The inverse data space: examples and practical aspects

Eric Verschuur [1]
A.J. Berkhout [2]
Jan Thorbecke [3]

It has been proposed that in the case of strong multiple scattering, the data matrix $\mathbf{P}$ *is replaced by its inverse* $\mathbf{P}^{-1}$ *before starting seismic processing. In the inverse data space, all surface-related multiples map onto the origin, whereas the inverse of the primaries map mostly at negative times. This interesting property opens up new ways to separate primaries from multiples and apply other preprocessing steps. In the previous chapter the concept was introduced and illustrated with two simple examples. In this chapter, the inversion of seismic data is demonstrated for a more complicated 2D example. Furthermore, a new strategy for obtaining the inverse of the data is developed, without involving explicit multi-dimensional inversion.*

4.1 The inverse data space, a review

In the previous chapter, the concept of the forward data space (FDS) and the inverse data space (IDS) has been introduced. The inverse data space provides a natural separation of primaries and surface-related multiples, where all surface multiples map into the origin and the inverse of the primaries maps mainly at negative times. According to the feedback model (Berkhout, 1982; Berkhout and Verschuur, 1997), multiple scattering data in the forward data

[1]E-mail: D.J.Verschuur@tudelft.nl
[2]E-mail: A.J.Berkhout@tudelft.nl
[3]E-mail: jan@cray.com

space (FDS) is given by

$$\mathbf{P} = \mathbf{P}_0 + \mathbf{P}_0\mathbf{A}\mathbf{P}, \tag{4.1a}$$

or

$$\mathbf{P} = [\mathbf{I} - \mathbf{P}_0\mathbf{A}]^{-1}\mathbf{P}_0, \tag{4.1b}$$

in which $\mathbf{P}$ contains the data with all surface multiples included and $\mathbf{P}_0$ represents the data without surface multiples (i.e. primaries and internal multiples). Surface operator $\mathbf{A}$ contains the inverse source and detector properties and the reflection operator at the free surface (Verschuur et al., 1992):

$$\mathbf{A} = \mathbf{S}^{-1}\mathbf{R}^{\cap}\mathbf{D}^{-1}. \tag{4.1c}$$

All bold quantities represent wavefields or wavefield operators for one frequency component organized in a matrix: one column contains data from one shot record and one row contains data belonging to a common receiver gather (see Berkhout, 1982). From expression 4.1b, multiple scattering data in the inverse data space (IDS) can be easily derived:

$$\mathbf{P}^{-1} = \mathbf{P}_0^{-1}[\mathbf{I} - \mathbf{P}_0\mathbf{A}]$$

or

$$\mathbf{P}^{-1} = \mathbf{P}_0^{-1} - \mathbf{A}. \tag{4.2}$$

Equation 4.2 may be referred to as the multiple scattering equation in the inverse data space. It shows that the *inverse* data space is very simple with respect to the *forward* data space: it consists of the inverse surface-free response, (primarily) situated at negative times, and the surface-related properties at and around zero time. This can be well understood if we bear in mind that the inversion process transforms the poles in the reverberant forward data to zeros in the non-reverberant inverse data.

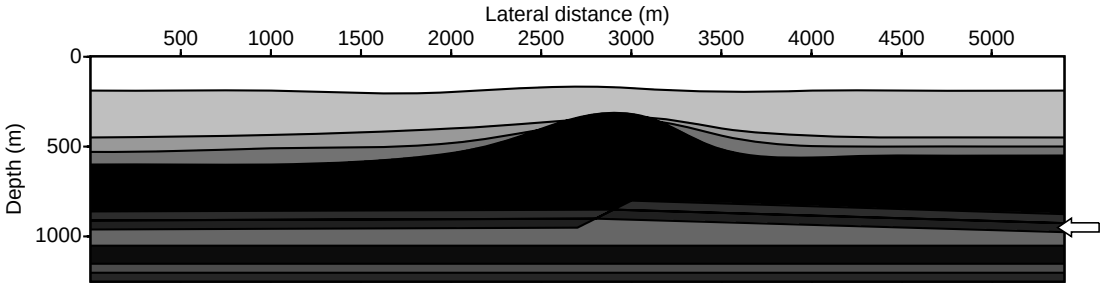
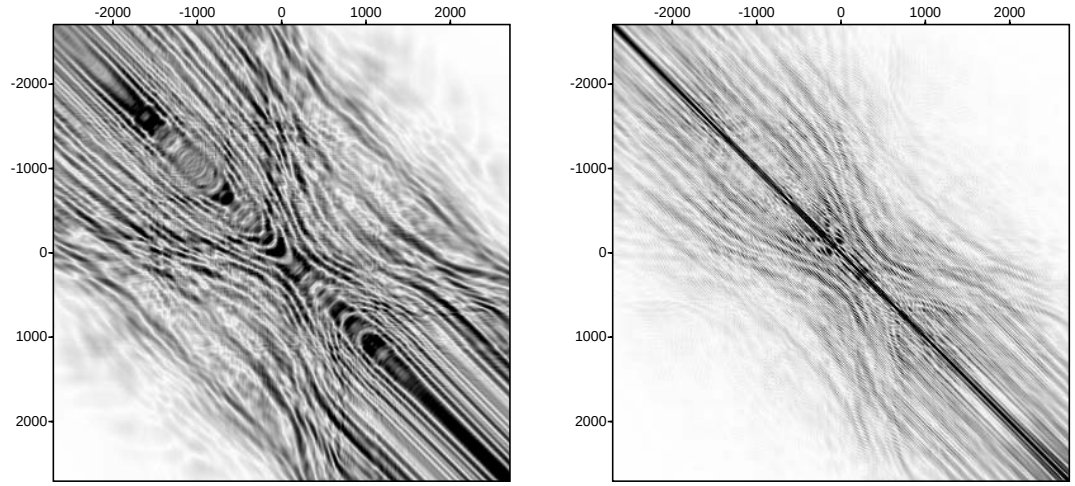


