Imaging using optimal rational approximation to the paraxial wave equation and generalized Bremner series

Edward Jenner, Maarten V. de Hoop, Ken Larner and
Center for Wave Phenomena, Colorado School of Mines, Golden CO 80401

Mattheus van Stralen
Laboratory of Electromagnetic Research, Delft University of Technology, The Netherlands

ABSTRACT

Migration in areas of complex structure requires that multi-pathing be taken into account to obtain a good image of the subsurface. In any scheme that involves ray-tracing (e.g., Kirchhoff and two-way reverse time migration), properly accounting for all ray paths and computing the correct amplitudes is difficult to implement. One-way wave-propagation algorithms, based on the paraxial wave equation, offer computationally reasonable alternatives to the asymptotic ray-based migration techniques. These propagation schemes require approximations to the square root in the paraxial wave equation. The use of higher-order terms in the square-root expansion results in greater accuracy but at increased computational cost. The larger number of coefficients available in higher-order expansions give more degrees of freedom for optimal operator design without increasing the computation time. Van Stralen et al. (1997) formulate a relatively fast and accurate one-way propagation algorithm by using rational approximations to the vertical slowness and to the horizontal and vertical derivatives as they appear in the paraxial-equation method. Moreover, they cast the one-way wave equation in such a way as to better honor lateral velocity variations, obviating the need for the thin-lens term. In addition, they use a novel optimization procedure, simultaneously minimizing the differences between the group and phase slownesses and the group and true slownesses, in the high-frequency limit, for a given discretization in space. With this optimization, accuracy is maintained for relatively coarse computational grids, even where lateral heterogeneities in the medium are significant.

Here, we test the accuracy of their approach on the Marmousi dataset, obtaining a faithful depth image, including the deep target. Accuracy is maintained even for large vertical grid spacings (as few as 2.5 points per wavelength). Further, we have implemented use of the Bremner coupling series, which allows calculation of the backscattered wavefield in a systematic way, such that primary and multiple scattering are separated. By selecting terms in the series, one can control the constituents of the wavefield (e.g., primary, first-order multiples, second-order multiples) that are used in modeling or migration.

Key words: finite-difference, Bremner series, migration, imaging

Introduction

Being the critically important step in mapping seismic data acquired at the Earth’s surface into a representation of the depth structure of the Earth, migration is a vital step towards identifying potential targets for oil and gas exploration. In areas of complex structure, migration via prestack depth migration has the poten-
tial for obtaining the most accurate images. Naturally the choice of migration algorithm in this costly process will be a trade-off between computational cost and accuracy. Methods based on phase-shift in the Fourier-transform domain, including phase-shift with interpolation, split-step, and phase-screen approaches (Gazdag, 1978; Gazdag and Snieder, 1984b; Stoffa et al., 1990; Wu, 1994) generally require velocity that does not vary too rapidly in the lateral direction.

The Kirchhoff approach, the most commonly used method for prestack depth migration, relies on the high-frequency approximation of ray-tracing. Often, moreover, ray-tracing algorithms calculate only first-arrival traveltimes, and for complex structures where significant multi-pathing occurs this will be inaccurate. More accurate ray-tracing that takes into account multi-pathing and computes the correct amplitudes of each arrival is computationally more expensive and more difficult to implement.

Prestack depth migration using finite-difference solutions of the full (i.e., two-way) wave equation (Baysal et al., 1985; De Faria et al., 1986) are capable of complete wave solutions for arbitrary models. All reflections are modeled, including all multiples, and arbitrarily steep reflectors can be imaged. Unfortunately such approaches are also expensive, requiring use of computation in the time domain on a fine grid. Moreover, they lack control over multiple scattering. During the propagation of both the upcoming and downgoing wavefields reflections and multiples are created from boundaries in the velocity model. If the subsurface velocity were known accurately, then all multiples would be correctly imaged at reflector positions. Inaccuracies in the velocity model, however, will not only cause reflectors to be incorrectly positioned, the multiple energy will also be imaged at positions that differ from those of the primary energy, further contaminating the desired image.

