High-order finite-difference reverse time migration with the 2-way non-reflecting wave equation

John T. Etgen

INTRODUCTION

Most conventional migration schemes impose limits on the generality of the earth velocity model and/or the dip of waves traveling through this model. Finite-difference migration schemes also face a limitation on the frequency content of the wavefield because the second order derivatives usually employed are accurate only if wavefields are heavily oversampled. Because the amount of core memory in computers is regrettably finite, the oversampling requirements also translate into a restriction on the size problem that can be solved. In this paper I will present a migration method that can be used for any general earth velocity, that can propagate waves in all directions, and reduces the amount of sampling required to propagate waves accurately.

REVERSE TIME MIGRATION

I will discuss a method of migration that uses a time-reversed stacked section (assumed zero-offset) as a time-variable boundary condition for wave propagation in an earth velocity model. This method, known as *reverse time migration* has been discussed before by (Whitmore,1984), (Sherwood, Kosloff, and Baysal; 1984) and (Levin, SEP-37). With this method, migration of zero-offset, stacked data propagates the observed wavefield back into the earth and chooses the position of wavefield at the time corresponding to the zero-offset reflection time as the earth image. This differs from conventional migration in that waves are propagating in time, rather than being extrapolated in depth. Like most post-stack migration procedures, this procedure relies on the *exploding reflector* model of zero-offset seismograms. Reverse time migration proceeds by using time slices of the stacked data, taken progressively shallower off the bottom of the stacked section as a boundary condition for wave propagation in the specified velocity model. We can think of this as using the geophones on the earth's surface to play the recorded wavefield backward in time, causing this wavefield to propagate back into the earth along the path it followed going

toward the surface when it was recorded. The stacked section is played back in reverse order, because the latest arrivals in general have the greatest distance to travel back into the earth. The earth velocities are halved to account for the one way propagation of the exploding reflector model versus the actual two way propagation of a true zero-offset experiment. Thus, when all of the stacked section has been "played back", and time equals zero has been reached, the wavefield is at the position of the reflecting interfaces, Assuming the correct velocity $v(x,z)$ was used in the model. There is some freedom to choose the method we wish to propagate the wavefield through the velocity model. Use of the scalar wave equation allows any general velocity model and propagates waves with all dips. However, the scalar wave equation will generate reflections from waves passing through interfaces in the velocity model. These extra internal reflections are annoying because they do not fit our primaries-only model of the stacked data, nor do they correspond to the position of any reflector when the time equals zero imaging condition is applied. Smoothing the velocity function $v(x,z)$ reduces the amplitude of the internal reflections, but this is not always physically realistic. John Sherwood et. al. (1983) developed a variation of the scalar wave equation that will not generate internal reflections off velocity interfaces for waves passing through velocity boundaries at near normal incidence. This equation is called *the 2-way non-reflecting wave equation.*

THE 2-WAY NON-REFLECTING WAVE EQUATION

To derive the 2-way non-reflecting scalar wave equation, examine the full 2-way scalar wave equation given in (Claerbout, 1984):

$$\frac{\partial^2 U}{\partial t^2} = K[\frac{\partial}{\partial x}\frac{1}{\rho(x,z)}\frac{\partial U}{\partial x} + \frac{\partial}{\partial z}\frac{1}{\rho(x,z)}\frac{\partial U}{\partial z}] \tag{1}$$

Where K (bulk modulus)$= \rho(x,z)v^2(x,z)$ Now introduce the variable:

$$Z = \rho(x,z)v(x,z)$$

which has physical meaning as the acoustic impedance; and rewrite the equation with $v(x,z)$ and Z as the space variable properties:

$$\frac{\partial^2 U}{\partial t^2} = Z\, v(x,z)[\frac{\partial}{\partial x}\frac{v(x,z)}{Z}\frac{\partial U}{\partial x} + \frac{\partial}{\partial z}\frac{v(x,z)}{Z}\frac{\partial U}{\partial z}] \tag{2}$$

If we set Z=constant, factor it out of the equation, obtaining:

$$\frac{\partial^2 U}{\partial t^2} = v(x,z)\,[\frac{\partial}{\partial x}v(x,z)\,\frac{\partial U}{\partial x} + \frac{\partial}{\partial z}v(x,z)\frac{\partial U}{\partial z}] \tag{3}$$

Or equivalently:

$$\frac{\partial^2 U}{\partial t^2} = v^2(x,z)\,[\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial z^2}] + v(x,z)[\frac{\partial v(x,z)}{\partial x}\,\frac{\partial U}{\partial x} + \frac{\partial v(x,z)}{\partial z}\,\frac{\partial U}{\partial z}] \tag{4}$$

Equation is the *2-way non-reflecting wave equation*. The derivation is often called impedance matching. Notice that $Z = \rho(x,z)\, v(x,z)$ causes the reflection coefficient to be zero, strictly speaking, for normal incidence only. The non-reflecting behavior of the equation is not perfect for non-normal incidence, but for most practical applications it works quite well.

HIGH-ORDER FINITE-DIFFERENCES

Spatial second derivatives

As noted before, most finite-difference migration schemes use second order difference operators to approximate the spatial derivatives in the scalar wave equation (or its parabolic approximation). However, this second order approximation :

$$\frac{1}{\Delta x^2}\left[U_{i+1} - 2U_i + U_{i-1}\right] + O(\Delta x^2) = \frac{\partial^2 U}{\partial x^2} \tag{5}$$

is only a close approximation to the second derivative for large spatial wavelengths. Grid dispersion commonly occurs when the spatial sample interval is not sufficiently small. A common rule of thumb with second order operators is to allow about 16 sample points per shortest wavelength to prevent grid dispersion. If we allow the use of the explicit method of finite-differencing, higher order finite-difference operators can be used to improve the accuracy of the spatial derivatives. To derive the centered finite-difference operator for the second derivative, write out the taylor series approximation to the function U centered at the i^{th} grid point for a point n grid points away:

$$U_{i+n\Delta x} = U_i + \frac{\partial}{\partial x}U_i\, n\Delta x + \frac{1}{2!}\frac{\partial^2}{\partial x^2}U_i\,(n\Delta x)^2 + \frac{1}{3!}\frac{\partial^3}{\partial x^3}U_i\,(n\Delta x)^3 + \cdots \tag{6}$$

