FREQUENCY DISPERSION AND WAVE-MIGRATION ACCURACY

Frequency dispersion is a result of 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 changing pops into percussive notes. Even while frequency dispersion is a barely perceptible phenomenon in reflection seismology, it is a substantial nuisance in seismic data processing, where we often see artful displays of the difference between differential operators and difference operators. As such, it is an embarrassment to process builders. Dispersion is not an essential feature of data processed by finite differences: it 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 depicts some dispersed pulses.

But dispersion can be a useful warning to the seismologist that the data itself is in danger of transgression over the boundary into aliased space. Frequency-domain methods do not depend on difference operators so they have the advantage that they do not show dispersion. Penalties for this advantage are: (1) limitation to constant material properties, (2) wraparound, (3) occurrence of spatial aliasing without the warning of dispersion.

Spatial Aliasing

Aliasing can occur on the axes of time, depth, geophone, shot, midpoint, offset, or crossline. We will begin on the horizontal space axis where the problem is worst. 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 by the semicircle relation $k_z(\omega,k_x) = (\omega^2/v^2 - k_x^2)^{\frac{1}{2}}$. Sampling on the x-axis gives an upper limit to k_x equal to the Nyquist frequency $\pi/\Delta x$. Both frequency-domain methods and finite-difference methods treat higher frequencies as if they were folded at the Nyquist frequency. Thus the semicircle dispersion relation is replicated above the Nyquist frequency as shown in figure 2. Angle considerations cause an association between the temporal frequencies and spatial

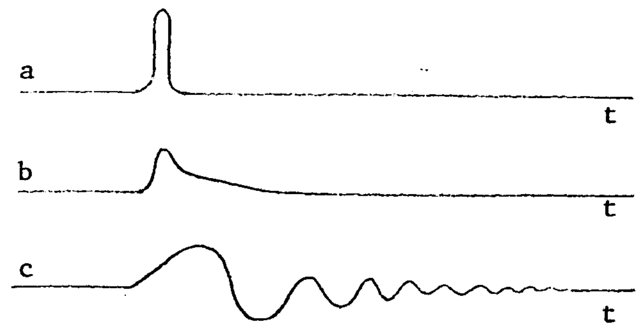

FIG. 1. (a) A pulse. (b) A pulse slightly dispersed as by the physical dissipation of high frequencies. (c) A pulse with a substantial amount of frequency dispersion, as could result from data processing.

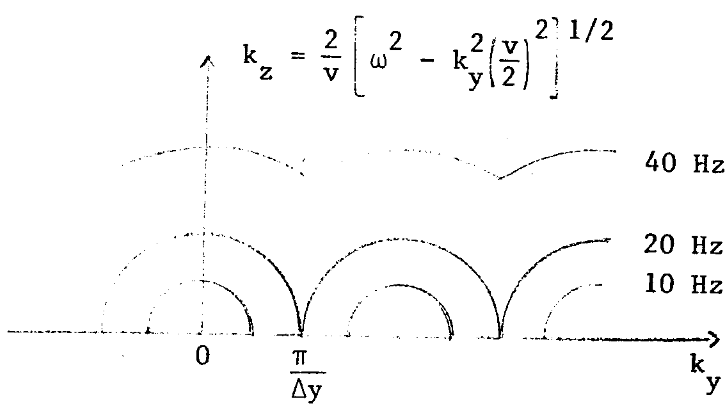

FIG. 2. The effective dispersion relation of the wave equation when the horizontal axis is sampled. Values are given for zero-offset migration where Δy = 25 meters and velocity v = 2000 m/sec. The semicircular arcs correspond to frequencies of 40, 20, and 10 Hz.

