Implementation of Transparent Sources in FDTD Simulations

John B. Schneider, Christopher L. Wagner,* and Omar M. Ramahi†

August 2, 1998

Abstract

Sources can be embedded in a finite-difference time-domain (FDTD) grid in any one of several ways. Depending on the particular implementation, the embedded source corresponds physically to a hard field source (applied field), a transparent current source (impressed current), a finite-impedance voltage source, or some other physical excitation. While the implementation of any of these sources is a straightforward procedure in FDTD simulations, ensuring an accurate correspondence between the physical source and its numerical implementation is challenging. In this work, we describe the implementation of a new field source, referred to as a transparent field source, that couples the same fields into the FDTD grid as a hard field source. However, unlike the hard field source, the transparent source does not scatter energy, i.e., the usual FDTD update equation applies to the source node. The implementation is described both in terms of a single node and in terms of an array of nodes. The latter is discussed in the context of parallel-plate waveguide excitation.

*School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA 99164-2752

†Digital Equipment Corporation PKO3-1/R11, 129 Parker St., Maynard, MA 01754

1 Introduction

Energy can be coupled into a finite-difference time-domain (FDTD) grid by external or internal sources. When a source is external to the FDTD grid, the energy radiated by that source, i.e., the incident field, is coupled into the grid usually by means of a total-field/scattered-field formulation or a scattered-field approach. In the total-field/scattered-field formulation [1] the grid is divided into a total-field region and a scattered-field region and the incident field is introduced over the boundary between the two. In the scattered-field formulation [1],[2], the scattered field radiates directly from any material that differs from the background medium.

For many simulations, however, the source of energy must be embedded within the grid. For example, excitation of resonators or antennas requires that the source be positioned within the FDTD grid. A popular source implementation is known as a "hard" source (e.g., section 6.4 of [1]) which is implemented by specifying the field at a given node with a temporal driving function. Since the update equation does not apply to this source node and its value is fixed solely by the driving function, it scatters any energy incident upon it. In certain applications, scattering from the source node is a spurious artifact of the source implementation which degrades the quality of the simulation. One approach to eliminating source scattering requires the use of a pulsed driving function that goes to zero after a finite duration. Once the driving function is zero, the value of the source node is set by the update equation. For this approach to succeed, the duration of the driving function must be shorter than the time it takes for energy to travel from the source node to any material discontinuity and back again. However, in many circumstances this requirement is overly restrictive. In this paper the implementation of a source that radiates the same energy as a hard source, but that does not scatter energy, is presented. We call such a source a transparent field source.

A node in an FDTD grid that has the same material properties as its neighbors and that is governed by the standard FDTD update equation does not, per se, scatter energy. Therefore, it appears that one may simply implement a transparent field source by setting the value of the source node equal to the sum of the value returned by the update equation and the value of the driving function. Unfortunately, although this approach yields a node that does not act as a scatterer (and in that sense is transparent), the energy that it couples into the grid may bear little resemblance to that of a hard field node. A source implemented in this way is, in fact, an injected current and we label such a source a current source. In applications where one measures the energy coupled into the grid and then normalizes by that measured value (as is done when obtaining S parameters), the distinction between a field source and a current source is inconsequential. Thus, in applications where one merely wishes to characterize the spectral properties of a given device and where one properly normalizes by the measured signals, a current source may provide sufficient excitation to characterize the system under test. Nevertheless, one should be aware that the source of energy corresponds physically to a current source rather than a field source. (In [3],[4] such an additive source was described but the source was mistakenly identified as a field source. However, additive and hard source were correctly identified in [5].)

The transparent field sources described here have their greatest utility when one needs to specify the field (either electric or magnetic) that excites a structure. For example, consider a problem in which one must determine the radiation from a horn antenna that is excited via a waveguide feed. Further assume that, in the absence of any discontinuities, the field in the feeding waveguide is given. Thus, the field *incident* on the "throat" of the horn is fixed. The simplest way to model such a problem is to "hard-wire" the known field at the horn throat. Unfortunately, such a model does not accurately reflect the true physical system because the hard-wired field source, which is perfectly reflecting, masks the effects the feeding waveguide has on the radiation. Another approach to modeling this problem is to include a section of the feeding waveguide in the model and again employ a hard field source to excite the waveguide. However, this can be computationally costly since one must ensure that the source is causally or spatially isolated from any reflections that occur at the horn. An alternative approach is provided by the transparent field source. The field can be specified right at the termination of the waveguide, i.e., at the horn throat. To one side of the source would be the horn and to the other would be a small section of the feeding waveguide which is then terminated with an absorbing boundary condition. This approach ensures the model will include the loading of the feed while not significantly increasing the cost.

