Chapter 7

Velocity and dip moveout

7.1 STACKING AND VELOCITY ANALYSIS

Hyperbolic stacking over offset may be the most important computer process in the prospecting industry. It is more important than migration because it reduces the data base from a volume in (s, g, t)-space to a plane in (y, t)-space. At the present time few people who interpret seismic data have computerized seismic data movies, so most interpreters must have their data stacked before they can even look at it. Migration merely converts one plane to another plane. Furthermore, migration has the disadvantage that it sometimes compounds the mess made by near-surface lateral velocity variation and multiple reflections. Stacking can compound the mess too, but in bad areas nothing can be seen until the data is stacked. In addition to its other drawing points, stacking gives as a byproduct estimates of rock velocity.

Historically, stacking has been done using ray methods, and it is still being done almost exclusively in this way. Migration, on the other hand, is more often done using wave-equation methods, that is to say, by Fourier or finite-difference methods. Both migration and stacking are hyperbola-recognition processes. The advantages of wave-equation methods in migration have been many. Shouldn't these advantages apply equally to stacking? It would seem so, but current industrial practice does not bear this out. The reasons are not yet clear. So the latter part of this section really belongs to a research monograph with the facetious title "Theory That Should Work Out Soon." More advanced ideas of velocity estimation are later. Wave-equation stacking and velocity-determination methods are ingenious. Perhaps they have not yet been satisfactorily tested, or perhaps they are just imperfectly assembled. The reader can guess, and time will tell.

One possible reason why much of this theory is not in routine industrial use is that the issue of stacking to remove redundancy may be more appropriately a statistical problem than a physical one. To allow for this contingency I have included a bit on "wave-equation moveout," a way of deferring statistical analysis until after downward continuation. Another possibility is that the problems of missing data off the ends of the recording cable and spatial aliasing within the cable may be more flexibly attacked by ray methods than by wave-equation methods. For this contingency I have included a brief subsection on data restoration. Whatever the case, the data-manipulation procedures in this chapter should be helpful.

7.1.1 Normal moveout (NMO)

Normal moveout correction (NMO) is a stretching of the time axis to make all seismograms look like zero-offset seismograms. In its simplest form, NMO is based on the Pythagorean relation $t_{NMO}^2 = t^2 - x^2/v^2$. In a constant velocity earth, the NMO correction would take the asymptote of the hyperbola family and move it up to $t = 0$. This abandons anything on the time axis before the first arrival, and stretches the remainder of the seismogram. The stretching is most severe near the first arrival, and diminishes at later times. In the NMO example in Figure 1.1 you will notice the low frequencies caused by the stretch.

NMO correction may be done to common-shot field profiles or to CMP gathers. NMO applied to a field profile makes it resemble a small portion of a zero-offset section. Then geologic structure is prominently exhibited. NMO on a CMP gather is the principal means of determining the earth's velocity-depth function. This is because CMP gathers are insensitive to earth dip.

Mathematically, the NMO transformation is a *linear* operation. It may seem paradoxical that a non-uniform axis-stretching operation is a linear operation, but axis stretching does satisfy the mathematical conditions of linearity. Do not confuse the widespread linearity condition with the less common condition of time invariance. Linearity requires only that for any decomposition of the original data P into parts (say P_1 and P_2) the sum of the NMOed parts is equal the NMO of the sum. Examples of decompositions include: (1) separation into early times and late times, (2) separation into even and odd time points, (3) separation into high frequencies and low frequencies, and (4) separation into big signal values and small ones.

To envision NMO as a linear operator, think of a seismogram as a vector. The NMO operator resembles a diagonal matrix, but the matrix contains *interpolation filters* along its diagonal, and the interpolation filters are shifted off from the diagonal to create the desired time delay.

7.1.2 Conventional velocity analysis

A conventional velocity analysis uses a collection of trial velocities. Each trial velocity is taken to be a constant function of depth and is used to moveout correct the data. Figure 1.2 (left) exhibits the CMP gather of Figure 1.1 (left) after moveout correction by a constant velocity. Notice that the events in the middle of the gather are nearly flattened, whereas the early events are undercorrected and later events are overcor-

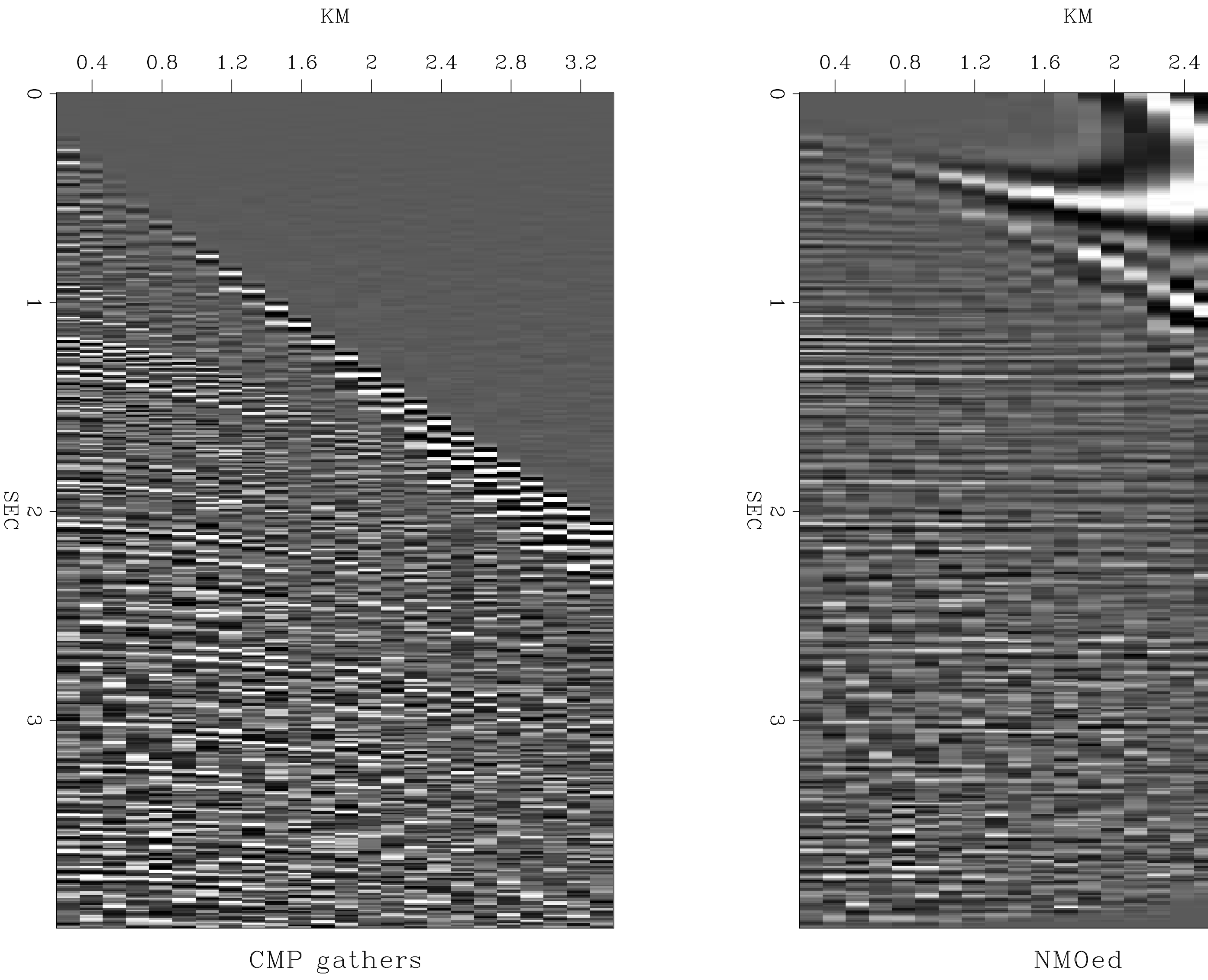

Figure 7.1: CMP gather (Western Geophysical) from the Gulf coast shown at the left was NMO corrected and displayed at the right. vdmo-cmpnmo [ER]

rected. This is typical because the amount of moveout correction varies inversely with velocity (by Pythagoras), and the earth's velocity normally increases with depth. A measure of the goodness of fit of the NMO velocity to the earth velocity is found by summing the CDP gather over offset. Presumably, the better the velocities match, the better (bigger) will be the sum. The process is repeated for many velocities. The amplitude of the sum, contoured as a function of time and velocity, is shown in Figure 1.2 (right).

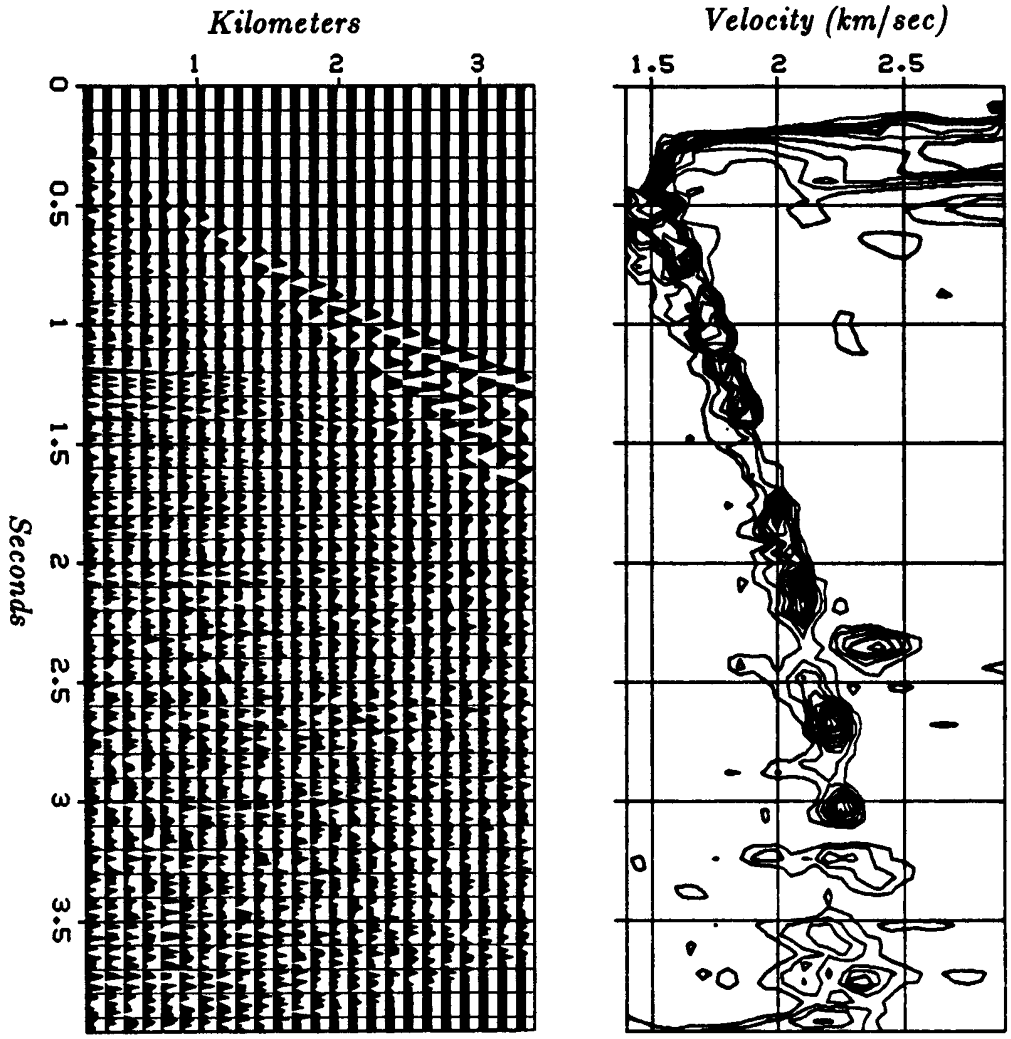

Figure 7.2: NMO at constant velocity with velocity analysis. (Hale) vdmo-cmphale [NR]

In practice additional steps may be taken before summing. The traces may be balanced (scaled to be equal) in their powers and in their spectra. Likewise the amplitude of the sum may be normalized and smoothed. (See Taner and Koehler [1969]). Also the data may be *edited* and *weighted* as explained in the next subsection.

The velocity giving the best stack is an average of the earth's velocity above the reflector. The precise definition of this average is deferred till later.

7.1.3 Mutes and weights

An important part of conventional processing is the definition of a *mute.* A mute is a weighting function used to suppress some undesirable portions of the data. Figure 1.3 shows an example of a muted field profile. Weights and mutes have a substantial effect

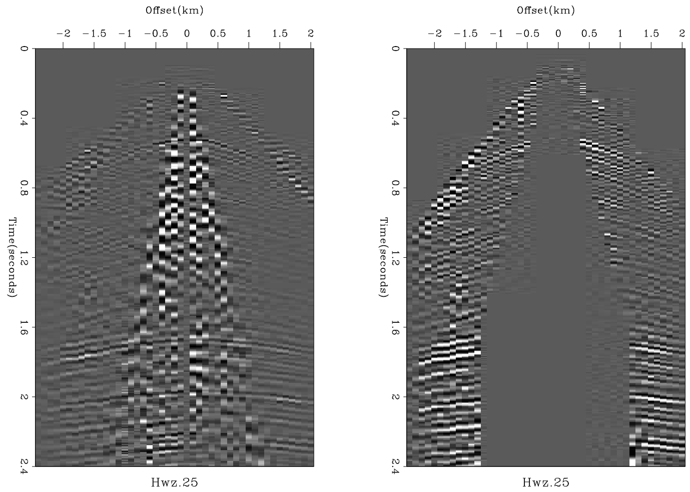

Figure 7.3: Left is a land profile from Alberta (Western Geophysical). On the right it is muted to remove ground roll (at center) and head waves (the first arrivals). vdmo-mute [ER]

on the quality of a stack. So it is not surprising that in practice, they are the subject of much theorizing and experimentation.