Fig. 4.1 *Subsurface model containing a high-velocity salt layer overlaying the target area with a fault structure. The arrow indicates the target boundaries.*



a) Data matrix for $\mathbf{P}$ *at 30 Hz* *b) Autocorrelation matrix* $\mathbf{P}^H\mathbf{P}$ *at 30 Hz*

Fig. 4.2 *Matrix at 30 Hz for the synthetic data with multiples modeled in the subsurface model of Figure 4.1. b) The autocorrelation matrix of the data with multiples. This matrix is inverted in the least-squares implementation of* $\mathbf{P}^{-1}$*. Note that the matrices are of size 361x361.*

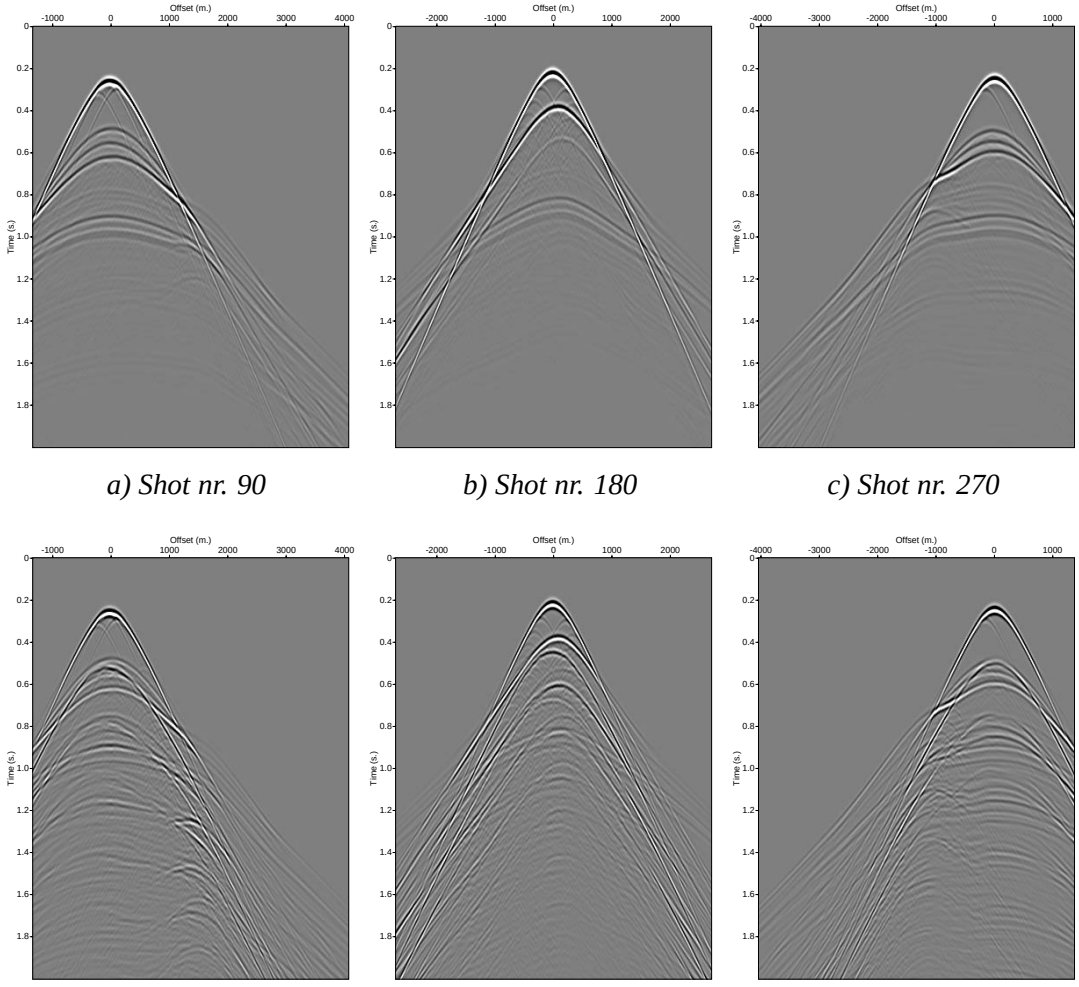
4.2 Synthetic 2D data example

For a demonstration of transforming data to the inverse data space, a relatively complex subsurface model is considered. Figure 4.1 shows the subsurface model containing vertical as well as lateral velocity and density variations. The model is used to simulate seismic data with an acoustic finite difference scheme. An anticlinal salt structure is overlaying the target - a fault structure - and strong surface-related multiples are expected related to the water bottom and the top of the salt. Data are modeled in a fixed spread configuration, with sources and receivers positioned between $x = 0$ and $x = 5400$ m with a step-size of 15m. This results in a prestack dataset of 361x361 traces. Note that the same dataset has been used in the next chapter of this DELPHI Volume for demonstrating the transformation of multiples into primaries. For this example the data matrix, $\mathbf{P}$, is far from Toeplitz (see Figure 4.2a), hence the inverse of $\Delta\mathbf{P}$ and $\mathbf{P}$ was computed in a least-squares sense with the aid of equation 4.3 (see the next section). Note that the autocorrelation matrix $\mathbf{P}^H\mathbf{P}$ has a more diagonal structure than $\mathbf{P}$, which stabilizes the inversion (see Figure 4.2b).

