Iterative methods for 3-D finite-difference migration and modeling

Steve Cole

ABSTRACT

Splitting wave extrapolation operators introduces some inaccuracy in 3-D finite difference modeling and migration. While this splitting can be avoided, traditional methods for doing so are prohibitively expensive. Iterative methods such as Gauss-Seidel and successive over-relaxation (SOR) are less expensive than direct solution methods and just as accurate. An additional benefit is that iterative methods can be interrupted when a desired level of accuracy is reached, making the methods even cheaper. A serious drawback is that the iterative methods I use here have stability problems for the 45 degree approximation, which is precisely where they are needed.

INTRODUCTION

Techniques for finite difference migration in two and three dimensions based on approximations to the the paraxial wave equation are developed by Claerbout (1985). In two dimensions, both the 15 and 45 degree approximations yield tridiagonal linear systems of equations. Tridiagonal systems can be solved economically, with a cost proportional to the size of the computational grid.

In three dimensions, the situation is more complicated. Both the 15 and 45 degree approximations yield systems of equations that are no longer tridiagonal. Claerbout points out that the best direct methods for solving the resulting systems have a cost proportional to the cube of the size of the computational grid. For reasonably sized grids, such a cost is prohibitively expensive, so alternative solution methods must be found.

In the case of the 15 degree approximation, the equation can be split into x and y dependent parts. This results in our having to solve two tridiagonal systems for each extrapolation step rather than one more complicated system. For the 45
degree approximation, however, splitting is not possible, as discussed by Brown (1983). Brown describes an approximation due to Francis Muir that does allow splitting to be performed in the 45 degree case, but with some loss of accuracy for certain directions of energy propagation.

While Claerbout's assertion that splitting is necessary because of the prohibitive cost of solving the full 3-D problem is correct, it refers to direct solution methods such as those based on Gaussian elimination. In this paper, I experiment with using iterative methods to solve the full 3-D problem. The three iterative methods to be examined are Jacobi's method, the Gauss-Seidel method, and the method of Simultaneous Over-Relaxation (SOR). These methods are closely related, with SOR being the most difficult to implement and the fastest to converge. I claim that for reasonably sized grids, iterative methods can solve the full 3-D problem rapidly enough that we can consider using them in 3-D migrations rather than accepting the limitations introduced by splitting.

Unfortunately these iterative methods can have stability problems. It turns out that for the 15 degree approximation (where we do not need to use iterative methods because splitting is perfectly accurate) the iterative methods I will discuss are guaranteed to be stable. For the 45 degree case, where splitting isn't accurate and we could benefit from these iterative methods, they are not guaranteed to be stable. I derive the stability criterion and discuss how we might work around it. More work needs to be done to determine whether these stability concerns can be reduced.

FINITE DIFFERENCING IN TWO AND THREE DIMENSIONS

The 15 degree diffraction equation in two dimensions is given by:

\[
\frac{\partial Q}{\partial z} = \frac{v(x,z)}{-2i\omega} \frac{\partial^2 Q}{\partial x^2}
\]

(1)

Using the Crank-Nicolson implicit method, this equation results in a linear system of equations that is tridiagonal. Claerbout gives a tridiagonal solver that is very efficient. Its cost is proportional to the size of the computation grid. He shows that even for the more accurate 45 degree approximation, the system of equations is still tridiagonal.

In three dimensions, the 15 degree diffraction equation has the following form:

\[
\frac{\partial Q}{\partial z} = \frac{v(x,y,z)}{-2i\omega} \left\{ \frac{\partial^2 Q}{\partial x^2} + \frac{\partial^2 Q}{\partial y^2} \right\}
\]

(2)

This gives a system of equations that is no longer tridiagonal. Claerbout states that the cost of solving such a system is proportional to the cube of the grid size, which for large grids means that direct solution methods are prohibitively expensive.
But since the left side of the equation does not involve $x$ or $y$, a splitting method can be used to solve it. For each depth step, we first solve the equation:

$$\frac{\partial Q}{\partial z} = \frac{v}{-2i\omega} \frac{\partial^2 Q}{\partial x^2}$$  \hspace{1cm} (3)

And then using the solution to that equation we solve:

$$\frac{\partial Q}{\partial z} = \frac{v}{-2i\omega} \frac{\partial^2 Q}{\partial y^2}$$  \hspace{1cm} (4)

