GEOPHYSICS, VOL. 51, NO. 1 (JANUARY 1986); P. 54-66, 13 FIGS., 1 TABLE.

The application of high-order differencing to the scalar wave equation

M. A. Dablain*

ABSTRACT

The second-order central difference is often used to approximate the derivatives of the wave equation. It is demonstrated that gains in computational efficiency can be made by using high-order approximations for these derivatives. A one-dimensional model is used to illustrate the relative accuracy of $O(\Delta t^2, \Delta x^2)$, $O(\Delta t^2, \Delta x^4)$, $O(\Delta t^4, \Delta x^4)$, and $O(\Delta t^4, \Delta x^{10})$ central-difference schemes. For comparison, $O(\Delta t^2)$ and $O(\Delta t^4)$ pseudospectral schemes are used as an additional measure of performance. The results indicate that $O(\Delta t^4, \Delta x^{10})$ differencing can achieve similar accuracy as the $O(\Delta t^4)$ spectral scheme. For practical illustration, a two-dimensional form of the $O(\Delta t^4, \Delta x^{10})$ algorithm is used to compute the exploding reflector response of a salt-dome model and compared with a fine-grid $O(\Delta t^2, \Delta x^4)$ result. Transmissive sponge-like boundary conditions are also examined and shown to be effective.

INTRODUCTION

Finite-difference solutions of the wave equation and its parabolic approximation have proven useful tools in exploration seismology. In the forward mode (Kelly et al., 1976), the wave equation can be used as an interpretive aid in complex geology or as a research tool for testing processing algorithms. In the inverse mode, differencing techniques can be applied to the parabolic form of the wave equation for the purpose of reflector mapping (Claerbout, 1976). As a deterministic processing tool or as an interpretive aid, the finite-difference form of the scalar wave equation is a versatile algorithm. The direct large-scale application of this technique has met with resistance, however, because of its intensive use of central processing unit (CPU) time and its need for large amounts of direct-access memory. Recently, these constraints have become more manageable with the use of supercomputers. The supercomputers generally offer large amounts of in-core memory, a fast clock-cycle time, and vector instructions. Although supercomputers make the task of numerical modeling appear less formidable, the standard second-order operator applied to large-scale seismic models still represents a challenge. A simple example for the CRAY 1 supercomputer is enough to illustrate this point.

Consider a two-dimensional (2-D) earth model, 10 000 m × 3 000 m, with a velocity range from 1 800 m/s to 4 500 m/s. Using a second-order difference operator and a 25 Hz Ricker wavelet source, an exploding reflector synthetic would require about 2.25 m sampling. At this sampling rate a general velocity model would require 18 million words of storage. This is well outside the range of in-core computations for a CRAY 1. Moreover, even if the CRAY had this resident memory available, it would require about 1.7 hours of CPU time for a second-order scheme to complete this model, assuming that 3.0 s of data are collected. This is a great deal of computational power. To put this performance into perspective, a VAX 11/780 would compute this same problem about 180 to 200 times more slowly, or in 13 days of CPU time. With this perspective it is seen that large-scale, wave-equation computations are still challenging. One possible solution for this problem can be found by recognizing that the standard second-order difference operator requires a heavy oversampling of the model. Kosloff and Baysal (1982) proposed a solution to this oversampling with the use of a pseudospectral operator. In theory the pseudospectral scheme should reduce the sampling rate from about 8 grid points per wavelength at the Nyquist frequency to 2 points. Because the pseudospectral scheme represents a theoretical lower bound on the sampling rate, it is used here as a baseline for comparison with high-order difference schemes.

As a prelude to this study, a few comments are appropraite concerning the large-scale application of discrete numerical methods. Because of the practical constraints inherent in the implementation of large-scale computing methods, judgments with regard to the efficiency of an algorithm must consider both the CPU demands and the memory demands of an algorithm. For example, a high-order differencing scheme which breaks even in computing time with a low-order method may be judged more efficient if its memory savings are significant. In general, an out-of-core solution to a problem is slower than its in-core version, unless the data transfers can be overlapped

Manuscript received by the Editor November 19, 1984; revised manuscript received June 14, 1985.
*Formerly Mobil Research and Development Corporation, 13777 Midway Road, Dallas, TX 75234; currently Exxon Production Research, P. O. Box 2189, Houston, TX 77001.

with the computing. The faster an algorithm or an arithmetic unit, the more difficult it becomes to hide the data transfers behind the computing. In fact, this problem is compounded when applying finite-difference solutions to the scalar wave equation. The faster an algorithm, i.e., the shorter the operator, the more oversampling in the problem, and hence more data must be transferred per grid update. Because of these real constraints in large-scale computing, a good algorithm balances input/output and memory needs with the speed of the arithmetic unit. Within this context, note that simply counting floating-point operations as a measure of efficiency with no regard to storage needs or the speed of the available input/output devices can result in misleading estimates for algorithm performance.

TEMPORAL AND SPATIAL DISPERSION

Grid dispersion is the most significant numerical problem limiting the usefulness of point-wise discretization schemes for the wave equation. As its name indicates, this numerical artifact causes the phase speed to become a function of the discretization interval. The relative computational merit of most discretization schemes hinges on their ability to minimize this effect. Fortunately, in exploration seismology the observations of interest are linear and band-limited. It is always possible, therefore, to design the sampling rate so the numerical dispersion is outside the band of interest. In theory, the only limitation with regard to this sampling rate is the finite storage capacity of the computational machine and the patience of the user. In practice, the limitations on this sampling rate are the computational performance of the arithmetic unit and the availability of high-speed memory. From a practical viewpoint it is optimal to use the maximum allowed sampling rate $2\Delta x$ on the minimum wavelength of interest. Numerical dispersion, however, usually forces the use of about $8\Delta x$ to cover the Nyquist frequency. The requirement of $8\Delta x$ at Nyquist is a function of the low accuracy with which the standard second-order difference scheme approximates derivatives. The exact cause for this low accuracy is illustrated in a simple but effective way.

The one-dimensional (1-D) scalar wave equation may be written

$$\frac{\partial^2 \phi}{\partial t^2} = c_0^2 \frac{\partial^2 \phi}{\partial x^2}, \tag{1}$$

where

c_0 is the sound speed and ϕ is a pressure-like variable.

A second-order difference of the time portion of equation (1) would discretize as

$$\frac{\phi^{n+1} - 2\phi^n + \phi^{n-1}}{\Delta t^2} = c_0^2 \frac{\partial^2 \phi}{\partial x^2} \tag{2}$$

where

$$\phi^n = \phi(x, n\Delta t).$$

Substituting the harmonic form $e^{i(kx+\omega t)}$ into equation (2) produces

$$\frac{e^{i\omega\Delta t} + e^{-i\omega\Delta t} - 2}{\Delta t^2} = -c_0^2 k^2,$$

or

$$\frac{2[\cos(\omega\Delta t) - 1]}{\Delta t^2} = -c_0^2 k^2. \tag{3}$$

The phase speed c is by definition (Whitham, 1974)

$$c = \frac{\omega}{k}. \tag{4}$$

The definition of the phase speed in equation (4) is now substituted into equation (3) to give the dispersion relationship for temporal discretization.

$$2[\cos(\omega\Delta t) - 1] = \frac{-\omega^2}{c^2} c_0^2 \Delta t^2,$$

or

$$\frac{c^2}{c_0^2} = \frac{-(\omega\Delta t)^2}{2[\cos(\omega\Delta t) - 1]}, \tag{5}$$

or

$$\frac{c^2}{c_0^2} = \frac{-\theta^2}{2(\cos\theta - 1)},$$

where

$$\theta = \omega\Delta t.$$

The behavior of the relative phase speed c/c_0 in equation (5) is of interest, especially when compared to its spatial counterpart. Ideally, the relative phase speed should be unity in a region of constant sound speed. It is evident, however, in equation (5) that this is not true in the discretized case. Theoretically, the time step Δt can be designed such that it adequately samples the highest frequency in the band-limited signal. In this case θ would range between 0 and π. In practice, however, sampling can be done at finer intervals than that dictated by the Nyquist sampling limit. This finer sampling with regard to the temporal sampling has very important practical consequences. It is generally preferable to design difference schemes for the wave equation which require finer temporal sampling but allow coarse spatial sampling. This is because the computational resources (CPU time and memory) needed for finer spatial sampling increase geometrically, whereas the resources for finer temporal sampling increase only linearly.

Returning to the discussion of relative phase speed, the gain of some intuitive feeling for the consequences of equation (5) is useful. Assuming that low frequencies or very small time steps are taken, i.e., $\theta \sim 0$, then equation (5) can be approximated by

$$\frac{c^2}{c_0^2} = \frac{-\theta^2}{\left(-\frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \cdots\right)} \simeq 1 + \frac{\theta^2}{12}. \tag{6}$$

Equation (6) indicates that as Δt increases or the frequencies become greater, the tendency is for parts of the wavelet to be advanced with respect to the true phase speed. Because it is the higher frequencies which are advanced, this behavior may be identified as anomalous dispersion. In the numerical exam-

ples which follow, I show that this dispersion indeed dominates when the temporal error exceeds the spatial error.