To overcome some of the problems associated with the full solution of the acoustic wave equation, methods based on the paraxial (one-way) wave equation have been used. Such methods can be implemented in the frequency domain, allowing individual frequency components to be computed in parallel. These methods, first introduced by Claerbout (1970), have been used extensively for the migration of seismic data (Claerbout, 1985). Where velocity does not change too rapidly in the lateral direction, the paraxial operators used to approximate the one-way wave equation correctly model waves traveling within a cone centered about a particular axis of the problem. These operators are usually referred to by the extent of their angular accuracy (Claerbout, 1985) and, for inhomogeneous media, are formulated in the \( \omega - z \) domain. The most commonly used are the so-called 15° and 45° approximations. Higher-order approximations are more accurate, but are also more costly.

The method we describe (from Van Stralen et al., 1997) starts with a paraxial wave equation that honors lateral velocity variations without the need of a separate thin-lens term. Then, rational approximations are applied to the operators, i.e. the horizontal and vertical derivatives, and the vertical slowness operator, of that paraxial equation. Furthermore Van Stralen et al. (1997) used an optimization procedure to simultaneously minimize the differences between the group and phase slownesses and the group and true slownesses, in the high-frequency limit, for a given discretization in space. This results in accuracy being maintained on coarse grids, down to 2.5 points per wavelength, and up to large angles of propagation (less than 2.5° error up to approximately 80°).

The Physical Basis

The details of the method are given in Van Stralen et al. (1997); here we give an outline of the physical basis of the theory.

The main operators in the paraxial equation method are the vertical slowness operator, the transverse derivatives (the Laplace operator) and the vertical derivative. Rather than using phase-shift or screen-type approaches, Van Stralen et al. (1997) use rational approximations for the horizontal and vertical derivatives, as well as for the vertical slowness operator. This allows a matrix representation of the operators that is optimally sparse. The propagator matrix has five non-vanishing bands as compared to the three non-vanishing bands in the 15° implicit finite-difference approach.

We use the notation of \( P \) for the wavefield, \( v \) for velocity and \( \partial_z \), \( \partial_x \) and \( \partial_t \) for the vertical, horizontal and temporal derivatives, respectively. The constant-density, two-dimensional scalar wave equation in the \((x, z, t)\) domain where \( x \) and \( z \) are horizontal and vertical position, respectively, and \( t \) is time, can be written as (Claerbout, 1976)

\[
\partial_t^2 P = \frac{1}{v^2} \partial_x^2 P - \partial_z^2 P.
\]

After Fourier transforming over \( x \), \( z \) and \( t \), we get the dispersion relation

\[
k^2 = \frac{\omega^2}{v^2} - k_z^2,
\]

where \( \omega \) is frequency and \( k_x \), \( k_z \) are the vertical and horizontal wavenumbers, respectively. Solving the dispersion relation for \( k_z \) gives
The dispersion relation for the paraxial wave equation, factoring \( \omega \) out of the square root, inverse transforming in the vertical and horizontal directions, and thus relating \( ik_z \) to \( \partial_t \) and \( ik_x \) to \( \partial_x \), results in an expression for the paraxial wave equation in the frequency-space domain,

\[
\partial_t P = \pm i \omega \left[ \frac{1}{v^2} + \frac{\partial_x^2}{\omega^2} \right] \frac{1}{v} P.
\]

Selection of the positive square root corresponds to downward propagating waves while, conversely, the negative root corresponds to upward propagating waves.

