1.3 Four Wide-Angle Migration Methods[1]

The four methods of migration of reflection seismic data that are described here may all be found in modern production environments. As a group they are all strong in their ability to handle wide-angle rays. As a group they are all weak in their ability to deal with lateral velocity variation.

Traveltime Depth

Conceptually, the output of a migration program is a picture in the (x,z)-plane. In practice the vertical axis is almost never depth z; it is the ***vertical traveltime*** τ. In a constant-velocity earth the time and the depth are related by a simple scale factor. The meaning of the scale factor is that the (x,τ)-plane has a vertical exaggeration compared to the (x,z)-plane. In reconnaissance work, the vertical is often exaggerated by about a factor of five. By the time prospects have been narrowed to the point where a drill site is being selected, the vertical exaggeration factor in use is likely to be about unity (no exaggeration).

The traveltime depth τ is usually defined to include both the time for the wave going down and for the wave coming up. The factor of 2 thus introduced quickly disappears into the rock velocity. Recall that zero-offset data sections are generally interpreted in terms of exploding-reflector wave fields. To make the correspondence, the rock velocity is halved for the wave analysis:

$$\tau \;=\; \frac{2z}{v_{true}} \;=\; \frac{z}{v_{half}} \qquad (1)$$

The first task in interpretation of seismic data is to figure out the approximate numerical value of the vertical exaggeration. It is doubtful that it will be printed on the data header for the simple reason that it is not exactly known. This is because the seismic velocity is not exactly known. Furthermore, the velocity usually ***increases*** with depth, which means that the vertical exaggeration ***decreases*** with depth. For velocity-stratified media, we may write the time-to-depth conversion formula

$$\tau(z) \;=\; \int_0^z \frac{dz}{v(z)} \quad or \quad \frac{d\tau}{dz} \;=\; \frac{1}{v} \qquad (2)$$

[1] Adapted from SEP-25, pp 209-220.

Hyperbola Summation and Semicircle Superposition Methods

These methods are the most comprehensible of all known methods. Conceptually, at least, they seem to predate the use of computers. Computer implementations of these methods seem to predate the exploding-reflector concept, and they certainly predate the idea of downward extrapolating a wave field with $exp(ik_z z)$ followed by imaging at $t=0$.

First of all, recall the equation for a conic section, a circle in (x,z)-space or a hyperbola in (x,t)-space. With traveltime depth τ, we get

$$x^2 + z^2 \;=\; v^2 t^2 \qquad (3a)$$

$$\frac{x^2}{v^2} + \tau^2 \;=\; t^2 \qquad (3b)$$

Figure 1 illustrates the ***semicircle-superposition*** method. Taking the data field to contain a few impulse functions, then the output should be a superposition of the appropriate semicircles. Each semicircle denotes the spherical reflector earth model, which would be implied by a dataset with a single pulse. Taking the data field to be a thousand seismograms of a thousand points each, then the output is a superposition of a million semicircles. Since a seismogram has both positive and negative polarities, about half the semicircles will be superposed with negative polarities. The resulting superposition could look like almost anything. Indeed, the semicircles might mutually destroy one another almost everywhere except at one isolated impulse in (x,τ)-space. Should this happen you might rightly suspect that the input data section in (x,t)-space is a Huygens secondary source, namely energy concentrated along a hyperbola. This leads us to the ***hyperbola summation*** method.

The ***hyperbola summation*** method of migration is depicted in figure 2. The idea is to create one point in (x,τ)-space at a time, unlike the semicircle method, where each point in (x,τ)-space is built up bit by bit as the one million semicircles are stacked together. To create one fixed point in the output (x,τ)-space, a hyperbola, equation (3b), is imagined set down with its top upon the corresponding position of (x,t)-space. All data values touching the hyperbola are added together to produce a value for the output at the appropriate place in (x,τ)-space. In the same way, all other locations in (x,τ)-space are filled.

