Chapter 2

Formulation

The object of this chapter is to formulate a multichannel linear inversion for overcoming aliasing, first in a simple one-dimensional, time-invariant case and then in reflection seismology.

§ 2.1 Aliasing revisited

When a function $f(x)$ is sampled with a certain sampling frequency Δx, its Fourier transform is replicated: if $f(x)$ has the Fourier transform

$$f(x) \supset F(k), \tag{2.1.1}$$

then

$$\text{sampled}[f(x)] \supset \sum_{n=-\infty}^{\infty} F(k - n\kappa) . \tag{2.1.2}$$

The sampling frequency κ (kappa) is defined by

$$\kappa \equiv 2\pi/\Delta x . \tag{2.1.3}$$

Nyquist (1928) showed that if the function f is band-limited,

$$F(k) = 0 \quad \text{for } |k| \geq W , \tag{2.1.4}$$

then sampling is adequate if the sampling rate is at least twice the band limit

$$\kappa \geq 2W . \tag{2.1.5}$$

As shown in Figure 2.1.1f, if the condition (2.1.5) is true, then $F(k)$ can be easily restored from its samples simply by looking at the spectrum for $|k| < \kappa/2$. However, if W is larger than $\kappa/2$, replicated transforms overlap as shown in Figure 2.1.1h, and the result is aliasing: high frequencies appear as low frequencies. Unique restoration of f from its samples is impossible.

If the band limit is $W = N\kappa/2$, there are N contributions of the form $F(k - n\kappa)$ to the Fourier transform of the sampled data. If the aliasing fold N is odd,

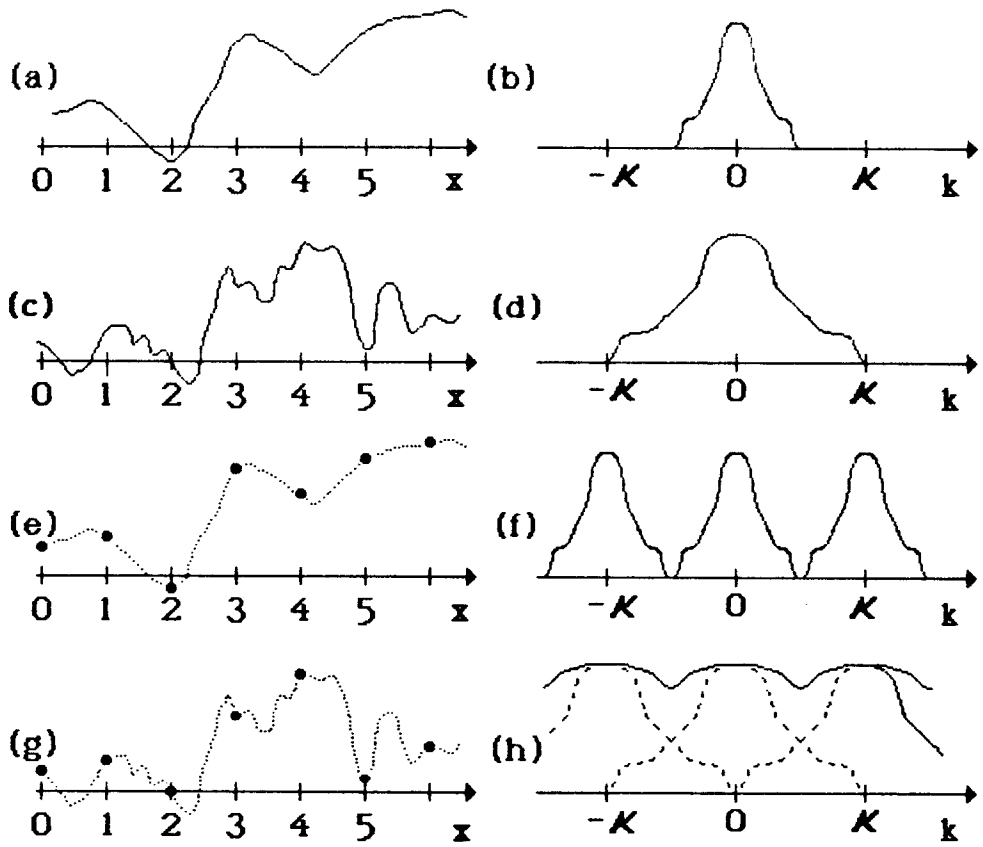


FIG. 2.1.1. Spectral interpretation of sampling and aliasing. (a) A low-frequency signal, and (b) its schematic Fourier transform. (c) A high-frequency signal, and (d) its Fourier transform. (e) Samples from the function in (a), and (f) their corresponding Fourier transform. (g) Samples from the function in (c) and (h) their corresponding Fourier transform. The solid line is the absolute value of the sum of two contribution whose absolute values are denoted by the dashed lines. The sampling rate in (g) is inadequate: there is no way to recover the function in (c) from the sequence of samples in (g).

$$\text{sampled}[f(x)] \supset \sum_{n=-(N-1)/2}^{(N-1)/2} F(k-n\kappa) . \tag{2.1.6a}$$

If N is even,

$$\text{sampled}[f(x)] \supset \sum_{n=-N/2+1}^{N/2} F(k-n\kappa) . \tag{2.1.6b}$$

In a matrix form, for $N=3$, we have

$$\text{sampled}[f(x)] \supset \sum_{n=-1}^{1} F(k-n\kappa) = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} F(k-\kappa) \\ F(k) \\ F(k+\kappa) \end{pmatrix} . \tag{2.1.7}$$

Equation (2.1.7) is under determined: the data are not sufficient to restore f .