Conventionally the velocity is factored from the square root in equation (1) to give

\[
\partial_t P = \pm \frac{i \omega}{v} \left[ 1 + \frac{v^2 \partial_x^2}{\omega^2} \right] \frac{1}{v} P,
\]

and the square root is approximated by a polynomial or rational function. For instance, the 15\(^o\) equation uses the approximation

\[
\partial_t P = \pm \frac{i \omega}{v} \left[ 1 + \frac{v^2 \partial_x^2}{2 \omega^2} \right] \frac{1}{v} P.
\]

To show how the square root is approximated in the method of Van Stralen et al. (1997) we introduce the vertical slowness operator \( \Gamma \) as

\[
\Gamma = \sqrt{\left( \frac{1}{v^2} + \frac{\partial_x^2}{\omega^2} \right)}.
\]

so from equation (1),

\[
\partial_t P = \pm i \omega \Gamma P.
\]

Equation (3) is a high-frequency approximation to the exact vertical slowness operator in case the medium is laterally heterogeneous. Applying \( \Gamma \) to itself, thus yielding \( \Gamma^2 \), should result in the horizontal two-way Helmholtz operator. In this process, the horizontal derivative in the first vertical slowness operator acts on the velocity contained within the second. The meaning of the square-root operator is not obvious. It must be derived from the horizontal two-way Helmholtz operator. To higher-order in frequency, Van Stralen et al. (1997) modify \( \Gamma \) by splitting the local wave speed such that the paraxial wave equation becomes,

\[
\hat{\Gamma} P = v^{-1/2} \left[ 1 + \frac{v \partial_x^2(v)}{\omega^2} \right]^{1/2} v^{-1/2} P,
\]

where \( \hat{\Gamma} \) represents the approximation to the true vertical slowness operator. Note that the operator \( \partial_x^2(v) \) inside the square root acts on all quantities to the right. That includes the field \( v^{-1/2} P \) and, because of the square root, repeatedly on itself. By taking into account lateral derivatives of velocity, use of this operator obviates the need for a separate thin-lens computation. Numerically, we avoid the infinite repetitions by approximating the square root. First we define the operator

\[
X = \frac{v}{\omega \partial_x^2(v)}.
\]

The vertical slowness operator in equation (5) can then be written

\[
\hat{\Gamma} = v^{-1/2} (1 + X)^{1/2} v^{-1/2}.
\]

Using the third-order Thiele approximation (Serafini and De Hoop, 1991; De Hoop and De Hoop, 1992), we obtain the rational approximation to the vertical slowness,

\[
\hat{\Gamma}_{r.o.} = v^{-1/2} \left[ 1 + \frac{\beta_1 X + \beta_2 X^2}{(1 + \beta_3 X)} \right] v^{-1/2},
\]

where \( \beta_1, \beta_2 \) and \( \beta_3 \) are constants that are determined through optimization. Substituting \( \hat{\Gamma}_{r.o.} \) for \( \hat{\Gamma} \) in equation (4) gives the expression for the vertical derivative

\[
\partial_t P = \pm \frac{i \omega}{v^{1/2}} \left[ 1 + \frac{\beta_1 X + \beta_2 X^2}{(1 + \beta_3 X)} \right] \frac{1}{v^{1/2}} P.
\]

The odd-order (i.e., third-order) expansion is chosen, because unlike even-order expansions, its use does not lead to undesirable, singular artificial waves at zero-offset. This problem, which occurs, for instance, with the 45\(^o\) equation, results in artifacts that are difficult to attenuate without also attenuating the physical wavefield. (Some artificial wave are created for complex propagation angles, even for odd-order expansions. Suppression of these artificial slow waves is discussed below.) Use of the third-order expansion, as opposed to first- or fifth-order, is a trade-off between accuracy and computational speed.

