Interpolation techniques to improve the accuracy of the plane wave excitations in the finite difference time domain method

Uğur Öğuz and Levent Gürel
Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey

Abstract. The importance of matching the phase velocity of the incident plane wave to the numerical phase velocity imposed by the numerical dispersion of the three-dimensional (3-D) finite difference time domain (FDTD) grid is demonstrated. In separate-field formulation of the FDTD method, a plane wave may be introduced to the 3-D computational domain either by evaluating closed-form incident-field expressions or by interpolating from a 1-D incident-field array (IFA), which is a 1-D FDTD grid simulating the propagation of the plane wave. The relative accuracies and efficiencies of these two excitation schemes are compared, and it has been shown that higher-order interpolation techniques can be used to improve the accuracy of the IFA scheme, which is already quite efficient.

1. Introduction

The finite difference time domain (FDTD) method was suggested three decades ago as a numerical technique to solve time-dependent Maxwell's equations [Yee, 1966], whose general solution could not be obtained otherwise. With the increase of computing power available to the scientists in recent years and owing to the efficiency, flexibility, and the ease of implementation of the FDTD method, it has become one of the most popular solution techniques in the area of computational electromagnetics.

New extensions and enhancements of the FDTD method are continuously being introduced in order to employ the technique in the solution of new problems [Taflove, 1988; Taflove and Umashankar, 1989; Kunz and Luebbers, 1993; Taflove, 1995; Shlager and Schneider, 1995], perhaps as never envisaged by the original developers. Electromagnetic scattering problems, where the objects are placed in unbounded media and illuminated by plane waves, are also among the wide variety of problems solved by using the FDTD method. The FDTD method does not automatically incorporate the radiation boundary condition; therefore it is more suitable for the solution of problems involving geometries enclosed in conducting or otherwise impenetrable boxes, such as closed waveguides, shielded microwave integrated circuits (MICs) and cavities [DePourcq, 1985]. However, due to the importance of the scattering problems in the computational electromagnetics, they were the first problems to be solved using the FDTD method [Yee, 1966], even before the concept of absorbing boundary conditions (ABCs) was introduced. Later on, the development of the early ABCs [Merewether, 1971; Lindman, 1975] and the solution of scattering problems [Taflove and Brodwin, 1975; Mur, 1981] progressed hand in hand.

Although the FDTD method was originally developed [Yee, 1966] as a “time domain” method and other electromagnetic solution techniques exist for “frequency domain” problems [Harrington, 1982; Miller et al., 1992], the FDTD method is frequently employed in obtaining the sinusoidal steady state solutions of complicated electromagnetics problems excited with time-harmonic sources [Taflove and Brodwin, 1975; Taflove, 1980; Mur, 1981; Umashankar and Taflove, 1982; Taflove and Umashankar, 1983; Taflove et al., 1985]. This is mainly due to the simplicity of the FDTD method and its ability to model complicated inhomogeneities with ease and at no extra cost. Sinusoidal excitation within the FDTD method is used even for some problems that need to
be solved over a finite frequency band. An example is the computation of the radar cross section (RCS) of an object at multiple frequencies.

Being a computational method, the FDTD method produces results with finite accuracy. If this accuracy is sufficient for a given application, the results are considered to be reliable. In the past, 1–2 dB of accuracy was targeted with the FDTD method [Taflove, 1980; Taflove et al., 1985], and this accuracy was sufficient for the engineering applications considered at that time. Noting that 1-dB accuracy corresponds to 12% error in the signal and 26% error in the power, we can conclude that the range of applications, where this much error is still acceptable, is getting narrower. Recently, FDTD solutions with subdecibel accuracy have become possible due to the progress in the following areas:

1. More accurate ABCs have been developed to reduce the reflection error while keeping the problem size at reasonable levels [Berenger, 1994, 1996a, b].

2. High-performance computers have become available. Hence larger problems corresponding to denser FDTD meshes and higher-order FDTD algorithms [Fang, 1989; Deveze et al., 1992; Omick and Castillo, 1993; Manry et al., 1995] can be solved to reduce the dispersion error due to the coarseness of the discretization.

3. Signal-engineering techniques have been introduced to condition the time dependence of the excitation to reduce the errors due to the high-frequency components of the excitation [Guo and Oğuz, 1997].