§ 2.2 Multichannel inversion in one dimension

Suppose a function $m(x)$ is filtered by J convolution filters, h_j, prior to sampling as shown in Figure 2.2.1. For each channel we can use equation (2.1.2) with $F(k) = h_j(k)\, m(k)$:

$$d_j(k) = \sum_n h_j(k - n\kappa)\, m(k - n\kappa), \qquad \text{for } j = 1,2,\cdots,J \ . \tag{2.2.1}$$

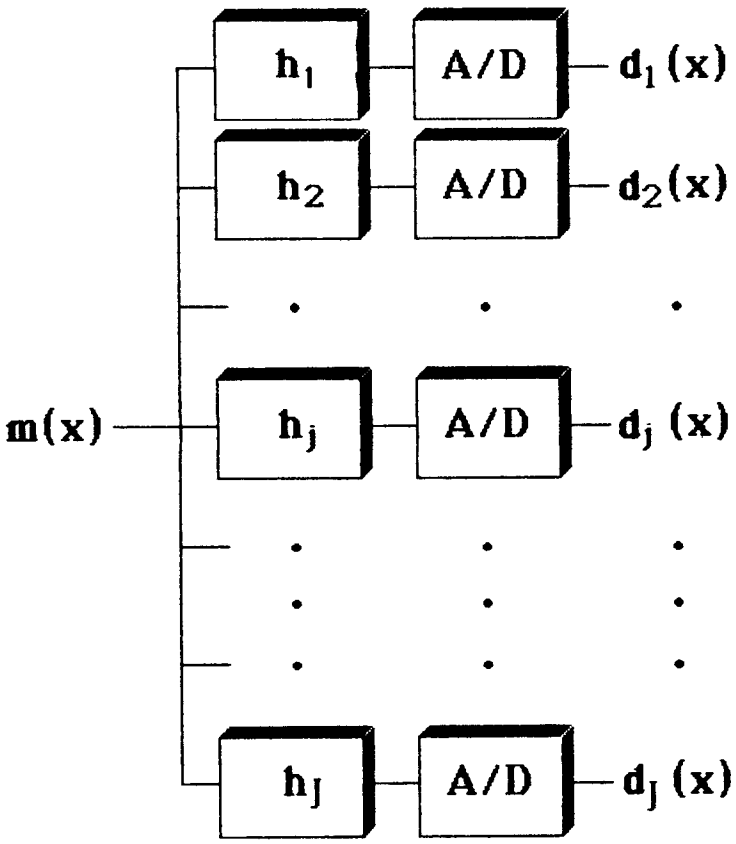


FIG. 2.2.1. Multichannel data. The model m is filtered by J filters, in parallel, before sampling.

In a matrix form, for $J=5$ and $N=3$, we have

$$\begin{pmatrix} d_1(k) \\ d_2(k) \\ d_3(k) \\ d_4(k) \\ d_5(k) \end{pmatrix} = \begin{pmatrix} h_1(k-\kappa) & h_1(k) & h_1(k+\kappa) \\ h_2(k-\kappa) & h_2(k) & h_2(k+\kappa) \\ h_3(k-\kappa) & h_3(k) & h_3(k+\kappa) \\ h_4(k-\kappa) & h_4(k) & h_4(k+\kappa) \\ h_5(k-\kappa) & h_5(k) & h_5(k+\kappa) \end{pmatrix} \begin{pmatrix} m(k-\kappa) \\ m(k) \\ m(k+\kappa) \end{pmatrix} . \tag{2.2.2}$$

Multichannel inversion is the solution to this system of equations. It interpolates on the x axis because it extrapolates on the k axis: the data are known up to the frequency $\kappa/2$, because $d(k)$ is replicated: $d(k-\kappa)=d(k)$; the model, however, is found for $|k|<3\kappa/2$, because $m(k)$, $m(k-\kappa)$, and $m(k+\kappa)$ are found for $|k|<\kappa/2$.

A classical example of multichannel inversion is Shannon's note on sampling a function and its derivative quoted in §1.3. A schematic Di-channel ($J=2$) experiment is described in Figure 2.2.2. The data d_1 are samples from the model m. The data d_2 are samples from the first derivative of the model, $ik\ m(k)$.

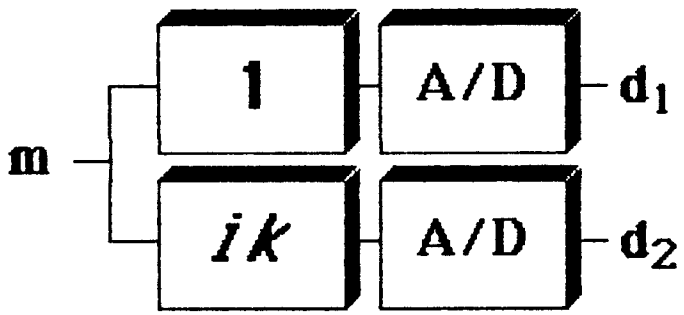


FIG. 2.2.2. Di-channel data of a function and its derivative. ik is the differentiation operator.

If there is no aliasing ($N=1$), the data are

$$\begin{pmatrix} d_1(k) \\ d_2(k) \end{pmatrix} = \begin{pmatrix} 1 \\ ik \end{pmatrix} m(k) \ . \tag{2.2.3}$$

The least-squares solution for this overdetermined system is (for $k \neq 0$)

$$m(k) = \frac{1}{2} \begin{pmatrix} 1 & 1/ik \end{pmatrix} \begin{pmatrix} d_1(k) \\ d_2(k) \end{pmatrix} . \tag{2.2.4}$$

The processing implied by equation (2.2.4) is a separate-channels inversion, performed as follows:

(1) Separately invert each channel.

(2) Sum.

This is the approach taken in geophysical processing:

(1) Map each common-offset section to the zero offset (by NMO and DMO).

(2) Stack.

One may want to allow weighted stacking to reduce the expected noise near $k=0$ in channel 2 in equation (2.2.4) (to mute noisy or stretched data in the geophysical case), but essentially the separate-channels inversion is adequate when there is no aliasing.

