Chapter 8

Dispersion relations

This chapter attends to those details that enable us to do a *high-quality* job of downward continuing wavefields. There will be few new seismic imaging concepts here. There will, however, be interesting examples of pitfalls. And in order to improve the quality of seismic images of the earth, several new and interesting mathematical concepts will be introduced. Toward the end of the chapter a program is prepared to simulate and compare various migration methods.

8.0.1 The magic of color

The first thing we will consider in this chapter is signal strength. Echoes get weaker with time. This affects the images, and requires compensation.

Next, seismic data is colored by filtering. This filtering can be done in space as well as time. Time-series analysis involves the concept of enhancing the signal-to-noise ratio by filtering to suppress some spectral regions and enhance others. Spectral weighting can also be used on wavefields in the space of ω and k. In the absence of noise, wave-equation theory tells us what filters to use. Loosely, the wave equation is a filter with a flat amplitude response in (ω, k)-space and a phase response that corresponds to the time delays of propagation. The different regions of (ω, k)-space have different amounts of noise. But the different regions need not all be displayed at the strength proposed by the wave equation, any more than data must be displayed with $\Delta x = \Delta z$.

An example of the mixture of filter theory and migration theory is provided by the behavior of the spatial Nyquist frequency. Because seismic data is often spatially aliased, this example is not without practical significance. Think of an impulse function with its Nyquist frequency removed. The removal has little relative effect on the impulse, but a massive *relative* effect on the zeroes surrounding the impulse. When migrating an impulse by frequency domain methods, spatial frequencies just below the spatial Nyquist are treated much differently from frequencies just above it.

One is treated as left dip, the other as right dip. This discontinuity in the spatial frequency domain causes a spurious, spread-out response in the space domain shown in Figure 1.1.

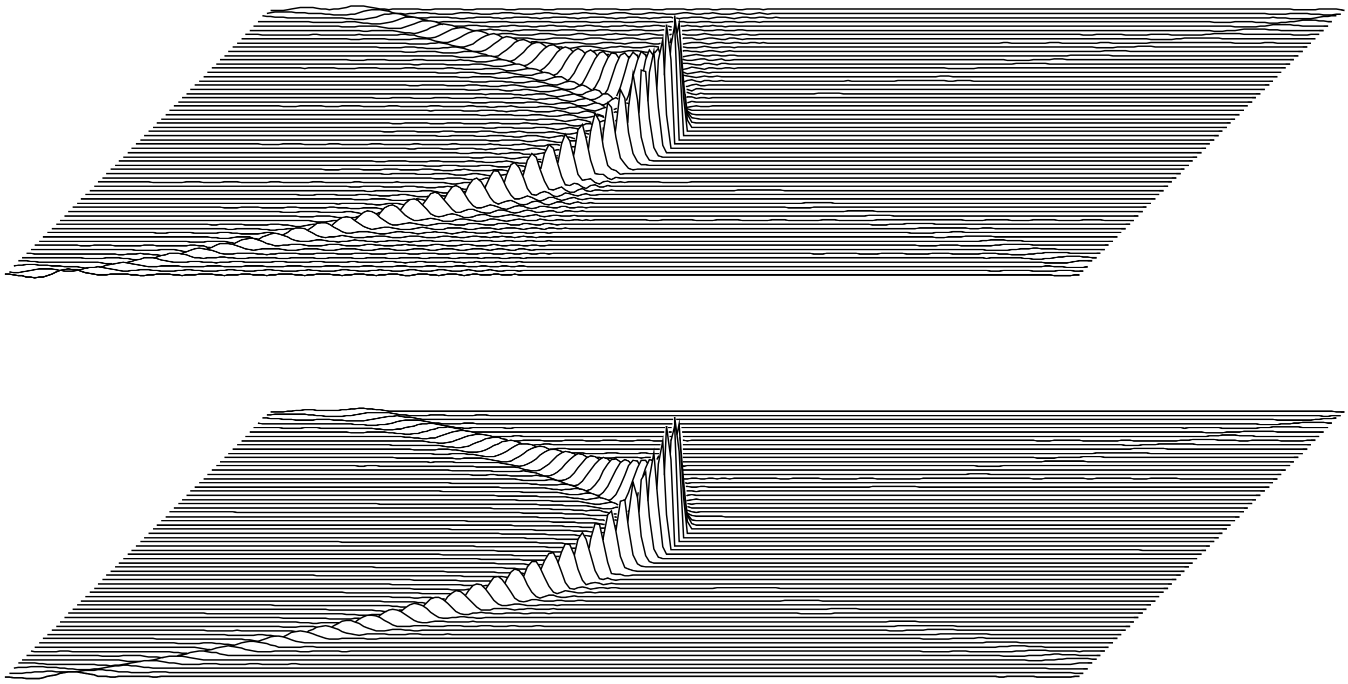


Figure 8.1: Hyperbola amplified to exhibit surrounding Nyquist noise (top) removed by filtering (bottom). dspr-hypnoise [ER]

The spurious Nyquist noise is readily suppressed, not by excluding the Nyquist frequency from the display, but by a narrow band filter such as used in the display, namely $(1 + \cos\, k_x \Delta x)/(1 + .85\, \cos\, k_x \Delta x)$ which goes smoothly to zero at the spatial Nyquist frequency. This filter has a simple tridiagonal representation in the x-domain.

8.0.2 Survey of migration technique enhancements

In our quest for quality, we will also recall various approximations as we go. Now is the time to see how the use of approximations degrades results, and to discover how to improve those results. Five specific problems will be considered:

1 The frequency dispersion that results from the approximation of differential operators by difference operators

2 The anisotropy distortion of phase and group velocity that results from square-root approximations

3 The effect of truncation at the end of the survey line

4 Dips greater than ninety degrees

5 Wraparound problems of Fourier transformation

6 The effect of $v(z)$ upon the Stolt method and how to improve the result by stretching

Following study of these approximations, is a penetrating study of causality, which covers much ground including how Fourier domain migration can simulate the causality intrinsic to time domain migrations and a grand summary of techniques. A single program is presented that can simulate diffraction hyperbolas from many different migration methods. This facilitates comparison of techniques and optimization of parameters. Figure 1.1 and many of the other figures in this chapter were produced with this program, so you should be able to reproduce them.

8.0.3 A production pitfall: weak instability from v(x)

Some quality problems cannot be understood in the Fourier domain. Unless carefully handled, lateral velocity variation can create instability.

The existence of lateral velocity jumps causes reflections from steep faults. A more serious problem is that the extrapolation equations themselves have not yet been carefully stated. The most accurate derivation of extrapolation equations included in this book so far was done from dispersion relations, which themselves imply velocity constant in x. The question of how a dispersion relation containing a $v\,k_x^2$ term should be represented was never answered. It might be represented by $v(x,z)\partial_{xx}$, $\partial_x v(x,z)\partial_x$, $\partial_{xx}v(x,z)$ or any combination of these. Each of these expressions, however, implies a different numerical value for the internal reflection coefficient. Worse still, by the time all the axes are discretized, it turns out that one of the most sensible representations leads to reflection coefficients greater than unity and to numerical instability.

A weak instability is worse than a strong one. A strong instability will be noticed immediately, but a weak instability might escape notice and later lead to incorrect geophysical conclusions. Fortunately, a stability analysis leads to a *bulletproof* method at the end of chapter 9.

8.1 COSMETIC ASPECT OF WAVE EXTRAPOLATION

Frequency filtering, dip filtering, and gain control are three processes whose purposes seem to be largely cosmetic: they are used to improve the appearance of data. The criteria used to choose the quantitative parameters of these and similar processes are often vague and relate to human experience or visual perception. In principle, it should be possible to choose the parameters by invoking information theory and using objective criteria such as signal and noise dip spectra. But in routine practice this is not yet being done.

The importance of cosmetic processes is not to be underestimated. On many occasions, for example, a comparison of processing techniques (in order to choose a contractor perhaps?) has been frustrated by an accidental change in cosmetic parameters. These cosmetic processes arise naturally within wave-propagation theory. It seems best to first understand how they arise, and then to build them into the processing, rather than try to append them in some artificial way after the processing. The individual parts of the wave-extrapolation equations will now be examined to show their cosmetic effects.

8.1.1 t squared

Echos get weaker with time. To be able to see the data at late times, we generally increase data amplification with time. I have rarely been disappointed by my choice of the function t^2 for the scaling factor. The t^2 scaling function cannot always be expected to work, because it is based on a very simple model. But I find t^2 to be more satisfactory than a popular alternative, the growing exponential. The function t^2 has no parameters whereas the exponential function requires two parameters, one for the time constant, and one for the time at which you must stop the exponential because it gets too large.

The first of the two powers of t arises because we are transforming three dimensions to one. The seismic waves are spreading out in three dimensions, and the surface area on the expanding spherical wave increases in proportion to the radius squared. Thus the area on which the energy is distributed is increasing in proportion to time squared. But seismic amplitudes are proportional to the square root of the energy. So the basic geometry of energy spreading predicts only a single power of time for the spherical divergence correction.

An additional power of t arises from a simple absorption calculation. Absorption requires a model. The model I'll propose is too simple to explain everything about seismic absorption, but it nicely predicts the extra power of t that experience shows we need. For the model we assume:

1 One dimensional propagation
2 Constant velocity
3 Constant absorption Q^{-1}
4 Reflection coefficients random in depth
5 No multiple reflections
6 White source

These assumptions immediately tell us that a monochromatic wave would decrease exponentially with depth, say, as $\exp(-\alpha\,\omega\,t)$ where t is travel-time depth and α is a decay constant which is inversely proportional to the wave quality factor Q. Many people go astray when they model real seismic data by such a monochromatic wave. A better model is that the seismic source is broad band, for example an impulse function. Because of absorption, high frequencies decay rapidly, eventually leaving only low frequencies, hence a lower signal strength. At propagation time t the original white (constant) spectrum is replaced by the forementioned function $\exp(-\alpha\,\omega\,t)$ which is a damped exponential function of frequency. The seismic energy available for the creation of an impulsive time function is just proportional to the area under the damped exponential function of frequency. As for the phase, all frequencies will be in phase because the source is assumed impulsive and the velocity is assumed constant. Integrating the exponential from zero to infinite frequency provides us with an inverse power of t thus completing the justification of a t^2 divergence correction.

It is curious that the shape of the expected seismogram envelope t^{-2} does not depend on the dissipation constant α. But changing the spectrum of the seismic source does change the shape of the envelope. It is left for an exercise to show that a seismic source with spectrum $|\omega|^\beta$ implies a divergence correction $t^{2+\beta}$.

The seismic velocity increases with depth, so sometimes people who know the velocity may improve the divergence correction by making it a function of velocity (and hence offset) as well as time.

In reality it may be fortuitous that t^2 fits data so well. Actually, Q generally increases with depth whereas reflection coefficients generally decrease with depth.

8.1.2 Noise, surface waves and clip

If seismic data contained nothing but reflections, then there would be little trouble plotting it. You would simply multiply by t^2 and then scale so that the largest data values stayed in the available plotting area. In reality there are two problems: (1) noisy traces and (2) noise propagation modes. We have noisy traces because the people in the world won't all be quiet while we listen for echoes. Noise propagation modes are waves trapped in surface layers. So their divergence is in a two-dimensional space rather than the three-dimensional space for reflections. Water noises are additionally strong because of the homogeneity and low absorption of water.

Noises are handled by "clipping" data values at some level lower than the maximum. Clipping means that values larger than the clip value are replaced by the clip value. Since the size of the noise is generally unpredictable, the most reliable method is to use quantiles. Imagine the data points sorted in numerical order by the size of their absolute values. The $n^{\rm th}$ quantile is defined as the absolute value that is $n/100$ of the way between the smallest and largest absolute value. So if data is clipped at the 99^{th} percentile, then up to one percent of the data can be infinitely strong noise. I find that most field profiles have less than 10% noisy points. So I often clip at thrice the $90^{\rm th}$ percentile. To find the quantile, it is not necessary to fully sort the data. That would be slow. Hoare's algorithm is much faster (see FGDP or Claerbout and

Muir [1973] for full reference and more geophysical context).

Different plots have different purposes. It is often important to preserve linearity during processing, but at the last stage—plotting—linearity can be sacrificed to enable us to see all events, large and small. After all, human perceptions are generally logarithmic. In our lab we generally use power laws. I find that replacing data points by their signed square roots generally compresses all signals into a visible range. When plotting field profiles with a very close trace spacing, it may be better to use the signed cube roots. More generally, we do non-linear gain with

$$\text{Display} \quad = \quad \text{sgn(Data)} \quad |\text{Data}|^{\gamma} \tag{8.1}$$

Gamma is a term in photography to describe nonlinearity of photographic film. Most of the data plots in this book use $\gamma = 1$, t^2 gain, and clip at the 99$^{\text{th}}$ percentile.

The industry standard approach seems to be AGC (Automatic Gain Control). AGC means to average the data magnitude in some interval and then divide by the magnitude. Although AGC is nonlinear, it is more linear than using γ so it is presumably better if you plan later processing. But with AGC, you lose reversibility and the sense of absolute gain.

Figure 1.2 is an interesting example. Since it is a split spread, you assume it to be land data. Ships can't push cables in front of them. But the left panel clearly shows marine multiples. The reverberation period is uniform, and there are no reflections before the water bottom. It must be data collected on ice over deep water (375m). From the non-linear gain in the center panel we clearly see a water wave, and before it a fast wave in the ice. There is also weak low-velocity, low-frequency "ground roll" on the ice. There are also some good reflections.

8.1.3 Complex velocity in the 5 degree equation

The 5° equation, namely,

$$\frac{\partial P}{\partial z} \quad = \quad \frac{1}{v}\,\frac{\partial P}{\partial t} \tag{8.2}$$

$$i\,k_z \quad = \quad \frac{-i\omega}{v} \tag{8.3}$$

states that a wavefront will take some time to get from one depth to another. With velocity v being a real constant, waves controlled by (1.2) and (1.3) propagate without change in form. In practice waveform changes are observed. So v should not be a real constant. An imaginary part of the velocity would cause attenuation. A frequency-dependent velocity would cause frequency dispersion.