Returning to equation (1), the same analysis is performed for a second-order difference applied to the spatial portion of the wave equation. This difference is written

$$\frac{\partial^2\phi}{\partial t^2} = \frac{c_0^2}{\Delta x^2}(\phi_{i+1} - 2\phi_i + \phi_{i-1}), \tag{7}$$

where

$$\phi_i = \phi(i\Delta x, t).$$

Substituting the harmonic form $e^{i(kx+\omega t)}$ into this equation results in

$$-\omega^2 = \frac{2c_0^2}{\Delta x^2}[\cos(k\Delta x) - 1]. \tag{8}$$

By applying the definition in equation (4), the expression for the relative phase speed is

$$\frac{c^2}{c_0^2} = \frac{-2}{\psi^2}(\cos\psi - 1), \tag{9}$$

where

$$\psi = k\Delta x.$$

For very fine spatial sampling or long wavelengths, $\psi \sim 0$ and equation (9) may be approximated as

$$\frac{c^2}{c_0^2} = \frac{2}{\psi^2}\left(-\frac{\psi^2}{2!} + \frac{\psi^4}{4!} - \frac{\psi^6}{6!} + \cdots\right) \simeq 1 - \frac{\psi^2}{12}. \tag{10}$$

Comparing the relative phase speeds as derived in equations (6) and (10), note that spatial dispersion acts in the opposite sense as temporal dispersion. A difference scheme dominated by spatial dispersion therefore delays the higher frequencies. In summary, the numerical dispersion in a scheme dominated by temporal error leads the signal, whereas the dispersion in a signal dominated by spatial error follows the signal. This behavior is evident in the numerical examples.

ONE-DIMENSIONAL COMPARISONS

A pragmatic approach was used to investigate the relative performance of high-order difference schemes. A representative test case was formulated in one dimension and then computed with a variety of difference and spectral-based algorithms. The difference algorithms used explicit central differencing of the following orders of accuracy: $O(\Delta t^2, \Delta x^2)$, $O(\Delta t^2, \Delta x^4)$, $O(\Delta t^4, \Delta x^4)$, and $O(\Delta t^4, \Delta x^{10})$. The exact form of these algorithms is discussed in Appendix A. The spectral-based algorithms applied a fast Fourier transform to compute the spatial derivatives and used a central-difference scheme for the temporal integration. The temporal differences were $O(\Delta t^2)$ and $O(\Delta t^4)$. The velocity model used for this comparative study is illustrated in Figure 1. The distance from the source to the farthest receiver is 5 203 m. The total model is 10 406 m long. All other parameters of significance are

$\Delta x = 10.16$ m,

Minimum velocity = 1 524 m/s,

Maximum velocity = 3 048 m/s,

Input source = Ricker wavelet,

Dominant frequency = 25 Hz,

Source position = 2 602 m,

Receiver positions = 3 902, 6 504, and 7 805 m,

and

Total data length = 3.5 s.

The temporal sampling rate was scheme-dependent due to considerations of stability and accuracy. For comparison, each wavelet was normalized to its maximum amplitude at receiver 1. Also, for detailed display purposes, a time window of approximately 200 ms was used around the arrival time of the wavelet at each receiver.

Because the objective of this study is relative performance, it is not particularly significant how the Nyquist frequency of the source is defined, as long as it is done consistently. For the purposes of this investigation the Nyquist frequency of the source was taken to be twice the dominant frequency. In this case that implies a Nyquist frequency of 50 Hz. The rule of thumb in difference schemes for choosing an appropriate sampling rate based on the Nyquist frequency can be written

$$\Delta x = \frac{c_{\min}}{f_N \cdot S}, \tag{11}$$

where

f_N = Nyquist frequency,
$c_{\min}$ = minimum sound speed, and
S = number of points needed to cover the Nyquist frequency for nondispersive propagation.

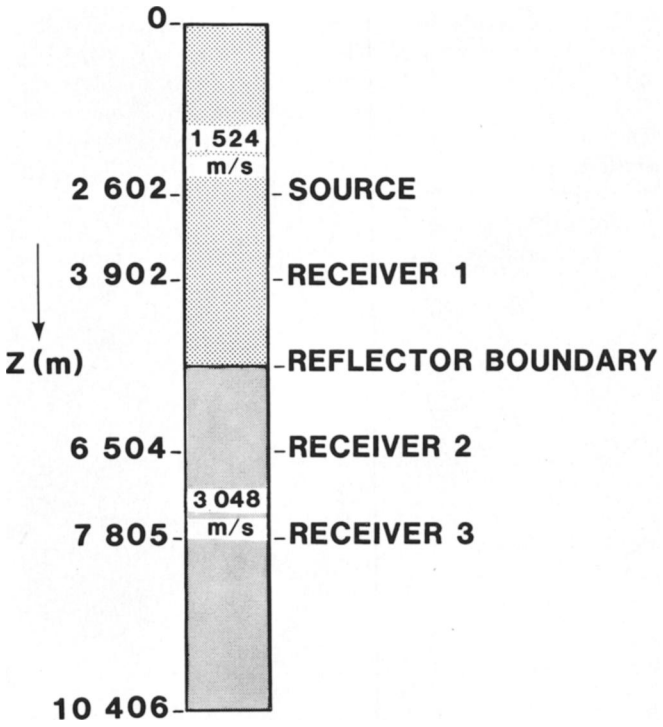

FIG. 1. One-dimensional model for difference and spectral comparisons.

For a Ricker wavelet source this rule of thumb is an effective criterion for nondispersive propagation. The quantity S is scheme-dependent. It has been found through numerical experience that $S = 8$ for $O(\Delta t^2, \Delta x^2)$, $S = 4$ for $O(\Delta t^2, \Delta x^4)$, and $S = 3$ for $O(\Delta t^4, \Delta x^{10})$. Two points should be made with regard to S: (1) these S values assume a reasonable amount of propagation in the lowest velocity medium; and (2) any discretization scheme with dispersive properties can eventually be made to display a dispersed wavelet. These S values are, therefore, quoted for propagation on the order of 30–40 wavelengths in the lowest velocity medium and for accuracies which are acceptable for the interpretive testing of large-scale models. Although the absolute value of S depends upon the fitting frequency, the source, and the application, there should be an inner consistency with regard to relative performance. Therefore, note that the relative performance indicated by the second-order spatial scheme to the fourth-order spatial scheme is essentially 2 : 1. This means that by using twice as large a sampling rate as the second-order scheme, similar accuracies should be achieved with the fourth-order scheme. This 2 : 1 performance agrees well with the numerical examples presented in Alford et al. (1974). Given that these S values are most useful as relative numbers, the value $S = 3$ is used for this model study. This choice is based on the advantage of hindsight which indicates that $S = 3$ is the minimum sampling requirement for comparable performance between the $O(\Delta t^4)$ pseudospectral scheme and the $O(\Delta t^4, \Delta x^{10})$ difference scheme.

Figure 2 shows the primary arrivals at receivers 1, 2, and 3. These arrivals are taken to be representative of the exact solution and were generated by refining the grid to $6\Delta x$ at Nyquist and using the $O(\Delta t^4, \Delta x^{10})$ scheme. The grid sampling in the refined example was 5.08 m. That the finite-difference solution does indeed converge on the analytical result is verified by realizing that the solution of the 1-D wave equation with a source is the first integral of the source shifted by $x = ct$ (Chester, 1971). Figure 3 shows the first integral of the Ricker wavelet source, which was computed from its analytical expression. It agrees very well with the wave recorded at receiver 1. Receivers 2 and 3 should, in theory, be the input wave multiplied by the transmission coefficient, which in this case is 4/3. Again, the results on the refined grid at receivers 2 and 3 show good agreement with the expected theoretical result. These considerations justify use of this experiment as a baseline for comparison, and at the same time verify the convergence of the $O(\Delta t^4, \Delta x^{10})$ algorithm. Figures 4 and 5 show the arrivals at receivers 1, 2, and 3 for the $O(\Delta t^2, \Delta x^2)$ and the $O(\Delta t^2, \Delta x^4)$ difference schemes on the coarse grid ($\Delta x = 10.16$ m). The second-order spatial difference is clearly unacceptable at this sampling rate. The nature of the numerical dispersion is made clear by noting that the higher frequencies are trailing the signal. An improvement with the fourth-order scheme is evident, but the waveform is still dispersive. The dispersion is again of the normal type, indicating that the spatial error is dominating the phase speed. The temporal sampling in the $O(\Delta t^2, \Delta x^2)$ example was

$$\Delta t = \frac{\Delta x}{c_{\max}} .95, \tag{12}$$

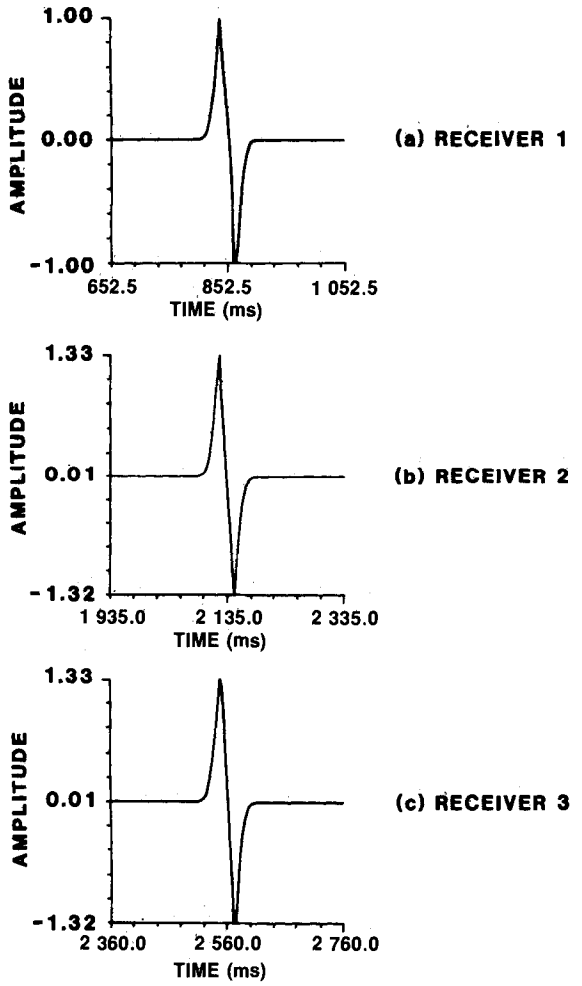

FIG. 2. The fine-grid solution $O(\Delta t^4, \Delta x^{10})$ used to verify convergence.