frequencies. The temporal frequencies are divided into low frequencies, which are always free from worries about spatial aliasing, and higher frequencies, which are safe only at increasingly small dips. To get some feeling for the seriousness of the practical problem, consider zero-offset migration, which implies two-way traveltime. Two-way time amounts to halving velocity or frequency so horizontally moving energy (as reflections from a vertical fault) has $\omega = vk_y/2$. Temporal frequencies which cannot be spatially aliased are $\omega \le v\pi/(2\Delta y)$ or $f \le v/4(\Delta y)$ Hz. Take as typical a velocity of 2 km/sec and midpoint samplings Δy of 25 meters (good), 50 meters (reconnaisance), and 100 meters (3-D crossline). Then the completely safe frequencies turn out to be 20, 10, and 5 Hz. The seriousness of this problem is often understated by a factor of two when two-way time is not considered.

Another view is that steeply dipping waves are suppressed by the geophone group. (This disregards shot-space aliasing.) In this view the limitation should be thought of in terms of angles at which energy is missing from the data. Taking the ray angle to be 30 degrees instead of 90 degrees doubles horizontal wavelengths. This allows the doubled frequencies (40,20,10) Hz to be safe from spatial aliasing. Some perspective on the significance of wide-angle processing is gained by realizing that data commonly exhibit good signals above 40 Hz.

Second Space-Derivatives

The defining equation for a second-derivative operator is

$$\frac{\delta^2}{\delta x^2} P = \frac{P(x + \Delta x) - 2P(x) + P(x - \Delta x)}{(\Delta x)^2} \qquad (1)$$

The second derivative operator is defined by taking the limit

$$\frac{\partial^2}{\partial x^2} P = \lim_{\Delta x \to 0} \frac{\delta^2}{\delta x^2} P \qquad (2)$$

Many different equations can all go to the same limit as Δx goes to zero. So the problem is to find an expression which is accurate when Δx is larger than zero. The practical problem is to find an accurate expression which is not too complicated. Our first objective is to see how to evaluate quantitatively the accuracy of equation (1). Second, we will look at an expression that is slightly more complicated but much more accurate. This expression is incorporated into nearly all production finite-difference programs in exploration seismology.

The basic method of analysis is Fourier transformation. More simply, we take derivatives of the complex exponential $P = P_0 \exp(ikx)$ and look at errors as a function of the spatial frequency k. For the second derivative we have

$$\partial_{xx} P = \frac{\partial^2}{\partial x^2} P = -k^2 P \tag{3}$$

We define $\hat{k}$ by an analagous expression with the difference operator:

$$\delta_{xx} P = \frac{\delta^2}{\delta x^2} P = -\hat{k}^2 P \tag{4}$$

Ideally $\hat{k}$ would equal k. Inserting the complex exponential into (1) we get an expression for $\hat{k}$ in terms of k

$$-\hat{k}^2 P = \frac{P_0}{\Delta x^2} \left[e^{ik(x+\Delta x)} - 2e^{ikx} + e^{ik(x-\Delta x)} \right] \tag{5a}$$

$$-\delta_{xx} = \hat{k}^2 = \frac{2}{\Delta x^2} [1 - \cos(k\Delta x)] \tag{5b}$$

It is a straightforward matter to make plots of $\hat{k}\Delta x$ versus $k\Delta x$ from (5). The half-angle trig formula allows an analytic square root of 5 which is

$$\hat{k}\Delta x = 2 \sin \frac{k\Delta x}{2} \tag{5c}$$

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.

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

$$\frac{\partial^2}{\partial x^2} \approx \frac{\delta^2}{\delta x^2} - \frac{\Delta x^2}{12} \frac{\delta^4}{\delta x^4} + \text{etc.} \tag{6}$$

As Δx tends to zero (6) tends to the basic definition (1) and (2). Coefficients like the 1/12 in (6) may be determined by the Taylor-series method if great accuracy is desired at small k. Or a somewhat different coefficient may be determined by curve-fitting techniques if accuracy is desired over some range of k. In practice I believe (6) is hardly ever used because there is a less obvious expression which offers much more accuracy at less cost! The idea is indicated by

$$\frac{\partial^2}{\partial x^2} \approx \frac{\dfrac{\delta^2}{\delta x^2}}{1 + \dfrac{\Delta x^2}{6} \dfrac{\delta^2}{\delta x^2}} \tag{7a}$$