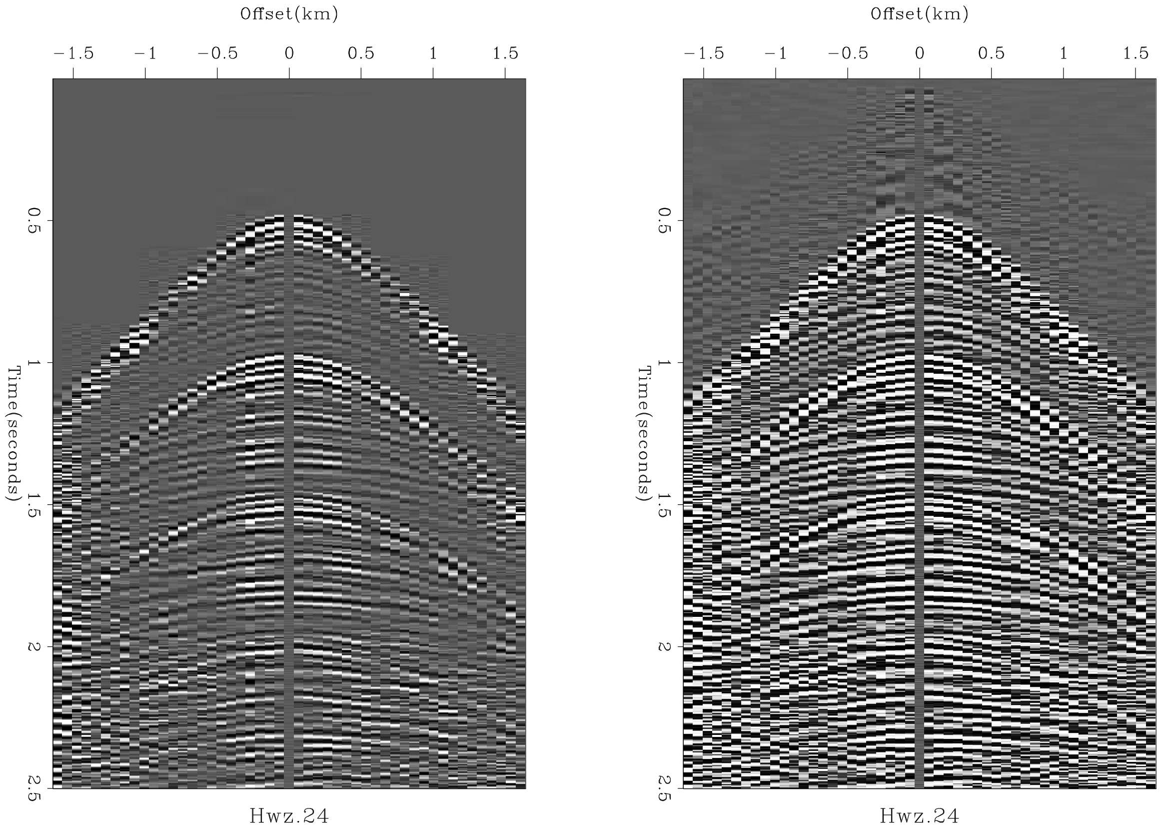


Figure 8.2: Arctic profile from Western Geophysical. Left, with t^2. Center with t^2 and $\gamma = 0.4$. Right, with Western's AGC. dspr-arctic [ER]

8.1.4 Absorption

A basic model arises when $v(\omega)$ is defined by the equation

$$\frac{-i\omega}{v(\omega)} \quad = \quad \frac{\omega_0}{v_0} \left(\frac{-i\omega}{\omega_0} \right)^{1-\epsilon} \tag{8.4}$$

For $\epsilon = 0$, equation (1.4) gives a constant velocity. Equation (1.4) models the so-called causal, constant Q attenuation where $Q^{-1} = \tan \pi\epsilon$. Figure 1.3 shows an example of a synthetic seismogram generated by the exploding-reflector model using equations (1.2), (1.3) and (1.4).

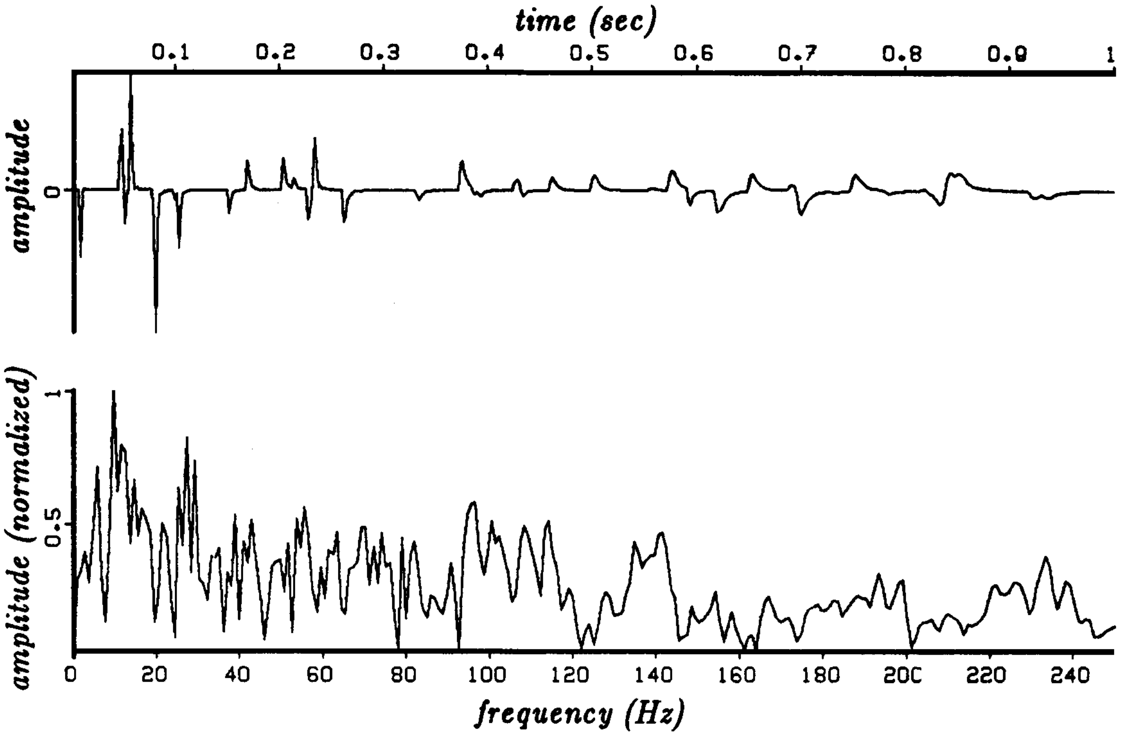


Figure 8.3: Synthetic seismogram for an earth with $Q \approx 100$. (Hale) dspr-qhale [NR]

Equation (1.4) creates attenuation by introducing an imaginary part into the velocity. The main effect of this attenuation is to weaken the arrivals at late time. A secondary effect is to make the frequency content of late arrivals lower. A tertiary effect is this: It happens that the requirement of causality forces the real part of the velocity to be slightly frequency-dependent. In the figure, this slight frequency-dependence is evidenced by the "rise time" on each pulse being faster than the "fall time." This means that the high frequencies are traveling slightly faster than the low frequencies. In practice, this tertiary effect is rarely noticeable.

In making earth images, earth dissipation might be compensated by amplifying high-frequency energy during downward continuation. This might be done just like migration, except that $k_z = -\sqrt{\omega^2/v^2 - k_x^2}$ would be replaced by something like $ik_z = (-i\omega)^{1-\epsilon}$. In practice, however, no one would really do this, since it would amplify high frequency noise. This raises the issue of signal-to-noise ratio.

Noise isn't simply an ambient random fluctuation. It is mainly repeatable if the data is reshot. Noise is anything for which we have no satisfactory model. On a practical level, time-variable filters are often used to select pleasing time-variable passbands. Equations (1.2), (1.3) and (1.4) could be used for this implementation of time-variable filters, but it would be an oversimplification to view their use as compensation for the earth Q.

8.1.5 Dispersion

The frequency-dependence of velocity in the case of surface waves is more dramatic. For example, a frequency-dependent velocity is given by the equation

$$\frac{-i\omega}{v(\omega)} \quad = \quad \frac{-i\omega}{v_0} \sqrt{1 \ + \ \omega_0^2/\omega^2} \tag{8.5}$$

The left side of Figure 1.4 contains some frequency-dispersive ground roll. On the right side of Figure 1.4, the dispersion has been backed out by a migration-like process. One difference between this process and migration is that migration extrapolates down the z-axis whereas on the right of Figure 1.4 the extrapolation is along the x-axis. (The extrapolation direction is really just in the computer). Each trace on the right of Figure 1.4 is processed separately. In migration, data $p(t, z = 0)$ is extrapolated to an image $p(t = 0, z)$ using a dispersion relation $k_z = -\sqrt{\omega^2/v^2 - k_x^2}$. In this process, data $p(t, x = 0)$ is extrapolated to an image $p(t = 0, x)$ using a dispersion relation like $k_x = f(\omega/v)$. After this pseudomigration a pseudodiffraction is done with a *constant* velocity. The total effect is to undo the frequency dispersion. Finally, it is possible to see that the noise consists of two separate events. Techniques resembling this one were first used to locate faults in coal seams (Beresford-Smith and Mason [1980]).

8.1.6 False semicircles in migrated data

Dip filtering can be used to suppress multiples. Chapter 11 will show that multiples are unlike primaries in one important respect: their strength may change rapidly in the horizontal direction. They need not be spread out into broad diffraction hyperbolas as primaries must. This difference arises because multiples often spend much time focusing themselves in the irregular, near-surface areas. Common evidence for this behavior is contained in the appearance of wide-angle migrated sections. Such sections often show semicircular arcs coming all the way up to the surface. These arcs warn that something is wrong. The arcs could result from multiples, statics, or unexplained impulsive noise. In any case, they could be partially suppressed without touching primaries.

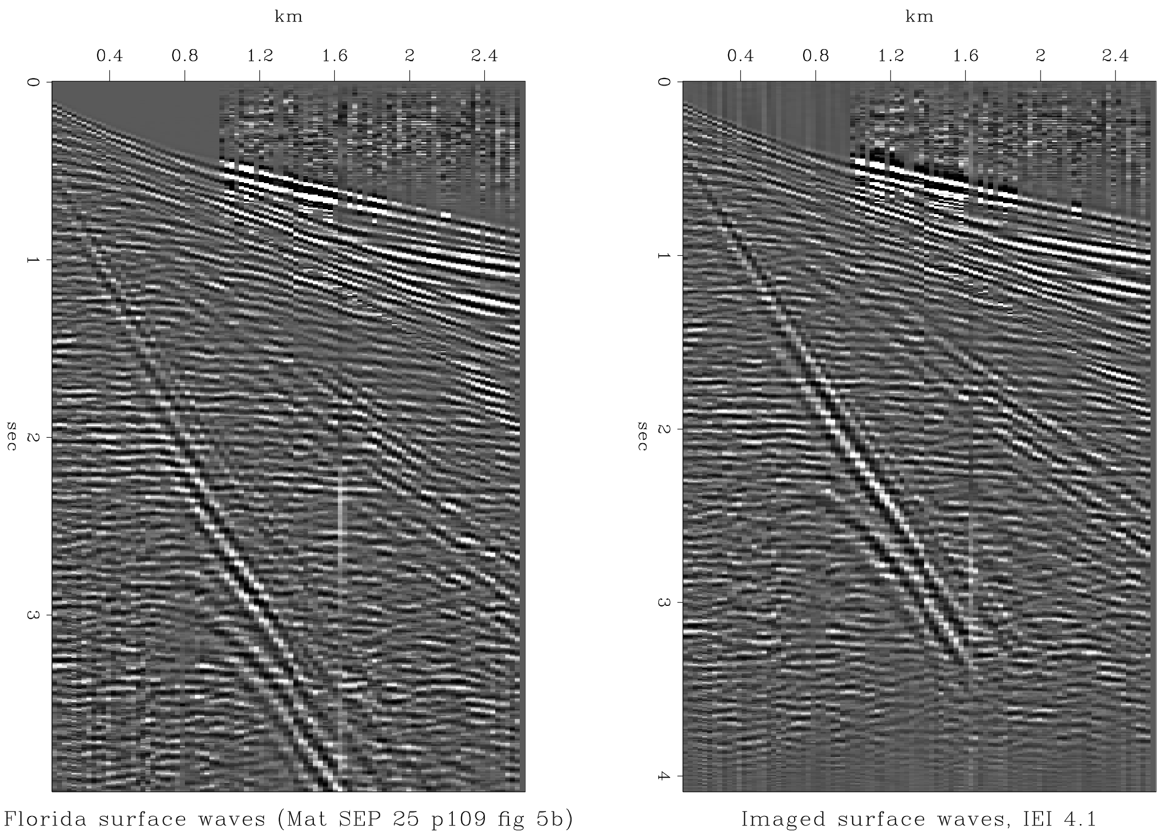


Figure 8.4: Dispersive surface wave (left), with the frequency dispersion backed out (right). Bottom shows two arrivals, the direct, straight-line arrival, and a hyperbola flank. The hyperbola represents sidescatter that must come from some object on the earth's surface off to the side of the survey line. (Conoco, Sword) $\boxed{\text{dspr-sword}}$ [ER]

8.1.7 Zapping multiples in dip space

Think of the migration of a common-depth-point stack as downward continuation in (ω, k_x, z)-space. Ordinarily, velocity increases with depth. As the downward continuation proceeds, the velocity cutoff along the evanescent line bites out more and more area from the (ω, k_x)-space. Energy beyond this cutoff does not fit the primary wave-propagation model, and it should be suppressed as soon as it is encountered. Such noise suppression can lead to a large drop in total power at late times.

8.1.8 Mixed appearance of dip-filtered data

An objection often raised against dip filtering is that it can give data a *mixed* appearance. *Mixed* means that adjacent channels appear to have been averaged and that they are no longer independent. This is indeed an effect of dip filtering, and it is inevitable at late times since the horizontal resolving power of reflection data decreases with time. There are two reasons for decreasing lateral resolution. First, dissipation causes high frequencies to disappear. Second, ray bending causes the angular aperture to decrease for deeper sources. It is unrealistic to ignore this fundamental limitation and imagine that adjacent channels should have an appearance of independence. If a mixed appearance is to be avoided for display purposes, then I advocate removing the low-velocity, coherent, signal-generated noise and replacing it by low-velocity, incoherent, Gaussian, random noise. Many plotters lose dynamic range at close trace spacing, and random noise can tend to restore it.

8.1.9 Accentuating faults

It often happens that the location of oil is controlled by faulting. But the dominating effect of stratified reflectors may overwhelm the weak diffraction evidence of faulting. A cosmetic process could weaken the zero and small dips, accentuating dips in the range of 10° to 60°, and then suppress the wide angles and evanescent energy. As with frequency filtering, sharp cutoffs are not desirable because of the implied long (and in space, wide) impulse response.