Doing better than the 15 degree approximation in 3-D poses some problems. The 3-D 45 degree equation is:

$$\left\{1 - \left[\frac{v(x, y, z)}{-2i\omega}\right]^2 \frac{\partial^2}{\partial x^2}\right\} \frac{\partial Q}{\partial z} = \frac{v(x, z)}{-2i\omega} \left\{\frac{\partial^2 Q}{\partial x^2} + \frac{\partial^2 Q}{\partial y^2}\right\}$$  \hspace{1cm} (5)

Note that $x$ and $y$ appear on the left side of the equation. A splitting approach is no longer possible.

That splitting is impossible for the 45 degree approximation follows directly from the continued-fraction expansion derivation of the 45 degree approximation to the dispersion relation. The dispersion relation in 3-D is:

$$k_x = \sqrt{\frac{\omega^2}{v^2} - k_x^2 - k_y^2}$$  \hspace{1cm} (6)

The 15 degree approximation is:

$$k_x = \frac{\omega}{v} - \frac{v(k_x^2 + k_y^2)}{2\omega}$$  \hspace{1cm} (7)

That splitting will be permitted with this dispersion relation follows directly from the fact that it can be separated into terms which involve either $x$ or $y$ but not both. The 45 degree approximation to the dispersion relation is:

$$k_x = \frac{\omega}{v} - \frac{k_x^2 + k_y^2}{2\omega - \frac{v(k_x^2 + k_y^2)}{2\omega}}$$  \hspace{1cm} (8)

The denominator in this expression makes it impossible to segregate the $x$ and $y$ terms. Thus splitting in 3-D for the 45 degree equation is not allowed.

As discussed by Brown (1983), Francis Muir came up with another way to approximate the dispersion relation that does allow splitting:

$$k_x = \frac{\omega}{v} \left[\sqrt{1 - \frac{v k_x^2}{\omega^2}} + \sqrt{1 - \frac{v k_y^2}{\omega^2}} - 1\right]$$  \hspace{1cm} (9)

Not only does this approximation allow splitting, but also for the cases $k_y = 0$ and $k_x = 0$ (energy travelling in the inline and crossline directions) this approximation
is *exact*. Practically speaking, to implement equation 9 we have to approximate the two square roots. If we use 45 degree approximations, we obtain the standard 45 degree dispersion curve for energy travelling in the inline and crossline directions.

For other directions of propagation, however, the approximation is poorer. The worst case is for energy travelling at 45 degrees to the axes. In this direction, the approximation given by equation 9 gives the effective dispersion relation shown in Figure 1. The approximation is better than the standard 15 degree approximation, but worse than the 45 degree approximation.

![Image: Dispersion curves for energy propagating at 45 degrees to the coordinate axes.](image)

**FIG. 1.** Dispersion curves for energy propagating at 45 degrees to the coordinate axes. Exact relation is given by solid line, standard 45 degree approximation by dotted line, Muir approximation by thick dashed line, and 15 degree approximation by thin dashed line. In this worst-case direction, the Muir approximation is better than 15 degrees but not as good as 45 degrees.

How serious is this limitation? It's difficult to say just from looking at dispersion curves. But the fact that there is this non-uniformity in data treatment with direction of propagation prompted me to look for ways to solve the full 3-D problem and avoid splitting altogether.

As mentioned earlier, to avoid splitting we have to solve a linear system of equations that is not tridiagonal. If we consider the system to be of the form $Ax = b$, then the matrix $A$ has the form shown on page 102 of Imaging the Earth's Interior. It is nearly pentadiagonal. Claerbout claims that unlike the tridiagonal case, where the cost of solving the system is proportional to the size of the computation grid, in the 3-D case the cost is proportional to the cube of the grid size. Golub and Van
Loan (1983) discuss using gaussian elimination and other direct methods to solve the problem and verify that Claerbout’s claim is correct.

Golub and Van Loan and Press et al. (1986) discuss a number of iterative techniques that are often used in solving large, sparse linear systems. I have experimented with some of them and found them to be faster than the predicted speed of direct methods (which I haven’t tried) and accurate. Stability is a concern, however. In the remainder of this paper, I will discuss the work that I have been doing with iterative schemes. I’ll begin with a brief review of the theory.