The opposite of data processing, building models from data, is constructing synthetic data from models. By means of a slight modification, the above two processing programs can be converted to modeling

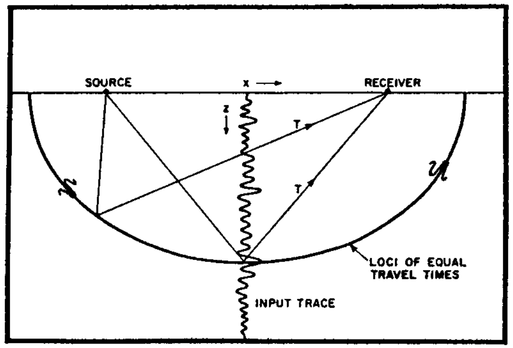


FIG. 1. [from Schneider, W. A., 1971, Developments in seismic data processing and analysis (1968-1970): Geophysics, v. 36, no. 6, p. 1043-1073] The -------- process may be described in numerous ways; however, two very simple and equally valid representations are indicated in figures 1 and 2. Shown here is a representation of the process in terms of what happens to a single input trace plotted in depth (time may also be used) midway between its source and receiver. Each amplitude value of this trace is mapped into the subsurface along a curve representing the loci of points for which the traveltime from source to reflection point to receiver is constant. If the velocity is constant, these curves are ellipses with source and receiver as foci. The picture produced by this operation is simply a wavefront chart modulated by the trace amplitude information. This clearly is not a useful image in itself, but when the map is composited with similar maps from neighboring traces (and common-depth-point traces of different offsets), useful subsurface images are produced by virtue of constructive and destructive interference between wavefronts in the classical Huygens sense. For example, wavefronts from neighboring traces will all intersect on a diffraction source, adding constructively to produce an image of the diffractor in the form of a high-amplitude blob whose (z,x) resolution is controlled by the pulse bandwidth and the horizontal aperture of the array of neighboring traces composited. For a reflecting surface, on the other hand, wavefronts from adjacent traces are tangent to the surface and produce an image of the reflector by constructive interference of overlapping portions of adjacent wavefronts. In subsurface regions devoid of reflecting and scattering bodies, the wavefronts tend to cancel by random addition.

programs. Instead of *hyperbola summation* or *semicircle superposition*, one does *hyperbola superposition* or *semicircle summation*. You might wonder whether the processing programs really are the inverse to the modeling programs. You might also wonder whether the two different methods of modeling or processing are equivalent. If they differ, which

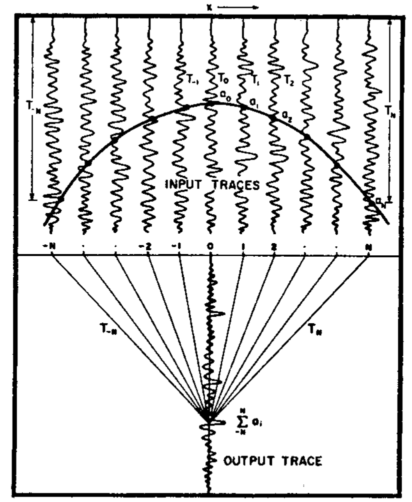


FIG. 2. (from Schneider, 1971 - see figure 1 caption for complete reference) A second description of the -------- process is provided here. The process is represented in terms of how an output trace is developed from an ensemble of input traces, shown as CDP-stacked traces in the upper half of the figure. The output in the lower half reflects how each amplitude value at (x,z) is obtained by summing input amplitudes along the traveltime curve shown. This curve defines a diffraction hyperbola, and if a diffraction source existed in the subsurface at the output point shown, a large amplitude would result. The process also works for reflectors since we may regard a reflector as a continuum of diffracting elements whose individual images merge to produce a smooth continuous boundary.

is better? Clearly some facts which have been glossed over are (1) the angle-dependence of amplitude (obliquity function) of the Huygens waveform, (2) spherical spreading of energy, and (3) the phase shift on the Huygens waveform. Actually, results are reasonably good even without these. Since proper migration is an all-pass filter, the inverse to the filter should be simply its time reverse. So cross-correlating the data (a superposition of hyperbolas) with another hyperbola should result in a filter that is a pretty good inverse to the filter which is a delta function on a hyperbola.