In addition to the reflection, dispersion, and high-frequency errors, there are several other factors (such as geometry modeling, excitation modeling, etc.) affecting the accuracy of the FDTD solutions. In this paper, we will investigate the errors introduced to the FDTD solution through plane wave excitations. The errors due to the numerical dispersion of an incident plane wave with sinusoidal time dependence are investigated in section 3 following Taflove [1988, 1995]. The errors due to the numerical inaccuracies encountered in the computation of an incident plane wave with arbitrary time dependence are investigated in section 4.

2. Plane Wave Excitation Schemes

For FDTD solutions of most scattering problems, an incident field, whose sources are outside the FDTD computational domain, needs to be simulated. This can be accomplished using either the total-field or the scattered-field formulations of Maxwell's curl equations. The total-field formulation has larger dynamic range compared with the scattered-field formulation [Taflove, 1980]. In the scattered-field formulation, only the outgoing scattered waves need to be absorbed by the ABCs. On the other hand, the scattered-field formulation requires the evaluation of the incident fields everywhere on the surfaces of the impenetrable structures (e.g., conducting objects) and everywhere in the volumes of the lossy or lossless penetrable structures (e.g., dielectric objects). In order to exploit the advantages of each method while keeping the number of the incident-field evaluations at a minimum, the separate-field formulation is suggested [Mur, 1981; Umashankar and Taflove, 1982; Merewether et al., 1980].

In the separate-field formulation, the computational domain is divided into two parts, as shown in Figure 1, such that the total-field region contains all of the inhomogeneities and the scattered-field region is a homogeneous medium surrounding the total-field region. The two regions are connected by a mathe-

![Figure 1. The incident-field array (IFA) excitation scheme in the separate-field formulation. The one-dimensional (1-D) source grid (IFA) points in the direction of propagation. The incident-field values in the 3-D computational domain are interpolated from the closest two elements of the 1-D source grid (when linear interpolation is used).](image-url)
matical boundary, on which a special set of "connecting equations" or "consistency equations" are used. These equations are related to Huygens' principle and the equivalence principle in electromagnetic theory [Merewether et al., 1980]. The incident fields are introduced to the computational domain in these consistency equations. Therefore the incident fields are computed only at the mathematical boundary that separates the two regions, and thus the number of incident-field evaluations is independent of the sizes and types of the scatterers, as opposed to the pure scattered-field formulation.

The accurate and efficient computation of the incident fields is important for the accuracy of the solution. Since the incident field is a known quantity, it is very practical to use closed-form expressions in the connecting equations. This simple method is called the "closed-form incident-field" (CFIF) computation scheme. The implementation of the CFIF scheme is simple, but it requires the computation of a very large number of complicated expressions, such as sinusoids or exponentials. An efficient, FDTD-based method of computing incident fields was proposed by Taflove [1995], which interpolates the incident-field values from a look-up table. The look-up table is a one-dimensional (1-D) grid excited by a hard source, which will be called the incident-field array (IFA) in this work. The incident wave is propagated in this source grid by the 1-D FDTD equations. This source grid or IFA is assumed to point in the direction of propagation of the incident wave, as shown in Figure 1.

The implementation of the IFA scheme is explained in detail by Taflove [1995, pp. 121-124]; however, it will also be outlined here for completeness. A reference point of the IFA, depicted as R in Figure 1, coincides with the initial-contact point on the total-field region. Then, a position vector \( \mathbf{r} \), extending from the reference point \( R \) to the point of interest, is defined. When an incident-field value has to be computed at a particular point in the 3-D computational domain, first the relative position of that point is determined on the source grid. This is done by computing the projection of \( \mathbf{r} \) on the direction of the wave vector \( \mathbf{k}_{\text{inc}} \), that is, \( p = \frac{\mathbf{k}_{\text{inc}} \cdot \mathbf{r}}{\left| \mathbf{k}_{\text{inc}} \right|} \). Let \( P \) denote the greatest integer that is less than or equal to \( p \). Then, the indexes of the closest 1-D vector elements are determined as \( P \) and \( P + 1 \). The desired incident-field value is interpolated from these 1-D vector elements. Figure 1 depicts the case of linear interpolation using the closest two points in the source grid as originally suggested by Taflove [1995]. In this work, well-known Lagrange's polynomial interpolation formula is used for both linear and higher-order interpolations. Higher-order interpolations require more than two points from the IFA, for example, second-order interpolation uses the points indexed as \( P - 1 \), \( P \), and \( P + 1 \). The efficiency of the IFA scheme is due to the fact that both the 1-D FDTD propagation in the IFA and the interpolation operations on the connecting boundary require simple multiplications and additions instead of the evaluation of complicated expressions.