The transparent field sources presented here should also prove useful in applications involving nonlinear materials. For linear systems, one has the option of using almost any excitation to measure a system transfer function. Then, from the system transfer function one can obtain the system response to other excitations that may be of interest. On the other hand, for most nonlinear systems one typically does not have this option and must use the excitation of interest to determine the system response. The fact that the transparent field sources described here permit one to introduce, in the time domain, a given excitation into a computational domain (without any concern for how the source nodes "load" the system) makes them ideal for such nonlinear applications.

As shown in this paper, it is possible to record a grid impulse response at the source node and then construct a transparent source that couples into the grid the same field as a hard field source. The impulse response is measured at the source node and is fundamentally different from the time-domain Green's function (which is, of course, itself an impulse response, but one for which the source and observation points are not collocated). In one dimension, the impulse response is of finite duration when using the Courant limit. In two and three dimensions, and in one dimension for Courant numbers other than the limit, the impulse response is infinite in duration.

Perhaps the simplest way to implement a transparent field source, and the one used in this study, is first to run an auxiliary simulation that records the impulse response of the grid. This simulation must use the number of dimensions and the Courant number that pertain in the problem of interest, but symmetry can be exploited to reduce the size of the auxiliary simulation. The transparent field source is then realized, in part, by convolving the impulse response with the driving function. Once found, the impulse response can be used for subsequent simulations that have similar geometries. To realize a transparent field *screen* (i.e., a multi-element source that couples the same energy into the grid as a hard multi-element source, but which does not scatter), impulse responses must be recorded over the set of nodes that are members of the screen.

In Sec. 2 implementation of a single node transparent field source in one dimension is described. In Sec. 3 implementation in two and three dimensions is described. Transparent field screens are described in Sec. 4. The transparent field screen is presented in the context of a parallel plate waveguide. The analytic solution for the field in the guide is presented for hard-source excitation. Numerical results obtained using FDTD simulations with either a hard field screen, a transparent field screen, or a transparent current screen are compared to the analytic solution.

2 One-Dimensional Transparent Field Source

Although a transparent source in one dimension is of little practical use, it is simpler first to consider implementation in one dimension. We assume propagation in the x direction and z polarization of the electric field so that the governing differential equations are

$$\frac{\partial E_z}{\partial t} = \frac{1}{\epsilon}\frac{\partial H_y}{dx} \tag{1}$$

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu}\frac{\partial E_z}{dx} \tag{2}$$

If a source current density J_z were present, it would be subtracted from the right side of (1) (and scaled by $1/\epsilon$). These equations lead to the standard update equations for the FDTD method in one dimension:

$$E_z^{n+1}(i) = E_z^n(i) + Zs\left[H_y^{n+1/2}(i) - H_y^{n+1/2}(i-1)\right] \quad (3)$$

$$H_y^{n+1/2}(i) = H_y^{n-1/2}(i) + \frac{1}{Z}s\left[E_z^n(i+1) - E_z^n(i)\right] \quad (4)$$

where $s = c\Delta t/\Delta x$ is the Courant number, $Z = \sqrt{\mu/\epsilon}$ is the impedance, and Δx and Δt are the spatial and temporal step sizes, respectively. For brevity, the spatial offset between the E_z and H_y nodes is suppressed in the arguments of the discrete forms, thus $H_y^{n+1/2}(i)$ is equivalent to $H_y([i+1/2]\Delta x, [n+1/2]\Delta t)$. The maximum value of the Courant number s yields the minimum amount of numerical dispersion and the longest simulation duration for a given number of time steps, but the Courant limit cannot be used throughout the computational domain for simulations of inhomogeneous regions.

