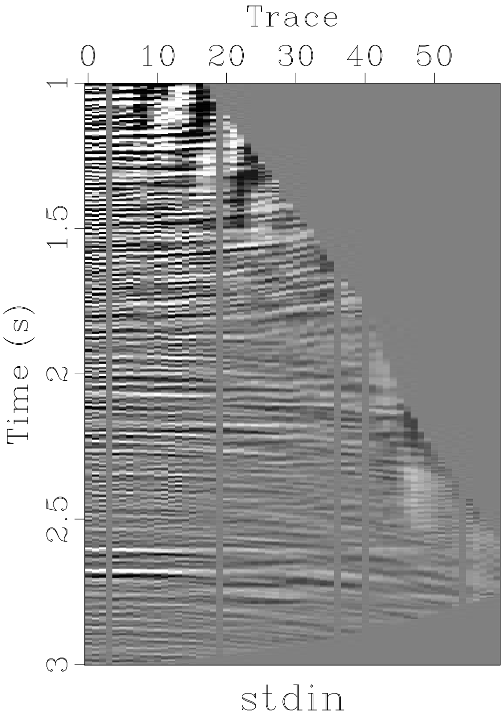

Figure 2.7: NMO corrected CDP 2000 with tnmo=0.0,1.2,1.5,1.8,2.1,2.6,2.75 vnmo=1500,1800,1800,1800,2200,2400,2700

2.5 Stack and Post-Stack Migration

The velocity-denominated quantity determined by flattening NMO-corrected CDP gathers is the so-called RMS velocity, or $v_{\rm rms}$, the root mean square of a local wave velocity expressed as a function of vertical travel time t_0:

$$v_{\rm rms}^2(t_0) = \frac{2}{t_0}\int_0^{t_0} d\tau \tilde{v}^2(\tau). \tag{2.1}$$

Vertical traveltime is related to depth by

$$z = \int_0^{t_0} d\tau \tilde{v}(\tau) \tag{2.2}$$

in which $\tilde{v}$ is the local (or *interval* wave velocity as a function of t_0. Velocity as a function of depth is then given by composing $\tilde{v}$ with the inverse of the tranformation (2.2).

This construction can be carried out for every midpoint. If the subsurface structure is sufficiently laterally homogeneous, then to good approximation the actual wave velocity is well-approximated by the interval velocity in the "well" below each midpoint. The interval velocity as function of t_0 or z can then be used to perform migration of zero-offset data, to good approximation. Also, to good approximation zero-offset data is similar to the stack of NMO-corrected CDPs.

This process requires a velocity at every CDP. So far we have only determined four such RMS velocity profiles. However the command `sunmo` will interpolate between midpoints, and extrapolate towards the ends of the line as necessary, and even output the interpolated $v_{\rm rms}$ as a function of t_0 and midpoint (keyword `voutfile`). This output is critical, as it can be converted later into other forms of velocity suitable for time and depth migration. Of course, the command also produces a stacked section (Figure 2.8).

Because the deeper events are much less energetically imaged than the shallow, automatic gain control (amplitude equalization) produces a much more informative Figure 2.9. This AGC'd stack will be input to all poststack processes.

You will note the precence of many *diffraction tails* especially below 2 s. These may be partially collapsed either by poststack time migration using Gazdag's algorithm (Figure 2.10) or, more effectively, by poststack depth migration using one of several methods. Gazdag poststack time migration assumes layered velocity structure, hence uses only the first column of the interval velocity as function of time. Nonetheless Figure 2.10 shows considerably less interference from diffraction tails; the graben structures that give this prospect its name are now clearly in evidence.

Poststack depth migration (a process that will be identified later as an approximation to the adjoint linearized zero-offset modeling operator) does even better - Figure 2.11

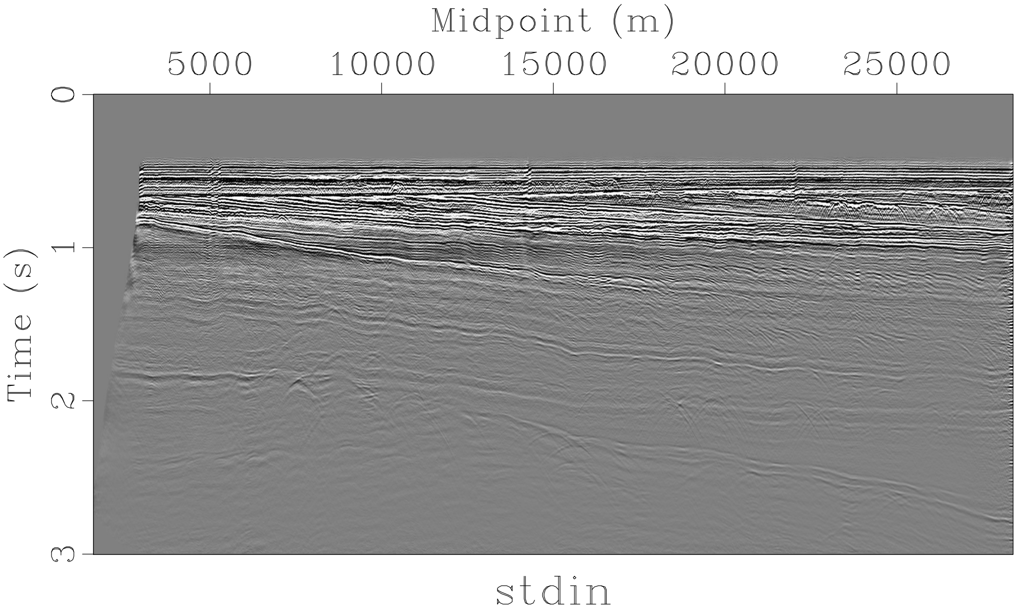

Figure 2.8: Stack with linearly interpolated, constant extrapolated v_{rms}

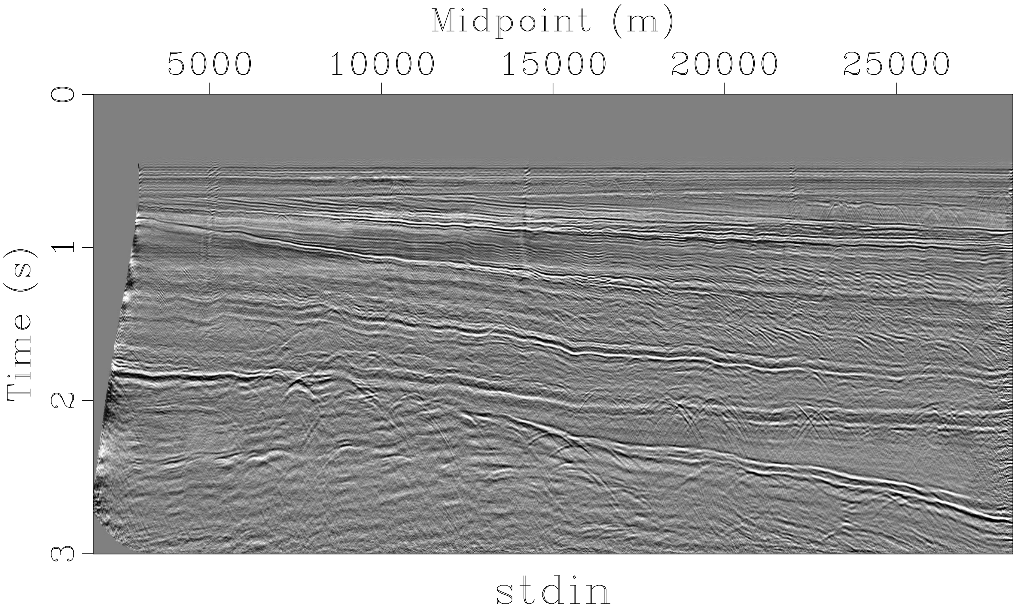

Figure 2.9: Automatic Gain Control (AGC) applied to data of Figure 2.8. Note pronouced diffraction hyperbolae in the deeper part of the secion.

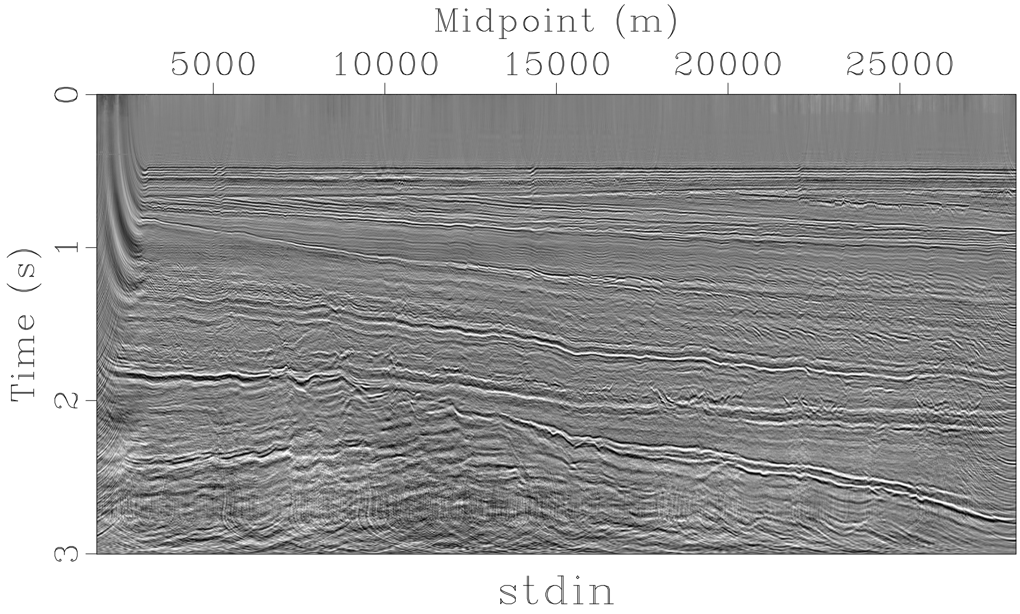

Figure 2.10: Gazdag's phase-shift time migration applied to the stack (Figure 2.8), followed by AGC. Diffraction artifacts are largely collapsed to the diffracting points that caused them, and the graben structures that give this prospect its name begin to be clearly visible. This migration assumes horzontal layering, which is not correct on the scale of 25 km.

displays the output of Phase-Shift-Plus-Interpolation poststack depth migration, using the velocity $v(z,x)$ depicted in Figure 2.12 (obtained via use of the SU utility `velconv`, see `project/SConstruct` for details). The image gives reasonable depths for the various horizons. Also, this velocity model is the beginning of the next, prestack phase of Viking Graben processing.

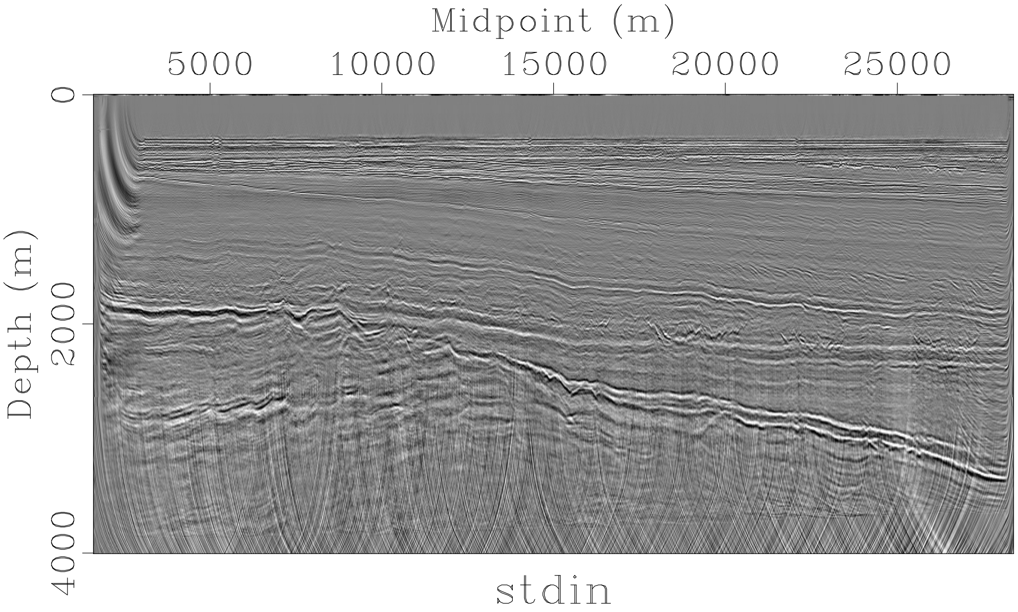

Figure 2.11: Gazdag's phase-shift-plus-interpolation post-stack depth migration applied to the stack (Figure 2.8), followed by AGC. Accounts for lateral velocity variation - uses interval velocity as function of midpoint and depth - more accurately than does the time migration of Figure 2.10. The geology is even more clearly delineated, and depths should approximate those to be obtained by more sophisticated imaging.

With an eye to the prestack processing to be discussed later in the course, I include analogous results for downfiltered data. Jie Hou suggested a 5-10-30-40 Hz bandpass filter. The resulting AGC'd stack (Figure 2.13) shows the expected decrease in resolution over the original data, which has significant energy over 50 Hz.

[A question: the process, as you will see from the SConstruct, consists in filtering the data then applying NMO. Since the stack is a collection of time traces, you might think that you could get the same result by applying NMO first, then filtering. Is this true?]

The PSPI post-stack depth migration (Figure 2.14) reveals the same structure as does the stack, with the same decreased resolution over Figure 2.11, but with diffraction tails suppressed and stretched to depth.

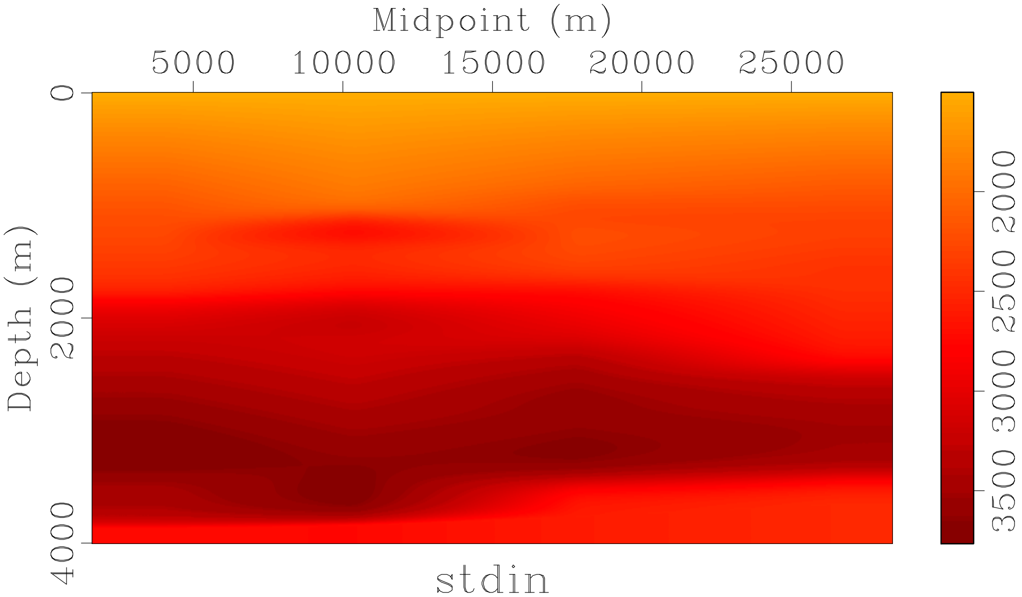

Figure 2.12: Interval velocity as function of depth, derived from NMO velocity analysis. Probably not to be taken seriously below 3 km - for initial MVA estimate should extend by contant from 3 km and smooth.

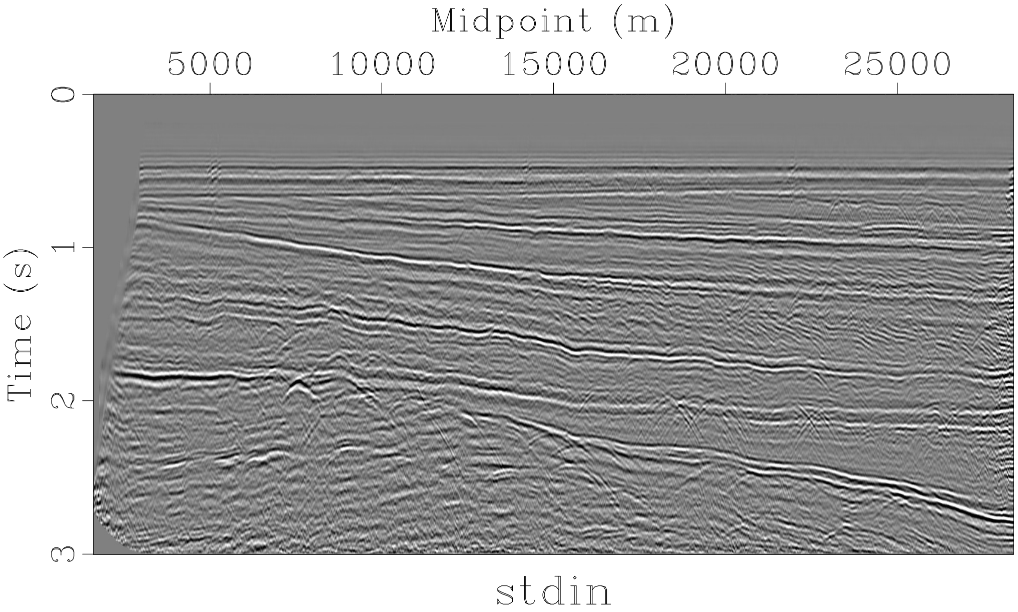

Figure 2.13: NMO stack of (5,10,30,40) Hz bandpass filtered data, after AGC. Compare to Figure 2.8.

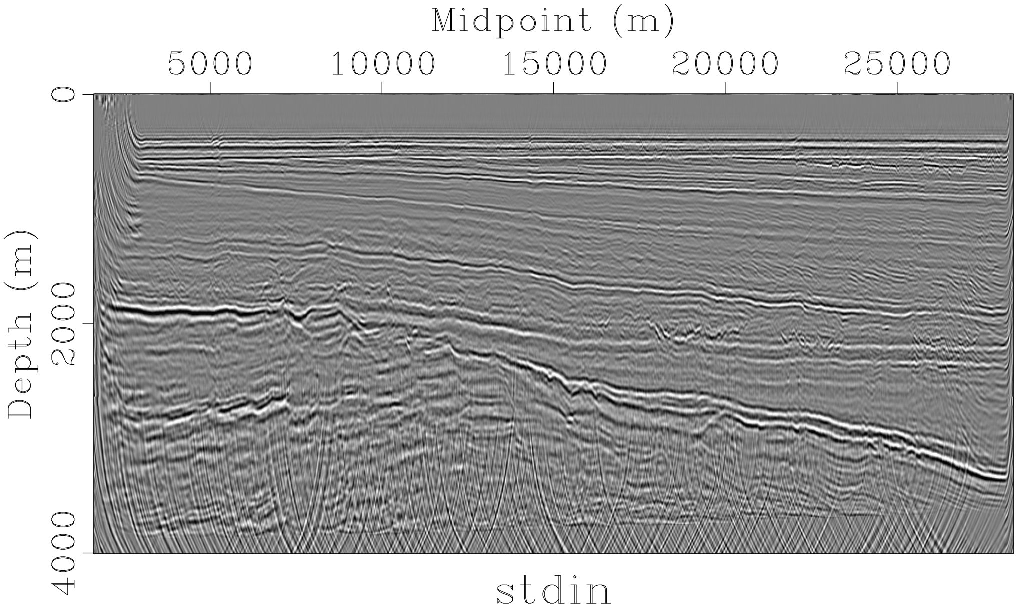

Figure 2.14: PSPI post stack migrated image from stack in Figure 2.13.

2.6 Conclusion

Use of simplified physics (NMO, stack, poststack migration) based on layered modeling produces plausible subsurface structure images from the Viking Graben data. Public domain software - in this exercise, mostly SU - provides enough functionality to carry out this imaging task, provided that some basic pre-processing is performed (multiple suppression, mute). Mobil provided a suitably preprocessed version of the data in 1994.

The sequence of processing steps takes us far enough from the basic physics of wave propagation to raise questions about whether the images produced here are actually images of anything real. Later in the course, we turn to prestack processing based more directly on the acoustic wave equation - only a step further, but meaningful nonetheless. To telegraph part of the punch: prestack processing supports the validity of the images obtained in this paper by poststack processes. It seems that in the case of this “tame” data, classic seismic data processing does not steer you wrong.

2.7 Suggested Projects

1. The velocity analysis presented above seems quite minimal - only four analysis CMPs selected, more or less evenly spaced. Is there any advantage in choosing more analysis points? Does the stack become cleaner, the interval velocity any more believable?
2. Find another field data set and carry out a similar analysis.
3. Taking another tack, find a synthetic model with more lateral variation in velocity - Marmousi is an obvious choice - and carry out NMO based velocity analysis. Does the interval velocity so obtained bear any relation to the actual velocity (which of course you know, since the example is synthetic)? What about the PSPI migrated image and the actual model?
4. AGC is a crude tool for equalizing amplitudes, that completely destroys the relation between event and reflector dynamics. Is it possible to use a simpler scaling, such as by a power of t or z, to achieve the same enhancement of important reflectors at depth while maintaining a simple relation between reflector and reflection amplitudes?

REFERENCES

Keys, R., and D. Foster, 1998, Comparison of seismic inversion methods on a single real dataset: Society of Exploration Geophysicists.

Chapter 3

Born Approximation

3.1 Introduction

Several features of the seismic wave field are obvious from field recordings:

- it consists largely of waves, that is, coherent space-time structures with definite apparent velocities;
- the velocity of seismic waves varies with position in the earth;
- some of the waves appear to be reflections, that is, the result of interaction with the earth's structure that changes (or even reverses) the direction of wave motion.

The simplest model of wave motion that predicts these features is *constant density acoustics*, connecting the acoustic pressure field $p(\mathbf{x},t)$, the compressional wave velocity $v(\mathbf{x})$, and a source wavefield $f(\mathbf{x},t)$ through the partial differential equation

$$\begin{aligned} \frac{\partial^2 p}{\partial t^2} - v^2 \nabla^2 p &= f \\ p &= 0, t << 0 \end{aligned} \tag{3.1}$$

The second condition guarantees that the field is causal, that is, vanishes before the onset of nonzero source wavefield values.

This section takes a first look at the relation between p and v implicit in the system 3.1, and introduces the linearization of this relation, often called (somewhat imprecisely) the Born approximation in the seismic literature. The Born approximation is both a powerful analytical tool and the foundation for much of seismic data processing. It is particularly accurate when the reference model is smooth, or slowly varying, on the wavelength scale, and the perturbation is oscillatory, that is, contains all wavelength-scale features. Under these circumstances, Born data represents the part of the wavefield that has interacted once with the wavelength-scale features in the model, that is, the singly-scattered field, and provides a rough definition of the primary reflection field. These features are illustrated via a model constructed from the NMO-based velocity analysis and PSPI depth image derived from the Mobil Viking Graben data in Chapter 2.

3.2 Refraction

Note that the wave equation 3.1 does not necessarily predict wave motion: if the velocity field v is sufficiently complex, not only may waves not be observable, but energy may not even flow as one would expect in a wave setting. So called Anderson localization, in which a complex material with highly spatially variable index of refraction ($1/v$) causes energy to remain for arbitrarily long periods in the same location, or to diffuse slowly rather than propagate, has been predicted theoretically and even observed in the lab (Hu et al., 2008; Wright and Weaver, 2010). Clear evidence of waves at seismic frequencies already tells us something about the structure of the earth's interior.

The classical radiation solution for the wave equation (Courant and Hilbert, 1962) is the pressure field $p(\mathbf{x},t)$ satisfying 3.1 with constant (space-independent) v and $f(\mathbf{x},t) = w(t)\delta(\mathbf{x})$. As is well-known, the solution in 3D is

$$p(\mathbf{x},t) = \frac{w(t-r/v)}{4\pi r}, \; r = |\mathbf{x}| = \sqrt{\mathbf{x}^T\mathbf{x}}. \tag{3.2}$$

The 2D case has a more complex expression:

$$p(\mathbf{x},t) = \frac{1}{\pi v^2}\int_0^{\sqrt{t-r/v}} d\sigma \frac{w(t-\sigma^2-r/v)}{\sqrt{\sigma^2+2r/c}}. \tag{3.3}$$

The constant v radiation solution predicts waves moving at speed v, but (of course) their velocity does not vary with position. Also, this model does not generate anything that looks like reflected waves. So, reflection (and even more so, variable wave velocity) implies spatially variable wave velocity.

Next examine wave propagation via numerical approximation. Throughout these notes, I use finite difference schemes to model wave propagation. A good overview of finite difference methods appears in (Moczo et al., 2006). These methods approximate the actual solution with an asymptotic rate called the scheme order. Actually only some solutions are approximated with this order, in general - see (Symes et al., 2008) or (Symes and Vdovina, 2009) for examples in which the formal order does not predict the actual convergence behaviour. In general, for smooth solutions and coefficients (like v), the formal order actually predicts the rate of convergence: if the scheme is of second order, the error will be roughly proportional to Δt^2 for small enough Δt, for instance. For the examples shown in this section of the notes, we use the classic centered difference scheme, with second order centered difference in time replacing the time derivatives in 3.2, and $2k$th order centered differences replacing the spatial derivatives (usually $k = 2$, so the scheme is fourth order in space). Singular right hand sides are represented by numerical delta functions.

A slightly more realistic propagating medium from the constant case (equation 3.2) is the interval velocity predicted by NMO velocity analysis from the Mobil AVO data (Symes, 2017a), which is slowly varying on the wavelength scale. [Average velocity = 2500 m/s, wavelength at 20 Hz = 125 m.] Figure 3.1 shows the square of this interval velocity. Since we have no density information in this model, this field may be regarded as proportional to bulk modulus, with constant density.

Up to this point boundary conditions have not been mentioned: the radiation solution does not interact with any boundaries. The problem solved numerically in the following examples takes place in a bounded rectangle in $\mathbf{R}^3$, namely the one shown in Figure 3.1. Every boundary is subject to the pressure-free condition $p = 0$. The domain is so chosen that for the point sources used in the numerical experiments, the interaction of the wavefield with the side and bottom boundaries does not affect the simulated trace data - as will be evident shortly. The free surface at the top of the model does interact with these solutions, however.

Figures 3.2, 3.3 and 3.4 show the computed pressure wavefield at $t = 0.5s, 1.0s$ and $1.5s$, for a shot at $x_s = 10012.5m$ (in the coordinate system of the trace headers). The wavelet is a 5-10-20-25 Hz zero-phase bandpass filter.

There is little evidence of reflection occuring in this example. The shot gather (all acquisition parameters as in the field data, reported in (Symes, 2017a)) shows the direct wave and a diving wave arrival - at longer offsets, the latter will overtake the former.

3.3 Reflectors and Reflection

From the homogeneous and slowly varying examples, one could infer that only transmitted (direct and diving wave) arrivals occur in models for which the wave velocity (and eventually other parameters) change slowly (are smooth) on the scale of a wavelength. This supposition is correct, and will be fully justified in the next section (Symes, 2017b). Therefore, to make reflections appear in model-derived synthetics requires a model having substantial variation on the wavelength scale.

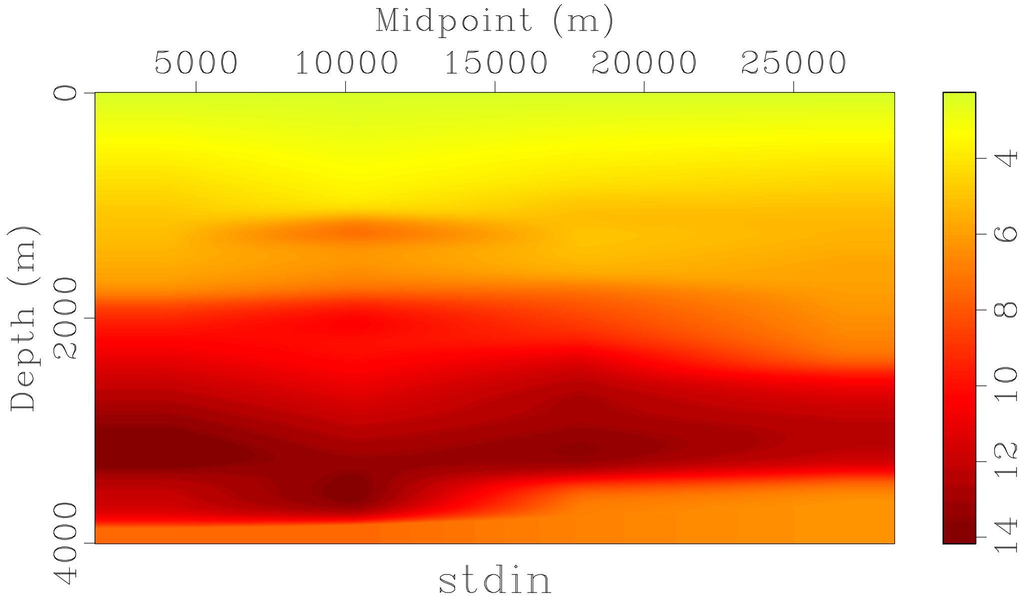

Figure 3.1: Interval velocity as function of depth, derived from NMO velocity analysis. Probably not to be taken seriously below 3 km.

We create such a model by adding a scaled version of the PSPI depth-migrated image from (Symes, 2017a), shown in Figure 3.6. Keeping the velocity between 1.0 and 5.0 km/s by using a scale factor of 0.15, we obtain the model shown in Figure 3.7. This model has wavelength scale features, and we would expect it to generate reflections. In fact it does, as can be seen in the snapshots at 0.5, 1.0, and 1.5 s (Figures 3.8, 3.9 and 3.10). The shot gather, Figure 3.11 appears to be rich in reflected events.

Recall that the reference data (Figure 3.5) created from the slowly varying model (Figure 3.1) contains only direct and refracted waves. These are also visible in the data (Figure 3.11) created from the model with reflectors (Figure 3.7), and the wavefield snapshots also suggest that the transmitted waves persist. It is natural to speculate that the difference between these two shot gathers might contain only reflections: we plot this difference in Figure 3.12.

3.4 Linearization

Of course the relation between v and p implied by 3.1 is nonlinear, so the change between the shot gathers of Figures 3.5 and 3.11 cannot be linear in the perturbation (Figure 3.6).

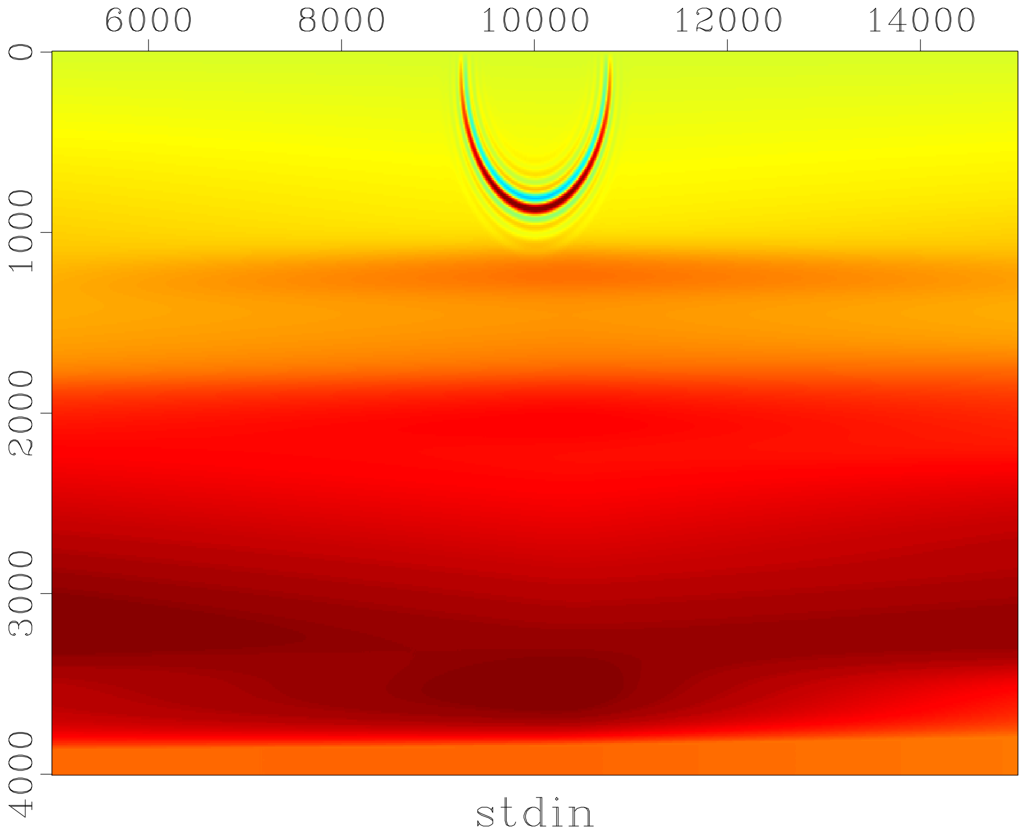

Figure 3.2: Pressure field at $t = 0.5s$ for a point source at $x = 10012$ m. Wavelet is 5-10-20-25 Hz zero-phase bandpass filter. Velocity field derived from NMO velocity analysis, square depicted in Figure 3.1.

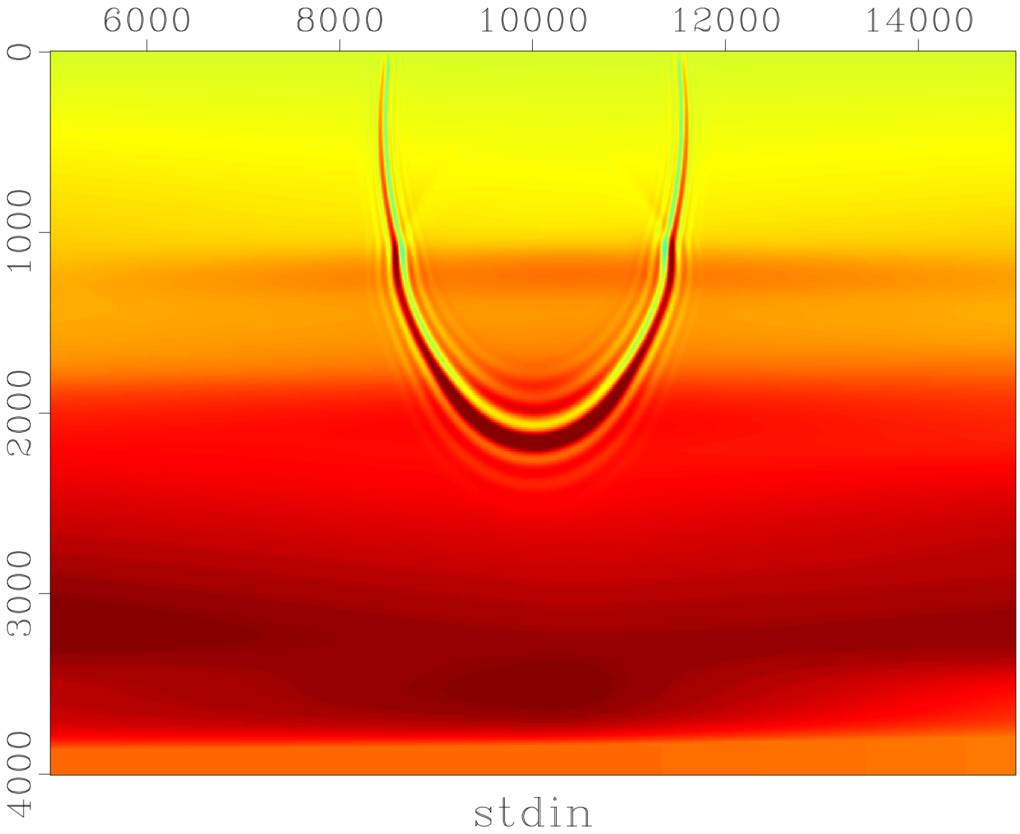

Figure 3.3: Pressure field at $t = 1.0s$ for a point source at $x = 10012$ m. Wavelet is 5-10-20-25 Hz zero-phase bandpass filter. Velocity field derived from NMO velocity analysis, square depicted in Figure 3.1.

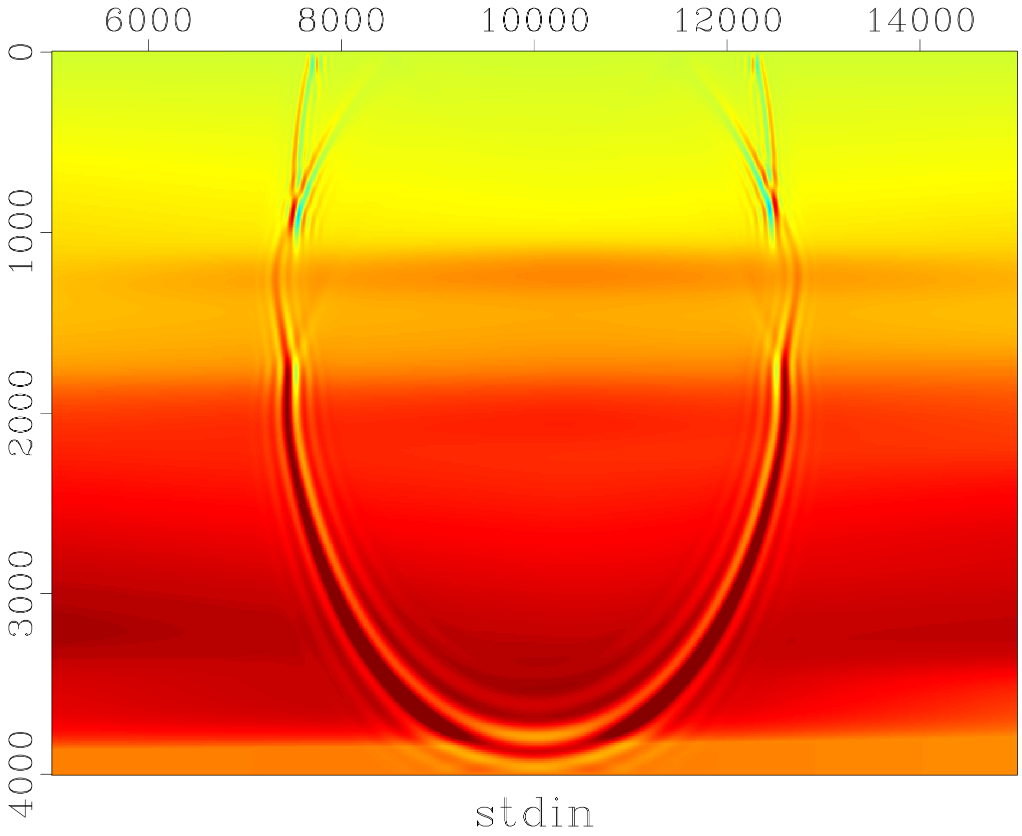

Figure 3.4: Pressure field at $t = 1.5s$ for a point source at $x = 10012$ m. Wavelet is 5-10-20-25 Hz zero-phase bandpass filter. Velocity field derived from NMO velocity analysis, square depicted in Figure 3.1.

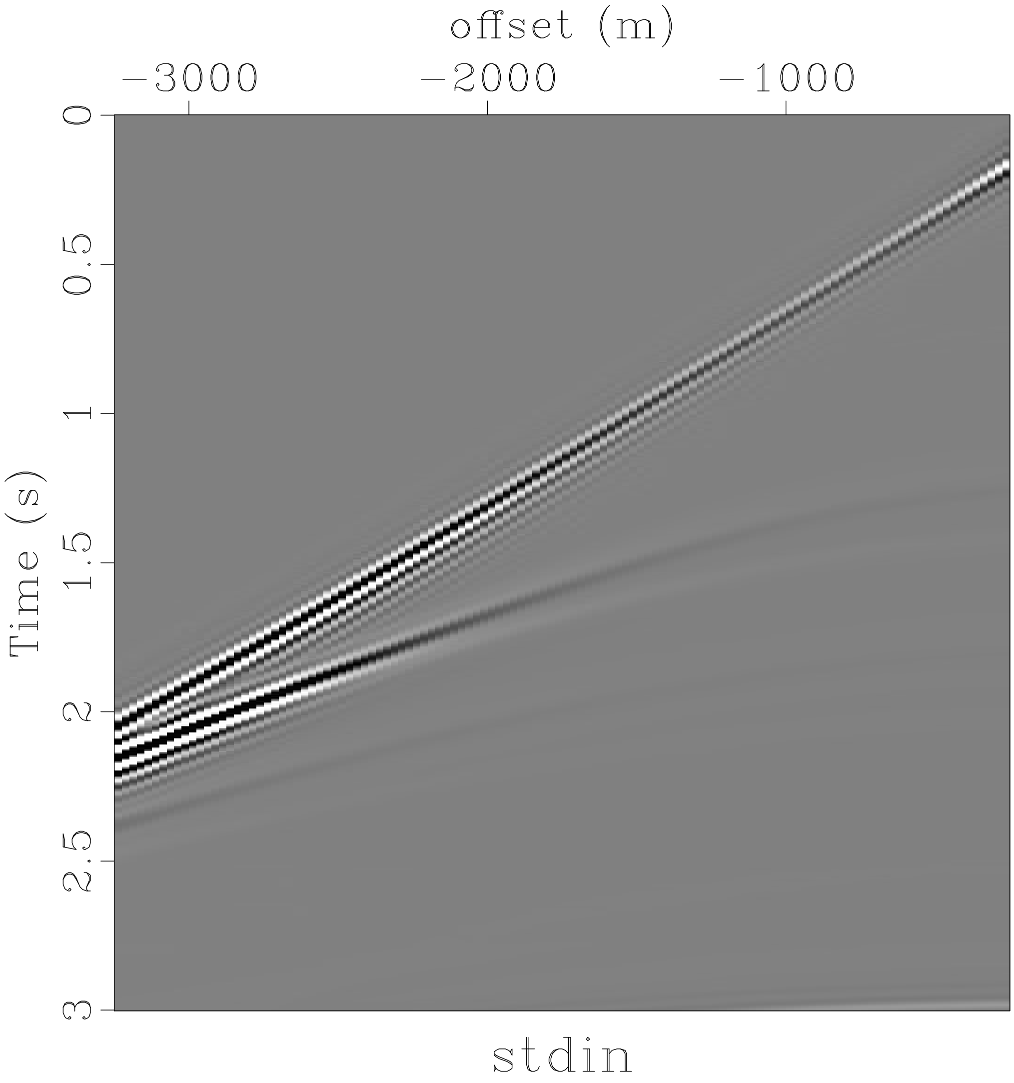

Figure 3.5: Pressure shot gather for shot at $x = 10012$ m. Wavelet is 5-10-20-25 Hz zero-phase bandpass filter. Velocity field derived from NMO velocity analysis, square depicted in Figure 3.1.