3. Numerical Dispersion in Plane Wave Excitation

In this section, FDTD errors caused by the numerical dispersion of an incident plane wave with sinusoidal time dependence are investigated following Taflove [1988, 1995]. Any plane wave with arbitrary incidence can be generated with the separate-field formulation but with a limited accuracy. One major source of error is the numerical dispersion. As the incident wave propagates through the 3-D computational grid, its phase velocity is changed due to the discretization. That is, the numerical phase velocity \( \tilde{v}_p \) of the wave is different than its theoretical phase velocity \( v_p \). For this reason, there exists a phase difference between the total and the incident fields on the total-field/scattered-field boundary, which produces an error signal in the FDTD equations.

The direction-dependent numerical phase velocity in the 3-D grid is related to the numerical wavenumber \( \tilde{k} \) through

\[
\tilde{v}_p(\theta, \phi) = \frac{\omega}{\tilde{k}(\theta, \phi)}.
\]

The numerical wavenumber \( \tilde{k} \) satisfies the discretized dispersion relation

\[
\left\{ \frac{1}{c \Delta t} \sin \left( \frac{\omega \Delta t}{2} \right) \right\}^2 = \left\{ \frac{1}{\Delta x} \sin \left( \frac{\tilde{k}_x \Delta x}{2} \right) \right\}^2 + \left\{ \frac{1}{\Delta y} \sin \left( \frac{\tilde{k}_y \Delta y}{2} \right) \right\}^2 + \left\{ \frac{1}{\Delta z} \sin \left( \frac{\tilde{k}_z \Delta z}{2} \right) \right\}^2
\]

in the 3-D FDTD grid, where \( \tilde{k}_x \), \( \tilde{k}_y \), and \( \tilde{k}_z \) are the \( x \), \( y \), and \( z \) components of \( \tilde{k} \). The theoretical wavenumber used in the expressions of the CFIF computation scheme should be replaced by the numerical wavenumber \( \tilde{k} \), which is the solution of equa-
tion (2). By doing so, the theoretical and numerical phase velocities are matched and the dispersion error is significantly decreased.

In order to quantify the errors created in the plane wave generation process, the excitation and propagation of waves in a homogeneous media are considered. A 3-D empty computational domain composed of \(30 \times 30 \times 30\) Yee cells and terminated by eight-cell-thick perfectly matched layers (PML) is set up for this purpose. The PML walls are designed to have a theoretical normal reflection ratio \(R(0) = 10^{-4}\) and parabolic conductivity profile. The space sampling period is \(\Delta = 0.625\) cm. The time step is selected at the Courant stability limit as \(\Delta t = 12.081\) ps. Separate-field formulation is employed with a total-field region of \(18 \times 18 \times 18\) cells and a six-cell-thick scattered-field region. The incident plane wave values are computed with the CFIF scheme. The plane wave is incident at \(\theta = 90^\circ\) and \(\phi = 45^\circ\). The incident electric field is polarized in the \(z\) direction, and its amplitude is unity. The incident magnetic field is polarized in the direction of \(\mathbf{\hat{x}} - \mathbf{\hat{y}}\). The time dependence of the incident plane wave is given by

\[
e(t) = w(t) \sin(2\pi f_0 t),
\]

where \(f_0 = 1\) GHz and \(w(t)\) is either the unit step function or a Hanning window defined as

\[
w(t) = \begin{cases} 
0, & t \leq 0, \\
0.5 - 0.5 \cos \left( \frac{\pi t}{L} \right), & 0 \leq t \leq L, \\
1, & \text{otherwise}.
\end{cases}
\]

Note that \(w(t)\) becomes a unit step function when \(L = 0\). For \(L > 0\), the Hanning windows help reduce the FDTD errors due to the high-frequency components of the excitation signal by smoothing the time dependence of the incident plane wave [Gürel and Öğuz, 1997].