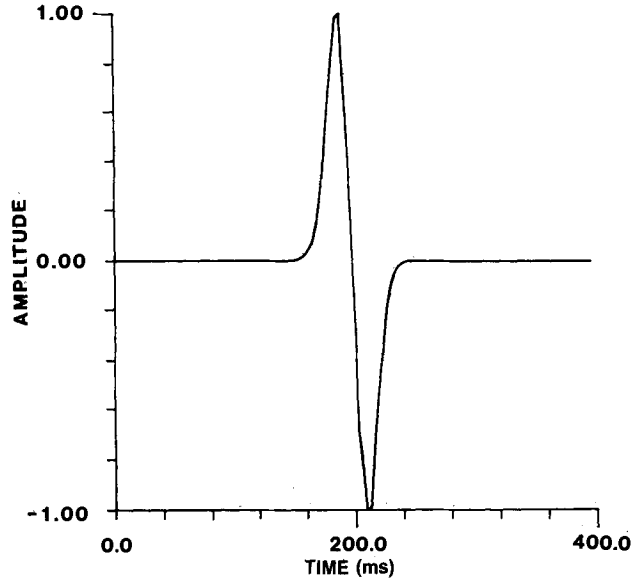

FIG. 3. The first integral of the Ricker wavelet, confirming the free-space solution of the 1-D wave equation as an integration of the source.

where c_{max} is the maximum sound speed. The reason for this is that the $O(\Delta t^2, \Delta x^2)$ scheme is error free at $c\Delta t/\Delta x = 1$ (Mitchell and Griffiths, 1980) which means that the propagation path through the highest velocity material would not have been representative of general scheme performance. The temporal sampling rate for the $O(\Delta t^2, \Delta x^4)$ scheme was chosen to be

$$\Delta t = \sqrt{3/4}\ .95\ \frac{\Delta x}{c_{max}}. \tag{13}$$

The factor $\sqrt{3/4}$ is the ratio of the stability limit for the $O(\Delta t^2, \Delta x^4)$ to that of the $O(\Delta t^2, \Delta x^2)$ scheme in two dimensions. The objective was to choose a sampling rate indicative of practical performance in applications. Figure 6 shows the results for the $O(\Delta t^4, \Delta x^4)$ scheme, which was computed at $\Delta t = .95\ \Delta x/c_{max}$. The signal is dispersive and seems to be worse by comparison with the $O(\Delta t^2, \Delta x^4)$ result. This impression, however, is misleading. A closer examination of the signal in Figure 6 indicates that the waveform from receivers 1 to 3 remains essentially the same. In fact, the waveforms at receivers 2 and 3 are indistinguishable. This suggests that propagation through the higher-velocity region was dispersion-free. This is not the case with the $O(\Delta t^2, \Delta x^4)$ scheme. A comparison between receivers 2 and 3 in this case indicates further distortion in both amplitude and phase. The relatively dispersion-free propagation in the higher velocity region can also be noted in the $O(\Delta t^2, \Delta x^2)$ result of Figure 4. Although analysis has not been performed, these observations suggest that by differencing the temporal and spatial derivatives to the same order of accuracy, the stability limit is pushed toward the maximum Courant number $c\Delta t/\Delta x = 1$, where there is typically less dispersion.

Figure 7 shows the results for the $O(\Delta t^2)$ spectral scheme, discussed in Kosloff and Baysal (1982). The temporal sampling rate was chosen by specifying $c_{max}\Delta t/\Delta x = .5$. Because the Fourier transform is used to compute the spatial derivative, $2\Delta x$ is expected to be sufficient at Nyquist. The problem is that the $O(\Delta t^2)$ temporal differencing defines an anomalous dispersion relationship. This causes the high frequencies to lead the signal as is apparent in Figure 7. Equation (5) indicates that this anomalous dispersion can be controlled by decreasing the temporal sampling rate. Finally, Figures 8 and 9 show the results for the $O(\Delta t^4, \Delta x^{10})$ differencing and the $O(\Delta t^4)$ spectral scheme. The $O(\Delta t^4)$ spectral scheme was computed with $c_{max}\Delta t/\Delta x = .5$ and the $O(\Delta t^4, \Delta x^{10})$ difference scheme with $c_{max}\Delta t/\Delta x = .75$. The results are equivalent for the practical performance expected from large-scale interpre-

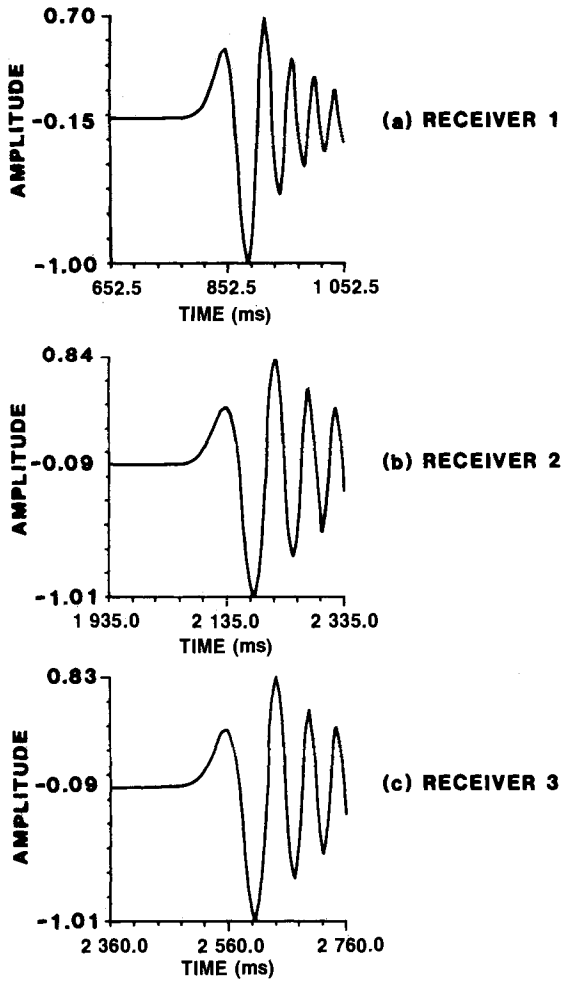

FIG. 4. A coarse-grid, finite-difference solution $O(\Delta t^2, \Delta x^2)$ of the 1-D model.

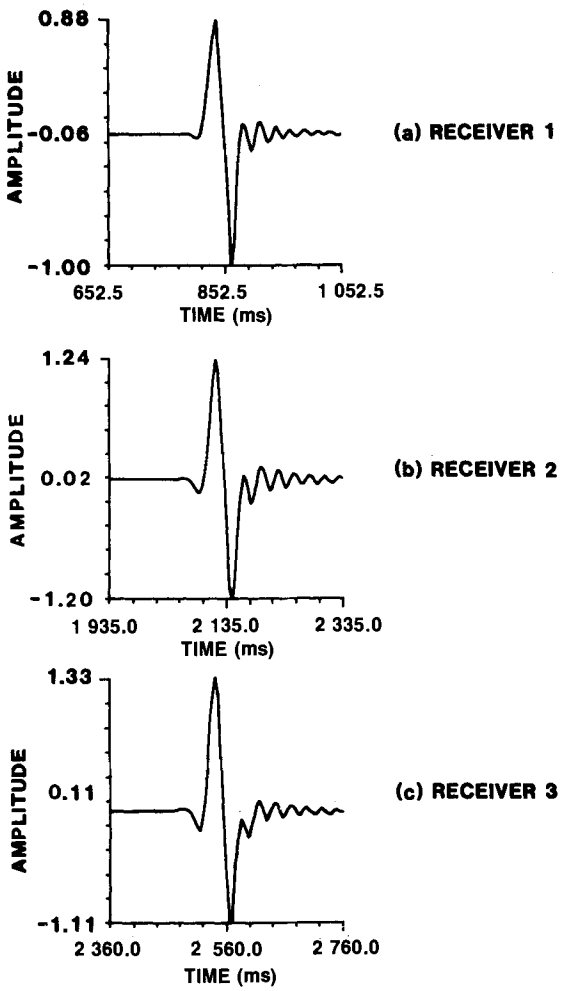

FIG. 5. A coarse-grid, finite-difference solution $O(\Delta t^2, \Delta x^4)$ of the 1-D model.

tive modeling. The agreement between the difference and the spectral schemes suggests that the largest part of the remaining error is due to temporal rather than spatial inaccuracies. In both cases a slight broadening in the negative half of the wavelet is noted, suggesting that the higher frequencies are still leading the signal. The results of Figures 8 and 9 suggest that $O(\Delta t^4, \Delta x^{10})$ differencing will give comparable results to the $O(\Delta t^4)$ spectral scheme for many *practical* applications. The use of *practical* is emphasized here.