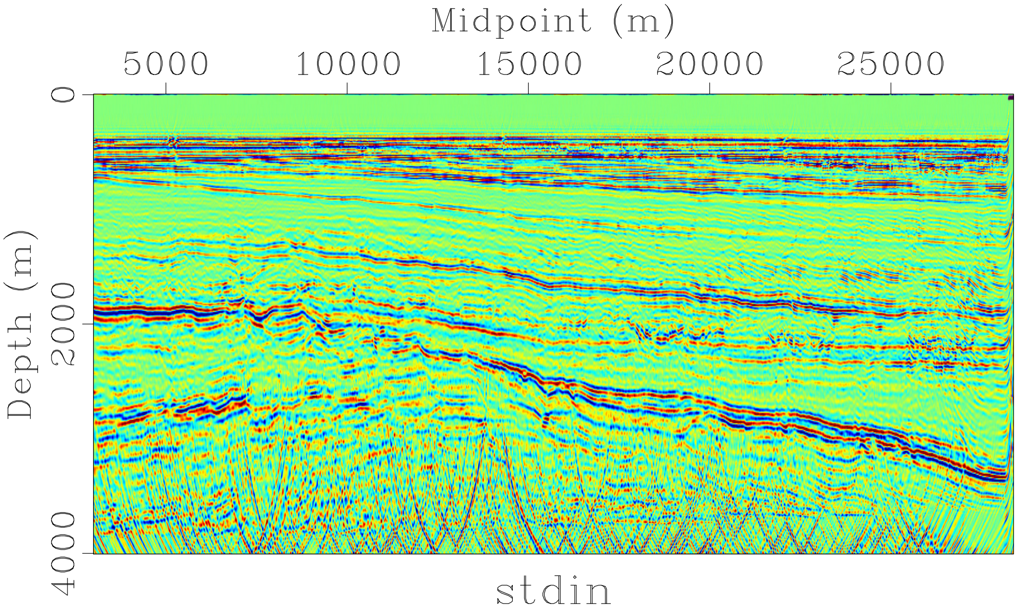

Figure 3.6: PSPI depth migration of Mobil AVO data using square velocity of Figure 3.1.

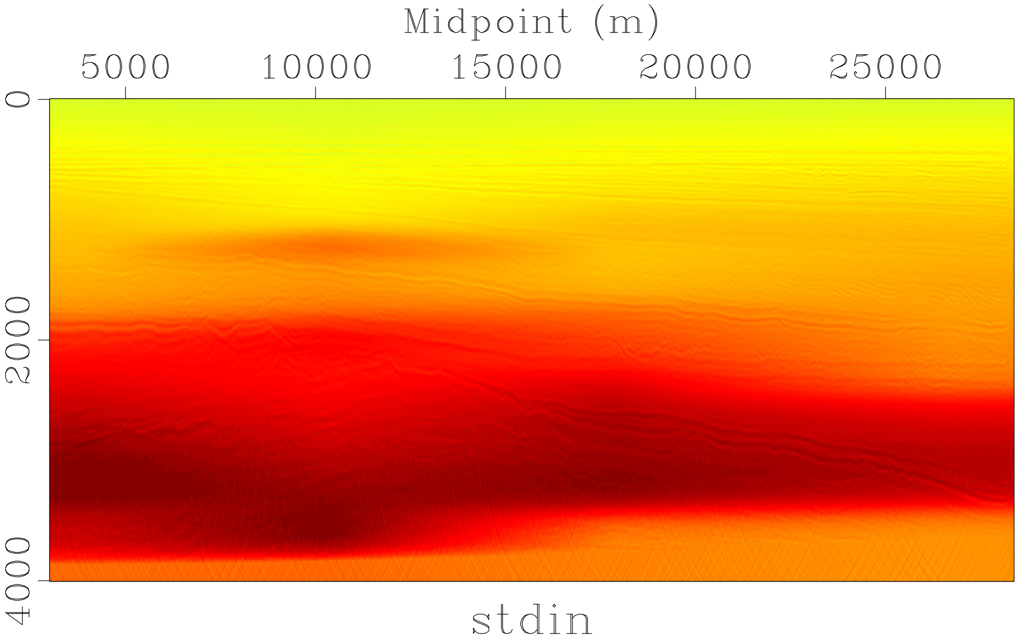

Figure 3.7: Interval squared velocity as function of depth, combines background model of Figure 3.1 with PSPI poststack image, scaled by 0.15.

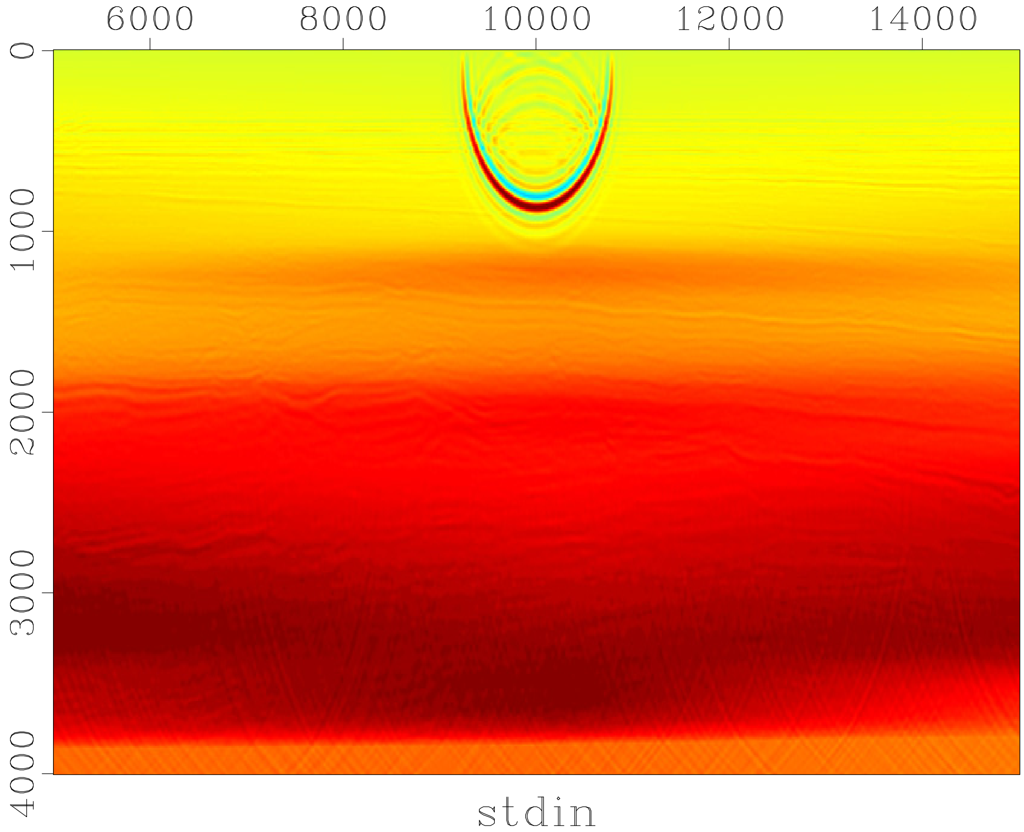

Figure 3.8: Pressure field at $t = 0.5s$ for a point source at $x = 10012$ m. Wavelet is 5-10-20-25 Hz zero-phase bandpass filter. Velocity field depicted in Figure 3.7.

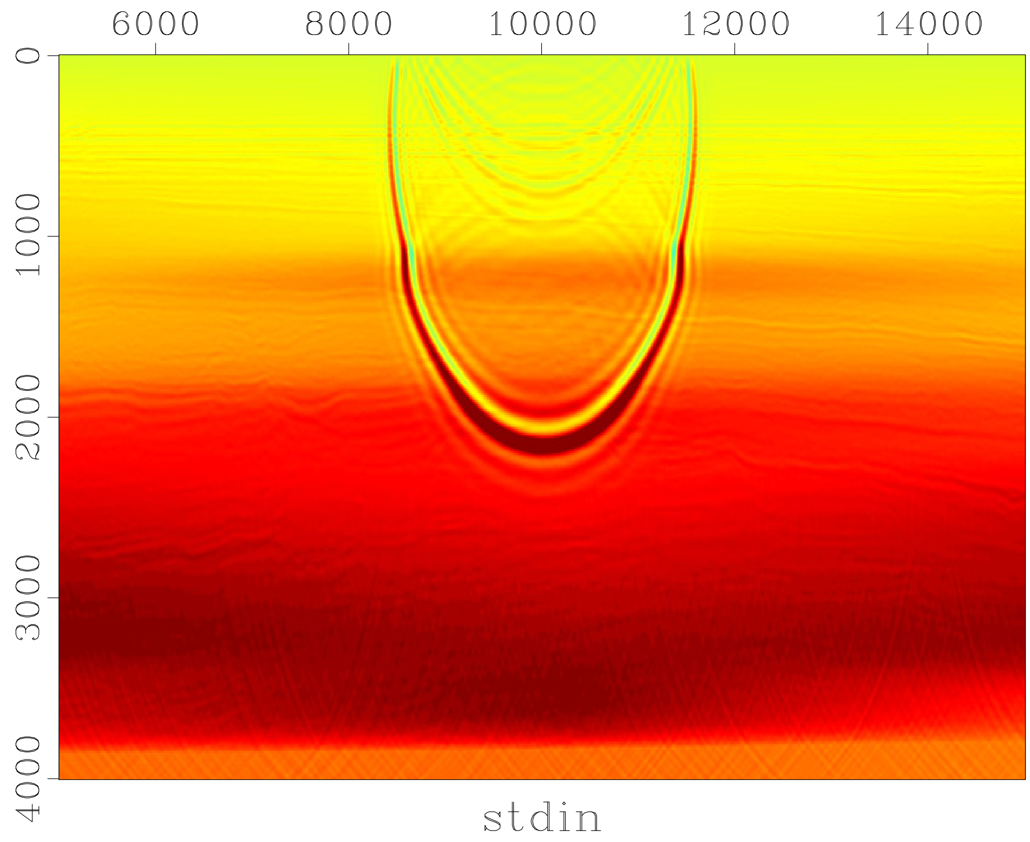

Figure 3.9: Pressure field at $t = 1.0s$ for a point source at $x = 10012$ m. Wavelet is 5-10-20-25 Hz zero-phase bandpass filter. Velocity field depicted in Figure 3.7.

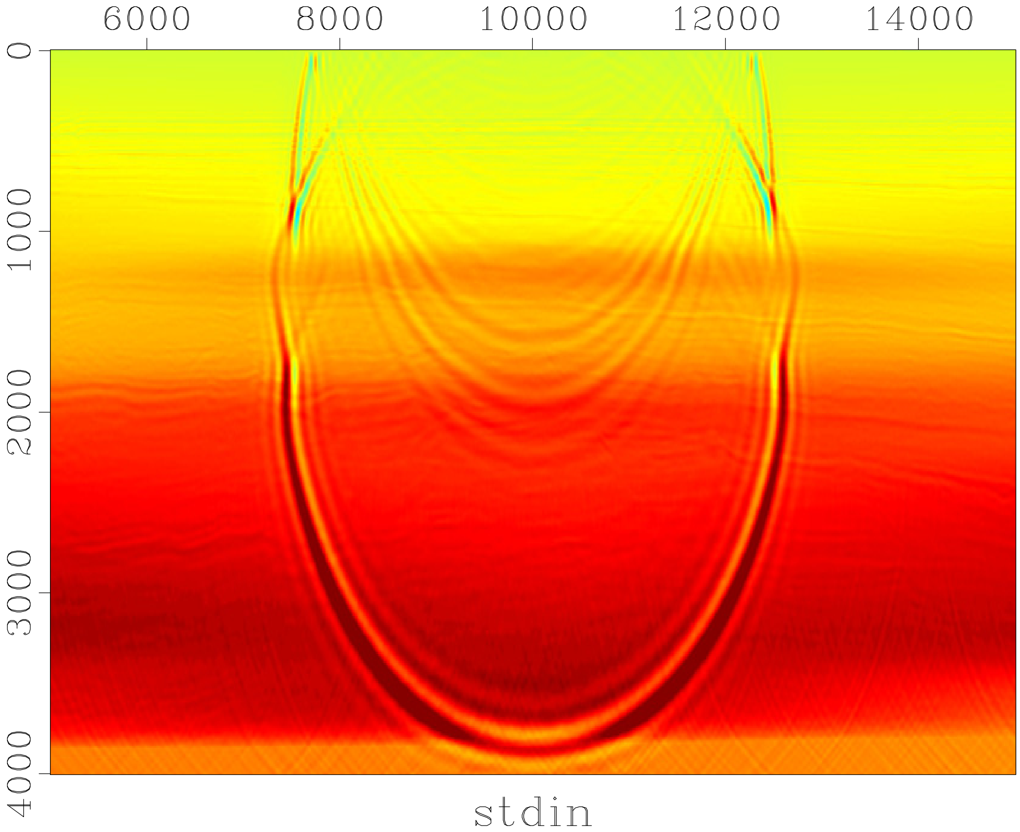

Figure 3.10: Pressure field at $t = 1.5s$ for a point source at $x = 10012$ m. Wavelet is 5-10-20-25 Hz zero-phase bandpass filter. Velocity field depicted in Figure 3.7.

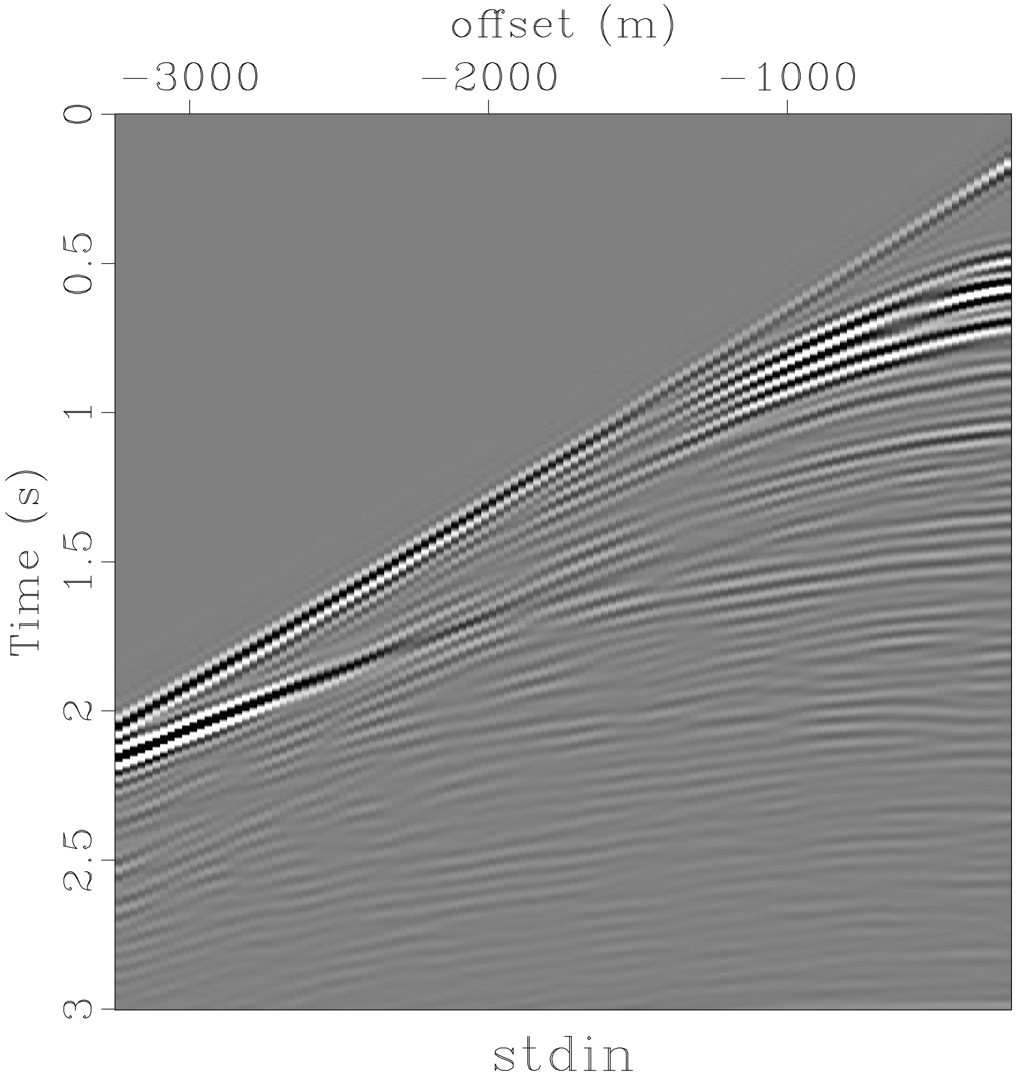

Figure 3.11: Pressure shot gather for shot at $x = 10012$ m. Wavelet is 5-10-20-25 Hz zero-phase bandpass filter. Velocity field depicted in Figure 3.7.

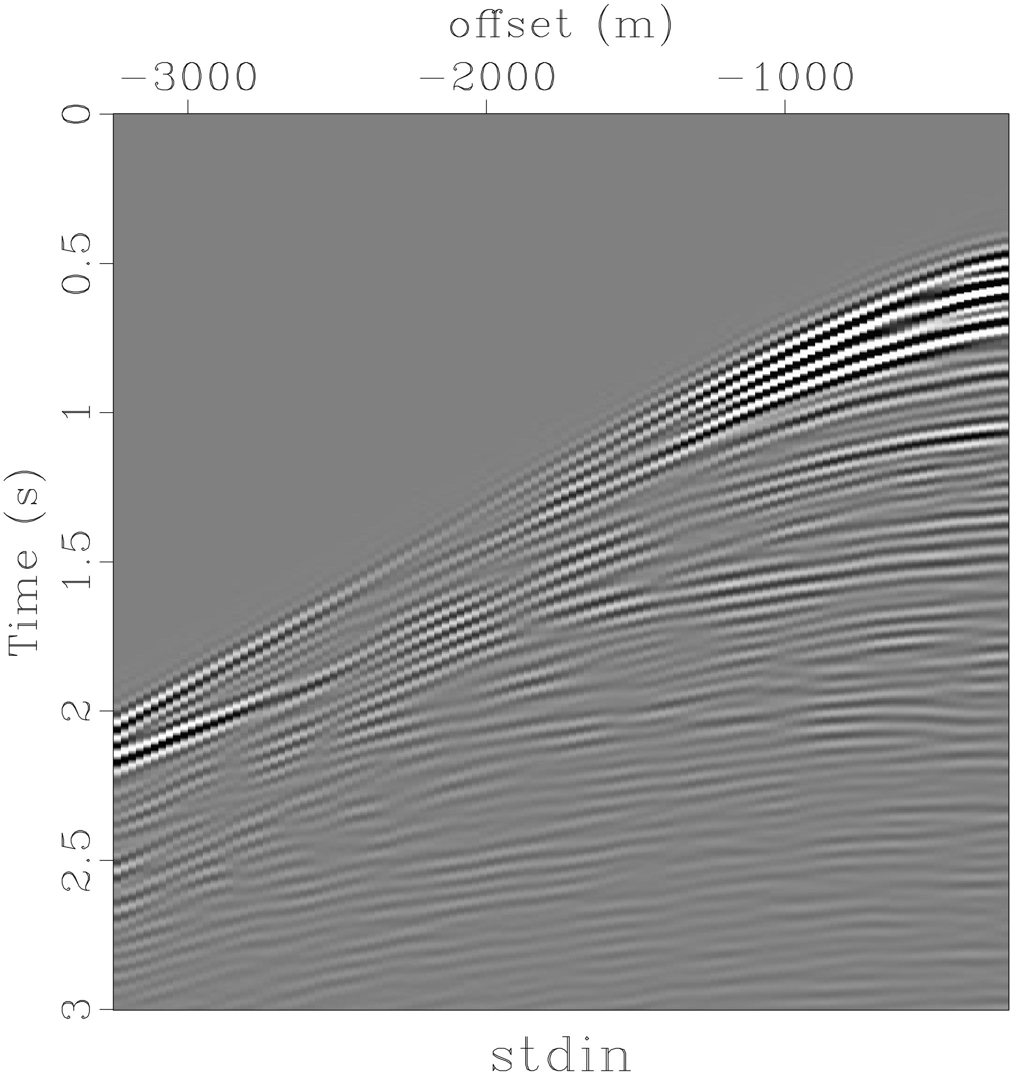

Figure 3.12: Difference between shot gathers for v^2 and $v^2 + \epsilon \delta v^2$, with $\epsilon = 0.15$ (shown in Figures 3.5 and 3.11).

On the other hand, the RMS (or L^2 norm) of the perturbation is approximately 2% of the L^2 norm of the square velocity, so in some sense a small perturbation, and it is natural to think that a linear approximation to the model-data relation would be predictive of the data change. Linearization, or (as it is somewhat imprecisely known in this business) the Born approximation. amounts to replacing both p and v by perturbed quantities $p + \delta p$, $v + \delta v$, substituting these in equation 3.1, throwing away all terms with two or more δs in them, and using 3.1 itself to eliminate a few more terms. The upshot is a relation between the reference fields p, v (presumed to solve 3.1) and perturbation fields $\delta p, \delta v$:

$$\begin{aligned} \frac{\partial^2 \delta p}{\partial t^2} - v^2 \nabla^2 \delta p &= \delta v^2 \nabla^2 p \\ \delta p &= 0, t << 0 \end{aligned} \tag{3.4}$$

This equation may be approximated numerically by the same kind of finite difference method explained above for the reference system. The perturbation field is sampled in exactly the same way as the other shot gathers shown here to produce Figure 3.13. If linearization is an accurate approximation, then this field should approximate the difference between the shot gathers produced from the model with reflectors (Figure 3.11 and the one without (Figure 3.5), plotted as Figure 3.12. Indeed the two look quite similar: the difference between the data in Figure 3.12 and FIgure 3.13 is displayed in Figure 3.14, all three figures plotted on the same grey scale.

The events appearing in Figure 3.13 are found in Figure 3.12, which however contains other events as well: careful examination shows that these other events tend to be slower than those retained in Figure 3.13. These slower events are multiple reflections, bouncing between the free surface at the top of the model (a perfect reflector) and the various reflectors evident in Figure 3.7, especially those near the surface.

To understand better the relation between the full model (equation 3.1) and the Born model (equation 3.4), and set the stage for what follows, introduce the *forward map* or *modeling operator* $\mathcal{F}$, mapping the square velocity v^2 to the trace data:

$$\mathcal{F}[v^2] = \{p(\mathbf{x}_r, t; \mathbf{x}_s)\} \tag{3.5}$$

in which $\mathbf{x}_r, \mathbf{x}_s$ are presumed to run over the combinations of receiver and source locations present in the data. Similarly the linearized forward map, or Born modeling operator, $D\mathcal{F}$ is defined by

$$D\mathcal{F}[v^2]\delta v^2 = \{\delta p(\mathbf{x}_r, t; \mathbf{x}_s)\}. \tag{3.6}$$

The Born modeling operator is actually the derivative of the modeling operator, as the notation suggests, in a suitable sense: this is proven in the theoretical part of this course. In fact, the modeling operator is twice differentiable, so long as the pulse is bandlimited as are the sources used here. One would therefore expect that the remainder of the first

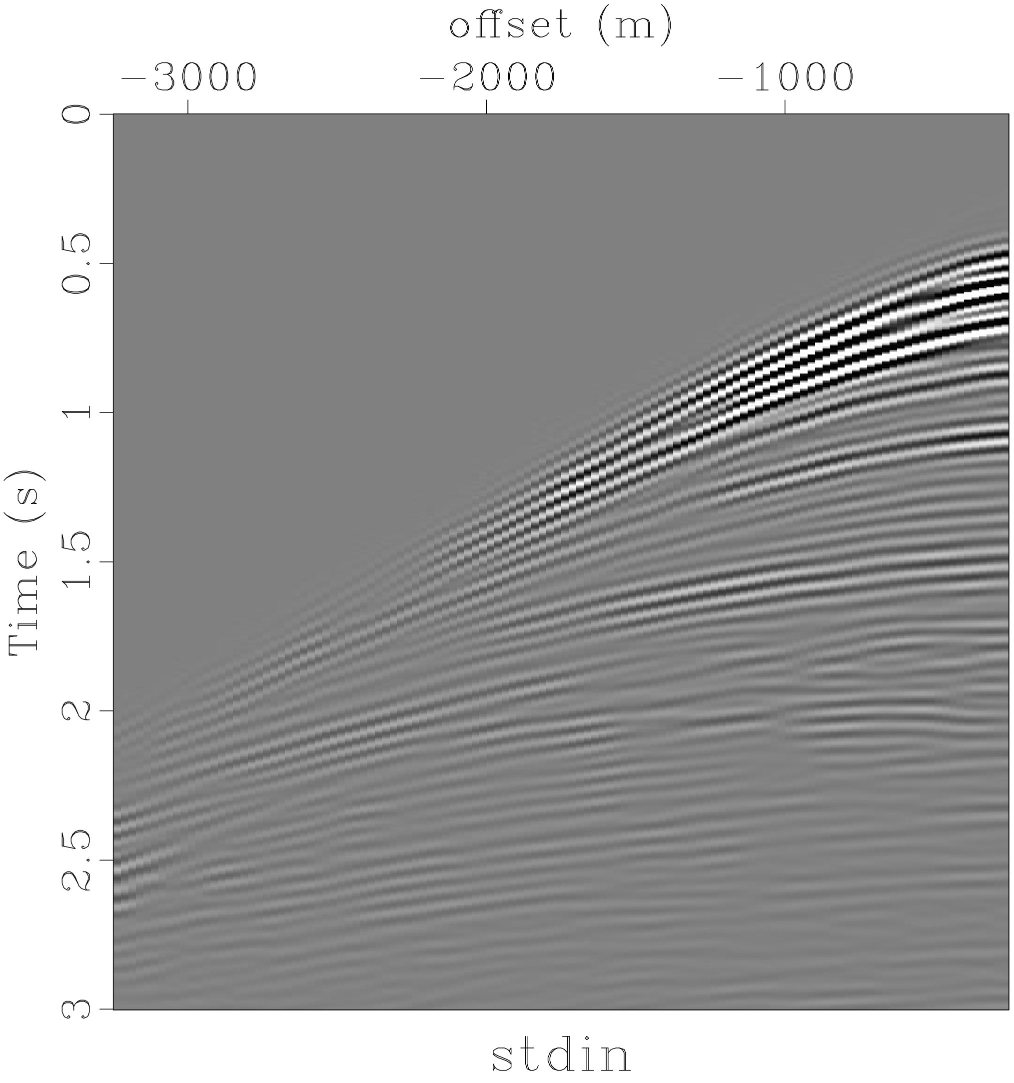

Figure 3.13: Born data: shot gather at shot position $x = 10012m$ with the reference model (v^2) of Figure 3.1 and the perturbation ($\delta(v^2)$) of Figure 3.6, scaled by $\epsilon = 0.15$.

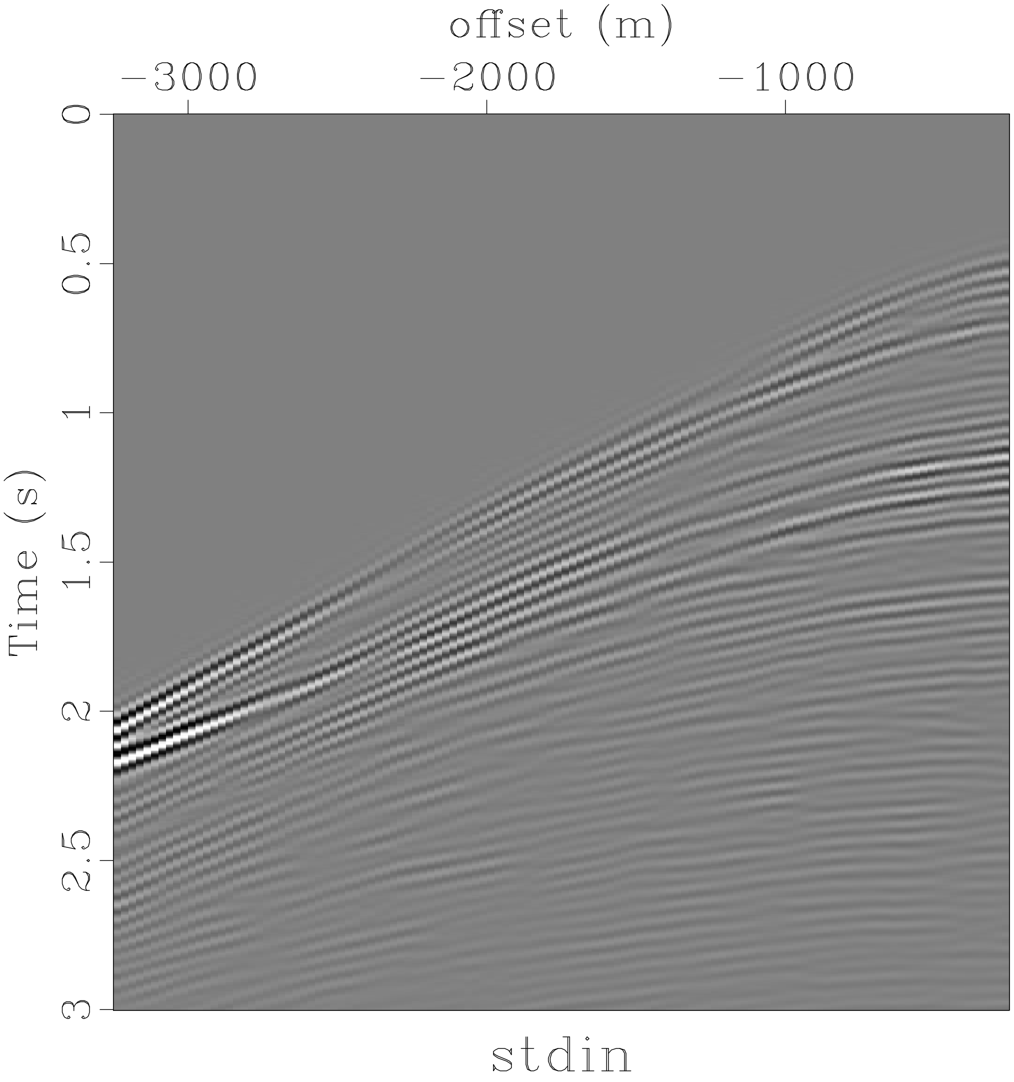

Figure 3.14: Difference between Born shot gather (Figure 3.13) and residual shot gather (Figure 3.12) for perturbation scaled by $\epsilon = 0.15$, all three plotted on same grey scale.

order Taylor series should decrease like the second power of the perturbation length:

$$\mathcal{F}[v^2+\epsilon\delta v^2]-\mathcal{F}[v^2]-\epsilon D\mathcal{F}[v^2]\delta v^2 = O(\epsilon^2). \tag{3.7}$$

Using the common notation $\|\cdot\|$ for the L^2 norm (RMS scaled by cell volume), this would imply

$$\|\mathcal{F}[v^2+\epsilon\delta v^2]-\mathcal{F}[v^2]-D\mathcal{F}[v^2]\delta v^2\| \approx K\epsilon^2 \tag{3.8}$$

for a constant K depending on everything in sight. We can see this relation by measuring the L^2 norms of the Taylor series remainder, which you have already seen, namely as Figure 3.14, for the choices of v^2 (Figure 3.1) and δv^2 (Figure 3.6) introduced above, and $\epsilon = 0.15$. If you replace ϵ by 0.05, then the norm of the first order Taylor series residual should drop by roughly a factor of 9, according to the formula 3.8. In fact the norms are respectively 0.0027 ($\epsilon = 0.15$) and 0.00029 ($\epsilon = 0.05$), and differ almost exactly a factor of 9, as predicted. The residual and Born approximation (linear term in the Taylor series) appear as Figures 3.15 and 3.16 respectively, plotted on the same scale as the remainder in the first order Taylor series in 3.17.

Another important fact to learn about the Born approximation is that its accuracy reflects *scale separation*, and the first order Taylor series remainder is much smaller when the perturbation contains little energy on the scale of a wavelength. This freedom from long-scale components is a feature of the PSPI image 3.6 used as δv^2 in the preceding examples, which accounts for the relative accuracy of the Born approximation. If instead one uses the perturbed model (Figure 3.7) as reference (v^2) and perturbs by a multiple of the smooth background model (Figure 3.1) (δv^2), a much larger first order Taylor remainder results. For example, with $\epsilon = 0.02$ to give a perturbation of the same L^2 size as the PSPI image, but with v^2 and δv^2 as just suggested, the residual shown in Figure 3.18 results, which may be compared to the Born approximation (linearization) shown in Figure 3.19. The larger difference is evident to the eye: the norm of the difference (first order Taylor series remainder) is 0.0089, actually larger than the residual norm of 0.0077. In contrast, the corresponding numbers in the preceding example were 0.0027 and 0.0048. This phenomonon - the more rapid variation of $\mathcal{F}$ in smooth directions - is key to understanding many aspects of the seismic inverse problem. The next section of this course will give some inkling about why it occurs, but the full mathematical story has yet to be told.

3.5 Project Notes

The computational results for this section depend on those of the previous section (Symes, 2017a), and the corresponding project must be built first.

Along with the displays in the paper, the project build (`scons`) creates two propagating wave movies, from which the snapshots in the paper were taken. To view these, use the `xtpen` command in the project directory:

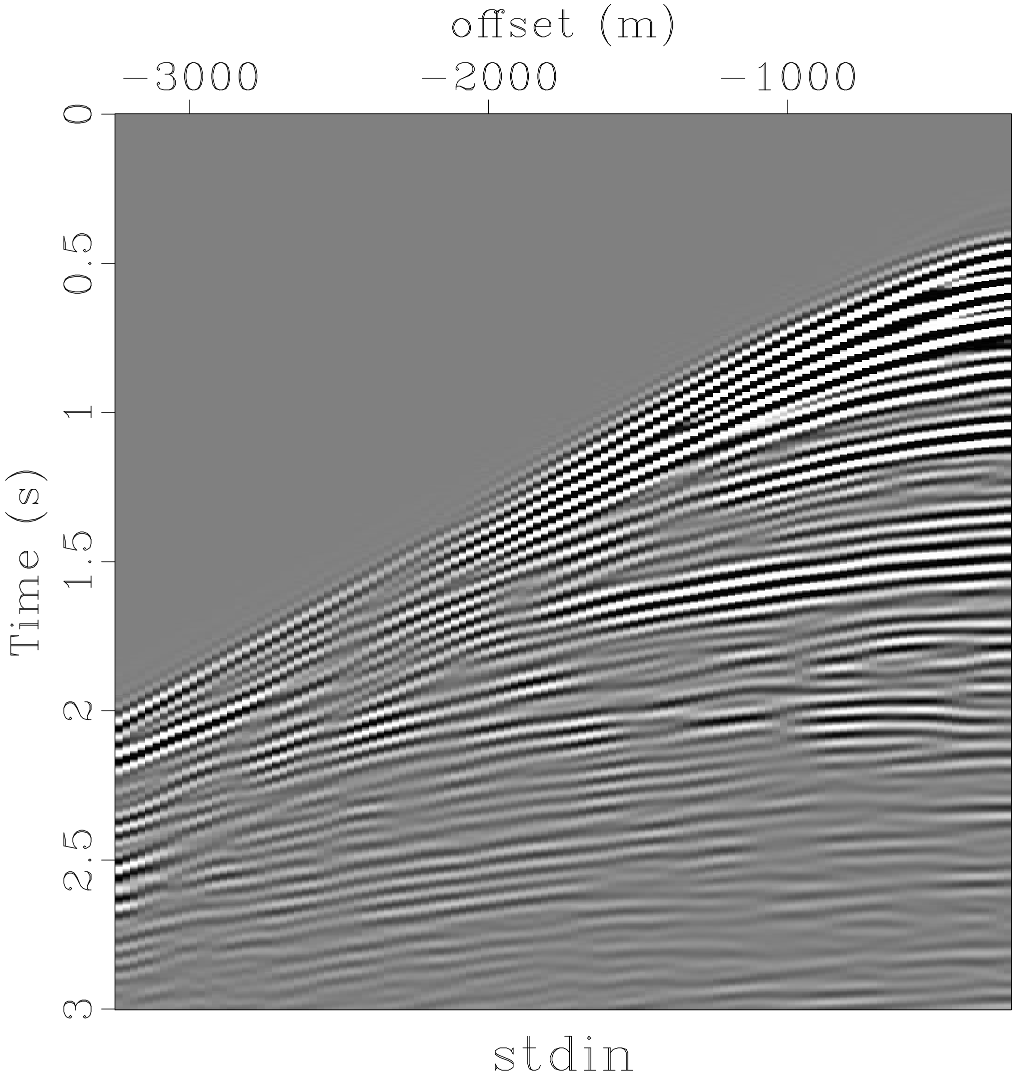

Figure 3.15: Difference between shot gathers for v^2 and $v^2 + \epsilon \delta v^2$, with $\epsilon = 0.05$.

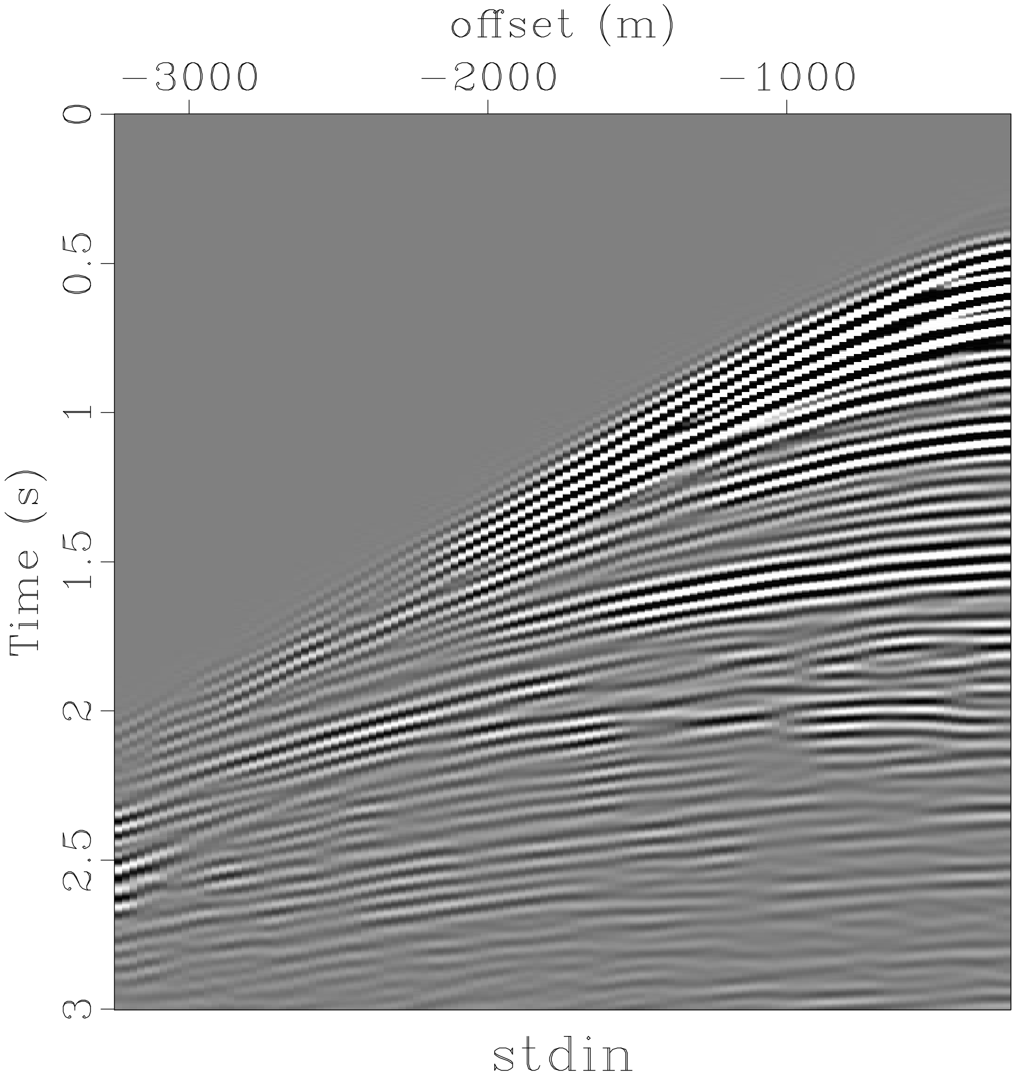

Figure 3.16: Born data: shot gather at shot position $x = 10012m$ with the reference model (v^2) of Figure 3.1 and the perturbation ($\delta(v^2)$) of Figure 3.6, scaled by $\epsilon = 0.05$.

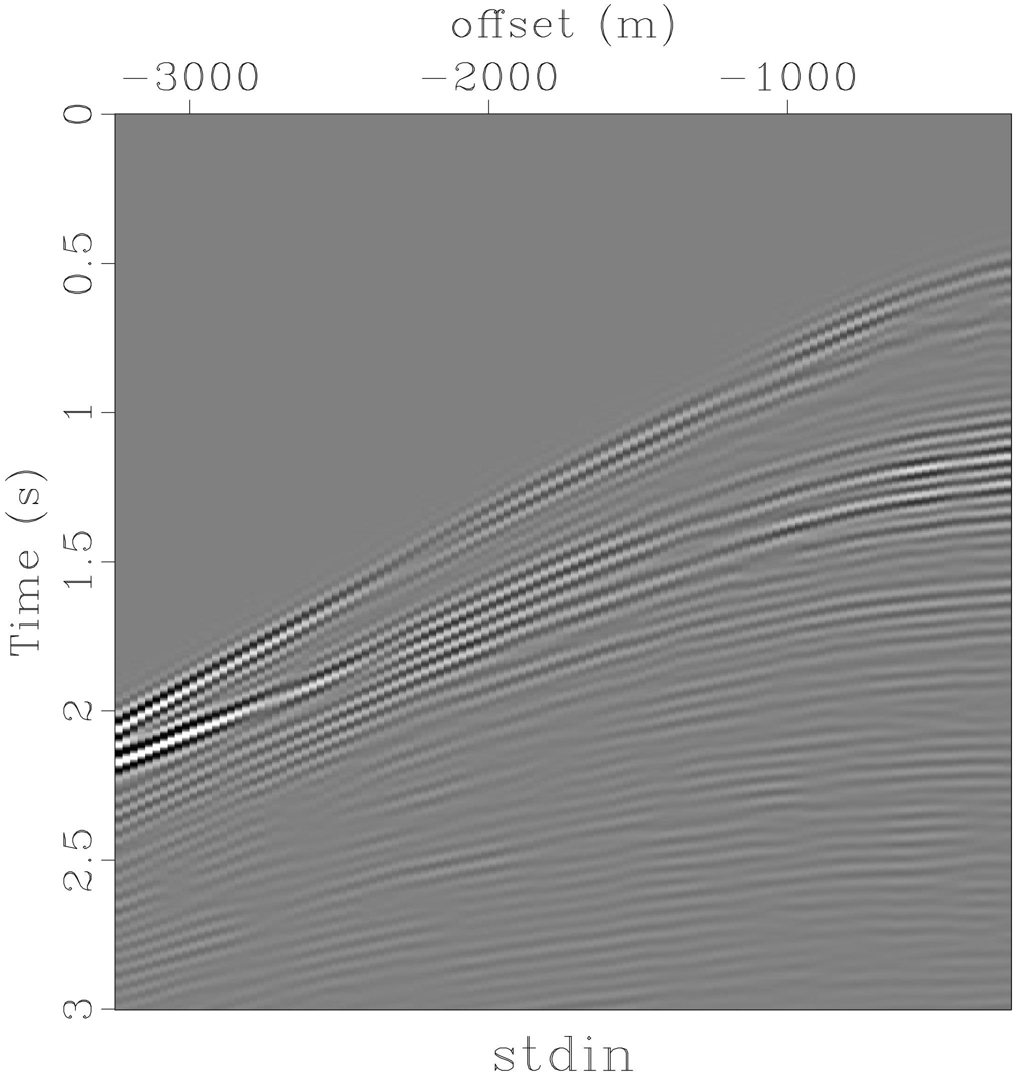

Figure 3.17: Difference between Born shot gather (Figure 3.16) and residual shot gather (Figure 3.15) for perturbation scaled by $\epsilon = 0.05$, all three plotted on same grey scale.