The parameters \( \beta_1, \beta_2 \) and \( \beta_3 \) are optimized so as to minimize the difference between the approximate vertical slowness operator given by equation (7) and the exact vertical slowness operator, for pre-critical propagation. An important advantage of splitting the velocity as was done in equations (5) and (7) is that this makes the approximate vertical slowness operator, \( \hat{\Gamma}_{r.o.} \), symmetric and self-adjoint. Moreover, this property is maintained after discretization. This property ensures conservation of vertical power flow, which means that the amplitudes of the computed wavefield are preserved during propagation. In the optimization procedure, the parameters to be optimized are allowed to become complex. This is done so that the artifacts resulting from the approximation to the square root in equation (1) can be effectively suppressed (discussed below). The numerical instability that could arise with use of complex parameters is avoided by our having made \( \hat{\Gamma}_{r.o.} \) symmetric and self-adjoint.

For the discrete approximation to the horizontal de-
rivative in the operation $Xv^{-1/2}P$ [which appears in equation (8)], we use
\[
(1 + \alpha_{2}^2 \delta_v^2 v^{-1/2} P) = (\Delta x)^{-2} \alpha_0 \delta_v^2 v^{-1/2} P, \tag{9}
\]
where $\Delta x$ is the horizontal sampling interval,
\[
\delta_v^2 v^{-1/2} P = v^{-1/2}(x + \Delta x)P(x + \Delta x) - 2v^{-1/2}(x)P(x) + v^{-1/2}(x - \Delta x)P(x - \Delta x),
\]
and $\delta_v^2 v^{-1/2} P$ denotes the discrete approximation to the second horizontal derivative of the product $v^{-1/2} P$. Equation (9) is obtained from the third-order Taylor series expansion of $P$ at $x \pm \Delta x$ from which we find $\alpha_1 = 1$ and $\alpha_2 = 1/12$ (Mitchell and Griffiths, 1985). Increased accuracy can be obtained, however, by optimizing $\alpha_2$ which like the betas in equation (7) can become complex.

The accuracy of the numerical vertical group slowness (as a function of the horizontal group slowness) in approximating the exact vertical group slowness governs the degree of numerical anisotropy (i.e., artificial variation of wave speed with propagation angle); the difference between the vertical phase and group slownesses introduces numerical dissipation (manifested as dispersion).

The numerical anisotropy is caused by approximating the square-root in the paraxial wave equation. In the $15^\circ$ approximation, equation (2) for instance, the dispersion relation is distorted from the exact dispersion relation
\[
k^2 = \frac{\omega^2}{v^2} - k_z^2,
\]
to
\[
k^2 = 2 \left[ \frac{\omega^2}{v^2} - \frac{\omega k_z}{v} \right],
\]
a manifestation of the fact that the $15^\circ$ approximation is an altered version of the wave equation. To see how this changes the computed wavefront, consider vertical propagation, $k_x = 0$. The $15^\circ$ approximation to the dispersion relation becomes
\[
k^2 = \frac{\omega^2}{v^2},
\]
which agrees with the true dispersion relation for $k_x = 0$. Thus the $15^\circ$ approximation is exact for vertical propagation. As we increase $k_x$, $k_z$ decreases and at $k_z = 0$ the approximate dispersion relation becomes
\[
k^2 = \frac{2\omega^2}{v^2},
\]
so $k_z$ is greater than in the true dispersion relation. Thus, away from the vertical the approximate wavefront propagates more slowly than does the true wavefront.

To ensure accuracy in traveltimes and amplitudes, Van Stralen et al. (1997) use an optimization procedure to match the discrete phase slowness to the solution given by the eikonal equation. The group slowness is then matched to the phase slowness. All three operators (horizontal and vertical derivatives and the vertical slowness) are simultaneously optimized in the high-frequency limit, for a fixed bandwidth of the wavefield and grid size in space. Because the optimization depends on the medium velocity, the optimal parameters will vary with both frequency and position; however, an average optimal-parameter set can be obtained for the range of medium velocities and frequencies of interest.