8.1.10 Dip filtering

Dip filtering is conveniently incorporated into the wave extrapolation equations. Instead of initializing the Muir expansion with $ik_z = -i\omega r_0$ we use $ik_z = \epsilon - i\omega r_0$. (Recall from chapter 3 that r_0 is the cosine of an exactly fitting angle). For the 15° equation we have

$$i\, k_z^{(15)}\, v \quad = \quad -\, i\, \omega \;+\; \frac{v^2\, k_x^2}{\epsilon \;-\; i\, \omega\, (r_0 \;+\; 1)} \tag{8.6}$$

For the 45° equation we have

$$i\, k_z^{(45)}\, v \quad = \quad -\, i\, \omega \;+\; \frac{v^2\, k_x^2}{-\, i\, \omega\, 2 \;+\; \frac{v^2\, k_x^2}{\epsilon \,-\, i\, \omega\, (r_0 \,+\, 1)}} \tag{8.7}$$

Figures 1.5 and 1.6 show hyperbolas of diffraction for the 15° and 45° equations with and without the dip filtering parameter ϵ.

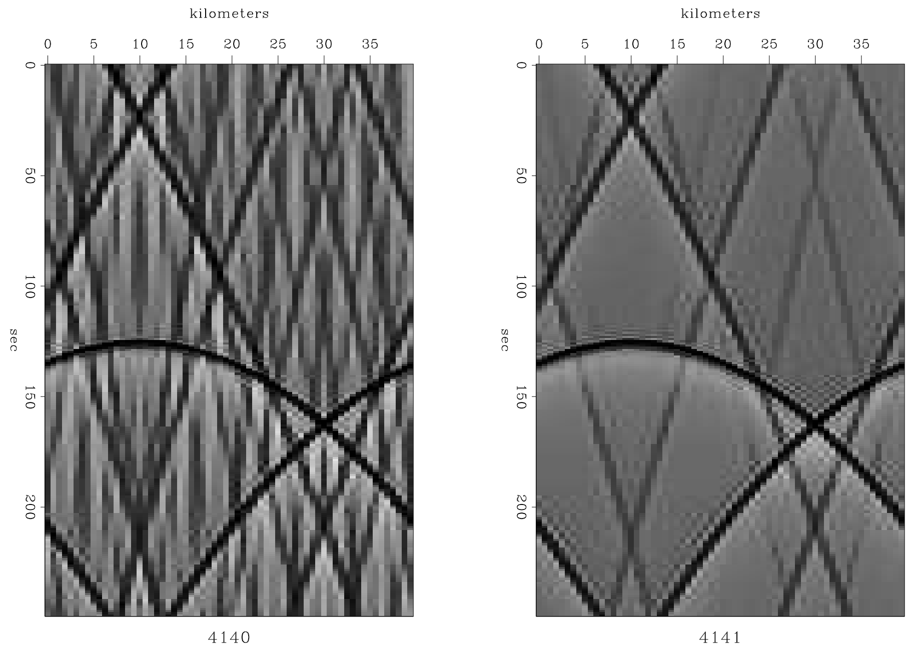


Figure 8.5: Diffraction hyperbolas of the 15° equation without dip filtering (left), and with dip filtering (right). dspr-hyp15 [ER]

8.1.11 Gain control does dip filtering too

Echoes arriving late are weaker than echoes arriving early. Thus data is ordinarily scaled for plotting using some time-variable scale. Should migration be done before or after this scaling? The results will differ in an interesting way. The top part of the hyperbola has flat dip, whereas the asymptotes, which come later, have steep dip. So, amplification of late information coincidentally amplifies the steep dips. I think the main effect of choosing to do migration before or after scaling is selection of the dip spectrum in the final display. A pedantically correct approach is to migrate first and scale second, but the result will be weaker in dip and fault information than the answer obtained by scaling first and migrating second. A side benefit of the latter method is that you can save computer memory by storing scaled values as short integers. I used 16 bit integer storage in my pioneering work. Computations and local storage used 32 bit floating point arithmetic. I see little justification for 32

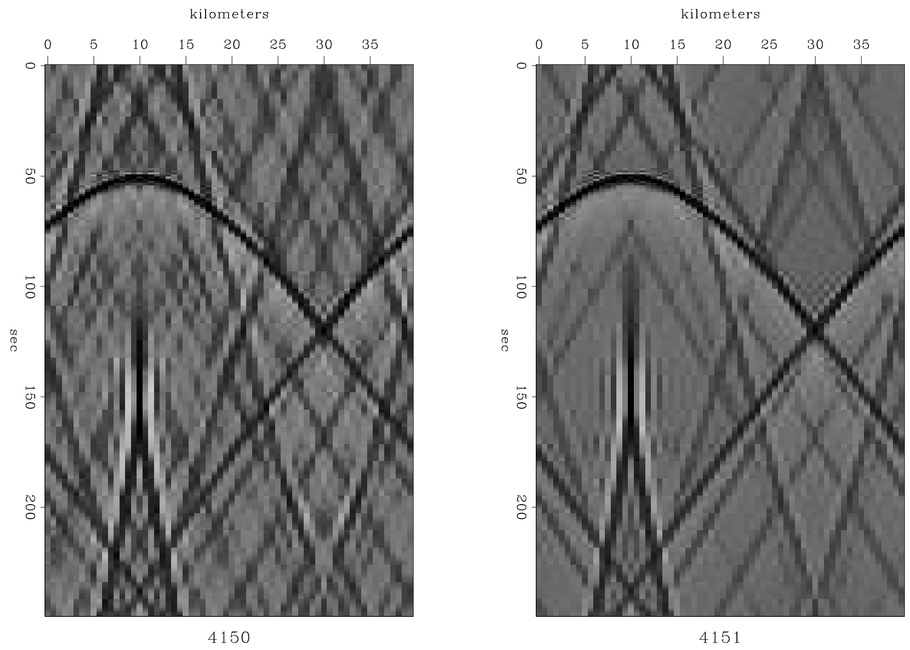


Figure 8.6: Diffraction hyperbolas of the 45° equation without dip filtering (left), and with dip filtering (right). dspr-hyp45 [ER]

bit storage generally used today. We can't interpolate between channels to 4 bits of accuracy.

8.1.12 Rejection by incoherence or rejection by filtering?

It is a pitfall to judge a supposed noncosmetic process by a cosmetic effect. I once got caught. The process was migration before stack. The feature that was deemed desirable was the relative strength of the steepest clear event on the record, a fault-plane reflection. But even gain control can affect dip spectra! I hoped the process was working by correctly eliminating some of the rejection of steep dips by CDP stack. Perhaps it was, but how could I know whether this was happening or whether dips were being accidentally enhanced by spatial filtering?

8.1.13 Spatial scaling before migration

Scaling on the time axis before migration can be advantageous. What about scaling on the space axis? The traditional methods of scaling that are called *automatic gain control* (AGC) deduce a scaling divisor by smoothing the data envelope (or its square or its absolute value) over some window. Such scaling can vary rapidly from trace to trace, so concern is justified that diffractions might be caused by lateral jumps in the scaling function. On the other hand, there might be good reasons for the scale to jump rapidly from trace to trace. The shots and geophones used to collect land data normally have variable strength and coupling, and these problems affect the entire trace.

A model must be found that respects both physics and statistics. I suggest allowing for gain that is slowly time-variable and shots and geophones of arbitrarily variable strength, but I also prefer to regard an impulse as evidence that the earth really can focus. For example, data processing with this model can be implemented by smoothing the scaling *envelope* with the filter

$$1 \quad - \quad \frac{\omega^2}{\alpha \ + \ \omega^2} \, \frac{k_x^2}{\beta \ + \ k_x^2} \tag{8.8}$$

Filter cutoff parameters are α and β. When the scaling envelope has been smoothed with this filter, it no longer varies rapidly with both x and t, although it can vary with either one or the other. This filter (1.8) can be economically implemented using the tridiagonal algorithm.

8.1.14 Exponential scaling

Exponential scaling functions have some ideal mathematical properties. Take the Z-transform of a time function a_t:

$$A(Z) \quad = \quad a_0 \ + \ a_1 \, Z \ + \ a_2 \, Z^2 \ + \ \ldots \tag{8.9}$$

The exponentially gained time function is defined by

$$\uparrow A(Z) \quad = \quad a_0 \ + \ a_1 \, e^{\alpha} \, Z \ + \ a_2 \, e^{2\alpha} \, Z^2 \ + \ \ldots \tag{8.10}$$

The symbol $\uparrow$ denotes exponential gain. Mathematically, $\uparrow$ means that Z is replaced by $e^{\alpha} Z$. Polynomial multiplication amounts to convolution of the coefficients:

$$C(Z) \quad = \quad A(Z) \; B(Z) \tag{8.11}$$

By direct substitution,

$$\uparrow C \quad = \quad (\uparrow A) \; (\uparrow B) \tag{8.12}$$

This means that exponential gain can be done either before or after convolution. You may recall from Fourier transform theory that multiplication of a time function by a decaying exponential $\exp(-\alpha t)$ is the equivalent of replacing $-i\omega$ by $-i\omega + \alpha$ in the transform domain.

Specialize the downward-continuation operator $\exp(i k_z z)$ to some fixed z and some fixed k_x. The operator has become a function of ω that may be expressed in the time domain as a filter a_t. Hyperbola flanks move *upward* on migration. So the filter is *anticausal*. This is denoted by

$$A(Z) \quad = \quad a_0 \quad + \quad a_{-1} \, \frac{1}{Z} \quad + \quad a_{-2} \, \frac{1}{Z^2} \quad + \quad \ldots \tag{8.13}$$

The large negative powers of Z are associated with the hyperbola flanks. Exponentially boosting the coefficients of positive powers of Z is associated with diminishing negative powers—so $\uparrow A$ is A with a weakened tail—and tends to attenuate flanks rather than move them. Thus $\uparrow A$ may be described as viscous.

From a purely physical point of view cosmetic functions like gain control and dip filtering should be done after processing, say, $\uparrow (AB)$. But $\uparrow (AB)$ is equivalent to $(\uparrow A)(\uparrow B)$, and the latter operation amounts to using a viscous operator on exponentially gained data. In practice, it is common to forget the viscosity and create $A(\uparrow B)$. Perhaps this means that dipping events carry more information than flat ones.

8.1.15 The substitution operator

The $\uparrow$ operator has been defined as the substitution $Z \rightarrow Z\,e^{\alpha}$. The main property of this operator is that if $C = AB$, then $C{=}(A)(B)$. This property would be shared by any algebraic substitution for Z, not just the one for exponential gain. Another simple substitution can be used to achieve time-axis stretching or compression. For example, replacing Z by Z^2 stretches the time axis by two. Yet another substitution, which has a deeper meaning than either of the previous two, is the substitution of the constant Q dissipation operator $(-i\omega)^{\gamma}$. In summary:

Substitutions for Z-Transform Variable Z [all preserve $C(Z) = A(Z)B(Z)$]	
Exponential growth	$Z \rightarrow Z_e^{\alpha}$ $(i\omega \rightarrow i\omega + \alpha)$
Time expansion $(\alpha > 1)$	$Z \rightarrow Z^{\alpha}$
(Inverse) Constant Q dissipation	$-i\omega \rightarrow (-i\omega)^{\gamma}$

EXERCISES:

1 Use a table of integrals to show that a seismic source with spectrum $|\omega|^{\beta}$ implies a divergence correction $t^{2+\beta}$.

2 Assuming that t^2 is a suitable divergence correction for field profiles, what divergence correction should be applied to CDP stacks?

3 How is the t^2 correction altered for water of travel time depth t_0? Assume the Q of water is infinite.

4 Consider a source spectrum $e^{-\beta|\omega|}$. How is the t^2 correction altered?

5 The spectrum in Figure 1.3 shows high frequencies smoother than low frequencies. Explain.

6 Propose some criteria that can be used in the selection of the cutoff parameters α and β for the filter (1.8).

8.2 ANISOTROPY DISPERSION IN MIGRATION

Two distinct types of errors are made in wave migration. Of greater practical importance is *frequency dispersion,* which occurs when different frequencies propagate at different speeds. This may be reduced by improving the accuracy of finite-difference approximations to differentials. Its cure is refinement of the differencing mesh.

Of secondary importance, and the subject of this section, is *anisotropy dispersion.* Anisotropic wave propagation is waves going different directions with different speeds. In principle, anisotropic dispersion is remedied by the Muir square-root expansion. In practice, the expansion is generally truncated at either the 15° or 45° term, creating anisotropy error in data processing. The reasons often given for truncating the series and causing the error are (1) the cost of processing and (2) the larger size of other errors in the overall data collection and processing activity. Anisotropy error should be studied in order to (1) recognize the problem when it occurs and (2) understand the basic trade-off between cost and accuracy.

Anisotropy is often associated with the propagation of light in crystals. In reflection seismology, anisotropy is occasionally invoked to explain small discrepancies between borehole velocity measurements (vertical propagation) and velocity determined by normal moveout (horizontal propagation). These fundamental, physical anisotropies and the subject of this section, anisotropy in data processing, share a common mathematical and conceptual basis.

8.2.1 Rays not perpendicular to fronts

Anisotropy means that waves propagating in different directions propagate at different speeds. Anisotropy does *not* mean that velocity is a function of spatial location, and thus anisotropy does not cause rays to bend. The peculiar thing about anisotropy is that rays are not perpendicular to wavefronts. Figure 1.7 illustrates this idea. The diagram on the left shows spherical wavefronts emanating from a point source at the origin. This is the usual, isotropic case. The diagram on the right shows the nonspherical wavefronts of the 15° migration equation. Note that near the z-axis they

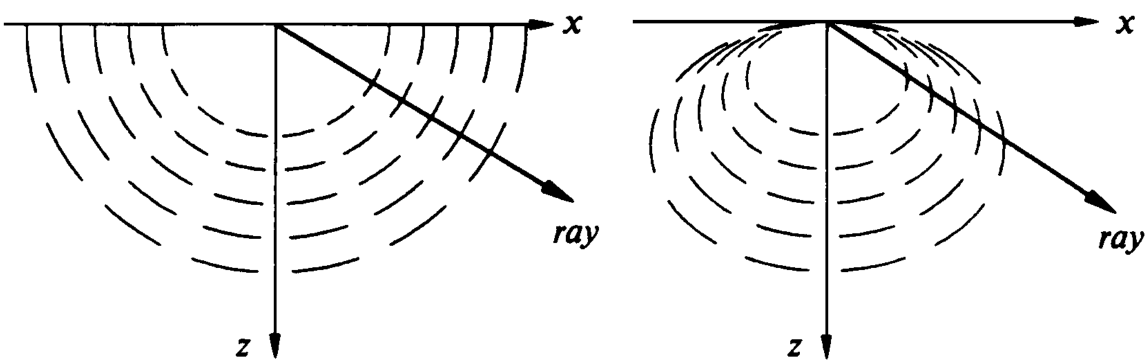


Figure 8.7: Wavefronts in an isotropic medium (left) and an anisotropic medium (right). Note that on the right, the rays are not perpendicular to the wavefronts. (Rothman) dspr-aniso [NR]

are nearly spherical, but further away they do a poor job of matching a sphere with its center at the origin.