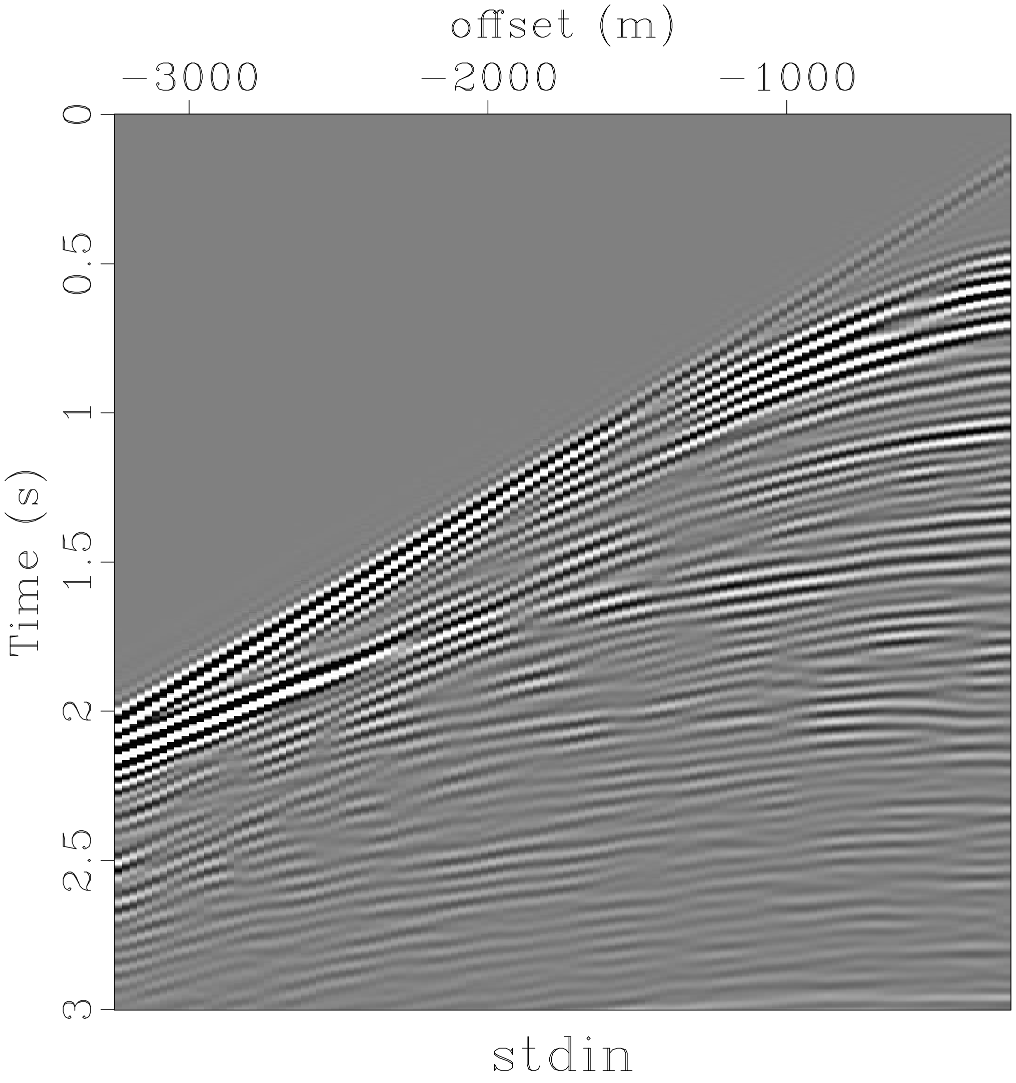

Figure 3.18: Difference between shot gathers for v^2 and $v^2 + \epsilon \delta v^2$, with v^2 = the perturbed model (Figure 3.7), δv^2 = the smooth background model (Figure 3.1), and $\epsilon = 0.02$, chosen to create a perturbation of the same L^2 size as the PSPI image (Figure 3.6).

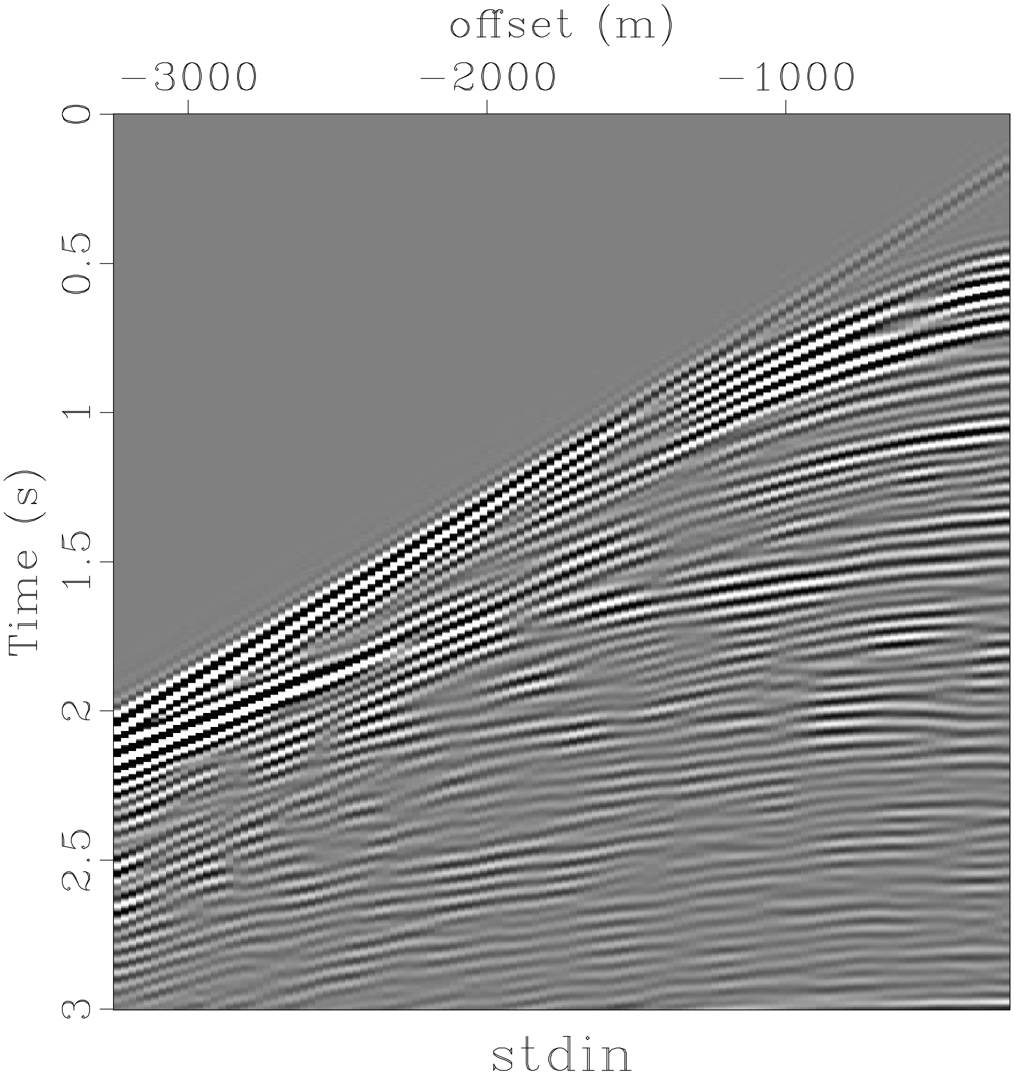

Figure 3.19: Born data: shot gather at shot position $x = 10012m$ with the reference, perturbation, and scale factor ϵ as in Figure 3.18, plotted on same scale.

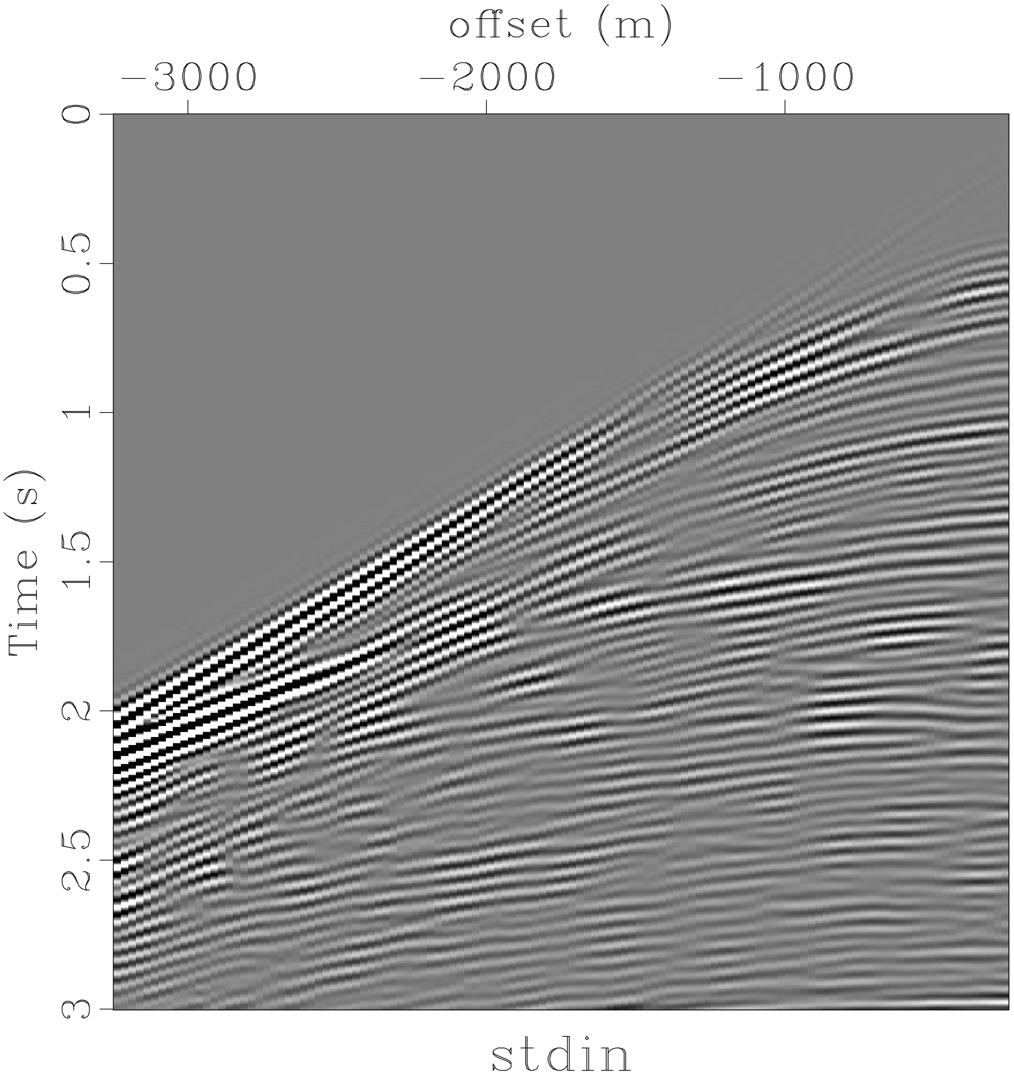

Figure 3.20: Difference between Born shot gather (Figure 3.19) and residual shot gather (Figure 3.18), plotted on same scale.

```
xtpen < Fig/moviereflooverlay.vpl
```

for the radiation solution from a point source in the smooth square velocity model (Figure 3.1), and

```
xtpen < Fig/moviereflopoverlay.vpl
```

for the radiation solution from a point source in the square velocity model with reflectors (Figure 3.7).

The L^2 norms quoted in the text were obtained using `sfattr` - flows are provided in `project/SConstruct` that write the relevant line out to a text file.

3.6 Suggested Projects

1. Just how separated are the scales between the "background" model (Figure 3.1) and reflectivity or perturbation (Figure 3.6)? This question can be explored using the 2D fft implemented in Madagascar.
2. create examples with well-separated scales and investigate the linearization error for perturbations in smooth background vs. rough reflectivity.

REFERENCES

Courant, R., and D. Hilbert, 1962, Methods of mathematical physics: Wiley-Interscience, **II**.

Hu, H., A. Strybulevych, J. H. Page, S. E. Skipetrov, and B. A. van Tiggelen, 2008, Localization of ultrasound in a three-dimensional elastic network: Nature Physics, **4**, 945–948.

Moczo, P., J. O. A. Robertsson, and L. Eisner, 2006, The finite-difference time-domain method for modeling of seismic wave propagation: Advances in Geophysics, **48**, 421–516.

Symes, W., 2017a, Basic processing applied to a North Sea survey: CAAM 641, spring 2017.

———, 2017b, High frequency asymptotics: CAAM 641, spring 2017.

Symes, W., I. Terentyev, and T. Vdovina, 2008, Gridding requirements for accurate finite difference simulation: 78th Annual International Meeting, Expanded Abstracts, Society of Exploration Geophysicists, SM1.5.

Symes, W., and T. Vdovina, 2009, Interface error analysis for numerical wave propagation: Computational Geosciences, **13**, 363–370.

Wright, M., and R. Weaver, eds., 2010, New directions in linear acoustics and vibrations: Cambridge University Press.

Chapter 4

Geometric Optics

4.1 Introduction

The "geometric optics" of the section title refers to a high frequency asymptotic approximation to wave motion. In combination with the Born approximation explained in the last section, this approximation leads to most of the intuition and many of the methods used in practical seismic data processing. Recall that the Born approximation is particularly accurate when the background is smooth and the perturbation oscillatory on the wavelength scale. This is exactly the setting in which geometric optics is effective in approximating the Born approximation.

4.2 Progressing Wave Expansion

Begin once again with the wave equation 3.1, which we repeat here for convenience:

$$\begin{aligned} \frac{\partial^2 p}{\partial t^2} - v^2 \nabla^2 p &= f \qquad (4.1) \\ p &= 0, t << 0 \end{aligned}$$

Denote by $G(\mathbf{x}, t; \mathbf{x}_s)$ the Green's function, that is, the solution of equation 4.1 for $f(\mathbf{x}, t) = \delta(t)\delta(\mathbf{x} - \mathbf{x}_s)$. In view of the linearity of the wave equation and the time-independence of the coefficients, the solution of 4.1 may be expressed as

$$p(\mathbf{x}, t) = \int dt_s \int dx_s \, f(\mathbf{x}_s, t - t_s) G(\mathbf{x}, t_s; \mathbf{x}_s) \qquad (4.2)$$

As noted in the last section, if v is constant, then in 3D

$$G(\mathbf{x}, t; \mathbf{x}_s) = \frac{\delta(t - r/v)}{4\pi r}, \; r = |\mathbf{x} - \mathbf{x}_s| = \sqrt{(\mathbf{x} - \mathbf{x}_s)^T (\mathbf{x} - \mathbf{x}_s)}. \qquad (4.3)$$

whereas in 2D,

$$G(\mathbf{x},t;\mathbf{x}_s) = \frac{1}{\pi v^2}\int_0^{\sqrt{t-r/v}} d\sigma \frac{\delta(t-\sigma^2-r/v)}{\sqrt{\sigma^2+2r/c}}. \tag{4.4}$$

While it is not possible to write such explicit expressions for the fundamental solution in , it is possible to describe the leading singularity of G, assuming smoothness of the wave velocity v (physically, separation of scales). This is accomplished *via* the *progressing wave expansion* (Courant and Hilbert (1962), Ch. VI). Each of the formulas for G above is of the form $a(\mathbf{x},\mathbf{x}_s)S(t-\tau(\mathbf{x},\mathbf{x}_s))$ where a and the travel time function τ are smooth except possibly at $\mathbf{x}=\mathbf{x}_s$, and $S(t)$ is singular at $t=0$. The progressing wave expansion allows the extension of this expression away from $\mathbf{x}=\mathbf{x}_s$, up to a limit signaled by a fundamental change in the nature of the wavefield, and with an error which is *smoother* than S: it takes the form

$$G(\mathbf{x},t;\mathbf{x}_s) = a(\mathbf{x},\mathbf{x}_s)S(t-\tau(\mathbf{x},\mathbf{x}_s)) + R(\mathbf{x},t;\mathbf{x}_s) \tag{4.5}$$

where R is in some sense to be smoother than S, and a and τ are assumed to be smooth in some as-yet unspecified region. The first term clearly expresses wave motion: if it is possible to show that it is truly identifiable as separate (more singular) than the remainder R, then the progressing wave expansion justifies the name of the wave equation.

Applying the wave operator to the right-hand side of equation 4.5, we obtain

$$\begin{aligned}
&\left(\frac{\partial^2}{\partial t^2} - v^2\nabla^2\right)G \\
&= a(1-v^2|\nabla\tau|^2)S''(t-\tau) \\
&\quad - v^2(2\nabla a\cdot\nabla\tau + a\nabla^2\tau)S'(t-\tau) \\
&\quad -v^2\nabla^2 a S(t-\tau) + \frac{\partial^2 R}{\partial t^2} - v^2\nabla^2 R \\
&= \delta(\mathbf{x}-\mathbf{x}_s)\delta(t)\,.
\end{aligned} \tag{4.6}$$

Formally, the terms written in the above order have decreasing orders of singularity, so that if G is to solve the wave equation for $\mathbf{x}\neq\mathbf{x}_s$, each of the coefficients above ought to vanish. Certainly, if

$$|\nabla\tau|^2 = \frac{1}{v^2} \tag{4.7}$$

$$2\nabla a\cdot\nabla\tau + a\nabla^2\tau = 0 \tag{4.8}$$

then the first two terms vanish. The last term vanishes if

$$\delta(\mathbf{x}-\mathbf{x}_s)\delta(t) + v^2\nabla^2 a S(t-\tau) = \frac{\partial^2 R}{\partial t^2} - v^2\nabla^2 R, \tag{4.9}$$

an inhomogeneous wave equation for the remainder R.

Equation 4.7 is the *eikonal equation* of geometric optics (of which the progressing wave expansion is a variant). Inspecting the Green's functions 4.4 and 4.3, evidently $\tau(\mathbf{x},\mathbf{x}_s) = |\mathbf{x}-\mathbf{x}_s|/v$ for constant v, and indeed it is easy to verify that this function satisfies the eikonal equation 4.7. The correct "initial" condition at $\mathbf{x} = \mathbf{x}_s$ forces τ to be asymptotic to the constant v solution there:

$$\tau(\mathbf{x},\mathbf{x}_s)/|\mathbf{x}-\mathbf{x}_s| = v(\mathbf{x}_s) \text{ as } |\mathbf{x}-\mathbf{x}_s| \to 0. \tag{4.10}$$

The second condition, the *transport equation* 4.8, may be rewritten as

$$\nabla \cdot (a^2 \nabla \tau) = 0 \tag{4.11}$$

Reference to the constant-coefficient Green's functions suggests that solutions will need a specified singularity at $\mathbf{x} = \mathbf{x}_s$. This topic will be addressed in the next section.

As stated above, the progressing wave expansion is not much of an expansion. In fact, the decompositionn $G = aS(t-\tau) + R$ is the first term of a series,

$$G(\mathbf{x},t;\mathbf{x}_s) = \sum_{n=0}^{N} a_n(\mathbf{x},\mathbf{x}_s) S_n(t-\tau(\mathbf{x},\mathbf{x}_s) + R_N(\mathbf{x},t;\mathbf{x}_s) \tag{4.12}$$

in which S_k is the kth primitive of S, that is,

$$S_0 = S;\ a_0 = a;\ \frac{dS_k}{dt} = S_{k-1}, k = 1,2,...$$

That is, 4.12 decomposes G into successively less singular terms. The case $N = 1$ will be important in understanding the singularity of G:

$$G(\mathbf{x},t;\mathbf{x}_s) = a_0(\mathbf{x},\mathbf{x}_s)S_0(t-\tau(bx,\mathbf{x}_s) + a_1)\mathbf{x},\mathbf{x}_s)S_1(t-\tau(\mathbf{x},\mathbf{x}_s) + R_1(\mathbf{x},t;\mathbf{x}_s). \tag{4.13}$$

Applying the wave operator, one sees that G solves the wave equation for $\mathbf{x} \neq \mathbf{x}_s$ provided that

$$\begin{aligned} 2\nabla a_1 \cdot \nabla \tau + a_1 \nabla^2 \tau &= \nabla^2 a_0 \\ v^2 \nabla^2 a_1 &= \frac{\partial^2 R_1}{\partial t^2} - v^2 \nabla^2 R_1 \end{aligned} \tag{4.14}$$

4.3 Ray Theory

The *method of characteristics* extrapolates τ away from $\mathbf{x}_s$. To see how this works, suppose first that τ solves the eikonal equation, and let $\mathbf{X}(t)$ be a solution of the system of ordinary differential equations

$$\frac{d\mathbf{X}}{dt} = v^2(\mathbf{X}) \nabla \tau(\mathbf{X}). \tag{4.15}$$

Then

$$\begin{aligned} \frac{d}{dt}\tau(\mathbf{X}(t)) &= \nabla\tau(\mathbf{X}(t))\cdot\frac{d\mathbf{X}}{dt}(t) && (4.16)\\ &= v^2(\mathbf{X}(t))|\nabla\tau(\mathbf{X}(t))|^2 = 1\ . && (4.17)\end{aligned}$$

Therefore we can identify τ with t: if the segment $\{\mathbf{X}(t') : t_0 \le t' \le t\}$ lies entirely in a domain in which τ is defined, then

$$\tau(\mathbf{X}(t)) = \tau(\mathbf{X}(t_0)) + t - t_0 \tag{4.18}$$

Thus from knowledge of the characteristic curves (rays) $\mathbf{X}(t)$, we can construct τ.

Somewhat more surprisingly, it is possible to construct the rays directly, which furnishes a construction of τ as well. Define the vector

$$\mathbf{P}(t) = \nabla\tau(\mathbf{X}(t)) \tag{4.19}$$

with units of time/length, or *slowness*. Note that you can write 4.15 in the form

$$\frac{d\mathbf{X}}{dt} = v^2(\mathbf{X})\nabla\tau(\mathbf{X}) = v^2(\mathbf{X})\mathbf{P}. \tag{4.20}$$

P also solves a differential equation in terms of τ:

$$\begin{aligned} \frac{d\mathbf{P}}{dt}(t) &= (\nabla\nabla\tau)(\mathbf{X}(t))\cdot\dot{\mathbf{X}}(t) && (4.21)\\ &= v^2(\mathbf{X}(t))\nabla\nabla\tau(\mathbf{X}(t))\cdot\nabla\tau(\mathbf{X}(t)) && (4.22)\\ &= \frac{1}{2}v^2(\mathbf{X}(t))\nabla|(\nabla\tau)(\mathbf{X}(t))|^2 && (4.23)\\ &= \frac{1}{2}v^2(x)(\nabla v^{-2})(\mathbf{X}(t)) && (4.24)\\ &= -\frac{1}{2}|\nabla\tau(x)|^2(\nabla v^2)(\mathbf{X}(t)). && (4.25)\end{aligned}$$

Define the *Hamiltonian*

$$H(\mathbf{x},\mathbf{P}) = \frac{1}{2}v^2(\mathbf{x})|\mathbf{P}|^2 \tag{4.26}$$

then the equation 4.20 for **X** and **??** for **P** can be written in the form

$$\begin{aligned} \dot{\mathbf{X}} &= \nabla_{\mathbf{P}}H(\mathbf{X},\mathbf{P}) && (4.27)\\ \dot{\mathbf{P}} &= -\nabla_{\mathbf{x}}H(\mathbf{X},\mathbf{P}) && (4.28)\end{aligned}$$

These are *Hamilton's equations* of classical mechanics, a system of $2n$ autonomous ordinary differential equations. If $(\mathbf{X},\mathbf{P})$ is a solution of the system 4.27, both

$$t \mapsto \mathbf{X}(t)$$

and

$$t \mapsto (\mathbf{X}(t), \mathbf{P}(t))$$

are called rays; in the mathematical literature, the latter trajectory, with values in phase space, is known as a *bicharacteristic strip* or just a *bicharacteristic*.

It's an easy exercise to see that the equations 4.27 implies that H is constant along rays, that is,

$$\frac{d}{dt} H(\mathbf{X}(t), \mathbf{P}(t)) = 0 \tag{4.29}$$

if $(\mathbf{X}, \mathbf{P})$ is a ray. The rays that figure in construction of the traveltime τ are related to τ through equations 4.15 and 4.19, so because of the iekonal equation satisfy $v(\mathbf{X})|\mathbf{P}| = 1$ - so along such rays, $H(\mathbf{X}, \mathbf{P}) = 1/2$. Since $H - 1/2$ generates the same Hamiltonian system, these rays are also called *null bicharacterisitics*.

The rays that play a role in the construction of the Green's function for a source at $\mathbf{x} = \mathbf{x}_s$ are those for which

$$\mathbf{X}(0) = \mathbf{x}_s, \, \mathbf{P}(0) = \mathbf{P}_s, \, v(\mathbf{X}_s)|\mathbf{P}_s| = 1. \tag{4.30}$$

Solving the system 4.27 and initial conditions 4.30 generates a mapping

$$(t, \mathbf{P}_s) \mapsto (\mathbf{X}(t, \mathbf{P}_s), \mathbf{P}(t, \mathbf{P}_s)) \tag{4.31}$$

called the ray (polar) coordinate system centered at $\mathbf{x}_s$. For sufficiently small $t > 0$, the Jacobian of the map 4.31 is nonsingular for any $\mathbf{P}_s$ with $v(\mathbf{x}_s)|\mathbf{P}_s| = 1$ (exercise!). It follows that ray polar coordinates define an invertible differentiable coordinate system on an open set $\Omega(\mathbf{x}_s)$ containing $\mathbf{x}_s$. For constant v, of course, $\Omega(\mathbf{x}_s)$ consists of all of Euclidean space, and

$$\begin{aligned} \mathbf{X}(t, \mathbf{P}_s) &= \mathbf{x}_s + t v(\mathbf{x}_s)^2 \mathbf{P}_s, && (4.32) \\ \mathbf{P}(t, \mathbf{P}_s)) &= \mathbf{P}_s. && (4.33) \end{aligned}$$

Given a ray coordinate system centered at $\mathbf{x}_s$ and valid in the set $\Omega(\mathbf{x}_s)$, define

$$\tau(\mathbf{x}, \mathbf{x}_s) = t \text{ for } \mathbf{x} = \mathbf{X}(t, \mathbf{P}_s). \tag{4.34}$$

That is, $\tau(\mathbf{x}, \mathbf{x}_s)$ is the "radius" in the polar coordinate system.

The first step in showing that τ, so defined, satisfies the eikonal equation in $\Omega(\mathbf{x}_s)$ is the proof of the

Claim The ray tangent vector $d\mathbf{X}/dt$ at $\mathbf{X}(t, \mathbf{P}_s)$ is perpindicular to the surface (isochron) $\{\mathbf{x} = \mathbf{X}(t, \mathbf{P}_s) : \text{ all } \mathbf{P}_s \text{ such that } v(\mathbf{x}_s)|\mathbf{P}_s| = 1\}$.

That is, for any vector $\delta\mathbf{P}_s \perp \mathbf{P}_s$,

$$(\nabla_{\mathbf{P}_s}\mathbf{X}(t,\mathbf{P}_s)\delta\mathbf{P}_s)\cdot d\mathbf{X}(t,\mathbf{P}_s)/dt = 0. \tag{4.35}$$

since all of the tangent vectors to the isochron take the form of the vector on the left. To see this, differentiate the left hand side of 4.35:

$$\frac{d}{dt}\left((\nabla_{\mathbf{P}_s}\mathbf{X}\delta\mathbf{P}_s)\cdot\frac{d\mathbf{X}}{dt}\right)$$

$$= \left(\nabla_{\mathbf{P}_s}\frac{d\mathbf{X}}{dt}\delta\mathbf{P}_s\right)\cdot\frac{d\mathbf{X}}{dt} + (\nabla_{\mathbf{P}_s}\mathbf{X}\delta\mathbf{P}_s)\cdot\frac{d^2\mathbf{X}}{dt^2} \tag{4.36}$$

Differentiation of the first equation in 4.27 and using both equations and the constant-H condition 4.29 yields the second-order ray equation, useful in its own right:

$$\frac{d^2\mathbf{X}}{dt^2} = 2\left(\frac{\nabla v}{v}\cdot\frac{d\mathbf{X}}{dt}\right)\frac{d\mathbf{X}}{dt} - v\nabla v \tag{4.37}$$

The first term in 4.35 may be rewritten as

$$\frac{1}{2}\left(\nabla_{\mathbf{P}_s}\left|\frac{d\mathbf{X}}{dt}\right|^2\right)\delta\mathbf{P}_s = \frac{1}{2}\nabla_{\mathbf{P}_s}v^2\delta\mathbf{P}_s$$

$$= v\nabla v\cdot\nabla_{\mathbf{P}_s}\mathbf{X}\delta\mathbf{P}_s. \tag{4.38}$$

Re-write equation 4.36 using the second-order ODE 4.37 in the first term and the identigy 4.38 in the second: net result is that the second term cancels, leaving

$$\frac{d}{dt}\left(\nabla_{\mathbf{P}_s}\mathbf{X}\delta\mathbf{P}_s\right)\cdot\frac{d\mathbf{X}}{dt} = 2\left(\frac{\nabla v}{v}\cdot\frac{d\mathbf{X}}{dt}\right)\frac{d\mathbf{X}}{dt}\cdot(\nabla_{\mathbf{P}_s}\mathbf{X}\delta\mathbf{P}_s) \tag{4.39}$$

This is a homogeneous linear ordinary differential equation for the quantity

$$\frac{d\mathbf{X}}{dt}\cdot(\nabla_{\mathbf{P}_s}\mathbf{X}\delta\mathbf{P}_s). \tag{4.40}$$

Rays are asymptotic to the frozen-coefficient trajectories specified in 4.32 as $t \to 0$, and for these, the dot product in 4.40 is zero identically - at the origin of the polar coordinates, the ray direction is perpindicular to the directions of constant $|\mathbf{P}_s|$. Therefore the effective initial condition for the ODE 4.39 is zero, and so the inner produce in 4.40 vanishes identically, establishing the Claim.

Choose any basis for the $d-1$-dimensional subspace of $\mathbf{R}^d$ perpindicular to $\mathbf{P}_s$ (reminder - d = space dimension). Think of $\mathbf{X}(t,\mathbf{P}_s$ as a function of t and the $d-1$ coefficients of this basis, which generate a vector perpindicular to $\mathbf{P}_s$. For a coefficient vector near the origin, this vector added to $\mathbf{P}_s$ projects uniquely onto the sphere of radius $1/v(\mathbf{x}_s)$. A restatement of the Claim is that the first column ($d\mathbf{X}/dt$) of the Jacobian of $(t, \mathbf{P}_s) \rightarrow \mathbf{X}$ is orthogonal to to the other columns. Therefore the first row of the inverse, which by the chain rule is $\nabla\tau(\mathbf{x}, \mathbf{x}_s)$, has length reciprocal to that of $d\mathbf{X}/dt$. Since the latter has length v, the eikonal equation follows.

The identification $\tau(\mathbf{x}, \mathbf{x}_s) = t$ when $\mathbf{x} = \mathbf{X}(t, \mathbf{P}_s)$ shows that τ is the time of arrival of the ray from $\mathbf{x}_s$ to $\mathbf{x}$. So τ is often called the travel- or arrival-time.

Note that this "ray-tracing" construction of τ works only in regions $\Omega(\mathbf{x}_s)$ within which ray polar coordinates are valid. The loss of validity comes when rays with two different values of takeoff slowness $\mathbf{P}_s$ intersect - that is, when two or more rays pass over both $\mathbf{x}_s$ and $\mathbf{x}$. If such *conjugate points* exist, then it is not hard to see that envelopes of ray families (defined by ranges of $\mathbf{P}_s$) must exist; these envelopes are known as *caustics*. Analysis of progressing waves near such caustics is somewhat harder than for the non-caustic case. The developments of this and the next few sections assumes that all of the important energy in the wavefield may be explained with rays that have not yet touched a caustic in their travel from the source point $\mathbf{x}_s$, that is, the analysis will be confined to the region $\Omega(\mathbf{x}_s)$ of *simple ray geometry*.

The Hamiltonian system 4.27 is the basis for the most common approach to computing τ, by solving the initial value problem for this system of ODEs. The main difficulty to be overcome in this approach is that the rays do not generally pass over grid points, so a gridded version of τ is not an immediate by-product. Many clever workarounds have been proposed (Chapman, 1985; Vinje et al., 1993).

It is also possible to treat τ as a solution of the eikonal equation 4.7, as an initial value problem for a partial differential equation. However τ appears to be no longer a function if one uses the ray construction above, outside of the region $\Omega(\mathbf{x}_s)$ in which each point is connected to $\mathbf{x}_s$ by a unique ray. One approach constructs a function that is equal to τ in $\Omega(\mathbf{x}_s)$ and elsewhere is the least value of t along all of the rays connecting to $\mathbf{x}_s$. This first-arrival-time construction may be approximated by solving special gridded finite difference or finite element discretizations of the eikonal equation (Vidale (1988), van Trier and Symes (1991)). For recent viewpoint and many references see (Luo et al., 2014). It has also been possible to extend this "Eulerian" point of view, via appropriately posed PDE problems, to computing global (multi-valued) travel time fields (Fomel and Sethian, 2002; Qian et al., 2003).

A simple 2D example exhibiting many of the features just described appears in Figure 4.1, showing a velocity field ($v(\mathbf{x})$) with a slow (low value) anomaly below the source point $\mathbf{x}_s$ at the top of the figure. Figure 4.2 shows both rays $t \mapsto \mathbf{X}(t, \mathbf{P}_s)$, and wavefronts τ = constant. The rays and wavefronts are perpindicular: this is precisely equation 4.35, established above. The rays bend towards the lower velocity region in the center, as follows directly from the 2nd order equation 4.37. The envelope of rays (caustic) are clearly visible. In the region beyond the caustics, three rays pass over each point.

4.4 Amplitudes

Having computed τ, it is easy to compute a. Indeed, the em transport equation 4.8 may be re-written as an ordinary differential equation along rays

$$\frac{d}{dt}a(\mathbf{X}(t)) - b(\mathbf{X}(t))a(\mathbf{X}(t)) = 0$$

where

$$b = \frac{1}{2}\nabla^2\tau.$$

Thus a may be computed by quadrature along the ray family associated with τ. Initial values (for small t) for $a(\mathbf{X}(t))$ cannot be read off directly for $t = 0$, since the amplitude is singular at $\mathbf{x} = \mathbf{x}_s$, but must be inferred from the $t \rightarrow 0$ asymptotics.

Rather than go through that exercise, note that the solution of the transport equation has a nice geometric interpretation. Construct a "ray tube" R (note - this is not the remainder term R, defined above and discussed in the next section!), whose lateral boundary surface B is swept out by rays, and whose ends are surfaces l and L on which $\tau =$ constant: $\partial R = B \cup l \cup L$. A 2D schematic appears in Figure 4.3.

Now integrate both sides of the re-written transport equation **??** over R. The divergence theorem implies that

$$\int_R \nabla \cdot (a^2 \nabla \tau) = \int_{l \cup L} a^2 \nabla \tau \cdot \mathbf{n} dA = 0 \tag{4.41}$$

where $\mathbf{n}$ is the outward normal vector on $\partial\Gamma$. The integral over the lateral boundary B vanishes as the normal is perpindicular to the rays that form it, hence to $\nabla\tau$. A second consequence of the orthogonality 4.35 is that $\nabla\tau = v^{-1}\mathbf{n}$ on L, $= -v^{-1}\mathbf{n}$ on l so the identity 4.41 implies that

$$\int_l a^2 v^{-1} = \int_L a^2 v^{-1}. \tag{4.42}$$

Shrinking the ray tube to a single central ray, one obtains in the limit

$$\frac{a^2(\mathbf{X}(t_0))}{v(\mathbf{X}(t_0))} = J(t_0, t_1)\frac{a^2(\mathbf{X}(t_1))}{v(\mathbf{X}(t_1))} \tag{4.43}$$

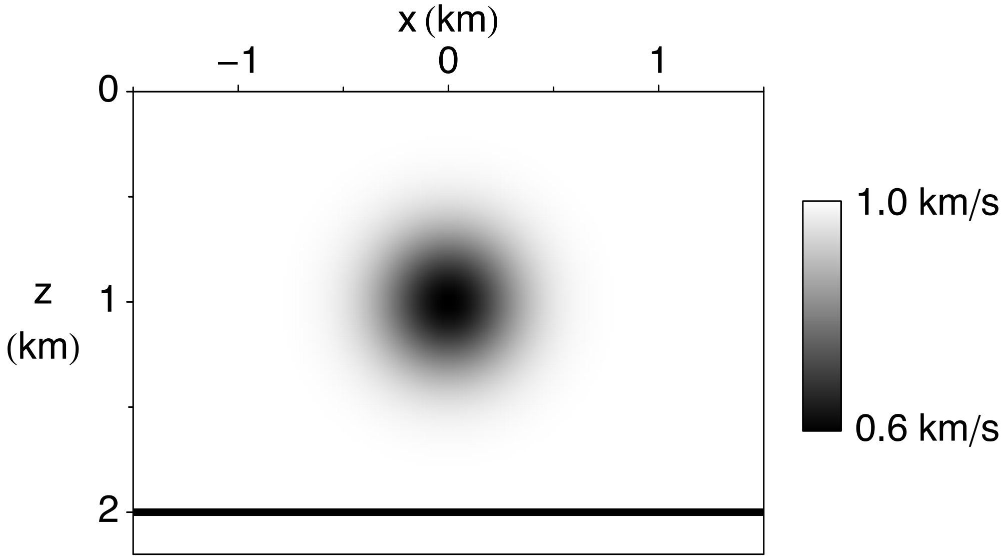

Figure 4.1: Low velocity lens in constant velocity (2 km/s) background. Dark bar at depth = 2 km is a reflector, used in later examples but not here. Borrowed from Stolk and Symes (2004).

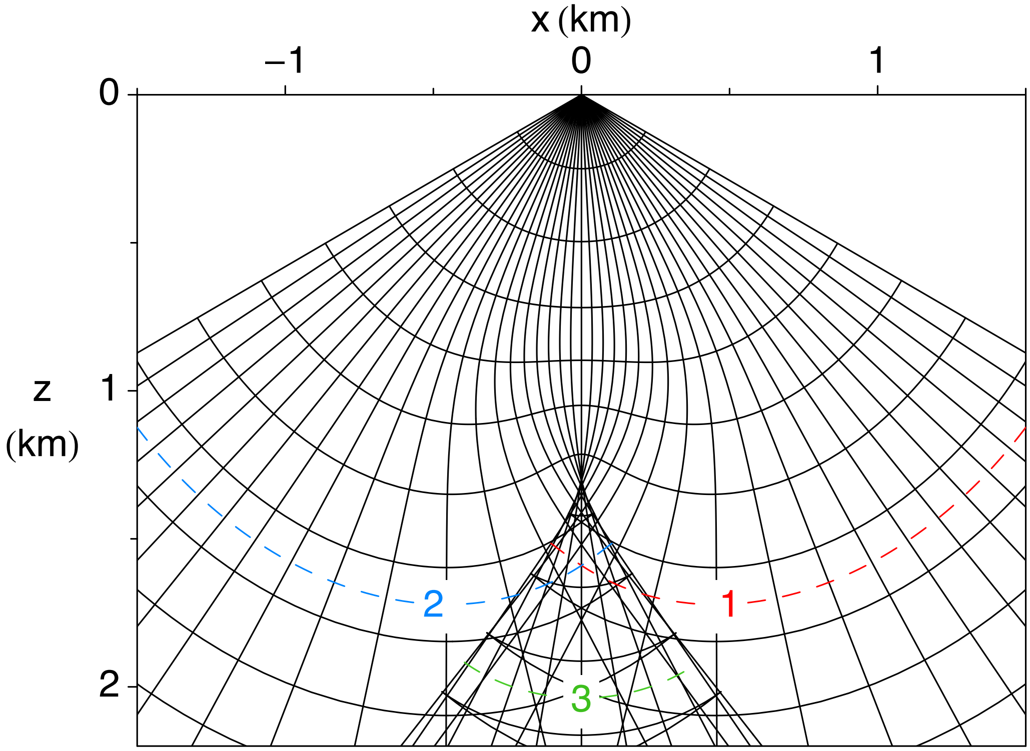

Figure 4.2: Rays and wavefronts emanating from a source point above the center of the lens. Note that rays are perpindicular to wavefronts, consistent with the theory developed in the text, and that the bending of rays towards low velocity leads to the development of a caustic (ray envelope), which itself has a cusp singularity just below the lens. This "butterfly" caustic joins two "fold" caustic branches at the cusp. Below the caustic, each point is joined to the source by three rays: one (with greater traveltime) passing through the slow lens, and two passing around the lens through the faster background medium. Borrowed from Stolk and Symes (2004).

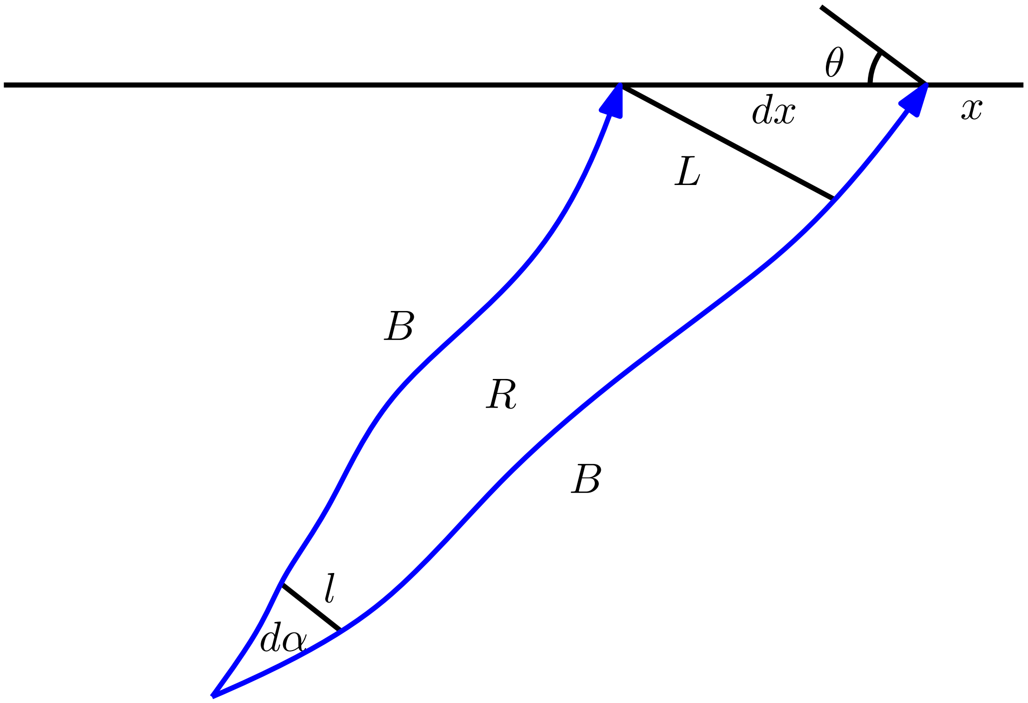

Figure 4.3: Sketch of “ray tube” construction. The segment “dx” is a projection onto the horizontal surface - it is useful in understanding prestack amplitudes, discussed in a later section. Borrowed from Hou and Symes (2015).

in which J is the transverse Jacobian determinant, that is, the Jacobian determinant of the map from l to L at the central ray intersection, defined by tracing rays from l to L. This determinant is the limit of the volume ratio $|L|/|l|$ as the tube shrinks to the central ray.