As later methods of migration were developed, the deficiencies of the earlier methods became more clearly understood, and were largely correctable by careful implementation. One advantage of the later methods was that they implemented true all-pass filters. Such migrations preserve the general appearance of the data. This suggests restoration of high frequencies, which tend to be destroyed by hyperbolic integrations. Work by Trorey, Schneider, Hilterman, and possibly others with the Kirchoff diffraction integral suggested quantitative means of bringing hyperbola methods into agreement with other methods, at least for constant velocity. Common terminology nowadays is to refer to any hyperbola or semicircular method as a Kirchoff method, although, strictly speaking, the Kirchoff integral applies only in the constant-velocity case.

Spatial Aliasing

The situation when data is insufficiently sampled on the space axis is called ***spatial aliasing***. This difficulty is so universal, that all all migration methods must consider it. It is commonly agreed that data should be sampled at more than about two points per wavelength. Otherwise the wave arrival direction becomes ambiguous. Figure 3 shows some synthetic data which is sampled insufficiently densely along the x-axis.

You can see that the problem becomes more acute at high frequencies and steep dips. There seems to be no automatic method for migrating data which is spatially aliased (a common problem). In such cases, human beings can usually do better than machines, because of their ability to recognized the true slopes. But when the data is adequately sampled, then computer migration of data based on the wave equation gives better results than manual methods. Contemporary surveys are usually adequately sampled along the line of the survey. But there is often considerable difficulty in the perpendicular direction.

The hyperbola-sum-type methods run the risk that the migration operator itself can become spatially aliased. This is a situation to be avoided by means of careful implementation. The first thing to realize is that you should be ***integrating*** along a hyperbolic trajectory. A summation incorporating only one point per trace is a poor approximation. It is better to incorporate more points, as depicted in figure 4. The likelihood of getting an aliased operator increases where the hyperbola is steep-sloped. In production examples, an aliased operator often stands out on the seafloor reflection where — although it may be perfectly flat — it acquires a noisy precursor from the steep-flanked hyperbola of the water path.

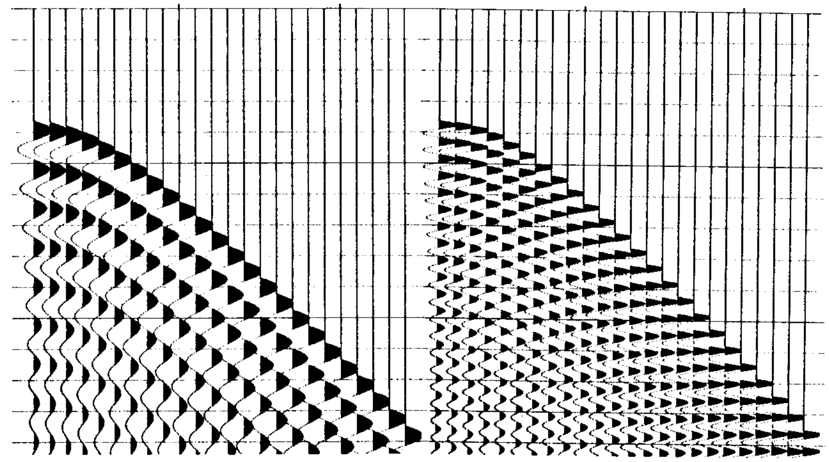
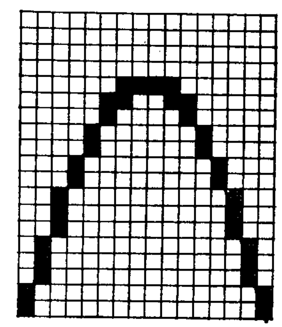


FIG. 3. Insufficient spatial sampling of synthetic data. To better perceive the ambiguity of arrival angle, view the figures at a grazing angle from the side.

FIG. 4. For a low velocity hyperbola, integration will require more than one point per channel.



The Phase-Shift Method (Gazdag)

The phase-shift method proceeds straightforwardly by extrapolating downward with $\exp(ik_z z)$ and subsequently evaluating the wave field at

$t=0$ (that is, when the reflectors explode). Of all the wide-angle methods it most easily incorporates depth variation in velocity. Even the phase angle and obliquity function are correctly included, automatically. Unlike Kirchoff methods, there is no danger of aliasing the operator. This method is also quite comprehensible.