The separate-channels inversion is inadequate if there is aliasing. If $m(k)$ is bandlimited by 2π, then adequate sampling requires a sampling interval of $\Delta x \leq 1/2$. If we sample each channel with $\Delta x = 1$, we have, according to equation (2.1.6b), with $N=2$ and $\kappa=2\pi$,

$$d_1(k) = m(k) + m(k-2\pi)$$

in channel 1, and

$$d_2(k) = ik\ m(k) + i(k-2\pi)\ m(k-2\pi)$$

in channel 2. In a matrix form, we have a special case of equation (2.2.1):

$$\begin{pmatrix} d_1(k) \\ d_2(k) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ ik & i(k-2\pi) \end{pmatrix} \begin{pmatrix} m(k) \\ m(k-2\pi) \end{pmatrix} . \tag{2.2.5}$$

The processing of equation (2.2.4) will give the incorrect result,

$$\begin{aligned} \tilde{m}(k) &= \frac{1}{2} \begin{pmatrix} 1 & 1/ik \end{pmatrix} \begin{pmatrix} d_1(k) \\ d_2(k) \end{pmatrix} \qquad (2.2.6) \\ &= \frac{1}{2} \begin{pmatrix} 1 & 1/ik \end{pmatrix} \begin{pmatrix} m(k) + & m(k-2\pi) \\ ik\ m(k) + & i(k-2\pi)\ m(k-2\pi) \end{pmatrix} \\ &= m(k) + \left(1 - \pi/k\right) m(k-2\pi) \\ &\neq m(k) . \end{aligned}$$

The correct processing is inversion of the matrix in equation (2.2.5):

$$\begin{pmatrix} m(k) \\ m(k-2\pi) \end{pmatrix} = \frac{1}{2\pi i} \begin{pmatrix} -i(k-2\pi) & 1 \\ ik & -1 \end{pmatrix} \begin{pmatrix} d_1(k) \\ d_2(k) \end{pmatrix} . \tag{2.2.7}$$

The multichannel inversion of equation (2.2.7) and the separate-channels inversion of equation (2.2.4) imply two different processing techniques. The conclusion for reflection seismology is that conventional processing, designed for well sampled data may become inadequate when the data are aliased.

§ 2.3 The choice of model and data

Formulating multichannel inversion in reflection seismology first requires the description of the relations between the physical properties of the earth (the model) and the seismic waves (the data).

The wave equation for an acoustic isotropic medium

$$\rho\, \ddot{\mathbf{u}} = \nabla\Big(K\, \nabla\cdot\mathbf{u}\Big)\ , \tag{2.3.1}$$

is obtained by substituting Hooke's law

$$P\ = -\,K\, \nabla\cdot\mathbf{u}\ , \tag{2.3.2}$$

into Newton's† law,

$$\rho\, \ddot{\mathbf{u}} = -\,\nabla P\ . \tag{2.3.3}$$

$\mathbf{u}(\mathbf{r},t\,)$ is the displacement vector, $P\,(\mathbf{r},t\,)$ is the presure, $\rho(\mathbf{r})$ is the density, and $K\,(\mathbf{r})$ is the incompressibility. t is the time, $\mathbf{r}$ is the space-location vector. ∇ is the del operator: ∇P is the presure gradient (force), $\nabla\cdot\mathbf{u}$ is the displacement divergence (strain). $\ddot{\mathbf{u}}$ is the second derivative in time of the vector $\mathbf{u}$.

If K and ρ are constant we have the familiar result:

$$\ddot{\mathbf{u}} = \frac{K}{\rho}\, \nabla\Big(\nabla\cdot\mathbf{u}\Big)\ . \tag{2.3.4}$$

Any wave, $\mathbf{u}(\mathbf{r},t\,) = f\ (t-\mathbf{s}\cdot\mathbf{r})\hat{\mathbf{e}}$, where $\mathbf{s}=\sqrt{\rho/K}\ \hat{\mathbf{e}}$ is the slowness vector, and $\hat{\mathbf{e}}$ is any unit length vector, is a solution to the wave equation (2.3.4).

The physical properties $K\,(\mathbf{r})$ and $\rho(\mathbf{r})$ seem to be a natural choice of a model for multichannel inversion. Unfortunately, the relations between seismic data and this model are nonlinear: if the elastic coefficient K is doubled the data will change in a complicated way, not just double in amplitude. In the context of multichannel inversion the h_j filters of equation (2.2.2) are nonlinear for this model.

$\rho(\mathbf{r})$ and $K\,(\mathbf{r})$ are seldom found by reflection seismology. *Reflectivity* is the physical property reflection seismology is good at finding, (Claerbout, 1984). Reflectivity is defined as the gradient of the log of the impedance, $\sqrt{K\,\rho}$. To see why, consider a one-dimensional example:

$$K\,(x\,) = \begin{cases} K_1 & \text{for } x<0 \\ K_2 & \text{for } x>0\ , \end{cases} \tag{2.3.5}$$

† Robert Hooke and Isaac Newton did not like each other, to say the least; nevertheless, seismology is based on a combination of the physical principles they formulated.

and

$$\rho(x) = \begin{cases} \rho_1 & \text{for } x < 0 \\ \rho_2 & \text{for } x > 0 \end{cases} . \tag{2.3.6}$$

The wavefield is

$$u(x,t) = \begin{cases} f(t - s_1 x) + R\ f(t + s_1 x) & \text{for } x < 0 \\ T\ f(t - s_2 x) & \text{for } x > 0 \end{cases} . \tag{2.3.7}$$

$f(t - s_1 x)$ is a general incident wave. R and T are the reflection and transmission coefficients. The slownesses are $s_j = \sqrt{\rho_j / K_j}$ for $j = 1,2$.

Continuity of displacement and presure at $x = 0$, can be used to determine T and R : continuity of displacement implies

$$1 + R = T \ ; \tag{2.3.8}$$

continuity of presure and Hooke's law give

$$\sqrt{K_1 \rho_1}\left(1 - R\right) = \sqrt{K_2 \rho_2}\ T \ . \tag{2.3.9}$$

From equations (2.3.8) and (2.3.9) we get

$$\begin{aligned} R &= \frac{\sqrt{K_1 \rho_1} - \sqrt{K_2 \rho_2}}{\sqrt{K_1 \rho_1} + \sqrt{K_2 \rho_2}} , \\ T &= \frac{2\sqrt{K_1 \rho_1}}{\sqrt{K_1 \rho_1} + \sqrt{K_2 \rho_2}} . \end{aligned} \tag{2.3.10}$$

The reflected wave $R\ f(t + x/v_1)$ is proportional to the reflectivity $R = \nabla \log \sqrt{K \rho}$.