The ideal wavefront from a Huygens secondary source is a semicircle. The secondary source that results from the 15° extrapolation equation is an ellipse. The secondary source that results from the 45° extrapolation equation is an interesting, heartlike shape. These are drawn in Figure 1.8. In practice, the top parts of the

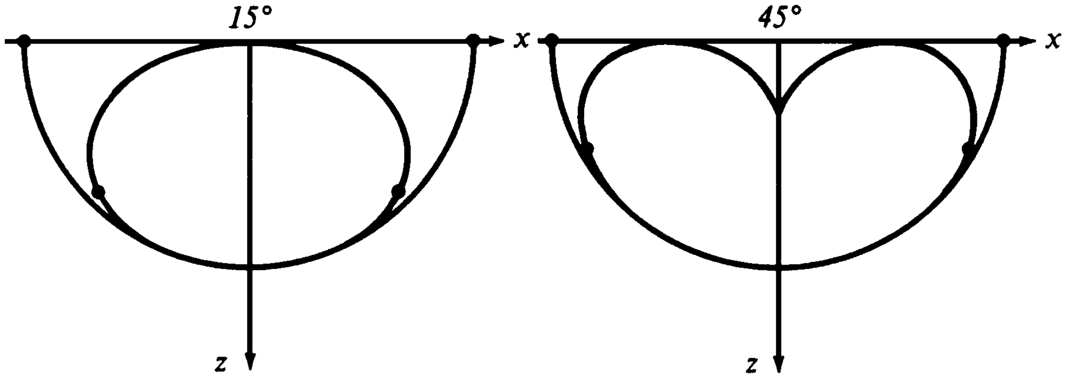


Figure 8.8: Wavefronts of 15° (left) and 45° (right) extrapolation equations, inscribed within the exact semicircle. Waves with $\sin\theta = vk_x/\omega = \pm 1$ are marked with small dots. Evanescent energy lies beyond the dots. (Rothman) dspr-anisoxz [NR]

ellipse and the heart are rarely observed because they are in the evanescent zone, and the x-axis is seldom refined enough for them to be below the aliasing frequency. The center of the heart is sometimes seen in the (x,t)-plane when the 45° program is used. It is shown by a line drawing in Figure 1.9 and shown using a 45° diffraction program in Figure 1.16.

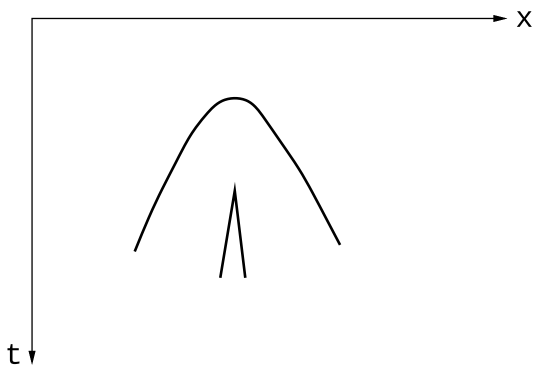


Figure 8.9: 45° heart theory. The cusp arises in the evanescent region. (Rothman) dspr-anisoxt [NR]

8.2.2 Wavefront direction and energy velocity

In ordinary wave propagation, energy propagates perpendicular to the wavefront. When there is anisotropy dispersion, the angle won't be perpendicular.

The apparent horizontal velocity seen along the earth's surface is dx/dt. The apparent velocity along a vertical, e.g., as seen in a borehole, is dz/dt. By geometry, both of these apparent speeds exceed the wave speed. The vector perpendicular to the wavefront with a magnitude inverse to the velocity is called the slowness vector:

$$\text{slowness vector} \quad = \quad \left(\frac{dt}{dx} \; , \; \frac{dt}{dz} \right) \tag{8.14}$$

The phase velocity vector is defined to go in the direction of the slowness vector, but have the speed of the wavefront normal. More precisely, the phase velocity vector is the slowness vector divided by its squared magnitude:

$$\text{phase velocity} \quad = \quad \frac{\left(\frac{dt}{dx} \; , \; \frac{dt}{dz} \right)}{\left(\frac{dt}{dx} \right)^2 \; + \; \left(\frac{dt}{dz} \right)^2} \tag{8.15}$$

For a disturbance of sinusoidal form, namely, $\exp(i\phi) = \exp(-i\omega t + ik_x x + ik_z z)$, the phase ϕ may be set equal to a constant:

$$0 \quad = \quad d\phi \quad = \quad -\,\omega\, dt \; + \; k_x\, dx \; + \; k_z\, dz \tag{8.16}$$

Thus, in Fourier space the slowness vector is

$$\text{slowness vector} \quad = \quad \left(\frac{k_x}{\omega} \; , \; \frac{k_z}{\omega} \right) \tag{8.17}$$

The direction of energy propagation is somewhat more difficult to derive, but it comes from the so-called group velocity vector:

$$\text{group velocity} \quad = \quad \left(\frac{\partial}{\partial k_x} \; , \; \frac{\partial}{\partial k_z} \right) \; \omega(k_x, k_z) \tag{8.18}$$

For the scalar wave equation $\omega^2/v^2 = k_x^2 + k_z^2$, the group velocity vector and the phase velocity vector turn out to be the same, as can be verified by differentiation and substitution. The most familiar type of dispersion is frequency dispersion, i.e. different frequencies travel at different speeds. Later in this section it will be shown that the familiar (15°, 45°, etc.) extrapolation equations do not exhibit frequency dispersion. That is, as functions of ω and angle k_x/ω, the velocities in these equations do not depend on ω. In other words, the elliptical and heart shapes in Figure 1.8 are not frequency-dependent.

An interesting aspect of anisotropy dispersion is that energy appears to be going in one direction when it is really going in another. An exaggerated instance of this occurs when the group velocity has a downward component and the phase velocity has an upward component. Figure 1.10, depicting the dispersion relation of the 45° extrapolation equation, shows an example. A slowness vector, which is in the direction of the wavefront normal, has been selected by drawing an arrow from the origin to the dispersion curve. The corresponding direction of group velocity may now be determined graphically by noting that group velocity is defined by the gradient operator in equation (1.18). Think of ω as the height of a hill from which k_z points south and k_x points east. Then the dispersion relation is a contour of constant altitude. Different numerical values of frequency result from drawing Figure 1.10 to different scales. The group velocity, in the direction of the gradient, is perpendicular to the contours of constant ω.

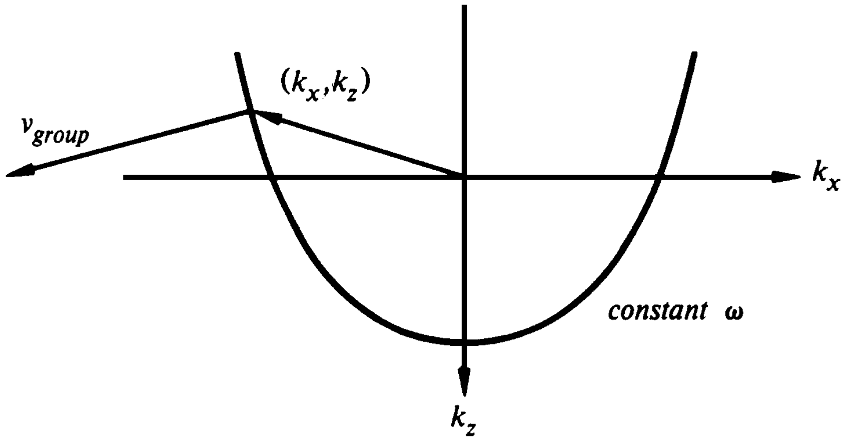


Figure 8.10: Dispersion relation for downgoing extrapolation equation showing group velocity vector and slowness vector. (Rothman) dspr-group15 [NR]

The anisotropy-dispersion phenomenon can be most clearly recognized in a movie, although it can be understood on a single frame, as in Figure 1.11. The line drawing interprets energy flow from the top, through the prism, reflecting at the 45° angle, reflecting from the side of the frame, and finally entering an area of the figure that is sufficiently large and uncluttered for the phase fronts to be recognizable as energy apparently propagating upward but really propagating downward.

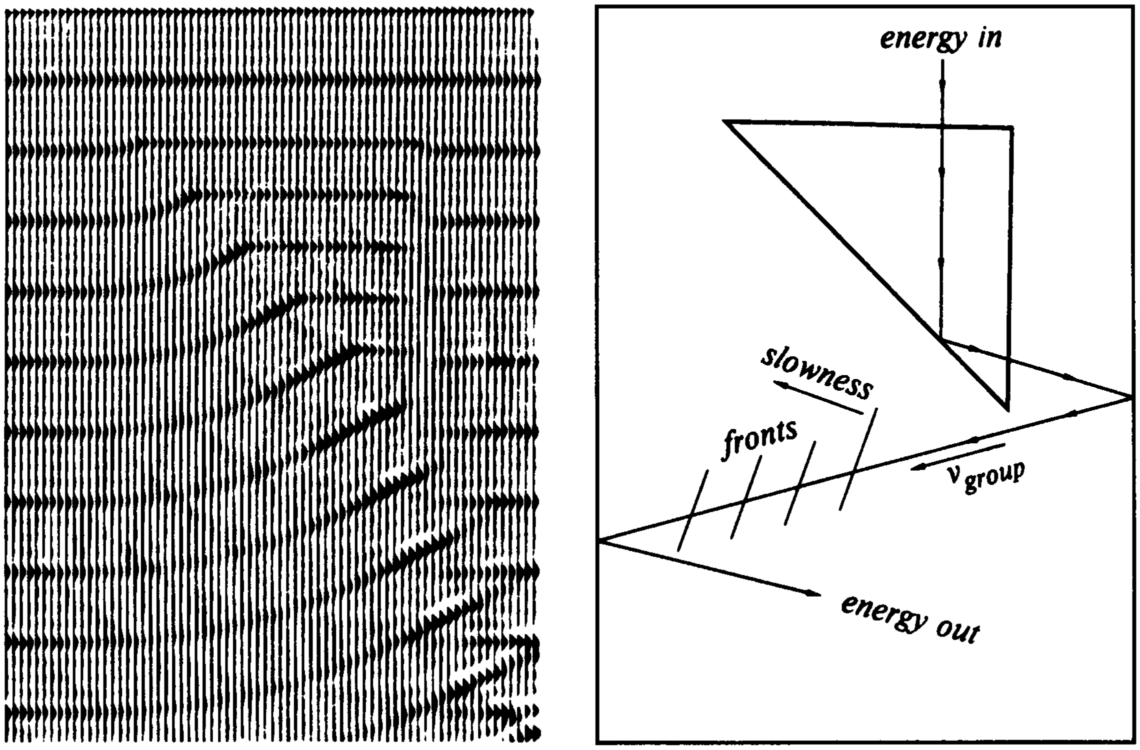


Figure 8.11: Plane waves of four different frequencies propagating through a right, 45° prism. Left is the wavefield. Right is a ray interpretation that illustrates different directions of energy and wavefront normal. (Estevez) dspr-prism [NR]

That neither energy nor information can propagate upward in Figure 1.11 should be obvious when you consider the program that calculates the wavefield. The program does not have the entire frame in memory; it produces one horizontal strip at a time from the strip just above. Thus the movie's phase fronts, which appear to be moving upward, seem curious. Theoretically, wave extrapolation using the 45° equation is not expected to handle angles to 90°. Yet the example in Figure 1.11 shows that these extreme cases are indeed handled, although in a somewhat perverted way.

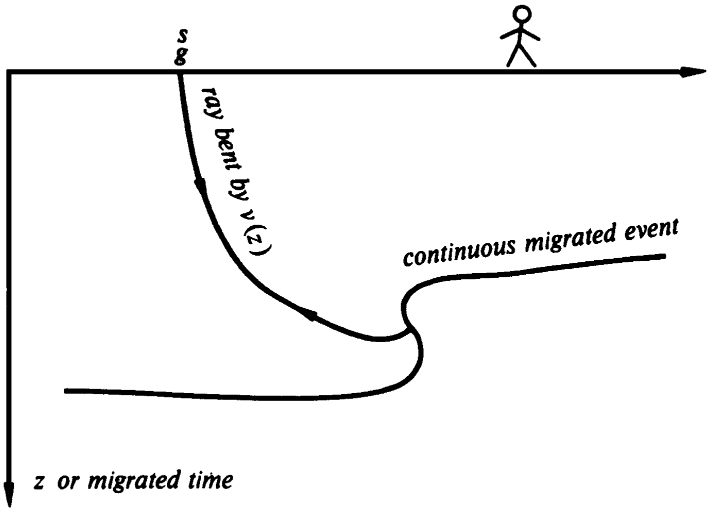


Figure 8.12: Ray reflected from the underside of an overthrust. dspr-overthrust [NR]

I once saw a similar circumstance on reflection seismic data from a geologically overthrusted area. The data could not be made available to me at the time, and by now is probably long lost in the owner's files, so I can only offer the line drawing in Figure 1.12, which is from memory. The increasing velocity with depth causes the ray to bend upward and reflect from the underside of the overthrust. To see what is happening in the wave equation, it is helpful to draw the dispersion curve at two different velocities, as in Figure 1.13.

Downward continuation of a bit of energy with some particular stepout $dt/dx = k_x/\omega$ begins at an ordinary angle on the near-surface, slow-velocity dispersion curve. But as deeper velocity material is encountered, that same stepout implies a negative phase velocity. Although the thrust angle is unlikely to be quantitatively correct, the general picture is appropriate. It is like Figure 1.11.