In Figure 4.3 three shot records from this dataset have been displayed that were modeled without and with surface-related multiples. Note the impact of the surface-related multiples in Figure 4.3 d-f.

Both datasets have been inverted using the least-squares matrix inversion. After the inversion in the space-frequency domain, the data is transformed back to the time domain and the gathers for the same source locations as in Figure 4.3 have been extracted, both for the case without and with surface multiples. Results are displayed in Figure 4.4. When comparing the inverse

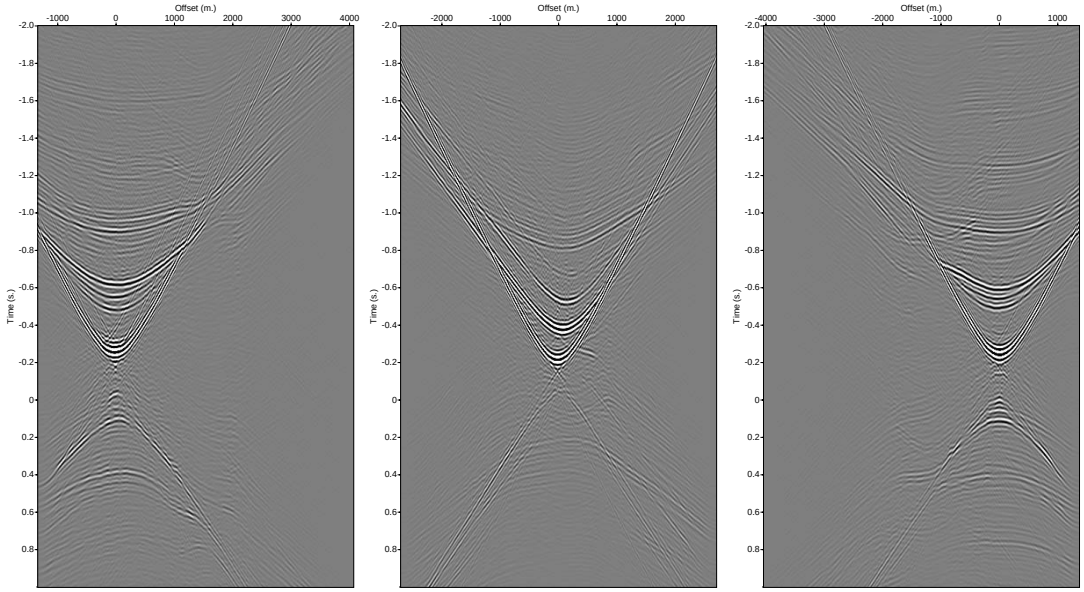
of the two datasets, we can see that all multiple energy is concentrated in and around the origin. In the inverse space, the two datasets are very similar (besides the event in and around the origin). Still, some differences can be observed away from the origin, especially at positive times. This is probably due to the imperfectness of the inversion procedure. This aspect is treated in the next section.



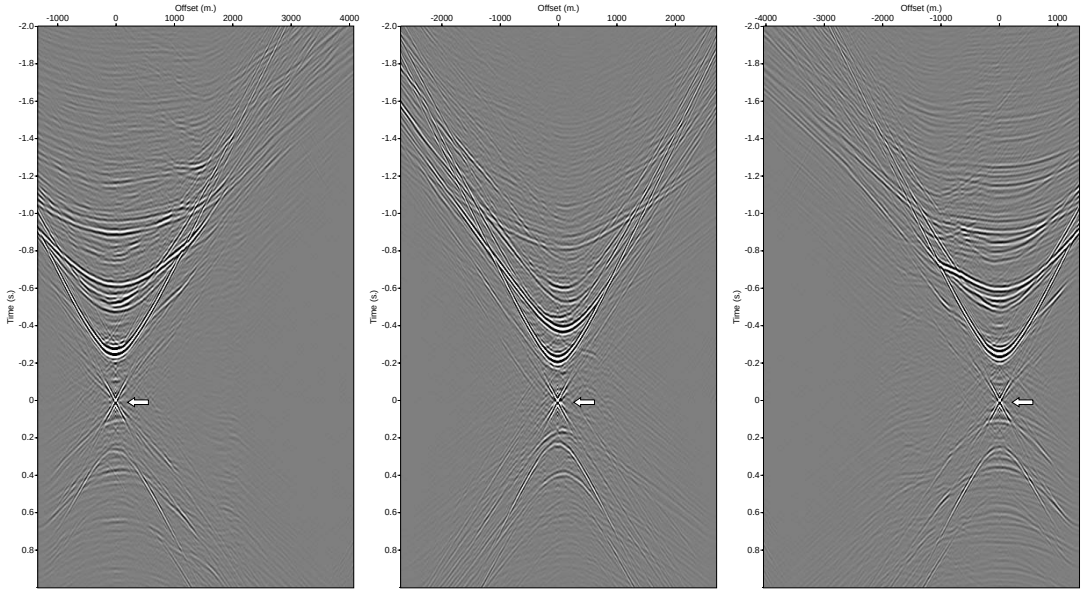
a) Shot nr. 90 *b) Shot nr. 180* *c) Shot nr. 270*