The accuracy may be numerically evaluated by substituting from (5) to get

$$\left(\frac{\hat{k}\Delta x}{2}\right)^2 = \frac{\sin^2 \dfrac{k\Delta x}{2}}{1 - \dfrac{1}{6} 4 \sin^2 \dfrac{k\Delta x}{2}} \tag{7b}$$

The square root is plotted in figure 3.

If the 1/6 in (7) were replaced by 1/12 then (7) and (6) would agree to second order in Δx. Actually the 1/12 comes from series expansion but the

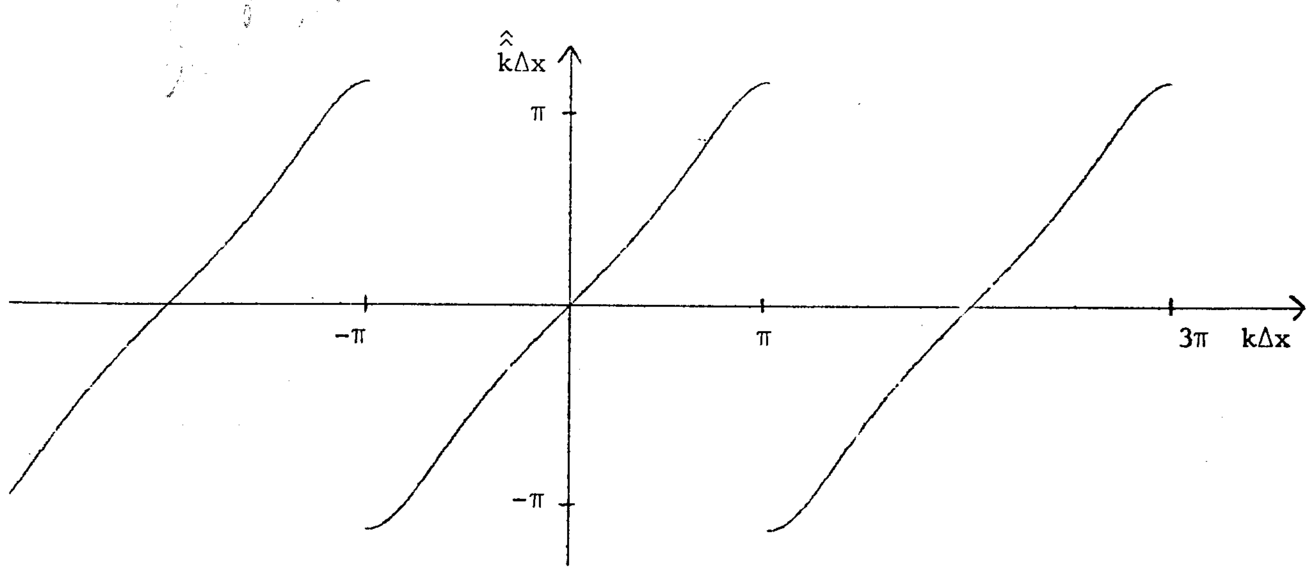

FIG. 3. Accuracy of the second-derivative representation of (7) as a function of spatial wavenumber. The sign of the square root of (7b) was chosen to agree with k in the range $-\pi$ to π and to be periodic outside the range. (Curve computed by Dave Hale.)

1/6 fits over a wider range and is a value in common use. F. Muir pointed out that the value $(1/\pi^2 - 1/4) \approx 1/6.7$ gives an *exact* fit at the Nyquist frequency and a quite accurate fit over all lower frequencies! Indeed, in 1980 it may be said that few explorationists consider the accuracy deficiency of (7) to be large enough to warrant interpolation of field-recorded values.