To start with, you need to do a two-dimensional Fourier transform (2D-FT) of your dataset. Some practical details about 2D-FT are described in a later section. Then you push the transformed data values, all in the (ω, k_x)-plane, downward to a depth Δz by means of multiplication by

$$e^{i k_z \Delta z} = \exp\left\{ -i \frac{\omega}{v} \left[1 - \left(\frac{v k_x}{\omega} \right)^2 \right]^{1/2} \Delta z \right\} \quad (4)$$

Ordinarily the time-sample interval $\Delta\tau$ for the output migrated section will be chosen equal to the time-sample rate of the input data (often 4 ms). Thus, choosing the depth $\Delta z = v\,\Delta\tau$, the downward-extrapolation operator for a single time unit is

$$\exp\left\{ -i\,\omega\,\Delta\tau \left[1 - \left(\frac{v k_x}{\omega} \right)^2 \right]^{1/2} \right\} \quad (5)$$

Next is the task of imaging. At each depth we imagine an inverse Fourier transform followed by selection of the value at $t=0$. Luckily, only the Fourier transform at one point, $t=0$, is needed, so that is all that need be computed. It is especially easy since the value at $t=0$ is merely a summation of each ω frequency component. Finally, inverse Fourier transform k_x to x. The overall migration process may be summarized as follows.

$P(\omega, k_x) = FT[p(t,x)]$
For $\tau = \Delta\tau, 2\Delta\tau$, ..., end of time axis on seismogram {
  For all k_x {
    For all ω {
      $P(\omega, k_x) = P(\omega, k_x) \exp[-i\,\omega\Delta\tau \cos\vartheta(\omega, k_x)]$
      }
    $Image(k_x, \tau) = \sum_\omega P$
    }
  $Image(x, \tau) = FT[Image(k_x, \tau)]$
  }

One hardly ever knows the velocity very precisely, so although the velocity may be increasing fairly steadily with depth, it is often approximated as constant in about 20 layers, rather than slowly changing at each of the thousand or so points on a seismogram. The advantage of velocity being constant in layers is one of economy. Once the square root and the sines and cosines in (5) have been computed, then the complex multiplier (5) can be re-used for all 20 layers.

The Stolt Method

On most computers the Stolt method of migration is the fastest method, by a considerable margin. For many applications, this will be the most important attribute to consider. In a constant-velocity earth the Huygens wave source is treated exactly correctly. Like the other methods, this migration method can be reversed and made into a modeling program. One drawback, a matter of principle, is that the method does not handle depth variation in velocity. This drawback is largely offset in practice by the existence of an approximate correction by an axis-stretching procedure. A practical drawback is the periodicity of all the Fourier transforms. In principle this is no problem at all, being solvable by a sufficient surrounding of the data by zeros. A single line sketch of the Stolt method is this:

$$P(x,t) \rightarrow P(k_x, \omega) \rightarrow P\left[k_x,\ k_z = \left(\frac{\omega^2}{v^2} - k_x^2 \right)^{1/2} \right] \rightarrow P(x,z)$$

To see why this works, begin with the input-output relation for downward extrapolation of wave fields:

$$P(\omega, k_x, z) = e^{i k_z z}\, P(\omega, k_x, z=0) \quad (6)$$

Perform a two-dimensional inverse Fourier transform:

$$p(t,x,z) = \int\int e^{i k_x x - i\omega t + i k_z z}\, P(\omega, k_x, 0)\, d\omega\, dk_x$$

Apply the idea that the image at (x,z) is the exploding reflector wave at time $t=0$:

$$Image(x,z) = \int\int e^{i k_x x}\, e^{i k_z(\omega, k_x) z}\, P(\omega, k_x, 0)\, d\omega\, dk_x \quad (7)$$

Equation (7) states the answer we want, but it is in a very unattractive form. The computational effort implied by (7) is that a two dimensional integration must be done for each and every z-level. The Stolt procedure will be to convert the three dimensional calculation implied by (7) to a single two dimensional Fourier transform.