d) Shot nr. 90 *e) Shot nr. 180* *f) Shot nr. 270*

Fig. 4.3 *Three shot records modeled in the subsurface model of Figure 4.1. The top row contains data modeled without surface multiples, and the bottom row contains data with surface multiples included. The 361 receivers are located from x=0 to x=5400m and the source locations are at x=1350 m (a,d), at x=2700 m (b,e) and at x=4050 m (c,f) respectively.*



a) Shot nr. 90 *b) Shot nr. 180* *c) Shot nr. 270*



d) Shot nr. 90 *e) Shot nr. 180* *f) Shot nr. 270*

Fig. 4.4 *Shot records extracted from the data in the inverse data space (IDS). The top row contains the inverse of the data modeled without surface multiples, and the bottom row contains the inverse of the data with surface multiples included. The source locations are at x=1350 m (a,d), at x=2700 m (b,e) and at x=4050 m (c,f) respectively. The arrows point at the multiple energy being mapped at and around the origin.*

4.3 From inversion to iterative modeling

The calculation of the inverse of the data matrix $\mathbf{P}$ is a process which is not easy to accomplish in a stable and artifact-free manner because of the following reasons:

- Seismic data is naturally band-limited, which imposes stabilization problems in the inversion process for high temporal and spatial frequencies.
- The relationship between primaries and multiples, as shown in equations 4.1 and 4.2, assumes an infinite time and offset registration. Any truncation in the FDS will show up as artifacts in the IDS.

To handle the first aspect, the inversion of the data matrices in the examples was done in a least-squares sense:

$$\mathbf{P}^{-1} \approx \mathbf{P}^H \mathbf{B}, \tag{4.3a}$$

where

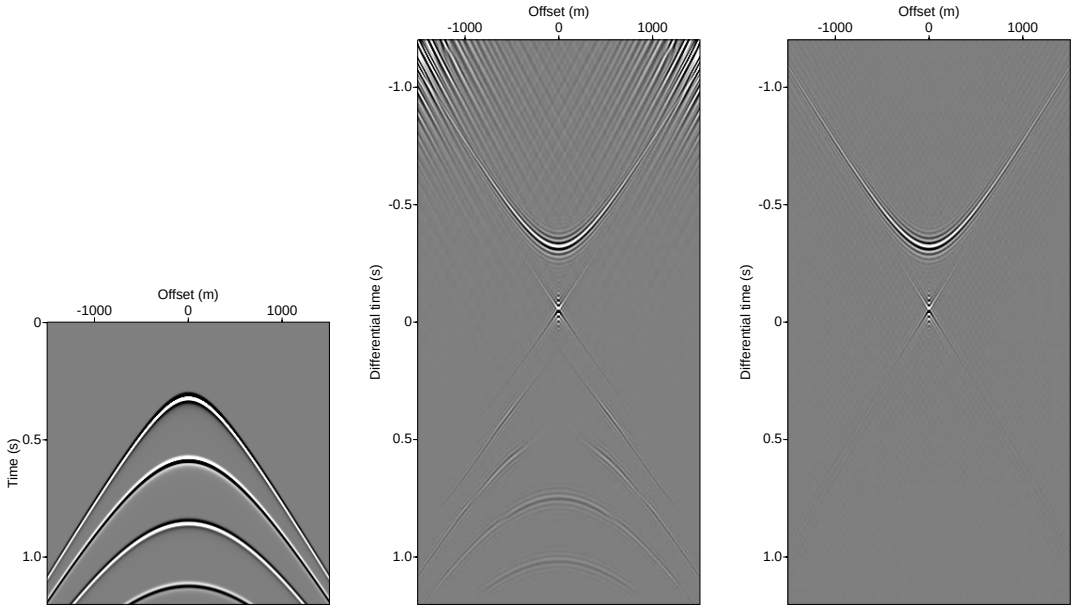
$$\mathbf{B} = \left[\Delta\mathbf{P}\Delta\mathbf{P}^H + \epsilon^2\mathbf{I}\right]^{-1}. \tag{4.3b}$$

In equation 4.3 superscript H denotes the Hermitian operator and the extra term ϵ^2 is a small positive constant that is used for stabilization purposes. Of course, equations 4.3a, b can be refined.

The second aspect was handled in the synthetic data examples by modeling the seismic data with a large offset range (e.g. 10 km) and long registration times (e.g. 8 seconds). In this way, inversion artifacts could be reduced. In Figure 4.5 an example is given where the inverse of a shot record from a model with one horizontal reflector is calculated for a small and large modeling window. For this simple dataset, the inverse should contain one event at negative times (i.e. $\mathbf{P}_0^{-1}$) and one band-limited spike at the origin. This is indeed the case when the large modeling window is used (see Figure 4.5c), but when a small window of input data is used to calculate the inverse, the truncation of the input data leads to an incomplete inversion result: several artifacts can be observed in Figure 4.5b.