**ITERATIVE METHODS FOR SOLVING LINEAR SYSTEMS**

Our goal is to solve iteratively the linear system of equations:

\[ Ax = b \]  \hfill (10)

I will discuss three closely related methods for solving this problem: Jacobi’s method, the Gauss-Seidel method, and the method of successive over-relaxation (SOR). Jacobi’s method is the simplest to understand and implement, and the slowest. Gauss-Seidel is essentially Jacobi’s method with updating in place. SOR is just Gauss-Seidel with an accelerating parameter whose optimum value depends on the eigenvalues of a matrix.

**JACOBI’S METHOD**

Re-write the matrix \( A \) as:

\[ A = L + D + U \]  \hfill (11)

where \( L \) is the lower triangular part of \( A \) (with zeros on the diagonal), \( U \) is the upper triangular part, and \( D \) is the diagonal part. In Jacobi’s method, we begin with a guess at the solution \( x^0 \) and iterate using the formula:

\[ D\varepsilon^r = -(L + U)x^{r-1} + b \]  \hfill (12)

Conceptually, we solve for each element of \( x \) assuming that the other elements of \( x \) are given by the values from the previous iteration. For the initial guess we can use zeros, but a more effective choice is to use the answer from the previous wave extrapolation step.

The *iteration matrix* that converts one value of \( x \) to the next is \(-D^{-1}(L + U)\). The speed with which the algorithm converges depends on the eigenvalues of this matrix, as discussed by Press et al. All the eigenvalues must have moduli less than 1 for the method to converge. The modulus of the largest eigenvalue is called the spectral radius \((\rho_s)\) of the iteration matrix. Press et al. state that the number of iterations \( r \) required to reduce the error in the solution by a factor of \( 10^{-p} \) is:

\[ r \approx \frac{p \ln 10}{\ln \rho_s} \]  \hfill (13)

*SEP-60*
The iteration matrix is easy to compute because calculating $D^{-1}$ means simply taking the reciprocal of each diagonal element. In another section I describe a method to obtain theoretical convergence rates for Jacobi's method by iteratively computing the spectral radius, the modulus of the largest eigenvalue of the iteration matrix.

GAUSS-SEIDEL METHOD

The Gauss-Seidel method is similar to Jacobi's method. The only difference is that the vector $x$ is updated in place. We march down the vector $x$ from the top, solving for each sample assuming that the elements of $x$ below the current point come from the previous iteration, while for the elements above the current point we make use of those values that have already been computed during this iteration. This corresponds to the matrix description:

$$(L + D)x^r = -Ux^{r-1} + b$$

(14)

The iteration matrix is $-(L + D)^{-1}U$. This would be more difficult to calculate than for Jacobi's method, but fortunately it has been shown that the spectral radius of this iteration matrix is just the square of that for Jacobi's method. Thus Gauss-Seidel is significantly faster.

SIMULTANEOUS OVER-RELAXATION (SOR)

The method of simultaneous over-relaxation overcorrects at each iteration to anticipate future changes. Following Press et al., we can rewrite equation 14 as:

$$x^r = x^{r-1} - (L + D)^{-1}x^{r-1}$$

(15)

where $\xi^{-1}$ is the residual vector given by:

$$\xi^{r-1} = Ax^{r-1} - b$$

(16)

The overcorrection of SOR is given by:

$$x^r = x^{r-1} - \omega(L + D)^{-1}x^{r-1}$$

(17)

Over-relaxation, which gives convergence faster than Gauss-Seidel, occurs for $1 < \omega < 2$. Press et al. state that the improved performance of SOR is strongly dependent upon precise knowledge of $\omega$. Fortunately they give $\omega$ in terms of the spectral radius of the Jacobi method, which is relatively easy to compute. If $\rho_j$ is the spectral radius of Jacobi's method, then $\omega$ is given by:

$$\omega = \frac{2}{1 + \sqrt{1 - \rho_j^2}}$$

(18)
and the spectral radius for SOR \( \rho_s \) is given by:

\[
\rho_s = \left( \frac{\rho_j}{1 + \sqrt{1 - \rho_j^2}} \right)^2
\]  

(19)

In the next section, I will describe a simple method for obtaining the spectral radius of Jacobi's method for a particular problem. From it the converge rates of Gauss-Seidel and SOR and the parameter \( \omega \) will all follow.