Now take the sum of the two taylor series for points symmetric about the i^{th} grid point, which eliminates all odd power terms:

$$\left[U_{i+n\Delta x} + U_{i-n\Delta x}\right] = 2U_i + \frac{2}{2!}\frac{\partial^2}{\partial x^2}U_i(n\Delta x)^2 + \frac{2}{4!}\frac{\partial^4}{\partial x^4}U_i(n\Delta x)^4 + \cdots \tag{7}$$

To create a high order finite-difference operator, weight terms with $n = 1, 2, \cdots nmax$ in a manner that causes the cancellation of terms greater than second order. Obtaining the correct weights involves solving a system of linear equations for the weights. If two terms are used corresponding to $n = 1$ and $n = 2$ the fourth order term can be eliminated by choosing the correct weights. To cancel up through the eighth order terms, weight and combine the terms corresponding to $n = 1, 2, 3, 4$ in equation (7). Then a weight is applied to the constant term U_i to cancel it from the resulting sum with the other weighted terms, leaving the second derivative term and a remainder of order Δx^{2n+2}

$$W_0 U_i + \sum_{n=1}^{nmax} W_n\left[U_{i+n\Delta x} + U_{i-n\Delta x}\right] = \frac{\partial^2 U}{\partial x^2}(\Delta x^2) + O(\Delta x^{2nmax+2}) \tag{8}$$

Subsequently, I used either the fourth order approximation or the eighth order approximation for all spatial second derivatives.

Spatial first derivatives

The first derivatives can be found in a fashion similar to that used for second derivatives. Instead of forming the sum of the symmetrical points about the i^{th} grid point, form their difference, causing all even order terms to vanish:

$$[U_{i+n\Delta x} - U_{i-n\Delta x}] = 2\frac{\partial}{\partial x}U_i(n\Delta x) + \frac{2}{3!}\frac{\partial^3}{\partial x^3}U_i(n\Delta x)^3 + \frac{2}{5!}\frac{\partial^5}{\partial x^5}U_i(n\Delta x)^5 \cdots \tag{9}$$

Again, Taking weighted combinations of these terms for $n = 1, 2, \cdots nmax$ will allow the elimination of higher order derivatives up to the term of order Δx^{2n+1} With the weights again determined by solving a system of linear equations.

$$\sum_{n=1}^{nmax} W_n[U_{i+n\Delta x} - U_{i-n\Delta x}] = \frac{\partial}{\partial x}U_i \cdot \Delta x + O(\Delta x^{2nmax+1}) \tag{10}$$

Dividing by Δx gives a first derivative operator, centered at the i^{th} grid point with error of order Δx^{2nmax}. This concludes the Taylor series derivation of high-order spatial finite-difference operators.

Time derivatives

It is possible to calculate high order approximations to time derivatives in the same fashion as spatial derivatives. However, when using an explicit method, this is undesirable because no information is available at time levels above the present, and operators that are not centered at the present time level are often unstable, or require more complicated updating schemes. A high-order centered difference for the time derivative would also require more memory locations per grid point than we would like. Mark Dablain (1986), presented a method to increase the order of the accuracy of the time derivatives without having to resort to more complicated operators for the time derivative. I will outline his approach using the scalar wave equation, then derive the very similar result I used for the non-reflecting wave equation.

Begin by noticing that the second derivative in time can be expressed using the taylor series as:

$$\frac{\partial^2}{\partial t^2}U = \frac{1}{\Delta t}\left\{[U_{t+\Delta t} - 2U_t + U_{t-\Delta t}] - \frac{1}{12}\frac{\partial^4 U}{\partial t^4} - \cdots\right\} \tag{11}$$

The first part is the standard second-order difference quotient that is commonly used. The second term involves the fourth derivative of U in time. The accuracy of (11) would be fourth-order if we

had a good estimate of the fourth derivative in time. Dablain noticed that the fourth derivative in time could be computed from the scalar wave equation :

$$\frac{\partial^4}{\partial t^4}U = \frac{\partial^2}{\partial t^2}[\frac{\partial^2 U}{\partial t^2}] = \frac{\partial^2}{\partial t^2}\Big\{v^2(x,z)[\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial z^2}]\Big\} \tag{12}$$

Or using the equivalence of time derivatives with spatial operators, rewrite this as:

$$\frac{\partial^4 U}{\partial t^4} = v^2(x,z)\Big\{\frac{\partial^2}{\partial x^2}v^2(x,z)[\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial z^2}] + \frac{\partial^2}{\partial z^2}v^2(x,z)[\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial z^2}]\Big\} \tag{13}$$

Therefore calculate the fourth derivative in time by cascading the wave equation upon itself. In practice, Dablain approximated this calculation by assuming that the velocity was largely constant except at isolated interfaces and pulled out the velocity from the cascade of derivatives:

Thus, the scheme updates the wavefield with the familiar explicit finite-difference method, with the addition of a correction term to the time derivative. I extended this method to the *2-way non-reflecting wave equation* in a similar fashion. Once again, the second derivative of U in time is the same as (11). Realize that the fourth derivative can once again be computed from the governing wave equation:

$$\frac{\partial^2 U}{\partial t^2} = v^2(x,z)\,[\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial z^2}] + v(x,z)[\frac{\partial v(x,z)}{\partial x}\frac{\partial U}{\partial x} + \frac{\partial v(x,z)}{\partial z}\frac{\partial U}{\partial z}] \tag{4}$$

By expressing the fourth derivative in time as:

$$\frac{\partial^4 U}{\partial t^4} = \frac{\partial^2}{\partial t^2}\Big[v^2(x,z)\,[\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial z^2}] + v(x,z)[\frac{\partial v(x,z)}{\partial x}\frac{\partial U}{\partial x} + \frac{\partial v(x,z)}{\partial z}\frac{\partial U}{\partial z}]\Big] \tag{14}$$

Then once again, use the relationship between $\frac{\partial^2}{\partial t^2}$ and the spatial derivatives to calculate the fourth derivative in time:

$$\frac{\partial^4 U}{\partial t^4} = \mathcal{L}\Big[\mathcal{L}[U]\Big] \tag{15}$$