The problem at hand is that inversion assumes that the truncated data is the complete dataset, i.e. that the input data is really zero beyond a certain time and offset value. However, this means for example that for some primaries the corresponding multiples are not registered, or for some multiples the higher-order multiples are not present. This renders the physical relationship between primaries and multiples to break down, being compensated by introducing spurious events. Note that most data transforms suffer from truncation effects. As an example, the parabolic Radon transform gives a smearing of each parabolic event along two events in Radon space, that correspond to the first and last offset in the data (see e.g. Maeland, 2003).

A solution to this problem is to redefine the inversion problem as a parameter estimation problem: find the inverse data domain such that it explains the input data within the offset and time window that is present, but does not assume anything for the area beyond this window. Such an inversion problem can have many solutions, and therefore an extra constraint need be



a) Data with multiples

b) Inverse of input with a small data window

c) Inverse of input with a large data window

Fig. 4.5 *a) Shot record with multiples from a medium with one horizontal reflector. b) Inverse of the shot record using the displayed offset and time window. c) Inverse of the data using 4 km offset and 8 seconds of registration time. Note the artifacts when inverting data with a small time-offset window.*

built in to obtain one inversion result. Thus, the inversion procedure can be formulated with the following main steps:

- Define an initial estimate of the inverse domain data $\hat{\mathbf{P}}^{-1}$;
- Check how well it explains the input data, e.g. by computing $\hat{\mathbf{P}}^{-1}\mathbf{P} - \mathbf{I} = \Delta\mathbf{I}$;
- Based on the error, update the inverse domain values $\hat{\mathbf{P}}^{-1}$ by making use of $\Delta\mathbf{I}$, until $\Delta\mathbf{I}$ is small enough.

4.3.1 Solution by a Conjugate Gradient scheme

One way of solving the above inversion problem is by a Conjugate Gradient scheme (Shewchuk, 1994), that solves a problem defined by the following equation:

$$\mathbf{L}\mathbf{x} = \mathbf{y}, \tag{4.4}$$

which is reformulated as a set of normal equations:

$$\mathbf{L}^H\mathbf{L}\mathbf{x} = \mathbf{L}^H\mathbf{y}. \tag{4.5}$$

If $\mathbf{A}$ is defined as $\mathbf{L}^H\mathbf{L}$ and $\mathbf{b}$ is defined as $\mathbf{L}^H\mathbf{y}$, this set of equations now reads:

$$\mathbf{A}\mathbf{x} = \mathbf{b}. \tag{4.6}$$

Optionally, the definition of $\mathbf{A}$ can be redefined as:

$$\mathbf{A} = \mathbf{L}^H\mathbf{L} + \lambda\mathbf{Q}^{-1}, \tag{4.7}$$

where the addition of matrix $\mathbf{Q}^{-1}$ yields a stabilization effect. If $\mathbf{Q}^{-1}$ is chosen to be a unit matrix, this describes the standard damped least-squares solution. However, this matrix can also contain information on the solution $\mathbf{x}$ itself, which makes the inversion process a non-linear one. An example of such stabilization is found by defining a sparseness constraint on the solution vector $\mathbf{x}$, as is used in the sparse Radon and Fourier transform work (see e.g. Sacchi and Ulrych, 1995; Sacchi et al., 1998). In the Conjugate Gradient solution to the problem defined above, the residual vector $\mathbf{r}$ in iteration i:

$$\mathbf{r}_{(i)} = \mathbf{b} - \mathbf{A}_{(i)}\mathbf{x}_{(i)}, \tag{4.8}$$

is used to update the Conjugate Gradient search direction $\mathbf{c}$:

$$\mathbf{c}_{(i)} = \mathbf{r}_{(i)} + \beta\mathbf{r}_{(i-1)}, \tag{4.9}$$

where β is a scalar that is related to the current and previous residual vectors $\mathbf{r}_{(i)}$ and $\mathbf{r}_{(i-1)}$. With the new Conjugate Gradient search direction, an update for the desired solution vector $\mathbf{x}$ is found by optimizing:

$$\mathbf{x}_{(i+1)} = \mathbf{x}_{(i)} + \alpha\mathbf{c}_{(i)} \tag{4.10}$$

as a function of parameter α such that the objective function is optimum. This describes a line search procedure. The Conjugate Gradient scheme is started by choosing a suitable initial estimate for $\mathbf{x}_{(0)}$ and setting the first gradient direction $\mathbf{c}$ vector to be zero. When the solution does not improve anymore, the iteration procedure is stopped. In Shewchuk (1994) a very comprehensive description of the CG algorithm can be found, as well as a list for further references in literature.

For the considered inversion problem, the model domain vector $\mathbf{x}$ is replaced by the desired inverse data domain $\mathbf{P}^{-1}$ and transform matrix $\mathbf{L}$ is taken as $\mathbf{P}$, yielding the following inversion problem to solve:

$$\mathbf{P}^{-1}\mathbf{P}\mathbf{P}^H = \mathbf{P}^H. \tag{4.11}$$

4.3.2 Solution by iterative convolution

Our approach is to go back to the original relations for the inverse data domain and derive an update scheme from them. We aim at finding an inverse of the data $\mathbf{P}^{-1}$ such that:

$$\mathbf{P}^{-1}\mathbf{P} = \mathbf{I}. \tag{4.12}$$

When an estimate of the inverse, $\hat{\mathbf{P}}^{-1}$, is found, then this relation can be written as:

$$\hat{\mathbf{P}}^{-1}\mathbf{P} = \mathbf{I} + \Delta\mathbf{I}, \tag{4.13}$$

meaning that the estimate contains an error. This can also be written as:

$$\Delta\mathbf{I} = \hat{\mathbf{P}}^{-1}\mathbf{P} - \mathbf{I}. \tag{4.14}$$

Multiplying both sides with $\mathbf{P}^{-1}$ gives:

$$\hat{\mathbf{P}}^{-1} = \mathbf{P}^{-1} + \Delta\mathbf{I}\mathbf{P}^{-1}, \tag{4.15}$$

or

$$\hat{\mathbf{P}}^{-1} \approx \mathbf{P}^{-1} + \Delta\mathbf{I}\hat{\mathbf{P}}^{-1} \tag{4.16}$$

and thus

$$\mathbf{P}^{-1} \approx \hat{\mathbf{P}}^{-1} - \Delta\mathbf{I}\hat{\mathbf{P}}^{-1}. \tag{4.17}$$

The latter can be seen as an update formula, where the current version of the inverse of $\hat{\mathbf{P}}^{-1}_{(i)}$ is modified according to the error in the reconstruction:

$$\hat{\mathbf{P}}^{-1}_{(i+1)} = \hat{\mathbf{P}}^{-1}_{(i)} - \Delta\mathbf{I}_{(i)}\hat{\mathbf{P}}^{-1}_{(i)}, \tag{4.18a}$$

in which

$$\Delta\mathbf{I}_{(i)} = \hat{\mathbf{P}}^{-1}_{(i)}\mathbf{P} - \mathbf{I}. \tag{4.18b}$$

The major issue in the above equations is that if the initial estimate of the data inverse $\hat{\mathbf{P}}^{-1}_0$ is far away from the solution, it will probably not converge. This is where the prior knowledge on the physical problem at hand can give a solution: in many cases it is possible to give a reasonable estimate of the inverse of the data, e.g. by :

$$\hat{\mathbf{P}}^{-1}_{(0)} = \alpha^2\hat{\mathbf{P}}^{H}_0 + \alpha\mathbf{I}, \tag{4.19}$$

where $\hat{\mathbf{P}}_0{}^{H}$ is the time reverse of an estimate of the primaries (e.g. the muted water bottom reflection) and α represents a factor that corresponds to the inverse of the seismic wavelet.

4.3.3 Example for a 1D earth

For a one-dimensional earth, all of the above matrix multiplications become scalar multiplications in the wavenumber-frequency domain, which can also be rewritten as convolutions in the time domain for one ray-parameter value. Therefore, we consider that we have a time signal $p(t)$ and that we are looking for the inverse $p^{-1}(t)$ such that

$$p^{-1}(t) * p(t) = \delta(t). \tag{4.20}$$

The iteration process as described in the previous section (Equation 4.18) can be rewritten as follows:

$$\hat{p}^{-1}_{(i+1)}(t) = \hat{p}^{-1}_{(i)}(t) - \epsilon_{(i)}(t) * \hat{p}^{-1}_{(i)}(t), \tag{4.21a}$$

in which

$$\epsilon_{(i)}(t) = \hat{p}^{-1}_{(i)}(t) * p(t) - \delta(t). \tag{4.21b}$$

Thus, finding the inverse of the data $p(t)$ requires a sequence of convolutions and additions. In a similar way, the Conjugate Gradient scheme can be rewritten in terms of iterative convolutions, additions and one line search.

This is investigated for two simple examples that contain the seismic response for a perfect impulsive source (i.e. the source wavelet is $\delta(t)$) and a horizontally layered earth for a horizontal plane wave component.

First we consider a model with one reflector with reflection coefficient $r = 0.5$, the primary response being given in Figure 4.6a. The response with all multiples is displayed in Figure 4.6b. For this model the inverse of the total data is known, and will consist of a unit spike at $t = 0$ and a spike with value $1/0.5 = 2.0$ at the negative primary time. As an initial estimate of the inverse data, two spikes with value 0.5 were put at the right locations. In Figure 4.6c and d results of the estimated inverse domain data for both inversion methods are displayed. Note that the first signal (iteration nr. 0) represents the initial estimate of $p^{-1}(t)$. Both the CG scheme as well as our proposed iteration scheme converge towards this inverse data result. Note, however, that the CG scheme needs more iterations to converge (about 12), whereas the alternative scheme needs about 6.

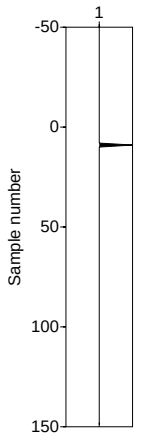
For a next example the response is extended with four more reflectors, as visible in the primary response in Figure 4.7a. The response with multiples becomes much more complex (Figure 4.7b). Taking 50 iterations of the CG scheme shows that the convergence is very slow: still after 50 iterations residuals in the order of 5% appear to be present when checking the inversion result by convolving the estimated inverse data with the input data $p(t)$. However, our update scheme reaches the final result quite fast (within 10 iterations). Note that our scheme has the property that the error is shifted towards positive times and thus is removed from the desired output window. This can be observed in Figure 4.6d for iteration 3-7 and in Figure 4.7d for the iteration results 4, 5 and 6.

Finally, it need be mentioned that these examples are somewhat theoretical, as no seismic wavelet was involved and only one ray parameter was considered. Furthermore, it is expected that both iterative inversion methods can be fine-tuned to increase the performance.