The Jacobian and its determinant satisfy evolution equations along rays - solving these equations is *dynamic ray tracing*, see Cerveny (1985). Equation 4.43 implies that this construction fails when the Jacobian becomes singular, which happens precisely at the intersection of the central ray with a caustic (envelope of rays), and at the source point itself (a sort of degenerate caustic). In principle, the amplitude grows without bound near such points.

The actual behaviour of the wavefield near caustics can't be inferred from the progressing wave expansion in the form presented so far: extensions are required, discussed in a later section. However, The analysis of the remainder term in the next section requires an explicit description of the behaviour of $a(\mathbf{x},\mathbf{x}_s)$ as $\mathbf{x} \rightarrow \mathbf{x}_s$.

The 3D case is slightly simpler. Introduce polar coordinates for the ray field $\mathbf{X}$ centered at $\mathbf{x}_s$:

$$\mathbf{X} = \mathbf{x}_s + \left(\begin{array}{c} r\cos\theta\cos\phi \\ r\sin\theta\cos\phi \\ r\sin\phi \end{array} \right), \; r > 0, -\pi \le \theta \le \pi, -\frac{\pi}{2} \le \phi \le \frac{\pi}{2}$$

and parametrize the initial slowness $\mathbf{P}_s$ in terms of initial azimuth and polar angles θ_s,ϕ_s:

$$\mathbf{P}_s(\theta_s,\phi_s) = \frac{1}{v(\mathbf{x}_s)} \left(\begin{array}{c} \cos\theta_s\cos\phi_s \\ \sin\theta_s\cos\phi_s \\ \sin\phi_s \end{array} \right).$$

Denote by $\mathbf{P}_s^0 = \mathbf{P}_s(\theta_s^0,\phi_s^0)$ the initial slowness for the central ray of the ray tube construction, and define the tube to have distorted square cross section, of radius $\delta > 0$:

$$R = \{\mathbf{X}(t,\theta_s,\phi_s) : 0 < t_0 \le t \le t_1, |\theta_s - \theta_s^0| \le \delta, |\phi_s - \phi_s^0| < \delta\}.$$

Then

$$\int_l a^2 v^{-1} = \int_L a^2 v^{-1}. \tag{4.44}$$

$$\begin{aligned} \int_l a^2 v^{-1} &= \int_{\theta_s^0-\delta}^{\theta_s^0+\delta} d\theta_s \int_{\phi_s^0-\delta}^{\phi_s^0+\delta} d\phi_s \, \mu(t_0,\theta_s,\phi_s)(a^2 v^{-1})(\mathbf{X}(t_0,\theta_s,\phi_s)) d\theta_s d\phi_s \\ \int_L a^2 v^{-1} &= \int_{\theta_s^0-\delta}^{\theta_s^0+\delta} d\theta_s \int_{\phi_s^0-\delta}^{\phi_s^0+\delta} d\phi_s \, \mu(t_1,\theta_s,\phi_s)(a^2 v^{-1})(\mathbf{X}(t_0,\theta_s,\phi_s)) d\theta_s d\phi_s \end{aligned} \tag{4.45}$$

The surface area density $\mu(t,\theta_s,\phi_s)$ satisfies

$$\mu(t,\theta_s,\phi_s) d\theta_s \wedge d\phi_s = \mathbf{X}(t,\theta_s,\phi_s)^* dS_t \tag{4.46}$$

in which dS_t is the area element on the isochron surface $S_t = \{\mathbf{x} : \tau(\mathbf{x}, \mathbf{x}_s) = t\}$. By definition, dS_t satisfies

$$dS_t(\mathbf{V}_1, \mathbf{V}_2) = 1 \tag{4.47}$$

for any vector fields $\mathbf{V}_1, \mathbf{V}_2$ on S_t that form an orthonormal oriented basis at every point. The orientation of S_t is given by choosing $v^{-1}\nabla\tau$ to be the unit normal vector.

Note that

$$\frac{\partial \mathbf{X}}{\partial t} = \frac{\partial r}{\partial t} e_r + r\cos\phi \frac{\partial \theta}{\partial t} e_\theta + r\frac{\partial \phi}{\partial t} e_\phi \tag{4.48}$$

where

$$e_r = \begin{pmatrix} \cos\theta\cos\phi \\ \sin\theta\cos\phi \\ \sin\phi \end{pmatrix},\ e_\theta = \begin{pmatrix} -\sin\theta \\ \cos\theta \\ 0 \end{pmatrix},\ e_\phi = \begin{pmatrix} -\cos\theta\sin\phi \\ -\sin\theta\sin\phi \\ \cos\phi \end{pmatrix},$$

form an orthonormal basis of $\mathbf{R}^3$. Similarly,

$$\frac{\partial \mathbf{X}}{\partial \theta_s} = \frac{\partial r}{\partial \theta_s} e_r + r\cos\phi \frac{\partial \theta}{\partial \theta_s} e_\theta + r\frac{\partial \phi}{\partial \theta_s} e_\phi \tag{4.49}$$

$$\frac{\partial \mathbf{X}}{\partial \phi_s} = \frac{\partial r}{\partial \phi_s} e_r + r\cos\phi \frac{\partial \theta}{\partial \phi_s} e_\theta + r\frac{\partial \phi}{\partial \phi_s} e_\phi \tag{4.50}$$

Suppose $\mathbf{W_1}, \mathbf{W_2}$ are arbitrary vector fields on S_t: orthogonalizing them and using the algebraic properties of the 2-form dS_t, it follows from the volume element property 4.47 that

$$dS_t(\mathbf{W}_1, \mathbf{W}_2) = \sqrt{|\mathbf{W}_1|^2|\mathbf{W}_2|^2 - \langle \mathbf{W}_1, \mathbf{W}_2\rangle^2}$$

So

$$\begin{aligned}\mu_t &= (\mu_t d\theta_s \wedge d\phi_s)\left(\frac{\partial}{\partial\theta_s}, \frac{\partial}{\partial\phi_s}\right) \\ &= \mathbf{X}^* dS_t\left(\frac{\partial}{\partial\theta_s}, \frac{\partial}{\partial\phi_s}\right) = dS_t\left(\frac{\partial \mathbf{X}}{\partial\theta_s}, \frac{\partial \mathbf{X}}{\partial\phi_s}\right) \\ &= \left(\left|\frac{\partial \mathbf{X}}{\partial\theta_s}\right|^2 \left|\frac{\partial \mathbf{X}}{\partial\phi_s}\right|^2 - \left\langle \frac{\partial \mathbf{X}}{\partial\theta_s}, \frac{\partial \mathbf{X}}{\partial\phi_s}\right\rangle^2\right)^{\frac{1}{2}}\end{aligned}$$

Using the orthonormality of e_r, e_θ, and e_ϕ and the chain rules 4.49 and 4.50, see that this is

$$= \left\{\left[\left(\frac{\partial r}{\partial\theta_s}\right)^2 + r^2\left(\cos^2\phi\left(\frac{\partial\theta}{\partial\theta_s}\right)^2 + \left(\frac{\partial\phi}{\partial\theta_s}\right)^2\right)\right]\left[\left(\frac{\partial r}{\partial\phi_s}\right)^2 + r^2\left(\cos^2\phi\left(\frac{\partial\theta}{\partial\phi_s}\right)^2 + \left(\frac{\partial\phi}{\partial\phi_s}\right)^2\right)\right]\right.$$

$$-\left[\left(\frac{\partial r}{\partial \theta_s}\right)\left(\frac{\partial r}{\partial \phi_s}\right)+r^2\left(\cos^2\phi\left(\frac{\partial \theta}{\partial \theta_s}\right)\left(\frac{\partial \theta}{\partial \phi_s}\right)+\left(\frac{\partial \phi}{\partial \theta_s}\right)\left(\frac{\partial \phi}{\partial \phi_s}\right)\right)\right]^2\Bigg\}^{\frac{1}{2}} \tag{4.51}$$

So asymptotic behaviour of the density μ as $t \to 0$ follow from the asymptotics of r, θ, and ϕ and their angular derivatives. I have implicitly assumed that the velocity field v has as many derivatives as are convenient; for the following argument, three continuous derivatives are certainly sufficient. Then the solutions of Hamilton's equations 4.27 are several times differentiable as functions of t and of the initial conditions, that is, of the azimuth and polar angles, so 4.30 imply that

$$\mathbf{X}(t,\theta_s,\phi_s) = \mathbf{x}_s + tv(\mathbf{x}_s)^2\mathbf{P}_s(\theta_s,\phi_s) + t^2\mathbf{Y}(t,\theta_s,\phi_s)$$

in which $\mathbf{Y}$ is smooth in $t \geq 0, \theta_s, \phi_s$. Therefore

$$r(t,\theta_s,\phi_s) = v(x_s)t + t^2 y(t,\theta_s,\phi_s), \tag{4.52}$$

where y is several times differentiable for $t \geq 0$ and real-valued. This equation in turn implies that for a non-vanishing range of times $0 \leq t < t^*$, there are constants $0 < v_* \leq v^*$ so that

$$v_* t \leq r(t,\theta_s,\phi_s) \leq v^* t, \tag{4.53}$$

uniformly in θ_s, ϕ_s.

Under these conditions, the map $(t_0,\theta_s,\phi_s) \mapsto (r,\theta,\phi)$ is a C^k diffeomorphism on $[0,t^*)\times(-\pi,\pi)\times(-\pi/2,\pi/2)$, into the slit $\mathbf{R}^3$. Of course part of $\mathbf{R}^3$ is left out, but by using several versions of polar coordinates with different polar axes, it is possible to cover all of $\mathbf{R}^3$ except the origin. For all of these rotated coordinates, the coordinate change is of class C^k up to $t_0 = 0$, so derivatives are bounded uniformly on compact subsets of the domain. I will assume that derivatives of order at least two are continuous.

Equation 4.52 also implies that

$$\begin{aligned} \frac{\partial r}{\partial \theta_s} &= O(t^2) \\ \frac{\partial r}{\partial \phi_s} &= O(t^2) \end{aligned} \tag{4.54}$$

Together with equation 4.53, this shows that

$$\begin{aligned} \frac{1}{r}\frac{\partial r}{\partial \theta_s} &= O(t) \\ \frac{1}{r}\frac{\partial r}{\partial \phi_s} &= O(t) \end{aligned} \tag{4.55}$$

A similar examination of the Hamilton equation for the slowness yields the conclusion that

$$\begin{aligned} \theta &= \theta_s + O(t) \\ \phi &= \phi_s + O(t) \\ \frac{\partial\theta}{\partial\theta_s} &= 1 + O(t) \\ \frac{\partial\theta}{\partial\phi_s} &= O(t) \\ \frac{\partial\phi}{\partial\theta_s} &= O(t) \\ \frac{\partial\phi}{\partial\phi_s} &= 1 + O(t) \end{aligned} \tag{4.56}$$

In all cases, the remainders hiding inside the big Os are several times differentiable in $t \geq 0, \theta_s, \phi_s$.

Putting together equations 4.51, 4.55, and 4.56, conclude that

$$\mu((t,\theta_s,\phi_s) = r^2 \cos\phi_s(1 + t^2 \nu(t,\theta_s,\phi_s)), \tag{4.57}$$

with ν of class $C^k([0,t^*) \times (-\pi,\pi) \times (-\pi/2,\pi/2))$, $k \geq 2$.

That is, μ is to leading order the same as the Euclidean area element on the sphere of radius r. So the relations 4.45 imply

$$\int_{\theta_s^0-\delta}^{\theta_s^0+\delta} d\theta_s \int_{\phi_s^0-\delta}^{\phi_s^0+\delta} d\phi_s\, r(t_0,\theta_s,\phi_s)^2 \cos\phi_s(1 + t_0\nu(t_0,\theta_s,\phi_s))(a^2v^{-1})(\mathbf{X}(t_0,\theta_s,\phi_s))$$

$$= \int_{\theta_s^0-\delta}^{\theta_s^0+\delta} d\theta_s \int_{\phi_s^0-\delta}^{\phi_s^0+\delta} d\phi_s\, r(t_1,\theta_s,\phi_s)^2 \cos\phi_s(1 + t_1\nu(t_1,\theta_s,\phi_s))(a^2v^{-1})(\mathbf{X}(t_1,\theta_s,\phi_s)) \tag{4.58}$$

Dividing by $4\delta^2$ and letting $\delta \to 0$, obtain

$$r(t_0,\theta_s,\phi_s)^2 \cos\phi_s(1+t_0^2\nu(t_0,\theta_s,\phi_s))(a^2v^{-1})(\mathbf{X}(t_0,\theta_s,\phi_s)) = r(t_1,\theta_s,\phi_s)^2 \cos\phi_s(1+t_1^2\nu(t_1,\theta_s,\phi_s))(a^2v^{-1})(\mathbf{X}(t_1,\theta_s,\phi_s)) \tag{4.59}$$

Changing variables to r,θ,ϕ and discarding the factor $\cos\phi_s$,

$$r_0^2(1 + \tau(r_0,\theta,\phi)^2\nu(r_0,\theta,\phi))(a^2v^{-1})(r_0,\theta,\phi) = r_1^2(1 + \tau(r_1,\theta,\phi)^2\nu(r_1,\theta,\phi))(a^2v^{-1})(r_1,\theta,\phi) \tag{4.60}$$

Divide both sides of 4.58 by $v^{-1}(r_0,\theta,\phi)1+\tau(r_0,\theta,\phi)^2\nu(r_0,\theta,\phi)$ and denote by $\gamma(\theta,\phi)$ the limit of the modiffiedLHS as $r_0 \to 0$:

$$\gamma(\theta,\phi) = r_1^2(1+\tau(r_1,\theta,\phi)^2\nu(r_1,\theta,\phi))(a^2v(\mathbf{x}_s)v^{-1})(r_1,\theta,\phi) \tag{4.61}$$

Since the RHS is > 0, so is γ. Since the RHS is actually the same as the LHS of 4.60 with $r_0 \leftrightarrow r_1$, there is $r^* > 0$ so that

$$a(r,\theta,\phi)\sqrt{\frac{v(\mathbf{x}_s)}{v(r,\theta,\phi)}} = \frac{\sqrt{\gamma(\theta,\phi)}}{r} + r\tilde{b}(r,\theta,\phi),\ \tilde{b} \in C^k([0,r^*]\times(-\pi,\pi)\times(-\pi/2,\pi/2))$$

To put this relation in the form that will be useful in analyzing the remainder term in the progressing wave expansion, Taylor-expand $(v(\mathbf{x})v(\mathbf{x}_s))^{1/2}$ to second order and amalgamate the second-order remainder with $\tilde{b}$ above to obtain

$$a(\mathbf{x},\mathbf{x}_s) = \frac{\sqrt{\gamma(\theta,\phi)}}{r} + \frac{\sqrt{\gamma(\theta,\phi)}}{r}\mathbf{g}\cdot(\mathbf{x}-\mathbf{x}_s) + rb(r,\theta,\phi),\ b \in C^k([0,r^*]\times(-\pi,\pi)\times(-\pi/2,\pi/2)) \tag{4.62}$$

where $\mathbf{g} = (v(\mathbf{x}_s)^{-1/2}(\nabla v^{1/2})(\mathbf{x}_s)$.

4.5 Remainder

The entire construction is justified by the final step, showing that the remainder R in equation 4.5 is actually less singular than the leading term $aS(t-\tau)$, which justifies separating the two. The remainder satisfies equation 4.9, which I repeat for convenience:

$$(v^2\nabla^2 a)S(t-\tau) + \delta(\mathbf{x}-\mathbf{x}_s)\delta(t) = \frac{\partial^2 R}{\partial t^2} - v^2\nabla^2 R). \tag{4.63}$$

Introduce the *energy form*: for (any) function $p(\mathbf{x},t)$,

$$E(t) = \frac{1}{2}\int dx \left[\frac{1}{v^2(\mathbf{x})}\left(\frac{\partial p}{\partial t}(\mathbf{x},t)\right)^2 + |\nabla p(\mathbf{x},t)|^2\right]. \tag{4.64}$$

For solutions of the wave equation 4.1, the *energy identity* holds:

$$E(t_1) - E(t_0) = \int_{t_0}^{t_1} dt \int dx \frac{1}{v^2(\mathbf{x})}\frac{\partial p}{\partial t}(\mathbf{x},t)f(\mathbf{x},t). \tag{4.65}$$

To see this, differntiate the definition 4.64 with respect to time. Passing the time derivative under the integral sign, obtain

$$\frac{dE}{dt}(t) = \int dx \left[\frac{1}{v^2(\mathbf{x})}\left(\frac{\partial p}{\partial t}\frac{\partial^2 p}{\partial t^2}(\mathbf{x},t)\right) + \nabla p(\mathbf{x},t)\cdot\nabla\frac{\partial p}{\partial t}\right]$$

$$= \int dx \frac{\partial p}{\partial t}(\mathbf{x},t)\left(\frac{1}{v^2(\mathbf{x})}\frac{\partial^2 p}{\partial t^2}(\mathbf{x},t) - \nabla^2 p(\mathbf{x},t)\right)$$

after integrating by parts to shift ∇ (this step assumes that $p = 0$ for large $|\mathbf{x}|$). Use the wave equation satisfied by p to substitute f for the factor in parantheses, and so obtain

$$= \int dx \frac{1}{v^2(\mathbf{x})}\frac{\partial p}{\partial t}(\mathbf{x},t) f(\mathbf{x},t). \tag{4.66}$$

Integrate 4.66 in time from t_0 to t_1 to obtain 4.65.

The Cauchy-Schwarz inequatity and the observation that the square of $\partial p/\partial t$ is part of the energy form 4.64 leads to the inequality

$$E(t_1) - E(t_0) \le C\left(\int_{t_0}^{t_1} dt\, E(t)\right)^{\frac{1}{2}} \|f\|_{[t_0,t_1]} \tag{4.67}$$

in which $\|\cdot\|$ denotes the space-time L^2 norm:

$$\|f\|_{[t_0,t_1]} = \left(\int_{t_0}^{t_1} dt \int dx\, f(\mathbf{x},t)^2\right)^{\frac{1}{2}}, \tag{4.68}$$

and $C > 0$ depends on the minimum of v. Using the famous inequality $ab \le \frac{\alpha}{2}a^2 + \frac{1}{2\alpha}b^2$ with proper choice of $\alpha > 0$, inequality 4.67 implies in turn that

$$E(t_1) - E(t_0) \le \int_{t_0}^{t_1} dt\, E(t) + C\|f\|^2_{[t_0,t_1]} \tag{4.69}$$

This integral inequality can be solved (a fact that goes under the name "Gronwall's inequality"): the result is a bound on the growth of E in terms of the

$$E(t_1) - E(t_0) \le Ce^{t_1 - t_0}((t_1 - t_0)E(t_0) + \|f\|_{[t_0,t_1]}). \tag{4.70}$$

Apply this bound to R_1. From equation 4.73, for this case $f = (v^2\nabla^2 a)H(t-\tau)$, which is square-integrable (but see remark below), take $t_0 \le 0$ so that $E(t_0) = 0$. Note that $\partial R_1/\partial t_0 = R$ is part of E, so the specialization of inequality 4.70 to this case leads to

$$\|R\|_{[0,t]} \le Ce^t \tag{4.71}$$

in which C depends on the transport coefficient a and the traveltime τ (which determines the size of the set $\{(\mathbf{x},t) : \tau(\mathbf{x},\mathbf{x}_s) \le t\}$).

Continuing as in the last section, I will discuss explicitly the 3D case, $S = \delta$ per 4.3. In order to apply the energy estimates just explained to R, the main requirement that the

the inhomogeneous term in 4.63 be square-integrable. That is obviously not the case, no matter what the properties of a, because $\delta(t-\tau)$ is not. So the first step is to modify 4.63 to correct this deficiency.

Recall that the Green's function defined here is the causal one, therefore zero for t¡0. Integrating both sides of equation 4.63 in t from $-\infty$ to t, obtain for

$$R_1(\mathbf{x},t) = \int_{-\infty}^{t} dt' R(\mathbf{x},t') \tag{4.72}$$

$$-(v^2\nabla^2 a)H(t-\tau) + \delta(\mathbf{x}-\mathbf{x}_S)H(t-\tau) = \frac{\partial^2 R_1}{\partial t^2} - v^2\nabla^2 R_1)\,. \tag{4.73}$$

with H being the Heaviside function. Square-integrability of the RHS clearly depends on the properties of a near $\mathbf{x}=\mathbf{x}_s$, given by equaiton 4.62.

From 4.62, a is the sum of three terms. The Laplacian of the first term is

$$\sqrt{\gamma(\theta,\phi)}\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\frac{1}{r} + \frac{1}{r^3}\Delta_S\sqrt{\gamma(\theta,\phi)}$$

in which

$$\Delta_S = \frac{1}{\cos^2\phi}\frac{\partial^2}{\partial\theta^2} + \frac{1}{\cos\phi}\frac{\partial}{\partial\phi}\cos\phi\frac{\partial}{\partial\phi}$$

is the spherical Laplacian. In order for the second term to define a disribution of order zero (like $\delta(\mathbf{x}-\mathbf{x}_s)$), the coefficient of $1/r^3$ must vanish, implying that γ is constant, and the second term drops out. The first term is the same as

$$\sqrt{\gamma}\nabla^2 1r = \gamma 4\pi\delta(\mathbf{x}-\mathbf{x}_s)$$

So if we chose

$$\gamma = \frac{1}{4\pi v^2(\mathbf{x}_s)}$$

then the Laplacian of the first term in 4.62 cancels the space-time delta in 4.63.

Assume for the moment that the second term in 4.62 vanishes, that is, ∇v vanishes at $\mathbf{x}=\mathbf{x}_s$. The Laplacian of the third term is

$$\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2\cos^2\phi}\frac{\partial^2}{\partial\theta^2} + \frac{1}{r^2\cos\phi}\frac{\partial}{\partial\phi}\cos\phi\frac{\partial}{\partial\phi}\right) r b(r,\theta,\phi)$$

$$\frac{1}{r}h(r,\theta,\phi),\; h \in C^0([0,r^*]\times(-\pi,\pi)\times(-\pi/2,\pi/2))$$

The singularity is square-integrable in 3D, and remains so in 4D after multiplication by $H(t-\tau)$. That is, equation 4.73 reduces to

$$\frac{1}{r}h(r,\theta,\phi)H(t-\tau) = \frac{\partial^2 R_1}{\partial t^2} - v^2\nabla^2 R_1) \,. \tag{4.74}$$

with a square-integrable left-hand side. The energy estimate 4.70 applies: the first derivatives of R_1 are square-integrable. In particular, $R = \partial R_1/\partial t$ is square integrable, justifying the progressing wave expansion in this case.

For the general case in which ∇v does not vanish at $\mathbf{x} = \mathbf{x}_s$, the left-hand side of 4.63 has the additional summand

$$-v^2(\mathbf{x})\nabla^2 \frac{1}{4\pi v^2(\mathbf{x}_s)r}\mathbf{g}\cdot(\mathbf{x}-\mathbf{x}_s) = v^2(\mathbf{x})\frac{1}{4\pi v^2(\mathbf{x}_s)}\frac{\mathbf{g}\cdot(\mathbf{x}-\mathbf{x}_s)}{r^3} = O(r^-2). \tag{4.75}$$

This function is not square-integrable, so the energy argument does not apply directly.

4.6 Suggested Projects

1. In the ray theory section, I used the idea that the traveltime should be asymptotic to the constant velocity specializations as $\mathbf{x} \rightarrow \mathbf{x}_s$. I made a similar implicit argument for the amplitude in the following section. However both traveltime and amplitude are singular at $\mathbf{x} = \mathbf{x}_s$. Supply the missing justification.

2. Singularity is measured by the rate of growth/decay of the Fourier transform at infinite frequency. Recast the progressing wave expansion in terms of frequency:

   $$u(\mathbf{x},\omega) = a(\mathbf{x})e^{i\omega(t-\tau(\mathbf{x}))} + \frac{1}{i\omega}R(\mathbf{x},\omega)$$

   with R bounded as $\omega \rightarrow \infty$. A classic reference from this point of view is Lax (1957).

3. What really happens at a caustic? The actual behaviour of the wavefield was first understood in the 1960's. Major contributors, whose papers are well worth a read, are Ludwig (1966) and Kravtsov (1968), and Stickler et al. (1981) provide a very readable account.

4. Having worked that out, create an example like the low-velocity lens (Figure 4.1) and compute the wavefield with a finite difference code, for example IWave. What actually happens to the wavefield in the vicinity of the caustic shown in Figure 4.2? Why? How is this consistent with theory?

5. Work out all the missing details in the argument presented in the last section, and in particular answer the question posed in the last paragraph.

REFERENCES

Cerveny, V., 1985, Ray synthetic seismograms for complex two- and three-dimensional structures: J. Geophys., **58**, 2–26.

Chapman, C., 1985, Ray theory and its extensions: WKBJ and Maslov seismograms: Journal of Geophysics, **58**, 27–43.

Courant, R., and D. Hilbert, 1962, Methods of mathematical physics: Wiley-Interscience, **II**.

Fomel, S., and J. Sethian, 2002, Fast phase space computation of multiple arrivals: Proceedings of the National Academy of Sciences, **99**, 7329–7334.

Hou, J., and W. Symes, 2015, An approximate inverse to the extended Born modeling operator: Geophysics, **80**, no. 6, R331–R349.

Kravtsov, Y., 1968, Two new asymptotic methods in the theory of wave propagation in inhomogeneous media: Soviet Physics–Acoustics, **14**, 1–17.

Lax, P., 1957, Asymptotic solutions of oscillatory initial value problems: Duke Math. J., **24**, 627–646.

Ludwig, D., 1966, Uniform asymptotic expansions at a caustic: Comm. Pure Appl. Math., **29**, 215–250.

Luo, S., J. Qian, and R. Burridge, 2014, High-order factorization based high-order hybrid fast sweeping methods for point-source eikonal equations: SIAM Journal on Numerical Analysis, **52**, 23–44.

Qian, J., L.-T. Cheng, and S. Osher, 2003, Level set based Eulerian method for multivalued traveltimes in both isotropic and anisotropic wave propagations: 73rd Annual International Meeting, Expanded Abstracts, Society of Exploration Geophysicists, 1801–1804.

Stickler, D., D. Ahluwalia, and L. Ting, 1981, Application of ludwigs's uniform progressing wave ansatz to a smooth caustic: J. Acoust. Soc. Amer., **69**, 1673–1681.

Stolk, C. C., and W. Symes, 2004, Kinematic artifacts in prestack depth migration: Geophysics, **69**, 562–575.

van Trier, J., and W. Symes, 1991, Upwind finite-difference calculation of travel-times: Geophysics, **56**, 812–821.

Vidale, J., 1988, Finite-difference calculation of travel times: Bull., Seis. Soc. Am., **78**, 2062–2076.

Vinje, V., E. Iversen, and H. Gjoystal, 1993, Traveltime and amplitude estimation using wavefront construction: Geophysics, **58**, 1157–1166.

Chapter 5

The Convolutional Model

5.1 Introduction

The model of the title relates a seismic data trace $d(t)$, a *reflectivity* $r(t)$, a wavelet $w(t)$, and a noise function $n(t)$ by

$$d = w * r + n \tag{5.1}$$

An enormous literature addresses various properties of this model - see Robinson and Treitel (1980) and many works cited there. A large part of this literature is deveoted to suppressing or otherwise managing the influence of the noise term n on the recovery of r from d given w and just some properties of n.

Equation 5.1 is understood to hold for every data trace, after NMO correction as carried out as in Chapter 2. Two questions suggest themselves:

- what is NMO correction, and
- why does it lead to the very simple relation 5.1, presumed to hold for every source-receiver combination?

The aim of this chapter is to answer these two questions, *starting with the wave equation*. The approximations developed in the preceding two chapters play central roles in the answer.

5.2 Integral Representation of the Singly Scatterered Field

At first, one might be tempted to say, well *of course* the convolution relation 5.1 holds. Assuming that the pressure field $p(\mathbf{x}, t; \mathbf{x}_s)$ is a solution of the wave equation 4.1, with isotropic point radiator source $f(\mathbf{x},t) = w(t)\delta(\mathbf{x}-\mathbf{x}_s)$, then the representaion of the solution by the Green's function (equation 4.2) implies that

$$p(\mathbf{x}_r,t;\mathbf{x}_s) = \int dt_r\, w(t-t_r)G(\mathbf{x}_r,t_r;\mathbf{x}_s). \tag{5.2}$$

If we model the data as traces of the pressure field, then this would appear to give us the convolutional model 5.1, though without the noise term on the right-hand side. Apparently $r(t)$ is idenfied as the same trace of the Green's function. So Q. E. D.

There are two big things wrong with this point of view. First, the Green's function is a pressure field - a particular solution of the wave equation. It is not a property of the material. However, the reflectivity $r(t)$ is meant to correspond *directly and locally* to properties of earth material - and there is no obvious reason why the Green's function should have any relation to a local measurement of these properties (density, stiffness, etc.). In fact, $r(t)$ must somehow encode properties at depth, whereas the Green's function is evaluated at the source and receiver points in equation 5.2, and at best encodes all of the properties that the field encounters in propagating over the time interval $[0,t_r]$. Second, nominallly G is a pressure, and even if it were a local earth property (which it is not), it is not clear what combination of the basic parameters of the acoustic model it represents, or why.

We can make some progress towards a resolution of this conundrum by assuming a *known background square velocity* v^2 and regarding the data as traces of the perturbational field due to a perterbation δv^2. The perturbational pressure field δp satisfies the inhomogeneous wave equation 3.4, reproduced for convenience here:

$$\begin{aligned}\frac{\partial^2 \delta p}{\partial t^2} - v^2\nabla^2\delta p &= \delta v^2 \nabla^2 p \\ \delta p &= 0, t << 0\end{aligned} \tag{5.3}$$

The traces of δp make up the image of the linearized map $D\mathcal{F}[v^2]$. Recalling that the solution of an inhomogenous wave equation may be expressed in terms of the Green's function, obtain

$$D\mathcal{F}[v^2]\delta v^2(\mathbf{x}_r,t;\mathbf{x}_s) = \int dt_r \int dx\, G(\mathbf{x}_r,t_r;\mathbf{x})\delta v^2(\mathbf{x})\nabla^2 p(\mathbf{x},t-t_r;\mathbf{x}_s).$$

Using equation 5.2 to replace p results in

$$= w *_t \int dt_r \int dx\, G(\mathbf{x}_r,t_r;\mathbf{x})\delta v^2(\mathbf{x})\nabla^2 G(\mathbf{x},t-t_r;\mathbf{x}_s). \tag{5.4}$$

It is convenient to use the wave equation for G to replace the space derivatives in the integrand 5.4 with time derivatives. This is possible under the assumption:

$$\delta v^2(\mathbf{x}) = 0 \text{ for } \mathbf{x} \text{ near } \mathbf{x}_s. \tag{5.5}$$

Granted assumption 5.5,

$$D\mathcal{F}[v^2]\delta v^2(\mathbf{x}_r,t;\mathbf{x}_s) = w *_t \frac{\partial^2}{\partial t^2}\int dt_r \int dx\, G(\mathbf{x}_r,t_r;\mathbf{x})\frac{\delta v^2}{v^2}(\mathbf{x})G(\mathbf{x},t-t_r;\mathbf{x}_s). \quad (5.6)$$

This relation has the great advantage over equation 5.2 that a physical parameter, namely the perturbation in square velocity, appears directly. The influence of this parameter, however still appears to be non-local. To see that it is in fact local, one more level of approximation is needed.

5.3 Generalized Radon Transform

Add to the accumulating hypotheses the assumption that the background square velocity v^2 is slowly varying on the wavelength scale, and δv^2 is rapidly varying (oscillatory) with small mean over regions a wavelength in diameter. Also assume that no caustics develop in the ray fields centered at either source or receiver, at least in the region where δv^2 is non-zero. Then the result of Chapter 4 suggests that we might write G in terms of its progressing wave expansion 4.5, with $S = \delta$ (3D):

$$G(\mathbf{x},t;\mathbf{x}_s) = a(\mathbf{x},\mathbf{x}_s)\delta(t-\tau(\mathbf{x},\mathbf{x}_s)) + R(\mathbf{x},t;\mathbf{x}_s) \quad (5.7)$$

Since the first term will be most significant for high frequency waves, keep it and drop the second, insert in 5.6 to obtain

$$D\mathcal{F}[v^2]\delta v^2(\mathbf{x}_r,t;\mathbf{x}_s) \approx$$

$$w *_t \frac{\partial^2}{\partial t^2}\int dt_r \int dx\, a(\mathbf{x}_r,\mathbf{x})\delta(t_r-\tau(\mathbf{x}_r,\mathbf{x})\frac{\delta v^2}{v^2}(\mathbf{x})a(\mathbf{x},\mathbf{x}_s)\delta(t-t_r-\tau(\mathbf{x},\mathbf{x}_s))$$

(since τ and a are symmetric functions of their arguments)

$$= w *_t \frac{\partial^2}{\partial t^2}\int dx\, a(\mathbf{x}_r,\mathbf{x})a(\mathbf{x}_s,\mathbf{x})\delta(t-\tau(\mathbf{x}_r,\mathbf{x})-\tau(\mathbf{x}_s,\mathbf{x}))\frac{\delta v^2}{v^2}(\mathbf{x}) \quad (5.8)$$

That is, apart from the convolution with the wavelet w, $D\mathcal{F}$ acts (approximately) by integrating the relative perturbation $\delta v^2/v^2$ over the *isochron surface of equal two-way traveltime* $\{\mathbf{x} : t = \tau(\mathbf{x}_r,\mathbf{x}) + \tau(\mathbf{x}_s,\mathbf{x})\}$.

One additional assumption needs to be made for this reasoning to be correct: the last step in the above argument was this identity:

$$\int dt_r\delta(t_r-\tau(\mathbf{x},\mathbf{x}_r))\delta(t-t_r-\tau(\mathbf{x},\mathbf{x}_s)) = \delta(t-\tau(\mathbf{x}_r,\mathbf{x})-\tau(\mathbf{x}_s,\mathbf{x})). \quad (5.9)$$

This identity does not hold in general. To take an extreme example, if $t-\tau(\mathbf{x},\mathbf{x}_s)=\tau(\mathbf{x},\mathbf{x}_r)$ over a small region of $\mathbf{x}$, then in that region the integrand is the square of a delta function, which is not a legal distribution. This is only a hint of what goes wrong, which I am not going to discuss in detail. Suffice it to quote a result from (Gel'fand and Shilov, 1958): the left-hand side of equation 5.9 is well-defined, and equal to the right-hand side, when

$$\nabla_{\mathbf{x}}\tau(\mathbf{x},\mathbf{x}_r)\neq -\nabla_{\mathbf{x}}\tau(\mathbf{x},\mathbf{x}_s) \tag{5.10}$$

for all $\mathbf{x}_r,\mathbf{x}_s$ in the survey geometry and $\mathbf{x}$ in the support of δv^2, that is, the region where $\delta v^2\neq 0$.

Condition 5.10 has a significant physical interpretation: according to the ray equation 4.15, the gradient of τ is proportional to the ray velocity vector. Thus condition 5.10 means that the ray from $\mathbf{x}_r$ to $\mathbf{x}$ should arriva at $\mathbf{x}$ with a velocity vector unequal to the *negative* of the velocity vector of the ray from $\mathbf{x}_s$ to $\mathbf{x}$, which is the velocity vector of the ray from $\mathbf{x}$ to $\mathbf{x}_s$. If this condition is violated, then the curve got by concatenating the rays from $\mathbf{x}_s$ to $\mathbf{x}$ and from $\mathbf{x}$ to $\mathbf{x}_r$ is itself a ray (thanks to the uniqueness theory for ordinary differential equations). Thus in that case there is a ray from $\mathbf{x}_s$ to $\mathbf{x}_r$, that is, a direct or diving ray connecting source and receiver. This circumstance must be ruled out if equation 5.9 is to hold: specifically, *no diving ray may intersect the support of* δv^2.

Add this no-diving-rays hypothesis to the standing assumptions. Two remarks are worth making:

- A more refined analysis shows that it's OK for diving rays to pass over the support of δv^2, so long as directions in which δv^2 oscillates strongly are not perpindicular to the rays. The ideas necessary to properly formulate and justify this assertion will appear later in the book.
- It is ironic that contemporary FWI relies almost entirely on diving waves, that is, waves associated to the very rays outlawed by condition 5.10. This point will also get some attention later in the book.

The formula 5.8 is surprisingly accurate under the conditions assumed, as I showed by explicit example in my old notes (Symes, 1998). With numerical methods for computation of traveltimes and amplitudes, 5.8 turns into a so-called Kirchhoff integral operator approximation to $D\mathcal{F}$. Such numerical integral operator representations are very useful, and widely used.

5.4 Layered Models

The convolutional model is the result of all of the assumptions and approximations made so far, plus a few more. The first of these, in its strictest form, is that *the earth is layered*:

that is that the phsyical properties of rock affecting seismic wave propagation (meaning v^2, in the constant density acoustic model) depend only on $x_3 = z$. Also, the sources and receivers are arrayed on horizontal surfaces, that is, $z = z_s$ and $z = z_r$ respectively.

In fact, the earth is not layered (even locallly) - if it were, then the results of seismic experiments would be translation-invariant, and they are not. For example, Figure 5.1 shows every 50th *Common Midpoint* gather. Common midpoint ("CMP") gathers are the collections of traces sharing a common *midpoint* $x_m = (x_r + x_s)/2, y_m = (y_r + y_s)/2$, more on that later). The half-offset $h = |(x_r - x_s, y_r - y_s)|/2$ indexes traces in a CMP gather. The scalar half offset defined here suffices for present purposes, since the layered hypothesis implies rotational symmetry of the traces in a CMP about the midpoint.

The CMP gathers in Figure 5.1 change quite a bit one to another, but they are 625 m apart. Figure 5.2 on the other hand shows the same number of CMP gathers, but chosen contiguously, with spacing of 12.5 m. Over this range of CMPs, the data changes quite slowly, so perhaps it is not so dreadfully inaccurate to model the earth as layered on this scale.

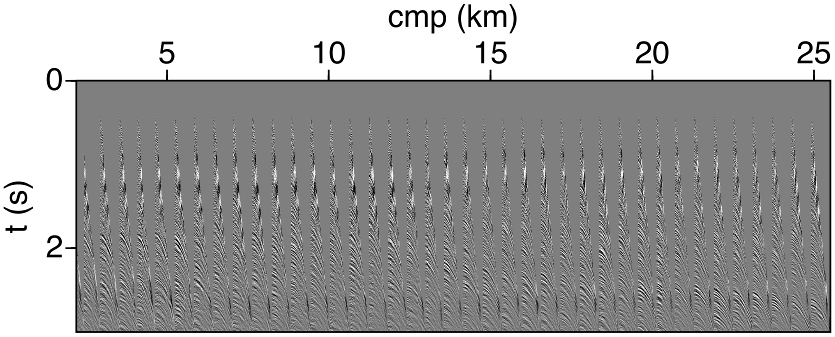

Figure 5.1: Every 50th Common Midpoint Gather from the Viking Graben line introduced in Chapter 2. Midpoint spacing is 625 m.

Due to the rotational symmetry already noted,

$$D\mathcal{F}[v^2]\delta v^2(\mathbf{x}_r, t; \mathbf{x}_s) = D\mathcal{F}[v^2]\delta v^2(\mathbf{x}_m + (h,0,0), t; \mathbf{x}_m - (h,0,0)) \tag{5.11}$$

This identity suggests a change of coordinates: replace

$$D\mathcal{F}[v^2]\delta v^2(x_m, y_m h, t) \leftarrow D\mathcal{F}[v^2]\delta v^2((x_m + h, y_m, z_r)t; (x_m - h, y_m, z_s)). \tag{5.12}$$

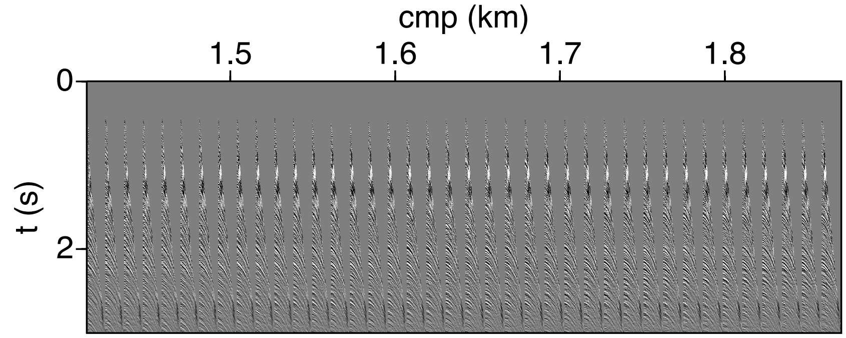

Figure 5.2: Common Midpoint Gathers 1002-1040 from the Viking Graben line introduced in Chapter 2. Midpoint spacing is 12.5 m.

These *midpoint-offset* coordinates are in common use. To honor the actual non-layered nature of the earth, a vector half-offset $h_x = (x_r - x_s)/2, h_y = (y_r - y_s)/2$ may be used instead.

5.5 Stationary Phase

Analysis of the layered case uses yet another approximation. Write the integral representation 5.8 in terms of the Fourier transform in time:

$$D\mathcal{F}[v^2]\delta v^2(\mathbf{x}_r,t;\mathbf{x}_s) \approx -\frac{1}{2\pi}\int d\omega\,\omega^2\hat{w}(\omega)\int dz\left(\int dx\int dy\right.$$

$$\left. a(\mathbf{x}_r,(x,y,z))a(\mathbf{x}_s,(x,y,z))e^{i\omega(t-\tau(\mathbf{x}_r,(x,y,z))-\tau(\mathbf{x}_s,(x,y,z)))}\right)\frac{\delta v^2}{v^2}(z) \qquad (5.13)$$

Since w is assumed highly oscillatory on the wavelength scale, its Fourier transform is concentrated a "large" frequencies. The principle of stationary phase (Guillemin and Sternberg, 1979; Bleistein et al., 2001) gives an asymptotic evaluation of integrals such as the dx integral in 5.13 for large frequency. To state the general principle: the integral of a rapidly fluctuating function $g(\mathbf{y})e^{i\omega\psi(\mathbf{y})}$ for large ω is approximated to every order in ω

by a sum of terms, one for each stationary phase point $\mathbf{y}^*$ (that is $\nabla\psi(\mathbf{y}^*) = 0$):

$$\begin{aligned}&\int_{\mathbf{R}^m} d\mathbf{y} g(\mathbf{y}) e^{i\omega\psi(\mathbf{y})}\\ &\approx \sum_{\nabla\psi(\mathbf{y}^*)=0} \left(\frac{2\pi}{\omega}\right)^{\frac{m}{2}} e^{\frac{\pi i}{4}\text{sgn Hess }\psi(\mathbf{y}^*)} |\text{det Hess}\psi(\mathbf{y}^*)|^{-\frac{1}{2}} g(\mathbf{y}^*) e^{i\omega\psi(\mathbf{y}^*)}\end{aligned} \tag{5.14}$$

The main condition for validity of this approximation is that the Hessian determinant not vanish at the stationary points (that is, that the stationary points are *nondegenerate*).