So far we have done nothing to specify that we have an *upgoing* wave instead of a downgoing wave. The direction of the wave is defined by the relationship of z and t required to keep the phase constant in the expression $\exp(-i\omega t + ik_z z)$. If ω were always positive, then $+k_z$ would always refer to a downgoing wave and $-k_z$ to an upgoing wave. We need negative frequencies ω as well as positive frequencies in order to describe waves that have real values (not complex). So the proper description for a downgoing wave is that the signs of ω and k_z must agree. For an upgoing wave it is the reverse. With this clarification we prepare to change the integration variable in (7) from ω to k_z.

$$\omega \;=\; -sgn(k_z)\, v \sqrt{|k_x^2 + k_z^2|} \tag{8a}$$

$$\frac{d\omega}{dk_z} \;=\; -sgn(k_z)\, v \,\frac{k_z}{+\sqrt{k_x^2 + k_z^2}} \tag{8b}$$

$$\frac{d\omega}{dk_z} \;=\; \frac{-v\,|k_z|}{+\sqrt{k_x^2 + k_z^2}} \tag{8c}$$

Now we will introduce (8) into (7) including also a minus sign so that the integration on k_z may be taken from minus infinity to plus infinity as was the integration on ω.

$$Image(x,z) = \int\int e^{ik_z z + ik_x x} \left\{ P[\omega(k_x,k_z),k_x,0]\, \frac{v\,|k_z|}{\sqrt{k_x^2 + k_z^2}} \right\} dk_z\, dk_x \tag{9}$$

Equation (9) states the final result as a two-dimensional inverse Fourier transform. The Stolt migration method is a direct implementation of (9). The steps of the algorithm are:

1. Double Fourier transform data from $p(t,x,0)$ to $P(\omega,k_x,0)$.
2. Reinterpolate P onto a new mesh so that it is a function of k_x and k_z. Multiply P by the scale factor (which has the interpretation $\cos\vartheta$).
3. Inverse Fourier transform to (x,z)-space.

Samples of Stolt migration of impulses are shown in figure 5. You can see the expected semi-circular smiles. You can also see a semicircular frown hanging from the bottom of each semicircle. The worst frown is on the deepest spike. What seems to be happening is that our semicircular mirrors are actually circular, with centers not only at the earth's surface $z=0$ but also at the bottom of the model $z=z_{\max}$. The obvious practical solution is to stay away from the bottom of the model.

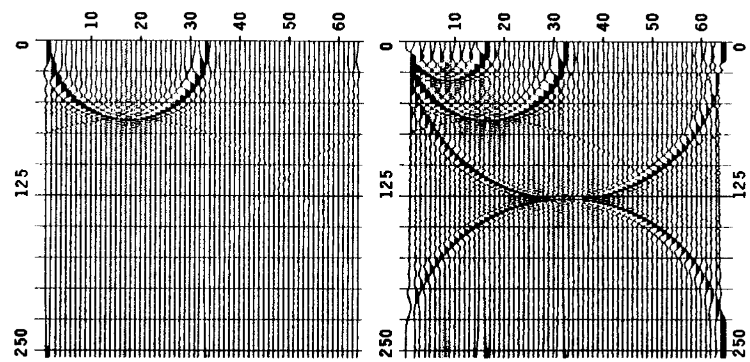


FIG. 5. Response of Stolt method to data with impulses. Semicircles are seen, along with computation artifacts.

It seems that we must use an extraordinary amount of zero padding on the time axis. To keep memory costs reasonable, the algorithm can be reorganized. We need storage space for a long vector, say $u(t)$, (about four times as much as for a typical seismogram). If you have an array processor, that is where this vector belongs. The Stolt algorithm becomes:

$P(k_x,t) = FT[p(x,t)]$
For all k_x {

$u(t) = P(k_x,t)$

Pad out remaining length of u with zeros.

$U(\omega) = FT[u(t)]$

$$U'(k_z) = U[-sgn(k_z) v \sqrt{k_x^2 + k_z^2}] \frac{v\,|k_z|}{\sqrt{k_x^2 + k_z^2}}$$

$u(z) = FT[U'(k_z)]$

$P(k_x,z) = u(z)$

}

$p(x,z) = FT[P(k_x,z)]$

Even this improved algorithm is not trouble-free. The periodicity in x still requires padding with lots of zeros on x.