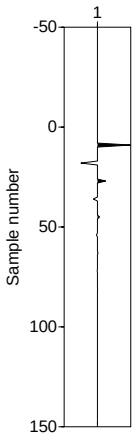
4.4 Conclusions

The inverse data space opens new possibilities in the case of multiple scattering, as primaries and multiples are separated in a very natural way, without any prior information. Furthermore, the source and receiver properties can be extracted in the inverse data space.

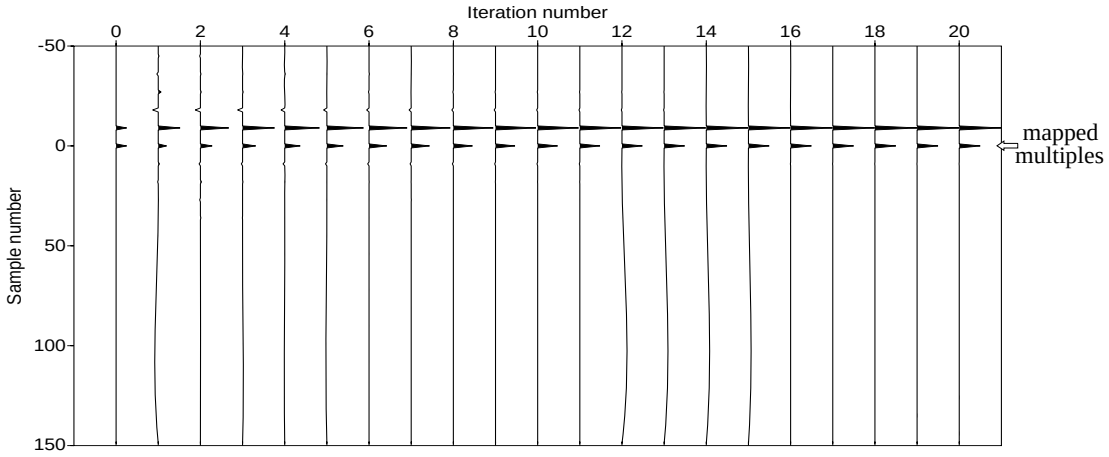
The inversion procedure itself requires matrix inversion in the frequency domain that has the tendency to create artifacts in the space-time domain, because of truncation effects. Therefore, it is proposed to redefine the inversion process as an iterative parametric inversion procedure, in which an initial solution of the inverse data domain is updated based on the inversion error.