Figure 1 shows the vertical phase and group slownesses associated with the non-optimized one-way equation, while Figure 2 shows the same graphs for the optimized case. The upper two panels in Figures 1 and 2 show the normalized vertical phase and group slownesses as functions of the normalized horizontal slowness. Figure 1 shows that, for the non-optimized one-way propagation, computed vertical group and phase slownesses deviate significantly from the true vertical slowness, particularly at large propagation angles ($\theta \sim 50^\circ$ in a homogeneous medium; $\nu\alpha = 0.75$). The optimized one-way equation, on the other hand, is accurate for larger propagation angles ($\theta \sim 60^\circ$, $\nu\alpha = 0.98$), and deviates significantly from the true slowness at angles only beyond the critical angle ($\nu\alpha > 1$ in Figure 2).

The rational approximations to the vertical slowness operator and to the transverse and the vertical derivatives result in a modified wave equation that yields artificial slow waves, which must be attenuated. These are represented by the portion of the solid curves in Figures 1 and 2 for which $\nu\alpha > 1$, corresponding to a complex propagation angle. Unlike the true slowness (dashed lines in Figures 1 and 2), the vertical group slowness for these waves is not zero; hence, in the approximate medium, waves for which $\nu\alpha > 1$ propagate as slow homogeneous waves. In order to achieve attenuation of these waves, the real frequency $\omega$ in the vertical slowness operator is replaced by a complex one,
\[
\omega' = \omega(1 - i\Omega), \text{ with } 0 < \Omega \ll 1.
\]

Use of a complex frequency has the effect of creating a pseudo-medium in which the artificial slow waves are attenuated, while leaving the real waves relatively untouched. $\Omega$ is kept small because it causes the medium to have frequency-dependent attenuation, with the high frequencies being more rapidly attenuated than low ones. This dissipation device causes the optimal parameters, such as the square root expansion coefficients ($\beta_1$, $\beta_2$ and $\beta_3$), to become complex. In this process the post-
critical amplitude is attenuated strongly, while the change in amplitude of the pre-critical regime is less than 5%.

If desired, the one-way equations for downgoing and upgoing waves can be coupled to produce reflections. This is accomplished via the generalized Bremmer series. Again a detailed mathematical description is beyond the scope of this paper, but can be found in Van Straalen et al. (1997), Collins and Westwood (1991) and Collins (1989). The Bremmer series is an expansion whose terms successively contain the \( j \)-times-scattered wavefield \( j \geq 0 \) with the zero-order term being the first term in the series. Thus the zero-order term simply gives one-way propagation, and contributes only to the solid lines in Figure 3. The first-order term gives the primary reflections (depicted by the dashed lines in Figure 3) and also contributes to the solid lines. However, the main contribution to the transmitted wave is in the zero-order term; higher-order terms contribute progressively less to the transmitted wave. Successively higher-order terms produce successively higher-order reflections. In the process, lower-order terms are used to calculate the successive terms in the series, but only the terms of interest need be selected during imaging. Thus to calculate, say, first-order multiples, the first three terms of the Bremmer series must be computed in order to compute the fourth term, which contains the first-order multiples traveling upwards. It is possible, however, to output only a particular term and thus separate the modeling (or imaging) of primary reflections from that of multiples.

Note that all terms of the Bremmer series contain the multi-pathing due to left and right scattering. Only the scattering in the vertical direction is decomposed into different terms.

Figure 1. Approximate vertical phase slowness (upper left) and the vertical group slowness (upper right) associated with the discretized, approximate, one-way wave equation with no optimization, and a sample rate of five points per wavelength. The exact vertical slownesses are given by the dashed lines. In the bottom row, the difference between the vertical phase slowness and the vertical group slowness, and the difference between the vertical group slowness and the exact vertical slowness are plotted (after Van Straalen et al., 1997). \( v \) is the medium velocity, \( \gamma \) is the true vertical slowness, \( \alpha \) is the horizontal slowness, and \( \gamma^{ph} \) and \( \gamma^{gr} \) are numerical vertical phase and group slowness.

Figure 2. Same as Figure 1 but here the vertical phase slowness (upper left) and the vertical group slowness (upper right) associated with the discretized, approximate, one-way wave equation are computed with (real) optimization parameters (after Van Straalen et al. 1997).