Where :

$$\mathcal{L}[\,U\,] = v^2(x,z)[\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial z^2}] + v(x,z)[\frac{\partial v(x,z)}{\partial x}\frac{\partial U}{\partial x} + \frac{\partial v(x,z)}{\partial z}\frac{\partial U}{\partial z}] \tag{16}$$

I made the assumption that the first derivative term is small, and approximated the fourth derivative correction term (15) using only the second derivatives in $\mathcal{L}$ for the outer operator, giving:

$$\frac{\partial^4 U}{\partial t^4} \approx v^2(x,z)\Big[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\Big]\,\Big[\,\mathcal{L}(U)\,\Big] \tag{17}$$

The approximation I made was not essential, one could resubstitute the entire equation in itself, however, this would effectively double the computational cost of the method. I believe the approximation is justified, because for most geologic models, the velocity is constant or slowly varying over most of the model, with the large changes in velocity being distinct boundaries. This causes the term with the first derivatives to be small, or zero over most of the geologic model. Another

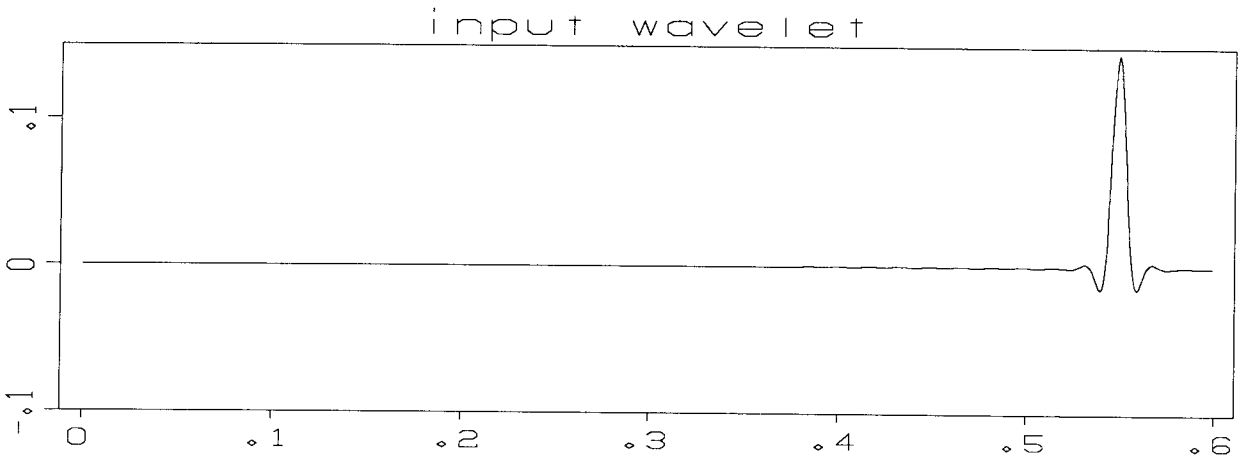

Figure 1: Input wavelet to the dispersion tests of the various order algorithms. Horizontal axis is time (sec), vertical axis is amplitude.

view of this is the fact that the term containing second derivatives controls the actual wave propagation, while the term containing the first derivatives merely acts to suppress reflections. So the most important part of the behavior of $\mathcal{L}$ is contained in the second derivatives. To further speed the computation of the correction term, the two spatial second derivatives in the correction term (17) were computed using a fourth- order scheme when the eighth-order spatial derivatives were used to compute the spatial derivatives in the main equation.

RESULTS

Wavelet tests

Three algorithms were run to examine the computational and kinematic behavior of the concepts I presented. The most simple algorithm used the standard second-order time derivatives and used fourth-order spatial derivatives. The second algorithm used the fourth-order time derivative and the eighth-order spatial derivatives with the fourth-order correction term calculated with fourth order spatial derivatives. An algorithm using an alternating direction implicit (ADI) method (second-order derivatives in x, z and time) was run for comparison. The first series of figures shows the input waveform and the waveform after 300 time steps were taken for each method. The velocity in the first 600 meters was 3500 m/sec and the velocity in the last 600 meters was 2500 m/sec. For each case, the grid point spacing was $\Delta x = \Delta z = 6m$, and the

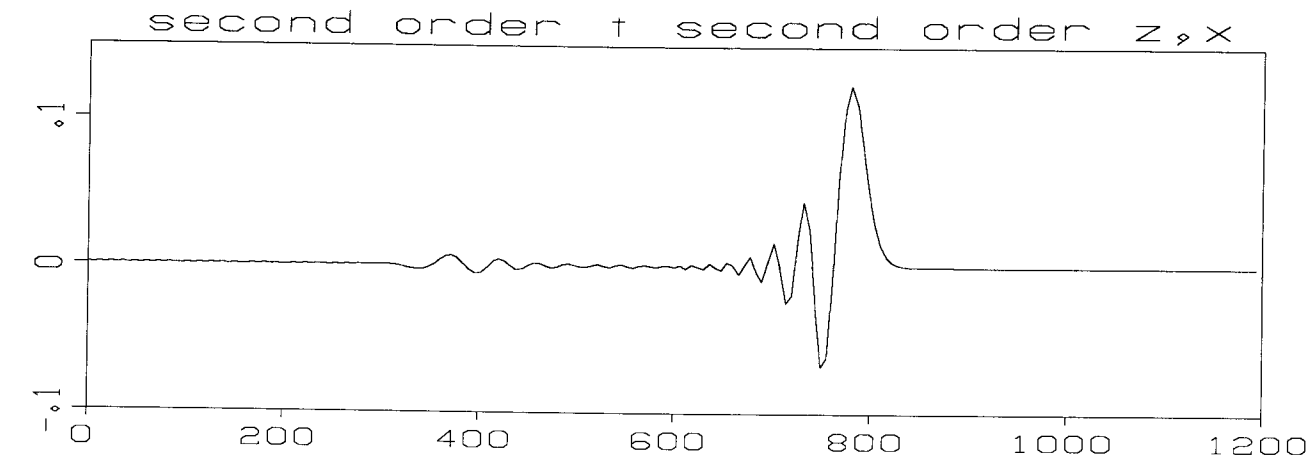

Figure 2: Waveform in depth for second-order differencing in time and x, z algorithm after 300 time steps . Grid dispersion distorts the wavelet. Horizontal axis is depth; vertical axis is amplitude.