Equation (2.3.10) was obtained for a one-dimensional example but it holds also in two and three dimensions for normally incident waves. Since the primary reflections on a zero-offset section are in normal incidence, the relationship between the seismic wavefield, as recorded on the zero-offset section and the earth's reflectivity is approximately linear: the approximation is in neglecting transmission loss and multiples; the linear operator is poststack migration. Non-zero-offset data are also approximately in linear relationship to the reflectivity: the linear operator is prestack migration; the approximation is in multiples and amplitudes; the travel time of the primary reflections is exact. Since both the zero-offset section and the non-zero-offset section are linearly related to the reflectivity, they are linearly related to each other, the linear operator is prestack-partial-migration also known as dip-moveout (DMO).

As shown in Figure 2.3.1, the model is the zero-offset section which we would record if we performed an adequately sampled zero-offset experiment.

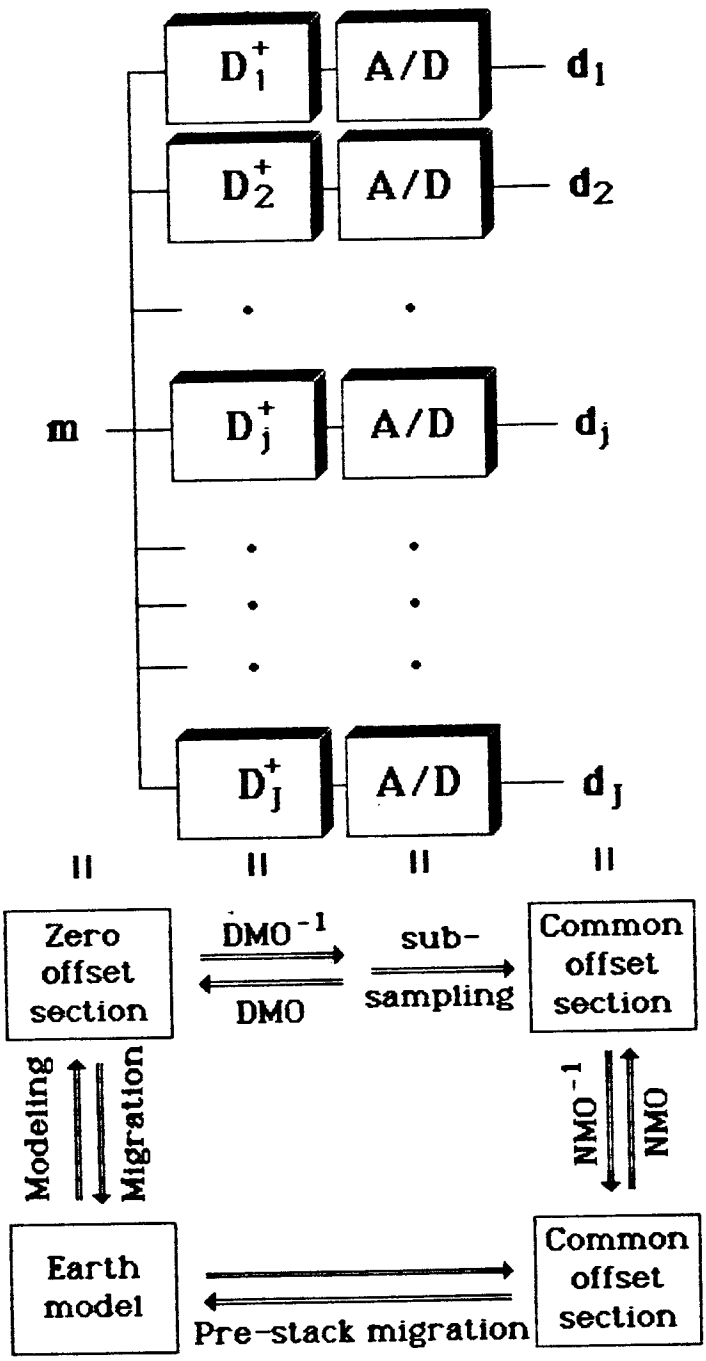


FIG. 2.3.1. The choice of model and data. $\mathbf{m}$ is the zero-offset section; each $\mathbf{d}_j$ is a common-offset section after NMO. The $\mathbf{D}^+{}_j$ operators are inverse DMO.

§ 2.4 Velocity variations

The data are after NMO and the model is unmigrated, as shown in Figure 2.3.1. This leaves only the DMO (and the sampling, **A/D**) between the data and the model. This is to decouple the process of overcoming spatial aliasing, from estimation of velocities: NMO and migration strongly depend on the velocity; DMO depends on velocity only slightly.

The DMO is velocity independent for any constant velocity. In the presence of velocity variations the DMO requires corrections. Rocca (1982) suggested to use a velocity-corrected effective offset in the DMO. Rocca's correction for depth varying velocity is easily applied in both DMO and inverse DMO. Incorporating a correction for lateral velocity variations is also possible (Ronen, 1983). Hale (1983) followed and developed Rocca's suggestion, but showed a minor effect of the correction when it was applied to field data, he (1983) and Bolondi et al (1984) showed that velocity variations have two distinct and opposite effects in connection to DMO, and the overall effect is mild.