**CONVERGENCE**

To determine the spectral radius for Jacobi's method, and hence the rate of convergence of the method, we need to compute the eigenvalue with the largest modulus. While computing all the eigenvalues and eigenvectors of one of these iteration matrices would be very expensive due to their large size, fortunately there is a relatively inexpensive method called the **power method** to obtain just the value that we are interested in. Following Golub and Van Loan, it can be described in a few steps as follows:

For \( k = 1, 2, \ldots \)

\[
z^k = A \nu^{k-1}
\]

\[
\lambda^k = z_1^k \text{ where } |z_1^k| \text{ is a maximum}
\]

\[
\nu^k = z^k / \lambda^k
\]

Here \( \nu^k \) is an estimate of the dominant eigenvector, which emerges after many iterations because its eigenvalue is large compared to all the others. Convergence is most rapid when this dominant eigenvalue is large compared to the next-largest eigenvalue. If the dominant eigenvalue has a modulus that is large compared to that of the next-largest eigenvalue, then this method rapidly converges, yielding for \( z^k \) the dominant eigenvector. The dominant eigenvalue is then computed by solving the problem \( A x = \lambda x \).

With the power method we can determine the spectral radius for Jacobi's method, and using the equations given above, the spectral radii for the Gauss-Seidel and SOR methods, as well as the optimum relaxation parameter \( \omega \) for SOR. Thus we can predict the relative costs of the three methods for a given problem.

**STABILITY**

For any of these iterative methods to be stable, its spectral radius must be less than 1. Golub and Van Loan state that the spectral radius of the iteration matrix for Gauss-Seidel is guaranteed to be smaller than 1 if the matrix \( A \) is diagonally dominant. That is, if the sum of the moduli of the off-diagonal elements in a given row of \( A \) is smaller than the modulus of the diagonal element. This criterion shows that the 15 degree approximation is guaranteed to be stable, but that the
45 degree approximation can be unstable for certain combinations of parameters. This stability analysis is described in the following sections.

**Stability analysis of 15 degree equation**

The 15 degree diffraction equation is:

\[
\frac{\partial Q}{\partial z} = \frac{v(x, y, z)}{-2i\omega} \left\{ \frac{\partial^2 Q}{\partial x^2} + \frac{\partial^2 Q}{\partial y^2} \right\}
\]

(20)

Following Claerbout, an implicit finite-difference implementation of this equation yields a matrix \( A \) with diagonal elements \( 1 - 4\alpha \) where \( \alpha \) is an imaginary value given by:

\[
\alpha = \frac{v\Delta z}{-i\omega 4\Delta x^2}
\]

(21)

Here I have assumed that there is no lateral velocity variation, and that we are using zero-value boundaries. Otherwise the coefficients would not be constant over all the rows. Also we will assume that \( \Delta x = \Delta y \).

For the 3-D problem there are up to four off-diagonal elements in a given row. If \( \alpha = ai \), the modulus of the diagonal element is:

\[
|1 - 4ai| = \sqrt{1 + 16a^2}
\]

(22)

while the sum of the moduli of the off-diagonal elements is at most:

\[
4|ai| = 4a
\]

(23)

It is clear that for the modulus of the diagonal element is always larger than the sum of the moduli of the off-diagonal elements. The matrix \( A \) is diagonally dominant, and all three of the iterative methods are guaranteed to converge in all cases.

**Stability analysis of 45 degree equation**

In 3-D, the 45 degree equation is:

\[
\left\{1 - \frac{v(x, y, z)}{-2i\omega} \right\} \frac{\partial^2 Q}{\partial x^2} \frac{\partial Q}{\partial z} = \frac{v(x, z)}{-2i\omega} \left\{ \frac{\partial^2 Q}{\partial x^2} + \frac{\partial^2 Q}{\partial y^2} \right\}
\]

(24)

Like the 15 degree case, this equation, when implemented in an implicit finite-difference scheme, gives a nearly pentadiagonal system. The extension to 45 degrees simply adds an extra real term to all the elements of the matrix. The diagonal terms are now \( 1 - 4\alpha - 4b \) and the off-diagonal terms \( \alpha + b \) where \( b \) is given by:

\[
b = \frac{v^2}{4\omega^2 \Delta x^2}
\]