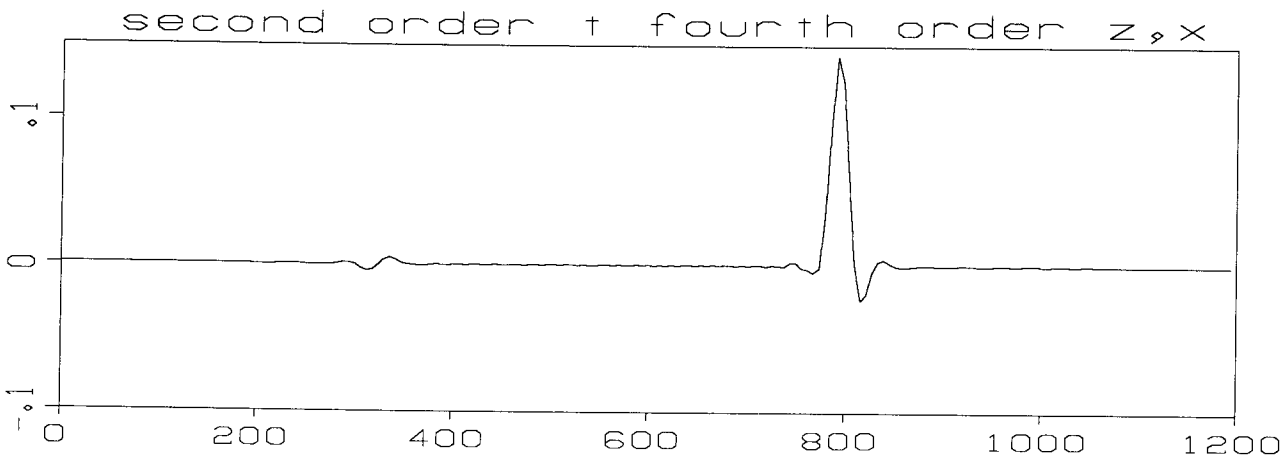

Figure 3: Waveform in depth for second-order differencing in time and fourth-order differencing in x, z algorithm after 300 time steps. Almost no dispersion is present, although a slight phase-shift is apparent. Horizontal axis is depth; vertical axis is amplitude.

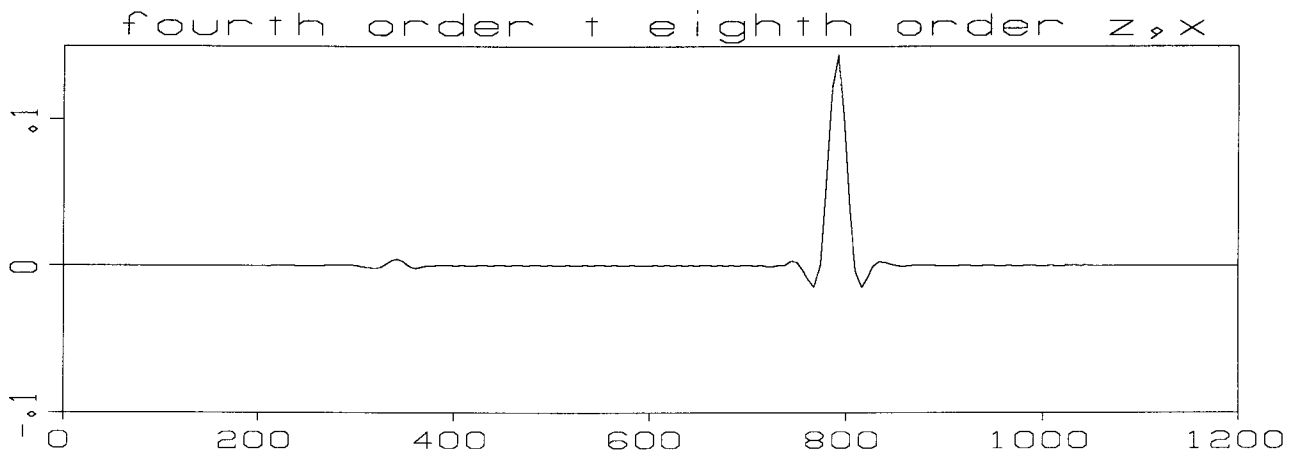

Figure 4: Waveform in depth for fourth-order in time and eighth-order in x, z algorithm after 300 time steps. The wavelet is faithfully propagated; there is no dispersion. Horizontal axis is depth; vertical axis is amplitude.

time step size was 1 msec. Figure 1 shows the input waveform as a vertical section through a horizontal plane-wave. The input waveform has a high-cutoff frequency of 70 Hz. Therefore the number of sample points per highest wavelength in the upper medium is 8.33 and in the lower medium is approximately 6. Figure 2 shows the waveform for the second-order in time and x, z ADI method after 300 time steps. The waveform is suffering from a great deal of grid dispersion. The high frequencies are traveling slower than the low frequencies which gives rise to the phase distortion and the long tail. Figure 3 shows the waveform for the second-order in time, fourth-order in x, z algorithm. The waveform is reasonably well preserved, there is only a slight amount of phase distortion . Figure 4 shows the waveform of the fourth-order in time, eighth-order in x, z algorithm. The waveform exhibits essentially no distortion at all. All spatial wavelengths travel at the correct velocity.