TRANSMISSIVE SPONGE BOUNDARIES

When implementing discrete modeling schemes for the wave equation, one consideration is the simulation of infinite boundaries on a finite computational grid. A variety of nonreflecting numerical boundaries have been discussed in Lysmer and Kuhlmeyer (1969), Smith (1974), and Clayton and Engquist (1977). In the high-order difference schemes discussed here, the method is that of Israeli and Orszag (1981). Like others, the idea of Israeli and Orszag was to approximate radiation conditions at the numerical boundaries; however, it was recognized that these transmission operators would be imperfect and would result in reflected waves. Their contribution to the solution of this problem was to propose the transmission operator in such a way that it radiated in one direction and damped in the opposite, thereby attenuating any residual boundary reflections. The simplest operator designed to implement this idea is written in one dimension as

$$\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = \frac{\partial^2\phi}{\partial x^2} - \frac{\varepsilon}{c}\left(\frac{\partial\phi}{\partial t} + c\frac{\partial\phi}{\partial x}\right), \tag{14}$$

with

$$\varepsilon > 0.$$

Intuition suggests that when the wave propagation is to the right, $\phi(x - ct)$, then $\partial\phi/\partial t + c(\partial\phi/\partial x)$ vanishes and the wave equation is left intact. When the wave propagation is to the left, however, then $\phi = \phi(x + ct)$ and

$$\frac{\partial\phi}{\partial t} = c\frac{\partial\phi}{\partial\xi},$$

$$\frac{\partial\phi}{\partial x} = \frac{\partial\phi}{\partial\xi},$$

where

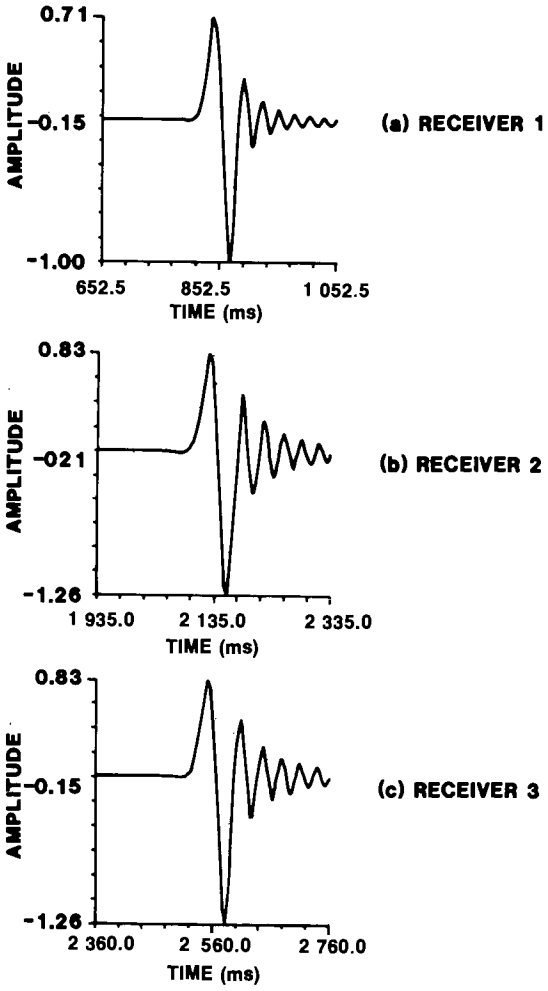

FIG. 6. A coarse-grid, finite-difference solution $O(\Delta t^4, \Delta x^4)$ of the 1-D model.

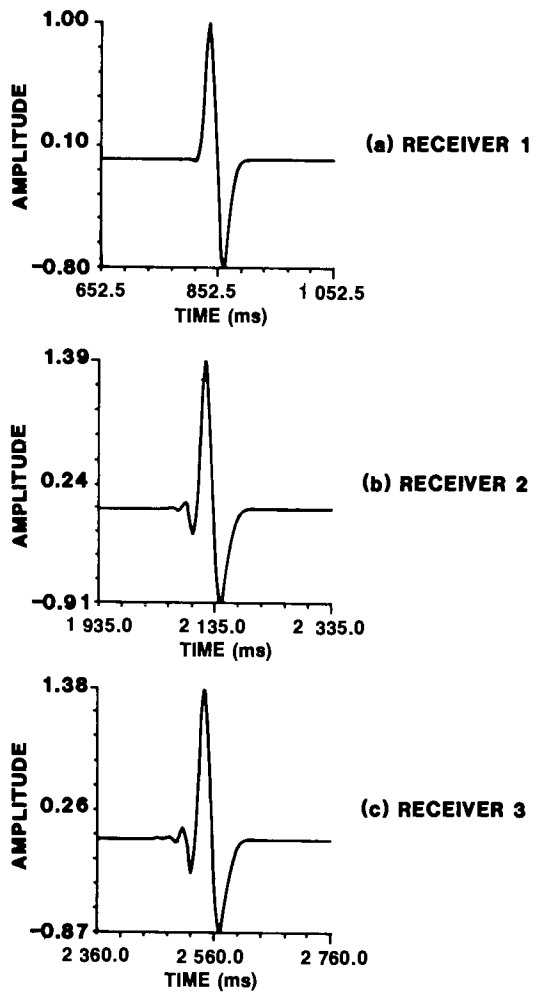

FIG. 7. A coarse-grid, pseudospectral solution $O(\Delta t^2)$ of the 1-D model.

$$\xi = x + ct,$$

leaving an error term of the form

$$2c\frac{\partial \phi}{\partial x}. \tag{15}$$

The effect of adding a first-order derivative to the wave equation is that of a damping term. The result then of equation (14) is to transmit waves in one direction and attenuate them in the other.

Thus, to first order, equation (14) produces the desired effect. As suspected, this first-order behavior is not perfect and some energy leaks back into the solution. It is interesting to examine in what sense propagation is achieved in the outgoing direction and damping for the incoming waves. For an incoming wave the proposed solution is

$$\phi = \bar{\phi}(\omega, k)e^{i(kx+\omega t)}. \tag{16}$$

Substituting equation (16) into equation (14) produces the dispersion relationship

$$\left(\frac{\omega}{c}\right)^2 = k^2 + i\varepsilon\left(\frac{\omega}{c} + k\right),$$

or

$$\frac{\omega}{c} = k\left[1 + \frac{i\varepsilon}{k^2}\left(\frac{\omega}{c} + k\right)\right]^{1/2}.$$

By assuming that ε is small, equation (17) is approximated as

$$\frac{\omega}{c} = k\left[1 + \frac{i\varepsilon}{2k^2}\left(\frac{\omega}{c} + k\right)\right]. \tag{18}$$

Collecting ω/c on the left-hand side of equation (18) and keeping terms to first order in ε leaves

$$\frac{\omega}{c} \simeq k + i\varepsilon. \tag{19}$$

The approximate dispersion relationship in equation (19) is substituted into the assumed spectral solution in equation (16), so that

$$\phi = \bar{\phi}(\omega, k)e^{-\varepsilon ct}e^{ik(x+ct)}. \tag{20}$$

By construction, $\varepsilon > 0$ so equation (20) represents a damped wave propagating toward the left, i.e., away from the boundary. If a right-going wave is now proposed, it is shown to first order in ε that the dispersion relationship reduces to

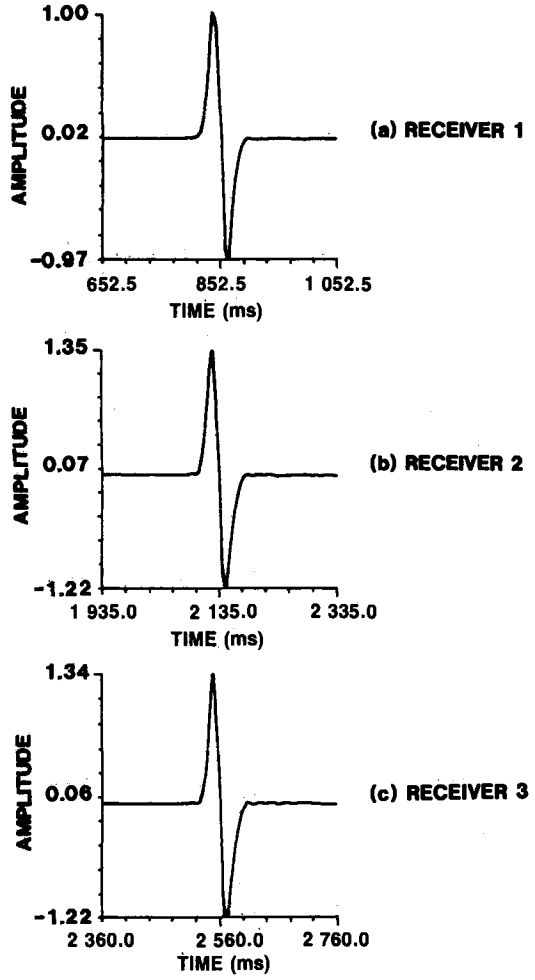

FIG. 8. A coarse-grid, finite-difference solution $O(\Delta t^4, \Delta x^{10})$ of the 1-D model.