Often the mute is a one-dimensional function of $r = h/t$. Reasons can be given to mute data at both large and small values of r.

At small values of r, energy is found that remains near the shot, such as falling dirt or water or slow ground roll.

At large values of r, there are problems with the first arrival. Here the NMO stretch is largest and most sensitive to the presumed velocity. The first arrival is often called a *head wave* or *refraction* . Experimentally, a head wave is a wave whose travel time appears to be a linear function of distance. Theoretically, a head wave is readily defined for layered media. The head wave has a ray that propagates horizontally along a layer boundary. In practice, a head wave may be weaker or stronger than the reflections. A strong head wave may be explained by the fact that reflected waves spread in three dimensions, while head waves spread in only two dimensions.

Muting may be regarded as weighting by zero. More general weights may be chosen to produce the most favorable CDP stack. A sophisticated analysis would certainly include noise and truncation. Let us do a simplified analysis. It leads to the most basic weighting function.

Ordinarily we integrate over offset along a hyperbola. Instead, think of the *three* -dimensional problem. You really wish to integrate over a hyperbola of revolution. Assume that the hyperboloid is radially symmetric. Weighting the integrand by h allows the usual line integral to simulate integration over the hyperboloid of revolution. A second justification for scaling data by offset h before stacking is that there is less velocity information near zero offset, where there is little moveout, and more velocity information at wider offset where $\Delta t / \Delta h$ is larger.

7.1.4 NMO equations

The earth's velocity typically ranges over a factor of two or more within the depth range of a given data set. Thus the Pythagorean analysis needs reexamination. In practice, depth variable velocity is often handled by inserting a time variable velocity into the Pythagorean relation. (The classic reference, Taner and Koehler [1969], includes many helpful details). This approximation is much used, although it is not difficult to compute the correct nonhyperbolic moveout. Let us see how the velocity function $v(z)$ is mathematically related to the NMO. Ignoring dip, NMO converts common-midpoint gathers, one of which, say, is denoted by $P(h,t)$, to an earth model, say,

$$Q(h,z) \quad = \quad \mathrm{earth}(z) \ \times \ \mathrm{const}(h) \tag{7.1}$$

Actually, $Q(h,z)$ doesn't turn out to be a constant function of h, but that is the goal.

The NMO procedure can be regarded as a simple copying. Conceptually, it is easy to think of copying every point of the (h,t)-plane to its appropriate place in the (h,z)-plane. Such a copying process could be denoted as

$$Q[h, z(h,t)] \quad = \quad P(h,t) \tag{7.2}$$

Care must be taken to avoid leaving holes in the (h,z)-plane. It is better to scan every point in the output (h,z)-plane and find its source in the (h,t)-plane. With a table $t(h,z)$, data can be moveout corrected by the copying operation

$$Q(h,z) \quad = \quad P[h, t(h,z)] \tag{7.3}$$

Using the terminology of this book, the input $P(h,t)$ to the moveout correction is called a CMP gather, and the output Q is called a CDP gather.

In practice, the first step in generating the travel-time tables is to change the depth-variable z to a vertical travel-time-variable τ. So the required table is $t(h,\tau)$. To get the output data for location (h,τ) you take the input data at location (h,t). The most straightforward and reliable way to produce this table seems to be to march down in steps of z, really τ, and trace rays. That is, for various fixed values of Snell's parameter p, you compute $t(p,\tau)$ and $h(p,\tau)$ from $v(\tau)$ by integrating the following equations over τ:

$$\frac{dt}{d\tau} \quad = \quad \frac{dz}{d\tau}\,\frac{dt}{dz} \quad = \quad v\,\frac{1}{v\ \cos\ \theta} \quad = \quad \frac{1}{\sqrt{1\ -\ p^2\,v(\tau)^2}} \tag{7.4}$$

$$\frac{dh}{d\tau} \quad = \quad \frac{dz}{d\tau}\,\frac{dh}{dz} \quad = \quad v\ \tan\ \theta \quad = \quad \frac{p\,v(\tau)^2}{\sqrt{1\ -\ p^2\,v(\tau)^2}} \tag{7.5}$$

(In equations (1.4) and (1.5) dt/dz and dh/dz are based on rays, not wavefronts). Given $t(p,\tau)$ and $h(p,\tau)$, iteration and interpolation are required to eliminate p and find $t(h,\tau)$. It sounds awkward—and it is—because at wide angles there usually are head waves arriving in the middle of the reflections. But once the job is done you can save the table and reuse it many times. The multibranching of the travel time curves at wide offset motivates a wave-equation based velocity analysis. The greatest velocity sensitivity occurs just where the classic hyperbolic assumption and the single-arrival assumption break down.

7.1.5 Linearity allows postponing statistical estimation

The linearity of wave-equation data processing allows us to decompose a dataset into parts, process the parts separately, then recombine them. The result is the same as if they were never separated.

For example, suppose a CMP gather is divided into two parts, say, inner traces A and outer traces B. Let $(A,0)$ denote a CMP gather where the *outer* traces have been replaced by zeroes. Likewise, $(0,B)$ could be another copy of the gather where the *inner* traces have been replaced by zeroes. We could downward continue $(A,0)$ and separately downward continue $(0,B)$. After downward continuation, $(A,0)$ and $(0,B)$ could be added. Alternately, we could pause, do some thinking about statistics, and then choose to combine them with some weighting function. Figure 1.4 shows a dataset of three traces decomposed into three datasets, one for each trace. Semicircles depict the separate downward continuation of each trace. Each semicircle goes through zero offset, giving the appropriately stretched, NMOed trace.

The idea of using a weighting function is a drastic departure from our previous style of analysis. It represents a disturbing recognition that we have been neglecting something important in all scientific analysis, namely, statistics! What are the ingredients that go into the choice of a weighting function? They are many. Signal and noise variances play a role. Some channels may be noisy or absent. When final display is contemplated, it is necessary to consider human perception and the need to compress the dynamic range, so that small values can be perceived. Dynamic-range compression must be considered not only in the obvious (h,t)-space, but also in frequency space, dip space, or any other space in which the wavefields may get too far out of balance.

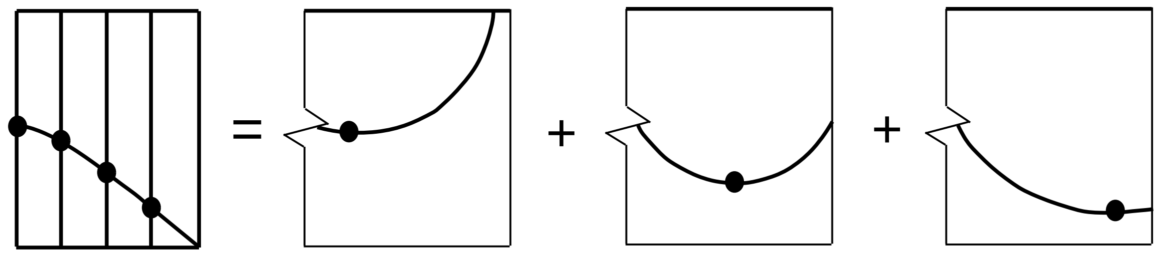

Figure 7.4: A three-trace CMP gather decomposed by traces. At the left, impulses on the data are interpolated, depicting a hyperbola. At the right, data points are expanded into migration semicircles, each of which goes through zero offset at the apex of the hyperbola. vdmo-decomp [NR]

There are many ways to decompose a dataset. The choice depends on your statistical model and your willingness to repeat the processing many times. Perhaps the parts of the data gather should be decomposed not by their h values but by their values of $r = h/t$. Clearly, there is a lot to think about.

7.1.6 Lateral interpolation and extrapolation of a CMP gather

Practical problems dealing with common-midpoint gathers arise because of an insufficient number of traces. *Truncation* problems are those that arise because the geophone cable has a fixed length that is not as long as the distance over which seismic energy propagates. Figure 1.5 shows why cable truncations are a problem for conventional, ray-trace, stacking methods as well as for wave-equation methods. *Aliasing* problems are those that arise because shots and geophones are not close enough together. Spatial aliasing of data on the offset axis seems to be a more serious problem for wave-equation methods than it is for ray-trace methods. The reason is that normal-moveout correction reduces the spatial frequencies. *Gaps* in the data, resulting from practical problems with the geophones, cable, and access to the terrain, are also frequently a snag.

Here these problems will all be attacked together with a systematic approach to estimating missing traces. The technique to be described is the simplest member of a more general family of *missing data* estimation procedures currently being developed at the Stanford Exploration Project.

First do normal-moveout correction, that is, stretch the time axis to flatten hyperbolas. The initial question is what velocity to use for the normal-moveout correction. For trace interpolation the appropriate moveout velocity turns out to be that of the *dominating energy on the gather.* On a given dataset this velocity could be primary velocity at some times and multiple velocity at other times. The reason for such a nonphysical velocity is this: the strong events must be handled well, in order to save the weak ones. Truncations of weak events can be ignored as a "second-order"

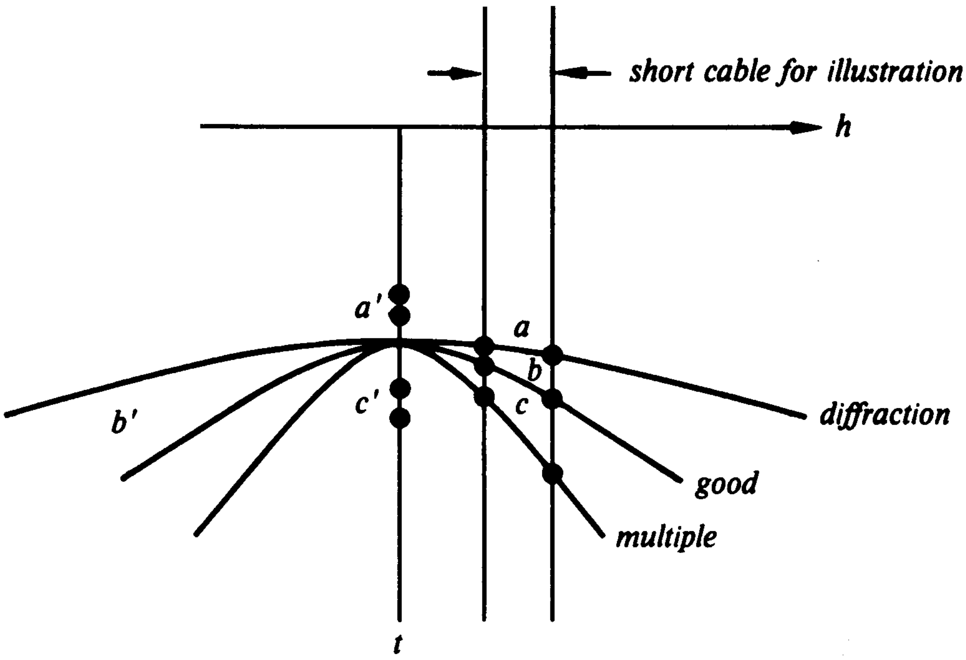

Figure 7.5: Normal moveout at the earth velocity brings the cable truncations on good events to a good place, causing no problems. The cable truncations of diffractions and multiples, however, move to a' and c', where they could be objectionable. Such corruption could make folly of sophisticated time-series analysis of the waveform found on a CDP stack. vdmo-nmotrunc [NR]

problem. The practical problem is usually to suppress strong water-velocity events in the presence of weak sedimentary reflections, particularly at high frequencies. In principle, we might be seeking weak $P-SV$ waves in the presence of strong $P-P$ waves.

After NMO, the residual energy should have little dip, except of course where missing data, now replaced by zeroes, forces the existing data to be broad-banded in spatial frequency. In order to improve our view of this badly behaved energy, we pass the data through a "badpass" filter, such as the high-pass recursive dip filter.

$$\frac{\frac{k^2}{-i\omega}}{\alpha \quad + \quad \frac{k^2}{-i\omega}} \tag{7.6}$$

Notice that this filter greatly weakens the energy with small k, that is, the energy that was properly moveout corrected. On the other hand, near the missing traces, notice that the spectrum should be broad-band with k and that such energy passes through the filter with almost unit gain.

The output from the "badpass" filter is now ready to be subtracted from the data. The subtraction is done selectively. Where recorded data exists, nothing is subtracted. This completes the first iteration. Next the steps are repeated, and iterated. Convergence is finally achieved when nothing comes out of the badpass filter at the locations where data was not recorded. An example of this process can be found in Figure 1.6.

The above procedure has ignored the possibility of dip in the midpoint direction.

This procedure is also limited because it ignores the possibility that several velocities may be simultaneously present on a dataset. To really do a good job of extending such a dataset may require a parsimonious model and a velocity spectral concept such as the ones developed later.

7.1.7 In and out of velocity space

Summing a common-midpoint gather on a hyperbolic trajectory over offset yields a stack called a *constant-velocity* or a *CV* stack. A velocity space may be defined as a family of CV stacks, one stack for each of many velocities. CV stacking is a transformation from *offset* space to *velocity* space. CV stacking creates a (t, v)-space velocity panel from a (t, h)-space common-midpoint gather. Conventional industrial velocity estimation amounts to CV stacking supplemented by squaring and normalizing. Linear transformations such as CV stack are generally *invertible,* but the transformation to velocity space is of very high dimension. Forty-eight channels and 1000 time points make the transformation 48,000-dimensional. With present computer technology, matrices this large cannot be inverted by algebraic means. However, there are some excellent approximate means of inversion.