(25)

The modulus of the diagonal elements is:

\[
|1 - 4ai - 4b| = \sqrt{1 + 16a^2 + 16b^2 - 8b}
\]

(26)
while the modulus of the off-diagonal elements is at most:

\[ 4|ai + b| = 4\sqrt{a^2 + b^2} \]  \( (27) \)

The matrix will be diagonally dominant if the result of equation refmodd is greater than the result of equation 27. Squaring the two equations and comparing gives us the following inequality which must be satisfied if the Gauss-Seidel and other algorithms are to converge:

\[ 1 - 8b > 0 \]  \( (28) \)

or

\[ b < 0.125 \]  \( (29) \)

Using the definition of \( b \) from equation 25 this translates to:

\[ \omega^2 > \frac{v^2}{\Delta x^2} \]  \( (30) \)

Now consider a typical case. Suppose that the velocity is 2000 meters/second, the trace spacing in \( x \) and \( y \) is 25 meters, and the time sampling interval is 4 milliseconds. This expression tells us that Gauss-Seidel will be stable only for frequencies greater than about 12 Hz. Independent of this theoretical estimate of stability I coded the Gauss-Seidel algorithm for the 45 degree case and, using the parameters I listed above, I found that the algorithm did in fact become unstable at about this frequency. In the following section I describe my numerical studies in more detail.

**NUMERICAL TESTS**

My numerical studies proceeded in two directions. First, I used the power method to assess stability and convergence rates for a typical problem. Then I ran a 3-D migration program using the iterative methods to try out the method and to see if the theoretical estimates of cost were correct.

**Eigenvalue studies**

The eigenvalues of the Jacobi method iteration matrix \(-D^{-1}(L + U)\) tell us the convergence rates of all three iterative methods (including whether they will converge at all) and for the SOR method, the optimum relaxation parameter to use. For a typical case, I will compute the spectral radius (the modulus of the largest eigenvalue) of this iteration matrix, and use that value to predict convergence rates.

Figure 2 shows the variation in spectral radius as a function of computational grid size for the three methods at a fixed frequency of 20 Hz, using the 15 degree equation. Other relevant parameters are a grid spacing of 25 meters in \( x \) and \( y \), and a constant velocity of 2000 meters/second. Note that the spectral radius always increases with grid size. We will have to use more iterations of a given method as
we increase the size of the grid. Figure 3 shows the number of iterations that each method would require (as a function of grid size) to reduce the error in the solution by a factor of $10^{-6}$. Note that if we are going to use the factor by which errors are reduced as our criterion for stopping, it makes sense to use the best initial guess possible. Rather than using a zero wavefield as our first guess, a better choice would be to use the wavefield from the previous extrapolation step. Or as Dave Nichols has suggested, we might even envision using a cheap explicit operator to create a first guess that is in the right ballpark.

We’ve seen that more iterations will be required as we increase the grid size. How are the convergence properties affected by other parameters? Figure 4 shows the number of iterations required for the three methods as a function of frequency for a 64 by 64 computational grid. All other factors being constant, the convergence is affected strongly by frequency. Low frequencies take the longest to converge. We might be able to save a great deal of computation time by excluding some of the lowest, most expensive frequencies. Figure 5 gives the number of iterations required as a function of frequency for an error reduction by a factor of $10^{-6}$.

All of the above tests were performed using the 15 degree equation. In the discussion on stability, I showed that the 45 degree equation may have problems
FIG. 3. Number of iterations required by each of the three methods to reduce errors by a factor of $10^{-6}$, as a function of grid size, using the 15 degree equation. The computational grid size is $2^n$ by $2^n$.

for low frequencies, depending on other parameters. In Figure 6 the spectral radius (modulus of the largest eigenvalue of the iteration matrix) for Jacobi’s method and the 45 degree equation is plotted as a function of frequency. One can see that for low frequencies the spectral radius is larger than 1; for such frequencies none of the three iterative methods presented here will converge.