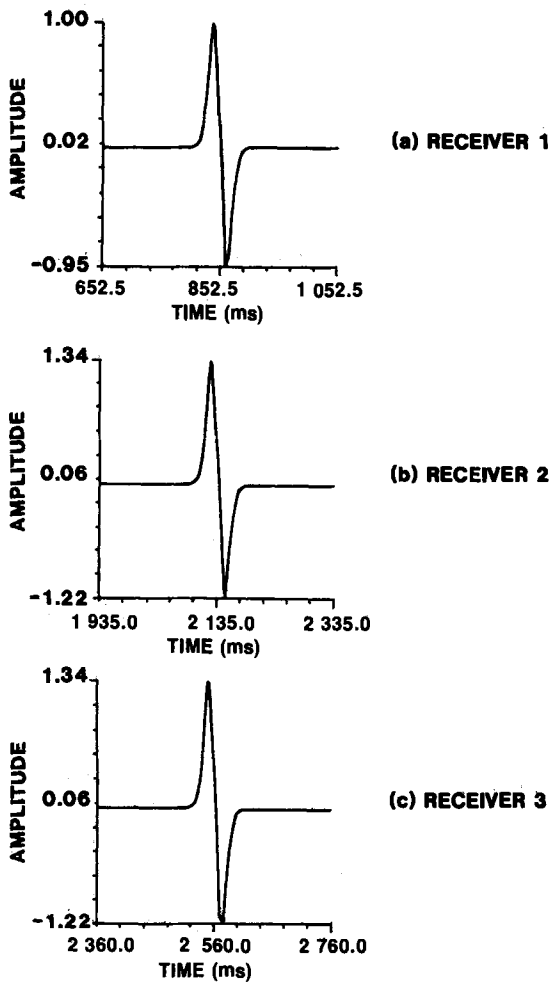

FIG. 9. A coarse-grid, pseudospectral solution $O(\Delta t^4)$ of the 1-D model.

$$\frac{\omega}{c} \simeq \frac{k - \frac{i\varepsilon}{2}}{1 - \frac{i\varepsilon}{2k}},$$

or (21)

$$\frac{\omega}{c} \simeq k + O(\varepsilon^2).$$

This indicates that the right-going wave, i.e., propagation into the sponge boundary, remains undisturbed at least to $O(\varepsilon^2)$. Implementation of these ideas in two dimensions is a straightforward application of the additional term in equation (14). The details of this implementation are provided in Appendix B.

EXAMPLE AND EFFICIENCY OF TWO-DIMENSIONAL, HIGH-ORDER DIFFERENCING

The $O(\Delta t^2, \Delta x^4)$ and $O(\Delta t^4, \Delta x^{10})$ difference schemes were implemented in two dimensions with the sponge boundaries of the preceding section. The algorithms were tested on a CRAY 1. To demonstrate the advantage of the higher-order $O(\Delta t^4, \Delta x^{10})$ scheme, a large model was computed and verified by comparison. The velocity model is indicated in Figure 10. The model extends 10 000 m laterally and 3 000 m deep and was intended to represent a fairly complex geology including pinchouts, grabens, a salt mass with an overhang, and truncated beds on the salt flanks simulating hydrocarbon potential.

In Figure 11 are the exploding reflector computations for this model. The exploding reflector model was initiated by lining the velocity boundaries with a Ricker wavelet source function. A Ricker wavelet with a dominant frequency of 50 Hz was used. The amplitude of each Ricker wavelet source was also weighted by the local value for normal incidence reflectivity. The exploding reflector response was computed two ways. Figure 11a is the result for the $O(\Delta t^2, \Delta x^4)$ algorithm and Figure 11b for the $O(\Delta t^4, \Delta x^{10})$ algorithm. The two results agree in great detail, thus verifying the adequacy of the spatial sampling rates which were 6.25 m and 8.33 m, respectively. It is satisfying to see that the algorithms converge as predicted by theory when sampled properly. There is, however, an important practical distinction between these two results. The $O(\Delta t^2, \Delta x^4)$ algorithm could not compute the entire model within the available 1.8 megawords of storage due to its finer sampling rate. A scaled down version of the original model was therefore computed. To compute the full model, the memory requirement for the $O(\Delta t^2, \Delta x^4)$ scheme was approximately 2.3 Mwords. The memory requirement for the $O(\Delta t^4, \Delta x^{10})$ scheme was approximately 1.3 Mwords. This amounts to a savings for this example of 1 million words of storage. The significance of this savings is even more dramatic if it is extrapolated to three dimensions and this economy is multiplied by the possible 200 or so extra planes of grid data which would be needed.

Analysis of CPU efficiency for these schemes was carried out in a very empirical way. A number of large problems were run with the $O(\Delta t^2, \Delta x^2)$, $O(\Delta t^2, \Delta x^4)$, and $O(\Delta t^4, \Delta x^{10})$ schemes to derive a reliable parameter which I refer to as the whiz W factor for the scheme. The whiz factor is derived by taking the sum of the observed CPU times for the difference scheme plus boundary conditions and normalizing this sum to a single grid cell carried through one update cycle or time step. Representative whiz factors on a CRAY 1 for $O(\Delta t^2, \Delta x^2)$, $O(\Delta t^2, \Delta x^4)$, and $O(\Delta t^4, \Delta x^{10})$ difference schemes with the sponge boundaries discussed were found to be

$$W(\Delta t^2, \Delta x^2) = 1.7 \times 10^{-7} \quad \text{CPU } s\text{/cell-cycle},$$

$$W(\Delta t^2, \Delta x^4) = 2.5 \times 10^{-7},$$

and

$$W(\Delta t^4, \Delta x^{10}) = 5.5 \times 10^{-7}.$$

To determine the CPU efficiency of a scheme, consider the total number of operations needed to achieve a given accuracy for a given model. This estimate should include any changes in the stability criteria needed to achieve the desired result. Computational experience has shown that changes in the stability criteria for the three schemes considered are not of practical significance. Based on the empirically derived whiz factors, the total CPU times (TCPU) for $S = 8$, $O(\Delta t^2, \Delta x^2)$, $S = 4$, $O(\Delta t^2,$

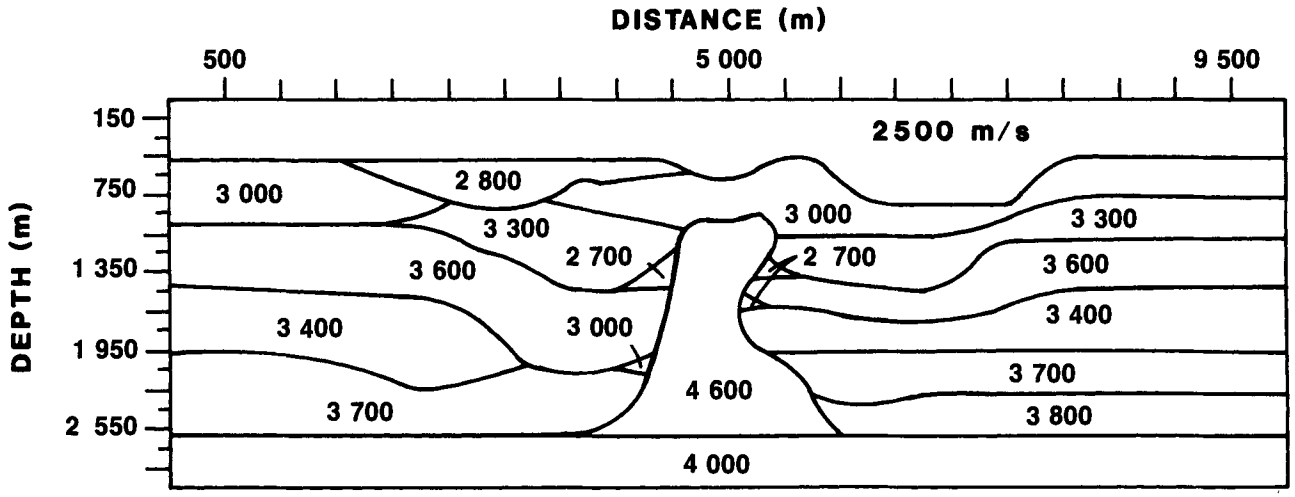

FIG. 10. A 2-D model for high-order difference comparisons.