Note: The symbol sgn in 6.56 means the *signature* of the symmetric matrix Hess$\psi(\mathbf{y}^*$, that is, the number of positive eigenvalues minus the number of negative eigenvalues.

The stationary phase principle 6.56 may be applied to approximate the integral in 5.13 over the horizontal variables $x_1 = x, x_2 = y$, identifying $m = 2$, $\psi(\mathbf{x}) = t - \tau(\mathbf{x}, \mathbf{x}_r) - \tau(\mathbf{x}, \mathbf{x}_s)$ and $g(\mathbf{x}) = -(a(\mathbf{x}, \mathbf{x}_r)a(\mathbf{x}, \mathbf{x}_s)/2\pi)(\delta v^2(\mathbf{x})/v^2(\mathbf{x}))$. This requires (a) finding the stationary points, and (b) verifying that the Hessian is nonsingular there.

The rotational symmetry implicit in the layered assumption strictly constrains the locatino of stationary points. To see how, note that rotational symmetry implies that $\tau(\mathbf{x}, \mathbf{x}_s)$ depends only $|(x - x_s, y - y_s)| = r_s$:

$$\tau(\mathbf{x}, \mathbf{x}_s) = \bar{\tau}(r_s, z, z_s), \tag{5.15}$$

and similarly $\tau(\mathbf{x}, \mathbf{x}_r) = \bar{\tau}(r_r, z, z_r)$. Note that $\bar{\tau}$ also solves the eikonal equation, in 2D:

$$\left(\frac{\partial\bar{\tau}}{\partial r}\right)^2 (r, z, z_s) + \left(\frac{\partial\bar{\tau}}{\partial z}\right)^2 (r, z, z_s) = \frac{1}{v^2(z)} \tag{5.16}$$

The locations of the stationary points of ψ, and their non-degeneracy, follows from several properties of $\bar{\tau}$:

- There exists a decreasing positive function $z_s < z \mapsto r_{\max}(z)$ so that a unique solution $\bar{\tau}$ of 5.16 exists on $\{(r, z) : z_s < z, |r| < r_{\max}\}$ satisfying

  $$\bar{\tau}(r, z, z_s) \sim \frac{\sqrt{(z - z_s)^2 + r^2}}{v(z_s)} \text{ as } z \to z_s,\ r \to 0. \tag{5.17}$$

  and similarly for z_r;

- in $\{(r, z) : z_s < z, |r| < r_{\max}\}$,

  - $$r\frac{\partial\bar{\tau}}{\partial r} > 0 \text{ if } |r| > 0 \tag{5.18}$$

–

$$\frac{\partial^2 \bar{\tau}}{\partial r^2}(r,z,z_s) > 0, \tag{5.19}$$

and similarly for z_r.

Granted these results, note that

$$\nabla_{x,y}\psi(\mathbf{x}) = -\nabla_{x,y}\tau(\mathbf{x},\mathbf{x}_r) - \nabla_{x,y}\tau(\mathbf{x},\mathbf{x}_s)$$

$$= -\frac{(x-x_r, y-y_r)}{|(x-x_r,y-y_r)|}\frac{\partial \bar{\tau}}{\partial r}(r_r,z,z_r) - \frac{(x-x_s,y-y_s)}{|(x-x_s,y-y_s)|}\frac{\partial \bar{\tau}}{\partial r}(r_s,z,z_s) \tag{5.20}$$

For $\mathbf{x}$ to be a stationary point of ψ, it follows from 5.20 that the two unit vectors appearing there must be colinear, that is, (x,y) must be on the line joining (x_r,y_r) and (x_s,y_s). Let

$$r_{sr} = |(x_r - x_s, y_r - y_s)|,\ (e_x,e_y) = \frac{(x_r - x_s, y_r - y_s)}{r_{sr}}$$

Then for a suitable r,

$$(x-x_s, y-y_s) = r(e_x,e_y),\ (x-x_r,y-y_r) = (r-r_{sr})(e_x,e_y).$$

and the stationary phase condition 5.20 hold if and only if

$$\mathrm{sgn}(r)\frac{\partial \bar{\tau}}{\partial r}(r_{sr}-r_s,z,z_r) + \mathrm{sgn}(r-r_{sr})\frac{\partial \bar{\tau}}{\partial r}(r_s,z,z_s) = 0$$

Either the two signs are the same, or they are different. In fact, it follows from the condition 5.18 that they cannot be the same (exercise!). Thus stationary phase 5.20 implies that

$$\frac{\partial \bar{\tau}}{\partial r}(r_{sr}-r,z,z_r) = \frac{\partial \bar{\tau}}{\partial r}(r,z,z_s). \tag{5.21}$$

The meaning of condition 5.21 is easiest to understand when $z_s = z_r$: in that case, the two sides of 5.21 are the same function with reversed argument. Since this function of r is increasing (result 5.18), the unique solution is $r = r_{sr}/2$, from which it follows that

$$x = x_m = \frac{x_r - x_s}{2},\ y = y_m = \frac{y_r - y_s}{2}. \tag{5.22}$$

That is, the stationary point $\mathbf{x}$ lies beneath the *midpoint* (x_m,y_m) of the line segment between source and receiver. The horizontal distance from either is $r = r_{rs}/2 = h$, the half-offset, so write from now on h instead of r.

It is not hard to see that the for typical z_r, z_s in towed streamer marine surveys - both a few meters - and typical target depth z - kilometers - the error caused by assuming $z_s = z_r$

in this calculation is negligible. Therefore 5.22 is assumed to hold in the following, and I will also assume $z_s = z_r$ to simplify various expressions.

Having found the only stationary points, the other issue that must be settled in order to apply the stationary phase approximation 6.56 is the nondegeneracy of the phase Hessian. From the definition 5.15 of $t\bar{a}u$ and its eikonal 5.16 it follows that at the midpoint 5.22,

$$\left(\begin{array}{cc} \dfrac{\partial^2\psi}{\partial x^2} & \dfrac{\partial^2\psi}{\partial x\partial y} \\ \dfrac{\partial^2\psi}{\partial x\partial y} & \dfrac{\partial^2\psi}{\partial y^2} \end{array}\right)(\mathbf{x}^*) =$$

$$-\left(\begin{array}{cc} e_y^2\dfrac{2}{r}\dfrac{\partial\bar{\tau}}{\partial r} + e_x^2\dfrac{\partial^2\bar{\tau}}{\partial r^2} & -e_xe_y\left(\dfrac{2}{r}\dfrac{\partial\bar{\tau}}{\partial r} + \dfrac{\partial^2\bar{\tau}}{\partial r^2}\right) \\ -e_xe_y\left(\dfrac{2}{r}\dfrac{\partial\bar{\tau}}{\partial r} + \dfrac{\partial^2\bar{\tau}}{\partial r^2}\right) & e_x^2\dfrac{2}{r}\dfrac{\partial\bar{\tau}}{\partial r} + e_y^2\dfrac{\partial^2\bar{\tau}}{\partial r^2} \end{array}\right)(h,z) \tag{5.23}$$

In view of the inequalities 5.18 and 5.19, the determinant of this matrix is nonzero (in fact, positive). Thus the stationary point of ψ (for each z,t) is nondegenerate.

The identity 5.23 implies that the Hessian has two negative eigenvalues. The easiest way to see this is to check the case in which the offset vector is parallel to the x-axis, so that $|e_x| = 1, e_y = 0$. Then the Hessian is diagonal, and the diagonal entries are exactly twice the negatives of the quantities estimated in 5.18 and 5.19. Since the Hessian for other orientations of the offset is related to the x-axis case by rotational similarity, the eigenvalues are always these quantities. Conclude that

$$\mathrm{sgnHess}\psi(\mathbf{x}^*) = -2 \Rightarrow e^{\frac{i\pi}{4}\mathrm{sgnHess}\psi(\mathbf{x}^*)} = -i \tag{5.24}$$

Since the stationary point is below the midpoint, rotational symmetry implies that this determinant is a function of h and z: abuse notation by writing

$$\mathrm{Hess}(h,z) = \left(\begin{array}{cc} \dfrac{\partial^2\psi}{\partial x^2} & \dfrac{\partial^2\psi}{\partial x\partial y} \\ \dfrac{\partial^2\psi}{\partial x\partial y} & \dfrac{\partial^2\psi}{\partial y^2} \end{array}\right)(\mathbf{x}^*) \tag{5.25}$$

where $h = r_{rs}/2 = |(x_r - x_s, y_r - y_s)|/2$ and $\mathbf{x}^* = (x_m, y_m, z)$.

The preceding paragraphs have supplied all of the ingredients necessary to apply the stationary phase approximation 6.56 to the frequency-domain expression of the Generalized Radon Transform approximation 5.13. As the reader will verify, the result is

$$D\mathcal{F}[v^2]\delta v^2(x_m,y_m,h,t) \approx\approx \int d\omega\, i\omega\hat{w}(\omega) \int dz\, e^{i\omega(t-2\bar{\tau}(h,z,z_s))} a(\mathbf{x}_s,\mathbf{x})a(\mathbf{x}_r,\mathbf{x})|Hess(h,z)|^{-1/2}\delta v^2(z) \tag{5.26}$$

(including a minus sign from the signature of the Hessian, which is 2).

Rotational symmetry also applies to the transport equation 4.8, and shows that $a(\mathbf{x}_s,\mathbf{x}_m) = a(\mathbf{x}_r,\mathbf{x}_m) = \bar{a}(h,z)$ (defining $\bar{a}$). Insert this identity and use the Fourier representation of the delta function to arrive at

$$D\mathcal{F}[v^2]\delta v^2(x_m,y_m,h,t) \approx w *_t \frac{\partial}{\partial t}\int dz\,\delta(t-2\bar{\tau}(h,z,z_s))\bar{a}^2(h,z)|\det \mathrm{Hess}(h,z)|^{-1/2}\frac{\delta v^2}{v^2}(z) \tag{5.27}$$

This approximation is very close to the desired model - the remaining steps will be taken in the next section.

The key inequalities 5.18 and 5.19 still require justification. The key observation in this argument is that $\bar{\tau}$ can also be constructed via rays, as shown in Chapter 4. The assocated Hamilton's equations 4.27 for the phase space vector (r,z,p,q) read

$$\frac{dr}{dt} = v^2(z)p \tag{5.28}$$

$$\frac{dz}{dt} = v^2(z)q \tag{5.29}$$

$$\frac{dp}{dt} = 0 \tag{5.30}$$

$$\frac{dq}{dt} = \frac{d\log v}{dz}(z) \tag{5.31}$$

In particular, p is constant along rays. Divding the first equation by the second and using the eikonal equation $v^2(z)(p^2+q^2)=1$, obtain

$$\frac{dr}{dz} = \frac{v(z)p}{\sqrt{1-v(z)^2p^2}} \tag{5.32}$$

which gives the evolution of r as a function of z along the ray so long as $q > 0$, that is, $v(z)|p| < 1$. The latter condition characterizes rays for which $dz/dt > 0$, implying also that along such rays, $\partial\tau/\partial z > 0$ .Rays violate this condition by turning horizontal. It is not too hard to see that a horizontal ray velocity vector implies the presence of a caustic, violating the conditions that ensure that the approximations 5.8 and 5.13 are accurate. Therefore, we assume that over the depth range of interest, and over the rays that are important in this construction, $v(z)|p| < 1$. This assumption will be justified ex post facto.

Integrating 5.32, obtain

$$r(z,p) = \int_{z_s}^{z} dz'\,v(z')p(1-v(z')^2p^2)^{-\frac{1}{2}} \tag{5.33}$$

Then

$$\frac{\partial r}{\partial p}(z,p) = \int_{z_s}^{z} dz'\,v(z')(1-v(z')^2p^2)^{-\frac{3}{2}} > 0 \tag{5.34}$$

under the assumptions made above. That is, define $p_{\max}(z) = \min\{1/v(z') : z_s \le z' \le z\}$. Then the function $p \mapsto r(z,p)$ is strictly increasing on $(-p_{\max}(z), p_{\max}(z))$ hence invertible: write $p(r,z)$ for the inverse, which is well-defined on $(-r_{\max}(z), r_{\max}(z))$ with $r_{\max}(z) = r(p_{\max}(z), z)$.

Of course, in equation 5.32 and following, p is the x-component of the slowness vector $\mathbf{P}$ in Hamilton's equations (that is, the system 5.28). So from 4.19 and the strict increase of $r \mapsto p(r,z)$ follows inequality 5.19. Symmetry implies that

$$\frac{\partial \bar{\tau}}{\partial r}(0,z) = 0;$$

together with the convexity result, that's enough to establish inequality 5.18.

5.6 Hyperbolic Normal Moveout

Transform the integral in equation 5.27 by the change of variable $z \to T(h,z,z_s) = 2\bar{\tau}(h,z,z_s)$. T is the *two-way time* from $(0,z_s)$ to (h,z). Because of the standing hypothesis that $dz/dt > 0$ along all rays from source or receiver to (h,z) in the region of defined by the inequalities $h < r_{\max}(z,p)$, this change of variables is invertible, and defines depth as a function of two-way time: $Z(h, T(h,z,z_s), z_s) = z$. Thus equation 5.27 can be re-written as

$$D\mathcal{F}[v^2]\delta v^2(x_m, y_m, h, t) \approx$$

$$w *_t \frac{\partial}{\partial t}\int dT \approx \int dz\,\delta(t-T)\left(\frac{\partial Z}{\partial T}\bar{a}^2|\det \mathrm{Hess}|^{-1/2}\right)(h, Z(h,T,,z_s)\frac{\delta v^2}{v^2}(Z(h,T,z_s))$$

$$= w *_t \frac{d}{dt}A(h,t,z_s)R(h,t,z_s) \tag{5.35}$$

in which

$$A(h,t,z_s) = \left(\frac{\partial Z}{\partial t}\bar{a}^2|\det \mathrm{Hess}|^{-1/2}\right)(h, Z(h,t,,z_s),\; R(h,t,z_s) = \frac{\delta v^2}{v^2}(Z(h,t,z_s))$$

Now rely on the separation of scales assumed to justify the use of geometric optics. Since v^2 is much smoother than δv^2, which is oscillatory on the scale of a wavelength, derivatives of A (depending on traveltime and amplitude, which are as smooth as v^2 or smoother) are much smaller than corresponding derivatives of R. Make another approximation of the same quality as those that have been used to arrive at 5.35 to arrive at

$$D\mathcal{F}[v^2]\delta v^2(x_m, y_m, h, t) \approx w *_t A(h,t,z_s)\frac{dR}{dt}(h,t,z_s) \tag{5.36}$$

The hyperbolic normal moveout approximation takes two further steps:

- replacing the depth coordinate z by the zero-offset two-way time $T_0(z,z_s) = T(0,z,z_s)$ with inverse $Z_0(t_0,z_s)$ (so that $Z_0(T_0(z,z_s),z_s) = z$), and
- replacing the non-zero offset time $T(h,z,z_s)$ with the lowest order nontrivial Taylor series in h, which may be expressed as

$$\bar{T}(h,t_0) = \sqrt{t_0^2 + \frac{h^2}{v_{\mathrm{RMS}}^2(t_0)}} = T(h,Z_0(t_0,z_s),z_s) + O(h^4). \tag{5.37}$$

To prove 5.37, begin by noting that

$$\left(\frac{\partial T}{\partial z}\right)^2 + \left(\frac{\partial T}{\partial h}\right)^2 = \frac{4}{v^2} \tag{5.38}$$

Differentiate this twice with respect to h and use the vanishing of odd-order h derivatives at $h = 0$ (implied by symmetry) to conclude that

$$\partial T(0,z,z_s)/\partial z = \frac{2}{v(z)} \tag{5.39}$$

so the second h derivative

$$q(z) = \frac{\partial^2 T}{\partial h^2}(0,z)$$

satisfies

$$\frac{2}{v}\frac{dq}{dz} + q^2 = 0$$

Introduce temporarily a new depth coordinate

$$\sigma(z) = \frac{1}{2}\int_{z_s}^{z} v$$

Then in terms of σ, q satisfies the Ricatti equation

$$\frac{dq}{d\sigma} + q^2 = 0$$

The solution which is singular at $\sigma = 0$, i.e. $z = z_s$, is

$$q(\sigma) = \frac{1}{\sigma} = \frac{2}{\int_{z_s}^{z} dz\, v}$$

From equation 5.39, you can also write this as

$$\frac{\partial^2 \bar{T}}{\partial h^2}(0,t_0) = \frac{\partial^2 T}{\partial h^2}(0,Z_0(t_0,z_s),z_s) q(\sigma(Z_0(t_0,z_s)) = \frac{1}{\int_0^{t_0} dt_0\, \bar{v}^2}$$

where $\bar{v}(t_0) = v(Z_0(t_0, z_s))$.

Note: In the seismic literature, $v(z)$ is called the interval velocity as a function of z, $\bar{v}(t_0)$ is called the interval velocity as a function of time - see self-doc for the SU utility `velconv`).

Thus

$$\begin{aligned}\bar{T}(h,t_0) &= t_0 + \frac{h^2}{2}\frac{\partial^2 \bar{T}}{\partial h^2}(h,t_0) + \ldots \\ &= t_0 + \frac{h^2}{2\int_0^{t_0} dt_0\,\bar{v}^2} + \ldots\end{aligned}$$

Since

$$\frac{\partial^2}{\partial h^2}(\bar{T}(h,t_0))^2_{h=0} = 2t_0\frac{\partial^2 \bar{T}}{\partial h^2}(0,t_0)$$

the above can be rewritten as

$$\begin{aligned}\bar{T}(h,t_0)^2 &= t_0^2 + \frac{h^2}{\frac{1}{t_0}\int_0^{t_0} dt_0\,\bar{v}^2} + \ldots \\ &= t_0^2 + \frac{h^2}{v_{\mathrm{RMS}}^2(t_0)} + \ldots \qquad (5.40)\end{aligned}$$

where

$$v_{\mathrm{RMS}}(t_0) = \sqrt{\frac{1}{t_0}\int_0^{t_0} dt_0\,\bar{v}^2} \qquad (5.41)$$

is the Root-Mean Square (RMS) velocity as function of t_0. Evidently equation 5.40 is equivalent to 5.37.

5.7 NMO Correction

The NMO correction is a change of variable in the data, namely

$$d(x_m, y_m, h, t) \rightarrow d(x_m, y_m, h, \bar{T}(h,t_0)). \qquad (5.42)$$

This operation has a simple interpretation under the assumption that an effective source signature deconvolution has been performed in the convolutional model 5.52, that is, that an inverse to the convolution with the source wavelet w has been performed. This would remove $w*_t$ from the right-hand side of 5.52. Actually exact source signature deconvolution is impossible, as realistic source wavelets are band-limited. Therefore there must be a residual source signature built into the data even after any sort of deconvolution is applied, consisting (at best!!) of a bandpass filter. I will comment further on this matter at the end of the chapter, but for now presume that deconvolution has removed $w*_t$ from the right-hand side of 5.52, to yield the *deconvolved convolutional model* of the NMO-corrected data:

$$d(x_m, y_m, h, \bar{T}(h, t_0)) \approx A(h, t, z_s)\frac{dR}{dt}(h, t, z_s) \tag{5.43}$$

Recall that

$$R(h, t, z_s) = \frac{\delta v^2}{v^2}(Z(h, t, z_s)),$$

so

$$R(h, \bar{T}(h, t_0), z_s) = \frac{\delta v^2}{v^2}(Z(h, \bar{T}(h, t_0), z_s)) = \frac{\delta v^2}{v^2}(Z(h, T(h, Z_0(, t_0, z_s), z_s))) = \frac{\delta v^2}{v^2}(Z_0(t_0, z_s))$$

since T and Z are an inverse pair. Note that $R(h, \bar{T}(h, t_0), z_s)$ is *independent of* t. Define

$$\bar{R}(t_0) = \frac{\delta v^2}{v^2}(Z_0(t_0, z_s)), \tag{5.44}$$

$$barA(h, t_0) = A(n, \bar{T}(h, t_0), z_s)\left(\frac{\partial \bar{T}}{\partial t_0}(h, t_0)\right)^{-1} \tag{5.45}$$

In terms of of $\bar{R}$ and $\bar{A}$, the deconvolved convolutional model 5.43 of a CMP gather is fantastically simple:

$$d(x_m, y_m, h, \bar{T}(h, t_0)) \approx \bar{A}(h, t_0)\frac{d\bar{R}}{dt_0}(t_0) \tag{5.46}$$

The second factor is the *reflectivity* - simply the derivative of the relative perturbation in v^2, parametrized by vertical two-way time t_0. The first is a combination of various amplitude effects, including the geometric amplitudes from source/receiver at offset h, the reciprolcal square root of the phase Hessian (measuring wavefront curvature), and the rate of change of two-way time at offset h with respect to two-way time at offset zero.

In the case of constant v, $\bar{A}$ can certainly be computed explicitly. However, its most important feature is that it is *smooth*, that is slowly varying on the wavelength scale, as it combines various geometric optics quantities that are smooth under the standing assumptions on v. The reflectivity on the other hand is not, so a grey-scale plot of NMO-corrected will have *flat wavefronts* - rapidly oscillating vertically, slowly varying horizontally. That is what you see in Figures 2.5a, 2.5b, 2.6a, and 2.6b. This feature is an indication that the data d and the velocity v are compatible, that is, that v adequately explains the kinematic features of the data.

On the other hand, if the NMO-corrected CMP gathers are not flat, then the velocity used to compute the NMO correction is not compatible with the data. Figure 2.7 showed an example of this.

Careful inspection of the relation 5.48 reveals the reason. To clarify the role of velocity, include it as an argument of $\bar{T}$: that is, $\bar{T}[v](t_0,h)$

$$\bar{T}[v](h,t_0) = \sqrt{t_0^2 + \frac{h^2}{v_{\mathrm{RMS}}^2(t_0)}} \tag{5.47}$$

Denote by $\bar{T}_0[v](h,t)$ its inverse function, that is,

$$\bar{T}[v](h,\bar{T}_0[v](h,t)) = t$$

Suppose that the data are compatible with v, that is that relation 5.48 holds. Choose another velocity v', and construct the NMO-corrected CMP for this velocity:

$$d(x_m,y_m,h,\bar{T}[v'](h,t_0)) \approx \bar{A}(h,\bar{T}_0[v](h,\bar{T}[v'](h,t_0)))\frac{d\bar{R}}{dt_0}(\bar{T}_0[v](h,\bar{T}[v'](h,t_0))) \tag{5.48}$$

The surfaces of equal phase for the last factor - these delineate the visible wavefronts in the NMO corrected gather, since the other factor is smooth - are the curves defined by

$$h \mapsto \bar{T}_0[v](h,\bar{T}[v'](h,t_0)).$$

To describe these curves conveniently, introduce the *two way time discrepancy* between v and v', denoted $\Delta T[v,v'](h,t_0)$, by

$$\sqrt{t_0^2 + \Delta T[v,v'](h,t_0)^2} = \bar{T}_0[v](h,\bar{T}[v'](h,t_0)) \tag{5.49}$$

Since $\bar{T}$ and $\bar{T}_0$ are an inverse pair, evaluating $\bar{T}[v](h,\cdot)$ at the right-hand-side of equation 5.49 gives

$$t_0^2 + \Delta T[v,v'](h,t_0)^2 + \frac{h^2}{v_{\mathrm{RMS}}^2(\bar{T}_0[v](h,\bar{T}[v'](h,t_0)))} = t_0^2 + \frac{h^2}{(v')_{\mathrm{RMS}}^2(t_0)}$$

so

$$\Delta T[v,v'](h,t_0)^2 = h^2\left(\frac{1}{(v')_{\mathrm{RMS}}^2(t_0)} - \frac{1}{v_{\mathrm{RMS}}^2(\bar{T}_0[v](h,\bar{T}[v'](h,t_0)))}\right) \tag{5.50}$$

It is worth taking a moment to understand the geometric meaning of equation 5.50. Equal phase surfaces $t = \bar{T}[v](h,t_0)$ are called *moveout curves* (or surfaces, in 3D). The inverse function $\bar{T}_0[v](h,t)$ gives the zero-offset two-way time for the moveout curve passing through (h,t), for the velocity v. So the *residual moveout* $\bar{T}_0[v](h,\bar{T}[v'](h,t_0))$ is the zero-offset intercept of the moveout curve for v, passing through the point $(h,\bar{T}[v'](h,t_0))$ on the moveout curve for v'. $\Delta T[v,v'](h,t_0)$ is the difference between t_0^2 and this two-way time (squared) obtained by starting at $(0,t_0)$ on the moveout curve for v', traveling out to offset h, and then back to $h = 0$ on the moveout curve for v. Residual moveout is a direct measure of the difference in the NMO-corrected phase. Equation 5.50 relates residual moveout directly to the values of the two velocities v and v'.

Since the residual moveout formula samples the two velocities at two different places, one of those depending on the velocities themselves, it is not entirely obvious exactly what equation 5.50 tells us about the relation with the velocity difference. The easy case to understand is when both v and v' are constant. Then $v_{\rm RMS} = v$, $v'_{\rm RMS} = v'$, and 5.50 becomes

$$\Delta T[v,v'](h,t_0)^2 = h^2\left(\frac{1}{(v')^2} - \frac{1}{v^2}\right) \tag{5.51}$$

This tells you that if you NMO-correct with a wrong constant velocity, the gathers will definitely not be flat, in fact residual moveout curvature is proportional to the error between $1/v^2$ and $1/(v')^2$. This in this case at least gather flatness directly measures error in velocity.

The non-constant velocity case is harder, but not impossible, to analyse, and the conclusion is the same: velocity errors large enough to cause significant traveltime errors on the scale of a wavelength show up in non-flat NMO corrected CMP gathers. See (Symes, 1999) for details.

5.8 NMO Stretch

As mentioned earlier, the "deconvolved convolutional model" is necessarily a myth, as ideal source signature deconvolution is impossible. Even granted strict validity of the acoustic constant density Born approximation with the isotropic point radiator source model, the optimal result still bandpass filters the Born impulse response (equation **??**). This fact has a striking consequence for NMO correction.

Compose both sides of the deconvolved convolutional model equation 5.48 with the *inverse NMO* function $\bar{T}_0$ to obtain tthe perfectly deconvolved trace in terms of the amplitude $\bar{A}$ and the reflectivity $\bar{R}$, both parametrized by offset h and vertical two-way time t_0:

$$d_{\rm imp}(x_m,y_m,h,t) \approx \bar{A}(h,t_0)\frac{d\bar{R}}{dt_0}(\bar{T}_0(h,t)). \tag{5.52}$$

I've renamed the left-hand side as $d_{\rm imp}$, for the impulsive (perfectly deconvolved) trace. Then admit that there is indeed a wavelet w in the mix:

$$d(x_m,y_m,h,t) = w *_t d_{\rm imp}(x_m,y_m,h,t)$$

$$\approx \int dt'\, w(t-t')\bar{A}(h,\bar{T}_0(h,t'))\frac{d\bar{R}}{dt_0}(\bar{T}_0(h,t'))$$

So the NMO-corrected trace is NOT the convolution of w with its perfectly-deconvolved cousin:

$$d(x_m, y_m, h, \bar{T}(h,t_0)) \approx \int dt_0'\, w(\bar{T}(h,t_0) - \bar{T}(h,t_0'))\frac{\partial \bar{T}}{\partial t_0}(h,t_0')\bar{A}(h,t_0')\frac{d\bar{R}}{dt_0}(h,t_0') \qquad (5.53)$$

Note that the argument of w inside the integral is (generally) nonlinear in t and t': this is not a convolution. The ratio

$$\frac{\bar{T}(h,t_0) - \bar{T}(h,t_0')}{t_0 - t_0'}$$

measures the distortion of convolution: it is roughly the Jacobian factor inside the integral, namely the rate of change of time with respect to zero-offset time. This factor is known as *NMO stretch*, in honor of its graphical effect.

To see how this works, start with the idealized convolutional model seismogram depicted in Figure 5.3a. This is a "pure event", that is, $\bar{R}$ is a Heaviside function so that its t_0 derivative is a delta function, and the amplitude $\bar{A}$ is set to 1. The SU command `suaddevent` is very convenient for creating such idealized reflections with hyperbolic moveout. The RMS velocity is chosen as 3000 m/s (the default), and the zero-offset intercept (of the moveout curve) as $t_0 = 0.8$ s - these two choices completely determine the moveout curve. The source wavelet is a [2,5,25,35] Hz trapezoidal bandpass filter.

Figure 5.3b shows the NMO correction. Notice that the wavelet seems to spread as offset increases: that's why the non-convolution of equation 5.53 is called "NMO stretch". If you were to add more events to the data, you would find that the degree of stretch varies with t_0.

Note that Viking Graben NMO-corrected CMP gathers (Figures 2.5a, 2.5b, 2.6a, and 2.6b) inherit a *mute* from the CMP gathers themselves, and this mute is chosen so that the NMO-corrected events are cut off before the offset becomes large enough to exhibit severe NMO stretch as in Figure 5.3b. However careful inspection will reveal that even the Viking Graben events are stretched a bit.

The implication of NMO stretch is that NMO correction does not actually reproduce a material parameter, the reflectivity, even accounting for amplitude. Instead, it produces a nonlinearly filtered version of a material parameter, with the degree of filtering varying with offset and vertical travel time. This feature is a fundamental defect of the convolutional model and NMO correction, with which the community has struggled for decades. See for example Harlan (2014) for a clever use of optimization to choose a *time-varying wavelet* so model, and eventually remove, NMO stretch and similar effects, also many references to older work.

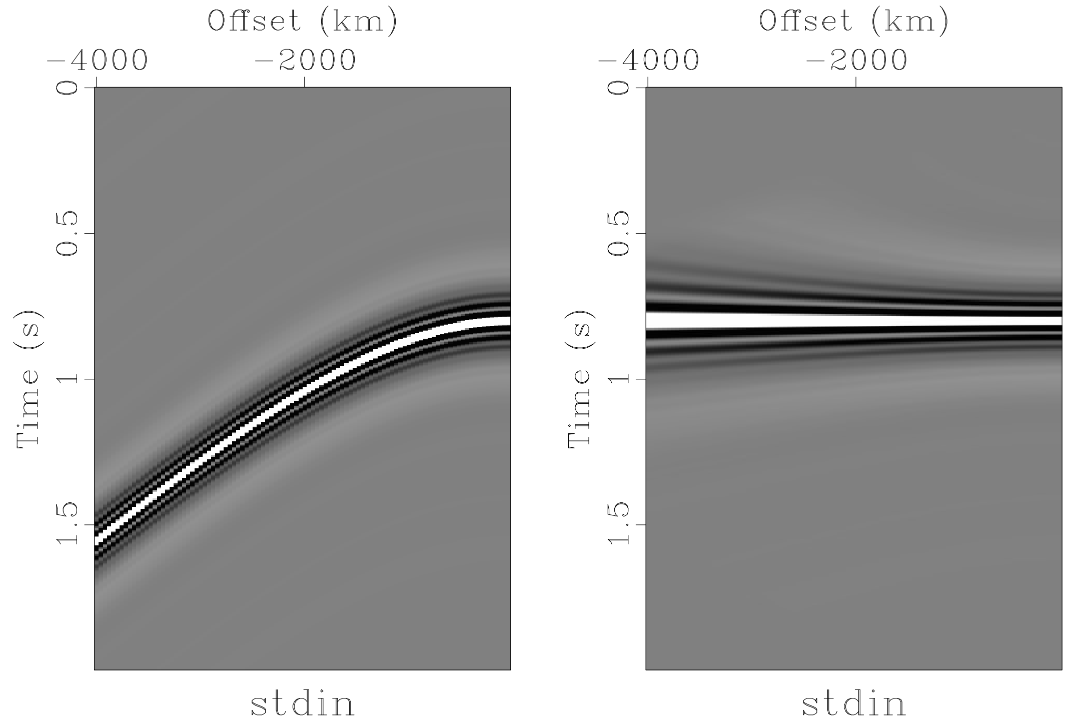

Figure 5.3: H
z bandpass wavelet; Right: result of NMO correction]Left: idealized event created with `suaddevent`, with [2,5,25,35] Hz bandpass wavelet; Right: result of NMO correction

5.9 Suggested Projects

REFERENCES

Bleistein, N., J. Cohen, and J. Stockwell, 2001, Mathematics of multidimensional seismic imaging, migration, and inversion: Springer.

Gel'fand, I., and G. Shilov, 1958, Generalized Functions: volume 1: Academic Press.

Guillemin, V., and S. Sternberg, 1979, Geometric Asymptotics: American Mathematical Society.

Harlan, W., 2014, Wavelet balancing for residual moveouts: Geophysics, **79**, V217–V225.

Robinson, E., and S. Treitel, 1980, Geophysical signal analysis: Prentice-Hall.

Symes, W., 1998, Mathematics of reflection seismology. (available at http://www.trip.caam.rice.edu).

——–, 1999, All stationary points of differential semblance are asymptotic global minimizers: layered acoustics: Technical Report 99-09, Department of Computational and Applied Mathematics, Rice University, Houston, Texas, USA.

Chapter 6

Post-stack Migration

6.1 Introduction

The term “migration”, as used in exploration geophysics, has a complicated origin. The reader should consult some of the standard texts on the subject, such as Dobrin and Savit (1988), or Sheriff and Geldart (1995), for an account of this history. For these notes, the term “migration” is a synonym for “imaging” or “image formation”, that is, a process of creating a subsurface image from data.

This chapter gives an idiosyncratic overview of post-stack migration, that is to say, creation of an image from the stack. The reader has already seen examples in Chapter 2. The term “post-stack migration” refers to techniques that convert zero-offset data to subsurface images. Since zero-offset data is seldom acquired, it is important that the stack and zero-offset data are roughly equivalent. I will explain this equivalence, and derive the *exploding reflector* model of zero-offset data. In terms of this model, I will pose a simple linear inverse problem for recovery of the square velocity (or bulk modulus) perturbation from the stacked data, by minimizing the mean square misfit between the stack and its exploding reflector prediction. From the various approximations and replacements necessary to get to this point, it will be clear that the field extracted from the data by solving this inverse problem cannot be expected to represent the actual values of a physical parameter - but perhaps can be expected to show the same geometric structure, that is, to be an *image* of the subsurface. Like all remotely reasonable formulations of the seismic inverse problem, this one is so large computationally that only iterative solution methods are remotely feasible. I introduce the *Conjugate Gradient* algorithm for iterative solution of linear systems, and apply it to the exploding reflector inverse problem. The first iterate, a multiple of the gradient of the mean squares error, is already a reasonably good image. I will explain why this is to be expected. Computation of the gradient is essentially equivalent to applying the *adjoint* or transpose of the simulation operator. I introduce the *adjoint state method* for computation of this adjoint, which is a verion of *reverse time migaration*, introduced into seismology in the early 80's and a key ingredient in many advanced processing methods.

6.2 Zero-Offset Modeling

Recall that in case of simple ray geometry (unique rays connecting sources and receivers to scattering points), geometric optics approximates the Born approximation, equation 5.8, as:

$$D\mathcal{F}[v^2]\delta v^2(\mathbf{x}_r,t;\mathbf{x}_s) \approx$$

$$= w *_t \frac{\partial^2}{\partial t^2} \int dx\, a(\mathbf{x}_r,\mathbf{x})a(\mathbf{x}_s,\mathbf{x})\delta(t-\tau(\mathbf{x}_r,\mathbf{x})-\tau(\mathbf{x}_s,\mathbf{x}))\frac{\delta v^2}{v^2}(\mathbf{x}) \tag{6.1}$$

"Zero offset" means that $\mathbf{x}_s = \mathbf{x}_r = \mathbf{x}_m$. So the zero-offset trace at receiver location $\mathbf{x}_r$ is

$$D\mathcal{F}_{\text{ZO}}[v^2]\delta v^2(\mathbf{x}_r,t) \approx$$

$$= w *_t \frac{\partial^2}{\partial t^2} \int dx\, a^2(\mathbf{x}_r,\mathbf{x})\delta(t-2\tau(\mathbf{x}_r,\mathbf{x}))\frac{\delta v^2}{v^2}(\mathbf{x}) \tag{6.2}$$

Recall that τ is a solution of the eikonal equation 4.7, written concisely as $v(\mathbf{x})|\nabla\tau(\mathbf{x},\mathbf{x}_r)| = 1$. Obviously $2\tau(\mathbf{x},\mathbf{x}_r)$ solves the same eikonal equation with the velocity $v/2$: "if you go half as fast, it takes twice as long to get there".

I claim that in fact the quantity $a\delta(t-2\tau)$ is the most singular (high frequency asymptotic) part of the Green's function $G_{v/2}(\mathbf{x},t;\mathbf{x}_r)$ for the velocity $v/2$, with "source" at $\mathbf{x}_r$, up to a factor of 2. Specifically, the relationship is

$$G_{v/2}(\mathbf{x},t;\mathbf{x}_r) = \frac{1}{2}G_v\left(\mathbf{x},\frac{t}{2};\mathbf{x}_r\right). \tag{6.3}$$

To see this, substitute the right-hand side of 6.3 in the wave equation for velocity field $v/2$:

$$\left(\frac{1}{\left(\frac{v}{2}(\mathbf{x})\right)^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\frac{1}{2}G_v\left(\mathbf{x},\frac{t}{2};\mathbf{x}_r\right)$$

$$= \left(\frac{4}{v(\mathbf{x})^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\frac{1}{2}G_v\left(\mathbf{x},\frac{t}{2};\mathbf{x}_r\right)$$

$$= \frac{1}{2}\left(\frac{1}{v(\mathbf{x})^2}\frac{\partial^2 G_v}{\partial t^2} - \nabla^2 G_v\right)\left(\mathbf{x},\frac{t}{2};\mathbf{x}_r\right)$$

$$= \frac{1}{2}\delta(\mathbf{x}-\mathbf{x}_r)\delta\left(\frac{t}{2}\right) = \delta(\mathbf{x}-\mathbf{x}_r)\delta(t). \tag{6.4}$$

The last step in the sequence of equations 6.4 is the identity $\delta(t) = c\delta(ct)$, which follows from the definition of the delta function and the substitution rule for integrals.

Equation 6.4 shows that the right-hand side of 6.3 solves the defining wave equation of the Green's function $G_{v/2}$. Since both sides are causal, 6.3 is established.

The geometric optics computation from Chapter 5 shows that the leading singularity of $G_{v/2}$ takes the form

$$G_{v/2}(\mathbf{x},t;\mathbf{x}_r) \approx a_{v/2}(\mathbf{x},\mathbf{x}_r)\delta(t-2\tau(\mathbf{x},\mathbf{x}_r)) = \frac{1}{2}a_{v/2}(\mathbf{x},\mathbf{x}_r)\delta\left(\frac{t}{2}-\tau(\mathbf{x},\mathbf{x}_r)\right) \tag{6.5}$$

However the same geometric optics computation shows that

$$\frac{1}{2}G_v\left(\mathbf{x},\frac{t}{2};\mathbf{x}_r\right) \approx \frac{1}{2}a_v(\mathbf{x},\mathbf{x}_r)\delta\left(\frac{t}{2}-\tau(\mathbf{x},\mathbf{x}_r)\right) = a_v(\mathbf{x},\mathbf{x}_r)\delta(t-2\tau(\mathbf{x},\mathbf{x}_r)) \tag{6.6}$$

Combining equations 6.3, 6.5, and 6.6, conclude that $a_{v/2} = a_v$, and that the integrand in the RHS of equation 6.2 contains exactly the leading order approximation to the Green's function $G_{v/2}$:

$$D\mathcal{F}_{\mathrm{ZO}}[v^2]\delta v^2(\mathbf{x}_r,t) \approx w *_t \frac{\partial^2}{\partial t^2}\int dx\, a_v(\mathbf{x},\mathbf{x}_r)G_{v/2}(\mathbf{x},t;\mathbf{x}_r)\frac{\delta v^2}{v^2}(\mathbf{x})$$

$$= w *_t \frac{\partial^2}{\partial t^2}\int dt' \int dx\, a(\mathbf{x},\mathbf{x}_r)G_{v/2}(\mathbf{x}_r,t-t';\mathbf{x})\frac{\delta v^2}{v^2}(\mathbf{x})\delta(t')$$

$$= \int dt' \int dx\, a(\mathbf{x},\mathbf{x}_r)G_{v/2}(\mathbf{x}_r,t-t';\mathbf{x})\frac{\delta v^2}{v^2}(\mathbf{x})\frac{d^2w}{dt^2}(t'). \tag{6.7}$$

Equation 6.7 unfortunately does *not* expresses the zero-offset data gather as sampling the solution of an acoustic wave equation, because of the extra amplitude factor under the integral sign. The very creative step from zero-offset to exploding reflector modeling consists dropping a and the factor $1/v^2\partial^2/\partial t^2$. Both are a and $1/v^2$ positive and (under standing assumptions) slowly varying on the wavelength scale, so that the overall effect ought to be scaling oscillatory reflection events without modifying their local relative amplitudes or locations. The second time derivative modifies the phase also without modifying locations and orientations of events, and combines with the other factors to yield an overall dimensionless change. A complete mathematical justification requires tools developed later in these notes.

Accepting these changes, re-interpret zero-offset data as represented by the *exploding reflector model*

$$D\mathcal{F}_{\rm ZO}[v^2]\delta v^2(\mathbf{x}_r,t) \approx \int dt' \int dx\, G_{v/2}(\mathbf{x}_r,t-t';\mathbf{x})\delta v^2(\mathbf{x})w(t'). \tag{6.8}$$

The Green's function $G_{v/2}$ is the solution operator of the initial value problem

$$\begin{aligned} \frac{\partial^2 p}{\partial t^2} - \frac{v^2}{4}\nabla^2 p &= f \\ p &= 0, t << 0 \end{aligned} \tag{6.9}$$

$D\mathcal{F}_{\rm ZO}$ applies this solution operator to the right-hand side

$$f(\mathbf{x},t) = \frac{\delta v^2}{(}\mathbf{x})w(t) \tag{6.10}$$

This observation explains the term *exploding reflector model* for the representation 6.7: the reflectors (δv^2) “explode” with time dependence $w(t)$.

6.3 Exploding Reflectors in Pressure-Velocity Acoustics

The pressure-velocity form of acoustics relates pressure p and the particle velocity vector field $\bar{v}$ through a pair of time-dependent PDEs:

$$\begin{aligned} \frac{\partial p}{\partial t} &= -\kappa(\nabla\cdot\bar{v}+g) \\ \frac{\partial \bar{v}}{\partial t} &= -\beta\nabla p \end{aligned} \tag{6.11}$$

In this system, $\kappa = \rho v^2$ where ρ is the material density, and $\beta = 1/\rho$ is the *buoyancy*.

Each equation in 6.11 has clear physical meaning. The first is the constituitve law of linear acoustics: the rate of change of pressure is proportional to the rat of change of volume (the divergence of the velocity field), with the constant of proportionality being the bulk modulus. The the energy source is represented as a defect (g) in this constitutive law: more specifically, a defect in the infinitesimal volume rate field. The second equation is Newton's law.