Synthetic examples

The next sequence of figures shows time snapshots of the wavefield during reverse time migration for a data set supplied by Zhiming Li. The time section has 128 traces of 4 seconds of data. The spatial sampling is $\Delta x = 40m$. The time sampling rate is .002 seconds. Figure 5 shows the zero-offset section input to the reverse time migration programs. The travel times were determined analytically by Li, et. al.(1984). The correct velocity is $v(z) = 1000 + z\ m/sec.$ The profile is a model of a high-angle reverse fault, and the object of the migration is to image

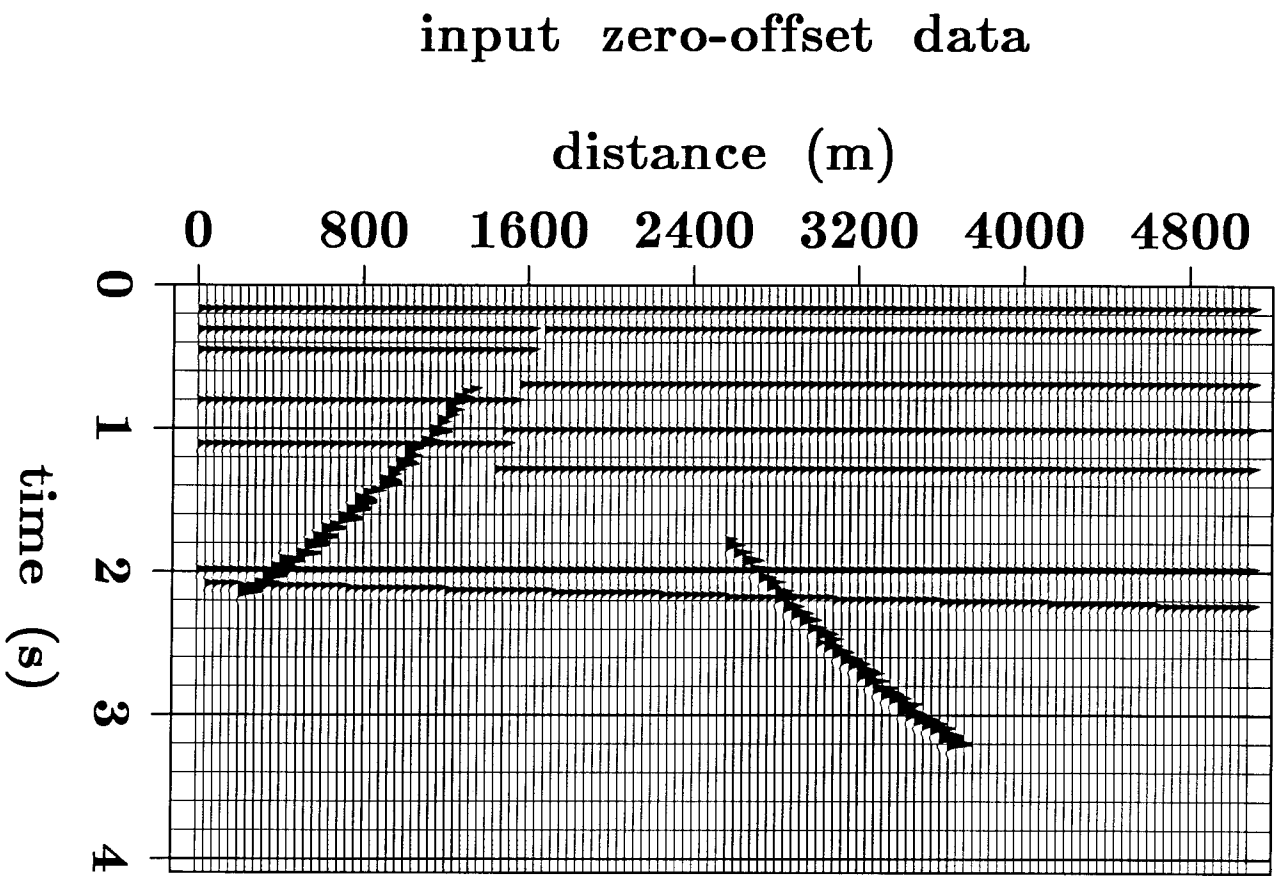

Figure 5: Synthetic zero-offset data used as input to the migration algorithms. Note the steeply dipping reflections from the fault interface.

both sides of the steeply dipping fault interface. The velocity model with the correct velocity was discretized with $\Delta x = 40m$, $\Delta z = 10m$. The grid size was $xmax = 128$ by $zmax = 200$. Both algorithms, second-order in time, fourth-order in x, z, and the fourth-order in time, eighth-order in x, z were run on the input data of Figure 5. Figure 6 shows a snapshot of the wavefield at time=1.0 sec, for the fourth-order in time, eighth-order in x, z algorithm. Three seconds of the input data have been propagated back into the model. The wavefield is propagating back toward the position of the reflecting interfaces in the model. In particular note the two steeply dipping events that correspond to the reflections of each side of the steeply dipping fault. Figure 7 is a snapshot of the wavefield at time=.6 sec. More events have been reintroduced into the model, while the previous events have propagated closer to the corresponding reflector positions. At this point in time, the wave from the underside of the fault interface is vertical. A migration program that did not account for the possibility of two-way propagation or did not save the evanescent