Consider a one-dimensional (1-D) computational domain in which the source is an electric field node at i_{src}. A hard source is realized by setting the source node equal to a given driving function $f(n\Delta t) = f^n$. The electric field at the source node is then $E_z^n(i_{\text{src}}) = f^n$, but all other nodes are governed by the update equations (3) and (4). Assuming that the driving function is zero prior to $n = 0$, Fig. 1 shows the values of E_z and H_y in the vicinity of the source for the first two time steps. At the Courant limit, a 1-D FDTD simulation of propagation in a homogeneous medium is equivalent to a series of shift operations. Hence, the electric field at node i and time step n is given by

$$E_z^n(i) = f^{n-|i-i_{\text{src}}|} \quad (5)$$

Since a hard source only depends on the driving function and is independent of other propagating fields, it is effectively perfectly reflecting. Therefore, if a space is inhomogeneous and a reflected field propagates back to the source, the source will, in turn, reflect that field.

With the goal of creating a transparent field source, let us implement the source as the sum of the driving function and the update equation that pertains at that node. The value of the source node is then given by

$$E_z^{n+1}(i_{\text{src}}) = E_z^n(i_{\text{src}}) + Zs\left[H_y^{n+1/2}(i_{\text{src}}) - H_y^{n+1/2}(i_{\text{src}} - 1)\right] + f^{n+1} \tag{6}$$

(This implementation is essentially that which was described in [3],[4].) Figure 2 shows the values of E_z and H_y in the vicinity of the source for the first two time steps for this source implementation. Significantly, the field throughout the grid cannot be obtained simply by a shifted (or delayed) value of the driving function. Instead, the field at an arbitrary node is given by

$$E_z^n(i) = \sum_{m=0}^{n-|i-i_{\text{src}}|} (-1)^{m+n-|i-i_{\text{src}}|} f^m \tag{7}$$

In contrast to the hard source, any field that is reflected back to the source node will pass through it. In this sense the source node is "transparent." However, the field that is coupled into the grid by the source node will not resemble that of the hard source since the source physically corresponds to a current source (as is evident by retaining the current density term in (1)). To demonstrate this distinction, consider the case $f^n = \delta[n]$ (the Kronecker delta function) for which f^0 is unity and all other values of f^n are zero. In this case the field that propagates away from the source node is a series of ones with alternating signs. Since, at the Courant limit, the field propagates without error, this result is directly attributable to the source implementation and is not indicative of any error inherent in the FDTD simulation.

Inspection of Fig. 2 shows that the field coupled into the grid can be made identical to that of the hard source with the addition of a delayed sample of the driving function. This delayed term, which is added to the update equation and the undelayed driving function as given by (6), cancels the "echo" of the previous source term (i.e., f^n) caused by using the update equation at the source node. Thus, a truly transparent field source that couples the same field into the grid as the hard source can be achieved using

$$E_z^{n+1}(i_{\text{src}}) = E_z^n(i_{\text{src}}) + Zs\left[H_y^{n+1/2}(i_{\text{src}}) - H_y^{n+1/2}(i_{\text{src}}-1)\right] + f^{n+1} + f^n \tag{8}$$

This source implementation produces the fields shown in Fig. 1, but the source node does not scatter (or reflect) any field incident upon it.

One-dimensional FDTD simulations performed using Courant numbers other than the limit do not permit such a simple implementation of a transparent field source. At the Courant limit, the term that is echoed by the update equation back onto the source node depends only on the value of the driving function at the previous time step. When the Courant number is less than unity, the FDTD algorithm is not equivalent to a set of simple shift operations nor can it provide an exact solution because of inherent numerical dispersion.

To facilitate the construction of transparent field sources that will work for any Courant number, we define a grid impulse response. First, consider a grid in which the source node is implemented as a hard source and the driving function is a Kronecker delta function. We define the grid impulse response as the values that are obtained using the update equation at the source node. (The update equation is calculated at the source node and the returned value is recorded as part of the impulse response. However, the value of the source node is not set to this value—as dictated by the hard Kronecker delta function, the source node is initially one and zero thereafter.) Thus, the impulse response is calculated from the previous value of the source node and its surrounding magnetic field nodes, but the impulse response does not couple back to the source node because the node is "hard" and its value is fixed by the Kronecker delta function. Therefore, the source node is given by $E_z^n(i_{\text{src}}) = \delta[n]$ while the impulse response is

$$I^n = E_z^{n-1}(i_{\text{src}}) + Zs\left[H_y^{n-1/2}(i_{\text{src}}) - H_y^{n-1/2}(i_{\text{src}}-1)\right] \tag{9}$$