§ 2.5 The channel operator — $\mathbf{D}^+$

I show in appendix A that the relation between the time-space Fourier transform of the zero-offset section $m\,(k\,,\omega)$ and the space Fourier transform of a common-offset section $d_j\,(t\,,k\,)$ is the following:

$$d_j\,(t\,,k\,) = \int d\,\omega A^{-1} e^{-i\,\omega A t}\, m\,(\omega,k\,)\,, \tag{2.5.1}$$

where

$$A\ = \sqrt{1 + [h_j\,k\,/\omega t\,]^2}\,. \tag{2.5.2}$$

A discrete form of equation (2.5.1) is

$$d_j\,(t\,,k\,) = \sum_{\omega} A^{-1} e^{-i\,\omega A t}\, m\,(\omega,k\,)\,, \tag{2.5.3}$$

which is a matrix-vector multiplication:

$$\mathbf{d}_j\,(k\,) = \mathbf{D}^+{}_j\,(k\,)\mathbf{m}(k\,)\,. \tag{2.5.4}$$

The vectors are the common-offset section

$$\mathbf{d}_j(k) = \begin{pmatrix} d_j(1,k) \\ \cdot \\ d_j(t,k) \\ \cdot \\ d_j(nt,k) \end{pmatrix}, \tag{2.5.5}$$

and the zero-offset section

$$\mathbf{m}(k) = \begin{pmatrix} m(1,k) \\ \cdot \\ m(\omega,k) \\ \cdot \\ m(n\omega,k) \end{pmatrix}. \tag{2.5.6}$$

The matrix $\mathbf{D}^+$ is the inverse DMO operator: an $nt \times n\omega$ matrix, whose value in the t-th row and the ω-th column is

$$\left[\mathbf{D}^+{}_j(k)\right]_{\omega,t} = A^{-1}e^{-i\omega At} . \tag{2.5.7}$$

§ 2.6 Multichannel inversion in reflection seismology

We can now merge sampling theory (equation 2.1.6) and migration (equation 2.5.4) to formulate multichannel inversion in reflection seismology.

A sampled common-offset section is

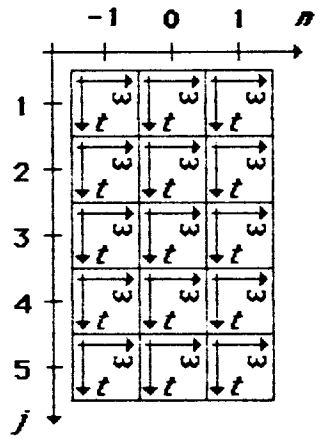
$$\mathbf{d}_j(k) = \sum_n \mathbf{D}^+{}_j(k-n\kappa)\mathbf{m}(k-n\kappa) . \tag{2.6.1}$$

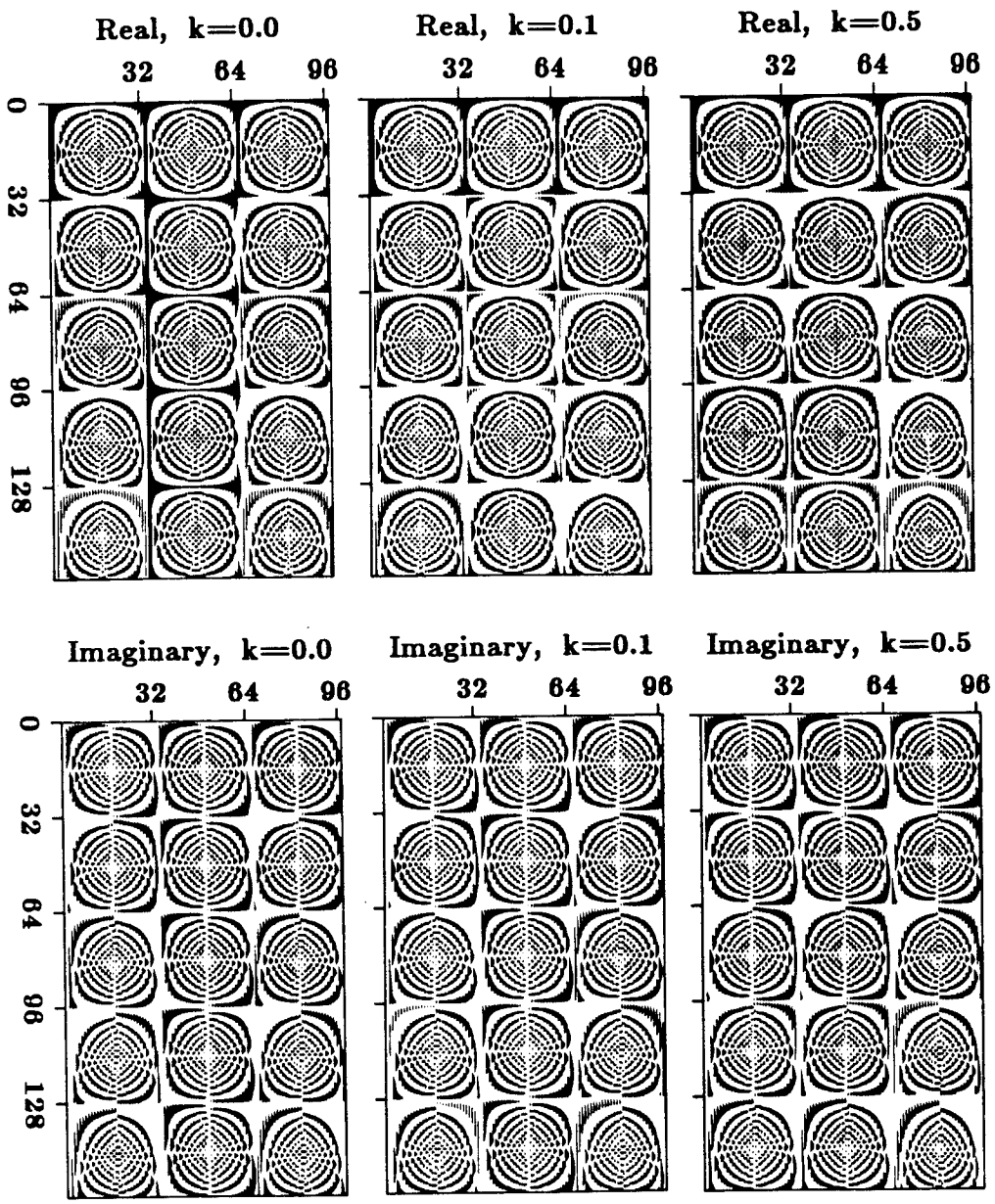
For $J=5$ offsets and aliasing fold $N=3$, the system is