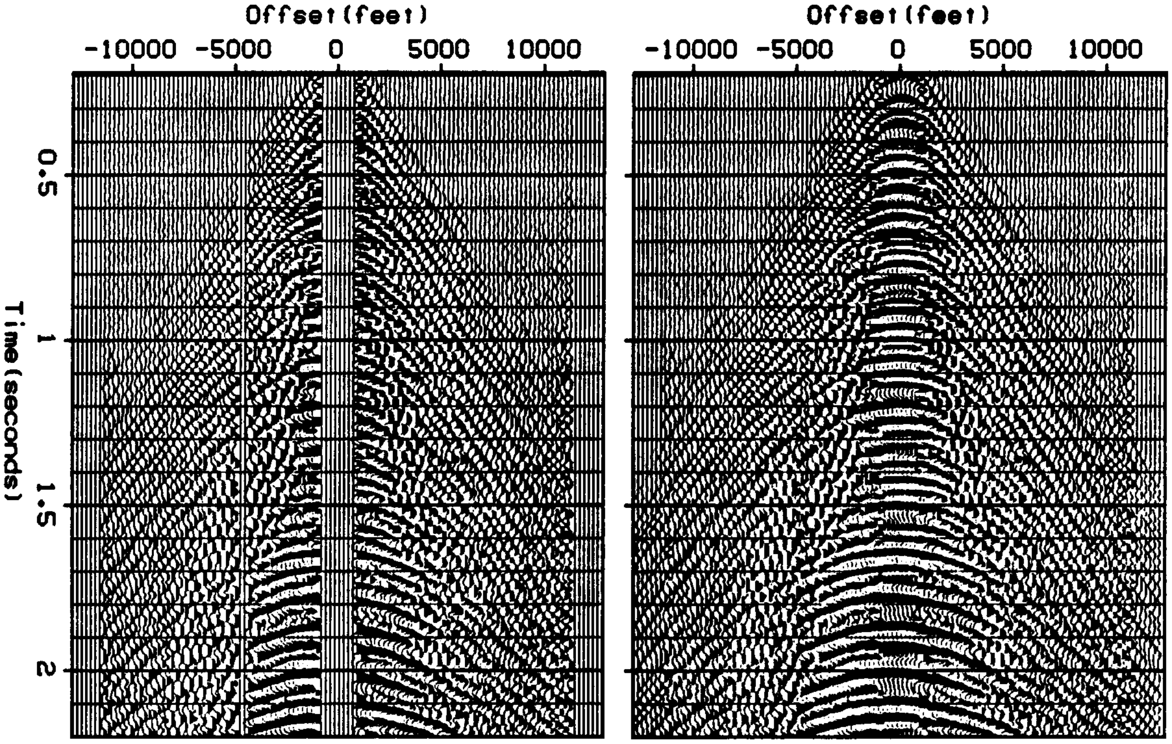

Figure 7.6: Field profile from Alaska with missing channels on the left (Western Geophysical), restored by iterative spatial filtering on the right. (Harlan) vdmo-missfil [NR]

For unitary matrices, the transpose matrix equals the inverse matrix. In wave-propagation theory, a *transpose* operator is often a good approximation to an *inverse* operator. Thorson [1984] pointed out that the transpose operation to CV stacking is just about the same thing as CV stacking itself. To do the operation transposed to CV stacking, begin with a velocity panel, that is, a panel in (t, v)-space. To create some given offset h, each trace in the (t, v)-panel must be first compressed to undo the original NMO stretch. That is, events must be pushed from the zero-offset time that they have in the (t, v)-panel to the time appropriate for the given h. Then stack the (t, v)-panel over v to produce the seismogram for the given h. Repeat the process for all desired values of h. The program for transpose CV stack is like the program for CV stack itself, except that the stretch formula is changed to a compensating compression.

The inversion of a CV stack is analogous to inversion of slant stack or Radon transformation. That is, the CV stack is almost its own inverse, but you need to change a sign, and at the end, a filtering operation, like rho filtering, is also needed to touch up the spectrum, thereby finishing the job. It is the rho filtering that distinguishes *inverse* CV stack from the *transpose* of CV stack.

The word *transpose* refers to *matrix transpose.* It is difficult to visualize why the word transpose is appropriate in this case because we are discussing data spaces that are two-dimensional and operators that are four-dimensional. But if you will map these two- and four-dimensional objects to familiar one- and two-dimensional objects by a transformation then you will see that the word *transpose* is entirely appropriate. The rho-type filtering required for CV stacks is slightly more complicated than ordinary rho filtering—refer to Thorson [1984].

Figure 1.7 shows an example of Thorson's velocity space inversion. Panel D is the original common-midpoint gather. Next is panel LU, the approximate reconstruction of D from velocity space. The hyperbolic events are reconstructed much better than the random noise. The random noise was not reconstructed so well because the range of velocities in the CV stack was limited between water velocity and 3.5 km/sec. The next two panels (U) and (LTD) are theoretically related by the "rho" filtering. LTD is the CV stack of D. LU is the transpose CV stack of U.

It is worth noting that there is a substantial amount of work in computing a velocity panel. A stack must be computed for each velocity. Velocity discrimination by wave-equation methods will be described next. The wave-equation methods are generally cheaper, though not fully comparable in effect.

7.1.8 The (z,t)-plane method

In the 15° continuation equation $U_{zt} = -\frac{1}{2}\, v\, U_{hh}$, scaling the depth z is indistinguishable from scaling the velocity. Thus, downward continuation with the wrong velocity is like downward continuation to the wrong depth. Stephen M. Doherty [1975] used this idea in a velocity-estimation scheme—see Figure 1.8.

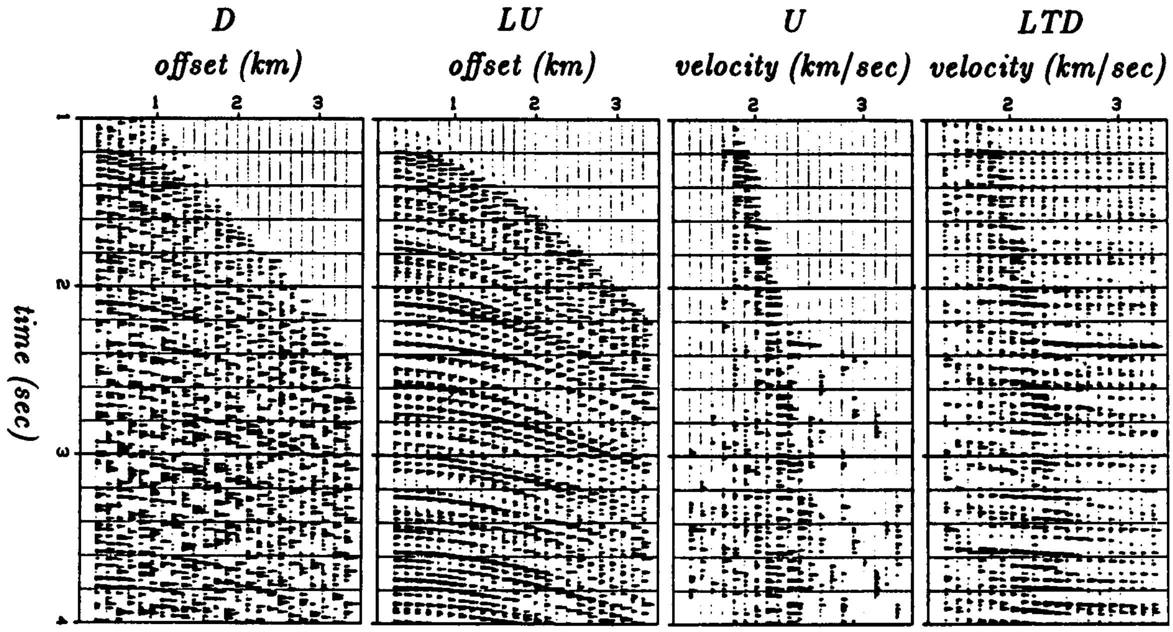

Figure 7.7: Panel D at the left is a CMP gather from the Gulf of Mexico (Western Geophysical). The second panel (LU) is reconstructed data obtained from the third panel (U) by inverse NMO and stack. The last panel (LTD) is a CV stack of the first panel. (Thorson) vdmo-thor [NR]

The idea is to downward continue with a preliminary velocity model and to display the zero-offset trace, a function of t', at all travel-time depths τ. If the maximum amplitude occurs at $t' = \tau$, then your preliminary model is good. If the maximum is shifted, then you have some analysis to do before you can say what velocity should be used on the next iteration.

7.1.9 Decomposing a gather by velocity

A process will be defined that can partition a CMP gather, *both reflections and head waves,* into one part with RMS velocity greater than that of some given model $\bar{v}(z)$ and another part with velocity less than $\bar{v}(z)$.

After such a partitioning, the low-velocity noise could be abandoned. Or the earth velocity could be found through iteration, by making the usual assumption that the velocity spectrum has a peak at earth velocity. As will be seen later, various data interpolation, lateral extrapolation, and other statistical procedures are also made possible by the linearity and invertibility of the partitioning of the data by velocity.

The procedure is simple. Begin with a common-midpoint gather, zero the negative offsets, and then downward continue according to the velocity model $\bar{v}(z)$. The components of the data with velocity less than $\bar{v}(z)$ will overmigrate through zero

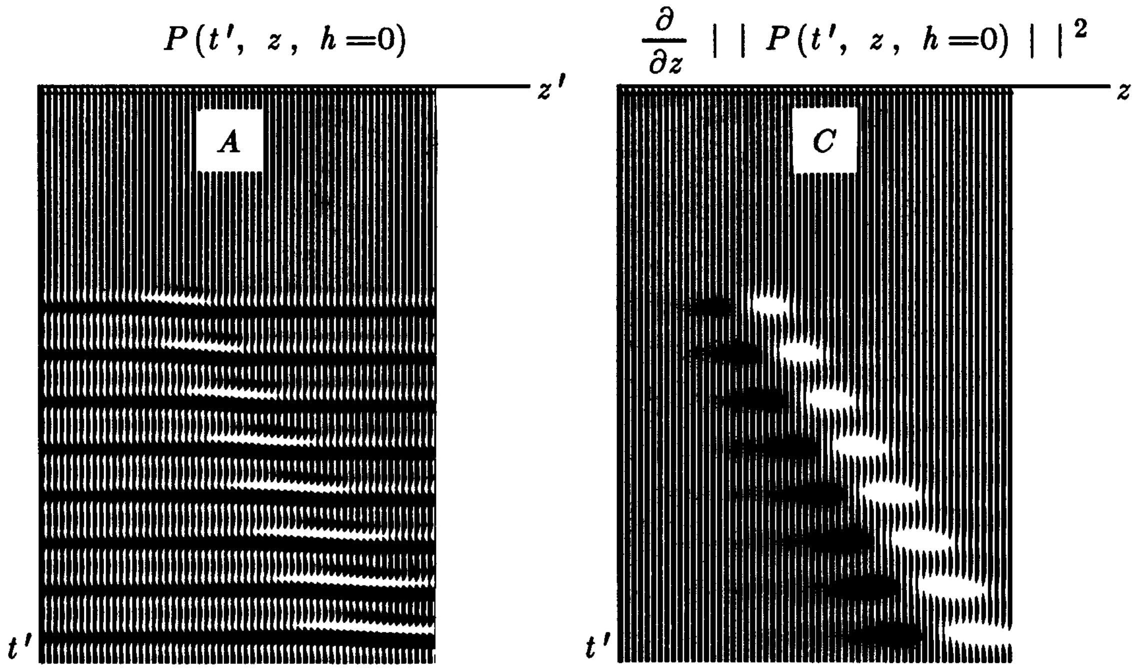

Figure 7.8: Two displays of the (z, t)-plane at zero offset. The earth model is eight uniformly spaced reflectors under a water layer (a family of hyperboloids in (h, t) at $z = 0$). The left display is the zero-offset trace. The amplitude maximum at the focus is not visually striking, but the phase shift is apparent. The right display is the z-derivative of the envelope of the zero-offset trace. A linear alignment along $z' = vt'$ is more apparent. (Doherty) vdmo-doherty [NR]

offset to negative offsets. The components of the data with velocity greater than $\bar{v}(z)$ will undermigrate. They will move toward zero offset but they will not go through. So the low-velocity part is at negative offset and the high-velocity part is at positive offset. If you wish, the process can then be reversed to bring the two parts back to the space of the original data.

Obviously, the process of multiplying data by a step function may create some undesirable diffractions, but then, you wouldn't expect to find an infinitely sharp velocity cutoff filter. Clearly, the false diffractions could be reduced by using a ramp instead of a step. An alternative to zeroing negative h would be to go into (k_h, ω)-space and zero the two quadrants of sign disagreement between k_h and ω.

This partitioning method unfortunately does not, by itself, provide a velocity spectrum. Energy away from $h = 0$ is unfocused and not obviously related to velocity. The need for a velocity spectrum motivates the development of other processes.

7.1.10 Reflected head waves on sections

It is common for an interpreter looking at a stacked section to identify a reflected head wave. Experimentally it is just a hyperbolic asymptote seen in (y, t)-space. Theoretically, it is a ray that moves away from a source along a horizontal interface until it encounters an irregularity, a fault perhaps, from which it reflects and returns toward the source. Reflected head waves are sometimes called reflected refractions. This event provides an easy velocity estimate, namely, $v = 2\,dy/dt$. From a processing point of view, such a velocity measurement is unexpected, because automatic processing extracts all velocity information in offset space, a space which many interpreters prefer to leave inside the computer. Of course, for a reflected head wave to be identified, a special geological circumstance must be present—a scatterer strong enough to have its hyperbolic asymptote visible. The point scatterer must also be strong enough to get through the typical suppression effects of shot and geophone patterns and CDP stacking. The most highly suppressed events, water velocity and ground roll, are just those whose velocities are most often apparent on stacked sections. (Recall Larner's streaks).