To make this system equivalent to the constant density second order equation 6.9, simply assume $\beta = \rho = 1$ and set

$$\bar{v} = \nabla\int_{-\infty}^{t} p,\ \kappa\frac{\partial g}{\partial t} = f.$$

In terms of the pressure-velocity form of acoustics, the exploding reflector model from the previous section is expressed as

$$D\mathcal{F}_{\mathrm{ZO}}[v^2]\delta v^2(\mathbf{x}_r, t) \approx p(x_r, t),$$

where $(p, \bar{v})$ solve the system 6.11 with $\kappa = v^2/4$, that is, the physical bulk modulus is divided by 4, and the right-hand-side

$$g(\mathbf{x}, t) = \frac{\delta v^2(\mathbf{x})}{v^2(\mathbf{x})} w_1(t). \tag{6.12}$$

where w_1 is the indefinite integral of w.

At this point, I will take possibly unfair advantage of a point I have not raised until now - we have actually no knowledge of the wavelet $w(t)$. Of course, whatever it is (if indeed the isotropic point radiator is even a fair model of the seismic source) it can be changed by filtering. Therefore I will from now on assume that w is the derivative of an approximate delta, filtered appropriately and with appropriate units, and therefore that w_1 is a filtered delta function. For convenience, I will even ignore the filter. Therefore the model of the right-hand side in the exploding reflector model becomes

$$g(\mathbf{x}, t) = \frac{\delta v^2(\mathbf{x})}{v^2(\mathbf{x})} \delta(t). \tag{6.13}$$

6.4 An Example

In Chapter 3, I used the stack as input to various forms of migration to produce images in both time and depth, without explaining why the stack is appropriate input (or, indeed, what the migration algorithms did). This chapter will clear up some of those mysteries. However, as I have done in earlier chapters, I will begin by building an example based on migration of a stack.

I construct an example similar to those in Chapter 3, by using the v^2 model estimated by flattening NMO-corrected gathers for the Viking Graben data (Figure 6.1), and for the perturbation δv^2 the PSPI post-stack depth migrated image Figure 6.3. The input data for the PSPI image is shown in Figure 6.2: it is the result of applying AGC to the stack of the NMO-corrected data from Chapter 3, then applying a (2,5,20,25) Hz bandpass filter. The additinoal low-pass filter allows accurate finite difference calculations with a coarser grid than is possible with the original data, and is used only to make the computations a bit less time-consuming.

The Viking Graben survey does not contain zero-offset data. The closest approximation is the near offset, at $h = -262$ m. With AGC to make the events more visible, this

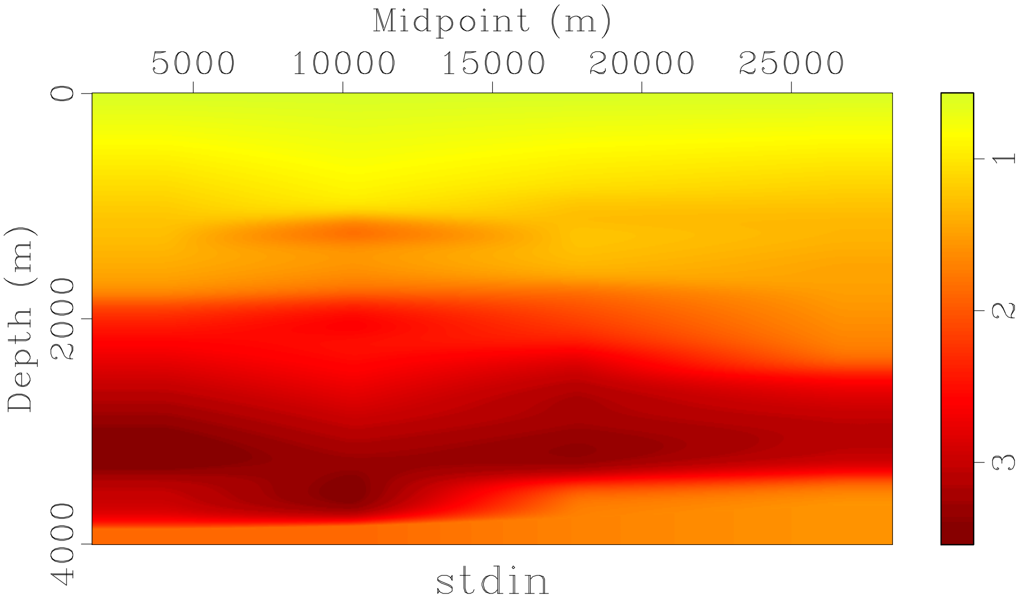

Figure 6.1: Scaled bulk modulus $v^2/4$ ($\rho = 1$) for exploding reflector simulation: obtained from Viking Graben data by flattening NMO gathers, as in Chapter 2.

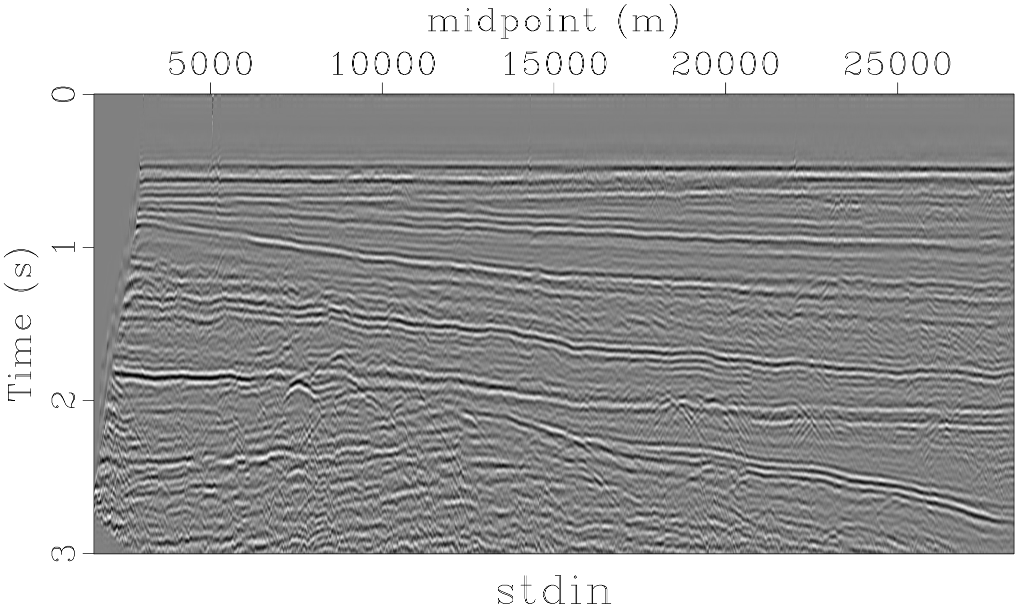

Figure 6.2: AGC'd and filtered stack from Chapter 3. Filter is (2, 5, 20, 25) Hz bandpass.

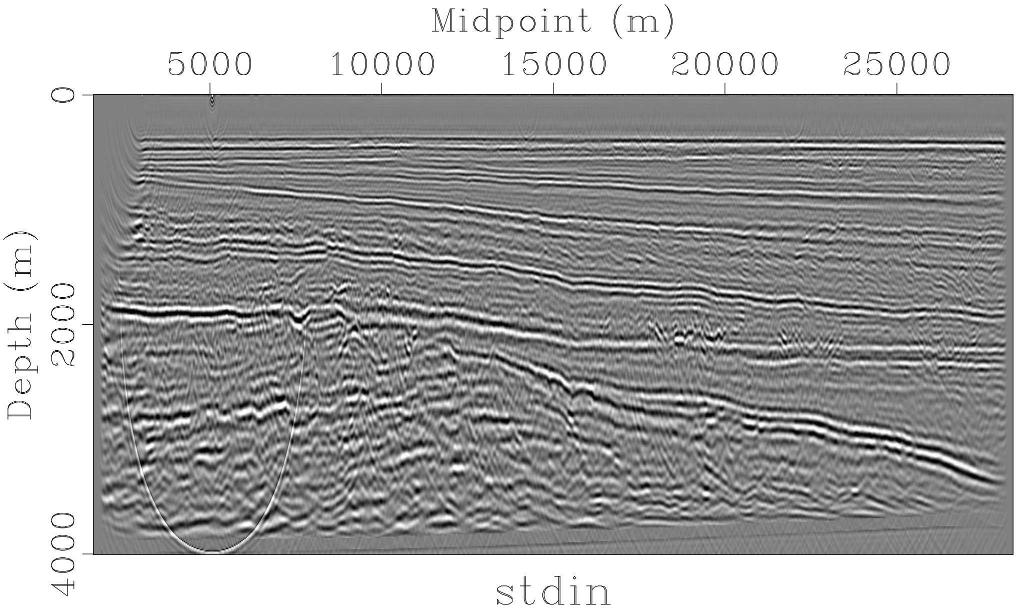

Figure 6.3: Bulk modulus perturbation δv^2 for exploding reflector simulation, obtained from Viking Graben data by Gazdag PSPI depth migration of filtered stack (Figure 6.2) using velocity model of Figure 6.1.

gather appears as Figure 6.4. Comparison with the result of the exploding reflector simulation, Figure **??**, shows general agreement of positioning and orientation of events. I have plotted both field data and simulated data after AGC, to emphasize the similarity in apparent structure.

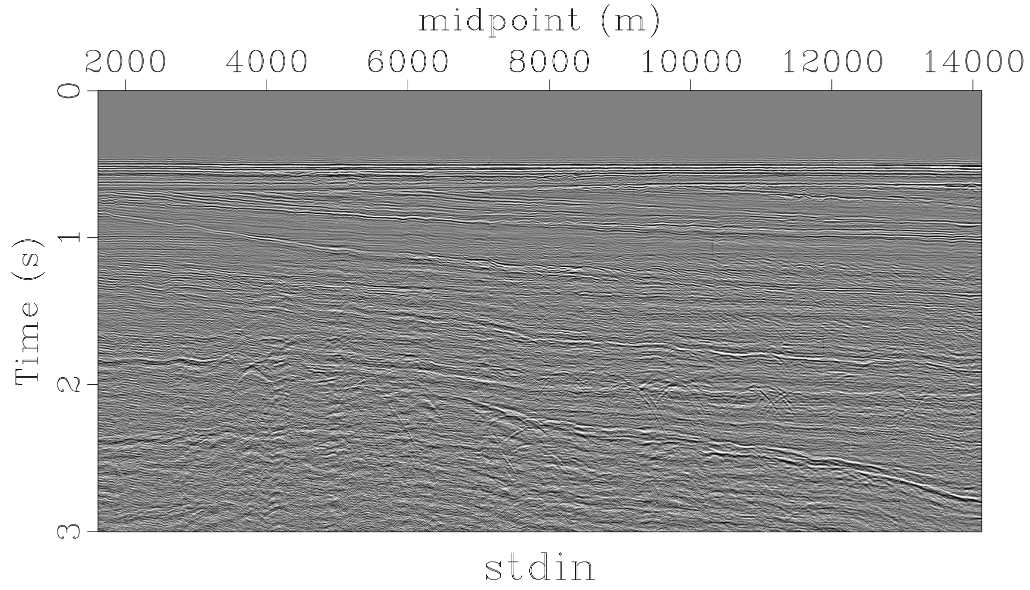

Figure 6.4: Near offset section, after AGC.

6.5 The Stack As a Stand-In for ZO Data

Recall from the last chapter that the convolutional model predicts a simple relation between perfectly deconvolved CMP gathers $d(x_m,y_m,h,t)$ after NMO correction using event traveltime $T(t_0,h)$ as function of zero offset time t_0 and (half) offset h, and a reflectivity function $R(t_0)$:

$$d(x_m,y_m,h,\bar{T}(h,t_0)) \approx \bar{A}(h,t_0)\frac{d\bar{R}}{dt_0}(x_m,y_m,t_0) \tag{6.14}$$

SInce $\bar{T}(0,t_0)=t_0$, this relation implies that for zero offset,

$$d(x_m,y_m,0,t_0) \approx \bar{A}(0,t_0)\frac{d\bar{R}}{dt_0}(x_m,y_m,t_0). \tag{6.15}$$

It also yields an approximation to the stack:

$$d_{\text{stack}}(x_m,y_m,t_0) = \int dh\, d(x_m,y_m,h,\bar{T}(h,t_0)) \approx \bar{A}_{\text{stack}}(t_0)\frac{d\bar{R}}{dt_0}(x_m,y_m,t_0). \tag{6.16}$$

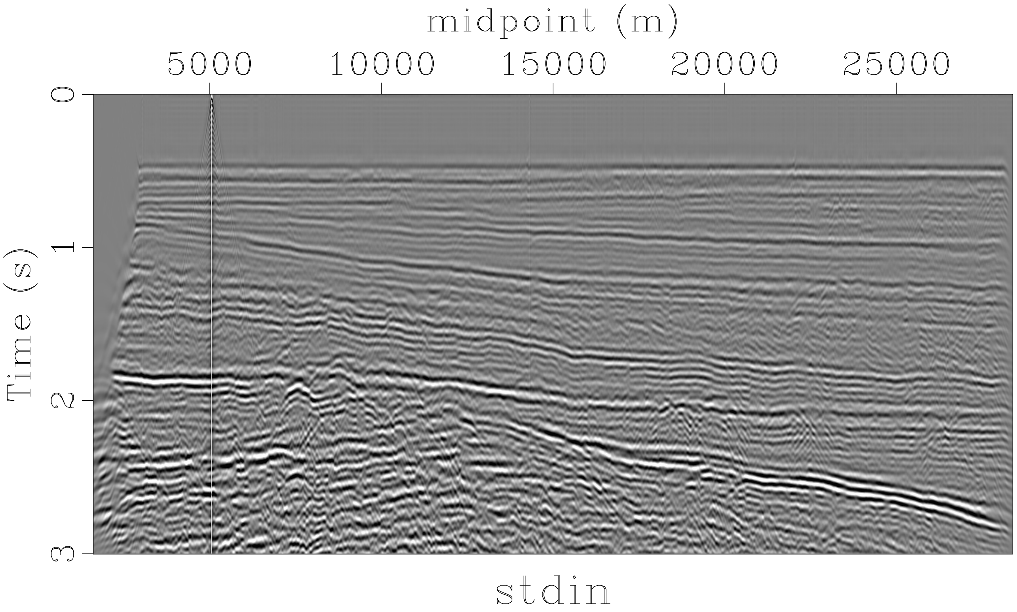

Figure 6.5: Exploding reflector simulated zero offset data based on PSPI depth image (Figure 6.2) as input reflectivity (right-hand side in equation 6.13, and scaled NMO-derived velocity model (Figure 6.1).

where the stacked amplitude is defined by

$$\bar{A}_{\text{stack}}(t_0) = \int dh \bar{A}(h, t_0). \tag{6.17}$$

Comparison of 6.15 and 6.16 shows that the stack is proportional to zero-offset data by the ratio of the zero-offset and stacked amplitudes, both of which are positive slow-varying functions of t_0. Thus if the deconvolved convolutional model 6.14 is to be believed, the stack has precisely the same *locations in time* of high-frequency events as zero-offset data, so that as input to imaging for structure, the two are interchangeable.

Of course, the deconvolved convolutional model is not entirely correct, as discussed in the last chapter, for (at least) three reasons. First, since perfect deconvolution is impossible, NMO stretch will decrease the frequency content of far offsets, hence lower decrease the energy content of the stack at high frequencies, hence its apparent resolution. Second, imperfectly defined NMO velocity will result in some averaging of the reflectivity, again depressing frequency content. Both of these effects are observable in the stacked data, shown in Figure 6.7 (also with AGC, but without low-pass filter) for comparison with the near-offset section (Figure 6.6, repeated here for the reader's convenience).

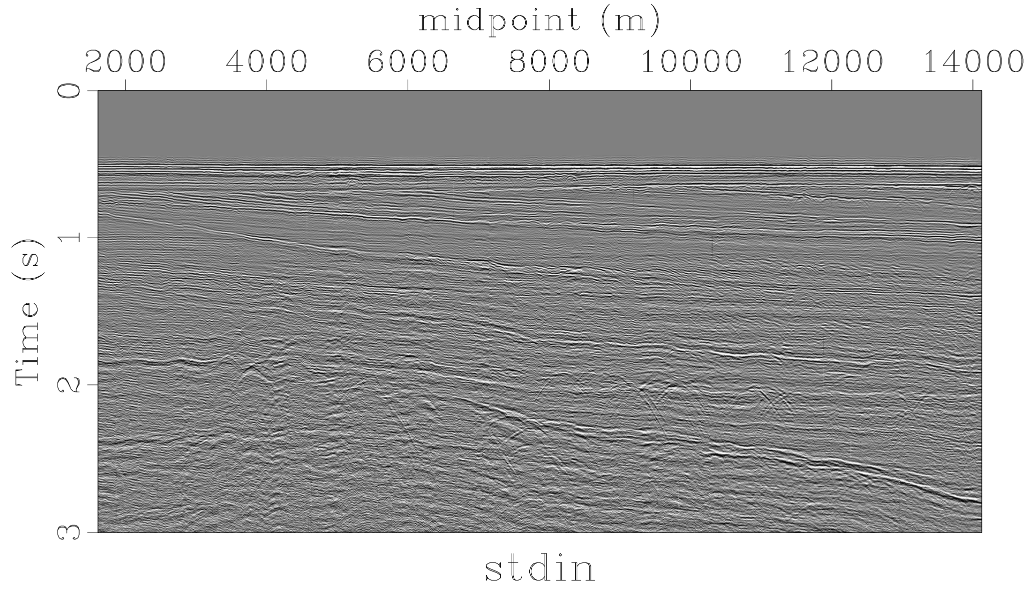

Figure 6.6: Near offset section, after AGC.

Finally, of course the convolutional model in any form is the result of several layers of approximation - linearization, high-frequency asymptotics, the "locally layered" assumption. Nontheless, in fact the stack is even more similar to the modeled exploding reflector data (Figure **??**) than is the zero offset section.

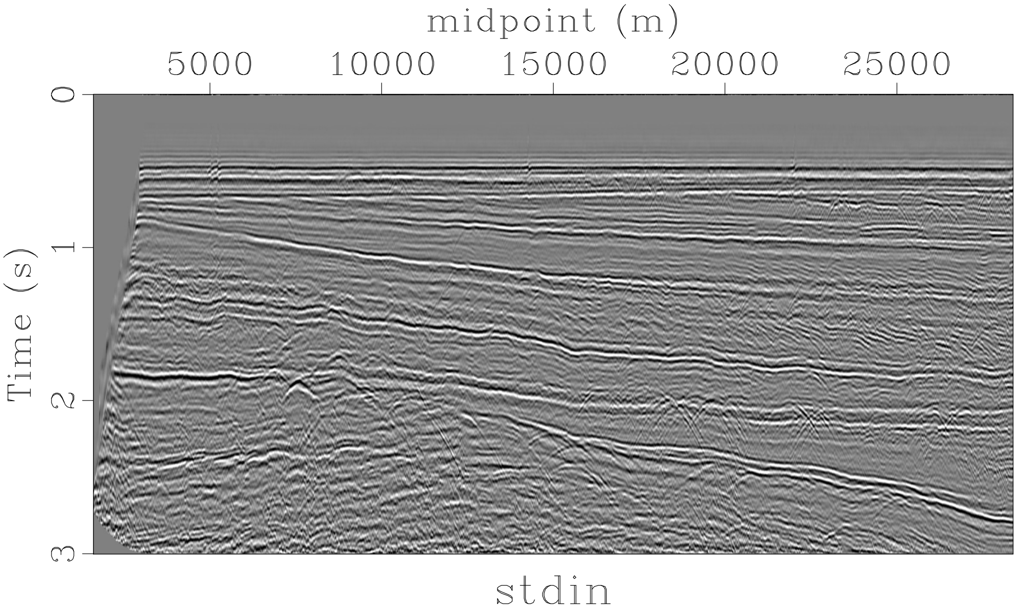

Figure 6.7: Stack of NMO-corrected Viking Graben data, AGC applied.

6.6 A Linear Inverse Problem

Seismic imaging is the construction of an image of the subsurface from seismic data - a statement which begs the question, “what is an image?”. [One could give the Justice Potter Stewart answer, of course - “I know it when I see it” - but it’s possible to do better.] To start with, I will pose the imaging problem in the simplest possible way: given the data, recover parameters that predict it. For the constant density acoustic model, “parameters” means the squared velocity field $v^2(\mathbf{x})$, and in principle the source model as well. Admitting the Born approximation, “parameters” means the source and the pair of squared velocity $v^2(\mathbf{x})$ and its perturbation $\delta v^2(\mathbf{x})$, with the former slowly varying on the wavelength scale hence transparent thanks to geometric optics, and the latter oscillatory and accounting for reflection dynamics. Recovery of these parameter fields must surely encompass producing their image, and therefore an image of earth structure, to the extent that the modeling assumptions underlying the formulation of the problem are valid.

I will specialize this parameter recovery or *inverse* problem to the zero-offset (or stacked) data case with the exploding reflector approximation replacing the acoustic model. As it turns out, this combination of data and model does not permit all parameters to be determined. If the source model is fixed as in equation 6.13, and the squared velocity is regarded as data, rather than as a parameter field to be determined, it turns out to be possible to estimate the remaining parameter, the square velocity perturbation $\delta v^2(\mathbf{x})$, so that the data is quite well predicted. The predicted exploding reflector data is linear in $\delta v^2(\mathbf{x})$, as the latter is essentially the right-hand side of a wave equation. Therefore the recovery of $\delta v^2(\mathbf{x})$ from $v^(\mathbf{x})$ and $d_{\text{stack}}(x_m, y_m, t_0)$ is a *linear inverse problem*. Formally, the problem amounts to the (approximate) solution of the linear system

$$D\mathcal{F}_{\text{ZO}}[v^2]\delta v^2(x_m, y_m, t_0) \approx d_{\text{stack}}(x_m, y_m, t_0). \tag{6.18}$$

Because of the very many approximations and unrealistic physical assumptions made in arriving at the system 6.18, actually solving it (even in the approximate sense that I will develop in this chapter) is overkill - asking too much of the data and modeling assumptions. The solution can't be confused with the "real" short-scale perturbations in compressional wave velocity occuring in the subsurface beneath the Norwegian North Sea. Nonetheless, one might hope first that some evidence of subsurface structure would emerge, and perhaps to see how the calculations might be short-circuited to provide just such image information without actually bothering to fit the data. Both hopes turn out to be realistic.

It is very likely that equation 6.18 cannot be solved by any choice of δv^2, that is, the data cannot be exactly fit, for example because the domain within which δv^2 is allowed to vary does not predict the full range of midpoints or times present in the stack. An example of exactly this type of behaviour appears below. Therefore it is natural to converte 6.18 to a *best fit* problem. The mean square error is a misfit measure with pleasant mathematical and computational properties. It also enjoys some physical justification, see Santosa and Symes (2000). Thus seek δv^2 to minimize (approximately)

$$J[\delta v^2] = \frac{1}{2}\sum_{x_m, y_m}\int dt_0\, |(D\mathcal{F}_{\text{ZO}}[v^2]\delta v^2)(x_m, y_m, t_0) - d_{\text{stack}}(x_m, y_m, t_0)|^2. \tag{6.19}$$

Since J is a positive semidefinite quadratic form, all of its stationary points are global minimizers. A stationary point is a zero of the gradient:

$$\nabla J[\delta v^2] = D\mathcal{F}_{\text{ZO}}[v^2]^T(D\mathcal{F}_{\text{ZO}}[v^2]\delta v^2 - d_{\text{stack}})(x_m, y_m, t_0)) = 0 \tag{6.20}$$

or alternatively in the form of a linear system, the so-called *normal equation*:

$$(D\mathcal{F}_{\text{ZO}}[v^2]^T D\mathcal{F}_{\text{ZO}}[v^2]\delta v^2)(x, y, z) = D\mathcal{F}_{\text{ZO}}[v^2]^T d_{\text{stack}})(x, y, z). \tag{6.21}$$

The transpose, or adjoint, operator $D\mathcal{F}_{\text{ZO}}[v^2]^T$ is the unique operator making the *dot product test* true:

$$\sum_{x_m, y_m}\int dt_0\,(D\mathcal{F}_{\text{ZO}}[v^2]\delta v^2)(x_m, y_m, t_0) d_{\text{stack}}(x_m, y_m, t_0)$$

$$= \int\int\int dx\,dy\,dz\,(D\mathcal{F}_{\rm ZO}[v^2]^T d_{\rm stack})(x,y,z)\delta v^2(x,y,z). \tag{6.22}$$

Note that the dot product test should hold for *any* choice of square velocity perturbation δv^2 and stacked data section $d_{\rm stack}$, not just the ones that you might be currently interested in.

After discretization, say by finite differences as are used in the IWAVE package, the fields δv^2 and $d_{\rm stack}$ become vectors, and the linear operators $D\mathcal{F}_{\rm ZO}[v^2]$, $D\mathcal{F}_{\rm ZO}[v^2]^T$, and $D\mathcal{F}_{\rm ZO}[v^2]^T D\mathcal{F}_{\rm ZO}[v^2]$ become matrices. The matrix of $D\mathcal{F}_{\rm ZO}[v^2]^T$ is the transpose of the matrix of $D\mathcal{F}_{\rm ZO}[v^2]$, and the matrix of the *normal operator* $D\mathcal{F}_{\rm ZO}[v^2]^T D\mathcal{F}_{\rm ZO}[v^2]$ is symmetric positive (semi-)definite.

The integrals in the equation 6.22 become scaled versions of the dot product, scaled by the cell volume on each side. On the left, I have continued to write a sum over midpoints, rather than integrals over midpoint coordinates, because midpoint coordinates may not be uniformly sampled and devising a quadrature method is difficult. Instead, I presume that the midpoint sampling is sufficiently uniform that I can treat it as perfectly uniform and ignore the midpoint area factor that would go into an approximation to an integral. So the only cell volume on the left hand side is the time step. On the other hand, the spatial fields - δv and the output of the transpose operator - are sampled on a regular finite difference grid, so the integral is well-approximated by the ordinary dot product multiplied by the volume of a grid cell. The consequence is the the results are stable against resampling in t and in x, y, z.

The inversion task boils down to solving (approximately, at least) the normal system 6.21. A first impediment to computational implementation is the size of this system. For the example created in this book from the Viking Graben survey, $d_{\rm stack}$ has 1.6 $\times 10^6$ samples, and δv^2 about 0.5 $\times 10^6$. That is, there are more equations than unknowns, by about a factor of 3: the system 6.18 is overdetermined, and is very unlikely to have a literal solution, justifying the least squares approach. Storing the matrix of $D\mathcal{F}_{\rm ZO}[v^2]$ requires $O(10^{12})$ words of memory - not inconceivable for vintage 2017 high-end computing equipment, but a LOT of storage nonetheless. The normal marix is smaller, $O(10^{11})$ words, especially if one stores only the upper or lower triangle, taking advantage of symmetry. It is still an unpleasantly large amount of data, and the vast number of loads and stores necessary to manipulate it make efficient implementation difficult. Finally, Gaussian elimination requires $O(10^{18})$ floating point operations, or more than 100 days at 100 Gflops - and this for a rather academic 2D problem.

6.7 Conjugate Gradient Iteration

Remarkably, it is possible to solve the least squares problem described in the last section (effectively, the normal equations 6.21) to an acceptable level of approximation, by means of an efficient *iterative* method, in a tiny fraction of the time estimated above for Gaussian elimination, and using very modest computing resources - for the examples to be shown in the next section, a few minutes on a vintage 2015 Apple laptop with a single thread of execution.

The efficient iterative method is Conjugate Gradient (CG) iteration (Nocedal and Wright, 1999; Golub and van Loan, 2012). For completeness, I describe the general form of this computation, in the context of the least squares problem defined in the last section. To make the notation tractable, I introduce the two abbreviations:

$$\begin{aligned} F &\leftarrow D\mathcal{F}_{\rm ZO}[v^2] \\ r &\leftarrow \frac{\delta v^2}{v^2} \end{aligned} \tag{6.23}$$

That is, v^2 is fixed for the course of the discussion, so suppress it from the notation, and abbreviate the notation for derivative. Also, the relative perturbation in v^2 is the source of reflection signal in the exploding reflector model: accordingly, abbreviate it as the reflectivity, denoted r, and regard the predicted signal as being a function of the relative perturbation rather than the perturbation itself - since v^2 is regarded as fixed, the two points of view are equivalent. Thus $r \mapsto Fr$ is a linear map, which would be represented after discretization by a (huge!) matrix if you were to be so foolish as to compute all of its entries. Instead, a perfectly good approach to computing the matrix-vector product Fr is available: use a finite difference method to approximately solve the exploding reflector problem 6.11, 6.12, which in the notation just introduced take the form

$$\begin{aligned} \frac{\partial p}{\partial t} &= -\kappa(\nabla \cdot \bar{v} - r\delta(t)) \\ \frac{\partial \bar{v}}{\partial t} &= -\beta \nabla p \\ p, \bar{v} &= 0, t < 0 \end{aligned} \tag{6.24}$$

The CG algorithm needs only the matrix-vector product, not the matrix entries themselves:

- Initialize:
  - $r_0 = 0$
  - $e_0 = d$

  - $s_0 = F^T d$
  - $g_0 = s_0$
  - $k = 0$
- Repeat:
  - $\alpha_k = \dfrac{\langle s_k, s_k \rangle}{\langle F s_k, F s_k \rangle}$
  - $r_{k+1} = r_k + \alpha_k s_k$
  - $e_{k+1} = e_k - \alpha_k F s_k$
  - $g_{k+1} = g_k - \alpha_k F^T F s_k$
  - $\beta_{k+1} = \dfrac{\langle g_{k+1}, g_{k+1} \rangle}{\langle g_k, g_k \rangle}$
  - $s_{k+1} = g_{k+1} + \beta_{k+1} s_k$
  - $k = k + 1$
- Until $\|e_k\|$ sufficiently small, or max iteration count exceeded

These relations follow from the statement of the algorithm:

$$\begin{aligned} e_k &= d - F r_k \\ g_k &= F^T (d - F r_k) \end{aligned} \tag{6.25}$$

That is, e_k is the data error (or residual) at the kth iteration, and g_k is the gradient (also the normal residual). s_k is the search direction, α_k is the correct multiple of s_k to attain the minimum of the quadratic objective on the line through r_k in the direction s_k. The justification for the direction update parameter (β_k) is more difficult to explain - see the cited references.

The corner brackets are the scaled dot products mentioned in the previous section: for instance, for data sections d_1 and d_2,

$$\langle d_1, d_2 \rangle = \sum_{x_m, y_m} \sum_{t_0} \Delta t_0 \, d_1(x_m, y_m, t_0) d_2(x_m, y_m, t_0) \tag{6.26}$$

and for depth fields $g_1(x,y,z), g_2(x,y,z)$

$$\langle g_1, g_2 \rangle = \sum_{x,y,z} \Delta x \Delta y \Delta z g_1(x,y,z) g_2(x,y,z) \tag{6.27}$$

For any inner product, the corresponding norm is the square root of the inner product ssquare: $\|g\| = \sqrt{\langle g, g \rangle}$.

Iteration	Residual Norm	Gradient Norm
1	2.18368799e+03	3.05936885e+00
2	9.36336853e+02	1.13921440e+00

Table 6.1: Conjugate Gradient first iteration.

Conjugate Gradient iteration has many fascinating properties, explained in the books cited earlier. I will mention just one: in its application to least squares problems as in the algorithm explained above, the residual norm $\|e_k\|$ decreases monotonically with k, whereas the gradient norm $\|g_k\|$ does not, necessarily.

6.8 Inversion of the Viking Graben Stack

I used CG iteration to solve the inverse problem set out in this chapter, assuming that the exploding reflector r is zero outside of the subrectangle $\{3000m \le x \le 27000m, 400m \le z \le 2800m\}$. The modeling is done in 2D, which means that some of the considerations developed above that depend on the nature of the leading singularity really ought to be changed. I ignored these conflicts, and simply used the system 6.24 without modification.

In order to make sure that all inputs were correct, I first calculated a single iteration of the CG algorithm described in the last section. The RVL implementation of CG outputs key information, as displayed in Table 1. Note that the RVL CG algorithm counts iterations Fortran-style, that is, starting from 1 instead of from 0, C style, as the algorithm pseudo-code in the last section would have it. I wonder why that is.

The first iteration has dropped the data residual norm to 75% of its initial value - not so impressive. The first iterate (r_1) is more interesting: Compare it with the PSPI migration in Figure 6.3). The resemblance is remarkable: in fact, the first iterate is a reasonably accurate *image*, in that it shows the same geometrical structures picked out in grey scale. Note from the CG listing that the first iterate is a multiple of the gradient of the least squares function at $r = 0$. So this observation poses a question:

Why is the gradient an image?

The answer comes in the section after next.

Note that the obvious diffraction hyperbolae in the stack (Figure 6.7) are absent in any of the depth images - collapsing of diffraction hyperbolae is a salient feature of migration (and indeed the first CG iterate is a migration).

Since the setup seemed to be working I went ahead with 6 steps. The residual record

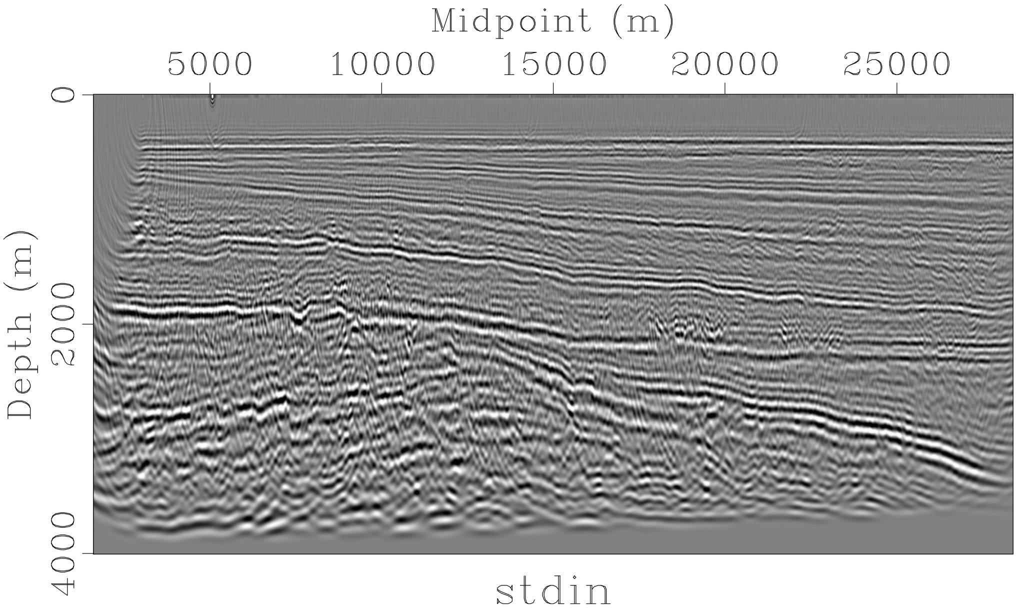

Figure 6.8: Viking Graben ZO inversion: CG iteration 1

appears in Table 2. Note that the next-to-last iteration actually increases the gradient norm, as is possible during CG iteration, but the data residual norm monotonically decreases as it should. The recovered reflectivity r after 6 iterations appears in Figure 6.9, and is somewhat less noisy than the 1 iteration result (Figure 6.8); both are very similar to the PSPI image in Figure 6.3.

The data predicted from this model (Figure 6.10) appears to match fairly well with the target data in Figure 6.2 - the two figures are plotted on the same grey scale. According to the table above, the iteration has reduced the data misfit to less than 15% of its initial size (RMS). Plotting the data residual on the same scale (Figure 6.12) shows how small the remaining error actually is.

The alert reader will have possibly formulated a critical question. I have explained that manipulation of F as a matrix is out of the question, even for this rather tame problem. If F were represented as a matrix, then the meaning of F^T would be clear, as would be its computation (swap the indices!). However, F is not represented as a matrix, but rather its action on a vector (representing an exploding reflecetor field) results from numerical solution of a wave equation. Accordingly,

How can F^T be computed without accessing its matrix?

Iteration	Residual Norm	Gradient Norm
1	2.18368799e+03	3.05936885e+00
2	9.36336853e+02	1.13921440e+00
3	5.54742188e+02	3.24299669e+00
4	5.08972382e+02	8.39157641e-01
5	4.04111511e+02	4.64890152e-01
6	3.84719269e+02	1.62544572e+00
7	3.25789581e+02	5.23918450e-01

Table 6.2: Residual and Gradient norms, six Conjugate Gradient iterations.

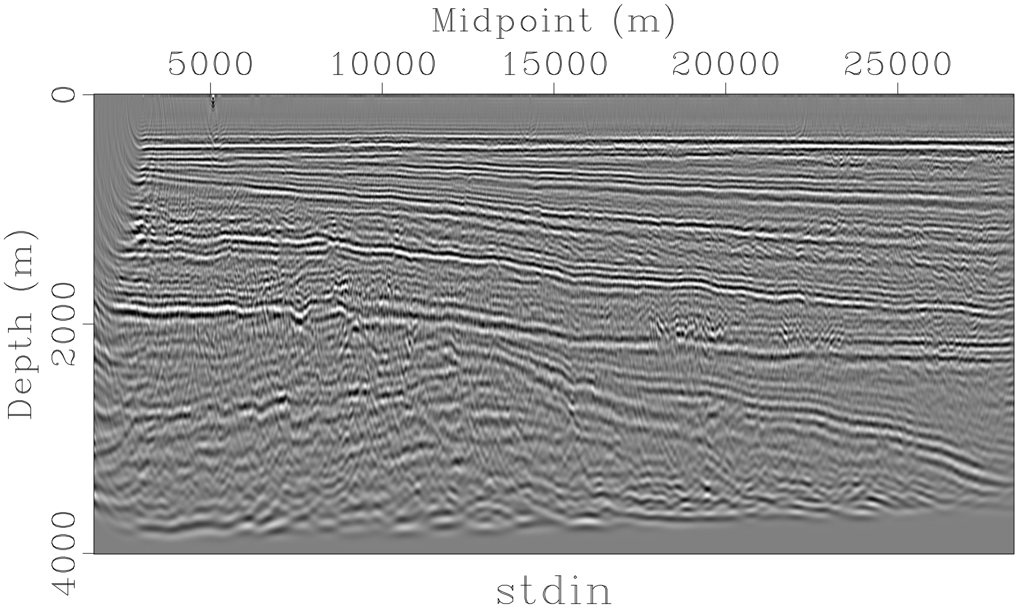

Figure 6.9: Viking Graben ZO inversion: CG iteration 6

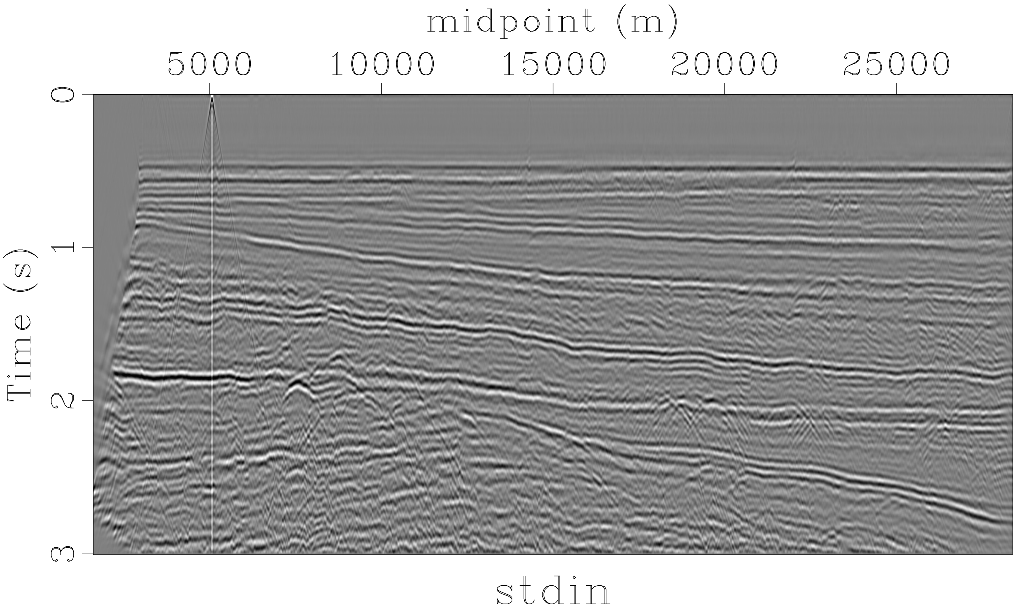

Figure 6.10: Simulated data from reflectivity produced by 6 CG iterations (Figure 6.9).

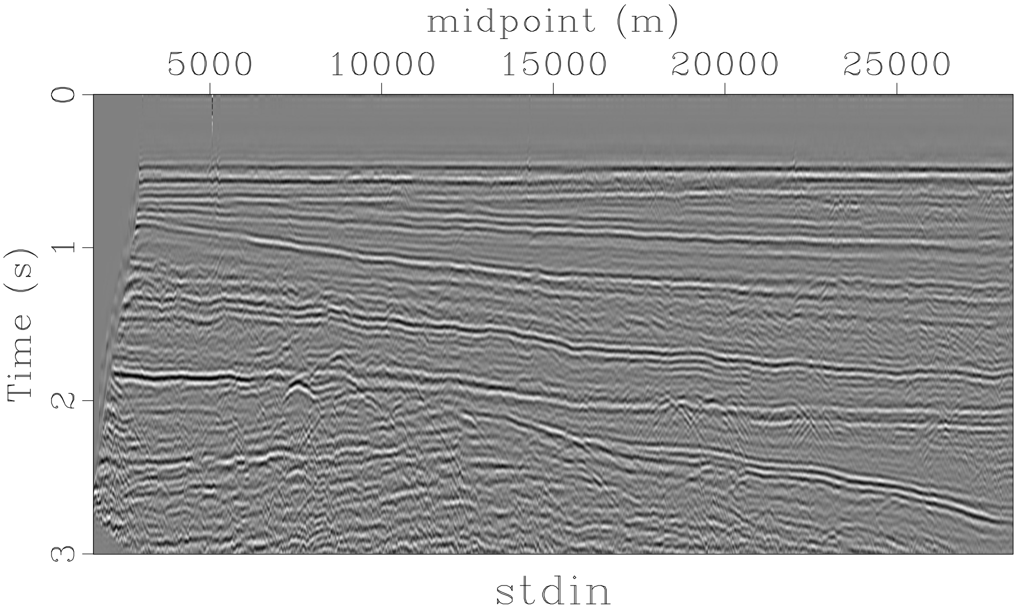

Figure 6.11: Target data for inversion: duplicates Figure 6.2, filtered AGC'd NMO-corrected stack of Viking Graben line.