$$\begin{pmatrix} \mathbf{d}_1(k) \\ \mathbf{d}_2(k) \\ \mathbf{d}_3(k) \\ \mathbf{d}_4(k) \\ \mathbf{d}_5(k) \end{pmatrix} = \begin{pmatrix} \mathbf{D}^+{}_1(k-\kappa) & \mathbf{D}^+{}_1(k) & \mathbf{D}^+{}_1(k+\kappa) \\ \mathbf{D}^+{}_2(k-\kappa) & \mathbf{D}^+{}_2(k) & \mathbf{D}^+{}_2(k+\kappa) \\ \mathbf{D}^+{}_3(k-\kappa) & \mathbf{D}^+{}_3(k) & \mathbf{D}^+{}_3(k+\kappa) \\ \mathbf{D}^+{}_4(k-\kappa) & \mathbf{D}^+{}_4(k) & \mathbf{D}^+{}_4(k+\kappa) \\ \mathbf{D}^+{}_5(k-\kappa) & \mathbf{D}^+{}_5(k) & \mathbf{D}^+{}_5(k+\kappa) \end{pmatrix} \begin{pmatrix} \mathbf{m}(k-\kappa) \\ \mathbf{m}(k) \\ \mathbf{m}(k+\kappa) \end{pmatrix} . \tag{2.6.2}$$

The only difference between equation (2.6.2) and the scalar case of equation (2.2.2) is that here there are vectors and matrices, $\mathbf{d}_j$, $\mathbf{m}$, and $\mathbf{D}^+{}_j$, while in equation (2.2.2) there were scalars, d_j, m, and h_j.

FIG. 2.6.1. The matrix of equation (2.6.2), for various spatial frequencies. The matrix has $J \times N$ blocks, the j,n-th block is the inverse DMO operator, $\mathbf{D}^{+}{}_{j}(k - n\kappa)$: a $nt \times n\omega$ matrix, (here $nt = n\omega = 32$, $J = 5$ and $N = 3$).





Equation (2.6.2) is a system of $J \times nt$ equations (one equation for every time sample of every common-offset section), with $N \times n\,\omega$ unknowns (that is N $m\,(k - n\,\kappa)$'s, each $n\,\omega$ long). The matrix has $J \times N$ blocks; each block $\mathbf{D}^{+}{}_{j}\,(k - n\,\kappa)$ is an $nt \times n\,\omega$ matrix. An example of the matrix in equation (2.6.2) is plotted in Figure 2.6.1.

Solving equation (2.6.2) not only continues the data in the *offset* direction as DMO should, but also continues the data in the *wavenumber* , k , direction from the low frequencies $\mid k \mid < \kappa/2$, on which the aliased data $\mathbf{d}_j$ are given, to the full range, $\mid k \mid < N\,\kappa/2$, required to describe the zero-offset section $\mathbf{m}$.

§ 2.7 Singular value decomposition

An obvious requirement for multichannel inversion is that the channels are independent. Whether they are can be studied by looking at the singular value decomposition of the matrix in equation (2.6.2). Singular values of that matrix for some frequencies k are shown in Figure 2.7.1. For $k = 0$ and the $k = \kappa/2$ there is apparently a serious rank deficiency. In fact, there is no problem with those frequencies because $m\,(x\,,t\,)$ is a real function and this imposes constraints on zero and $\kappa/2$ frequencies.

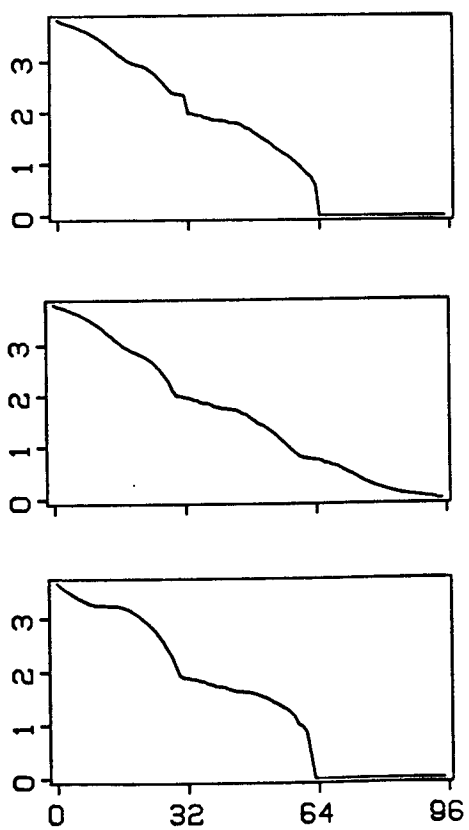


FIG. 2.7.1. Singular values of $\mathbf{G}(k\,)$ for various k 's. Top: $k = 0$. Middle: $k = \kappa/10$. Bottom: $k = \kappa/2$. The 96 singular values are sorted in decreasing order. High value indicates a well resolved feature. Zero indicates a feature that cannot be resolved (null space).

§ 2.8 Time-invariant formulation

When the channel filters are time invariant, multichannel inversion is performed by inverting many small matrices of the size $J \times N$ (scalars), as in equation (2.2.2). If $\mathbf{D}^+$ were time invariant we would solve $nx \times n\omega$ systems of $J \times N$, instead of solving nx systems of $(N \times n\omega) \times (J \times nt)$, as implied by equation (2.6.2).

$\mathbf{D}^+$ is time varying because the shape of its impulse response is

$$\left(\frac{t}{t_0}\right)^2 = \left(1 - \frac{(x-x_0)^2}{h^2}\right)^{-1} , \tag{2.8.1}$$

from equation (A.4.2). The input impulse is at time t_0 and midpoint x_0. The shape (2.8.1) is space invariant because it depends on x_0 only as $x - x_0$, but it is time varying because it depends on t/t_0.