Figure 3. Schematic representing the ray-path contributions of the first and second terms of the Bremmer series. The first term contributes only to the transmitted wavefield (solid lines). The second contributes to both the reflected (dashed lines) wavefield and also partly contributes to the transmitted wavefield.
Snapshots of the Wavefield

The optimized one-way technique is capable of modeling wavefields in media with significant heterogeneity, both lateral and vertical. To demonstrate this, we use a portion of the Marmousi velocity model (Versteeg and Grau, 1990) shown in Figure 4. A vertical displacement line source, represented by the star in Figure 4, is placed at the top of the target. As a benchmark for accuracy, Figure 5 shows a snapshot of the wavefield at 0.95 s using a fourth-order, two-way finite-difference algorithm. In the modeling, the source time function is a 25-Hz Ricker wavelet. To avoid numerical dispersion we used a fine grid spacing of 5 m in both the vertical and horizontal directions. While the first arrival is strong, secondary (i.e., multiple) arrivals have comparable amplitude, in some places stronger. Although the wavefield in Figure 5 also contains multiple scattering and head waves, the high-amplitude portion of the wavefield, above 1-km depth is due to the direct wave.

Figure 6 shows the wavefield computed using an implicit 15th finite-difference algorithm at times \( t = 0.5 \) s and \( t = 0.95 \) s. The highlighted artifacts below the true wavefront are a combination of numerical anisotropy, and numerical dispersion due to the large grid spacing and aliasing. At 0.5 s the wavefront shows some complexity, but no multi-pathing. At 0.95 s the wavefront is close to the surface and shows multi-pathing between 5.5 and 7.0 km. Even within this range of midpoints, however, the secondary arrivals are not as well defined as they are in the result of using the full wave equation (Figure 5) or the optimized one-way scheme (Figure 7).

Using the optimized one-way scheme with the same 8-m vertical and 25-m horizontal grid spacing gives the snapshots at 0.5 s and 0.95 s shown in Figure 7. The result is accurate to large angles of propagation, and the multi-pathing has been honored. Artifacts caused by the influences of the post-critical (slow) wave propagation are visible at 0.5 s, but are rapidly attenuated away from the source. Across the model, the first arrival is accurate in both amplitude and time; however, some of the secondary arrivals do not have the same amplitudes as in the two-way scheme. This may be the result of a number of factors, including inaccuracies in the one-way scheme at large scattering angles, and the presence of both post-critical phenomena (e.g., head waves) and multiple scattering, in the two-way scheme. Also, a phase shift exists between the wavelets in the results of the two-way and one-way schemes. It is not known why this occurs since the same source function was used for both snapshots. We are presently investigating this difference.

Instead of using only the first (zero-order) term in the Bremmer series, in Figure 8 we show the result of summing over the first three terms. The wavefield approximates the full wave equation (Figure 5) and, apart from the phase problems, the main differences are probably due to the modeling of head waves and other post-critical effects in the two-way formulation. Much of the backscattered wavefield, for instance at 2 to 2.5 km depth between 5 and 7.5 km, is accurately modeled, and many events not seen in Figure 7 are evident and coincide with those seen in Figure 5.

**Shot-gather Migration**

Here we focus on migration using one-way propagation as this is far more computationally efficient than using the Bremmer series to simulate a two-way scheme. Furthermore we wish to show that the algorithm is accurate in
regions of complex structure with significant lateral velocity variations. For this reason the Marmousi model was chosen as a test of accuracy for the prestack migration scheme.

The Marmousi dataset, described in Versteeg and Grau (1990), consists of a complex velocity model (part of which is shown in Figure 9) and 240 synthetic acoustic shot gathers each with 96 noise-free channels sampled at 4-ms interval. The first and last shot points are located at 3 and 9 km, respectively. Both the shotpoint spacing and receiver group interval are 25 m, and offsets range from -200 to -2575 m (negative because the receivers are to the left of the shot), simulating an off-end marine geometry. The target is the low-velocity wedge at 2.5-km depth, centered at midpoint 6.5 km.