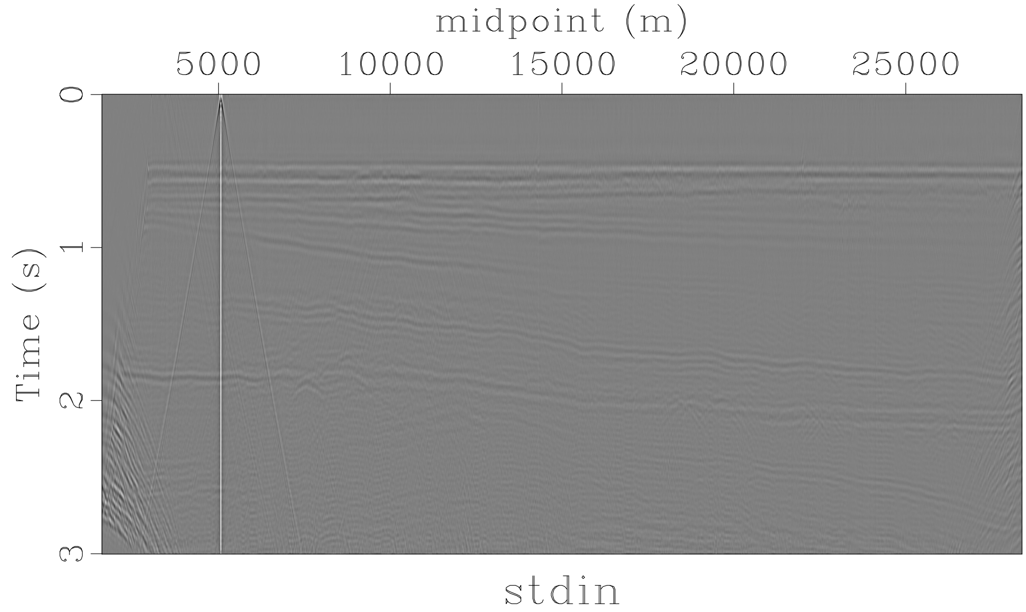

Figure 6.12: Data residual, Viking Graben ZO inversion: CG iteration 6

6.9 Computation of the Adjoint: Reverse Time Migration

The transpose or adjoint operator is *defined* by the dot product test: in the notation introduced in the last two sections,

$$\sum_{x_m,y_m}\int dt_0\,(Fr)(x_m,y_m,t_0)d(x_m,y_m,t_0)=\int\int\int dx\,dy\,dz\,r(x,y,z)(F^Td)(x,y,z). \tag{6.28}$$

From the definition of F and the delta function, the left-hand side of equation 6.28 is the same as

$$=\sum_{x_m,y_m}\int dt_0\,p(x_m,y_m,,z_m,t)d(x_m,y_m,t)$$

$$=\int\int\int\int dx\,dy\,dz\,dt\,p(x,y,z,t)\left(\sum_{x_m,y_m}d(x_m,y_m,t)\delta(x-x_m)\delta(y-y_m)\delta(z-z_m)\right) \tag{6.29}$$

In this last version of the dot product test, the data has been converted to a “data source”, a collection of point radiators with the data traces as their time dependence. Note that data is only recorded for a finite time (a few seconds, for exploration data) so make the very important presumption that d vanishes for t larger than a maximum time of recording $t_{\max}$,

If 6.28 can be re-written so that the integrand displays an explicit factor of r, then the $F^T d$ can be read off by inspection - whatever multiplies r must then be $F^T d$. From the system 6.24, r is the right-hand side of an equation satisfied by *derivatives* of $p, \bar{v}$. As it currently stands, p appears *without* derivatives in 6.28. However, derivatives can be "borrowed" from one factor in an integral to another via integration by parts. So if we can replace the "data source" in 6.28 by a suitable combination of derivatives of another field, then we could presumably "borrow" those derivatives to set up the left hand side of the first equation in 6.24 hence introduce r.

Suppose that the factor multiplying p in equation 6.29 were itself the right-hand side of a wave equation, one of a pair similar to 6.24:

$$\begin{aligned}
\frac{\partial q}{\partial t} &= -\kappa\left(\nabla\cdot\bar{w} - \left(\sum_{x_m,y_m} d(x_m,y_m,t)\delta(x-x_m)\delta(y-y_m)\delta(z-z_m)\right)\right)\\
\frac{\partial \bar{w}}{\partial t} &= -\beta\nabla q,\\
q,\bar{w} &= 0,\, t > t_{\max}.
\end{aligned} \tag{6.30}$$

It will become clear shortly why the fields $q, \bar{w}$ should vanish for large positive time, instead of for large negative time as is the case of $p, \bar{v}$. Then the right-hand side of equation 6.29 becomes

$$= \int\int\int\int dx\,dy\,dz\,dt\,p\left(\frac{\frac{1}{\kappa}\partial q}{\partial t} + \nabla\cdot\bar{w}\right)(x,y,z,t)$$

Integrate by parts in all space variables and time to move the derivatives onto p:

$$= -\int\int\int\int dx\,dy\,dz\,dt\left((\frac{1}{\kappa}\frac{\partial p}{\partial t}q + \nabla p\cdot\bar{w}\right)x,y,z,t)$$

There are no boundary terms from integraion by parts in t because p vanishes for large negative time, and q vanishes for large positive time. Since the midpoint (or receiver) locations occupy only a finite volume, and the data vanishes for $t > t_{\max}$ it follows from the finite speed of propagation of acoustic waves that at any fixed time, the fields $q, \bar{w}$ vanish outside of a large ball, so there all the boundary terms in space also vanish. Now you know why the $q, \bar{w}$ were required to vanish for large positive rather than negative time!

Use the equations 6.24 to replace the partial derivatives of p with those of $\bar{v}$:

$$= -\int\int\int\int dx\,dy\,dz\,dt\left((-\nabla\cdot\bar{v} + r\delta(t))q + \frac{1}{\beta}\frac{\partial\bar{v}}{\partial t}\cdot\bar{w}\right)(x,y,z,t)$$

then segregate the term involving r and integrate by parts again:

$$= -\int\int\int\int dx\,dy\,dz\,dt\,(r\delta(t)q)(x,y,z,t)$$

$$+\int\int\int\int dx\,dy\,dz\,dt\,\bar{v}\cdot\left(\nabla q + \frac{1}{\beta}\frac{\partial\bar{w}}{\partial t}\right)$$

The second summand vanishes because of the second equation in the system 6.30, so the end result is

$$= -\int\int\int\int dx\,dy\,dz\,dt\,r(x,y,z)q(x,y,z,0) \tag{6.31}$$

Comparing equations 6.28 and 6.31, conclude that

$$(F^T d)(x,y,z) = -q(x,y,z,0). \tag{6.32}$$

in which q is the first component of the solution of system 6.30.

This remarkable formula is a simple instance of the *adjoint state method*: it computes the adjoint of the modeling operator via solution of another system of wave equations. A discrete version also holds, and that is what is implemented in the IWAVE exploding reflector adjoint calculation.

6.10 Why the Gradient is an Image

Here is a slightly more ambitious statement about the meaning of "image": a function of 2 (or 3) variables is an image of another, if the two share the location and orientation of rapid oscillations. For example, this statement describes the relations between various subsurface images derived from the Viking Graben line in this text.

Location and rapid oscillation are to some extent contradictory properties, according to Heisenberg's uncertainty principle: it is not possible to localize a function in both space and frequency simultaneously with arbitrary precision. However it is possible to construct functions localized to some extent in space and spatial frequency, and use them as exploding reflector sources to see which parts of the data they influence. Figure 6.13 shows a localized rapid oscillation, oriented 45 degrees to the vertical. The Fourier transform is quite localized in the spatial frequency plane, and the function itself is nonzero onliy over an ellipse of horizontal diameter 2 km and vertical diameter 0.5 km. The center of the ellipse is $z = 2000, x = 10000$. Exploding reflector modeling with this data and constant velocity $v/2 = 1km/s$ produces the data shown in 6.14. You will note that the approximate center position of the enlongated blob of energy is $x = 12000$: it appears that the signal moved about 2 km laterally and 2 km vertically, as if it has traveled along a ray oriented at 45 degrees to the vertical, that is, perpindicularly to the equal phase surfaces of the oscillation.

In fact that is exactly what has happened, and it is possible to explain why with a variant of the geometric optics construction of Chapter 4.

A movie of the exploding reflector pressure field can be viewed with

```
scons pulsemovie45.rsf
sfgrey < pulsemovie45.rsf | xtpen
```

for example - the reader should stretch the window horizontally so that the aspect ratio is closer to 1:1, to begin with. The pulse clearly moves up and to the right, in the direction normal to the equal phase surfaces.

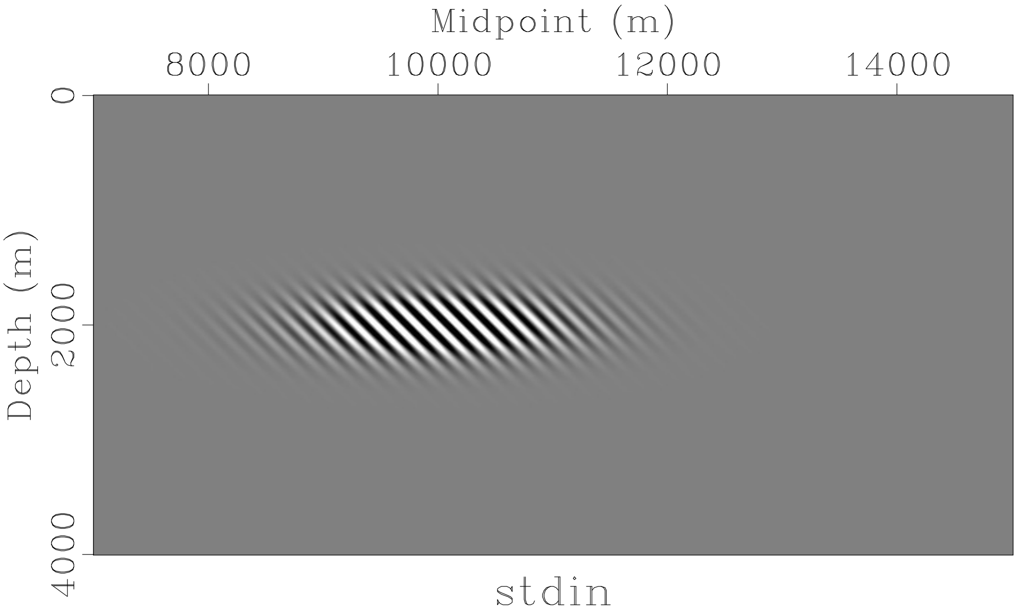

Figure 6.13: Exploding reflector oscillating pulse data.

The exploding reflector source has the form

$$f(x,z,t) = cos(k(x\sin\theta + z\cos\theta))\delta(t). \tag{6.33}$$

Since $\bar{v}(x,z,t) = 0$ for $t < 0$, it follows from the first equation of system 6.24 that

$$\lim_{t\to 0^+} p(z,x,t) = \kappa(x,z)cos(k(x\sin\theta + z\cos\theta)) \tag{6.34}$$

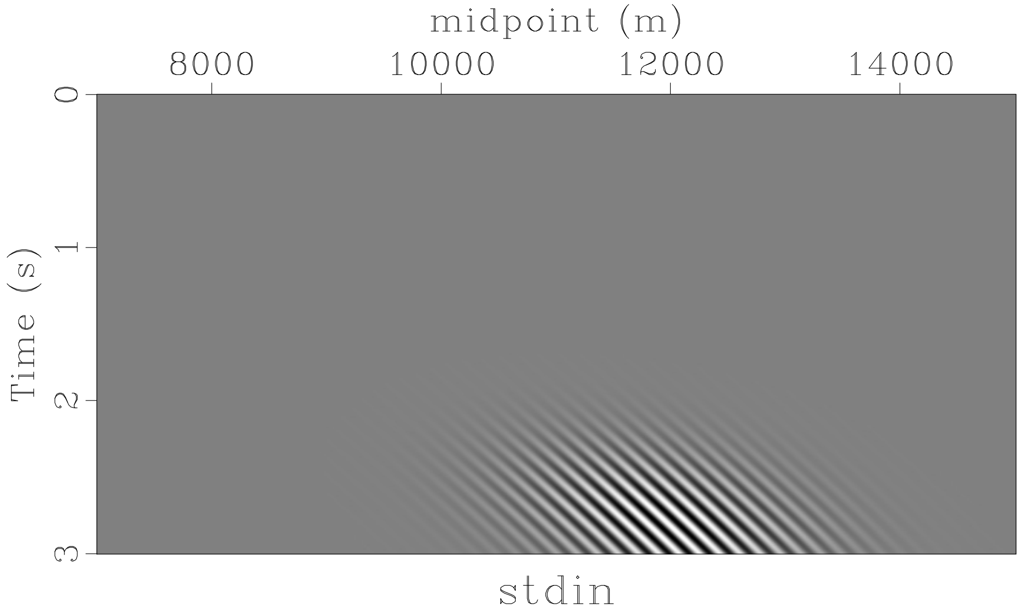

Figure 6.14: Data with same geometry as Viking Graben stack, from data of Figure 6.13 with constant (half-)velocity of 1 km/s, and density of 1 g/cm^3.

(in this case $\kappa = 1$). The data plot 6.14 and the movie mboth suggest that the solution takes the form

$$\begin{aligned} p(x,z,t) &\approx cos(k\phi(x,z,t))p_0(x,z,t) \\ \bar{v}(x,z,t) &\approx cos(k\phi(x,z,t))\bar{v}_0(x,z,t) \end{aligned} \tag{6.35}$$

with $\phi(x,z,t)$ roughly linear in x,z and shifting with some velocity as t changes. Also p_0 and $\bar{v}_0$ represent the observed envelopes of the solution (observed for p_0 - if you were to plot the components of $\bar{v}$ you would see the same), which appear to move along the direction normal to the oscillations at velocity $v/2 = 1$ km/s.

Plug the functional forms 6.35 into the acoustic system 6.24: for $t > 0$, obtain

$$\begin{aligned} -k\sin k\phi\frac{\partial\phi}{\partial t}p_0 + \cos k\phi\frac{\partial p_0}{\partial t} &\approx -\kappa(-k\sin k\phi\nabla\phi\cdot\bar{v}_0 + \cos k\phi\nabla\cdot\bar{v}_0) \\ -k\sin k\phi\frac{\partial\phi}{\partial t}\bar{v}_0 + \cos k\phi\frac{\partial\bar{v}_0}{\partial t} &\approx -\beta(-k\sin k\phi\nabla\phi p_0 + \cos k\phi\nabla p_0) \end{aligned} \tag{6.36}$$

The terms with factors of k are presumably dominant when $k >> 0$, so isolate them:

$$\begin{aligned} -k\sin k\phi \frac{\partial \phi}{\partial t} p_0 &\approx -\kappa(-k \sin k\phi \nabla\phi \cdot \bar{v}_0) \\ -k\sin k\phi \frac{\partial \phi}{\partial t} \bar{v}_0 &\approx -\beta(-k \sin k\phi \nabla\phi p_0) \end{aligned} \tag{6.37}$$

You can view this as a matrix system for the vector $u = (p_0, v_x, v_y, v_z)^T$,

$$k \sin k\phi M u = 0 \tag{6.38}$$

with matrix

$$M = \begin{pmatrix} \frac{\partial \phi}{\partial t} & \kappa\frac{\partial \phi}{\partial x} & \kappa\frac{\partial \phi}{\partial y} & \kappa\frac{\partial \phi}{\partial z} \\ \beta\frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial t} & 0 & 0 \\ \beta\frac{\partial \phi}{\partial y} & 0 & \frac{\partial \phi}{\partial t} & 0 \\ \beta\frac{\partial \phi}{\partial z} & 0 & 0 & \frac{\partial \phi}{\partial t} \end{pmatrix}$$

Equation 6.38 implies that (assuming that $p_0, \bar{v}_0$ are not all zero) the matrix M is singular. Its determinant is

$$\left(\frac{\partial \phi}{\partial t}\right)^2 \left[\left(\frac{\partial \phi}{\partial t}\right)^2 - \kappa\beta \|\nabla\phi\|^2\right] \tag{6.39}$$

Since $\kappa\beta = v^2$ and the first factor should be non-zero (why?), conclude that

$$\frac{1}{v^2}\left(\frac{\partial \phi}{\partial t}\right)^2 = \|\nabla\phi\|^2. \tag{6.40}$$

This is the time-dependent version of the eikonal equation 4.7. Just how close the relation is follows from the ray construction of the phase ϕ. As before, start with the solutions of the physical space part of the ray equations, posed in terms of the phase and a dimensional parameter σ:

$$\begin{aligned} \frac{dX}{d\sigma} &= -\nabla\phi(X,T) \\ \frac{dT}{d\sigma} &= \frac{1}{v^2(X)}\frac{\partial \phi}{\partial t}(X,T) \end{aligned} \tag{6.41}$$

Since $k\phi$ (the exponent, or argument of cos) must be dimensionless, and k is a spatial frequency, ϕ has units of length. T is a time, Therefore from the second equation in 6.41, σ must have dimensions of length. The first equation is then consistent, as both sides are dimensionless. In fact, it will turn out that σ is arc length.

The form of equations 6.41 suggest the definition of a Hamiltonian:

$$H(X,T,\xi,\omega) = \frac{1}{2}\left(\frac{\omega^2}{v^2(X)} - \|\xi\|^2\right)$$

Then the system 6.41 are the Hamilton equations for the evolution of X and T:

$$\frac{dX}{d\sigma} = \nabla_\xi H(X,T,\xi,\omega) = -\xi = -\nabla\phi(X,T) \tag{6.42}$$

and

$$\frac{dT}{d\sigma} = \frac{\partial}{\partial\omega}H(X,T,\xi,\omega) = \frac{\omega}{v(X)^2} = \frac{1}{v^2}\frac{\partial\phi}{\partial t}, \tag{6.43}$$

if you identify

$$\xi = \nabla\phi, \omega = \frac{\partial\phi}{\partial t} \tag{6.44}$$

Then

$$\begin{aligned}
\frac{d\omega}{d\sigma} &= \frac{d}{d\sigma}\frac{\partial\phi}{\partial t}(X,T) \\
= \nabla\frac{\partial\phi}{\partial t}\frac{dX}{d\sigma} + \frac{\partial^2\phi}{\partial t^2}\frac{dT}{d\sigma} &= -\nabla\frac{\partial\phi}{\partial t}\nabla\phi + \frac{1}{v^2}\frac{\partial^2\phi}{\partial t^2}\frac{\partial\phi}{\partial t} \\
&= \frac{1}{2}\frac{d}{dt}\left(-|\nabla\phi|^2 + \frac{1}{v^2}\left(\frac{\partial\phi}{\partial t}\right)^2\right) \\
&= 0 = -\frac{d}{dT}H(X,T,\xi,\omega).
\end{aligned} \tag{6.45}$$

Similarly,

$$\begin{aligned}
\frac{d\xi}{d\sigma} &= \frac{d}{d\sigma}\nabla\phi(X,T) \\
= \nabla\nabla\phi(X,T)\frac{dX}{d\sigma} + \nabla\frac{\partial\phi}{\partial T}\frac{dT}{d\sigma} &= -\nabla\nabla\phi X, T\nabla\phi(X,T) + \nabla\frac{\partial\phi}{\partial T}\frac{1}{v^2}\frac{\partial\phi}{\partial T} \\
&= \frac{1}{2}\nabla\left(-\|\nabla\phi\|^2 + \frac{1}{v^2}\left(\frac{\partial\phi}{\partial t}\right)^2\right) - \frac{1}{2}\nabla\frac{1}{v(X)^2}\left(\frac{\partial\phi}{\partial t}\right)^2 \\
&= \frac{1}{2}\nabla\frac{\omega}{v(X)^2} = -\nabla_X H(X,T,\xi,\omega)
\end{aligned} \tag{6.46}$$

The equations 6.42, 6.43. 6.46, and 6.45 form the Hamiltonian system for the Hamiltonian H. The foregoing calculations have shown that if ϕ is a solution of the time-dependent eikonal equation 6.40, then the trajectories X,T,ξ,ω obtained by solving 6.41 and defining ξ,ω by 6.44 are solutions of Hamilton's equation. In Chapter 4 I establihsed a similar set of identities for the time-independent eikonal 4.7, then showed that one could start with the solution of the Hamilton system and construct a solution of the eikonal equation from it - that is, of course, what is commonly termed ray-tracing. Similarly, it is possible start with family of solutions of the time-dependent Hamiltonian system 6.42, 6.43. 6.46, and 6.45, and construct from them a solution of the time-dependent eikonal 6.40. The apprriate initial conditions are (back to the 2D case!)

$$X(\mathbf{x}) = \mathbf{x},\, T(\mathbf{x}) = 0,\, \xi(\mathbf{x}) = -(\cos\theta, \sin\theta)^T,\, \omega(\mathbf{x}) = \frac{1}{v(\mathbf{x})}$$

Since $\frac{\partial T}{\partial \sigma} > 0$ (why?), you can parametrize the Hamiltonian trajectories by T, and As was the case with time-independent geometric optics in Chapter 4, the amplitudes p_0, $\bar{v}_0$ satisfy transport equations: they evolve along the rays, so that $p_0(X(\sigma), T(\sigma))/p_0(X(0), 0)$ is a positive slowly varying (smooth) function of $X(0)$.

For the constant velocity case illustrated in the figures, the Hamilton equations can be solved by inspection:

$$X = \mathbf{x} + v\xi\sigma, T = \frac{\sigma}{v}$$

with ξ and ω both constant along each ray. Note that $\sigma = vT$ and $dX/dT = v$, so that σ is arc length as stated earlier. The trajectories ("rays") $X(T)$ take off perpindicular to the equal phase surfaces of the exploding reflector source 6.35 and the amplitudes advect along the straight rays.

This construction completely explains the appearance of the examples, assuming that the geometric optics approximation 6.35 is accurate. In fact, you can show that the error is $O(1/k)$, by an argument similar to that explained in Chapter 4.

This reasoning does not quite explain why the first step of CG iteration produces an image. Instead of going further in this direction to construction such an explanation (as is certainly possible), I will take a different approach, that will prove useful in later chapters. This approach starts by noting that the first step of CG is proportional to the gradient ($F^T d$, in the notation of the last section) of the mean square error function J.

The goal is an analysis of the oscillatory behaviour of $F^T d$, in the case that the data d is consistent with the approximation being used, that is, $d = Fr$ for a suitable r. That is, assume that d is actually exploding reflector data for an exploding reflector source r. Then $F^T d = F^T Fr$, so the question becomes:

To what extent is $F^T Fr$ an "image" of r - that is, to what extent are its rapid oscillation present in the same locations and orientations as those of r?

The key to the answer lies in the integral representation of the exploding reflector modeling operator, derived in section 2 as equation 6.8 (2D version), reproduced here for convenience:

$$D\mathcal{F}_{\rm ZO}[v^2]\delta v^2(x_r,,t) \approx \int dt' \int dx\,dz\; G_{v/2}(x_r,t-t';x,z)\delta v^2(x,z)w(t'). \tag{6.47}$$

I have several times made changes in the time wavelet $w(t)$, not to mention amplitudes, and the justification for accepting these various changes will emerge finallly from the analysis to follow. So I will go ahead with another such change. Also I will suppress z_r from $\mathbf{x}_r = (x_r, z_r)$, using the simplifying assumption about acquisition geometry that all receivers are located at the same fixed depth. In the notation introduced in the last section, and with the choice $w(t) = \delta(t)$, 6.47 is equivalent to

$$Fr(x_r,t) \approx \int dx\, G_{v/2}(x_r,t;x,z)r(x,z). \tag{6.48}$$

This expression the analogue of a matrix representation of F, with $G_{v/2}$ as the matrix. Note that $G_{v/2}$ is not computationally accessible: the motivation for the construction of the last section is very much still active. However it is certainly possible to represent the adjoint F^T in terms of the "matrix" $G_{v/2}$, simply by reversing the roles of the arguments ("indices"):

$$F^T d(x,z) \approx \int dt \int dx_r\, G_{v/2}(x_r,t;x,z)d(\mathbf{x}_r,t). \tag{6.49}$$

Hence

$$F^T F r(x,z) \approx \int dt \int dx_r \int dx' \int dz'\, G_{v/2}(x_r,t;x,z)G_{v/2}(x_r,t;x',z')r(x',z') \tag{6.50}$$

Introduce the geometric optics approximation 6.6 for $G_{v/2}$ in 6.50:

$$F^T F r(\mathbf{x}) \approx$$

$$\int dt \int dx_r \int dx' \int dz'\, a(x_r,x,z)a(x_r,x',z')\delta(t-2\tau(x_r,x,z))\delta(t-2\tau(x_r,x',z'))r(x',z') \tag{6.51}$$

You can justify the evaluation of the delta functions by an argument similar to that mentioned in Chapter **??**. The upshot is

$$F^T F r(x,z) \approx \int dx_r \int dx' \int dz'\, a(x_r,x,z)a(x_r,x',z')\delta(2(\tau(x_r,x,z)-\tau(x_r,x',z')))r(x',z') \tag{6.52}$$

This integral is a distribution representation of a surface integral, over the family of surfaces $\{(x',z') : \tau(x_r,x',z') = \tau(x_r,x,z)\}$ parametrized by x_r, x, z. For constant velocity, these surfaces are simply spheres of radii $v\tau(x_r,x,z)$. For velocities not too far from constant, and x', z' not too far from x, z, it is possible to solve the equation $t = \tau(x_r, x', z')$ for z' in terms of the horizontal variables x', x_r. [Note that the 3D case is quite similar, just messier!] Express the solution of $\tau(x_r, x', z') = t$ as $z' = \zeta(x_r, x', t)$. Change variables from z' to $\tau(x_r, x', z')$ in the integral 6.52 requires the Jacobian factor

$$\frac{dz'}{dt}(x_r, x', t) = \left(\frac{\partial \tau}{\partial z'}(x_r, x', \zeta(x_r, x', t))\right)^{-1}.$$

Making the change of variable and integrating out the delta function gives

$$F^T F r(x,z) \approx \int dx_r \int dx'\, b(x_r, x, z, x') r(x', \zeta(x_r, x', \tau(x_r, x, z))) \tag{6.53}$$

in which I have abbreviated

$$b(x_r, x, z, x') = a(x_r, x, z) a(x_r, x', \zeta(x_r, x', \tau(x_r, x, z))) \left(\frac{\partial \tau}{\partial z'}(x_r, x', \zeta(x_r, x', t))\right)^{-1}. \tag{6.54}$$

Remember that the goal is to understand the relation between the high spatial frequency components of r and $F^T F r$. Introduce the frequency components of r into the relation 6.53 via the Fourier transform of r:

$$F^T F r(x,z) \approx \frac{1}{(2\pi)^2} \int dk_x \int dk_z \int dx_r \int dx'\, b(x_r, x, z, x') e^{i(k_x x' + k_z \zeta(x_r, x', \tau(x_r, x, z)))} \hat{r}(k_x, k_z) \tag{6.55}$$

The goal is to understand the behaviour of this oscillatory integral for large k_x, k_z. For this purpose, use the *Stationary Phase Principle*. It approximates an oscillatory integral, with *amplitude* g, smooth and vanishing outside of a ball in Euclidean n-space $\mathbf{R}^n$, and a real-valued *phase* ψ defined on the same ball and having the

Nondegenerate Stationary Point property: if $\nabla \psi(\mathbf{y}^*) = 0$, then $\det \nabla\nabla \psi(\mathbf{y}^*) \neq 0$

The approximation is

$$\begin{aligned} &\int_{\mathbf{R}^n} dy g(\mathbf{y}) e^{i\omega\psi(\mathbf{y})} \\ &\approx \sum_{\nabla\psi(\mathbf{y}^*)=0} \left(\frac{2\pi}{\omega}\right)^{\frac{n}{2}} e^{\frac{\pi i}{4} \mathrm{sgn}\, \nabla\nabla\psi(\mathbf{y}^*)} |\det \nabla\nabla\psi(\mathbf{y}^*)|^{-\frac{1}{2}} g(\mathbf{y}^*) e^{i\omega\psi(\mathbf{y}^*)} + O(\omega^{-n-1}). \end{aligned} \tag{6.56}$$

In the application to 6.55, $n = 2$, x, z are static parameters, and x_r, x are the integration variables. For the large parameter (ω in 6.56), take $k = \sqrt{k_x^2 + k_z^2}$, and write

$$\psi(x_r, x'; x, z, \theta) = \sin\theta x' + \cos\theta \zeta(x_r, x', \tau(x_r, x, z)) \tag{6.57}$$

in which

$$\sin\theta = \frac{k_x}{k}, \cos\theta = \frac{k_z}{k}$$

The amplitude is b as defined in 6.54, divided by $(2\pi)^2$.

The stationary phase condition is

$$\begin{aligned} \frac{\partial\psi}{\partial x'} = \sin\theta + \cos\theta \frac{\partial\zeta}{\partial x'} &= 0 \\ \frac{\partial\psi}{\partial x_r} = \cos\theta\left(\frac{\partial\zeta}{\partial x_r} + \frac{\partial\zeta}{\partial t}\frac{\partial\tau}{\partial x_r}\right) &= 0 \end{aligned} \tag{6.58}$$

From the defining equation of ζ and implicit differentiation, obtain

$$\frac{\partial\zeta}{\partial t} = \left(\frac{\partial\tau}{\partial z}\right)^{-1}; \frac{\partial\zeta}{\partial x'} = -\frac{\partial\tau}{\partial x'}\left(\frac{\partial\tau}{\partial z}\right)^{-1}; \frac{\partial\zeta}{\partial x_r} = -\frac{\partial\tau}{\partial x_r}\left(\frac{\partial\tau}{\partial z}\right)^{-1},$$

with appropriate arguments prescribed. Hence the first stationary phase equation implies that

$$\frac{\partial\tau}{\partial z'}\sin\theta - \frac{\partial\tau}{\partial x'}\cos\theta = 0,$$

that is $\nabla\tau(x_r, x', z')$ is perpindicular to $(-\cos\theta, \sin\theta)$, hence parallel to $(\sin\theta, \cos\theta)$ and therefore to (k_x, k_z) at a stationary point k_r, x', with $z' = \zeta(x_r, x', \tau(x_r, x, z))$ whence $\tau(x_r, x', z') = \tau(x_r, x, z)$. That is: (x, z) and (x', z') lie on the same isochron, and (k_z, k_z) is parallel to $\nabla\tau$ at (x', z').

The second stationary phase condition is equivalent to

$$\frac{\partial\tau}{\partial x_r}(x_r, x', z') = \frac{\partial\tau}{\partial x_r}(x_r, x, z). \tag{6.59}$$

From the discussion in Chapter **??**, the left hand side is the horizontal component of the ray slowness at $x = x_r, z = z_r$ for the ray connecting (x, z) with (x_r, z_r), and similarly for the right-hand side. A hidden hypothesis, implicit in our use of the convolutional model to justify replacing the zero-offset modeling with exploding reflector modeling, is that rays do not turn horizontal, so that both slownesses must have negative z-components at the receiver point. Therefore they are the same, by the eikonal equation. Since the two rays arrive at the same point, and with the same slownessess, all coordinates of the two rays are the same at the receiver point - therefore they are the same ray!!! Since both (x, z) and (x', z') lie on the same ray, and since they lie on the same isochron, as mentioned in the last paragraph, they are the same point. That is, we conclude that $x = x', z = z'$ and (k_x, k_z) is parallel to $\nabla\tau(x_r, x, z)$.

Having identified the stationary point (there is only one), it remains to compute the Hessian $\nabla\nabla\psi$ and its signature. In fact, the signature is 0, and I leave the Hessian as an exercise. Conclude that the oscillatory integral over x_r, x' in 6.55 is

$$\int dx_r \int dx'\, b(x_r,x,z,x')e^{i(k_x x'+k_z\zeta(x_r,x',\tau(x_r,x,z)))}\hat{r}(k_x,k_z) = \frac{2\pi}{k}B(x,z,k_x,k_z)e^{k_x x+k_z z}\hat{r}(k_x,k_z) \tag{6.60}$$

to leading order in k. The quantity $B(x,z,k_x,k_z)$ is the product of the amplitude $b(x_r,x,z,x')$ and the Hessian determinant factor from the stationary phase principle, evaluated at $x = x', z = x'$ and x_r equal to the receiver horizontal location at which the ray taking off from (x,z) with initial slowness (k_x,k_z) reaches the receiver depth $z = z_r$. Note that (k_x,k_z) enters into the construction of B only through the direction vector $(\sin\theta, \cos\theta)$, so that $B(x,z,k_x,k_z)$ is homogeneous in (k_x,k_z) of order 0.

Putting all of this together, obtain an approximation for highly oscillatory r (that is, large k):

$$F^T F(x,z) \approx \frac{1}{2\pi}\int dk_x\, dk_z\, \frac{1}{k}B(x,z,k_x,k_z)e^{k_x x+k_z z}\hat{r}(k_x,k_z) \tag{6.61}$$

Oscillatory integral operators of this type have come to be called *pseudodifferential*, and will be discussed in some detail in the next chapter. The quantity multiplying the exponential, excluding the Fourier transform $\hat{r}$ is called the *symbol* of the operator, and in this case is homogeneous in (k_x,k_z) of order -1.

Note, 2/17 draft: The careful reader will see that I have committed two errors in the preceding derivation: (1) I used the leading order asymptotic form of the 3D Green's function in a 2D calculation, and (2) the exploding reflector source is in effect used as the RHS in the 2nd order wave equation, not the first order system. These errors affect the details but not the overall conclusion of the argument. The operator $F^T F$ is correctly identified as a pseudodifferential operator, but the order and the detailed calculation of the symbol is wrong. The effect of the first error is to make the Green's function 1/2 order more singular than it should be. The correct $F^T F$ should be of one order less, that is, of order -2 rather than -1. The second error has the opposite effect: if the definition of F rests on the system 6.24, then the exploding reflector source must be differentiated in time to be used as the RHS in the 2nd order wave equation. The net result is that $F^T F$ defined by solving 6.24 is 0th order, and this conclusion is independent of space dimension.

The chief properties of pseudodifferential operators important at this point are these:

- Preservation of high frequency asymptotics: if r does *not* oscillate rapidly in the direction $(\sin\theta, \cos\theta)$ near (x,z), then neither does $F^T F$. If the symbol is positive on at $(x,z,\sin\theta,\cos\theta)$, then the converse is also true.

- Insensitivity to amplitude and frequency factors: the property just enunciated is preserved if the symbol is scaled by additional factors of spatial frequency or positive smooth amplitude.

- Changes of amplitude or spatial or temporal frequency in the representation of F result in the sort of changes to the symbol mentioned in the previous bullet.

Taken together, these properties and equation 6.61 validate the assertions made throughout this chapter about the exploding reflector model. In particular, the first step of CG iteration, that is, the application of the adjoint modeling operator to the data, results in an image - assuming that the data is itself consistent with the exploding reflector model. In other words, exploding reflector migration creates an output with the same localized high frequency oscillations as the exploding reflector source used to create the data.

The Viking Graben example was consistent with this conclusion, however the data was "real", that is, extracted from field data tapes. To verify this assertion in vacuo, synthetic examples are useful. Figure 6.15 shows a oscillating pulse exploding reflector source, with the phase normal 15 degrees from vertical. Figure 6.16 shows that exploding reflector data computed from it, and Figure 6.17 the exploding reflector migration of, that is, application of the adjoint operator to, this data. The recovery of the location and orientation of the oscillations is virtually perfect - evidently the spatial frequency used in this example is high enough that the error in stationary phase approximation is unimportant.

6.11 A One-Step Approximate Inversion

A re-write of the linear inverse problem introduced earlier in this chapter (equation 6.18) in terms of the operator F would be: given trace data d, find an exploding reflector source r so that

$$Fr \approx d. \tag{6.62}$$

The analysis in the preceding section illuminates a feature of this problem that was not obvious before: in general, it does not have unique solutions. For purely dimensional reasons, 6.62 does not always have a solution, as indicated earlier: it is an overdetermined problem, at least for the data layout and discretization choices made in this chapter.

It is also an underdetermined problem. For example, the exploding reflector depicted in Figure 6.18, with horizontal equal-phase normals, produces the data depicted in Figure 6.19. This figure is not a misprint - the data is effectively zero. Closer examination reveals that it is not exactly zero, but contains a region of apparent noise with amplitude $O(10^{-6})$ smaller than the data in Figure 6.16, for instance. These nonzero samples are a combination of finite difference and round-off error. Since $d = 0$, a perfectly good solution to equation 6.62 is $r = 0$. The exploding reflector source in Figure 6.18 is a null

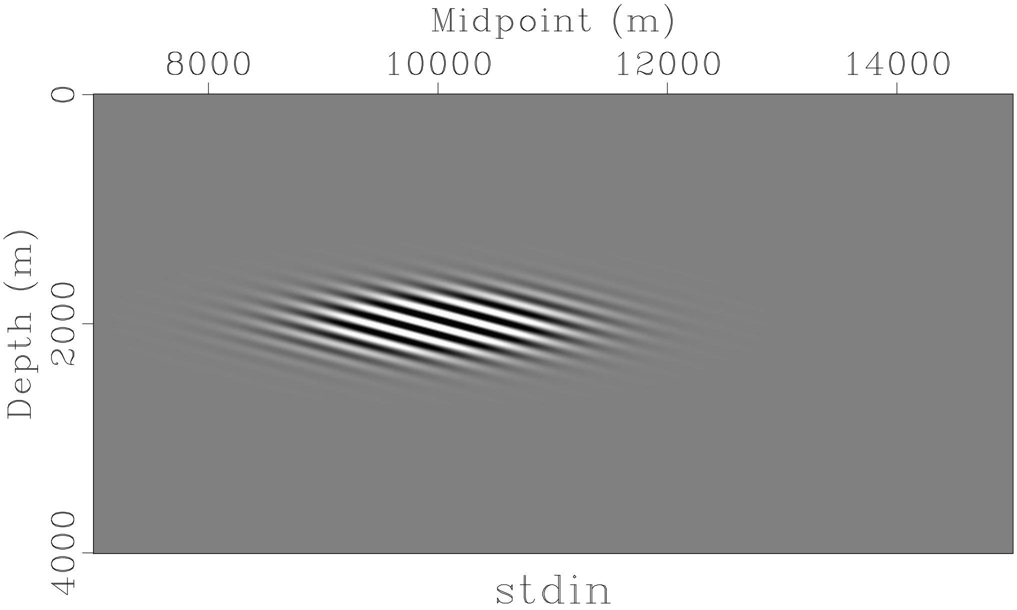

Figure 6.15: Exploding reflector oscillating pulse data - normal to phase surfaces 15 deg from vertical.

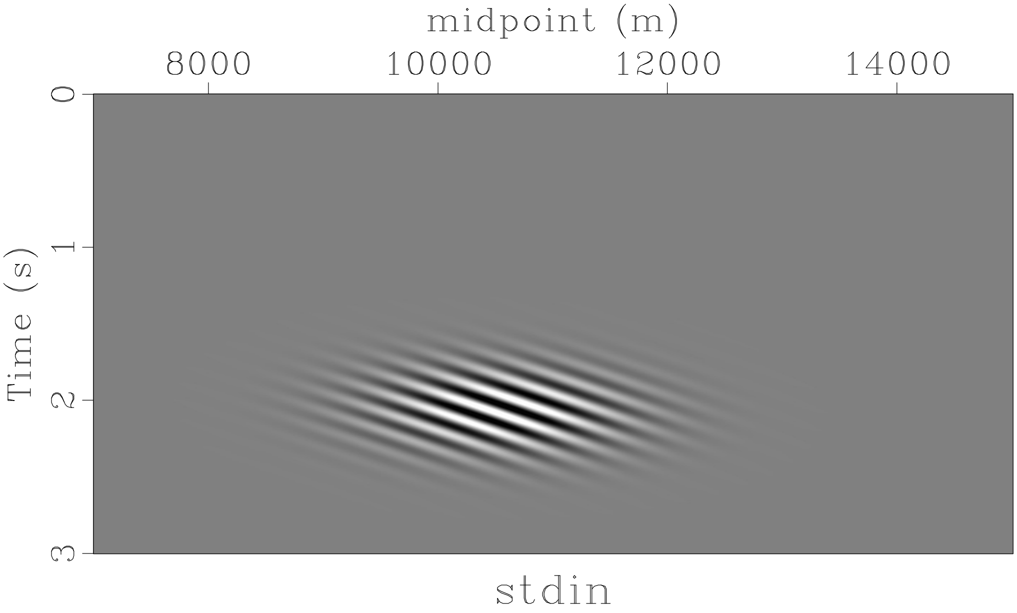

Figure 6.16: Data with same geometry as Viking Graben stack, from exploding reflector source of Figure 6.15 with constant (half-)velocity of 1 km/s, and density of 1 g/cm^3.

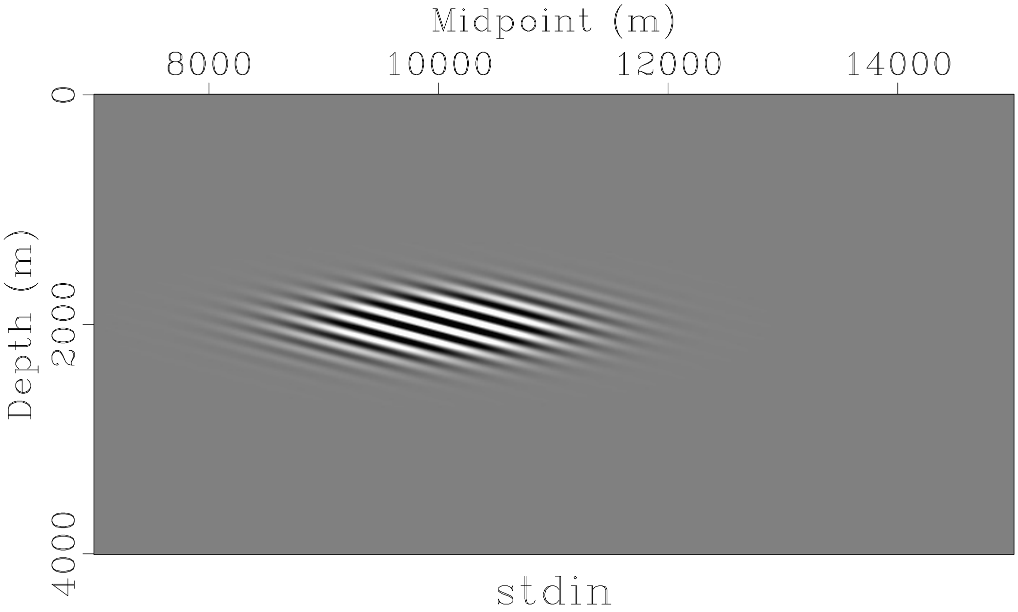

Figure 6.17: Exploding reflector migration of data in Figure 6.16.

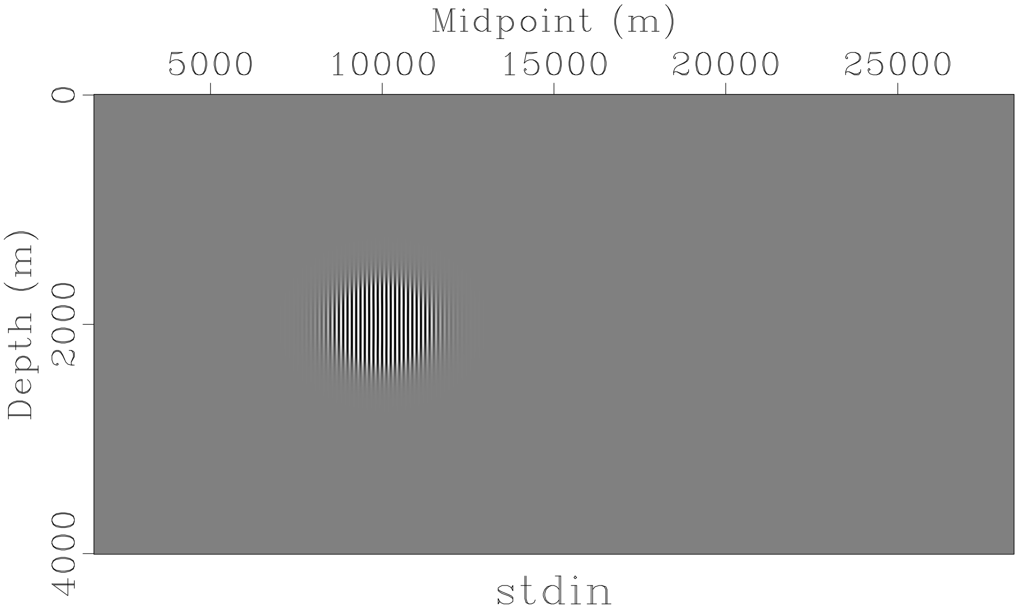

Figure 6.18: Exploding reflector oscillating pulse data - normal to phase surfaces 90 deg from vertical.