One can obtain the impulse response analytically—it is simply a polynomial whose order increases with each time step—but it quickly becomes unwieldy. For example, the first few terms of I^n are

$$
\begin{aligned}
I^0 &= 0 \\
I^1 &= 1 - 2s^2 \\
I^2 &= -2s^2 + 2s^4 \\
I^3 &= -2s^2 + 6s^4 - 4s^6 \\
I^4 &= -2s^2 + 12s^4 - 20s^6 + 10s^8 \\
I^5 &= -2s^2 + 20s^4 - 60s^6 + 70s^8 - 28s^{10}
\end{aligned}
$$

Note that when $s = 1$, the impulse response is $I^n = -\delta[n-1]$, but for Courant numbers less than unity, the impulse response is infinite in duration. Fortunately, it is not necessary, nor even desirable, to obtain the polynomial form of the impulse response. Instead, the impulse response can be obtained numerically via an FDTD simulation using a homogeneous grid that has the same material properties as those found at the source node in the problem of interest. In this simulation, a hard source is driven impulsively and the impulse response is recorded using (9). Symmetry can be exploited, since the magnetic field $H_y^{n-1/2}(i_{\text{src}} - 1)$ is the negative of $H_y^{n-1/2}(i_{\text{src}})$, so that only half the 1-D space need be simulated. The impulse response can then be found using

$$I^n = E_z^{n-1}(i_{\text{src}}) + 2ZsH_y^{n-1/2}(i_{\text{src}}) \tag{10}$$

where $I^n = 0$ for $n \leq 0$.

The impulse response can be used to give the field that will echo back to the source node if the source node is equal to the sum of the update equation and the driving function as given by (6). Assuming such a source and that the first non-zero value of the driving function is $E_z^0 = f^0$, the source node at the next time step is $E_z^1 = f^1 + I^1E_z^0$; at the next it is $E_z^2 = f^2 + I^1E_z^1 + I^2E_z^0$; and so on. Clearly, if a transparent source is to couple the same field into the grid as a hard source, the source node must, in the absence of any reflected field, take on the same values as those of a hard source, i.e., the source node must take on the values of the driving function and the echoed values must all be canceled so that $E_z^0 = f^0$, $E_z^1 = f^1$, $E_z^2 = f^2$, etc. The cancelation is realized by subtracting $I^1 f^0$ from the source node at the first update, subtracting $I^1 f^1 + I^2 f^0$ at the next update, subtracting $I^1 f^2 + I^2 f^1 + I^3 f^0$ at the next, and so on. Said another way, to implement a transparent field source, one must subtract the convolution of the impulse response and the driving function from the source node. Specifically, a transparent field source for an arbitrary Courant number is obtained using

$$E_z^{n+1}(i_{\text{src}}) = E_z^n(i_{\text{src}}) + Zs\left[H_y^{n+1/2}(i_{\text{src}}) - H_y^{n+1/2}(i_{\text{src}} - 1)\right] + f^{n+1} - \sum_{m=0}^{n} I^{n-m+1} f^m \tag{11}$$

In one dimension, the impulse response for a magnetic field node is the same as for an electric field node. In three dimensions the impulse response is different than in one dimension, but the impulse response is independent of the field component (i.e., all six field components have the same impulse response). In two dimensions, however, the impulse responses is polarization dependent. Field components in the plane (e.g., H_x and H_y in a TM problem or E_x and E_y in a TE problem) have the same impulse response, and out-of-plane components (e.g., E_z in a TM problem and H_z in a TE problem) have the same response, but the in-plane and out-of-plane responses differ. Nevertheless, in all dimensions and for all components, the transparent field source implementation, which is described in the next section, is basically the same.