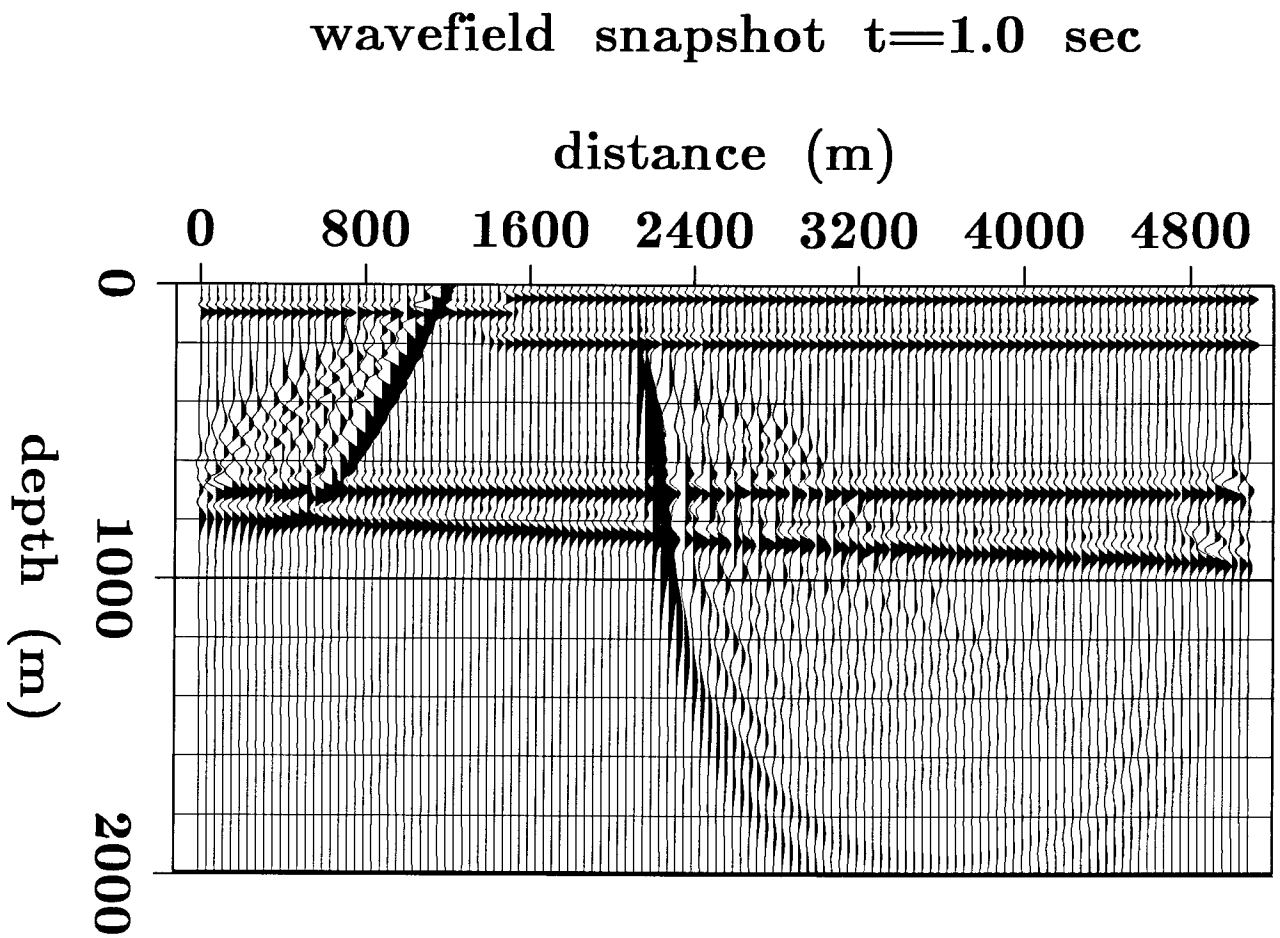

Figure 6: Snapshot of the wavefield in depth for time=1.0 sec. Migration with the fourth-order in time, eighth-order in x, z algorithm.

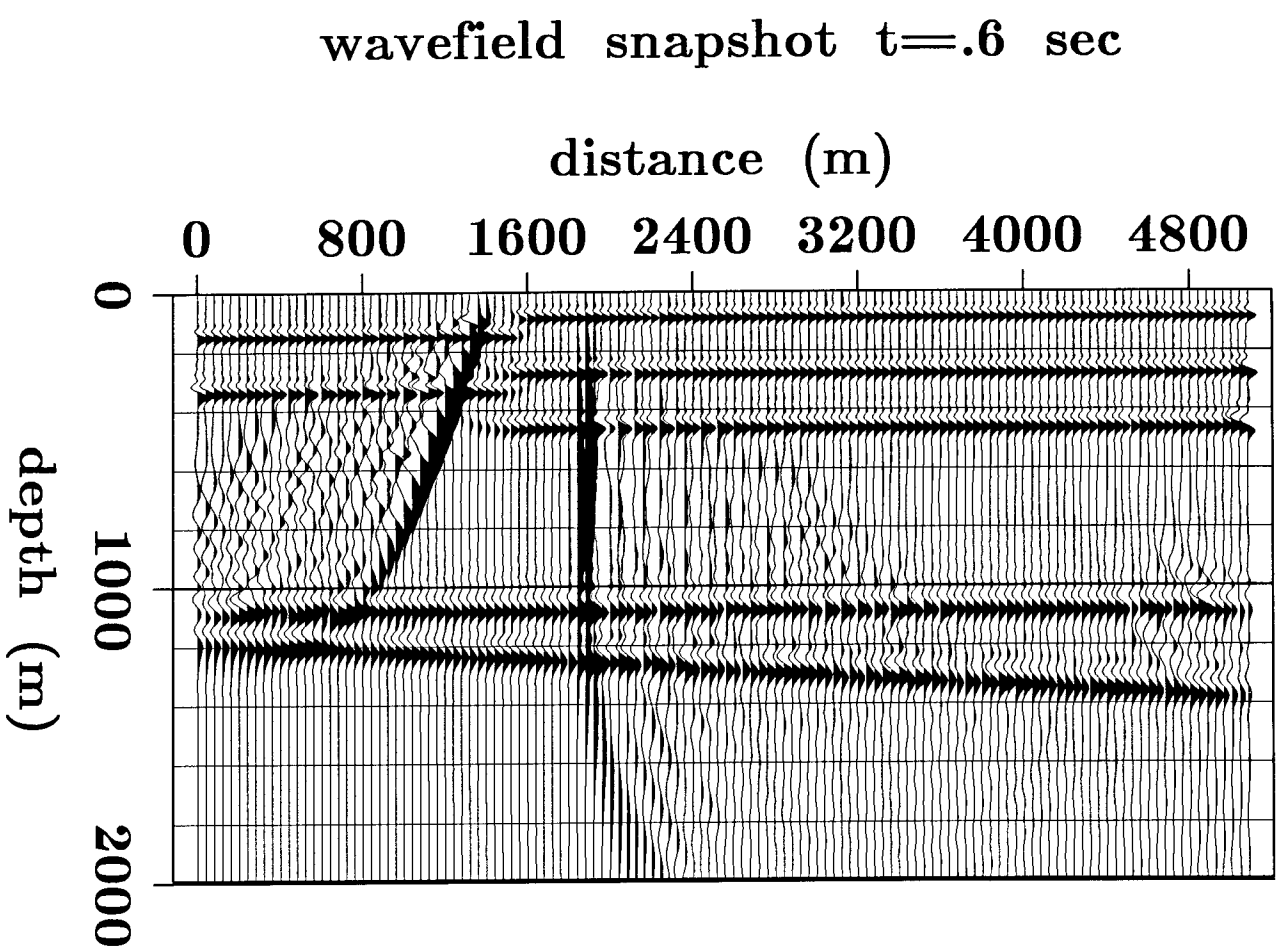

Figure 7: Snapshot of the wavefield in depth at time=.6 sec for the fourth-order in time, eighth-order in x, z algorithm. Note the vertical wave that will image the underside of the fault-plane reflection.