Sensitivity of Migration to Velocity Error

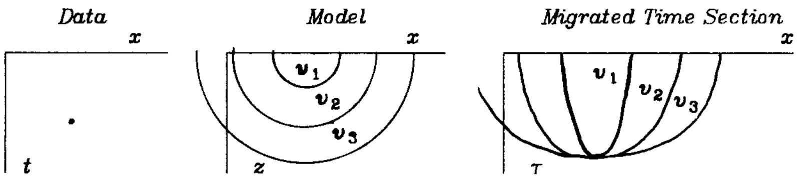
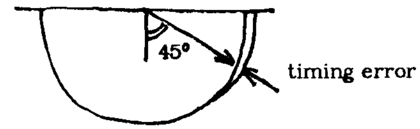


FIG. 6. Migration of a data impulse as a function of velocity.

Recall that migrated data is displayed as a *time* section. Figure 6 shows how the migration impulse response depends on velocity. Arbitrary velocity error makes no difference when processing horizontal bedding. Velocity error sensitivity increases with angle up to 90 degrees, where the accuracy needed to avoid destructive interference is in the ratio of half a seismic wavelength divided by the traveltime. Significant timing error may be assumed to be about a half-wavelength. Observationally the ratio of traveltime to wavelength is usually a hundred or less. [Exceptions occur (1) if much of the path is in water or (2) at time depths greater than about four seconds.] Figure 7 illustrates that the velocity accuracy required for 90-degree migration is about 1%. For 45-degree migration velocity error could be larger by the square root of 2.

FIG 7. Timing error of the wrong velocity increases with angle.



Velocities are rarely known this accurately. Is there any value in processing for which erroneous time shifts exceed a half-wavelength?

Subjective Comparison and Evaluation of Methods

The three basic methods of migration in this section are compared subjectively in table 1.

Finally, the perspective of later chapters allows some remarks on the overall quality of the wide-angle methods as a group. Their greatest weakness is their near inability to deal with lateral velocity variation. Their greatest strength, the wide-angle capability, is lessened in value by the weakness of other links in the data collection and processing chain.

Namely:

1. Shot-to-geophone offset angles are commonly large but ignored. A CDP stack is not a zero-offset section.
2. Why process to the very wide angles seen in the survey line when even tiny angles perpendicular to the line are being ignored?
3. Data is often insufficiently densely sampled to represent steeply dipping data without aliasing.
4. Accuracy in knowledge of velocity is seldom sufficient to justify processing to wide angles.

EXERCISES

1. Define the computer program for modeling with the phase shift method -- that is, create the surface data $P(x,z=0,t)$ from some exploding reflector distribution $P(x,z,t=0)$.

	Hyperbola Sum Semicircle Sup.	Phase Shift	Stolt
Speed	slow	average	very fast
Memory organization	awkward	good	good
$v(z)$	ray tracing	easily	approximately by stretching
wide angle?	Beware of data alias and operator alias.	Beware of data alias.	Beware of data alias.
Correct phase and obliquity?	possible with some effort for const v	easily for any $v(z)$	for const v
wraparound noise?	no	worst on x, but reasonable	(x,z,t) a problem
$v(x)$	Production programs have serious pitfalls.	approximately by iteration and interpolation	no known program
Side boundaries and irregular spacing	Excellent	Poor	Poor

TABLE 1. Subjective comparison of three wide-angle migration methods.

2. Define the computer program for the inverse to the Stolt algorithm - that is, create synthetic data from a given model.
3. Redesign the computer program of the phase shift migration method to save unnecessary computation when $v(z)$ is constant over a range of z.
4. The phase shift method tends to produce a migration which is periodic with z because of the periodicity of the Fourier transform over t. Ordinarily, this is not troublesome because we do not look at large z. The up-coming wave at great depth should be zero *before* $t=0$. Kjartansson pointed out that periodicity in z could be avoided if the wave at $t=0$ is subtracted from the wavefield before descending further. Thus information could never get to negative time to "wrap-around". Indicate how the program should be changed.