8.2.3 Analyzing errors of migration

A dipping reflector that is flat and regular can be analyzed in its entirety using the phase velocity concept. The group velocity concept is required only when more than

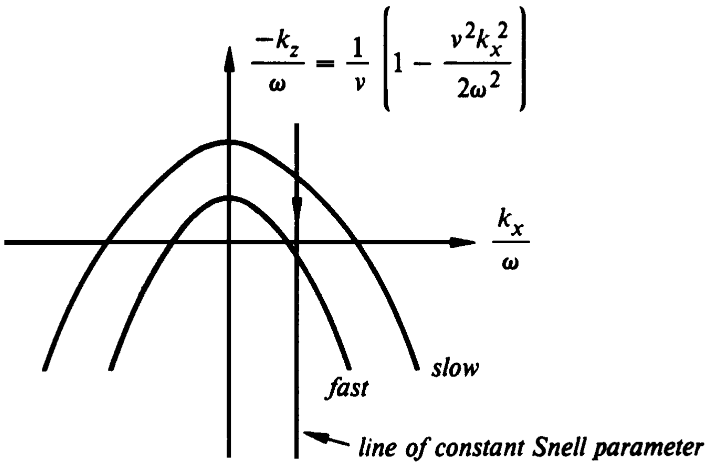
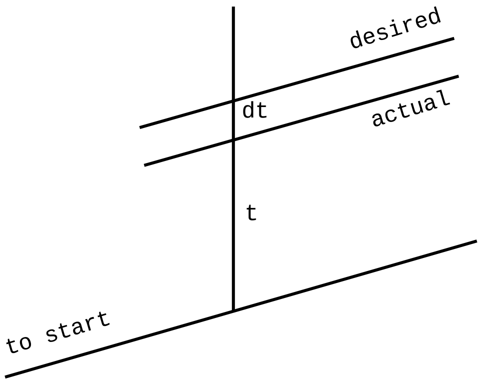


Figure 8.13: Dispersion curve at two different velocities, v_{fast} and v_{slow}. dspr-disper2z [NR]

one angle is simultaneously present. This simultaneity occurs with the point scatterer response. It also occurs when there is variable reflection amplitude along a dipping bed. The group velocity is needed because representation of either a curved event or an amplitude anomaly requires a range of plane-wave angles. Analogously, in time-series analysis the Fourier representation of an amplitude-modulated sinusoid requires a bandwidth of sinusoids.

Figure 1.14 depicts a smooth, flat, dipping bed that has been undermigrated because the $\hat{k}_z$ defined by some rational square-root approximation or some numerical approximation did not match the correct square-root value of k_z.

Figure 8.14: Undermigrated dipping reflector. dspr-undermig [NR]



The error in Figure 1.14 is entirely a time-shift error. Since the reflection coefficient is constant along the reflector, no lateral shift error can be recognized. The time error may be theoretically determined by

$$\frac{dt}{t} \quad \approx \quad \frac{dz}{z} \quad \approx \quad \frac{\hat{k}_z \ - \ k_z}{k_z} \tag{8.19}$$

For the so-called 15° equation, it turns out that about a half-percent phase error is made at 25°.

Next, the error in the collapse of a hyperbola will be determined. Figure 1.70 depicts the downward continuation of a hyperbola. For clarity, the downward continuation was not taken all the way to the focus. Select a ray of some Snell's parameter $p = dt/dx$ by choosing some slope p. Imagine a tangent line segment of slope p to each of the hyperboloids. If there were a little amplitude anomaly where the slope is p, you would be able to identify it on each of the hyperboloids.

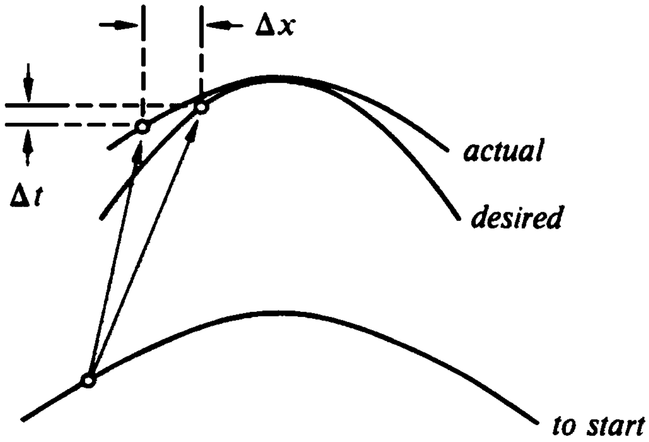


Figure 8.15: Error of hyperbola collapse. Note that the actual curve is above the desired curve, but the actual point is below the desired point. dspr-errcollapse [NR]

In Figure 1.7 the amount of time moved is too little; likewise, the lateral distance moved is too small. In practice, errors of the 15° equation with $r_{0=1}$ are sometimes compensated by an increase in either z or v of about 6%. The amounts of the errors may be calculated from

$$\frac{\Delta t}{t} = \frac{\frac{\partial}{\partial\omega}\,(\hat{k}_z - k_z)}{\frac{\partial}{\partial\omega}\,k_z} \tag{8.20}$$

$$\frac{\Delta x}{x} = \frac{\frac{\partial}{\partial k_x}\,(\hat{k}_z - k_z)}{\frac{\partial}{\partial k_x}\,k_z} \tag{8.21}$$

where k_z is taken to be a function of ω and k_x. It turns out that for the 15° equation, about a half-percent group velocity error occurs at 20°. Thus the group velocity error is generally worse than the phase velocity error.

8.2.4 Derivation of group velocity equation

An impulse function at the origin in (x, z)-space is a superposition of Fourier components:

$$\int\int e^{+ik_x x\,+\,ik_z z}\,dk_x\,dk_z \tag{8.22}$$

Physics (and perhaps numerical analysis) leads to a dispersion relation that is a functional relation between ω, k_x, and k_z, say, $\omega(k_x, k_z)$. The most common example of such a dispersion relation is the scalar wave equation $\omega^2 = v^2\,(k_x^2 + k_z^2)$. The solution to the equation is

$$e^{-i\omega t + i k_x x + i k_z z} \tag{8.23}$$

Integrating (1.23) over (k_x, k_z) produces a monochromatic time function that at $t = 0$ is an impulse at $(x, z) = (0, 0)$. This expression at some very large time t is

$$\int\int e^{-it[\omega(k_x,k_z) - k_x x/t - k_z z/t]}\, dk_x\, dk_z \tag{8.24}$$

At t very large, the integrand is a very rapidly oscillating function of unit magnitude. Thus the integral will be nearly zero unless the quantity in square brackets is found to be nearly independent of k_x and k_z for some sizable area in (k_x, k_z)-space. Such a flat spot can be found in the same way that the maximum or minimum of any two dimensional function is found, by setting derivatives equal to zero. This analytical approach is known as the stationary phase method. It gives

$$0 = \frac{\partial}{\partial k_x}[\quad] = \frac{\partial\omega}{\partial k_x} - \frac{x}{t} \tag{8.25}$$

$$0 = \frac{\partial}{\partial k_z}[\quad] = \frac{\partial\omega}{\partial k_z} - \frac{z}{t} \tag{8.26}$$

So, in conclusion, at time t the disturbances will be located at

$$(x, z) = t\left(\frac{\partial\omega}{\partial k_x}, \frac{\partial\omega}{\partial k_z}\right) \tag{8.27}$$

which justifies the definition of group velocity.

Now let us see how the left side of Figure 1.8 was calculated. The 15° dispersion relation was solved for ω and inserted into (1.27). The resulting (x, z) turned out to be a function of k_x/ω. Trying all possible values of k_x/ω gave the curve.

8.2.5 Derivation of energy migration equation

Energy migration in (x, t)-space is analyzed in a fashion similar to the way the group velocity was derived. Take depth to be large in the integral

$$\int\int e^{iz[k_z(\omega,k_x) - \omega t/z + k_x x/z]}\, d\omega\, dk_x \tag{8.28}$$

The result is that the energy goes to

$$(x, t) = z\left(-\frac{\partial k_z}{\partial k_x}, \frac{\partial k_z}{\partial\omega}\right) \tag{8.29}$$

This justifies our previous assertion that (1.20) and (1.21) can be used to analyze energy propagation errors. Equation (1.29) was also used to calculate the curve in Figure 1.9. The validity of the stationary phase concept is confirmed by Figure 1.16, which was produced using inverse Fourier transformation.

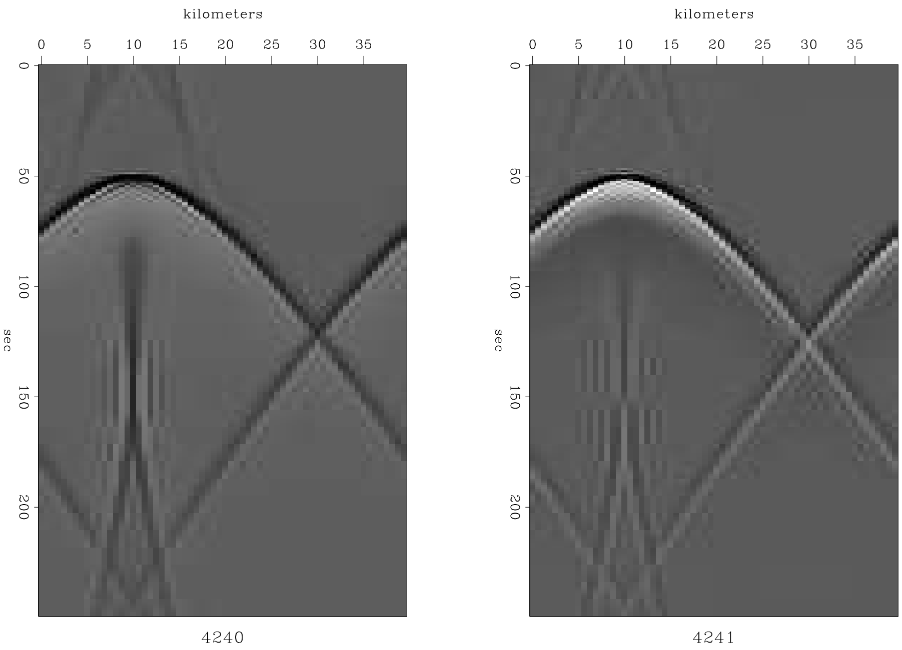


Figure 8.16: Impulse response of the 45° wave-extrapolation equation. The arrival before t_0 is a wraparound. dspr-aniso45 [ER]

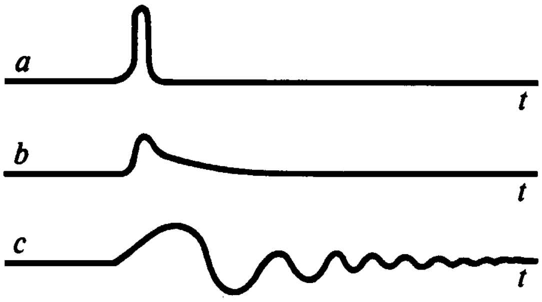
8.2.6 Extrapolation equations are not frequency-dispersive.

To prove that the familiar 15°, 45°, etc. wave extrapolators are not frequency-dispersive, recall from chapter 3 that the dispersion relations all have the form $k_z/\omega = f(k_x/\omega)$, where f is a semicircle approximation, say, 15° or 45°. No dispersion relation of this form can be frequency-dispersive. Performing the derivatives required by (1.29), you see that while the (x,t)-coordinates of a wavefront depend on the dip angle through the parameter vk_x/ω, they do not depend explicitly on ω. So any frequency dispersion observed in practice does not arise from a 15° or 45° approximation.

8.3 FREQUENCY DISPERSION IN WAVE-MIGRATION

Frequency dispersion results from different frequencies propagating at different speeds. The physical phenomenon of frequency dispersion is rarely heard in daily life, although many readers may have heard it while ice skating on lakes and rivers. Elastic waves caused by cracking ice propagate dispersively, causing pops to change into percussive notes. Frequency dispersion is generally observable on seismic waves that propagate along the earth's surface but frequency dispersion is hardly ever perceptible on internally reflected waves. In seismic data processing, frequency dispersion is a nuisance and an embarrassment to process designers. It arises mainly with the finite differencing method because differential operators and difference operators do not coincide at high frequencies. Frequency dispersion can always be suppressed by sampling more densely, and it is the job of the production analyst to see that this is done. Figure 1.17 depicts some dispersed pulses.

Figure 8.17: (a) A pulse. (b) A pulse slightly dispersed as by the physical dissipation of high frequencies. (c) A pulse with a substantial amount of frequency dispersion, such as could result from careless data processing. dspr-freqdisp [NR]



Frequency dispersion caused by data processing can be a useful warning that the data is in danger of being aliased. Frequency-domain methods do not depend on difference operators, so they have the advantage of not showing dispersion. The penalties that go along with this advantage are (1) limitation to constant material properties, (2) wraparound, and (3) the occurrence of spatial aliasing without the warning of dispersion.

Figure 1.18 shows an example of frequency dispersion in migrated data. At the top of the figure is a CDP stack. In the middle is the data after processing with no attempt to control frequency dispersion. The worst dispersion is near shot point 200 at 4 seconds. The bottom shows the data after reprocessing with greater attention to dispersion.

8.3.1 Spatial aliasing various space axes

Aliasing can occur on the axes of time, depth, geophone, shot, midpoint, offset, or crossline. In practice, aliasing is the worst on the horizontal space axes. Figure 1.19 provides an illustration. Looking at that figure, you get confused about whether the dip is to the left or right. Mathematical analysis has the same difficulty. The dispersion relation of the wave equation enables us to compute the vertical spatial frequency k_z from the temporal frequency ω, the velocity v, and the horizontal spatial frequency k_x using the semicircle relation $k_z(\omega, k_x) = \sqrt{\omega^2/v^2 - k_x^2}$. Sampling on the x-axis sets an upper limit on k_x equal to the Nyquist frequency $\pi/\Delta x$. Both frequency-domain methods and finite-difference methods treat higher frequencies as though they were folded back at the Nyquist frequency. Thus the semicircle dispersion relation is replicated above the Nyquist frequency, as shown in Figure 1.19.