In the frequency domain, imaging is performed by deconvolving the upgoing wavefield by the downgoing wavefield. This is a division in the frequency domain, with the division being carried out in some stable sense (Gazdag and Sguazzero, 1984a; Berkhout, 1986). We used the following imaging condition:

\[
P(x, z) = \sum_{\omega_i} \frac{S^*(x, z, \omega_i) R(x, z, \omega_i)}{N_z + \epsilon},
\]

where \(P(x, z)\) is the image as a function of \(x\) and \(z\), \(S(x, z, \omega_i)\) is the \(\omega_i\) frequency component of the downgoing wavefield, \(R(x, z, \omega_i)\) is the frequency component of the upgoing wavefield, \(\epsilon\) is a small real number and

\[N_z = \sum_{z=1}^{z_{\text{max}}} S^*(x, z, \omega_i) S(x, z, \omega_i)\]

is a normalization factor. \(S^*(x, z, \omega_i)\) is the complex conjugate of the source wavefield.

The Marmousi data were prestack migrated using a Kirchhoff algorithm (in the common-offset domain) and the optimized one-way finite-difference algorithm (in the shot-gather domain). The Marmousi velocity model is shown again in Figure 9 at the same scale as for the subsequent depth images. We used a velocity model, smoothed laterally and vertically with a 200-m operator, to obtain the Kirchhoff depth migrated section shown in Figure 10. The output midpoint spacing was 12.5 m, with a vertical sampling of 5 m. The faults and upper part of the section are well imaged and, although weak and noise-contaminated, the structure below the faults is evident. The target (between midpoints 6 and 7 km and at about 2.5-km depth), however, is less well imaged, and the reflectors below the faults and in the complex portion of the model (between midpoints 4 km and 8 km) are somewhat discontinuous and unfocused. These problems most likely occur because of the significant multi-pathing
in this portion of the model, not accounted for in the ray-tracing.

Figure 11 shows the result of finite-difference migration in the shot-gather domain using the optimized one-way wave equation with a 25-m horizontal and 8-m vertical grid spacing. The target is now clearly visible and correctly positioned (see Figure 12). The reflectors under the faults are also focused and correctly positioned. Compare, for instance in Figure 12, the image between midpoints 5 and 6-km and 1.5 to 2.0 s. The Kirchhoff migration has not focused the reflectors, even though velocity contrasts are significant, because the ray-tracing fails to give either the maximum-energy arrival times or those of other multiple arrivals. The artifacts seen below 2.5 km at midpoints less than 3.5 km and greater than 8
km are caused by a lack of data at the ends of the survey. They are seen in both Figures 10 and 11.

Also shown in Figure 12 are the results of one-way migration using a vertical grid spacing of 20 m instead of 8 m. Note that accuracy is maintained at this more coarse grid spacing, which results in fewer than 2.5 points per wavelength in the vertical direction, for the highest frequencies in the data (40 Hz). Any loss of vertical resolution (e.g., between midpoints 4.5 km and 5.5 km and at about 1.7-km depth in Figure 12d) is due to the coarser grid spacing, not to the accuracy of the algorithm. The strong reflector in this region has a step-like appearance because of the 20-m output sampling interval in depth. The positions of this and other reflectors in Figure 12d, however, are the same as for the 8-m output sampling interval (Figure 12c).

In shot-gather migration a source wavelet is used to compute the source wavefield $S(x, z, \omega_t)$ in equation (10). We used a 25-Hz Ricker wavelet and migrated frequencies in the range 10 to 40 Hz. Compared to the Kirchhoff migration, the optimized one-way scheme appears to have some loss of frequency content. This may be due to a difference, either in frequency content or in phase, between the (unknown) wavelet used in creating the synthetic data and that used in the migration.