Rather than pursue this further to provide a thorough analysis of errors, let us be sure it is clear how (7) is put into use. The simplest prototype equation is the heat-flow equation.

$$\frac{\partial}{\partial t} q = \frac{\partial^2}{\partial x^2} q \approx \frac{\delta_{xx}}{1 + \frac{\Delta x^2}{6} \delta_{xx}} q \tag{8a}$$

Just multiply through the denominator thus

$$\left(1 + \frac{\Delta x^2}{6} \delta_{xx}\right) \frac{\partial}{\partial t} q \approx \delta_{xx} q \tag{8b}$$

The apparently new aspect of this equation is the mixed δ_{xxt} derivative. But it is represented on the usual 6 point star as follows:

$$\xrightarrow{\quad\quad} x$$

$$\frac{1}{\Delta x^2 \Delta t} \begin{array}{|c|c|c|} \hline -1 & 2 & -1 \\ \hline 1 & -2 & 1 \\ \hline \end{array} \quad \downarrow t \qquad (9)$$

So, other than modifying the six coefficients on the star it adds nothing to the computational cost of solving the equation.

Time-Derivatives and the Bilinear Transform

You might be inclined to think that a second derivative is a second derivative and that there is no mathematical reason to do time-derivatives differently than space-derivatives. This is wrong. A hint of fundamental disparity is found by considering boundary conditions. With time-derivatives (and often with the depth z-derivative) we generally have a concept of causality, which means that the future is determined solely from the present and past. Appropriate boundary conditions on the time axis are initial conditions - that is, specification of the function (and perhaps some derivatives) at *one* point, the initial point in time. For depth z that special point is the earth 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 under consideration.

Causal differentiation is a deep subject. Its importance to stability in wave analysis merits more lengthy consideration in a later chapter on advanced wave extrapolation. But we have already seen the main idea, which is embedded in the Crank-Nicolson differencing scheme. It remains to examine accuracy and frequency dispersion.

Begin with a time function p_t. We define its Z-transform by

$$P(Z) \;=\; \cdots p_{-2}Z^{-2} + p_{-1}Z^{-1} + p_0 + p_1 Z + p_2 Z^2 + \cdots \qquad (10)$$

Define an operator $-i\hat{\omega}\Delta t$ by

$$\frac{1}{-i\hat{\omega}\Delta t} = \frac{1}{2}\frac{1+Z}{1-Z} \tag{11}$$

Let us apply this operator on P to get the Z-transform Q of another time function q_t.

$$Q(Z) = \frac{1}{2}\frac{1+Z}{1-Z}P(Z) \tag{12a}$$

Multiply both sides by $(1-Z)$:

$$(1-Z)Q(Z) = \frac{1}{2}(1+Z)P(Z) \tag{12b}$$

Equate the coefficient of Z^t on each side:

$$q_t - q_{t-1} = \frac{p_t + p_{t-1}}{2} \tag{12c}$$

Taking p_t to be an impulse function we see that q_t turns out to be a step function, that is,

$$p = \cdots 0,0,1,0,0,0,\cdots \tag{13a}$$

$$q = \cdots 0,0,\tfrac{1}{2},1,1,1,\cdots \tag{13b}$$

So q_t approximates the integral of p_t from minus infinity to time t. Taking $Z = \exp(i\omega\Delta t)$ equations (10), (11), and (12a,b) are expressions in the Fourier transform domain and the operation of (11) represents numerical integration by the Crank-Nicolson method. The accuracy of the integration (or differentiation) is evaluated by substituting $Z = \exp(i\omega\Delta t)$ into (11), say

$$-i\hat{\omega}\Delta t = 2\frac{1-e^{i\omega\Delta t}}{1+e^{i\omega\Delta t}} = 2\frac{e^{-i\omega\Delta t/2} - e^{i\omega\Delta t/2}}{e^{-i\omega\Delta t/2} + e^{i\omega\Delta t/2}}$$

$$= -2i\,\frac{\sin(\omega\Delta t/2)}{\cos(\omega\Delta t/2)}$$

$$\frac{\hat{\omega}\Delta t}{2} = \tan\left(\frac{\omega\Delta t}{2}\right) \tag{14}$$