In their derivation of finite differencing DMO, Bolondi et al (1982) used the log transform

$$\tau = \log t , \tag{2.8.2}$$

which makes the impulse response τ invariant:

$$\left(e^{\tau-\tau_0}\right)^2 = \left(1 - \frac{(x-x_0)^2}{h^2}\right)^{-1} , \tag{2.8.3}$$

where τ_0 is $\log t_0$.

To find the transfer function whose impulse-response shape is given by equation (2.8.3), start from equation (2.5.1):

$$\begin{aligned} d_j(t,k) &= \int_{-\infty}^{\infty} d\omega A^{-1} e^{-i\omega At}\, m(\omega,k) \qquad (2.8.4) \\ &= \int_{-\infty}^{\infty} d\omega A^{-1} e^{-i\omega At} \int_0^{\infty} dt'\; e^{i\omega t'}\; m(t',k) \\ &= \int_0^{\infty} dt'\; m(t',k) \int_{-\infty}^{\infty} d\omega\, A^{-1}\, e^{i\omega[t'-At]} . \end{aligned}$$

A is $\sqrt{1+(hk/\omega t)^2}$. Use the log transform (2.8.2) to obtain

$$d_j(e^{\tau},k) = \int_{-\infty}^{\infty} d\tau'\; e^{\tau'}\, m(e^{\tau'},k) \int_{-\infty}^{\infty} d\omega\, A^{-1}\, e^{i\omega[e^{\tau'}-Ae^{\tau}]} . \tag{2.8.5}$$

A is $\sqrt{1+(hk/\omega e^{\tau})^2}$. Following Owusu and Gardner (1983), use the log transform:

$$\hat{d}_j(\tau,k) = e^{\tau} d_j(e^{\tau},k) , \tag{2.8.6}$$

and

$$\hat{m}(\tau',k) = e^{\tau'} m(e^{\tau'},k)$$

(note that a gain term e^{τ} is included), to get

$$\hat{d}_j(\tau,k) = e^{\tau} \int_{-\infty}^{\infty} d\tau' \ \hat{m}(\tau',k) \int_{-\infty}^{\infty} d\omega \, A^{-1} \, e^{i\omega[e^{\tau'}-Ae^{\tau}]} \tag{2.8.7}$$

$$= \int_{-\infty}^{\infty} d\tau' \ \hat{m}(\tau',k) \, I(\tau,\tau',k) ,$$

where

$$I(\tau,\tau',k) = e^{\tau} \int_{-\infty}^{\infty} d\omega \, A^{-1} \, e^{i\omega[e^{\tau'}-Ae^{\tau}]} . \tag{2.8.8}$$

Change the integration variable

$$\theta = \omega e^{\tau} ,$$

then

$$I(\tau,\tau',k) = \int_{-\infty}^{\infty} d\theta \, A^{-1} \, e^{i\theta[e^{-(\tau-\tau')}-A]} ,$$

where $A = \sqrt{1+(hk/\theta)^2}$. $I(\tau,\tau',k)$ is therefore a function of $\tau-\tau'$ (and k). Define:

$$f_j(\tau,k) = I(\tau,\tau'=0,k) , \tag{2.8.9}$$

and then equation (2.8.7) can be written as the convolution

$$\hat{d}_j(\tau,k) = f_j(\tau,k) * \hat{m}(\tau,k) . \tag{2.8.10}$$

Fourier transforming over τ we obtain

$$\hat{d}_j(\Omega,k) = F_j(\Omega,k) \cdot \hat{m}(\Omega,k) , \tag{2.8.11}$$

where the transfer function F_j is the Fourier transform of f_j.

Including spatial aliasing, we have

$$\hat{d}_j(\Omega,k) = \sum_n F_j(\Omega,k-n\kappa) \, \hat{m}(\Omega,k-n\kappa) . \tag{2.8.12}$$

The advantage of the time-invariant formulation (2.8.12), compared to the block matrix inversion of equation (2.6.2), is efficiency. Some disadvantages are that this efficiency does not extent to 3-D, velocity variations cannot be incorporated, and some nasty integrals need to be evaluated.

I did not develop the time-invariant formulation (2.8.12), but applied the more straightforward block inversion of equation (2.6.2).

§ 2.9 Missing and muted data

Missing data are often treated as zero data in seismic processing. However, inversion fits a model to the data, and if an unreasonable part of the data are zero, the inversion will find an unreasonable model. The right way of treating missing data is to invert the operator that caused their loss.

Suppose there is a dead trace at x_d, due to a skipped shot or a dead receiver. The trace is dead because a filter

$$\text{miss}\,(x) = \begin{cases} 1 & \text{for } x \neq x_d \\ 0 & \text{for } x = x_d \end{cases} \tag{2.9.1}$$

was applied, in addition to the wave propagation filter $\mathbf{D}^+{}_j$ and the sampling $\text{III}(x/\Delta x)$. The data are

$$\mathbf{d}_j(x) = \text{III}(x/\Delta x)\ \text{miss}\,(x)\ \mathbf{D}^+{}_j(x)\ \mathbf{m}(x)\ . \tag{2.9.2}$$

In the frequency domain,

$$\mathbf{d}_j(k) = \sum_n \text{MISS}\,(k - n\kappa)\ \mathbf{D}^+{}_j(k - n\kappa)\ \mathbf{m}(k - n\kappa)\ , \tag{2.9.3}$$

where

$$\text{MISS}\,(k) = 1 - e^{ikx_d}\ . \tag{2.9.4}$$

Mute, due to NMO stretch, or any other reason should be treated as missing data. If the mute is the same in all midpoints there is no need to actually enter the MISS (k) filters. Instead, we allow a varying length of the $\mathbf{d}_j$ data vectors in equation (2.6.2). The NMO mute zone is not considered data and does not enter into the multichannel inversion.