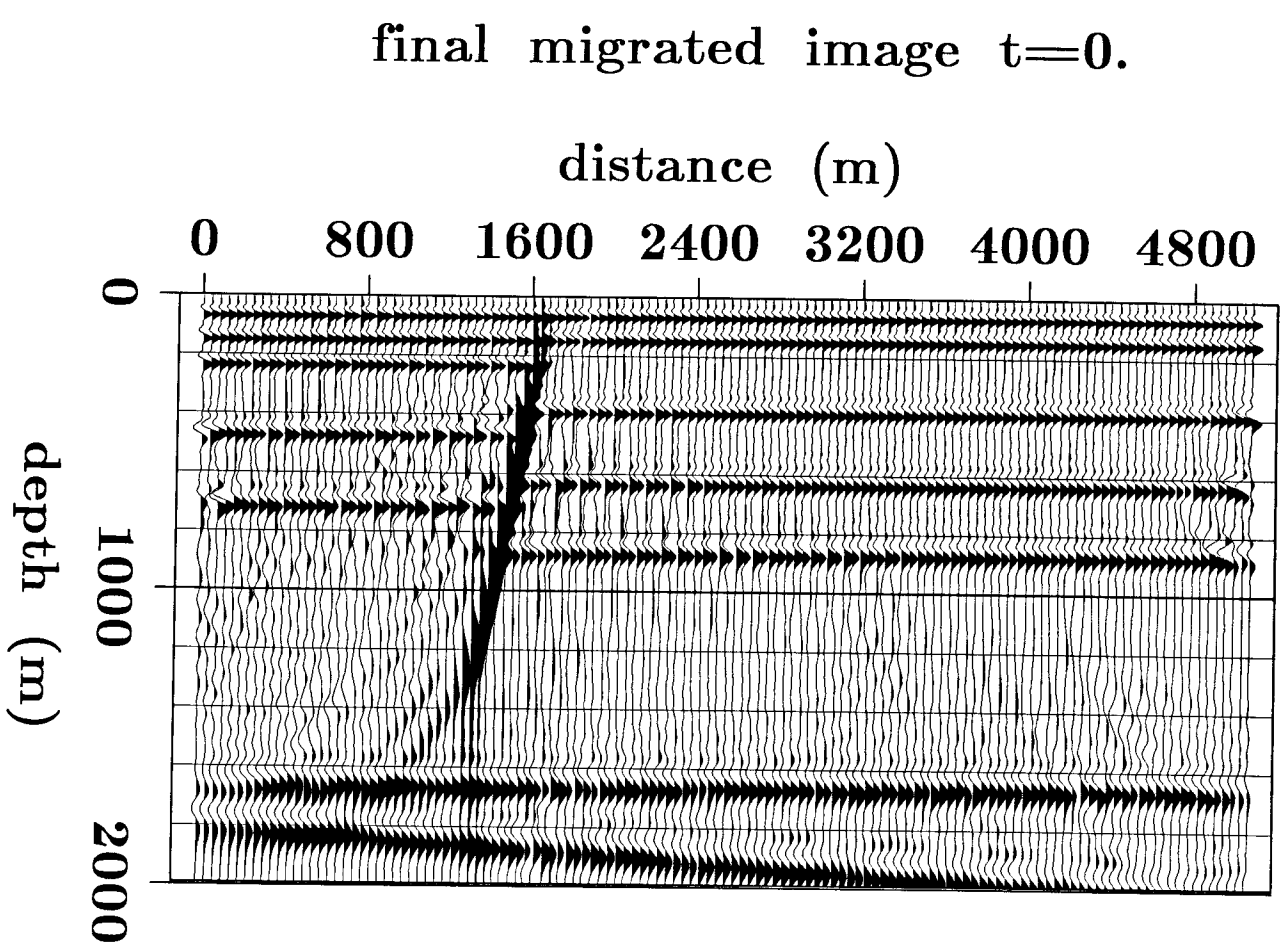

Figure 8: Final depth image of the wavefield at time=zero for the fourth-order in time eighth-order in x, z algorithm. The fault-plane is correctly imaged from both sides.

energy for a later procedure would fail to image the underside of the fault-plane correctly. Figure 8 is the final migrated depth image of the wavefield. The time equals zero imaging condition is applied now, and the position of the wave field is the position of the reflecting interfaces in the earth. Note that the underside of the steeply dipping reflector is imaged correctly. There are some artifacts present in each of the snapshots of the wavefield. The semicircular "smiles" are reflections off an incompletely absorbing boundary, or are due to the truncation of the data. There is also some grid dispersion from the waves that traveled horizontally. This is due to the fact that the sampling interval for the x axis was $40m$. The shortest wavelength in the middle of the section was approximately $66m$. The wavelengths are even shorter in the shallow part of the model due to the velocity gradient. The fact that the x axis is undersampled caused even the high-order operator to exhibit grid dispersion on the horizontally traveling waves . The flat dips exhibit little distortion since they are sampled at about 6 points per shortest wavelength