3 Transparent Field Sources in Two and Three Dimensions

A more general form of (11) that also holds in two and three dimensions is

$$F^{n+1}(\vec{r}_{\text{src}}) = (N\text{-D update equation}) + f^{n+1} - \sum_{m=0}^{n} I_N^{n-m+1} f^m \tag{12}$$

where F is any one of the six field components, N is the number of dimensions, $\vec{r}_{\rm src}$ is the source location, "N-D update equation" is the update equation appropriate for the node in the given number of dimensions, and I_N is the grid impulse response. With a change in dimension, the update equation changes and the values of the impulse response change, but the underlying approach does not change. The definition of the impulse response also remains unchanged: A hard source is driven impulsively and the impulse response is obtained using the update equation at the source node.

In three dimensions, the impulse response approaches zero and the rate at which it approaches zero depends on the Courant number. The closer the Courant number is to the limit, the more rapidly the impulse response approaches zero (here the rate of fall-off is discussed relative to the number of time steps, not absolute time). Thus, after a sufficient number of time steps the impulse response can be approximated by zero and the convolution with the source function in (12) does not necessarily have to be done over the entire previous history of the source function. Instead, it only needs to be done over the number of time steps the impulse response is treated as non-zero.

Unlike in one and three dimensions, the impulse response in two dimensions does not quickly converge to zero, rather it decays very slowly. This type of response can be problematic for simulations requiring a large number of time steps. There are ways, however, to work around this. For example, the decay rate is so slow that the convolution of the driving function and the impulse response eventually can be approximated by zero for a finite-duration driving function with no dc component. (The convolution of a constant with a signal that has no dc component is zero. The impulse response decay is slow enough that for signals of sufficiently short duration, the convolution may be well approximated by zero.)

Finally, to demonstrate the different behavior of hard and transparent field sources, consider a two-dimensional TM problem with a single-node E_z field source near a perfect electric conductor

(PEC) plane as shown in Fig. 3(a). The source driving function is a Ricker wavelet (details of this wavelet are discussed in the appendix). The spatial step is such that there are 32 points per wavelength at the peak frequency of the wavelet and the temporal step is set so that the Courant number is the limit. Figures 3(b) and (c) show the electric field in the vicinity of a hard and transparent field source, respectively, after 220 time steps. In these grayscale field maps, black corresponds to zero and the brightness of a pixel is indicative of the absolute value of the electric field found at the corresponding node. The hard source, Fig. 3(b), while radiating the same primary field as the transparent source, scatters the reflected field as evidenced by the non-black region between the reflected wave and the PEC surface. For the transparent source, Fig. 3(c), the primary wave is identical to that of the hard source but the source does not interfere with the field reflected by the PEC surface.

Unlike in two dimension, in three dimensions for a single-node source close to a PEC plane, the results obtained by implementing the source as either hard or transparent are not strikingly different. This is due to the more rapid decrease in the field as one moves away from the source and the smaller scattering cross section of the hard source. However, when the source consists of several nodes, the different implementations can yield profoundly different results. The implementation of a transparent multi-node source is discussed in the next section.

Finally, note that many individual transparent field sources can be used within a single simulation (this type of superposition is distinct from the application discussed in the next section). Each source introduces energy into the grid without scattering it. This could prove useful in the excitation of resonant structures or in applications where one is interested in mimicking the behavior of a "classic" phased-array with non-interfering elements (the total radiated field is the superposition of the individual radiated fields).

4 Excitation of a Parallel-Plate Waveguide

The individual transparent field sources considered thus far couple energy into the grid in the same way as would a hard field source in an unbounded medium. In this section we consider the construction of a multiple-element transparent source (or "screen") that couples energy into a bound structure in the same way as would a hard field screen. The general principal used to construct individual transparent field sources still holds for this case. However, impulse responses are now measured for each node in the screen when the entire screen is excited impulsively; thus, each source node has its own (unique) impulse response.

Two-dimensional parallel-plate waveguides have been used in many FDTD applications and have been especially prevalent in the evaluation of absorbing boundary conditions (ABCs) (e.g., [6],[7]). Thus, the use of transparent screens is demonstrated for such waveguides. We first provide the analytic solution for the field at an arbitrary point in the guide when the source is a hard screen. The solution is presented in terms of non-dimensional parameters so that it can easily be scaled to any waveguide and any excitation.