The problem of spatial aliasing begins when two circles touch each other, as shown at 20 Hz in Figure 1.19. This occurs when a half-wavelength $v/2f$ equals the spatial sample interval Δx. The exploding-reflector model implies that the velocity to use is half the rock velocity. Thus the aliasing problem is avoided if $2f\Delta x < \frac{1}{2}\, v_{\text{rock}}$. For a rock velocity equal to 2 km/sec, the safe frequencies are listed in the table below:

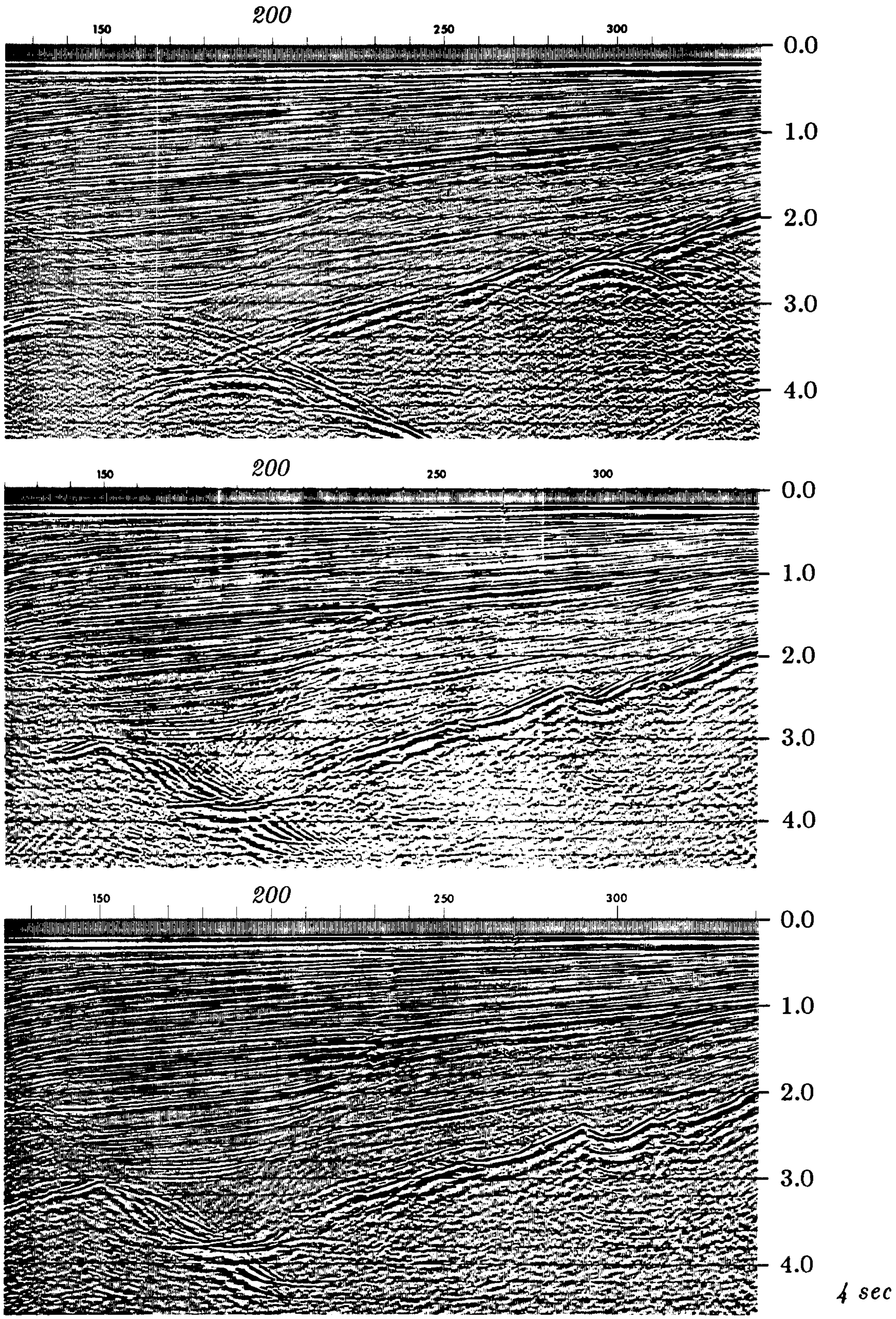


Figure 8.18: Conquering frequency dispersion. (Taner and Koehler, distributed by Seiscom Delta Inc.) $\boxed{\text{dspr-taner}}$ [NR]

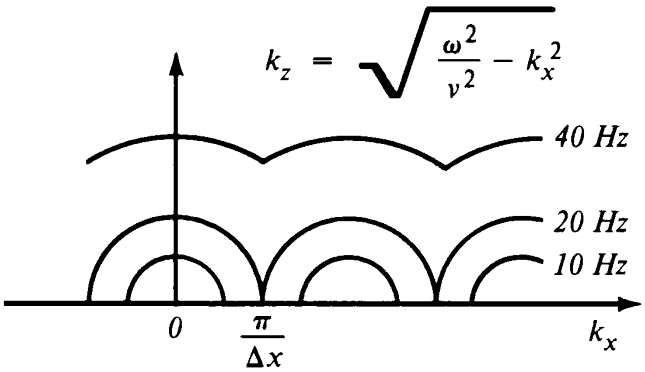


Figure 8.19: The effective dispersion relation of the wave equation when the horizontal axis is sampled. Frequencies are given for typical zero-offset migration. dspr-kxalias [NR]

	$\triangle x$	safe frequency
standard	25m	< 20Hz
reconnaissance	50m	< 10Hz
3-D cross line	100m	< 5Hz

Another view of the spatial aliasing problem is that steeply dipping waves are suppressed by the geophone group. (This disregards shot-space aliasing). From this standpoint the limit past which spatial aliasing begins should be thought of in terms of angles at which energy is missing from the data. Taking the ray angle to be 30° instead of 90° doubles horizontal wavelengths. Thus, for 30° and a rock velocity of 2 km/sec, to ensure safety from aliasing, frequencies should be in the ranges listed below:

	$\triangle x$	safe frequency
standard	25m	< 40Hz
reconnaissance	50m	< 20Hz
3-D cross line	100m	< 10Hz

Because data usually has good signal above 40 Hz, wide-angle processing is often frustrated by spatial aliasing.

The problem of spatial aliasing usually overshadows the difference between the 15° and the 90° equations. Aliased energy does not move between hyperbola flanks and the apex. Aliased energy tends to stay in place. This is illustrated on Figure 1.20 which shows a 90° hyperbola and a 15° hyperboloid from a finite difference equation. Overall, there is little difference. Look at the amplitude of the hyperbolic arrival. It is dropping off faster than predicted by spherical spreading and the obliquity function. This is because the dispersion curve semicircles overlap one another. There can be no angles of propagation beyond that which aliases x. Since waves can't go so steeply, they don't. The pulse doesn't spread properly.

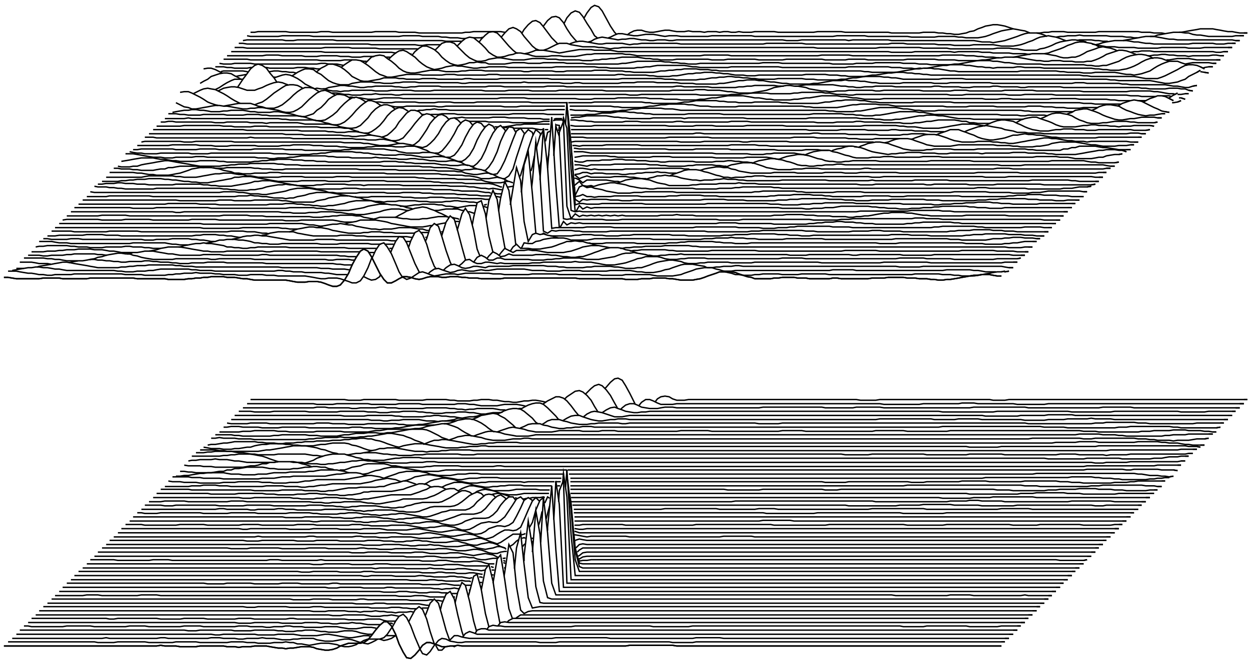


Figure 8.20: One second of synthetic hyperbolas with Δt = 4 ms, Δx = 25 meters, and velocity 2 km/sec. Fourier domain 90° hyperbola (top) and 15° finite difference hyperboloid (bottom). dspr-fdvrso [ER]

8.3.2 Second space derivatives

The defining equation for a second *difference* operator is

$$\frac{\delta^2}{\delta x^2} P \quad = \quad \frac{P(x + \Delta x) - 2P(x) + P(x - \Delta x)}{(\Delta x)^2} \tag{8.30}$$

The second *derivative* operator is defined by taking the limit

$$\frac{\partial^2}{\partial x^2} P \quad \underset{\lim \ \Delta x \to 0}{=} \quad \frac{\delta^2}{\delta x^2} P \tag{8.31}$$

Many different definitions can all go to the same limit as Δx goes to zero. The problem is to find an expression that is accurate when Δx is larger than zero and, on a practical level, is not too complicated. Our first objective is to see how the accuracy of equation (1.30) can be evaluated quantitatively. Second, we will look at an expression that is slightly more complicated than (1.30) but much more accurate.

The basic method of analysis we will use is Fourier transformation. Take the derivatives of the complex exponential $P = P_0 \ \exp\ (ikx)$ and look at any errors as functions of the spatial frequency k. For the second derivative,

$$\frac{\partial^2}{\partial x^2} P \quad = \quad - k^2 P \tag{8.32}$$

Define $\hat{k}$ by an expression analogous to the difference operator:

$$\frac{\delta^2}{\delta x^2} P \quad = \quad - \hat{k}^2 P \tag{8.33}$$

Ideally $\hat{k}$ would equal k. Inserting the complex exponential $P = P_0 \ \exp\ (ikx)$ into (1.30), we see that the definition (1.33) gives an expression for $\hat{k}$ in terms of k:

$$-\hat{k}^2 P \quad = \quad \frac{P_0}{\Delta x^2} \left[\, e^{ik(x+\Delta x)} \ - \ 2\, e^{ikx} + e^{ik(x-\Delta x)} \right] \tag{8.34}$$

$$-\,\frac{\delta^2}{\delta x^2} \quad = \quad \hat{k}^2 \quad = \quad \frac{2}{\Delta x^2}\, [1 \ - \ \cos(k \Delta x)] \tag{8.35}$$

It is a straightforward matter to make plots of $\hat{k}\Delta x$ versus $k\Delta x$ from (1.35). The half-angle trig formula allows us to take an analytic square root of (1.35), which is

$$\frac{\hat{k}\,\Delta x}{2} \quad = \quad \sin\ \frac{k\,\Delta x}{2} \tag{8.36}$$

Series expansion shows that for low frequencies $\hat{k}$ is a good approximation to k. At the Nyquist frequency, defined by $k\Delta x = \pi$, the approximation $\hat{k}\Delta x = 2$ is a poor approximation to π.

8.3.3 The 1/6 trick

Increased absolute accuracy may always be purchased by reducing Δx. Increased accuracy relative to the Nyquist frequency may be purchased at a cost of computer time and analytical clumsiness by adding higher-order terms, say,

$$\frac{\partial^2}{\partial x^2} \quad \approx \quad \frac{\delta^2}{\delta x^2} \ - \ \frac{\Delta x^2}{12}\ \frac{\delta^4}{\delta x^4} \ + \ \ \text{etc.} \tag{8.37}$$

As Δx tends to zero (1.37) tends to the basic definitions (1.30) and (1.31). Coefficients like the 1/12 in (1.37) may be determined by the Taylor-series method if great accuracy is desired at small k. Or somewhat different coefficients may be determined by curve-fitting techniques if accuracy is desired over some range of k. In practice (1.37) is hardly ever used, because there is a less obvious expression that offers much more accuracy at less cost! The idea is indicated by

$$\frac{\partial^2}{\partial x^2} \quad \approx \quad \frac{\frac{\delta^2}{\delta x^2}}{1 \ + \ b\,\Delta x^2 \frac{\delta^2}{\delta x^2}} \tag{8.38}$$

where b is an adjustable constant. The accuracy of (1.38) may be numerically evaluated by substituting from (1.35) to get

$$\left(\frac{\hat{k}\,\Delta x}{2} \right)^2 \quad = \quad \frac{\sin^2\ \frac{k\,\Delta x}{2}}{1 \ - \ b\,4\ \sin^2\ \frac{k\,\Delta x}{2}} \tag{8.39}$$

The square root of (1.39) is plotted in Figure 1.21 for a value of $b = 1/6$.

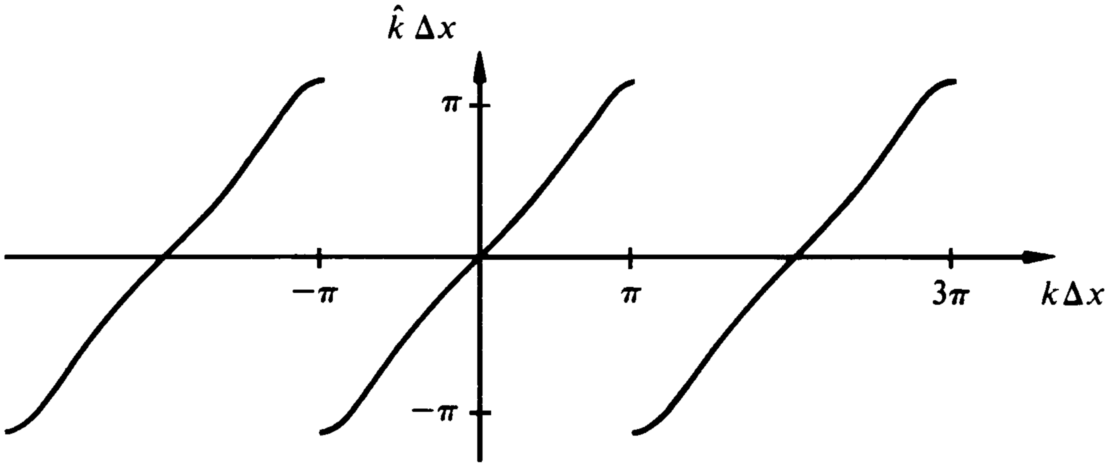


Figure 8.21: Accuracy of the second-derivative representation equation (1.38) and (1.39) (for $b = 1/6$) as a function of spatial wavenumber. The sign of the square root of (1.39) was chosen to agree with k in the range $-\pi$ to π and to be periodic outside the range. (Hale) dspr-sixth [NR]

Taking b in (1.38) and (1.39) to be 1/12, then (1.38), (1.39) and (1.37) would agree to second order in Δx. The 1/12 comes from series expansion, but the 1/6 fits over a wider range and is a value in common use. Francis Muir has pointed out that the value $1/4 - 1/\pi^2 \approx 1/6.726$ gives an *exact* fit at the Nyquist frequency and an accurate fit over all lower frequencies! Few explorationists consider the remaining accuracy deficiency of (1.38) and (1.39) to be sufficient to warrant interpolation of field-recorded values. Figure 1.22 compares hyperbolas for various values of b. Observe in Figure 1.22 that the longest wavelengths travel at the same speed regardless of b. The time axis in Figure 1.22 is only 256 points long, whereas in practice it would be a thousand or more. So Figure 1.22 exaggerates the frequency dispersion attributable in practice to finite differencing the x-axis.