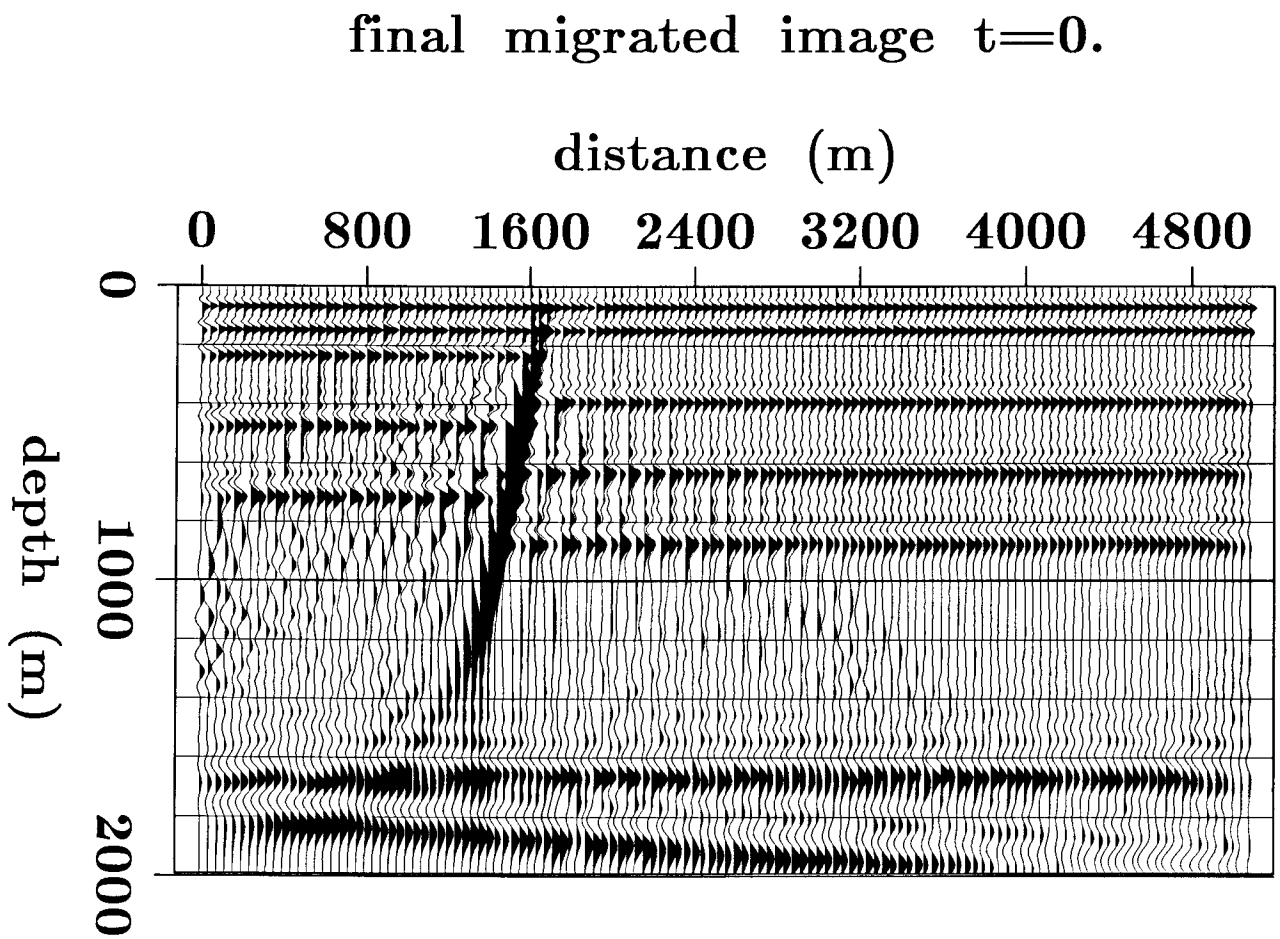

Figure 9: Final depth image at time=zero of the synthetic data for the second-order in time fourth-order in x, z algorithm. Compare to figure 8.

for the middle depths of the velocity model. It might seem puzzling that the algorithm does not break-down entirely at high frequencies for horizontally traveling waves, because the wavefield is sampled less than the nyquist criterion. I believe that the algorithm does not break-down, because of the peculiar nature of waves propagating through a turning point. The wavefronts are vertical at the turning point, but they propagate mostly in the vertical direction. All of the time, the wave has a significant component of vertical motion, so the wave is better sampled than the 1.5 points per wavelength implied by the horizontal sampling. Figure 9 is the final depth image obtained from the second-order in time, fourth-order in x, z algorithm. The result is very similar to Figure 8; there is slightly more grid dispersion, but the image is the same.

Computational highlights

The reverse time migration process with the *two-way non-reflecting wave equation* is highly

vectorizable on our **Convex C-1**. The calculations mainly consist of many convolutions with the derivative operators. To run the migration of the fault-plane model on a 128 by 200 grid required 171 CPU seconds for the second-order in time fourth-order in x, z algorithm. Counting both adds and multiplies, this translates to a little over 15 million floating point operations per second (Mflops). A similar problem on a 168 by 220 grid with 2048 time steps took the fourth-order in time eighth-order in x, z algorithm 406 CPU seconds. Once again, counting adds and multiplies, this translates to 15 Mflops as well. The reason the Mflops rate seems so high for the more complicated algorithm is that one grid-point update requires about twice as many computations compared to the second-order in time fourth-order in x, z.

CONCLUSIONS

High-order finite-difference operators offer the advantage of more faithful propagation of wavefields, while still retaining the simplicity of second-order operators. Course grid calculations with high-order finite-difference operators enable us to calculate more complicated and bigger models, more accurately. The *2-way non-reflecting wave equation* and reverse time migration provide a powerful tool to image the earth's subsurface in areas of complicated structure and heterogeneous velocity.

I would like to thank Zhiming Li for use of the data set with the overturned reflections, and Biondo Biondi for his critical reading of this paper.

REFERENCES

Baysal, E., Kosloff, D.D. and Sherwood, J.W.C., 1984, A two-way non-reflecting wave equation: Geophysics, **49**, 132-141.

Baysal, E., Kosloff, D.D. and Sherwood, J.W.C., 1983, Reverse time migration: Geophysics, **48**, 1514-1524.

Claerbout J. F., 1985, Imaging the earth's interior: Blackwell Scientific Publications., 47-49.

Dablain, M., 1986, The application of high-order differencing to the scalar wave equation: Geophysics, **51**, 54-66.

Levin, S., 1983, Principle of reverse time migration: SEP-37 ,37-42.

Li, Zhiming, Claerbout, J.F., Ottolini, R.A., 1984, Overturned- wave migration by two-way extrapolation: SEP-38, 141-150.

Whitmore, N.D., 1984, Interactive Depth Migration by Backward time propagation: Presented at the 53rd annual International SEG meeting September 13, 1984, in Las Vegas.