What can be done about this instability? One could avoid using frequencies for which the method is unstable. As will be seen in the next section, this is not always a satisfactory solution. Perhaps we can condition the problem in some way to reduce the instabilities. For example, I noticed that when I added the 1/6 trick to the 45 degree case that the borderline between stable and unstable frequencies shifted down from about 12 Hz to 10 Hz. Perhaps some other modification can do even better. I need to do more work on avoiding these instabilities if the method is to be worthwhile.

3-D migration with iterative methods

I coded versions of a 3-D zero-offset finite difference migration program using each of the three iterative methods, in order to look for differences in the results as compared to splitting, and also to see how the theoretical predictions of cost compared with what could be obtained in practice. Using these programs I generated impulse responses. Figure 7 shows depth slices taken from the impulse responses
FIG. 4. Spectral radius vs. frequency for Jacobi's method (solid line), Gauss-Seidel (dashed line), and SOR (dotted line) for a 64 by 64 computational grid, using the 15 degree equation. Low frequencies are particularly slow to converge.

for splitting and for Jacobi's method. I was surprised to see some improvement in the quality of the result using an iterative method, since splitting should be accurate for the 15 degree case. More interesting than the results themselves, perhaps, are the execution times. The splitting method implementation required 64 cpu seconds on the Convex C-1, the Jacobi method 246 seconds, the Gauss-Seidel method 334 seconds, and the SOR method 108 seconds. Jacobi's method was faster than Gauss-Seidel because it vectorizes completely on the Convex. Gauss-Seidel, with a recurrence due to the updating in place, does not. This could probably be overcome. The SOR method is less than twice as expensive as splitting, which is encouraging. Certainly all three methods are much cheaper than the predicted cost of direct solution methods.

I also inspected the iterations required for each frequency for each of the methods, and found good agreement between those results and the theoretical predictions based on the eigenvalue analysis.

For the 45 degree case, I generated impulse responses using only those frequencies known to be stable. There is some improvement in the quality of the result, as shown in Figure 8. The impulse response for the iterative method is more symmetric with respect to angle of propagation. However, in looking at the entire impulse response cube with the Movie program, I found that the quality was very poor due
FIG. 5. Number of iterations vs. frequency for Jacobi's method (solid line), Gauss-Seidel (dashed line), and SOR (dotted line) for a 64 by 64 computational grid, using the 15 degree equation, for errors to be reduced by a factor of $10^{-6}$.

to the missing low frequencies. So it seems that simply not using those frequencies for which the iterative methods are unstable will not be an acceptable solution.

CONCLUSIONS

Iterative methods are an inexpensive way to avoid the potential inaccuracies of splitting in 3-D finite difference wavefield extrapolations. The iterative methods described here (Jacobi's method, Gauss-Seidel, and simultaneous over-relaxation) are guaranteed to converge for the 15 degree approximation equation, for which splitting does not suffer from inaccuracy. Unfortunately for the 45 degree approximation, where splitting does have accuracy problems, these iterative methods are not guaranteed to be stable. Stability depends on such dataset-specific parameters as frequency, spatial sampling, velocity, and boundary conditions. At present all I can do is predict when instability will occur and process only those frequencies which are stable. Hopefully future work will allow us to overcome these stability concerns.

SEP-60
FIG. 6. Spectral radius vs. frequency for Jacobi’s method for a 64 by 64 computational grid, using the 45 degree equation. At low frequencies the spectral radius is greater than 1, indicating instability.

ACKNOWLEDGMENTS

Some of this work was done during my five-month stay at TOTAL-CFP in Paris. I’d like to thank Jean-Jacques Raoult of TOTAL for suggesting the applicability of iterative methods to 3-D extrapolation. Back at SEP, John Etgen pointed me in the direction of the power method. Dave Nichols made several helpful suggestions.

REFERENCES


FIG. 7. Depth slices from impulse responses for 15 degree migration using a splitting method (left) and using the Gauss-Seidel method (right). The other iterative methods give results comparable to Gauss-Seidel.

FIG. 8. Depth slices from impulse responses for 45 degree migration using a splitting method (left) and using Jacobi's method (right). The other iterative methods give results comparable to the Jacobi method. The iterative result is more uniform with respect to direction of propagation.
Sat

Computer crashes from heat

Room temp rises

Backup AC turns on

9:30AM Chilled water cuts off

Chilled water restored, computer rebooted

R = room temp

W = chilled water temp