Velocity estimates made from reflections are averages of all the layers above the reflection point. To get depth resolution, it is necessary to subtract velocity estimates of different depth levels. Because of the subtraction, accuracy is lost. So with reflected waves, there is naturally a trade-off between accuracy and depth resolution. On the other hand, velocity estimates from head waves naturally have a high resolution in depth.

Processing seems to ignore or discriminate against the backscattered head wave, yet it is often seen and used. There must be an explanation. Perhaps there is also a latent opportunity. From a theoretical point of view, Clayton's cosines showed that at wide angles the velocity and dip sensitivity of midpoint and offset are exchanged. At late times another factor becomes significant: the aperture of a cable length can be much less than the width of a migration hyperbola. So, although it is easy to find an asymptote in midpoint space, there is little time shift at the end of the cable in offset space.

What processing could take advantage of lateral reflectivity and could enhance, instead of suppress, our ability to determine velocity in this way? Start by stacking at a high velocity. Then use the idea that at any depth z, the power spectrum of the data $U(\omega, k_y)$ should have a cutoff at the evanescent stepout $p(z) = k_y/\omega = 1/v(z)$. This would show up in a plot of the power spectrum U^*U, or better yet of the dip spectrum, as a function of depth. Perhaps it would be still better to visually inspect the seismic section itself after filtering in dip space about the expected velocity.

The wave-extrapolation equation is an all-pass filter, so why does the power spectrum change with depth? It changes because at any depth z it is necessary to exclude all the seismic data before $t = 0$. This data should be zeroed before computing the dip spectrum. The procedure is depicted in Figure 1.9.

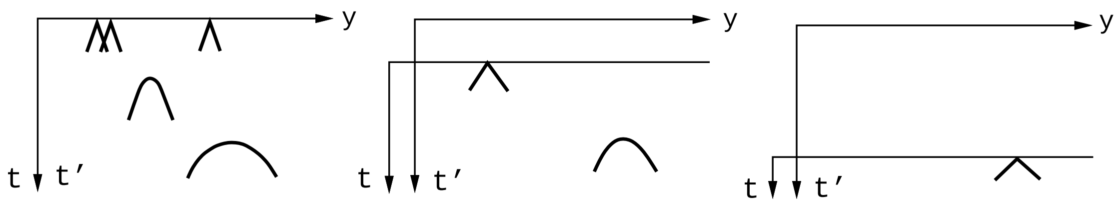

Figure 7.9: The dip-spectrum method of velocity determination. To find the velocity at any depth, seek the steepest dip on the section at that depth. On the left, at the earth's surface, you see the surface ground roll. In frames B and C the slowest events are the asymptotes of successively faster hyperbolas. vdmo-specz [NR]

To my knowledge this method has never been tried. I believe it is worth some serious testing. Even in the most layered of geological regions there are always faults and irregularities to illuminate the full available spectrum. Difficulty is unlikely to come from weak signals. More probably, the potential for difficulty lies in near-surface velocity irregularity.

EXERCISES:

1 Assume that the data $P(y, h, t)$ is constant with midpoint y. Given a common-midpoint gather $P(h, t, z = 0)$, define a Stolt-type integral transformation from $P(h, t, z = 0)$ to $P(h = 0, t, z)$ based on the double-square-root equation:

$$\frac{\partial}{\partial z} P \quad = \quad -i \frac{\omega}{v} \left[\sqrt{1 \; - \; (Y+H)^2} \; + \; \sqrt{1 \; - \; (Y-H)^2} \; \right] \; P$$

As with Stolt migration, your answer should be expressed as a 2-D inverse Fourier transform.

2 Start with a CDP gather $u(h,t)$ defined (by reciprocity) at both positive and negative values of h. Describe the effect of the following operations: Fourier transform to $U(k_h,\omega)$; multiply by $1 + \mathrm{sgn}(\omega)\,\mathrm{sgn}(k_h)$; transform back to (h,t)-space.

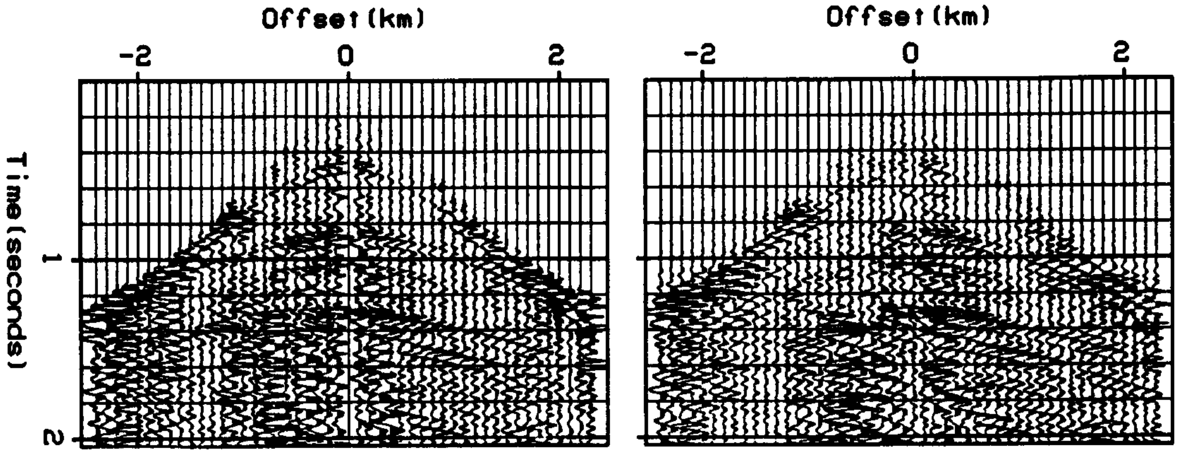

Figure 7.10: What is this? vdmo-exer [NR]

7.2 MIGRATION WITH VELOCITY ESTIMATION

We often face the three complications dip, offset, and unknown velocity at the same time. The double-square-root equation provides an attractive avenue when the velocity is known, but when it isn't, we are left with velocity-estimation procedures, such as that in the previous section, *which assume no dip.* In this section a means will be developed of estimating velocity in the presence of dip.

7.2.1 Dip moveout—Sherwood's devilish

Recall Levin's expression for the travel time of the reflection from a bed dipping at angle α from the horizontal:

$$t^2\, v^2 \quad = \quad 4\,(y-y_0)^2\,\sin^2\alpha \;+\; 4\,h^2\,\cos^2\alpha \tag{7.7}$$

In (h,t)-space this curve is a hyperbola. Scaling the velocity by $\cos\alpha$ makes the travel-time curve identical to the travel-time curve of the dip-free case. This is the conventional approach to stacking and velocity analysis. It is often satisfactory. Sometimes it is unsatisfactory because the dip angle is not a single-valued function of space. For example, near a fault plane there will be diffractions. They are a superposition of all dips, each usually being weaker than the reflections. Many dips are present in the same place. They blur the velocity estimate and the stack.

In principle, migration before stack—some kind of implementation of the full DSR equation—solves this general problem. But where do we get the velocity to use in the migration equations? Although migration is somewhat insensitive to velocity when only small angles are involved, migration becomes sensitive to velocity when wide angles are involved.

The migration process should be thought of as being interwoven with the velocity estimation process. J.W.C. Sherwood [1976] showed how the two processes, migration and velocity estimation, should be interwoven. The moveout correction should be considered in two parts, one depending on offset, the NMO, and the other depending on dip. This latter process was conceptually new. Sherwood described the process as a kind of filtering, but he did not provide implementation details. He called his process *Devilish,* an acronym for "dipping-event velocity inequalities licked." The process was later described more functionally by Yilmaz as *prestack partial migration,* but the process has finally come to be called simply *dip moveout* (DMO). We will first see Sherwood's results, then Rocca's conceptual model of the DMO process, and finally two conceptually distinct, quantitative specifications of the process.

Figure 1.11 contains a panel from a stacked section. The panel is shown several

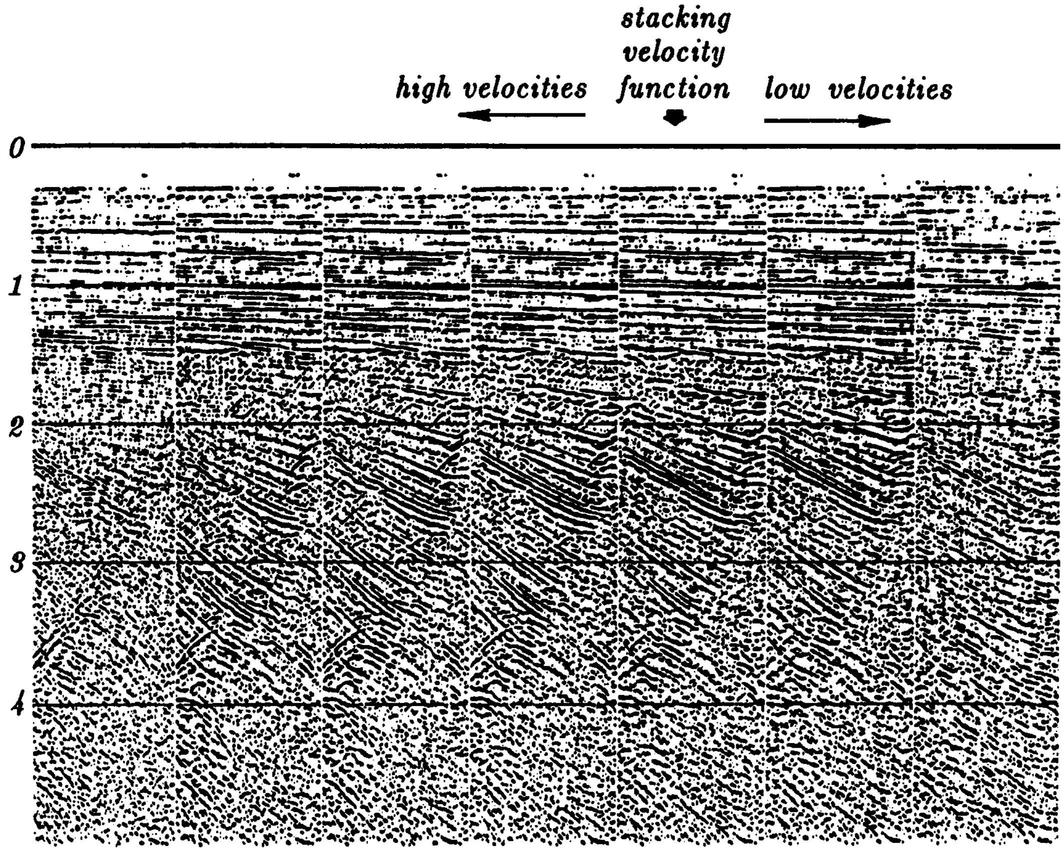

Figure 7.11: Conventional stacks with varying velocity. (distributed by Digicon, Inc.) vdmo-digicon [NR]

times; each time the stacking velocity is different. It should be noted that at the low velocities, the horizontal events dominate, whereas at the high velocities, the steeply dipping events dominate. After the *Devilish* correction was applied, the data was restacked as before. Figure 1.12 shows that the stacking velocity no longer depends on the dip. This means that after *Devilish,* the velocity may be determined without

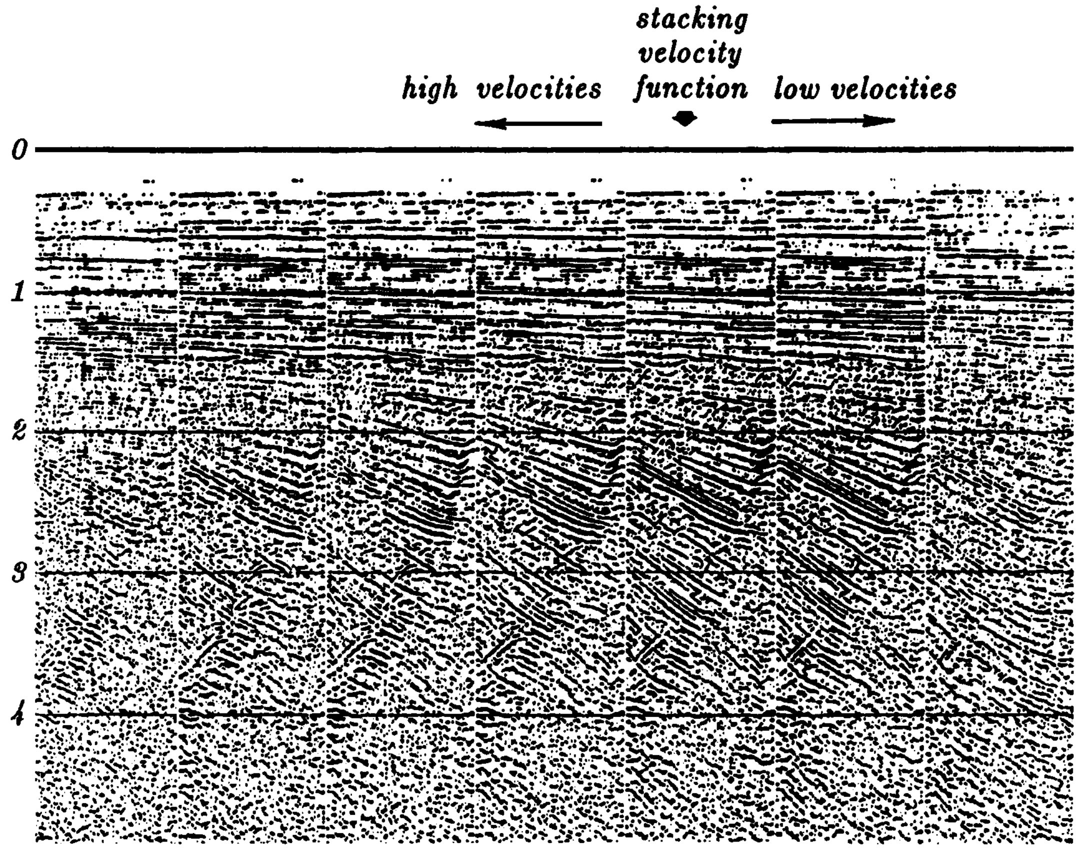

Figure 7.12: *Devilish* stacks with varying velocity. (distributed by Digicon, Inc.) vdmo-devlish [NR]

regard to dip. In other words, events with all dips contribute to the same consistent velocity rather than each dipping event predicting a different velocity. So the *Devilish* process should provide better velocities for data with conflicting dips. And we can expect a better final stack as well.