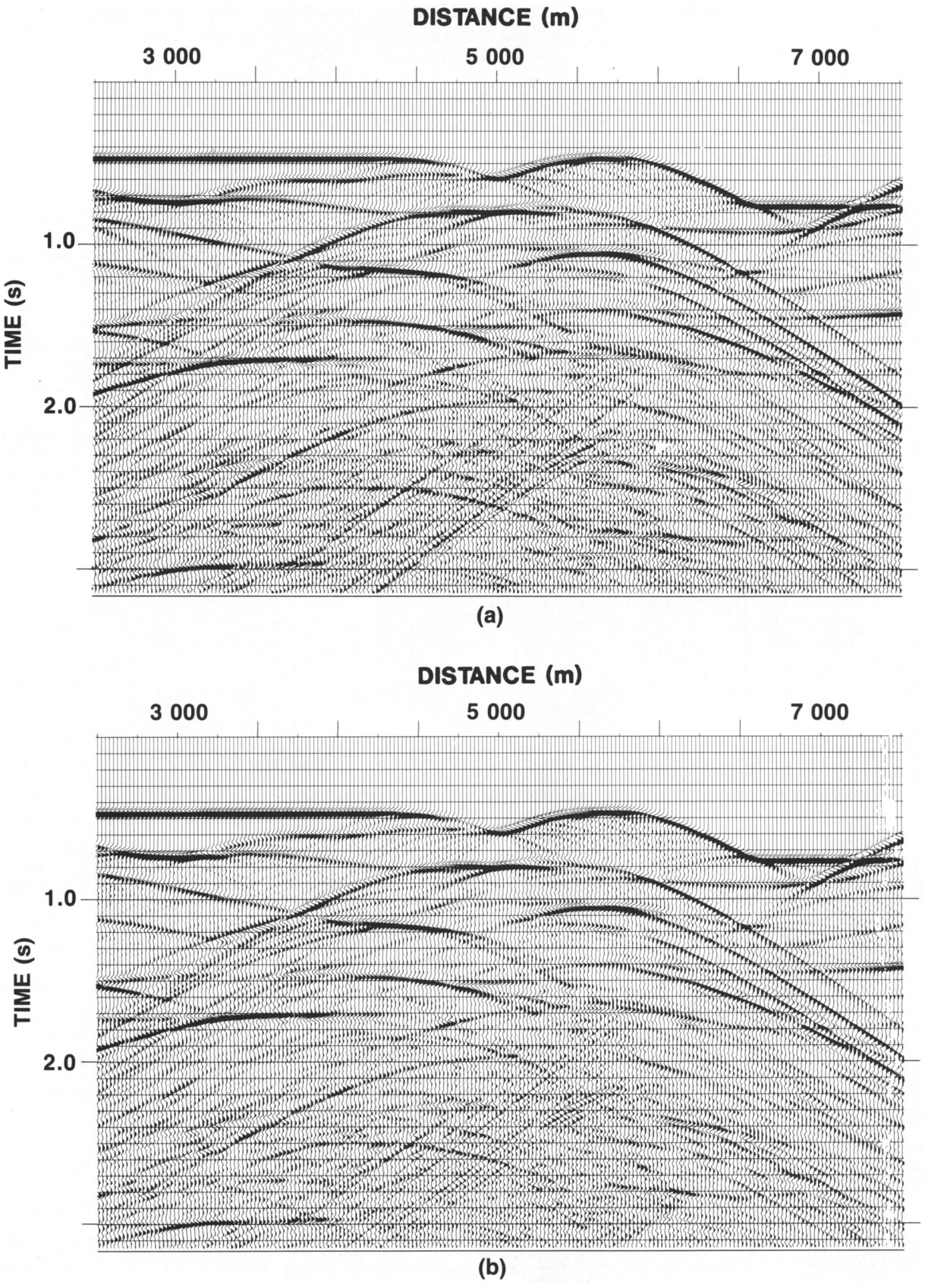

FIG. 11. (a) An exploding reflector synthetic of the 2-D model, computed with an $O(\Delta t^2, \Delta x^4)$ difference scheme. (b) An exploding reflector synthetic of the 2-D model, computed with an $O(\Delta t^4, \Delta x^{10})$ difference scheme. Compare to the $O(\Delta t^2, \Delta x^4)$ exploding reflector synthetic.

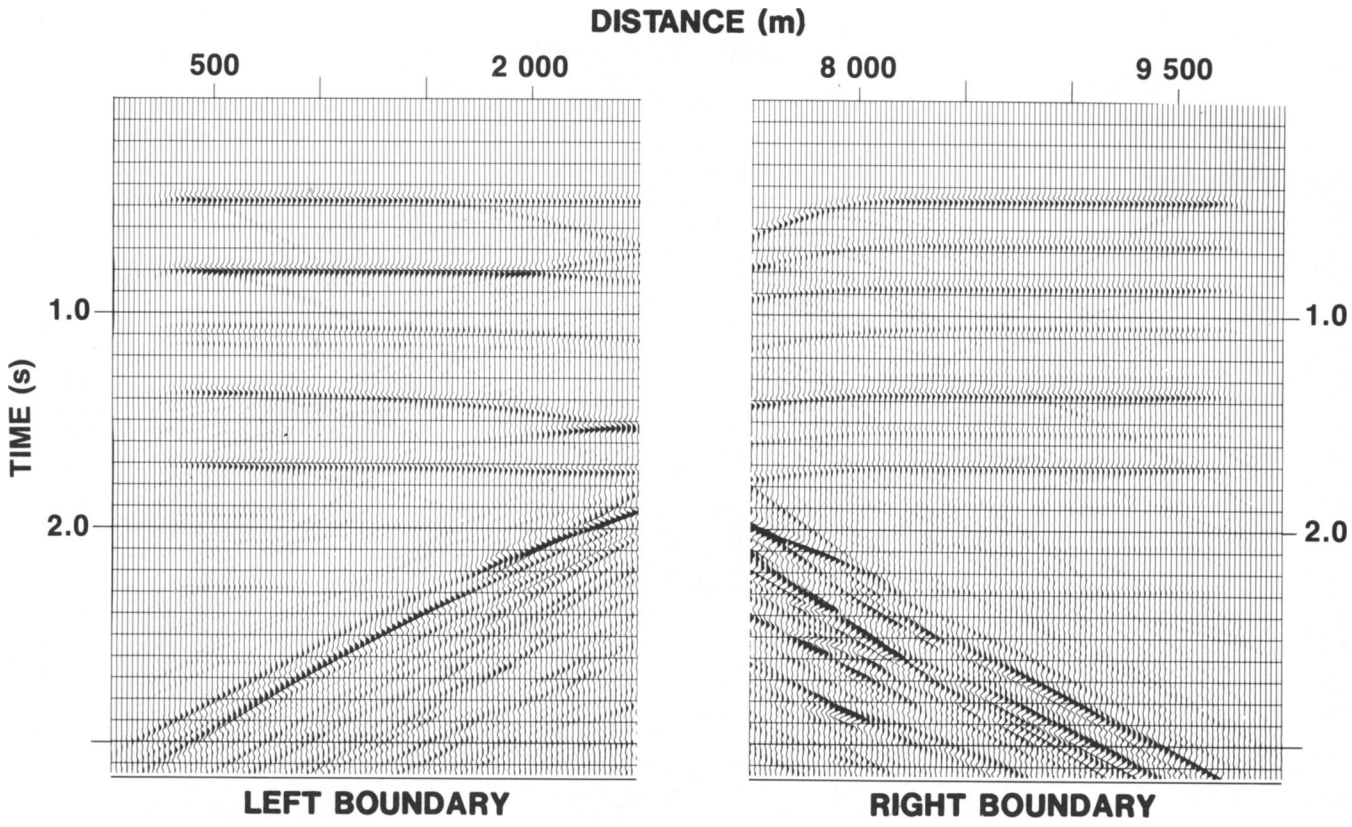

FIG. 12. Left and right transmissive-boundary performance displayed with a linear gain in time.