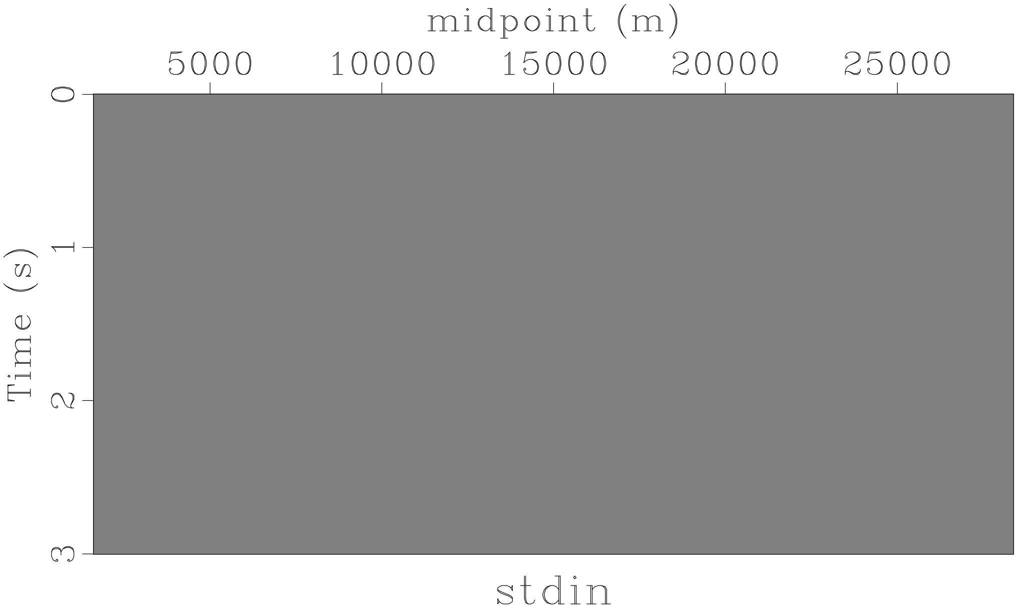

Figure 6.19: Data with same geometry as Viking Graben stack, from exploding reflector source of Figure 6.18 with constant (half-)velocity of 1 km/s, and density of 1 g/cm^3. Grey scale at same level as that of Figure 6.16.

vector for F. It must also be a null vector for F^TF, so solutions of the least-squares version of the inverse problem, or the normal equation 6.21, are also not unique.

We can however single out a solution of the least-squares problem, namely the solution with least length. In effect, this *pseudoinverse* to the normal operator produces the part of the exploding reflector source that fits as much of the data as possible. A standard approximation is the solution of the *regularized least-squares problem*: choose r to minimize

$$J_\lambda[r] = \|Fr - d\|^2 + \lambda^2\|r\|^2 \tag{6.63}$$

The minimizer is a solution of the regularized normal equation

$$(F^TF + \lambda^2 I)r = F^T d. \tag{6.64}$$

Since $F^TF + \lambda^2 I$ is positive definite for $\lambda \neq 0$, this problem has unique solutions. For small λ, the solution approximates the least squares solution of least length.

The regularized least squares problem can also be solved by conjugate gradient iteration, as before. [So why did we get away with solving the unregularized problem with CG? There is a good answer!]

I will next describe an approximation to this solution that only involves a single application of F^T. To begin with, observe that any r can be assembled to arbitrarily good approximation out of oscialling pulses similar to the last few examples: namely, choose a convenience partition of unity $\{\chi_i\}$ of $\mathbf{R}^2$ (we will stay with the 2D problem in this section, but 3D follows exactly the same pattern). Then

$$r(x,z) = \sum_i \frac{1}{(2\pi)^2}\int dx \int dz\, \hat{r}(k_x,k_z)\chi_i(x,z)\cos(k_x x + k_z z)$$

and the right-hand side consists of sums and integrals of oscillating pulses. The data for a typical summand

$$r(x,z) = \chi(x,z)\cos(k_x x + k_z z) \tag{6.65}$$

can be read off from the time-dependent geometric optics approximation developed in the last section. The pressure component of the solution looks like

$$p(x,z,t) = \cos(k\phi(x,z,t))p_0(x,z,t) + O\left(\frac{1}{k}\right),\ k = \sqrt{k_x^2 + k_z^2}. \tag{6.66}$$

so the modeled data is

$$Fr(x_r,t) = \cos(k\phi(x_r,z_r,t))p_0(x_r,z_r,t) + O\left(\frac{1}{k}\right). \tag{6.67}$$

The amplitude p_0 is transported along the ray field that begins at time $t = 0$. If all of the rays starting in the support of χ arrive at the receiver depth within the recording time interval, then it appears that there is enough information to recover r from the data traces: the amplitude p_0 can be reconstructed by solving the transport equation backwards in time, starting at points on the recording surface $z = z_r$ - the same goes for the phase, and then at $t = 0$ one recovers the initial values of ϕ and p_0 and therefore of r, to good approximation for large k.

This is correct, with the two important caveats. First, those who have viewed the propagating-pulse movies will understand that the soltion $(p, \bar{v})$ of the pressure-velocity system 6.24 contains not one, but two propagating pulses, that move in opposite directsions $\pm(k_x, k_z)$ from a monochromatic pulse source like that defined in 6.65. Correspondingly, two geometric optics solutions of the form 6.66 add together to approximate the pressure field. It is easy to see that the two amplitudes must be the same for $t = 0$, therefore each is half of the initial pressure data:

$$p_0(x,z,0) = \frac{1}{2}\chi(x,z).$$

In the half-space geometry of the synthetic survey, if one of the two traveling pulses generated by r travels upward toward the surface at an angle close enough to vertical that the entire signal (except for the $O(1/k)$ error) arrives at the receivers in the recorded interval, then the other pulse travels in the positive z direction and is not recorded. [The 45 degree pulse movie shows this behaviour clearly.] So the recovered initial value of the amplitude must be doubled.

The other caveat is that separating the recorded data into pulses and picking apart the amplitude and phase factors is a nontrivial task, best avoided if at all possible.

In fact, it turns out that it can be avoided. The basic idea is simple: equation 6.67 simply reads off the trace of the pressure field on the surface $z = z_r$. So you could regard the acoustic field $(p, \bar{v})$ as the solution of the boundary value problem

$$\begin{aligned}
\frac{\partial p}{\partial t} &= -\kappa(\nabla \cdot \bar{v} - r\delta(t)) \\
\frac{\partial \bar{v}}{\partial t} &= -\beta \nabla p \\
p(x, z_r, t) &= Fr(x, z_r, t) \\
p, \bar{v} &= 0, t < 0
\end{aligned} \tag{6.68}$$

The additional third equation makes sense if the receivers are dense enough to be regarded as continuous - for the example used here, the receiver sampling is the same as the $x-$ increment, so the discretization is compatible with that assumption.

The problem is however posed in a circular way. I have assumed that the acoustic system is posed in all of $\mathbf{R}^2$, and more than that, the region $z < z_r$ "above the surface" is filled with homogenous material that does not generate reflections. In effect, an *absorbing boundary condition* is posed at $z = z_r$ - in fact, in the computational implementation, that is exactlyl the setup. So it is not possible to specify another boundary condition at $z = z_r$ - any other boundary condition is redundant, and implies a compatibility condition with the absorbing condition that $(p, \bar{v})$ is already expected to satisfy.

The resolution of this paradox comes in two steps. First, observe that the initial conditions (last equations in 6.69) imply that $\lim_{t\to 0^+} \bar{v}(x,z,t) = 0$, but the form of the exploding reflector source implies that $\lim_{t\to 0^+} p(x,z,t) = \kappa(x,z)r(x,z)$. So you could reformulate the exploding reflector source as an initial condition:

$$\begin{aligned} \frac{\partial p}{\partial t} &= -\kappa \nabla \cdot \bar{v} \\ \frac{\partial \bar{v}}{\partial t} &= -\beta \nabla p \\ p(x,z,0) &= \kappa(x,z)r(x,z) \\ \bar{v}(x,z,0) &= 0. \end{aligned} \tag{6.69}$$

If you know the value of p at $z = z_r$, that is, $d = Fr$, then you can drop the absorbing boundary condition and substitute a Dirichlet condition:

$$\begin{aligned} \frac{\partial p}{\partial t} &= -\kappa \nabla \cdot \bar{v}, z > z_r \\ \frac{\partial \bar{v}}{\partial t} &= -\beta \nabla p, z > z_r \\ p(x,z_r,t) &= d(x,t) \\ p(x,z,0) &= \kappa(x,z)r(x,z), z > z_r \\ \bar{v}(x,z,0) &= 0, z > z_r. \end{aligned} \tag{6.70}$$

This problem has a unique solution, and it is the same as the solution of 6.69 if indeed d is the value at $z = z_r$ of p.

However the absorbing boundary condition has disappeared from the formulation of 6.70, so the relation between r and d is broken: it is possible to solve this system for *any* choice of r and d. The connection between the two lies in $(p, \bar{v})$ actually being the solution of (say) 6.69, with absorbing boundary at $z = z_r$, restricted to $z > z_r$. This condition is difficult to check directly, however one of its consequences is obvious. If r is the data of a pulse exploding reflector source of the form 6.65 with large k and wavenumber oriented sufficiently vertically that the part of the solution traveling in the negative z direction passes $z = z_r$ in the negative z direction, in the recording interval $0 \le t \le t_{\max}$, the the field $(p, \bar{v})$ at $t = t_{\max}$ in $z > z_r$ consists only of the other part of the pulse, propagating in the positive z direction. That part of the pulse propagates independently of the part propagating upward (= negative z direction), so you can *reproduce the upward propagating part of the pulse by ignoring the downward propagating part*: that is, you can simply zero out the field at $t = t_{\max}$. In so doing, you have committed two errors: an error of leaving out the entire downward propagating pulse, and an $O(1/k)$ error stemming from the use of the geometric optics approximation. Setting the field at this later time is inconsistent with initial conditions at time $t = 0$, so drop them:

$$\begin{aligned}
\frac{\partial p}{\partial t} &= -\kappa \nabla \cdot \bar{v}, z > z_r \\
\frac{\partial \bar{v}}{\partial t} &= -\beta \nabla p, z > z_r \\
p(x, z_r, t) &= d(x, t) \\
p(x, z, t_{\max}) &= 0, z > z_r \\
\bar{v}(x, z, t_{\max}) &= 0, z > z_r.
\end{aligned} \tag{6.71}$$

Under the hypotheses presented above, the solution $(p, \bar{v})$ of this system should be an $O(1/k)$ approximation the the upward propagating pulse part of the geometric optics approximation.

The relation to r is recovered by remembering that the two parts of the pulse, upward and downward propagating, are each equal to half of the exploding reflector source at time $t = 0$. Therefore

$$\frac{1}{2}\kappa(x,z)r(x,z) \approx p(x,z,0)$$

for the solution $(p, \bar{v})$ of 6.71, under these assumptions. For this solution, define

$$F^{\dagger} d(x,z) = \frac{2}{\kappa(x,z)} p(x,z,0). \tag{6.72}$$

Then the principle of superposition (linearity of the wave equation) implies:

For exploding reflector sources r that can be decomposed into oscillating pulses with wavenumbers sufficiently close to vertical,

$$Fr = d \Rightarrow r \approx F^{\dagger} d$$

Otherwise put, for exploding reflector sources r that generate "big" data d, $F^{\dagger} F r \approx r$. Note that if r generates near-zero data, that is, is in the approximate null space of F, then $F^{\dagger} F r \approx 0$. The properties indicate the $F^{\dagger}$ is an approximate pseudoinverse of F.

Because the initial data for 6.71 is posed at the maximum simulation time $t = t_{\max}$, the system is actually solved backwards in time. Such a backwards-in-time system has appeared earlier in this chapter, namely the adjoint state system for computation of F^T, which I reproduce here for convenience, in its 2D version:

$$
\begin{aligned}
\frac{\partial q}{\partial t} &= -\kappa\left(\nabla\cdot\bar{w} - \sum_{x_m,y_m} d(x_t,t)\delta(x-x_r)\delta(z-z_r)\right)\\
\frac{\partial \bar{w}}{\partial t} &= -\beta\nabla q,\\
q,\bar{w} &= 0, t > t_{\max}.
\end{aligned}
\tag{6.73}
$$

The difference between this system and the system 6.71 is that 6.73 is posed in all of $\mathbf{R}^3\times\mathbf{R}$, and its inhomogeneous term is a right hand side, whereas 6.71 is posed only in the half-space $\{z > z_r\}$ and its inhomogenous term is a boundary condition.

It is possible to convert 6.71 to a system in the whole space with inhomogeneous term, and thereby establish a relation between $F^\dagger$ and F^T. The convention in this book has been that the system 6.11 is defined in all of $\mathbf{R}^4$, in particular that κ and β are defined in all of $\mathbf{R}^3$. So it is possible replace $z > z_r$ in 6.71 with $z < z_r$ and so define a second system in the complementary half-space. Note that the boundary condition for p is the same in both cases, so if you regard p as being defined in $\mathbf{R}^4$, then it is *continuous* across $z = z_r$. It will turn out that $\bar{v}$ is not generally continuous, but still $(p,\bar{v})$ is a vector-valued function and solves the homogeneous wave equation at any point in $\mathbf{R}^4$ *except* on $z = z_r$.

Next observe that this extended version of $(p,\bar{v})$ solves a wave equation, but one with a non-zero right-hand side in the p equation, supported (non-zero) on $z = z_r$. To figure out what this right-hand side is, write the condition for a weak solution of the acoustic system 6.11 (the only kind of solution $(p,\bar{v})$ can be if $\bar{v}$ is discontinuous): for $(\phi,\psi) \in C_0^\infty(\mathbf{R}^3,\mathbf{R}\times\mathbf{R}^3)$,

$$
\int_{\mathbf{R}^3}\left(\frac{1}{\kappa}\frac{\partial\phi}{\partial t} + \nabla\cdot\psi\right)p + \left(\frac{1}{\beta}\frac{\partial\psi}{\partial t} + \nabla\phi\right)\cdot\bar{v}
$$

$$
= \int \phi g \tag{6.74}
$$

To identify the right choice of g to make 6.11 with initial condition at $t = t_{\max}$ (rather than $t = 0$) equivalent to 6.71, divide the integral on the left into two integrals, one from $z = -\infty$ to $z = z_r$, the other from $z = z_r$ to $z = \infty$. Assume that $(p,\bar{v})$ are regular, and that integration by parts is legitimate, except at $z = z_r$. Then the left hand side above is

$$
= -\left[\int_{z\le z_r} dzdxdydt + \int_{z\ge z_r} dzdxdydt\right]\left(\phi\frac{1}{\kappa}\frac{\partial p}{\partial t} + \psi\cdot\nabla p\right.
$$

$$
\left. +\psi\cdot\frac{1}{\beta}\frac{\partial\bar{v}}{\partial t} + \phi\nabla\cdot\bar{v}\right)
$$

$$-\int_{z=z_r} dxdydt(\psi_z[p]+\phi[v_z])(x,y,z_r,t) \tag{6.75}$$

in which the notation $[f]$ signifies the jump of a function f across $z=z_r$:

$$[f](x,y,z_r,t)=lim_{z\to z_r^+}f(x,y,z,t)-lim_{z\to z_r^-}f(x,y,z,t).$$

As mentioned before, $[p]=0$ as its limit from both sides is the (same) Dirichlet condition. Comparing 6.74 and 6.75 shows that in $z>z_r$, the solution $(p,\bar{v})$ of 6.71 is the same as the weak (distribution) solution of

$$\begin{aligned}
\frac{\partial p}{\partial t} &= -\kappa(\nabla\cdot\bar{v}-[v_z]\delta(z-z_r)),\\
\frac{\partial\bar{v}}{\partial t} &= -\beta\nabla p,\\
p(x,z_r,t) &= d(x,t)\\
p(x,z,t_{\max}) &= 0,\\
\bar{v}(x,z,t_{\max}) &= 0.
\end{aligned} \tag{6.76}$$

Comparing the systems 6.76 and 6.73, the solution $(q,\bar{w})$ of the latter must be equal to the solution $(p,\bar{v})$ if d is replaced by $[v_z]$. Since $F^Td=-q_{t=0}$,

$$F^{\dagger}d=-\frac{2}{\kappa}F^T[v_z] \tag{6.77}$$

6.12 Dirichlet-to-Neumann map

I have not said how κ and β are extended to the exterior domain $z<z_r$. A simple choice is evenly: that is, impose the condition

$$\kappa(x,y,z)=\kappa(x,y,2z_r-z),\ \beta(x,y,z)=\beta(x,y,2z_r-z).$$

This condition may render κ and β only continuous, not differentiable, at $z=z_r$, but the concept of weak solution is still applicable, and weak solutions may be constructed out of ordinary solutions in the two half-spaces that and right-hand sides supported on $z=z_r$.

Suppose $(p,\bar{v})$ is an ordinary (smooth) solution of 6.71 in $z\geq z_r$. Then extending p,v_x,v_y to be even and v_z to be odd generates a solution of the same system in $z<z_r$, as is easily verified. Viewed as functions on all of $\mathbf{R}^4$, the evenly-extended p,v_x,v_y are continuous, whereas, v_z generally acquires a discontinuity. Since the extension is odd, $[v_z](x,y,z_r,t)=2\lim_{z\to z_r^+}v_z(x,y,z,t)$.

Newton's law (second equation in 6.11 implies - for anti-causal solutions such as specified in 6.71 - that

$$v_z(\mathbf{x},t) = \int_t^\infty ds \, \frac{\partial p}{\partial z}(\mathbf{x},s)$$

so knowledge of the limiting value of v_z at $z = z_r$ is equivalent to knowledge of $\partial p/\partial t$. Since the boundary value of p (that is, d in the system 6.71) determines the anti-causal solution in z_r hence the boundary value of $\partial p/\partial z$ or equivalently v_z, in fact 6.71 defines a (clearly linear) map

$$\Lambda d = -2 \lim_{z \to z_r^+} v_z. \tag{6.78}$$

Since the boundary value of p is conventionally termed *Dirichlet data*, and that of $\partial p/\partial z$ *Neumann data*, on $z = z_r$, the map Λ is a version of the *Dirchlet-to-Neumann map*, which has figured in many works on inverse problems for partial differential equations (for example Sylvester and Uhlmann (1990)).

In terms of the Dirichlet-to-Neumann map, the relations 6.72 and 6.77 imply that

$$F^\dagger = \frac{2}{\kappa} F^T \Lambda. \tag{6.79}$$

The Dirichlet-to-Neumann map has two very important properties: applied to higly oscillatory data,

- it is positive definite, at least when restricted to data that can be decomposed into upcoming pulses;
- it depends only on the properties of the material parameters κ and ρ in an arbitrarily thin layer $z_r \le z < z_{\max}$.

As a consequence of the first property, 6.79 can be interpreted as meaning that $F^\dagger$ is the adjoint of F in the weighted inner product

$$\langle d_1, d_2 \rangle_{\text{data}} = \langle d_1, \Lambda d_2 \rangle_{L^2}$$

SInce F^{dagger} is an approximate inverse of F, 6.79 implies tht F is approimately unitary with resepct to the inner product $\langle \cdot, \cdot \rangle_{\text{data}}$. That in turn implies that the CG algorithm, for instance, will converge rapidly if reformulated in terms of $\langle \cdot, \cdot \rangle_{\text{data}}$.

The second property implies that, insofar as highly osicllatory data is concerned, the choice of extension of the density and bulk modulus does not matter: only the behaviour of these coefficients inside the half-space $z > z_r$, and in fact near the boundary $z = z_r$, are significant.

Both properties follow from the high-frequency approximation for pulse data introduced earier. Recall the form of the the time-dependent geometric optics approximate solution 6.35, which in 3D reads

$$\begin{aligned} p_{go}(x,y,z,t) &\approx cos(k\phi(x,y,z,t))p_0(x,y,z,t) \\ \bar{v}_{go}(x,y,z,t) &\approx cos(k\phi(x,y,z,t))\bar{v}_0(x,y,z,t) \end{aligned} \tag{6.80}$$

If the phase ϕ satisfies the time-dependent eikonal equation, and the amplitudes p_0, $\bar{v}_0$ the corresponding transport equations, then $(p_{go}, \bar{v}_{go})$ differs from a solution of the wave equation by a field that is $O(1/k)$ in any compact subset of space-time, as shown in Chapter 4. It is possible to construct an approximate solution for which the phase is a given linear phase, when restricted to $z = z_r$:

$$\phi(x,y,z_r,t) = t + \hat{k}_x x + \hat{k}_y y$$

Then the eikonal equation determines the normal derivative of ϕ at the boundary:

$$\frac{\partial \phi}{\partial z}(x,y,z_r,t) = \pm\sqrt{\frac{1}{v(x,y,z_r)^2} - \hat{k}_x^2 - \hat{k}_y^2} \tag{6.81}$$

Evidently, for this prescription to make sense, a necessary condition is that

$$1 > v(x,y,z_r)^2(\hat{k}_x^2 + \hat{k}_y^2) \tag{6.82}$$

over the support of $p_0, \bar{v}_0$. We will come back to the significance of this condition below.

Assuming that condition 6.82 is satisfied, construct an upcoming pulse, i.e. one associated with rays along which z decreases as t increases by solving the Hamilton equations 6.42 and 6.43. These equations imply that the correct sign above is +.

These relations imply that for an approximate solution of 6.76 of the form 6.80, representing an upcoming pulse,

$$v_z(x,y,z_r,t) = \int_t^\infty ds\,\beta(x,y,z_r)\frac{\partial p}{\partial z}(x,y,z_r,t)$$

$$= \int_t^\infty ds\,\beta(x,y,z_r)\left(\frac{\partial p_0}{\partial z}\cos(k\phi) - p_0(x,y,z_r,t)\sin(k\phi)k\frac{\partial \phi}{\partial z}\right)(x,y,z_r,t)$$

$$= \int_t^\infty ds\,\beta(x,y,z_r)\left(\frac{\partial p_0}{\partial z}\cos(k\phi) + p_0(x,y,z_r,t)\frac{\partial}{\partial t}\cos(k\phi)\frac{\partial \phi}{\partial z}\left(\frac{\partial \phi}{\partial t}\right)^{-1}\right)(x,y,z_r,t)$$

$$= -p_0(x,y,z_r,t)\cos(k\phi(x,y,z_r,t))\left(\frac{\partial \phi}{\partial z}\left(\frac{\partial \phi}{\partial t}\right)^{-1}\right)(x,y,z_r,t) + O\left(\frac{1}{k}\right)$$

after integrating by parts, which due to 6.81 becomes

$$= -p_0(x,y,z_r,t)\cos(k(t+\hat{k}_x x+\hat{k}_y y))\sqrt{\frac{1}{v(x,y,z_r)^2}-\hat{k}_x^2-\hat{k}_y^2}+O\left(\frac{1}{k}\right) \tag{6.83}$$

With this result, it is possible to write a high-frequency approximation to the Dirichlet-to-Neumann map. First construct a local approximation: presume that d is supported near a single point on $z=z_r$, and let $p_0 \in C_0^\infty$ be $=1$ on the support of d. Then

$$d(x,y,z_r,t) = p_0(x,y,z_r,t)d(x,y,z_r,t)$$

$$= \frac{1}{(2\pi)^3}\int dk\,dk_x\,dk_y p_0(x,y,z_r,t)\hat{d}(k,k_x,k_y)cos(k(t+\hat{k}_x x+\hat{k}_y y))$$

where $\hat{k}_x = k_x/k, \hat{k}_y = k_y/k$. From 6.83, the corresponding v_z is

$$v_z(x,y,z_r,t) = -\frac{1}{(2\pi)^3}\int dk\,dk_x\,dk_y$$

$$\times p_0(x,y,z_r,t)\hat{d}(k,k_x,k_y)cos(k(t+\hat{k}_x x+\hat{k}_y y))\sqrt{\frac{1}{v(x,y,z_r)^2}-\hat{k}_x^2-\hat{k}_y^2}+\ldots$$

where the elided terms have lower frequency content than the leading term displayed, as they derive from the $O(1/k)$ remainder in 6.83. Recalling the definition 6.78, and the fact that p_0 is $=1$ on the support of d, we obtain the approximation

$$\Lambda d(x,y,z_r,t) \approx$$

$$\frac{1}{(2\pi)^3}\int dk\,dk_x\,dk_y\hat{d}(k,k_x,k_y)cos(k(t+k_x x+k_y y))\sqrt{\frac{1}{v(x,y,z_r)^2}-\left(\frac{k_x}{k}\right)^2-\left(\frac{k_y}{k}\right)^2}+\ldots \tag{6.84}$$

exhibiting Λ in the form of a pseudodifferential operator on the boundary $z=z_r$.

This derivation is incomplete in several ways. To begin with, I have dispensed with the constraint 6.82. It must be reintroduced, via a factor under the integral sign in 6.84 that enforces it. The identity with the reconstructed pressure field at $t=0$ is then only correct if the data d really does consist (at least in high-frequency limit) of Fourier components satisfying 6.82. That and other gaps in the reasoning above may be filled by invoking the theory of pseudodifferential operators more completely than I have done here.

Note that the multiplier inside the integral in 6.84 is positive (when it is well-defined). This means that Λ, as defined here, is a positive operator, modulo low frequency (relatively smoothing) error. Also, operators of this form are symmetric, again modulo low frequency error. Therefore Λ may be used to define a norm in the data space, as noted above.

6.13 Computational Notes

6.14 Suggested Projects

REFERENCES

Dobrin, M., and C. Savit, 1988, Introduction to geophysical prospecting, fourth ed.: McGraw-Hill.

Golub, G. H., and C. F. van Loan, 2012, Matrix computations, 4th ed.: Johns Hopkins University Press.

Nocedal, J., and S. Wright, 1999, Numerical Optimization: Springer Verlag.

Santosa, F., and W. Symes, 2000, Multipole representation of small acoustic sources: Chinese Journal of Mechanics, **16**, 15–21.

Sheriff, R. E., and L. P. Geldart, 1995, Exploration seismology: Cambridge University Press.

Sylvester, J., and G. Uhlmann, 1990, The dirichlet to neumann map and applications: Inverse problems in partial differential equations, Society for Industrial and Applied Mathematics, 101–139.

Chapter 7

Resources: Reading and Software

7.1 Introduction

This book rests on a foundation of science and software provided by many talented researchers. This chapter describes some of the key references, most of which appear at one place or another in the preceding text, and discusses the book's software environment and its use.

7.2 Reading

Lots of good background reading exists for the topics of this course, some with links to codes and computational exercises, and I will refer participants to them whenever appropriate. Probably the closest in "Courant and Hilbert" spirit is Laurent Demanet's 18.325 class notes (Demanet, 2015). These were written for a graduate course in mathematics at MIT, and it shows: topics such as NMO do not appear, but the foundations of imaging theory and the adjoint state method, amongst other things, receive an elegant and complete treatment. The now-classic book by Norm Bleistein, Jack Cohen, and John Stockwell (Bleistein et al., 2001) gives a very thorough treatment of high-frequency asymptotics and "Kirchhoff" imaging. Jerry Schuster's book on seismic imaging basics (Schuster, 2010) is an excellent overview of fundamental geophysical and data processing concepts, with Matlab exercises. Gary Margrave's book (Margrave, 2016) gives a comprehensive account of seismic data processing, with links to the extensive CREWES Matlab codes.

Full waveform inversion (FWI) is a new enough (as a topic of widespread interest) that comprehensive references to complement the classic 1987 Tarantola book (republished by SIAM as (Tarantola, 2005)) are only recently beginning to appear. See for example Fichtner (2010). FWI is well-known to be only locallly convergent (a phenomenon that this course will illustrate, and to some extent explain). In some sense you have to "know the answer before you ask the question". Amongst many approaches that have been suggested for globalizing FWI convergence, I have dwelt on extended inversion, an idea borrowed from velocity analysis. The literature on extended inversion has become fairly large. The review paper Symes (2008) includes an extensive bibliography up to about 2007. The subject is overdue for another review at this writing (2020).

Excellent general texts on geophysical prospecting, for example those by Yilmaz (2001) and Dobrin and Savit (1988), set the topics discussed here in a wider context.

For the fundamental properties of wave equations and their solutions, the main reference is Blazek et al. (2013), an open access paper available through the journal's web site, `http://iopscience.iop.org/journal/0266-5611`. For finite difference methods, the classic is Richtmyer and Morton (1967), which (so far as I know) remains the only mathematically complete account of convergence in book form, half a century after its publication. I provide notes on a special case using energy estimates, that suffices for for the FD methods used in this book (and most of those used in computational seismology). Excellent books on other aspects of finite difference methods include Cohen (2002) and Leveque (2007). x

7.3 Software

A project like this book would simply be impossible without a high quality open source foundation. In particular, I make extensive use of Seismic Unix (SU) and Madagascar, and of the TRIP package develped by my group. These three packages are the results of multi-decade efforts by dozens of talented people, to whom I am deeply grateful. Both SU and Madagascar are community projects, with many contributors. Both have chief contributor/maintainers: for SU, John Stockwell, and for Madagascar, Sergey Fomel (who is responsible for the concept, design, and a great deal of the code).

I presume that the active reader will work in a Un*x-like operating environment, either a flavor of Linux or Mac OSX. The packages on which the book's examples depend will need to be built from C/C++ source code, so you will need C/C++ compilers. Several parts of the code take advantage of relatively recent language innovations: you compilers will need to support key features of the C99 and C++11 standards. Gnu (`gcc, g++`) version 4.4 and later, and Intel (`icc, icpc`) version 13 and later, are adequate and widely available. You will also need Python (version 2.7 or later) and SConstruct (any recent version). Recent Linux distributions generally include installed Python, and so do at least the latest Mac OSX releases. Otherwise you will need to install Python from the `www.python.org`. SConstruct will be installed automatically as the first step in installing

Madagascar, if you haven't already installed it, and is needed for installation of TRIP. Seismic Unix (SU) uses Gnu make as its build system. All of the environments I have mentioned come with a suitable version of Gnu make installed.

Given the prerequisites outlined in the preceding paragraph, you should be able to download and install SU, Madagascar, and TRIP from their web sites:

`reproducibility.org`

`cwp.mines.edu/cwpcodes`

`www.trip.caam.rice.edu/software`

In each source tree you will find installation instructions at the top level. Madagascar and SU have well-established release procedures and should install with minimal headaches. TRIP is newer, and has received less systematic testing; the release from the web site has installed correctly on Linux and Mac OSX under recent versions of Gnu and Intel suites. TRIP defines several modes of parallel execution via MPI. To enable these, it will be necessary to supply also a compatibly installed MPI.

That much will get you through building intermediate data from raw data and the figure files (in Madagascar `.vpl` format) from intermediate data. Inclusion in finished papers, or in this book, requires $\LaTeX$. The Tutorial on the Madagascar web site explains how to obtain and install $\LaTeX$, if you haven't already, and how to make it available to the Madagascar framework.

7.4 Seismic Unix

Installation of SU is straightforward - just remember to comment out `XDRFLAG` in Makefile.config, otherwise you will have to set XDR flags in TRIP software as well and various other complications will ensue. The world is little endian now, and XDR solves a problem that went away.

There is lots of tutorial material about SU on the web, beginning with

`http://www.cwp.mines.edu/sututor/sututor.html`

7.5 Madagascar

Madagascar is both a processing system, with hundreds of core and user-contributed utilities, and a framework for relatively straightforward publication of reproducible research.

While the processing system is very useful (and I use it a lot in the examples to follow), the reproducible research framework is unique: to my knowledge, no other available framework is as well-adapted to merging large-scale computation with text describing its significance. However, the system documentation tends to leave the novice user in the dark about a number of important issues. Most current documentation consists of oversimplified or incomplete examples, with little guidance about incorporating high-performance computing environments. This section gives some additional information on the steps necessary to produce a paper via Madagascar, *beyond the information provided on the Madagascar web site*. I strongly recommend that you

- read `http://reproducibility.org/wiki/Tutorial` and try some of the examples
- read `http://reproducibility.org/wiki/Reproducible_computational_experiments_using_SC`

I will only add a few remarks about aspects glossed over or not treated in the Madagascar web site tutorial materials. To illustrate these points, I'll use the Chapter 2 project directory `caam641/basic/project`. The `SConstruct` file found there implements simple processing workflow.

There are five major steps, starting with raw data and ending with a finished book chapter.

7.5.1 Import Raw Data: `Fetch`

`Fetch` is well-described in the "Reproducibility with Scons" page on the Madagascar site. Its syntax is

```
Fetch(file, dir, server=URL)
```

Here `file` is the name of the file to be fetched from the directory `dir`. The key point is that `dir` must be a subdirectory of `htdocs/data`, where `htdocs` is the usual doc directory under the root accessed by the URL. For example, in `caam641/basic/project/SConstruct` you will find the URL `http://www.trip.caam.rice.edu`. The sysadmin at our site has aliased the root directory (under `/`) accessed by this URL as `/www.trip`. So the file accessed is `/www.trip/htdocs/data/dir/file`.

With this understood, you can now set up your own `Fetch` data sources.

One common plaint heard from Madagascar users is "where did my data go?". The answer is usually "to the `DATAPATH`". Madagascar puts binary data in a directory signified by the `DATAPATH` environment variable, and the output of `Fetch` is assumed to be a binary data file. The default value of `DATAPATH` is `/var/tmp`, and that's what you'll get unless you set it yourself. Madagascar expects all working data process directories to be `project` subdirectories of a paper directory (say, `paper`, and each of those to be a subdirectory of a book directory (either a real book, like this one, in which case the papers are actually chapters, or an anthology or multi-author report, in which case the papers are stand-alone documents. In any case, the canonical directory tree `book/paper/project` is how Madagascar organizes practically everything. For `Fetch` executed in `book/paper/project`, Madagascar creates a subdirectory of `DATAPATH` named `book/paper/project` and places the actual file retrieved by `Fetch` there. A link is placed in the working directory, that is, `book/paper/project`.

7.5.2 Generate Intermediate Data: `Flow`

`Flow` is well-explained in the Madagascar documentation. The only point I wish to add is that it's entirely possible to use commands other than those provided by Madagascar in implementing `Flow`, and it's even possible to override the filter design of `Flow`. For example, in `basic/project/SConstruct` I use the SU command `segyread` to convert the SEGY data file (with standard SEGY structure and IBM binary word order) to an SU file (stripped of ebcidic and reel header, trace data converted to native four-byte floats). `segyread` is not a filter: it expects the input file to be specified as the vallue of `tape=`. `segyread` does dump its output to `stdout`, but I wanted to clean up the binary and real headers after returning from `segyread`, so the command I chose to embed in the flow is not even half a filter. To signify that data is *not* read from `stdin`, add `, stdin=0` after the command string. To signify that data is `not` written to `stdout`, add `, stdout=-1`. In the `segyread` example, I use both. You will see lots of similar constructions in the examples to come.

The command in a flow is a Python string, with any literal parts enclosed in single quotes. If you have defined non-Madagascar commands as string variables, as I recommend, these must be outside the quoted string literals: for example,

```
.. segyread + ' tape=${SOURCE}...'
```

Here `segyread` has been assigned the value `<path to CWP bin dir>/segyread` at the top of the SConstruct file: in the above fragment, it is a Python string variable, concatenated (Python addition of strings) with a literal that defines its arguments. Notice that the source file (second argument of `Flow`) is referenced as `${SOURCE}`, a convention that makes it easy to generate a number of similar commands with minimal typing. See the Tutorial for more on this.

7.5.3 Generate Figures: `Result`

Chapter 2 LaTeXsource resides in the subdirectory `basic` of the `caam641` book directory. As for every other such chapter or paper organized according to the Madagascar framework, each figure to be included in the paper (or chapter) is recorded as a `.vpl` file in the `project/Fig` subdirectory, by evaluation of the `Result` function.

`Result` takes three arguments: the name of a target figure file, the name of a source data file, and a command for producing the former from the latter. A typical example from `basic/project/SConstruct`:

```
Result('parastack','parastack.su','suread endian=0  read=data | put label1=Time label2=
```

Note that the graphics file is in Madagascar graphics format (`.vpl`), created in `project/Fig`, but the source data file is SU (SEGY) format. Even though the rest of the workflow is implemented using SU commands, the graphics commands in our examples use Madagascar utilities only, in order to conform to the reproducible research design. `sfsuread` converts the SEGY traces in `parastack.su` into RSF format, `sfput` adds header words for labels and units, and `sfgrey` converts the RSF data to graphics format. As usual, the prefix `sf` that is part of the name of every Madagascar command may be left off within the command argument for `Result` or `Flow`. Also, the figure file suffix `.vpl` is understood, and the name of the target is simply the filename root. If the input data file (`parastack.su` in this case) were an `.rsf` file, then the figure file root name and the data file root name can be presumed to be the same, and the souce data file name (second argument of `Result`) may be left out.

For example, if we were to convert `parastack.su` to RSF format first, in a separate `Flow`, then the prduction of this figure could read

```
Flow('parastack','parastack.su','suread endian=0  read=data | 'put label1=Time label2=T
Result('parastack', 'grey clip=1.e+7 xinch=10 yinch=5')
```

The final result (`basic/project/Fig/parastack.vpl`) is exactly the same, but now there is an extra intermediate file, `parastack.rsf` in `basic/project` (and its corresponding data file, `parastack.rsf@`, residing in the data path). Note that the suffice `.rsf` is also presumed in any file not containing a period `.` in its name, and may also be left off, as `.vpl` was left off in the specification of the `Result` target..

SU commands (and other non-Madagascar commands, such as compiled-in-place IWAVE) must be defined, most reliably by full pathnames. See the `SConstruct` file in `basic/project` for a good way to do this.

Otherwise, only the standard RSF project framework is required (`from rsf.proj import *` at the top of the `SConstruct`) for this part of the process.

You can do this at home: in your copy of `caam641/basic/project`, run `scons -c` to clean everything up, then run `scons` to regenerate everything.

7.5.4 Archive the Figures

Madagascar defines a sort of figure repository - actually, just a directory with a specific subdirectory tree structure in which copies of figure files are stored for future use. You control the location of this directory with the environment variable `RSFFIGS`. This is a local (platform-specific) definition - for example, on one of my laptops, `.bashrc` contains the line

```
export RSFFIGS=/Users/symes/Documents/tripbooks/figs
```

and you would do something similar if you use `tcsh`.

The figures get into the RSFFIGS figure "repository" by the `scons lock` command, run in the `project` subdirectory.

For example, if you have successfully run `scons` in `basic/project`, then execute `scons lock` to "lock" the figure files, i.e. copy them to the RSFFIGS directory.

After you do this, inspect `RSFFIGS`: you will see a `caam641` subdirectory, i.e. one with the same name as the parent directory of `basic` - this would be created, if it did not exist before. In `$RSFFIGS/caam641`, you will see a `basic` subdir, and within that a `project` subdir, containing copies of all the figure files from `caam641/basic/project/Fig`. This is Fomel's device for keeping the figures for each project in a unique place: create a subdirectory path in `RSFFIGS` mimicing the directory path under the "book" (`scratch`, in this case) and put the figures there. The construction follows the standard `book/paper/project` structure of Madagascar reproducible research.

7.5.5 Build the Paper

You can now build the paper in the `caam641/basic`, by typing `scons`. You will see the `.vpl` files from the `RSFFIGS` subdir converted to `.pdf` and incorporated in the paper `.pdf` file. The rigid choice of directory structure and environment variables makes all this work: you have to follow `book/paper/project`!

If you visit all of the `caam641/chapter/project` directories and execute `scons lock` successfully in each, then you can execute `scons` in `caam641` and reproduce the book that you are now reading!

Once again, only the standard RSF reproducible research framework is needed to make all of this work.

A bit reflection may suggest that actually the procedure outlined here and on the Madagascar web site really only makes sense when all of the work takes place in a single address space. At the moment, Madagascar does not provide adequate tools for incorporating the results of remote computations, for example at supercomputer centers. Later in the book, I'll outline a workaround.

7.6 TRIP

TRIP contains two subpackages:

- Rice Vector Library: a middleware package for connecting complex simulations to linear algebra and optimization algorithms, along with a number of linear algebra and optimization algorithms implemented in terms of this middleware;
- IWAVE, a framework for regular grid finite difference simulation, supplied with RVL interfaces, with several acoustic and elastic simulators implements in terms of the framework. Includes commands invoking the simulators, and their associated maps (Born approximation, adjoint, etc.).

TRIP's user interface slavishly imitates Madagascar and SU (and SEPLib, their immediate ancestor), insofar as possible: commands self-doc, and parameters are passed by keyword=value pairs. The chief difference is that TRIP commands are not filters (reading from stdin, writing to stdout). That unavoidable, as they represent interactions amongst data stored in many files: there are usually multiple inputs, and often multiple outputs. Note that SU commands are also not all structured as filters, for the same reasons, and one uses the same devices to include TRIP commands in Madagascar `Flows`. The first chapter to include examples of this is Chapter 3, Born Approximation, and there are many examples in following chapters.

In the current release, commands read/write trace data from/to SU files, and regularly gridded spatial data from/to RSF file pairs.

REFERENCES

Blazek, K., C. C. Stolk, and W. Symes, 2013, A mathematical framework for inverse wave problems in heterogeneous media: Inverse Problems, **29**, 065001:1–34.

Bleistein, N., J. Cohen, and J. Stockwell, 2001, Mathematics of multidimensional seismic imaging, migration, and inversion: Springer.

Cohen, G. C., 2002, Higher order numerical methods for transient wave equations: Springer.

Demanet, L., 2015, Waves and imaging. (http://math.mit.edu/icg/resources/notes325.pdf).

Dobrin, M., and C. Savit, 1988, Introduction to geophysical prospecting, 4th ed.: McGraw-Hill Book Company.

Fichtner, A., 2010, Full seismic waveform modelling and inversion: Springer Verlag.

Leveque, R., 2007, Finite difference methods for ordinary and partial differential equations: SIAM.

Margrave, G., 2016, Methods of seismic data processing. (https://www.crewes.org/ResearchLinks/FreeSoftware/).

Richtmyer, R., and K. W. Morton, 1967, Difference methods for initial value problems: Wiley/Interscience.

Schuster, G., 2010, Basics of seismic imaging. (http://csim.kaust.edu.sa/files/Short.Courses/SeismicImage.pdf).

Symes, W., 2008, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, no. 6, 765–790.

Tarantola, A., 2005, Inverse Problem Theory and Model Parameter Estimation: Society for Industrial and Applied Mathematics.

Yilmaz, O., 2001, Seismic data processing, *in* Investigations in Geophysics No. 10: Society of Exploration Geophysicists.