7.2.2 Rocca's smear operator

Fabio Rocca developed a clear conceptual model for Sherwood's dip corrections. Figure 1.13 illustrates Rocca's concept of a prestack partial-migration operator. Imagine a constant-offset section $P(t, y, h = h_0)$ containing an impulse function at some particular (t_0, y_0). The earth model implied by this data is a reflector shaped like an ellipse, with the shot point at one focus and the receiver at the other. Starting from

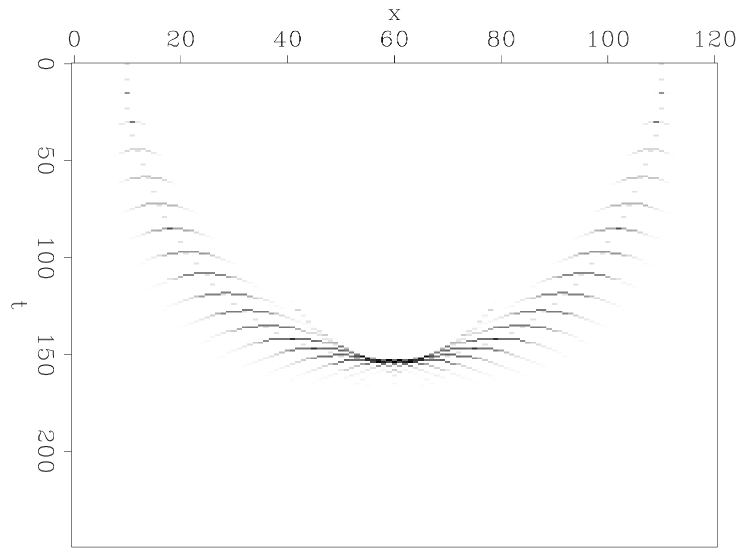

Figure 7.13: Rocca's prestack partial-migration operator is a superposition of hyperbolas, each with its top on an ellipse. Convolving (over midpoint) Rocca's operator onto a constant-offset section converts the constant-offset section into a zero-offset section. (Gonzalez) vdmo-rocca [ER]

this earth model, a zero-offset section is made by forward modeling—that is, each point on the ellipse is expanded into a hyperbola. Combining the two operations—constant-offset migration and zero-offset diffraction—gives the Rocca operator.

The Rocca operator is the curve of osculation in Figure 1.13, i.e., the smile-shaped curve where the hyperbolas reinforce one another. If the hyperbolas in Figure 1.13 had been placed everywhere on the ellipse instead of at isolated points, then the osculation curve would be the only thing visible (and you wouldn't be able to see where it came from).

The analytic expression for the travel time on the Rocca smile is the end of a narrow ellipse, shown in Figure 1.14. We will omit the derivation of the equation for this curve which turns out to be

$$1 \quad = \quad \frac{(y \ - \ y_0)^2}{h^2} \quad + \quad \frac{t^2}{t_0^2} \tag{7.8}$$

The Rocca operator appears to be velocity independent, but it is not completely so because the curve cuts off at $dt/dy \ = \ 2/v$.

The Rocca operator transforms a constant-offset section into a zero-offset section. This transformation achieves two objectives: first, it does normal-moveout correction; second, it does Sherwood's dip corrections. The operator of Figure 1.13 is convolved across the midpoint axis of the constant-offset section, giving as output a zero-offset section at just one time, say, t_0. For each t_0 a different Rocca operator must be designed. The outputs for all t_0 values must be superposed. Figure 1.15 shows a superposition of several Rocca smiles for several values of t_0.

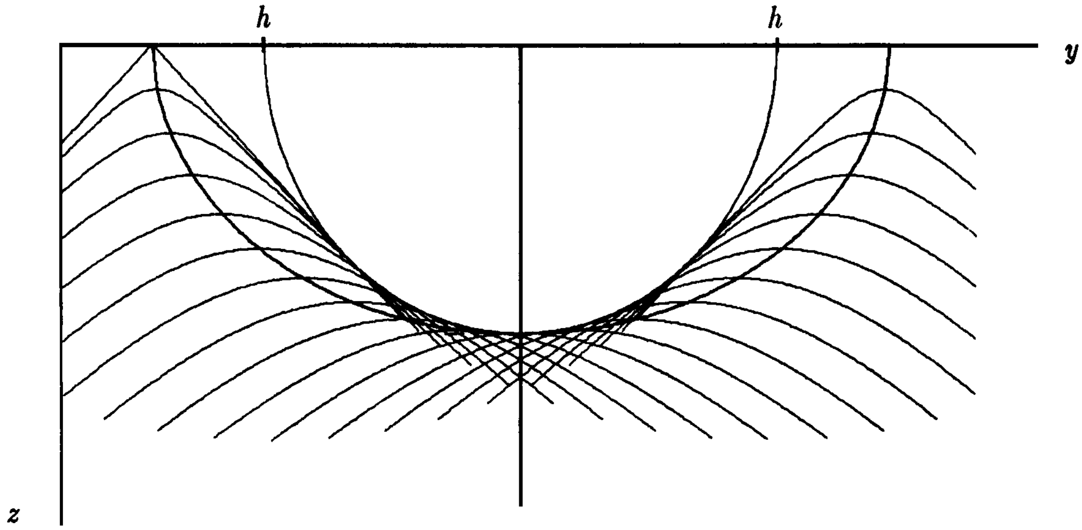

Figure 7.14: Rocca's smile. (Ronen) vdmo-rocca2 [NR]

This operator is particularly attractive from a practical point of view. Instead of using a big, wide ellipse and doing the big job of migrating each constant-offset section, only the narrow, little Rocca operator is needed. From Figure 1.15 we see that the energy in the dip moveout operator concentrates narrowly near the bottom. In the limiting case that h/vt_0 is small, the energy all goes to the bottom. When all the energy is concentrated near the bottom point, the Rocca operator is effectively a delta function. After compensating each offset to zero offset, velocity is determined by the normal-moveout residual; then data is stacked and migrated.

The *narrowness* of the Rocca ellipsoid is an advantage in two senses. Practically, it means that not many midpoints need to be brought into the computer main memory before velocity estimation and stacking are done. More fundamentally, since the operator is so compact, it does not do a lot to the data. This is important because the operation is done at an early stage, before the velocity is well known. So it may be satisfactory to choose the velocity for the Rocca operator as a constant, regional value, say, 2.5 km/sec.

An expression for the travel-time curve of the dip moveout operator might be helpful for Kirchhoff-style implementations. More to the point is a Fourier representation for the operator itself, which we will find next.

7.2.3 Hale's constant-offset dip moveout

Hale [1983] found a Fourier representation of dip moveout. Refer to the defining equations in Table 1.

To use the dip-dependent equations in Table 7.1 it is necessary to know the earth

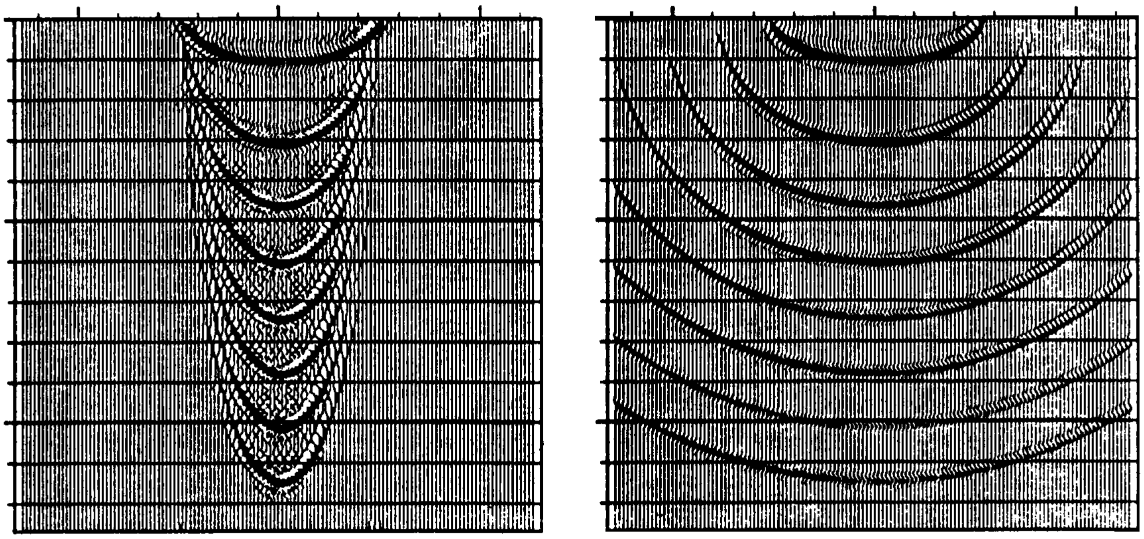

Figure 7.15: Point response of dip moveout (left) compared to constant-offset migration (right). (Hale) vdmo-dmopoint [NR]

NMO	$t \to t_n$	$t \;=\; \sqrt{t_n^2 + 4h^2v^{-2}}$
Levin's NMO	$t \to t_0$	$t \;=\; \sqrt{t_0^2 + 4h^2v^{-2}\cos^2\alpha}$
DMO	$t_n \to t_0$	$t_n \;=\; \sqrt{t_0^2 - 4h^2v^{-2}\sin^2\alpha}$

Table 7.1: Equations for normal moveout and dip moveout. Substituting the DMO equation into the NMO equation yields Levin's dip-corrected NMO.

dip α. The dip can be measured from a zero-offset section. On the zero-offset section in Fourier space, the sine of the dip is $vk_y/2\omega$. To stress that this measurement applies only on the *zero* -offset section, we shall always write ω_0.

$$\sin\,\alpha \quad = \quad \frac{v\,k_y}{2\,\omega_0} \tag{7.9}$$

In the absence of dip, NMO should convert any trace into a replica of the zero-offset trace. Likewise, in the presence of dip, the combination of NMO and DMO should convert any constant-offset section to a zero-offset section. Pseudo-zero-offset sections manufactured in this way from constant-offset sections will be denoted by $p_0(t_0, h, y)$. First take the midpoint coordinate y over to its Fourier dual k_y. Then take the Fourier transform over time letting ω_0 be Fourier dual to t_0.

$$P_0(\omega_0, h, k_y) \quad = \quad \int dt_0\; e^{i\omega_0 t_0}\; P_0(t_0, h, k_y) \tag{7.10}$$

Change the variable of integration from t_0 to t_n.

$$P_0(\omega_0, h, k_y) \quad = \quad \int\; dt_n\, \frac{dt_0}{dt_n}\quad e^{i\omega_0 t_0(t_n)}\; P_0(t_0(t_n), h, k_y) \tag{7.11}$$

Express the integrand in terms of NMOed data P_n. This is done by means of $P_n(t_n, h, k_y) \;= P_0(t_0(t_n), h, k_y)$.

$$P_0(\omega_0, h, k_y) \quad = \quad \int\; dt_n\, \frac{dt_0}{dt_n}\quad e^{i\omega_0 t_0(t_n)}\; P_n(t_n, h, k_y) \tag{7.12}$$

As with Stolt migration, the Jacobian of the transformation, dt_0/dt_n scales things but doesn't do time shifts. The DMO is really done by the exponential term.

Omitting the Jacobian (which does little), the over-all process may be envisioned with the program outline:

$P(k_y)$ =FT[$P(y)$]
$P_n(t_n)$ = NMO[$P(t)$]
for all k_y { # three nested loops, interchangeable
for all h { # three nested loops, interchangeable
for all ω_0 { # three nested loops, interchangeable
    sum = 0
    for all t_n {
        sum = sum + $\exp\left(i\omega_0\,\sqrt{t_n^2 + \frac{h^2k_y^2}{\omega_0^2}}\,\right) P_n(t_n, h, k_y)$
        }
    $P_0(\omega_0, h, k_y)$ = sum
    } } }
$p_0(t_0, h, y)$ =FT2D[$P_0(\omega_0, h, k_y)$]

Notice that the exponential in the inner loop in the program does not depend on velocity. The velocity in the DMO equation in Table 7.1 disappears on substitution of $\sin \alpha$ from equation (1.9). So dip moveout does not depend on velocity. An example of processing with dip moveout is in Figure 1.16.

The procedure outlined above requires NMO before DMO. To reverse the order would be an approximation. This is unfortunate because we would prefer to do the costly, velocity-independent DMO step once, before the iterative, velocity-estimating NMO step.

7.2.4 Ottolini's radial trace dip moveout

Ordinarily we regard a common-midpoint gather as a collection of seismic traces, that is, a collection of time functions, each one for some particular offset h. But this (h, t) data space could be represented in a different coordinate system. A system with some nice attributes is the radial-trace system introduced by Turhan Taner. In this system the traces are not taken at constant h, but at constant angle. The idea is illustrated in Figure 1.17.

Besides having some theoretical advantages, which will become apparent, this system also has some practical advantages, notably: (1) the traces neatly fill the space where data is nonzero; (2) the traces are close together at early times where wavelengths are short, and wider apart where wavelengths are long; and (3) the energy on a given trace tends to represent wave propagation at a fixed angle. The last characteristic is especially important with multiple reflections. But for our purposes the best attribute of radial traces is still another one.