Note that the transparent source screen can be brought arbitrarily close to a waveguide discontinuity. Any energy reflected back to the screen will pass through it. Therefore the computational domain (i.e., the waveguide) must extend to either side of the transparent source screen. Also, unlike in a total-field/scattered-field formulation, fields will radiate to either side of the screen. Thus, a practical problem might involve a transparent screen placed very close to a waveguide discontinuity (such as an iris, a flared horn, a bend, or a dielectric load). On the other side of the screen, an ABC would be used to absorb the field initially radiated by the screen and to absorb any field that is subsequently scattered by the discontinuity and propagated back through the screen.

Although the transparent screen is presented specifically for two-dimensional parallel-plate waveguides, the application to three-dimensional waveguides of any cross section follows a similar development. Furthermore, it is possible to implement a transparent screen over any aperture so that the aperture radiates as if the fields were "hard-wired." It is merely necessary to first determine the impulse responses for the source-screen nodes.

4.1 Analytic Solution for a Hard Screen

Consider a parallel-plate waveguide of thickness W as shown in the inset in Fig. 4. We restrict consideration to the TE_1 mode and assume E_z is specified over the cross section of the guide at $x = 0$. (In the context of a waveguide, TE implies the electric field is transverse to the guide axis so that the non-zero fields are E_z, H_x, and H_y. In the previous discussion, TE implied the electric field was transverse to the out-of-plane direction.) To facilitate application to guides of any dimension, a solution for the field at an arbitrary point is sought that is expressed in non-dimensional units. Assume that E_z at $x = 0$ is given by

$$E_z(x = 0, y, t) = \sin(\pi y/W)p(t) \tag{13}$$

where the "driving function" $p(t)$ can be chosen arbitrarily. The sinusoidal variation in y ensures fields only propagate in the TE_1 mode. The driving function used in the subsequent analysis is a Ricker wavelet which is described in detail in the appendix. (An alternate implementation of a source screen can be found in [8],[9] where an impedance is used to control the introduction of a given mode and the attenuation of undesired modes).

Assume, for the moment, that the field over the waveguide cross section at $x = 0$ is harmonic with phasor representation

$$\tilde{E}_z(x = 0, y) = Ae^{j\theta} \sin(\pi y/W) \tag{14}$$

where $\exp(j\omega t)$ time dependence is understood. The phasor at an arbitrary point is

$$\tilde{E}_z(x, y) = Ae^{j\theta} \sin(\pi y/W) e^{-\gamma x} \tag{15}$$

where γ is the guide propagation constant for the TE_1 mode given by

$$\gamma = j\frac{\omega}{c}\left(1 - \left(\frac{\lambda}{\lambda_C}\right)^2\right)^{1/2} \tag{16}$$

and $\lambda_C = 2W$ is the cut-off wavelength of the guide. Thus, for an arbitrarily located observation plane, the waveguide transfer function is

$$G(\omega) = e^{-\gamma x} = \exp\left[-j\frac{\omega x}{c}\left(1 - \left(\frac{\lambda}{\lambda_C}\right)^2\right)^{1/2}\right] \tag{17}$$

To express this in discrete form, assume a uniform spatial step size δ so that the distance from the $x = 0$ plane is $N_X\delta$, i.e., the distance from the "reference" plane to the observation plane is N_X cells. Let the cut-off wavelength be $\lambda_C = N_C\delta$, so that the waveguide is $N_C/2$ cells wide. Finally, consider a time-domain simulation in which the total number of time steps is N_T (this need not be the number of time steps used in an actual FDTD simulation). The discrete Fourier transform of a time series with N_T time steps has a spectral resolution of $\Delta f = 1/(N_T\Delta t)$ so that, assuming N_T is large enough so that there is no aliasing, the discrete spectrum $G[m]$ is obtained by setting ω equal to $2\pi m\Delta f$ in (17). Thus, $G[m] = G(2\pi m\Delta f)$ which can be written as

$$G[m] = \exp\left[-j2\pi\frac{mN_X}{sN_T}\left(1 - \left(\frac{sN_T}{mN_C}\right)^2\right)^{1/2}\right] \tag{18}$$

For pulsed excitation, the discrete time-domain field at an arbitrary location is obtained by taking the inverse discrete Fourier transform of the product $G[m]P[m]$ where $P[m]$ is the discrete spectrum of the driving function and where one must additionally incorporate appropriate scaling to reflect the vertical location of the observation point. Thus, the time-domain field N_X cells in front of the reference plane and N_Y cells above the bottom plate is given by

$$E_z[N_X, N_Y, n] = \sin\left(\frac{2\pi N_Y}{N_C}\right) \mathcal{F}^{-1}\left(G[m]P[m]\right) \tag{19}$$