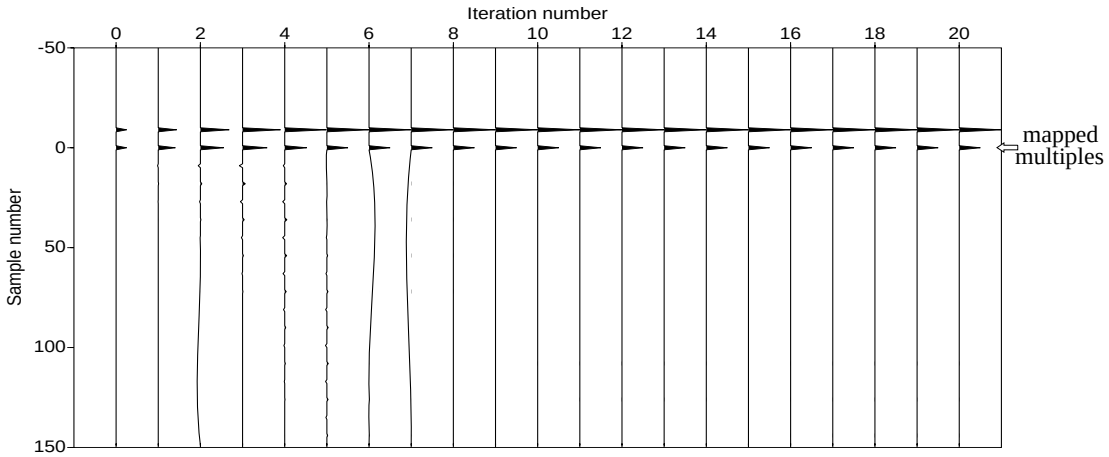
a) Data without multiples



b) Data with multiples

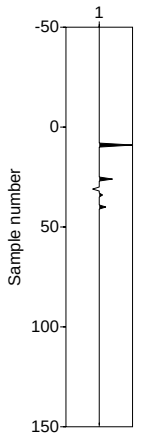


c) Inversion results in IDS with Conjugate Gradient method

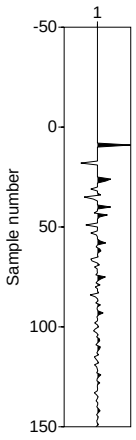


d) Inversion results in IDS with iterative convolution method

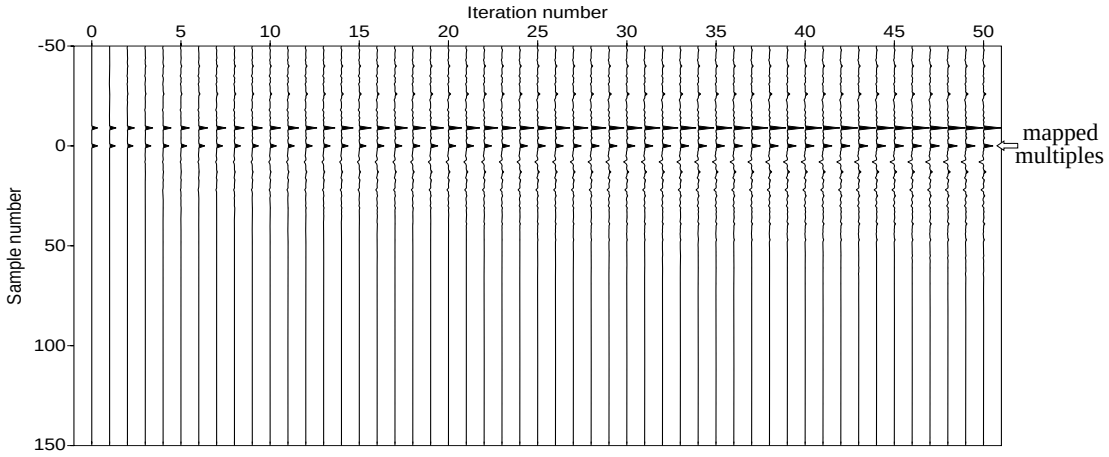
Fig. 4.6 *Inversion results for an earth model with one horizontal reflector displayed in the inverse data space (IDS). a) True primary response. b) Response with multiples. c) Inversion results for different iterations of a Conjugate Gradient scheme. d) Inversion results for different iterations of the alternative update scheme. Note that iteration nr. 0 represents the initial estimate of the inverse data. The arrows point at the multiples being mapped at zero time.*



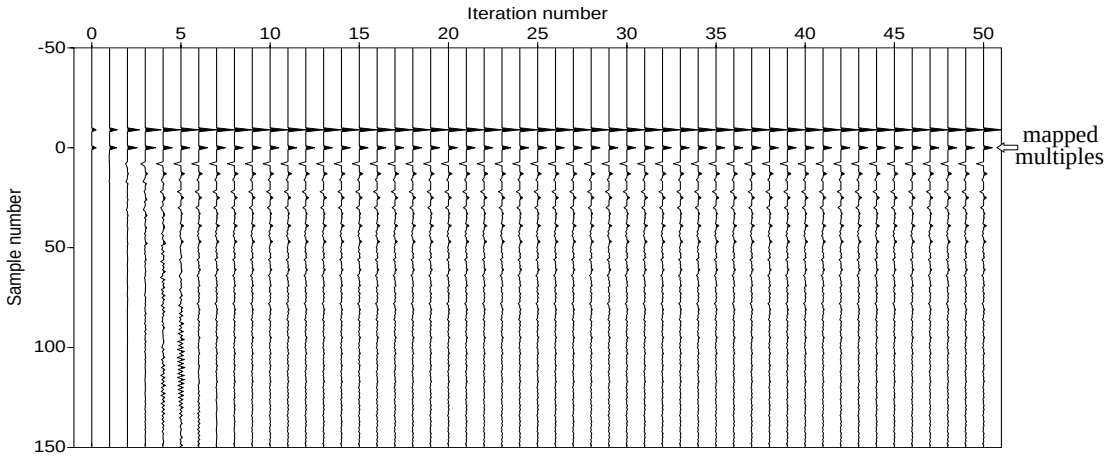
a) Data without multiples



b) Data with multiples



c) Inversion results in IDS with Conjugate Gradient method



d) Inversion results in IDS with iterative convolution method

Fig. 4.7 *Inversion results for an earth model with five horizontal reflectors displayed in the inverse data space (IDS). a) True primary response. b) Response with multiples. c) Inversion results for different iterations of a Conjugate Gradient scheme. d) Inversion results for different iterations of the alternative update scheme. Iteration nr. 0 represents the initial estimate of the inverse data. The arrows point at the multiples being mapped at zero time.*

Two iterative schemes have been proposed, where matrix inversion is replaced by a number of matrix multiplications. The advantage of such a scheme is that it becomes more efficient than a full inversion and that extra constraints can be built in to avoid artifacts in the inversion results. This is part of our current research. Some initial examples for 1D media show encouraging results. A multi-dimensional implementation of the iterative inversion is currently under investigation.

4.5 References

Berkhout, A. J., and Verschuur, D. J., 1997, Estimation of multiple scattering by iterative inversion, part I: theoretical considerations: Geophysics, **62**, no. 5, 1586–1595.

Berkhout, A. J., 1982, Seismic migration, imaging of acoustic energy by wave field extrapolation, a: theoretical aspects: Elsevier (second edition).

Maeland, E., 2003, Disrupton of seismic images by the parabolic radon transform: Geophysics, **68**, no. 4, 1060–1064.

Sacchi, M. D., and Ulrych, T. J., 1995, High-resolution velocity gathers and offset space reconstruction: Geophysics, **60**, 1169–1177.

Sacchi, M. D., Ulrych, T. J., and Walker, C. J., 1998, Interpolation and extrapolation using a high-resolution discrete fourier transform: IEEE Trans. Signal Processing, **46**, 31–38.

Shewchuk, J. R., 1994, An introduction to the conujugate gradient method without the agonizing pain: School of Computer Science, Carnegie Mellon Univ., Pittsburg.

Verschuur, D. J., Berkhout, A. J., and Wapenaar, C. P. A., 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, no. 9, 1166–1177.