Richard Ottolini noticed that a point scatterer in the earth appears on a radial-trace section as an *exact* hyperbola, not a flat-topped hyperboloid. The travel-time curve for a point scatterer, Cheops' pyramid, can be written as a "string length" equation, or a stretched-circle equation. Making the definition

$$\sin \psi \quad = \quad \frac{2\,h}{v\,t} \tag{7.13}$$

and substituting into equation (6.22) yields

$$v\,t \quad = \quad 2\sqrt{\frac{z^2}{\cos^2 \psi} \;+\; (y \;-\; y_0)^2} \tag{7.14}$$

Scaling the z-axis by $\cos \psi$ gives the circle and hyperbola case all over again! Figure 1.18 shows a three-dimensional sketch of the hidden hyperbola.

We will see that the radial hyperbola of Figure 1.18 is easier to handle than the flat-topped hyperboloid that is seen at constant h. Refer to the equations in Table 7.2.

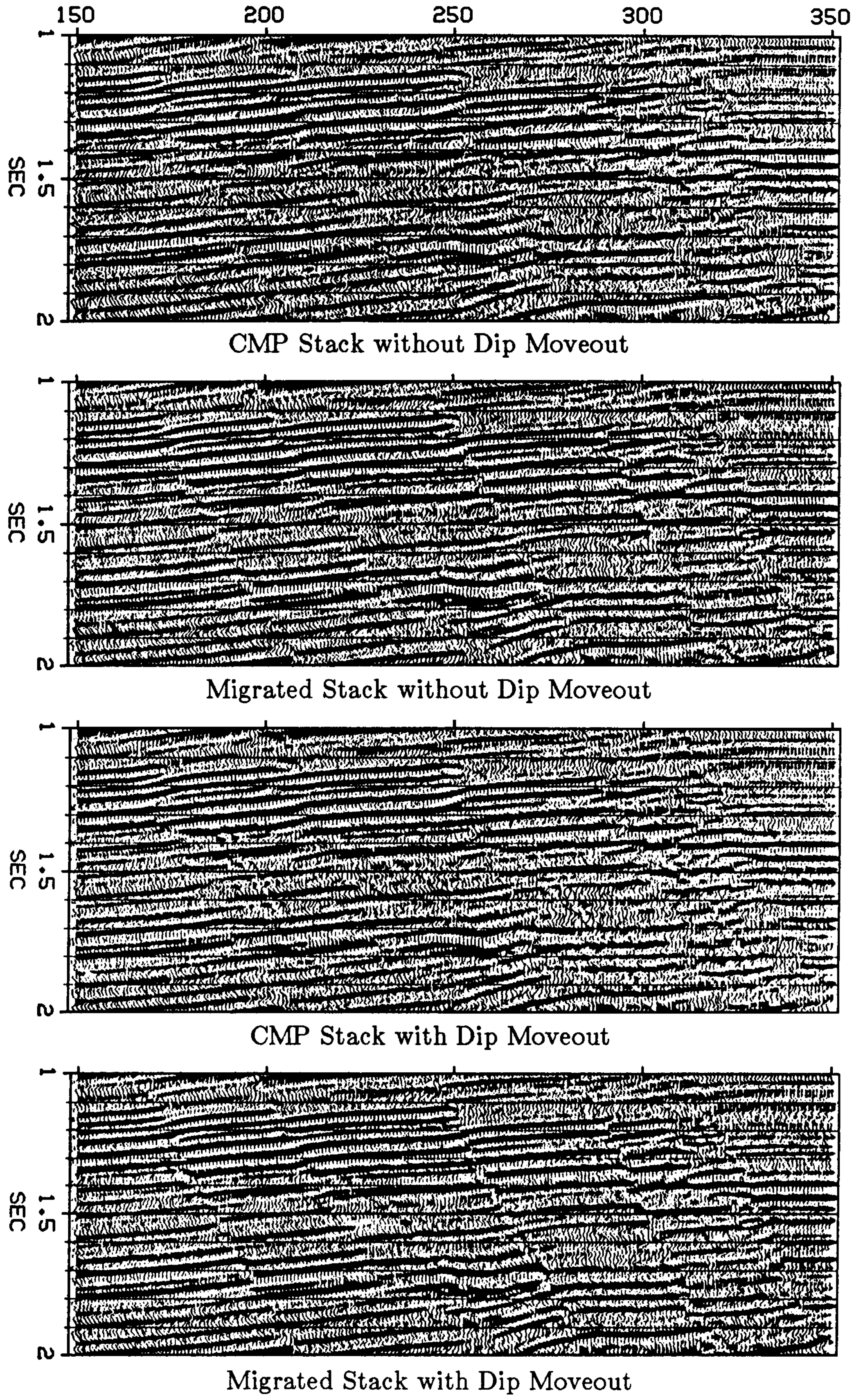

Figure 7.16: Processing with dip moveout. (Hale) $\boxed{\text{vdmo-dmoproc}}$ [NR]

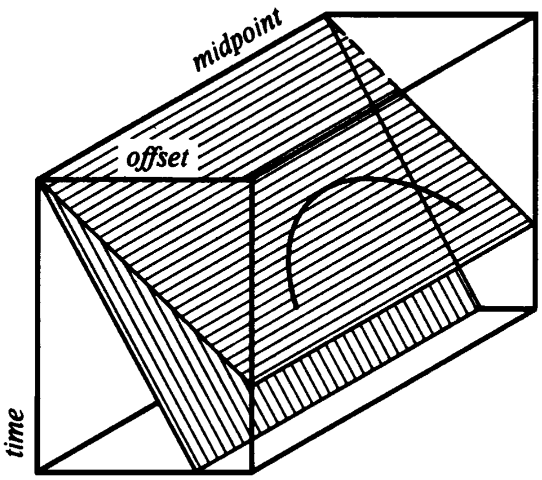

Figure 7.17: Inside the data volume of a reflection seismic traverse are planes called *radial-trace sections.* A point scatterer inside the earth puts a hyperbola on a radial-trace section. vdmo-otto [NR]

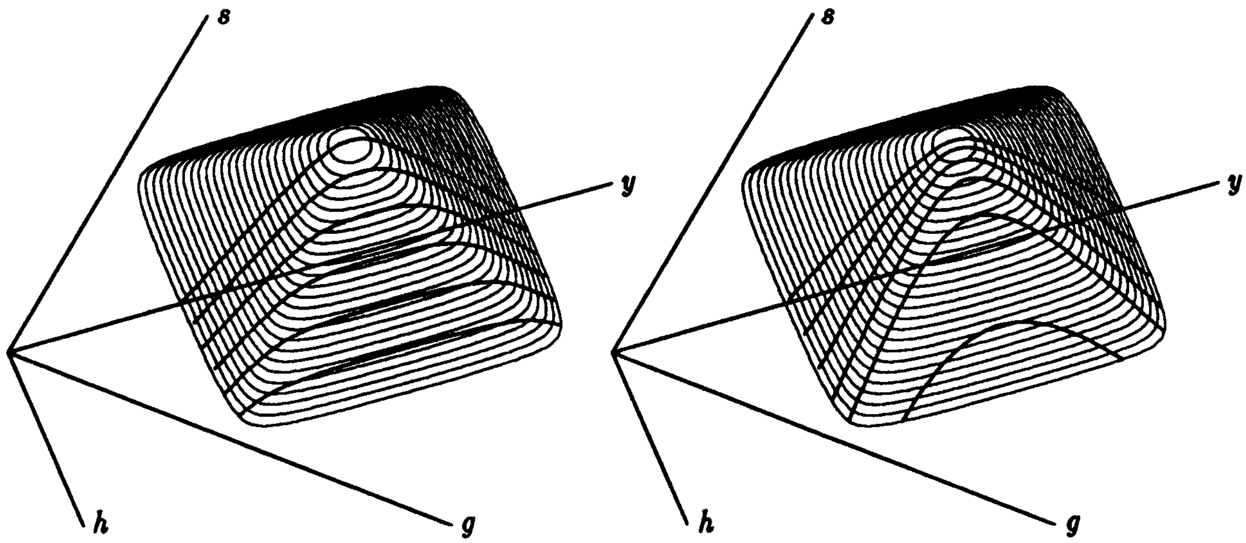

Figure 7.18: An unexpected hyperbola in Cheops' pyramid is the diffraction hyperbola on a radial-trace section. (Harlan) vdmo-ottohyp [NR]

migration	$\omega \to k_z$	$k_y^2 + k_z^2 \cos^2 \psi = 4\omega^2/v^2$
zero-offset diff.	$k_z \to \omega_0$	$k_y^2 + k_z^2 = 4\omega_0^2/v^2$
DMO+NMO	$\omega \to \omega_0$	$.25v^2 k_y^2 \sin^2 \psi + \omega_0^2 \cos^2 \psi = \omega^2$
radial DMO	$\omega \to \omega_s$	$.25v^2 k_y^2 \sin^2 \psi + \omega_s^2 = \omega^2$
radial NMO	$\omega_s \to \omega_0$	$\omega_0 \cos \psi = \omega_s$

Table 7.2: Equations defining dip moveout and ordinary moveout in radial trace coordinates.

The second equation in Table 7.2 is the usual exploding reflector equation for zero-offset migration. It may also be obtained from the DSR by setting $H = 0$. As written it contains the earth velocity, not the half velocity. Equation (1.14) says that hyperbolas of differing ψ values are related to one another by scaling the z-axis by cos ψ. According to Fourier transform theory, scaling z by a cos ψ divisor will scale k_z by a cos ψ multiplier. This means the first equation in Table 7.2 can be used for migrating and diffracting hyperbolas on radial-trace sections. Eliminating k_z from the first and second equations yields the middle equation $\omega \to \omega_0$ in Table 7.2. This middle equation combines the operation of migrating all offsets (really any radial angle) and then diffracting out to zero offset. Thus the total effect is that of *offset continuation* , i.e. NMO and DMO. The last two equations in Table 7.2 are a decomposition of the middle equation $\omega \to \omega_0$ into two sequential processes, $\omega \to \omega_s$ and $\omega_s \to \omega_0$. These two processes are like DMO and NMO, but the operations occur in radial space. Radial NMO is a simple time-invariant s tretch; hence the notation ω_s.

Unlike the constant-offset case, dip moveout is now done *before* the stretching, velocity-estimating step. Let us confirm that the dip moveout is truly velocity-independent. Substitute (1.13) into the radial DMO transformation in Table 7.2 to get the equation for transformation from time to stretched time:

$$\frac{h^2}{t^2}\, k_y^2 \quad + \quad \omega_s^2 \quad = \quad \omega^2 \tag{7.15}$$

We observe that the velocity v has dropped out of (1.15). Thus dip moveout in radial coordinates doesn't depend on velocity. Dip-moveout processing $\omega \to \omega_s$ does not require velocity knowledge. Radial coordinates offer the advantage that this comparatively costly process is done *before* the velocity is estimated $\omega_s \to \omega_0$.

The dip-moveout process, $\omega \to \omega_s$, can be conveniently implemented with a Stolt-type algorithm using (1.15).

The foregoing analysis has assumed a constant velocity. A useful practical approximation might be to revert to a $v(z)$ analysis after the dip moveout, just before conventional velocity analysis, stack, and zero-offset migration.

Both the radial-trace method and Hale's constant-offset method handle all angles exactly in a constant-velocity medium, but neither method treats velocity stratification exactly nor is it clear that this can be done—since neither method is rooted in the DSR. Yilmaz [1979] rooted his DMO work in the DSR, so his method can be expected to be exact for velocity stratification, but Yilmaz could not avoid angle-dependent approximations. So there remains theoretical work to be done.

7.2.5 Anti-alias characteristic of dip moveout

You might think that if (y, h, t)-space is sampled along the y-axis at a sample interval Δy, then any final migrated section $P(y, z)$ would have a spatial resolution no better than Δy. This is not the case.

The basic principle at work here has been known since the time of Shannon. If a time function and its derivative are sampled at a time interval $2\Delta t$, they can both be fully reconstructed provided that the original bandwidth of the signal is lower than $1/(2\Delta t)$. More generally, if a signal is filtered with m independent filters, and these m signals are sampled at an interval $m\Delta t$, then the signal can be recovered.

Here is how this concept applies to seismic data. The basic signal is the earth model. The various filtered versions of it are the constant-offset sections. Recall that the CDP reflection point moves up dip as the offset is increased. Further details can be found in a paper by Bolondi, Loinger, and Rocca [1982], who first pointed out the anti-alias properties of dip moveout. At a time of increasing interest in 3-D seismic data, special attention should be paid to the anti-alias character of dip moveout.

EXERCISES:

1 Describe the effect of the Jacobian in Hale's dip moveout process.

7.3 LATERAL VELOCITY IN BIGGER DOSES

To the interpreting *geologist,* lateral velocity variation produces a strange distortion in the seismic section. And the distortion is worse than it looks. The *geophysicist* is faced with the challenge of trying to deal with lateral velocity variation in a quantitative manner. First, how can reliable estimates of the amount of lateral velocity variation be arrived at? Then, do we dare use these estimates for reprocessing data?

Our studies of *dip* and *offset* have resulted in straightforward procedures to handle them, even when they are simultaneously present. Unfortunately, increasing lateral velocity variation leads to increasing confusion—confusion we must try to overcome. Strong lateral velocity variation overlies the largest oil field in North America, Prudhoe Bay. Luckily, however, we have many idealized examples that are easy to understand. Any "ultimate" theory would have to explain these examples as limiting cases.