Ideally, the fields in the total-field region of the FDTD grid should be exactly the same as the incident plane wave, and the field variables in the scattered-field region should be identically equal to zero. However, owing to the approximate nature of the FDTD method, numerical field variables are expected to deviate from their ideal counterparts. The deviation, that is, the error, can be computed at each time step, in every cell, and for any field component. Figures 2a and 2b show the maximum value of the error in the \(E_z\) component over both the total-field and scattered-field domains at each time step. Figure 2a is obtained using the theoretical wavenumber \(k\), whereas Figure 2b is obtained using the numerical wavenumber \(\tilde{k}\). These errors result are obtained by using Hanning windows of lengths \(L = 0, 0.25T_0, 0.5T_0, 0.75T_0, T_0, 1.5T_0,\) and \(2T_0\), where \(T_0 = 1/f_0\) is the period of the sinusoidal time dependence of the incident plane wave. The input signal is multiplied by a smoothing window at early times in order to decrease the errors due to the abrupt change at the onset of the input signal which has high-frequency components [Gürel and Öğuz, 1997]. The zero-length window corresponds to no smoothing at the input. Figures 2a and 2b show that the high-frequency errors are dominant when no window is used. Increasing the window length to \(L = 0.25T_0\) improves the results. Figure 2a shows that no further improvement is obtained as the window length is further increased using the theoretical wavenumber. This is due to the threshold of dispersion errors at this level. On the other hand, Figure 2b shows that threshold level due to the dispersion errors can be significantly reduced using the numerical wavenumber so that the improvements on the maximum error become visible as the window length is increased beyond \(L = 0.25T_0\). Note that the window length affects the high-frequency errors but not the dispersion errors, and these results testify to the importance of using the numerical wavenumber \(\tilde{k}\) in the CFIF computation scheme.

The numerical dispersion parameters are employed in a different way in the IFA computation scheme. The numerical phase velocities of the incident waves in both the 1-D and the 3-D grids are calculated using equations (1) and (2). Then, the ratio of these velocities is used to modify the permittivity and the permeability values used in the FDTD equations of the 1-D source grid as

\[
H_{inc, m+1/2}^{n+1/2} = H_{inc, m+1/2}^{n-1/2} + \frac{\Delta t}{\Delta \mu_0 \left[ \tilde{v}_p(\theta = 0, \phi = 0) \right]} \times \left( E_{inc, m}^n - E_{inc, m+1}^n \right),
\]

\[
E_{inc, m}^{n+1} = E_{inc, m}^n + \frac{\Delta t}{\Delta \epsilon_0 \left[ \tilde{v}_p(\theta = 0, \phi = 0) \right]} \times \left( H_{inc, m-1/2}^{n+1/2} - H_{inc, m+1/2}^{n+1/2} \right).
\]

This modification results in the equalization of the
Figure 2. Plots of maximum error on $E_z$ obtained by using different lengths of smoothing windows and (a) theoretical and (b) numerical wavenumbers.
numerical phase velocities in the two grids, which is crucial for an accurate excitation of the 3-D grid by the 1-D IFA [Taflove, 1995]. Results obtained using the IFA excitation scheme with the numerical wavenumber and employing various orders of interpolation are presented in the next section.