§ 2.10 Interlaced sampling

In most recording geometries different common-offset sections cover different midpoints. The j-th common-offset section can be displaced by x_j relative to some global reference:

$$\mathbf{d}_j(x) = \mathrm{III}[(x - x_j)/\Delta x]\, \mathbf{D}^+{}_j(x)\, \mathbf{m}(x) . \tag{2.10.1}$$

In the frequency domain, the shift operator $\exp[ikx_j]$ must be entered into the multichannel inversion as part of the operator

$$\mathbf{d}_j(k) = \sum_n \left[e^{i(k - n\kappa)x_j}\, \mathbf{D}^+{}_j(k - n\kappa) \right] \mathbf{m}(k - n\kappa) . \tag{2.10.2}$$

§ 2.11 Moving sources/receivers (marine)

Land data are collected with stationary sources and receivers. Boats, on the other hand, do not stand still in seismic surveys; the receivers move a considerable distance while recording. This can easily be entered into the operator by letting x_j in equation (2.10.2) be

$$x_j = x_{0j} + V_b\, t , \tag{2.11.1}$$

and the offset h_j in equation (2.6.1) be

$$h_j = h_{0j} - V_b\, t/2 , \tag{2.11.2}$$

where t is the recording time and V_b is the boat velocity.

§ 2.12 Deconvolving receiver groups

Each trace in reflection data is rarely recorded by a single receiver, rather it is the sum of a group of adjacent receivers. This improves the signal-to-noise ratio but the spatial resolution is reduced because the data are a moving average of the wavefield instead of the wavefield itself.

The smear effect of receiver groups can be partially removed by multichannel inversion. For example suppose there are three receivers in each group. The common-offset section $\mathbf{d}_j(x)$, recorded by the three receivers at half-offsets $h_j - \Delta$, h_j and $h_j + \Delta$ is the sum of the three common-offset (CO) sections:

$$\mathbf{d}_j(x) = \mathrm{CO}(h_j - \Delta, x + \Delta) + \mathrm{CO}(h_j, x) + \mathrm{CO}(h_j + \Delta, x - \Delta) . \tag{2.12.1}$$

Fourier transforming over x we obtain

$$\begin{aligned}\mathbf{d}_j(k) &= e^{ik\Delta}\,\mathrm{CO}(h_j-\Delta,k) + \mathrm{CO}(h_j,k) + e^{-ik\Delta}\,\mathrm{CO}(h_j+\Delta,k) \\ &= \mathbf{D}^+{}_j(k)\,\mathbf{m}(k)\,,\end{aligned} \tag{2.12.2}$$

where the operator is

$$\mathbf{D}^+{}_j = e^{ik\Delta}\left[\mathbf{D}^+(k)\right]_{h_j-\Delta} + \left[\mathbf{D}^+(k)\right]_{h_j} + e^{-ik\Delta}\left[\mathbf{D}^+(k)\right]_{h_j+\Delta}. \tag{2.12.3}$$

§ 2.13 Integral formulation

The matrix form of aliasing in §2.1 describes uniform undersampling. A more general form of undersampling is nonuniform.

Equation (2.5.1) can be written as

$$d_j(t,x) = \int dk e^{ikx} \int d\omega A^{-1} e^{-i\omega At}\, m(\omega,k)\,, \tag{2.13.1}$$

which is a linear system of equations with $m(\omega,k)$ as the unknowns. Compared with the uniform-sampling multichannel inversion of equation (2.6.2) we have here only one system of equations to solve instead of many (equation (2.6.2) had to be solved for every k), but equation (2.13.1) is nx times bigger than any one of (2.6.2).

We may solve (2.13.1) for $m(\omega,k)$, or use the Fourier transform:

$$m(\omega,k) = \int dt' e^{i\omega t'} \int dx' e^{-i\omega x'}\, m(t',x) \tag{2.13.2}$$

to solve directly for the model in the (t,x) domain, by substituting equation (2.13.2) into (2.13.1):

$$d_j(t,x) = \int dk e^{ikx} \int d\omega A^{-1} e^{-i\omega At} \int dt' e^{i\omega t'} \int dx' e^{-i\omega x'}\, m(t',x). \tag{2.13.3}$$

The boundaries of the integrals with respect to t' and x' correspond, more or less, to the size of the DMO impulse response.

§ 2.14 Multichannel inversion in three dimensions

In 3-D, there are two spatial axes, x and y. The half-offset $\mathbf{h}$ and the spatial frequency $\mathbf{k}$ are the vectors

$$\mathbf{h} = \begin{pmatrix} h_x \\ h_y \end{pmatrix} \quad \text{and} \quad \mathbf{k} = \begin{pmatrix} k_x \\ k_y \end{pmatrix}. \tag{2.14.1}$$

Aliasing may occur in both x and y directions, so the 3-D equivalent of equation (2.6.1) is

$$\mathbf{d_h}(k_x,k_y) = \sum_{n_x}\sum_{n_y}\mathbf{D_h^+}(k_x-n_x\kappa_x,k_y-n_y\kappa_y)\mathbf{m}(k_x-n_x\kappa_x,k_y-n_y\kappa_y)\,, \quad (2.14.2)$$

Where κ_x is $2\pi/\Delta x$, Δx is the sampling interval in x, κ_y is $2\pi/\Delta y$, and Δy is the sampling interval in y. $\mathbf{D_h^+}$ is the same as $\mathbf{D^+}_j$ in equation (2.6.1), only now

$$A^2 = 1 + \left(\frac{\mathbf{h}\cdot\mathbf{k}}{\omega t}\right)^2 \quad (2.14.3)$$

$$= 1 + \left(\frac{h_x(k_x-n_x\kappa_x)+h_y(k_y-n_y\kappa_y)}{\omega t}\right)^2 .$$

In the 2-D case (equation [2.5.2]) the offset vector was in the x direction ($h_y=0$).

§ 2.15 Summary

Processing that is optimal when there is no aliasing may be totally inadequate in the presence of aliasing.

Multichannel inversion in reflection seismology is based on sampling theory and the wave equation. The essential theory is summarized in equation (2.6.1).