Let us review. The double-square-root equation presumably works if the square roots are expanded and if we accept the usual limitation of accuracy with angle. Our problem with the DSR is that it merely tells us how to migrate and stack *once the velocity is known.* Kjartansson's method of determining the distribution of (some function of) $v(x, z)$ assumes straight rays, no dip, and a single, planar reflector. On the other hand, stacking along with prestack partial migration allows any scattering geometry but enables determination of $v(z)$ only under the presumption that there is no lateral variation of velocity. Clearly, there are many gaps. We begin with comprehensible, special cases but ultimately sink into a sea of confusion.

7.3.1 Replacement velocity: freezing the water

Sometimes you are lucky and you know the velocity. Maybe you know it because you are dealing with synthetic data. Maybe you know it because you have already drilled 300 shallow holes. Or maybe you can make a good estimate because you have a profile of water depth and you are willing to guess at the sediment velocity. Often the velocity problem is really a near-surface problem. Perhaps you have been dragging your seismic streamer over the occasional limestone reefs in the Red Sea.

Assuming that you know the velocity and that the lateral variations are near the surface, then you should think about the idea of a *replacement velocity.* For example, suppose you could freeze the water in the Red Sea, just until it is hard enough that the ice velocity and the velocity of the limestone reefs are equal. That would remove the unnecessary complexity of the reflections from deep targets. Of course you can't freeze the Red Sea, but you can reprocess the data to try to mimic what would be recorded if you could.

First, downward continue the data to some datum beneath the lateral variations. Then upward continue it back to the surface through the homogeneous replacement medium.

While in principle the DSR could be used for this job, in practice it would be expensive and impractical. The best approach is to study the two operations—going down, then going up—in combination. Since the two operations are largely in opposition to each other, whatever is done to the data should be just a function of the difference. For example, the equation

$$\frac{\partial P}{\partial z} \quad = \quad i\omega \left[\frac{1}{v(s)} \quad + \quad \frac{1}{v(g)} \quad - \quad \frac{2}{v_{\mathrm{avg}}} \right] P \tag{7.16}$$

combines the downward continuation with the upward continuation and makes little change to the wavefield P when the velocities are nearly the same. Equation (1.16) is basically a time-shifting equation. There is an industry process known as *static* corrections. The word *static* implies time-invariant—the amount of time shift does not depend on time. When the appropriate corrections are merely static shifts, then the earth model has lateral velocity variations in the near surface only. This is often the case. Equation (1.16) also has the ability to do time-variable time shifts because $v(s)$ and $v(g)$ can be any function of depth z. Because of the wide-offset angle normally used, it is desirable to extend (1.16) to a wider angle. Such extensions are found in Lynn [1979]. Lynn also shows how partial differential equations can be written to describe the influence of lateral velocity variation on stacking velocity. Berryhill [1979] illustrated the use of the Kirchhoff method for an irregular datum.

In practice, the problem of estimating lateral velocity variations is usually more troublesome than the application of these velocities during migration. Static time shifts are estimated from a variety of measurements including the elevation survey, travel times from the bottoms of shot holes to the surface, and crosscorrelation of reflection seismograms. Wiggins et al. [1976] provide an analysis to determine the static shifts from correlation measurements.

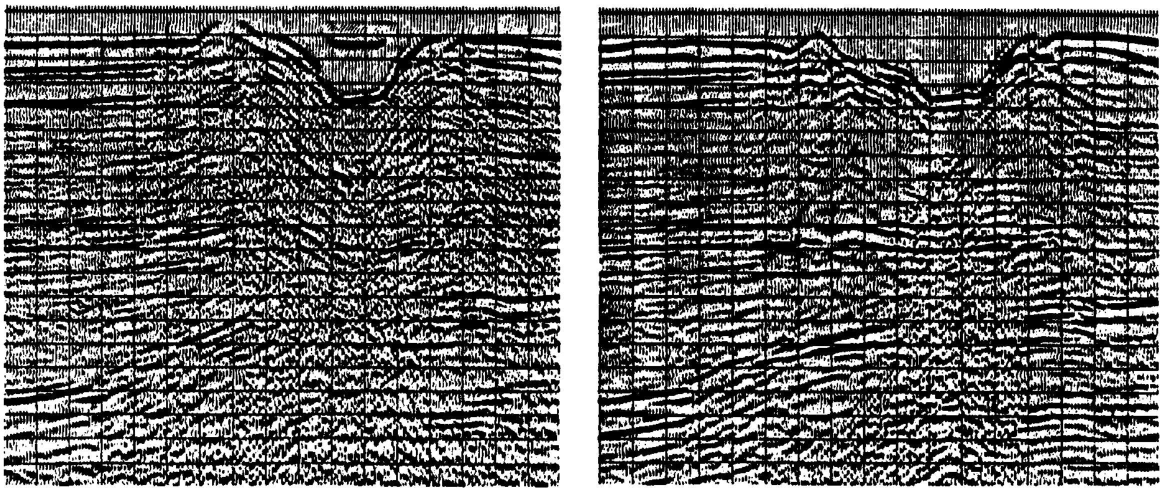

Figure 7.19:)
] Data (left) from Philippines with dynamic corrections (right). (by permission from *Geophysics,* Dent [1983]) vdmo-dent [NR]

Where the lateral variation runs deeper the time shifts become time-dependent. This is called the *dynamic* time-shift problem. To compute dynamic time shifts, dip is assumed to be zero. Rays are traced through a presumed model with laterally variable velocity. Rays are also traced through a reference model with laterally constant velocity. The difference of travel times of the two models defines the dynamic time shifts. See Figure 1.19. Where the lateral variation runs deeper still, the problem looks more like a migration problem. Figure 1.20 illustrates a process called *REVEAL* by Digicon, Inc., who have not *revealed* whether a time-shift method or a wave-equation method was used.

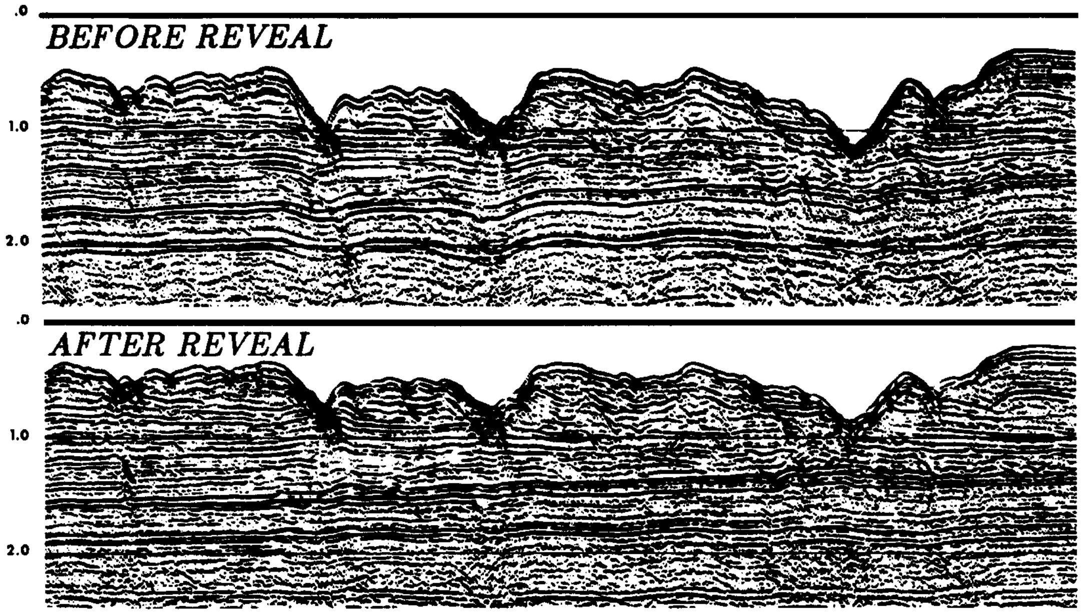

Figure 7.20: Example of processing with a replacement velocity. Observe that deeper bedding is now flatter and more continuous. (distributed by Digicon, Inc.) vdmo-reveal [NR]

7.3.2 Lateral shift of the hyperbola top

Figure 1.21 shows a point scatterer below a dipping interface. As usual there is a higher velocity below. This is a simple prototype for many lateral-velocity-variation problems. Surface arrival times will be roughly hyperbolic with distortion because of the velocity jump at the interface. The minimum travel time (hyperboloid top) has been displaced from its usual location directly above the point scatterer. Observe that

- At minimum time, the ray emerges going straight up.
- Minimum time is on the high-velocity side of the point scatterer.
- Minimum time is displaced further from the scatterer as offset increases.

Figure 7.21: Rays emerging from a point scatterer beneath a velocity wedge (left). Travel-time curve (right). The slope at a is the negative of that at b. The midpoint between a and b is at the top of the $h > 0$ curve. vdmo-lady [ER]

The travel-time curve is roughly hyperbolic, but the asymptote on the left side gives the velocity of the medium on the left side, and the asymptote on the right gives the velocity on the right.

Let $T(x)$ denote the travel time from the point scatterer to the surface point x. The travel time for a constant-offset section is then $t(y) = T(y+h) + T(y-h)$. To find the earliest arrival, set $dt/dy = 0$. This proves that the slope at a on Figure 1.21 is the negative of the slope at b. This shows why the displacement of the top of the hyperboloid from the scatterer increases with offset.

Lateral velocity variation causes hyperbolas to lose their symmetry. Computationally, it is the lens term that tilts hyperbolas, causing their tops to move laterally.

7.3.3 Phantom diffractor

A second example of lateral velocity variation is Figure 1.22, also taken from Kjartansson's dissertation. The physical model shown on the inset in Figure 1.22 is three constant velocity wedges separated by broken line segments representing reflectors. The bottom edge of the model also represents a reflector. The wavefield in Figure 1.22 was made using the exploding-reflector calculation, which Kjartansson regarded as a reasonable approximation to a zero-offset section. Notice that under the tip of the 4 km/sec wedge is a small diffraction on the bottom horizontal reflector. Because such a diffraction has nothing to do with the flat reflector on which it is seen, it is termed

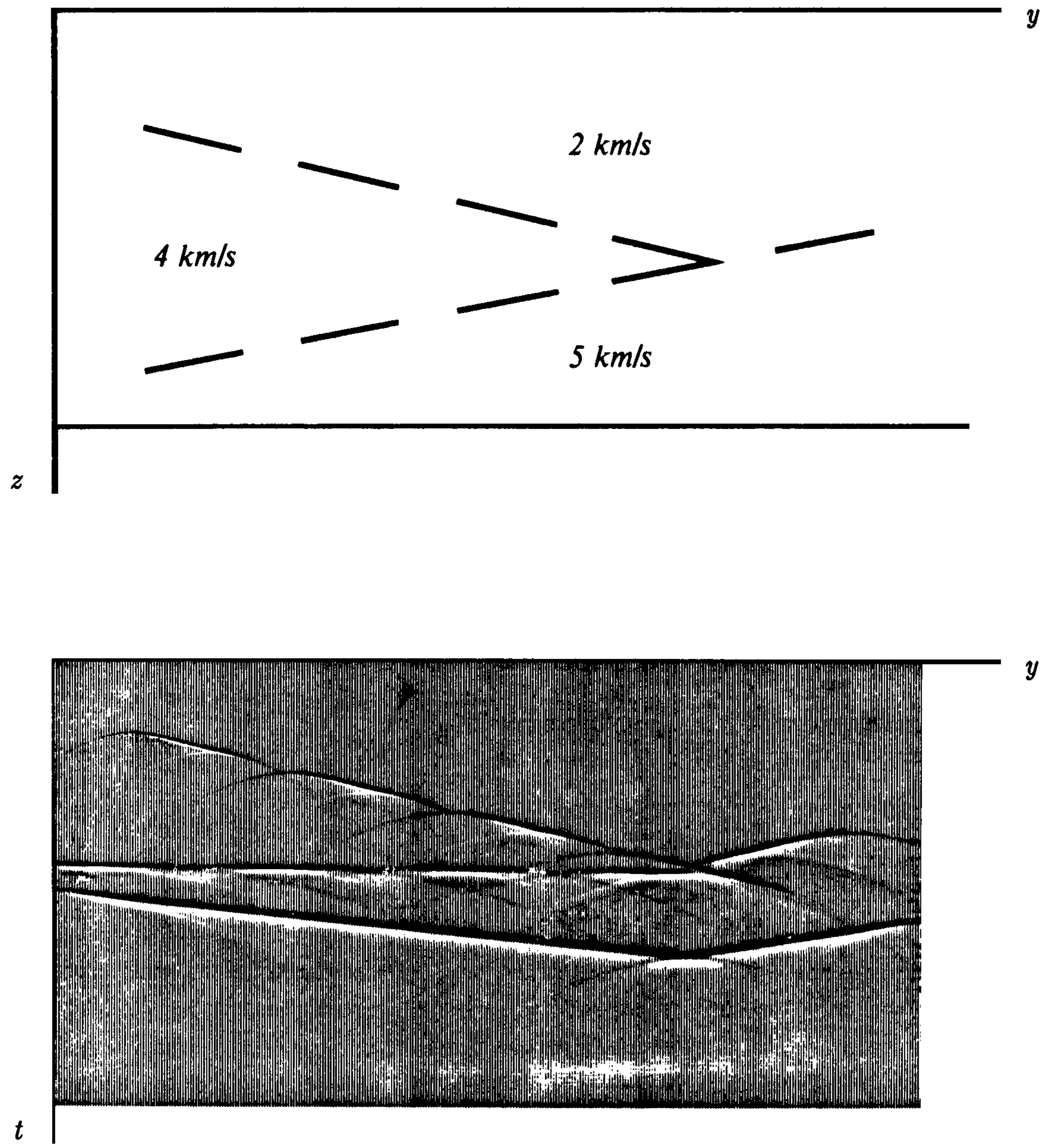

Figure 7.22: The model in the upper panel was taken from Western Geophysical's *Depth Migration* brochure. The model is not physical because of the segmenting of the interface; however, the segments make it a good case for the study of lateral shifts. The synthetic data is in the lower panel (from Kjartansson). The phantom diffraction is on the latest arrival just below the tip of the wedge. vdmo-westdepthmig [NR]

a "phantom" diffraction. Phantom diffractions are not easy to recognize, but they do occur. In reality, the "bright spots" in Kjartansson's section were probably phantom diffractions. It has been reported that phantom diffractions provide a means of prospecting for small, high-velocity, carbonate reefs.