4. Incident Field Array (IFA) With Higher-Order Interpolation

The top row of error plots shown in Figure 3 is obtained by using the numerical wavenumber and employing linear interpolation as originally suggested by Taflove [1995]. By comparing these relatively high error levels with those of Figure 2b for the same window lengths, one can easily conclude that although the IFA computation scheme is quite efficient, it is not as accurate as CFIF. However, it is possible to increase the accuracy of the IFA scheme by increasing the interpolation order in the computations. When the interpolation order is increased, the incident-field values are related to more points in the 1-D source grid. The geometry of the IFA scheme using cubic (4-point) interpolation is shown in Figure 4. As the IFA computation of an incident-field value uses more 1-D vector elements, the quality of the output also increases. In a linear (two-point) interpolation scheme, a straight line is assumed to pass through the two points. This is a rough estimate for curved functions such as sinusiodals. Higher-order polynomials, such as parabolas or cubic curves, are more suitable to model the variation of the incident wave. Therefore increasing the interpolation order decreases the error in the IFA computations. This improvement is depicted in Figure 3, where the results obtained by using Hanning windows of lengths $0.5T_0$, $T_0$, and $2T_0$ are shown for linear, quadratic, cubic, and fifth-order polynomial interpolation schemes. With the fifth-order interpolation, it is possible to achieve results close to the CFIF results. Using a half-period-long Hanning window and a fifth-order interpolation scheme, the resulting error level is very close to that of the CFIF result, for which the same length of smoothing window is used. Moreover, the IFA technique is still efficient with respect to the CFIF technique, as will be discussed in the next section.

The results of this section imply that together with the aliasing errors due to high-frequency components [Gürel and Oğuz, 1997] and the numerical dispersion [Taflove, 1995], the interpolation order has a significant role in determining the error level in the plane wave excitations. As long as the smoothing windows suppress the high-frequency components sufficiently and the phase velocities are matched, the maximum error level can be controlled by varying the interpolation order.

The usefulness of the error-reducing techniques presented in this section are demonstrated using plane wave excitations with sinusoidal time dependence. However, the applicability of these higher-order interpolation techniques is not limited to the sinusoidal time dependence; they are valid for any arbitrary time dependence of the plane waves.

5. Efficiency of the IFA

A simple experiment is set up to test the efficiency of the IFA technique with higher-order interpolation. One million sinusoidal functions are computed with closed-form expressions, and the computation time is compared with the time spent to perform one million linear, cubic, and fifth-order polynomial interpolations. The experiment is carried out on a SunSPARC10 workstation. The computation times are given in Table 1.

Clearly, even the fifth-order interpolation is more efficient than computing a simple sinusoidal function. If a smoothing window is used, which means that a second sinusoidal term has to be computed, or if an incident wave with a more complicated expression is propagated, the computation time for the CFIF scheme increases; however, the computation times remain the same for the IFA scheme regardless of the input expression.

By examining Table 1 and Figure 3 together, one can see the trade-off between the accuracy of the results and the efficiency with which they are computed. It is important to know that the IFA scheme is more efficient than the CFIF scheme even when fifth-order interpolation is used to obtain highly accurate results.

The computational complexity of an $n$-dimensional FDTD problem with a total of $N$ grid points is $O(N^{n+1/n})$, where the extra factor of $O(N^{1/n})$ is due to the number of time steps [Chew, 1990]. Therefore, for a 3-D FDTD problem, the computational complexity is $O(N^{4/3})$ since the number of time steps is proportional to $N^{1/3}$. The interpolation operations are carried out on a closed surface, namely, on the total-field/scattered-field boundary. On this connecting boundary, the interpolation operations are performed at $O(N^{2/3})$ grid points for $N^{1/3}$
Figure 3. Plots of maximum error levels obtained by using various window lengths and interpolation orders.
time steps. Hence the interpolation operations have a total computational complexity of $O(N)$, which is lower than the complexity of the 3-D FDTD algorithm. Therefore, no matter how high the order is, the interpolation scheme does not add any significant computational cost to the FDTD algorithm.

### 6. Scattering Results With Higher-Order Interpolation

The effects of using higher interpolation orders can be demonstrated by a scattering problem. A square metal plate of size $20 \times 1 \times 20$ Yee cells is modeled for this purpose. The plate lies on the $x$-$z$ plane, in the middle of a computational domain consisting of $40 \times 20 \times 40$ cells, which is divided into a total-field region of $28 \times 8 \times 28$ Yee cells and a six-cell-thick scattered-field region. The incident plane wave is identical to the one in section 3.