In a recent paper, Bevic (1997) describes a method of combining Kirchhoff migration and downward continuation in an attempt to make use of the most energetic portions of the wavefield. He calls this method semi-recursive Kirchhoff migration and compares the results to those of methods that use paraxial traveltimes and the band-limited traveltimes of Nichols (1994), and to those of reverse-time migration based on the full wave equation. His result compares favorably with other Kirchhoff migrations, and although slightly more expensive, gives significant improvement in imaging the target. The method, however, still has problems focusing reflectors in other portions of the model, for instance beneath the faults. The result in Figure 11 compares well with Bevic’s two-way result in both imaging and positioning of reflectors, and has the correct phase.

Relative Computational Costs of the Different Approaches

The Kirchhoff and finite-difference (one-way) migrations were performed on an SGI Power Challenge using a single processor. The Kirchhoff migration, with output sampling intervals $\Delta x = 25$ m and $\Delta z = 8$ m, takes approximately 2 hours of computation time. The ray-tracing portion of the computation is small (5-10 minutes) because the traveltimes are computed on a relatively coarse grid (50 m in the vertical and horizontal directions) and then interpolated onto the output grid. The optimized one-way finite-difference migration with the same output sampling intervals of $\Delta x = 25$ m and $\Delta z = 8$ m, shown in Figure 11, took about 40 hours. In both cases, total apertures of 5150 m were used. Specifically, for the Kirchhoff migration the aperture was 2575 m on either side of the output location, and, in the shot-gather migration the aperture was 2575 m on either side of the source location. The computational efforts for both the Kirchhoff and finite-difference schemes are approximately linear with aperture, number of time samples on input and the number of output samples. So, for instance, the finite-difference migration using a 20-m output spacing in the vertical direction (Figure 12d), took about 16 hours. In comparison, we estimate the time for a two-way prestack migration, using the two-way scheme to calculate the source wavefield and reverse-time extrapolating the receiver wavefield, to be approximately 430 hours. This is based on using a grid spacing of 5 m in the vertical and horizontal directions and a two-way code with the same 5150-m aperture as in the other migrations.

The optimized one-way code was not written to give the fastest possible CPU performance. For example, algorithms that are formulated in the frequency domain can be easily implemented on parallel processors, with each processor downward propagating separate frequencies. In addition, the grid size need not be constant for all frequencies. Although it would involve some interpolation, coarser grid spacings could be used for the lower frequencies than for the higher ones, thus reducing computational cost.

Since the implementations of the various codes differ, the CPU times given should be considered as only an approximate guide to the relative costs of the different migration schemes. For the same output spacing, the optimized one-way scheme is approximately 20 times slower than the Kirchhoff migration. The two-way scheme will suffer severely from numerical dispersion, particularly in the horizontal direction, if the same grid spacings ($\Delta x = 25$ m $\Delta z = 8$ m) were used. For a realistic spacing of 5 m in both directions, however, it would be more than 10 times slower than the optimized one-way scheme.

Conclusions

In complex structures such as the Marmousi model, multi-pathing must be taken into account to accurately image the subsurface below the faults and the target horizon. Kirchhoff migration can produce an accurate image in the upper portion of the model, but does not correctly
image those portions where significant multi-pathing occurs.

The one-way finite-difference approach developed by Van Stralen et al. (1997) is accurate enough to allow imaging of the complex portions of the Marmousi model, including the target region. The accuracy is maintained for large vertical as well as horizontal grid spacings (down to 2.5 or fewer points per wavelength).

Using the Bremmer coupling series, we can compute the backscattered waveform systematically in such a way that primary and multiple scattering can be separately computed and displayed. This allows the selection of a constituent or constituents (e.g., primary, first-order multiple, second-order multiple) to include in modeling or migration. With this method it is possible, for instance, to image only the multiple energy, either a specific multiple, e.g., first-order, or any combination of orders of multiples.

References