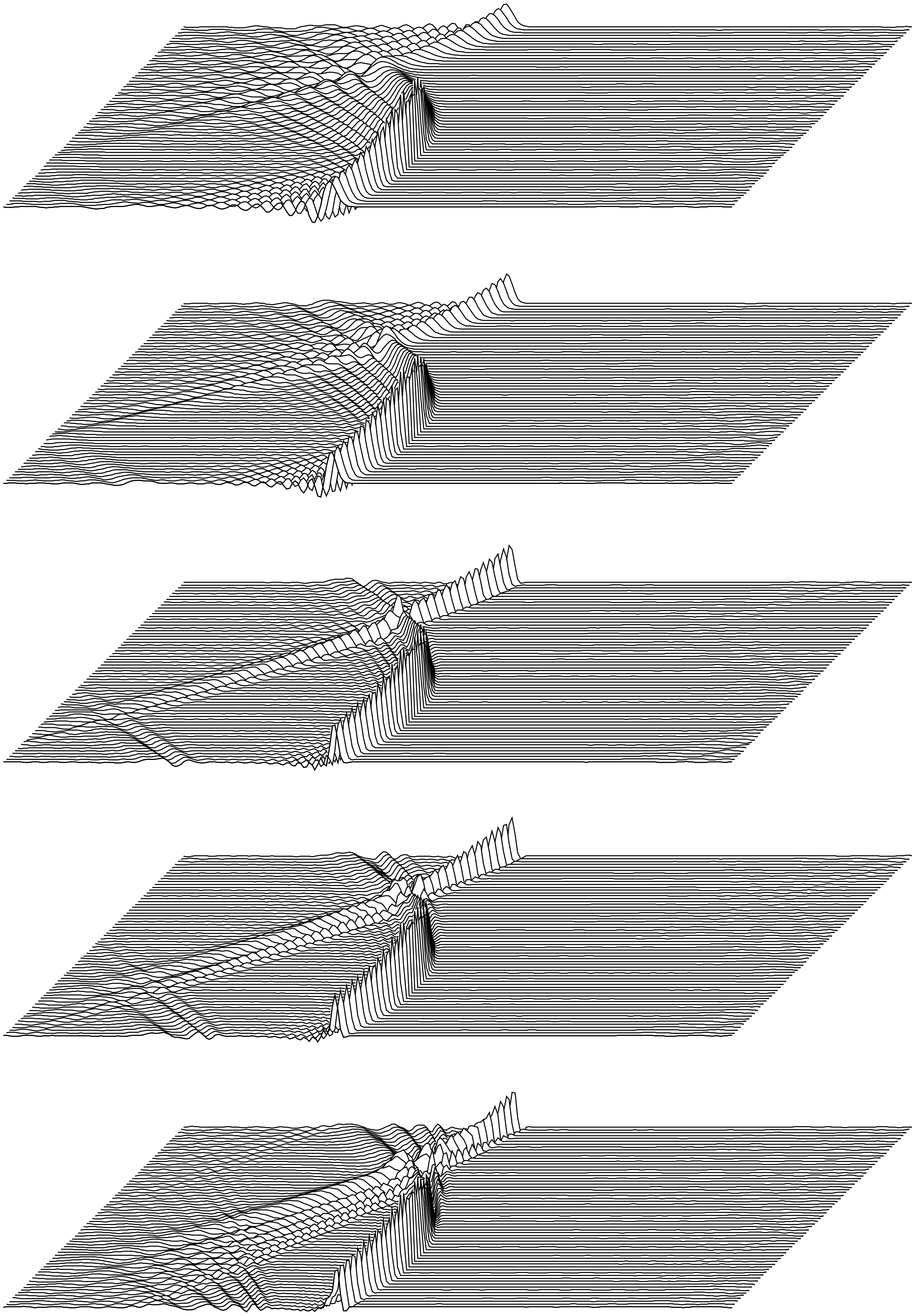


Figure 8.22: Hyperbolas for $b = 0, 1/12, 1/6.726, 1/6, 1/5$. $\boxed{\text{dspr-sixth5}}$ [ER]

Let us be sure it is clear how (1.38) and (1.39) are put into use. Take $b = 1/6$. The simplest prototype equation is the heat-flow equation:

$$\frac{\partial}{\partial t}\, q \quad = \quad \frac{\partial^2}{\partial x^2}\, q \quad \approx \quad \frac{\frac{\delta^2}{\delta x^2}}{1 \ + \ \frac{\Delta x^2}{6}\ \frac{\delta^2}{\delta x^2}}\ q \tag{8.40}$$

Multiply through the denominator:

$$\left(1 \ + \ \frac{\Delta x^2}{6}\ \frac{\delta^2}{\delta x^2}\right) \frac{\partial}{\partial t}\, q \quad \approx \quad \delta_{xx}\, q \tag{8.41}$$

8.3.4 Time and depth derivatives—the bilinear transform

You might be inclined to think a second derivative is a second derivative and that there is no mathematical reason to do time derivatives differently than space derivatives. This is not the case. A hint of disparity between t and x derivatives comes from boundary conditions. With time derivatives (and often with the depth z derivative) we must consider causality—which means the future is determined solely from the present and past. Appropriate boundary conditions on the time axis are initial conditions—the function (and perhaps some derivatives) is specified at *one* point, the initial point in time. For depth z that special point is the earth's surface at $z = 0$. But lateral space derivatives are different: they require boundary conditions at two widely separated points, usually at the left and right sides of the volume.

The differential equation

$$\frac{dq}{dz} \quad = \quad i\, k_z(\omega, k_x)\, q \tag{8.42}$$

is associated with the very definition of k_z. The analogous difference equation will define $\hat{k}_z$:

$$\frac{q_{z+\Delta z} \ - \ q_z}{\Delta z} \quad = \quad i\, \hat{k}_z\, \frac{q_{z+\Delta z} \ + \ q_z}{2} \tag{8.43}$$

Inserting the solution of (1.42) $q = q_0 \exp(ik_z z)$ into (1.43) gives us the relation between the desired k_z and the actual $\hat{k}_z$.

$$i\, \hat{k}_z\, \Delta z \quad = \quad 2\, \frac{e^{ik_z\, \Delta z} \ - \ 1}{e^{ik_z\, \Delta z} \ + \ 1} \quad = \quad 2\, \frac{e^{ik_z\, \Delta z/2} \ - \ e^{-ik_z\, \Delta z/2}}{e^{ik_z\, \Delta z/2} \ + \ e^{-ik_z\, \Delta z/2}} \tag{8.44}$$

This equation is known as the bilinear transform.

$$i\, \hat{k}_z\, \Delta z \quad = \quad 2\, i\, \frac{\sin\ k_z\, \Delta z/2}{\cos\ k_z\, \Delta z/2} \tag{8.45}$$

$$\frac{\hat{k}_z\, \Delta z}{2} \quad = \quad \tan\ \frac{k_z\, \Delta z}{2} \tag{8.46}$$

Equation (1.46) gives the accuracy of first derivatives obtained using the Crank-Nicolson method. Recall the migration differencing schemes in chapter 5. We did the time differencing in the same way that we did the depth differencing. So the same accuracy limitation must apply, namely,

$$\frac{\hat{\omega}\,\Delta t}{2} \quad = \quad \tan\frac{\omega\,\Delta t}{2} \tag{8.47}$$

Series expansion shows that $\hat{\omega}$ goes to ω as Δt goes to zero. Relative errors in ω at (4, 10, and 20) points per wavelength are (30%, 3%, and 1%). These errors are quite large, calling for either a choice of small Δt or a more accurate method than (1.46).

The bad news is that there does not seem to exist a representation of causal differentiation that is any more accurate than the Crank-Nicolson representation. There is nothing like the 1/6 trick. Thus the sample intervals of Δz and Δt must be reduced considerably from the Nyquist criterion. The practical picture may not be as bleak as the one I am painting. Many people are pleased with both the speed and accuracy of time-domain migrations at $\Delta t = 4$ milliseconds.

Stolt's classic paper [1978] besides introducing the fast Fourier transform migration method, points out that more accuracy can be achieved when the requirement of causality is dropped. Stolt shows how dropping causality at the known depth level while retaining it at the next level allows stable finite differencing. With the depth z-axis we are stuck with causal derivatives, although Fourier methods could be used for discrete layers. The depth axis is not so troublesome as the x- and t-axes, however, because it affects computer time only, not data storage.

Finite difference solutions don't just approximate the frequency—what they really do is to approximate $\exp ik_z\Delta z$. Solve (1.43) for the unknown.

$$q_{z+\Delta z} \quad = \quad \left(\frac{1 \;+\; i\,\hat{k}_z\Delta z/2}{1 \;-\; i\,\hat{k}_z\Delta z/2}\right)\; q_z \tag{8.48}$$

So for N_z layers in depth $z = N\Delta z$ we have the approximation

$$e^{ik_zN\Delta z} \quad \approx \quad \left(\frac{1 \;+\; i\,\hat{k}_z\Delta z/2}{1 \;-\; i\,\hat{k}_z\Delta z/2}\right)^N \tag{8.49}$$

which will be of later use for Fourier domain simulations of finite difference programs. Such simulations enable us to compare the accuracy of various migration methods.

8.4 ACCURACY—THE CONTRACTOR'S VIEW

A chain is no stronger than its weakest link. Economy dictates that all the links should be equally strong. Many broad questions merit study such as the errors associated with velocity uncertainty and with migration after, rather than before, stack. Having read this far, you are now qualified to attack these broad questions. Now we will narrow our focus and examine only the errors in downward continuation that result from familiar data processing approximations.

In the construction of a production program for wave-equation migration, weakness arises from approximations made in many different places. Economy dictates that funds to purchase accuracy should be distributed where they will do the most good. Geophysical contractors naturally become experts on accuracy/cost trade-offs in the migration of stacked data. Contractors will use the equations and program gathered below to obtain best results for the lowest cost. Users of reflection data are interested in learning to recognize imperfect migrations, so they may want to use the program to see the effect of various shortcuts.

In preparation for a big production job there are two general approaches to examining accuracy. The first approach, which gives the best insight into qualitative phenomena, is to make synthetic hyperbolas by the various methods. The synthetic hyperboloids can be be compared to the data at hand with a video movie system or by plotting on transparent paper. In the second approach you compute travel times of hyperboloids or spheroids of waves of different stepouts and frequencies for different mesh sizes, etc. Then an optimization program can be run to minimize the average error over the important range of parameters.

It is not necessary to write a time domain 45° finite difference migration program to see what its synthetic hyperboloids would look like. We can simply express all the formulas in the (ω, k_x)-domain and then do an inverse two-dimensional Fourier transformation. To facilitate comparisons between the many migration methods we will gather equations from different parts of the book in the order that they are needed. Then I'll present the program that makes the diffraction hyperbolas for many methods. Using the same equations you can compute travel times and solve the optimization problem as you wish.

8.4.1 Lateral derivatives

First, k_x will range over $\pm\pi/\Delta x$. If the x-axis is going to be handled by finite differencing then we will need equation (1.39):

$$\left(\frac{\hat{k}\,\Delta x}{2} \right)^2 \quad = \quad \frac{\sin^2 \frac{k\,\Delta x}{2}}{1 \ - \ b\,4\ \sin^2 \frac{k\,\Delta x}{2}}$$

So if the x-axis is going to be handled by finite differencing then subsequent reference to k_x should be replaced by $\hat{k}_x$. The finite differencing introduces the free parameter b. Likewise, you could also scale the whole expression by an adjustable parameter near unity. Also, Δx isn't necessarily fixed by the data collection. You could always interpolate the data before processing. A finite-difference method using interpolated data could be mandated by enough lateral velocity variation.

8.4.2 Viscosity and causality

The frequency ω will range over $\pm\pi/\Delta t$. If the t-axis is going to be handled by finite differencing then we will need the Z-transform variable equation (1.50)

$$Z \quad = \quad e^{i\,\omega\,\Delta t} \tag{8.50}$$

and the causal derivative (1.52)

$$-\,i\hat{\omega} \quad = \quad \frac{2}{\Delta t}\,\frac{1\ -\ \rho Z}{1\ +\ \rho Z} \qquad\qquad -1\ <<\ \rho\ <\ 1 \tag{8.51}$$

The data can be subsampled or supersampled before processing, so Δt is an adjustable parameter. The causality parameter ρ should be a small amount less than unity, say $\rho = 1-\epsilon$ where $\epsilon > 0$ is an adjustable parameter. You may want to introduce ρ even if you are migrating in the frequency domain because it reduces wraparound in the time domain—it is a kind of viscosity. The ϵ should be about inverse to the data length, say $1/N_t$ where N_t is the number of points on the time axis. (Because I made many plots of synthetic hyperbolas with square root gain, $\gamma = 1/2$, time wraparounds were larger than life. So I had the program default to ϵ four times larger). If you like to adjust free parameters, you could separately adjust numerator and denominator values of ρ. Subsequently, I'll distinguish between ω and $\hat{\omega}$, but you can take $\hat{\omega}$ to be ω if you don't care to introduce causality.

8.4.3 Retarded Muir recurrence

The k_z square root may be computed with the square root function in your computer or by Muir's expansion. For Fourier domain calculations incorporating causality, you must use a complex square root function. This will also take care of the evanescent region automatically—you no longer have the discontinuity between evanescent and nonevanescent regions. The square root of a complex number is multivalued, so you better first check that your computer chooses the phase as described in Figure **??**. Mine did. But I found that limited numerical accuracy prevented me from achieving strict positivity of the real part of the impedance until I replaced the expression $\sqrt{s^2\ +\ v^2k^2}\ -\ s$ by its algebraic equivalent $v^2k^2/(\sqrt{s^2\ +\ v^2k^2}\ +\ s)$.