The results presented in Figure 5 show the amplitude of the $z$ component of the induced surface current $J_z$ at an arbitrary point on the plate. The amplitudes of $J_z$ signals are computed at every time step with a two-point amplitude-detection scheme, as outlined in the appendix. The amplitude plots of Figures 5a–5d are obtained by using linear, quadratic, cubic, and fifth-order interpolation schemes, respectively. A Hanning window of length $0.5T_0$ is used in all simulations. Figure 5 shows the improvement in the convergence error of the $|J_z|$ variable as the interpolation order is increased. The results obtained with two-point interpolation in Figure 5a converge to a completely different value than the others. The amplitudes of the steady state oscillations in Figure 5b are decreased to a much lower level in Figure 5c by changing the interpolation from quadratic to cubic. Increasing the interpolation order to 5 in Figure 5d makes no significant improvement on the results obtained by cubic interpolation in Figure 5c. However, there is a slight difference in the amplitude levels that the two signals converge to. These results demonstrate various degrees of improvement obtained by using progressively higher orders of interpolation. Compared with the steady state value depicted in Figure 5d, the errors in Figures 5a–5c are 0.14%, 0.0021%, and 0.00062%, respectively, for the induced surface current on the patch. Similarly, the RCS values computed from the currents of Figure 5a are 0.13% in error compared with the that of the Figure 5d.

### 7. Conclusions

In this paper, we have demonstrated that matching the phase velocity of the incident plane wave to the numerical phase velocity imposed by the numerical dispersion of the 3-D FDTD grid is crucial for obtaining accurate plane wave excitation. This observation holds for both the CFIF and the IFA excitation schemes in the separate-field formulation. In general, the IFA scheme is more efficient than, but not as accurate as, the CFIF scheme. We have shown that it is possible to increase the accuracy of the IFA scheme by using higher-order interpolation techniques in the process of transferring the incident-field values from the 1-D IFA grid to the 3-D FDTD grid.
These higher-order interpolation techniques can be used for the excitation of plane waves with arbitrary time dependence.

**Appendix: Two-Point Amplitude and Phase Detection**

Figure 5 presents amplitude plots of a current component as the result of a scattering problem. The amplitudes and phases of the sinusoidal steady state signals are computed with a simple but accurate method. Assuming a discrete sequence obtained by sampling a pure sinusoidal signal, the amplitude and phase values can be extracted from two consecutive samples. The signal values at these time steps can be written as

\[ A \sin (\omega t_1 + \phi) = c_1, \quad (A1) \]
\[ A \sin (\omega t_2 + \phi) = c_2, \quad (A2) \]

where the time instants \( t_1 \) and \( t_2 \) are related to each other by

\[ t_2 = t_1 + \Delta t. \quad (A3) \]

Thus we have a nonlinear system of two equations.
with two unknowns. The unknown parameters are $A$ and $\phi$, the amplitude and phase of the signal. Solving equations (A1) and (A2) for these two unknowns yields

$$\phi = \arctan \left( \frac{\cot \omega t - \frac{c_2}{c_1}}{\csc \omega t} \right) - t_1, \quad (A4)$$

$$A = \csc \omega t \sqrt{(c_1 \sin \omega t)^2 + (c_1 \cos \omega t - c_2)^2}. \quad (A5)$$

Note that most of the trigonometric operations in the above are independent of $t_1, t_2, c_1,$ and $c_2$. Thus they can be computed only once and used several times. Then, equations (A4) and (A5) require the computation of one inverse tangent and one square root operators. Therefore equations (A4) and (A5) can be efficiently used at any two consecutive time instants, perhaps at every FDTD time step. By doing so, one can easily and efficiently keep track of the convergence of the signals to their steady state values, without having to wait several periods of the signal after the convergence.

Equations (A4) and (A5) are derived assuming perfect sinusoids. Thus any perturbations on the finite difference data distort the phase and amplitude computations. However, this distortion is in the same order as the amplitude of the error on the input signal. Therefore the two-point algorithm does not decrease the accuracy set by the FDTD method.

Acknowledgments. The authors would like to acknowledge an anonymous reviewer for his/her meticulous review, which significantly improved the manuscript, and Bijan Houšhmand for his useful suggestions and careful review of the manuscript.

References


L. Gürel, Department of Electrical and Electronics Engineering, Bilkent University, TR-06533, Bilkent, Ankara, Turkey. (e-mail: lgurel@bilkent.edu.tr)

U. Oğuz, Department of Electrical and Electronics Engineering, Bilkent University, TR-06533, Bilkent, Ankara, Turkey.

(Received April 9, 1997; revised August 7, 1997; accepted September 10, 1997.)