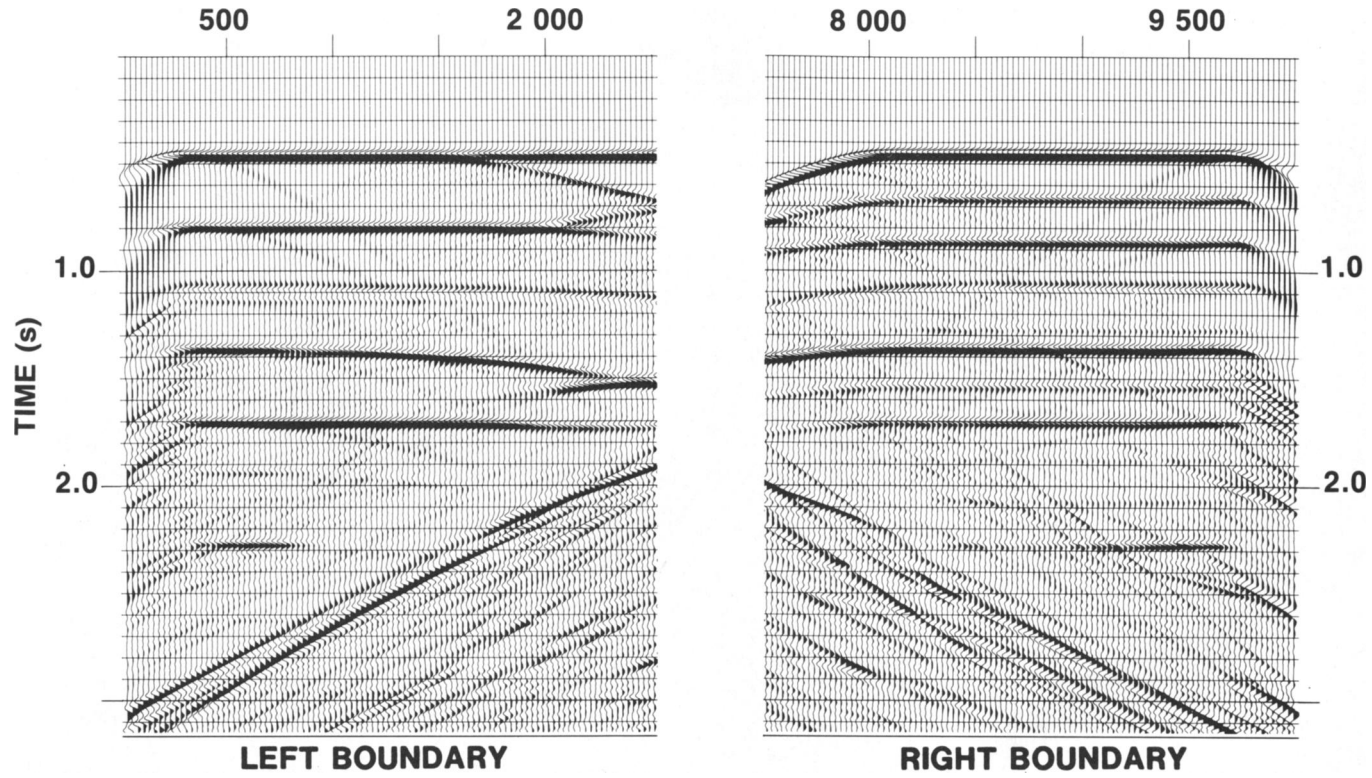

FIG. 13. Left and right transmissive-boundary performance displayed with an 800 ms AGC gate.

Δx^4), $S = 3$, $O(\Delta t^4, \Delta x^{10})$, and the salt-dome model of Figure 10 are

$$\text{TCPU } (\Delta t^2, \Delta x^2) = 3\,075 \text{ s},$$

$$\text{TCPU } (\Delta t^2, \Delta x^4) = 565 \text{ s},$$

and

$$\text{TCPU } (\Delta t^4, \Delta x^{10}) = 525 \text{ s}.$$

These estimates were derived assuming 4 s of data and a Courant number of 0.5. It is evident in moving from second-order to fourth-order accuracy in the spatial difference that a significant advantage is gained in CPU time. In going from the fourth-order to the tenth-order scheme, I essentially broke even in CPU time. The real significance of the high-order algorithms beyond about fourth order is not the CPU time but the memory savings, as illustrated in two dimensions by the salt-dome example.

Figures 12 and 13 display the left- and right-hand boundaries from the $O(\Delta t^4, \Delta x^{10})$ exploding reflector computation. In Figure 12 the boundaries are displayed with a linear gain applied to time. The boundary reflections are effectively nonexistent with only a small amount of boundary diffraction for the shallower reflectors. In Figure 13 these same boundaries are displayed with an 800 ms AGC. The dissipation effects are clear on this scale from the absence of the higher frequencies. Even at this scale it is apparent that the reflections have been practically eliminated.

CONCLUSIONS

As demonstrated, high-order difference schemes have practical advantages when applied to the scalar wave equation. $O(\Delta t^4, \Delta x^{10})$ differencing can provide effective dispersion-free wave models using about three points at Nyquist. One distinct advantage of high-order differencing is the economy achieved in storage requirements at no extra cost in total computational time. A consequence of this reduced storage requirement is the capacity to compute larger models or higher-frequency solutions within the available core memory on the computer. This is especially important when considering vector computations on supercomputers where it can be difficult to overlap input/output operations with the high-speed arithmetic unit. Transmissive sponge boundaries of the Israeli-Orszag type are also shown to be effective in controlling boundary reflections. Finally, it was demonstrated by example that large-scale exploding reflector models can be computed by high-order finite-difference methods on the order of 10 CPU minutes of CRAY 1 time.

ACKNOWLEDGMENTS

I extend my appreciation to A. F. Linville and S. J. Laster for their comments and suggestions regarding this work. Appreciation is also extended to Mobil Research and Development Corporation for permission to release these results. Finally, I thank my reviewers for their timely response and constructive suggestions for an improved final version.

REFERENCES

Alford, R. M., Kelly, K. R., and Boore, D. M., 1974, Accuracy of finite-difference modeling of the acoustic wave equation: Geophysics, **39**, 834–842.
Chester, C. R., 1971, Techniques in partial differential equations: McGraw-Hill Book Company.
Claerbout, J. F., 1976, Fundamentals of geophysical data processing: McGraw-Hill Book Company.
Clayton, R., and Engquist, B., 1977, Absorbing boundary conditions for acoustic and elastic wave equations: Bull. Seis. Soc. Am., **67**, 1529–1540.
Israeli, M., and Orszag, S. A., 1981, Approximation of radiation boundary conditions: J. Comp. Phys., **41**, 115–135.
Kelly, K. R., Ward, R. W., and Treitel, S., 1976, Synthetic seismograms: A finite-difference approach: Geophysics, **41**, 2–27.
Kosloff, D. D., and Baysal, E., 1982, Forward modeling by a Fourier method: Geophysics, **47**, 1402–1412.
Lysmer, J., and Kuhlmeyer, R. L., 1969, Finite dynamic model for infinite media: J. Eng. Mech. Div., **95**, 859–877.
Mitchell, A. R., and Griffiths, D. F., 1980, The finite difference method in partial differential equations: John Wiley and Sons.
Smith, W. D., 1974, A nonreflecting plane boundary for wave propagation problems: J. Comp. Phys., **15**, 492–508.
Whitham, G. B., 1974, Linear and nonlinear waves: John Wiley and Sons.

APPENDIX A

DIFFERENCE OPERATORS

The difference operators in this study were all derived in the same fashion. To illustrate the derivation procedure, the $O(\Delta t^4, \Delta x^{10})$ scheme is reviewed, thereby making all the lower order schemes a subset of this derivation procedure.

One means for introducing a higher-order time estimate for $\partial^2\phi/\partial t^2$ is to write the following power series and inspect it for the necessary information:

$$\phi^{n+1} = \phi^n + \frac{\partial \phi^n}{\partial t}\Delta t + \frac{\partial^2 \phi^n}{\partial t^2}\frac{\Delta t^2}{2!} + \frac{\partial^2 \phi^n}{\partial t^3}\frac{\Delta t^3}{3!} + \frac{\partial^4 \phi^n}{\partial t^4}\frac{\Delta t^4}{4!} + \cdots, \tag{A-1}$$

and

$$\phi^{n-1} = \phi^n - \frac{\partial \phi^n}{\partial t}\Delta t + \frac{\partial^2 \phi^n}{\partial t^2}\frac{\Delta t^2}{2!} - \frac{\partial^3 \phi^n}{\partial t^3}\frac{\partial t^3}{3!} + \frac{\partial^4 \phi^n}{\partial t^4}\frac{\Delta t^4}{4!} - \cdots, \tag{A-2}$$

where

$$\phi^n = \phi(n\Delta t).$$

Summing equations (A-1) and (A-2) and rearranging to bring the second derivative over to the left-hand side of the equation leaves

$$\frac{\partial^2 \phi^n}{\partial t^2} = \frac{1}{\Delta t^2}\left(\phi^{n+1} + \phi^{n-1} - 2\phi^n - \frac{\partial^4 \phi^n}{\partial t^4}\frac{\Delta t^4}{12} - \cdots\right). \tag{A-3}$$

The first three terms on the right-hand side of equation (A-3) form the usual $O(\Delta t^2)$ estimate for $\partial^2\phi/\partial t^2$. To make $\partial^2\phi/\partial t^2$ accurate through fourth-order terms, some estimate for $\partial^4\phi/\partial t^4$ must be introduced. $\partial^4\phi/\partial t^4$ could be estimated directly through a central-difference scheme; however, this would be very inefficient from a memory or storage point of view. It is also unsatisfying from a physically intuitive point of view because it would imply the need for additional initial

conditions. From the storage viewpoint a direct central difference in the time derivative would require additional grid planes which, in three dimensions, would translate into added grid cubes. The need for more time levels can be avoided by realizing that the higher-order even-time derivatives are equivalent to spatial derivatives. This equivalence is provided by the scalar wave equation. In order to construct $\partial^4\phi/\partial t^4$ from spatial derivatives, the following differentiation of the scalar wave equation can be made:

$$\frac{\partial^4\phi}{\partial t^4} = c^2\,\frac{\partial^2}{\partial t^2}\left(\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial z^2}\right),$$

or (A-4)

$$\frac{\partial^4\phi}{\partial t^4} = c^2\left[\frac{\partial^2}{\partial x^2}\left(\frac{\partial^2\phi}{\partial t^2}\right) + \frac{\partial^2}{\partial z^2}\left(\frac{\partial^2\phi}{\partial t^2}\right)\right].$$

The spatial derivatives of the scalar wave equation can now be substituted for the time derivatives on the right-hand side of equation (A-4), producing

$$\frac{\partial^4\phi}{\partial t^4} = c^2\left\{\frac{\partial^2}{\partial x^2}\left[c^2\left(\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial z^2}\right)\right] + \frac{\partial^2}{\partial z^2}\left[c^2\left(\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial z^2}\right)\right]\right\}. \quad \text{(A-5)}$$

If the velocity is constant, equation (A-5) reduces to

$$\frac{\partial^4\phi}{\partial t^4} = c^4\left(\frac{\partial^4\phi}{\partial x^4} + 2\,\frac{\partial^4\phi}{\partial x^2\partial z^2} + \frac{\partial^4\phi}{\partial z^4}\right). \quad \text{(A-6)}$$

For the majority of exploration problems the models consist of constant-velocity bodies of arbitrary shape. The wave field therefore spends most of its time propagating across constant-velocity space. As a practical decision, equation (A-6) was implemented instead of equation (A-5). Although there is certainly need for work on this point, the apparent convergence of the $O(\Delta t^4, \Delta x^{10})$ scheme for the 1-D case and its agreement with the $O(\Delta t^2, \Delta x^4)$ result in the 2-D case, as presented in the examples, suggests that the wave field may not be sensitive to this approximation. Substituting equation (A-6) into equation (A-3) allows a fourth-order accurate time integration for the constant-velocity scalar wave equation to be written,

$$\phi^{n+1} = -\phi^{n-1} + 2\phi^n + c^2\Delta t^2\left(\frac{\partial^2\phi^n}{\partial x^2} + \frac{\partial^2\phi^n}{\partial z^2}\right) + \frac{c^4\Delta t^4}{12}\left(\frac{\partial^4\phi^n}{\partial x^4} + 2\,\frac{\partial^4\phi}{\partial x^2\partial z^2} + \frac{\partial^4\phi}{\partial z^4}\right). \quad \text{(A-7)}$$

This completes the review of the fourth-order time integration scheme.