Equation (14) is the fundamental statement of the accuracy of approximation of the first-derivative operator by the Crank-Nicolson method. 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 (14). The bad news is that there does not seem to exist a representation of causal differentiation which is any more accurate than Crank-Nicolson. In other words we have nothing like the 1/6 trick for time-derivatives. So we must reduce the time sample interval Δt considerably from the Nyquist criterion.

On the other hand most geophysical differential equations have time-invariant coefficients so we can solve them in the ω-domain rather than in the time domain. Or we might find we do not need to have *causal* time-derivatives: antisymmetric derivatives might do. With the depth z-axis we are largely stuck with causal derivatives, although we could use Fourier methods over layers. But the depth axis is not so troublesome as the x- and t-axes because it usually affects computer time only, not data storage. The practical picture may not be as dreary as the one I am painting. Many people are very pleased with both the speed and accuracy of time-domain migrations at $\Delta t = 4$ milliseconds.

Accuracy - The Contractor's View

A chain is said to be no stronger than its weakest link. Economy dictates that the links should all be made of equal strength. Likewise, in the construction of a production program for wave-equation migration, weakness arises from approximations made in many different places. Again, economy dictates that funds to purchase accuracy should be distributed to where they will do the most good. Geophysical contractors naturally become experts on accuracy-cost trade-offs in the migration of stacked data. Certain broader questions also merit study, such as the error associated with velocity uncertainty and the error associated with migration after stack rather than before. Here we will just look at some of the more obvious considerations.

Migration is basically just a process of downward extrapolating surface data. All the various approximations imply timing errors, which for a single frequency are just phase errors. It is easy to write equations for these phase errors. The true phase ϕ at a depth z is given by the integral of k_z with depth. But we may as well replace the integral by the average integrand times the depth:

$$\phi_{true} = z\, k_z \tag{15a}$$

Discretizing the z-axis into N levels,

$$\phi_{diff} = N\,\Delta z\, \hat{k}_z = N\,\Delta z\, \frac{z}{\Delta z} \tan\left(\frac{k_z\,\Delta z}{z}\right) \tag{15b}$$

Specializing for scalar waves,

$$k_z = \left(\frac{\omega^2}{v^2} - k_x^2\right)^{\frac{1}{2}} \tag{16}$$

or retarded scalar waves

$$k_z = \frac{1}{v}\left[\omega - (\omega^2 - v^2 k_x^2)^{\frac{1}{2}}\right] \tag{17}$$

or from the lecture on the DSR equation, the phase for the zero-offset migration is obtained by replacement, midpoint y for x, and v/2 for z:

$$k_z = \frac{2}{v}\left\{\omega - \left[\omega - \left(\frac{v}{2}\right)^2 k_y^2\right]^{\frac{1}{2}}\right\} \tag{18}$$

We may simplify the algebra with no conceptual loss by making the 15-degree approximation

$$k_z = \frac{v k_y^2}{4\omega} \tag{19}$$

Discretizing the x- and t-axes, k_z becomes

$$k_z = \frac{v\hat{k}_y^2}{4\hat{\omega}_2} \tag{20}$$

The worst errors will occur at the highest frequency ω and the steepest dips k_y/ω. You need to estimate those. Then you need to decide on an acceptable phase error. This is often taken to be a half-wavelength or about 1%. Then you choose each of Δz, Δt, Δx to keep $(\zeta_{true} - \zeta_{diff})/\zeta_{diff}$ less than about 1%. If I say any more than this you will either go to sleep or you will disagree with me, so I stop here.