For finite differencing we will need the Muir recurrence. Let r_0 define the cosine of the angle that starts the Muir recurrence, often 0° or 45°. This is another free parameter for optimization. This angle is also an angle of exact fit for all orders of the recurrence.

$$s \quad = \quad -\,i\,\hat{\omega} \tag{8.52}$$

Starting from $R_0 = r_0\,s$ the Muir recursion is equation (1.53):

$$R_{n+1} \quad = \quad s\ +\ \frac{v^2\,k_x^2}{s\ +\ R_n} \tag{8.53}$$

For a diffraction program we will be evaluating $\exp(-Rz)$. Since R can be proven to have a positive real part, the exponential should never grow. Finite difference calculations are normally done with retarded time. To retard time, $\exp(-Rz)$ is expressed as

$$e^{-\,R\,z/v} \quad = \quad e^{-\,(R\,-\,s)\,z/v} \quad e^{-s\,z/v} \tag{8.54}$$

As discussed earlier in this chapter, you probably don't want the time shift of retardation to be associated with viscous effects. So you will probably want to downward continue instead with

$$e^{-\,[R(-i\hat{\omega})\,+\,i\hat{\omega}]\,z/v} \quad e^{+i\omega\,z/v} \tag{8.55}$$

Notice the signs and distinction of ω from $\hat{\omega}$.

From equation (1.53) we see that $R-s$ should have a positive real part. I found that numerical roundoff sometimes prevented it. So the Muir recurrence was reorganized to incorporate the retardation. Let

$$R' \quad = \quad R \;-\; s \tag{8.56}$$

Equation (1.53) becomes

$$R'_{n+1} \quad = \quad \frac{v^2\,k_x^2}{2\,s \;+\; R'_n} \tag{8.57}$$

From Muir's rules, you can see that R' will always have a positive real part if we start it that way, so we start it from

$$R'_1 \quad = \quad \frac{v^2\,k_x^2}{s\,(1 \;+\; r_0)} \tag{8.58}$$

(Combining (1.58) and (1.56) gives the same 15° equation as does (1.53).) Mathematically (1.55) is identical to

$$e^{-\,R'\,z/v} \quad e^{+i\omega\,z/v} \tag{8.59}$$

but numerically the exponential in (1.59) is assured to decay in z.

8.4.4 Stepping in depth

If you are using finite differences on depth, or travel-time depth, then we have the relation (1.46):

$$\frac{\hat{k}_z\,\Delta z}{2} \quad = \quad \tan\,\frac{k_z\,\Delta z}{2}$$

which can be used in $\exp\, ik_z z$. Since dissipation may be present, the quantities above may be complex. Let N_z be the number of depth layers, generally equal or less than N_t. Adapting (1.49) to (1.59) gives

$$e^{-\,R'\,N\,\Delta z/v} \quad \approx \quad \left(\frac{1\;-\;R'\Delta z/2v}{1\;+\;R'\Delta z/2v}\right)^N \tag{8.60}$$

8.4.5 Lightning phase shift migration

Not only can the causality and viscosity features of time domain methods be incorporated in the frequency domain, but the square roots and complex exponentials of the phase shift method can be replaced by complex multiplies and divides. First note that the square root expansion need not begin from a starting guess or from a lower order iteration. It could begin from the square root previously obtained from the preceding ω or k. (I noticed this when an early version of the program had a bug that made all my 15° calculations look like 90° calculations)! Also, $\Delta z/v$ is necessarily small in the phase shift method to accommodate imaging at every time point. So the finite differencing form (1.60) is probably as good as the complex exponential. Why don't you try it?

8.4.6 Final appearances

The output of an impulse-response program always looks awful. The main reason is the large area of (ω, k_x)-space which is near the Nyquist frequency or above the evanescent cutoff. Since we rarely sample data in time as coarsely as the Nyquist criterion permits, the program below defaults to final filtering with the filter $(1+Z)/(1-.8Z)$. This filter still passes a lot of energy outside the usual bandwidth of seismic data. Since all land data and most marine data do not have the zero frequency component, the program contains an option to filter further with $(1-Z)/(1-.8Z)$. I haven't displayed anything with this extra filter because I wanted this book to show all the artifacts you might encounter. Furthermore, I deliberately enhanced the visibility of artifacts on wiggle-trace, variable-area plots by plotting with the nonlinear $\gamma = 1/2$ gain. (Perspective hidden-line drawings always have linear gain). Since my plots are necessarily about 10 cm square in this book and in practice you will look at plots of about a hundred times the area, I plotted only one second of travel time.

8.4.7 Program

Many of the figures in the book were made with the program presented here. To enable you to reproduce them I am including the complete program. The parameter input and data output calls are site dependent, but I include them anyway to help clarify the defaults, increasing the odds that you can get exactly what I did.

```
# representations of   e^{-sqrt{ (-i omega /v)^2 + k^2 } z}
#
subroutine allhyp()
integer fetch
integer iw,nw,ik,nk, omhat, kxhat, kzhat, degree, tfilt, xfilt
real v, dt, dx, dz, xf, x0, tf, tau0,  rho, bi, r0, eps,  pi, omega, k, vk2
complex cz, cs, cikz, cexp, cmplx, csqrt, cp(1024)
                call putch('esize','i', 8)      # complex numbers
```

```
nw = 256;       call putch(   'n1','i',nw)      # inner index is omega
nk = 64;        call putch(   'n2','i',nk)      # outer index is k_x
                call putch(   'n3','i', 1)      # one frame movie

if( fetch(     'v', 'f',     v) == 0 )  v    = 3.754    # rock vel
if( fetch(    'dt', 'f',    dt) == 0 )  dt   = .004     # Delta t, sec
if( fetch(    'dx', 'f',    dx) == 0 )  dx   = .025     # Delta x, km
if( fetch(    'dz', 'f',    dz) == 0 )  dz   = .004     # Delta z, sec
if( fetch(    'xf', 'f',    xf) == 0 )  xf   = .25;   x0  =xf*nk*dx
if( fetch(    'tf', 'f',    tf) == 0 )  tf   = .5;    tau0=tf*nw*dt

if( fetch( 'omhat', 'i', omhat)==0)     omhat = 0       # omega HAT
if( fetch( 'kxhat', 'i', kxhat) == 0 )  kxhat = 0       # {k HAT}_x
if( fetch( 'kzhat', 'i', kzhat) == 0 )  kzhat = 0       # {k HAT}_z
if( fetch('degree', 'i',degree) == 0 )  degree= 90
if( fetch( 'tfilt', 'i', tfilt) == 0 )  tfilt = 1
if( fetch( 'xfilt', 'i', xfilt) == 0 )  xfilt = 1

if( fetch(   'rho', 'f',   rho) == 0 )  rho   = 1-4./nw
if( fetch(    'bi', 'f',    bi) == 0 )  bi    = 6.726   # b^{-1}
if( fetch(    'r0', 'f',    r0) == 0 )  r0    = 0.7071
if( fetch(   'eps', 'f',   eps) == 0 )  eps   = 0.

call putch('d1','f', dt);       call putch('label1','s','sec')
call putch('d2','f', dx);       call putch('label2','s','kilometers')

call hclose()                   # close data description file
pi = 3.14159265

do ik= 1, nk {                  # loop over all k_x
        k = 2*pi * (ik-1.) / nk
        if( k > pi )    k = k - 2*pi
        k = k / dx
        if( kxhat == 0 )
                vk2 = (v/2)**2 * k*k
        else
                vk2 = (v/2)**2 * (2/dx)**2 * sin(k*dx/2)**2  / (1 -
                                      (4./bi) * sin(k*dx/2)**2 )

        do iw= 1, nw {                  # loop over all omega
                omega = 2*pi * (iw-1.) / nw
                if( omega > pi )        omega = omega - 2*pi
                omega = omega / dt
                cz = cexp( cmplx( 0., omega * dt ) )
                if( omhat == 0 )
                        cs = cmplx( 1.e-5 / dt, - omega)
                else
                        cs = (2./dt) * (1. - rho * cz) / (1. + rho * cz)
                if ( degree == 90 )
                        cikz = vk2 / ( csqrt( cs * cs + vk2 ) + cs )
                if ( degree == 15  ||  degree == 45 )
                        cikz = vk2 / ( eps + (r0+1.) * cs )
                if (degree == 45 )
                        cikz =  vk2 / ( 2.*cs + cikz)
```

```
            if( real( cikz) < 0.) call erexit('cikz not positive real')

            if( kzhat == 0 )
                    cp(iw)= cexp( - tau0  * cikz )
            else
                    cp(iw)= ((1.-cikz * dz/2) /(1.+cikz * dz/2)) **(tau0/dz)
            cp(iw) = cp(iw) * cexp(cmplx(0., omega * tau0 ))   # unretard

            if( tfilt >= 1 ) cp(iw) = cp(iw) * (1+cz) / (1 - .8 * cz)
            if( tfilt >= 2 ) cp(iw) = cp(iw) * (1-cz) / (1 - .8 * cz)
            if( xfilt == 1 ) cp(iw) = cp(iw) * (1+cos(k*dx)) / (1+.85*cos(k*dx))
            cp(iw) = cp(iw) * cexp( cmplx(0.,k*x0) )
            }
        call srite( 'out', cp, 8*nw )   # write
        }
call exit(0);   end
```

Twenty-two plots displayed in this book were made with this program. The output of the program is 2-D Fourier transformed, the real part taken, and plotted. Different input parameters for the different plots are in the table below.

Figure	Default parameter overrides
3.8a	tfilt=0
3.8b	
5.1a	
5.1b	$\hat{\omega}$
1.1a	$\hat{\omega}$, $v\,=\,2.0$, xf=.5, xfilt=0
1.1b	$\hat{\omega}$, $v\,=\,2.0$, xf=.5
1.5a	15°
1.5b	15°, $\epsilon\,=\,1$
1.6a	45°, tf=.2
1.6b	45°, tf=.2, $\epsilon\,=\,1$
1.16	45°, tf=.2, $\hat{\omega}$
1.20a	v=2.0, 90°
1.20b	v=2.0, 15°, $\hat{\omega}$, $\hat{k}_x$, $\hat{k}_z$
1.22a	$\hat{k}_x$, $\hat{\omega}$, b^{-1}=1000000.
1.22b	$\hat{k}_x$, $\hat{\omega}$, b^{-1}=12.
1.22c	$\hat{k}_x$, $\hat{\omega}$, b^{-1}=6.726
1.22d	$\hat{k}_x$, $\hat{\omega}$, b^{-1}=6.
1.22e	$\hat{k}_x$, $\hat{\omega}$, b^{-1}=5.
9.6a	
9.6b	$\hat{\omega}$
1.23a	Δz=.004, 45°, tf=.3, $\hat{\omega}$, $\hat{k}_x$, $\hat{k}_z$
1.23b	Δz=.012, 45°, tf=.3, $\hat{\omega}$, $\hat{k}_x$, $\hat{k}_z$

It is rumored that accuracy can be improved by making the z mesh coarser. This couldn't happen if x and t were in the continuum, but since they may be discretized, there is the possibility of errors fortuitously canceling. To test this rumor, I tried tripling Δz (actually, tripling the increment in travel-time depth). The result is in Figure 1.23. What do you think?

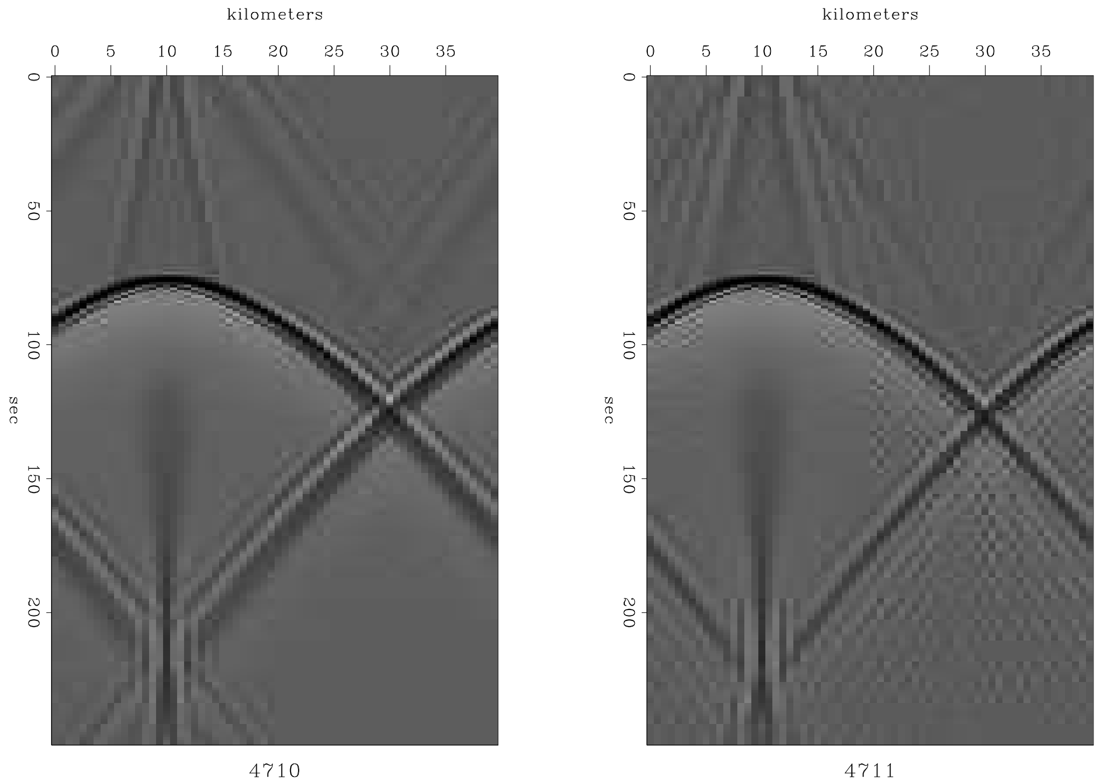


Figure 8.23: Left is a 45° point diffraction in (x, z, t)-space with $\Delta z = v\Delta t$. At the right, with $\Delta z = 3v\Delta t$. dspr-bigdz [ER]

Complete as this analysis must seem, it is limited by the assumption of Fourier analysis that velocity is constant laterally. To handle this problem we turn now to the final lecture on techniques.

Index