In all cases the spatial differencing in equation (A-7) was handled as a central difference in much the same manner as equations (A-1), (A-2), and (A-3). The second derivatives $\partial^2\phi/\partial x^2$ and $\partial^2\phi/\partial z^2$ were computed to $O(\Delta x^{10})$. The fourth derivatives were implemented by taking the tenth-order second derivatives and differencing them to second-order accuracy. The algebraic details for the tenth-order accurate derivative are very tedious and do not lend any insight into the scheme. For this reason only the procedure will be outlined with the final coefficients presented in Table A-1.

The procedure begins by taking the power series expressions for the sum of symmetric values about an (x, z), i.e., (i, j) point, and eliminating all nonessential derivatives until the error for the desired derivative is of $O(\Delta x^{12})$. For example, the points $(\phi_{i+1} + \phi_{i-1})$ and $(\phi_{i+5} + \phi_{i-5})$ can be combined to form

$$\frac{(\phi_{i+1} + \phi_{i-1})}{2} = \phi_i + \frac{\partial^2\phi_i}{\partial x^2}\,\frac{\Delta x^2}{2!} + \frac{\partial^4\phi_i}{\partial x^4}\,\frac{\Delta x^4}{4!} + \frac{\partial^6\phi_i}{\partial x^6}\,\frac{\Delta x^6}{6!} + \frac{\partial^8\phi_i}{\partial x^8}\,\frac{\Delta x^8}{8!} + \frac{\partial^{10}\phi_i}{\partial x^{10}}\,\frac{\Delta x^{10}}{10!} + \cdots, \quad \text{(A-8)}$$

and

$$\frac{(\phi_{i+5} + \phi_{i-5})}{2} = \phi_i + \frac{\partial^2\phi_i}{\partial x^2}\,\frac{(5\Delta x)^2}{2!} + \frac{\partial^4\phi_i}{\partial x^4}\,\frac{(5\Delta x)^4}{4!} + \frac{\partial^6\phi_i}{\partial x^6}\,\frac{(5\Delta x)^6}{6!} + \frac{\partial^8\phi_i}{\partial x^8}\,\frac{(5\Delta x)^8}{8!} + \frac{\partial^{10}\phi_i}{\partial x^{10}}\,\frac{(5\Delta x)^{10}}{10!} + \cdots. \quad \text{(A-9)}$$

In all, five equations can be formed from the sums $(\phi_{i+k} + \phi_{i-k})$, $k = 1$ to 5. These five equations can be combined through successive substitution to eliminate the derivatives $\partial^4\phi/\partial x^4$, $\partial^6\phi/\partial x^6$, $\partial^8\phi/\partial x^8$, $\partial^{10}\phi/\partial x^{10}$, thus leaving $\partial^2\phi/\partial x^2$ expressed as a linear combination of the points $(\phi_{i+k} + \phi_{i-k})$, $k = 0$ to 5. In short, the $\partial^2\phi/\partial x^2$ derivative can be expressed as a weighted sum or a convolution in the following form,

$$\frac{\partial^2\phi_i}{\partial x^2} = \frac{1}{\Delta x^2}\left[w_0\phi_i + \sum_{k=1}^{5} w_k(\phi_{i+k} + \phi_{i-k})\right]. \quad \text{(A-10)}$$

The weights as derived from the power series construction are provided in Table A-1. As already noted, the derivatives $\partial^4\phi/\partial x^4$, $\partial^4\phi/\partial x^2\partial z^2$ and $\partial^4\phi/\partial z^4$ were computed by differencing the $O(\Delta x^{10})$ accurate version of $\partial^2\phi/\partial x^2$ in the following way,

$$\frac{\partial^4\phi}{\partial x^4} \simeq \frac{1}{\Delta x^2}\left(\frac{\partial^2\phi_{i+1}}{\partial x^2} + \frac{\partial^2\phi_{i-1}}{\partial x^2} - 2\,\frac{\partial^2\phi_i}{\partial x^2}\right). \quad \text{(A-11)}$$

With regard to the stability of the $O(\Delta t^4, \Delta x^{10})$ operator in two dimensions, I can only offer the guideline of experience which has allowed stable computations at Courant numbers in the vicinity of 0.5. This observed stability includes the sponge layer boundary conditions as outlined in Appendix B.

Table A-1. Weights for a tenth-order accurate, second-difference operator.

k	w
0	-5.8544444444
1	3.3333333333
2	-0.4761904762
3	0.0793650794
4	-0.0099206349
5	0.0006349206

APPENDIX B
SPONGE LAYER IMPLEMENTATION

Two separate issues are considered with regard to the boundary conditions applied in this study. The first issue is producing the extra points at the boundaries which the higher-order operator cannot update. In the $O(\Delta t^4, \Delta x^{10})$ case there are six extra points needed. For example, at the right-hand boundary of the grid, IMAX, the points IMAX-k, $k = 0$ to 5, must be updated in time by some other means than the $O(\Delta t^4, \Delta x^{10})$ operator as outlined in Appendix A. In this case the procedure is to take the lowest order paraxial equation and apply it to the boundary zone. For example, the first point in the zone, IMAX-5, would be updated with the integration scheme

$$\phi^{n+1}_{\text{IMAX-5}} = \phi^{n}_{\text{IMAX-5}} - \frac{c\Delta t}{\Delta x}(\phi^{n}_{\text{IMAX-5}} - \phi^{n}_{\text{IMAX-6}}), \qquad \text{(B-1)}$$

which is one stable difference form for the simplest paraxial wave equation,

$$\frac{\partial \phi}{\partial t} + c\frac{\partial \phi}{\partial x} = 0. \qquad \text{(B-2)}$$

The second issue to consider from the viewpoint of this study is the question of the reflected waves produced by the crude paraxial condition just mentioned. This is where the Israeli-Orszag sponge boundaries become effective. Again, if the right-hand side of the grid is considered, one form of the sponge boundary condition would solve

$$\frac{\partial^2 \phi}{\partial t^2} = c^2\left(\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial z^2}\right) - \varepsilon(x)\left(\frac{\partial \phi}{\partial t} + c\frac{\partial \phi}{\partial x}\right). \qquad \text{(B-3)}$$

The damping coefficient $\varepsilon(x)$ should go from zero for the interior of the mesh, to its maximum value at the right-hand boundary. A simple linear ramp over twenty points with a flat top over ten points was used in this study. Equation (B-3) suggests that the sponge boundary term can be applied after the time update to the mesh has already taken place. If it is assumed that the update has already taken place, then the sponge damping in the right-hand boundary layer would appear as

$$\tilde{\phi}^{n+1}_i = \phi^{n+1}_i - e_i\left[\phi^{n+1}_i - \phi^{n}_i + \frac{c\Delta t}{\Delta x}(\phi^{n}_{i+1} - \phi^{n}_i)\right], \qquad \text{(B-4)}$$

where

ϕ^{n+1}_i is a time updated mesh point,

where

$\tilde{\phi}^{n+1}_i$ is a time updated mesh point with boundary reflections damped,

and

e_i is the damping coefficient in the boundary zone.

Two additional points should be made with regard to this sponge layer. The first point concerns the value of e_i. The maximum value of e_i used in any numerical experiments has been 0.1. This value was used in combination with a Courant number of 0.5. Heuristic reasoning suggests that any perturbing term added to the wave equation should be small in relationship to the physically important parameters in the zone of interest. The physically significant parameter in this case is the Courant number. This reasoning suggests that the sponge layer parameter should be adjusted up or down in relationship to the Courant number for the scheme and its proximity to the phenomenon of interest.

The second point concerns the effectiveness of the sponge layer for a fixed number of boundary points. The capacity of the sponge layer for damping is a function of how many wavelengths are traversed in that zone. This means that for a fixed number of grid cells the boundary layer becomes more effective if coarse-grid sampling can be used. The practical consequence is that a coarse-grid scheme will more effectively damp grid reflections than a fine-grid scheme for the same number of boundary points in a sponge-zone formulation.