7.3.4 Wavefront healing

Figure 1.23 (also in FGDP) shows another example of ray bending. The first frame

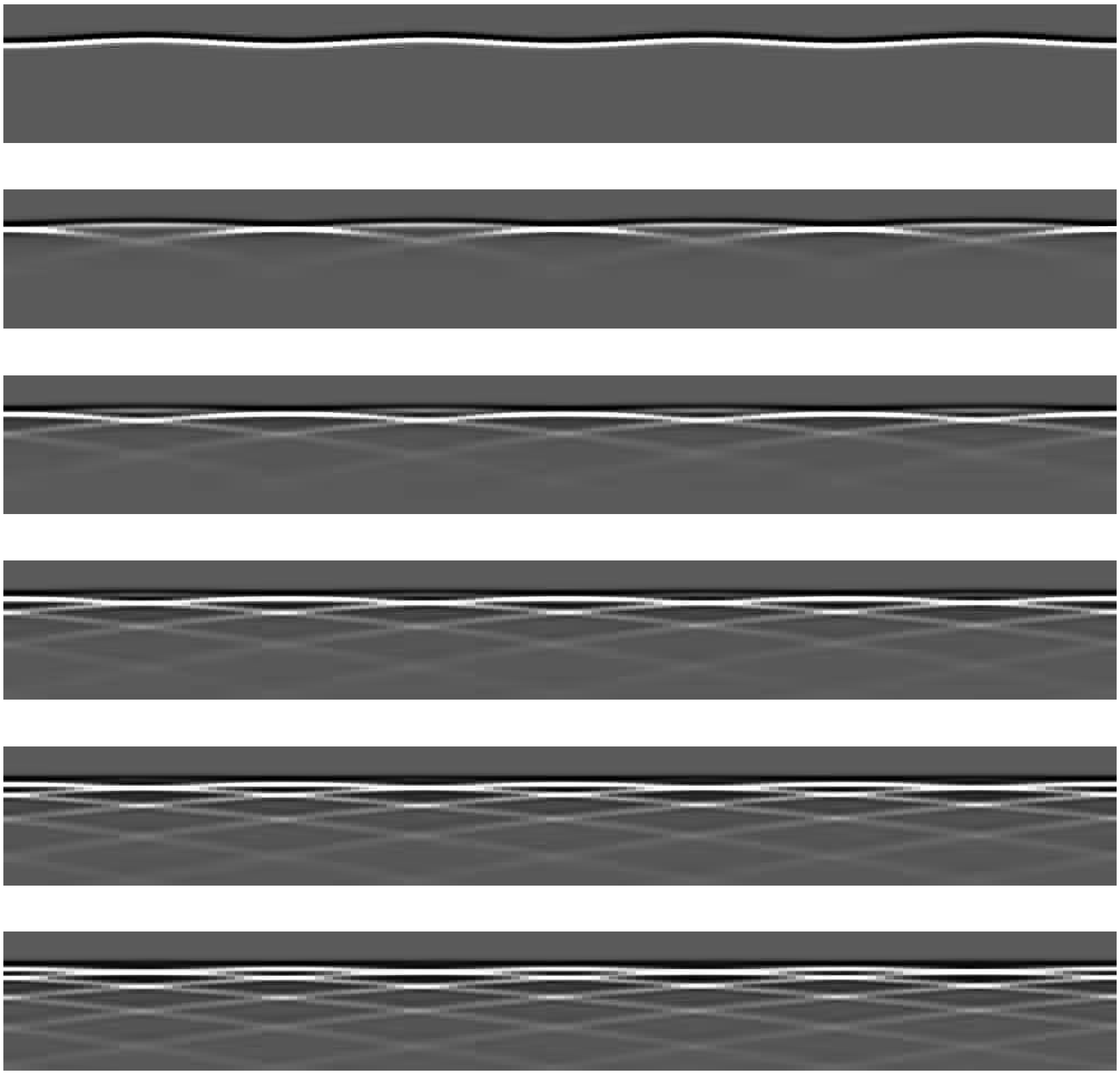

Figure 7.23: The first frame on the left shows a plane wave just after it has been distorted into a wavy shape by the thin-lens term. After this the thin-lens term vanishes. Subsequent frames show the effect of increasing amounts of diffraction. Notice the lengthening of the wave packet and the *healing* of the first arrival. (FGDP, p. 213, figure 10-22) vdmo-heal [ER]

on the left shows a plane wave just after it has been distorted into a wavy shape by the thin-lens term. After this the thin-lens term vanishes. Later frames show the effect of increasing amounts of diffraction.

7.3.5 Fault-plane reflection

Across a single vertical fault in the earth the velocity will be a simple step function of the horizontal coordinate. Rays traveling across such a fault suffer in amplitude because of reflection and transmission coefficients, depending on the angle. Since near-vertical rays are common, only small velocity contrasts are required to generate strong internal reflections. By this reasoning, steep faults should be more distorted, and hence more recognizable, on small-offset sections than on wide-offset sections or stacks.

This phenomenon is somewhat more confusing when seen in (x, t)-space. Figure 1.24 was computed by Kjartansson and used in a quiz. Study this figure and answer the questions in the caption. Here is a hint: A reflected ray beyond critical angle undergoes a phase shift. This will turn a pulse into a doublet that might easily be mistaken for two rays.

Figure 1.24 exhibits a geometry in which the exploding-reflector model fails to produce all the rays seen on a zero-offset section. The exploding-reflector model produces two types of rays: the ray that goes directly to the surface, and the ray that reflects from the fault plane before going to the surface. A zero-offset section has three rays: the two rays just mentioned, but moving at double travel time, once up, once down; and in addition the ray not present in Figure 1.24, which hits the fault plane going one way but not the other way.

There is a simple way to make constant-offset sections in laterally variable media when the reflector is just a point. The exploding-reflector seismogram recorded at $x = s$ is simply time convolved with the one recorded at $x = g$. Convolution causes the travel times to add. Even the non-exploding-reflector rays are generated. Too bad this technique doesn't work for reflector models that are more complicated than a simple reflecting point.

7.3.6 Misuse of v(x) for depth migration

The program that generated Figure 1.24 could be run in reverse to do a migration. All the energy from all the interesting rays would march back to the impulsive source. Would this be an effective migration program in a field environment? It is unlikely that it would. The process is far too sensitive to quantitative knowledge of the lateral velocity jump. It is the quantitative value that determines the reflection coefficient and ultimately the correct recombination of all the wavefronts back to a pulse. To see how an incorrect value can result in further error, imagine using

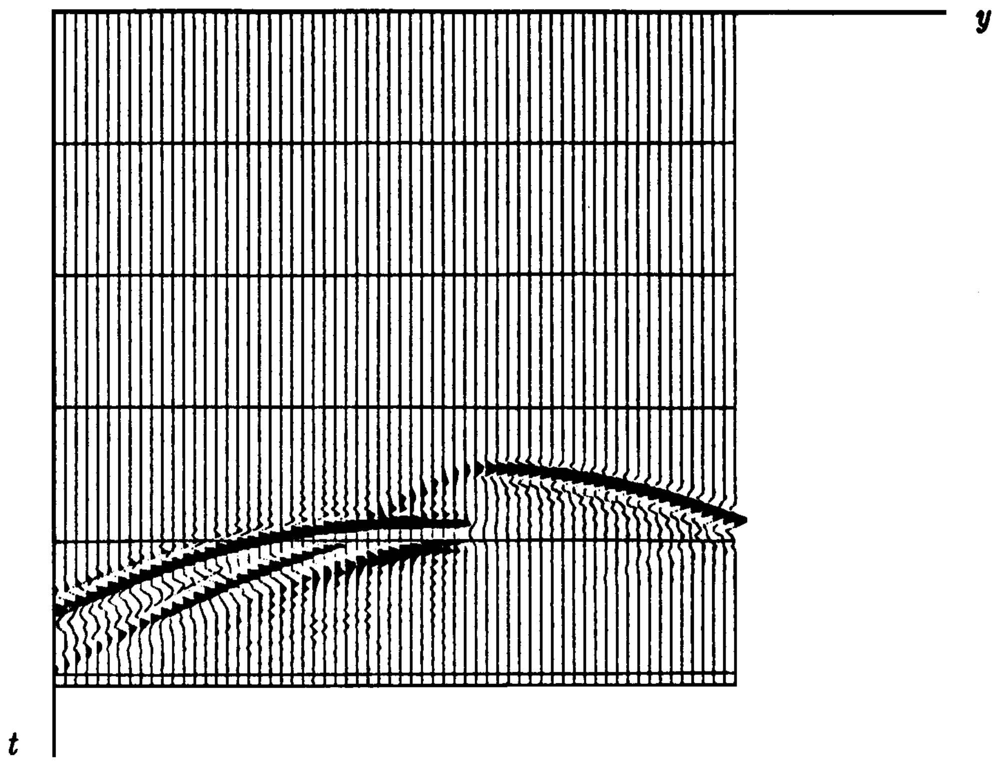

Figure 7.24: Synthetic data from an exploding-reflector calculation for an earth model containing a point scatterer and a velocity jump $v_1/v_2 = 1.2$ across a vertical contact. (Kjartansson) (a) Is the point scatterer in the slow or the fast medium? (b) Identify four arrivals and diagram their raypaths. (c) Identify and explain two kinds of computational artifacts. (d) Find an evanescent wave. (e) Find phase shift on a beyond-critical-angle reflection. (f) A zero-offset section has ray not shown above. Where? vdmo-veljump [NR]

the hyperbola-summation migration method. Applied to this geometry this method implies weighted summation over all the raypaths in the figure. The incorrect value would put erroneous amplitudes on various branches. An erroneous location for the fault would likewise mislocate several branches.

The lesson to be learned from this example is clear. Unnecessary bumps in the velocity function can create imaginary fault-plane reflections. Consistent with known information, a presumed migration velocity should be as smooth as possible in the lateral direction. Unskilled and uninformed staff at a processing center remote from the decision making should not have the freedom to introduce rapid lateral changes in the velocity model.

7.3.7 First-order effects, the lens term

Now let us be specific about what is meant by the lens term in the present context of before-stack migration in the presence of lateral velocity variation. Specializing the DSR equation to 15° angles gives

$$\frac{\partial U}{\partial z} \quad = \quad - \left\{ \left[\frac{i\omega}{v(s)} + \frac{v(s)}{2i\omega} \frac{\partial^2}{\partial s^2} \right] + \left[\frac{i\omega}{v(g)} + \frac{v(g)}{2i\omega} \frac{\partial^2}{\partial g^2} \right] \right\} U \tag{7.17}$$

Rearranging the terms to group by behavior gives

$$\frac{\partial U}{\partial z} \quad = \quad - \left[\frac{i\omega}{v(s)} + \frac{i\omega}{v(g)} \right] U - \left[\frac{v(s)}{2\,i\,\omega} \frac{\partial^2}{\partial s^2} + \frac{v(g)}{2\,i\,\omega} \frac{\partial^2}{\partial g^2} \right] U \tag{7.18}$$

$$\frac{\partial U}{\partial z} \quad = \quad \text{lens term} \quad + \quad \text{diffraction term} \tag{7.19}$$

So you see the familiar type of lens term, but it has two parts, one for shifting at the shot, and one for shifting at the geophone.

7.3.8 The migrated time section: an industry kludge

As geology becomes increasingly dramatic, reflection data gets more anomalous. The first thing noticeable is that the stacking velocity becomes unreasonable. In practice the available computer processes—based on inappropriate assumptions—will be tried anyway.

A stacking velocity will be chosen and a stack formed. How should the migration be done? Most basic migration programs omit the lens term. Although it is easy to include the lens term, the term is sensitive to lateral variation in velocity. Since estimates of lateral variation in velocity always have questionable reliability, use of a migration program with a lens term is usually limited to knowledgeable interpreters. The lens term is usually omitted from the basic migration utility program. Let us see what this means.

The migration equation is valid in some "local plane wave" sense, i.e.

$$k_z(y,z) \quad = \quad \frac{\omega}{v(y,z)} \ \sqrt{1 \ - \ \left(\frac{v(y,z) \ k_y(y,z)}{\omega} \right)^2} \tag{7.20}$$

A *migrated time section* is defined by transforming the depth variable z in (1.20) to a travel-time depth τ.

$$k_\tau(y,\tau) \quad = \quad \omega \ \sqrt{1 \ - \ \left(\frac{v(y,\tau) \ k_y(y,\tau)}{\omega} \right)^2} \tag{7.21}$$

The implementation of equation (1.21) requires no lens terms, so no large sensitivity to lateral velocity variation is expected. Unfortunately, there is a pitfall. The (y,z) coordinate system is an orthogonal coordinate system, but the (y,τ) system is not orthogonal [unless $v(y)$ =const]. So equation (1.20), which says that $\cos\theta = \sqrt{1-\sin^2\theta}$, is not correctly interpreted by (1.21). A hyperbola would migrate to its top when it should be migrating toward the low-velocity side.

In summary: In a production environment a great deal of data gets processed before anyone has a clear idea of how much lateral velocity variation is present. So the lens term is omitted. The results are OK if the lens term happens to commute with the diffraction term. The terms do commute when the lateral velocity variation is slow enough. Otherwise, you should reprocess with the lens term. The reprocessing will be sensitive to errors in velocity. Be careful!