Assuming the driving function is a Ricker wavelet and incorporating the discrete spectrum given in the appendix, this becomes

$$\begin{aligned} E_z[N_X, N_Y, n] \;=\; & \sin\left(\frac{2\pi N_Y}{N_C}\right) \mathcal{F}^{-1}\left(\frac{2}{f_R\sqrt{\pi}}\left(\frac{mN_R}{sN_T}\right)^2\right. \\ & \left.\exp\left[-\left(\frac{mN_R}{sN_T}\right)^2 - j2\pi\frac{m}{sN_T}\left\{N_R + N_X\left(1-\left(\frac{sN_T}{mN_C}\right)^2\right)^{1/2}\right\}\right]\right) \end{aligned} \tag{20}$$

where $\mathcal{F}^{-1}$ is the inverse discrete Fourier transform, f_R is the most energetic frequency of the Ricker wavelet, and N_R is the number of points per wavelength at this frequency.

Equation (20) can be used to determine analytically the field in the waveguide when the guide is excited by a hard screen (and the source function is a Ricker wavelet)—the hard-wired fields at the source screen are synonymous with the fields at the $x = 0$ reference plane. Fig. 4 shows the field at an observation point obtained from an FDTD simulation employing a hard screen and obtainted from the analytic solution given by (20). In this case, the width of the guide W was such that the cut off wavelength λ_C equaled λ_R, which is the wavelength corresponding to the peak frequency of the Ricker wavelet f_R. The observation point was $\lambda_R/2$ in front of the screen and displaced slightly from midway between the two plates. The relevant discrete parameters were $N_R = 40$, $N_C = 40$, $N_X = 20$, $N_Y = 9$, and $s = 0.95/\sqrt{2}$. For the analytic solution, N_T was 2048. Note that the FDTD simulation is only run for as many time steps as desired (400 steps in this case) and this number is independent of N_T (which must be large enough to ensure negligible frequency aliasing in the analytic solution). Fig. 4 shows excellent agreement between the analytic solution (circles) and the hard-screen FDTD results (solid line). Note that the field at the observation point bears little resemblance to the Ricker wavelet driving function. The persistent ringing is a consequence of having a significant portion of the driving function spectrum below cut off.

4.2 Transparent Field Screen

The transparent screen used to excite a two-dimensional parallel-plate waveguide consists of a column of E_z nodes. Similar to (12), each node is updated by the sum of the usual update equation, the desired source function (i.e., the same source function that is used when driving the hard screen), and the convolution of an impulse response with the source function. However, each node uses a unique impulse response that is measured specifically for the guiding structure under consideration—the homogeneous-media impulse response discussed in the previous section is no longer relevant. It is instructive first to describe how one obtains the appropriate impulse responses for simple multi-element structures before discussing the waveguide implementation.

Assume that one wants to radiate fields from a two-node source. The radiated fields are to be the same as those radiated in a homogeneous medium when both nodes are set by the same driving function, i.e., both nodes are hard and have the same values, but the source itself should be transparent to any reflected energy. To insure transparency, each node must be updated using the standard FDTD update equation; however, to obtain the same radiated field as would be present for hard nodes, each of the transparent source nodes must, in the absence of any scattered field, take on the same values as the hard nodes. As before, a convolution of an impulse response with the source function is used to cancel the "echoed" terms that results from using the update equation at the node. The impulse response for each source node is measured in a grid in which *both* source nodes are driven (in the hard sense) by Kronecker delta functions. And, the impulse response is recorded using an auxiliary simulation. At each time step, the value obtained from the source-node update equation is recorded, but the actual value of the source node is fixed by the delta function. Once the impulse responses have been recorded, they would be used in (12) to realize a transparent two-element source which could, for example, be used to illuminate a scatterer. Any field scattered back to the source would pass through it.

This approach can be generalized to any number of nodes. If all nodes are to be driven by the same source function, then the impulse responses for all the nodes can be obtained from a single auxiliary simulation in which each source node is driven by a unit-amplitude impulse. In the case of the waveguide source screen, the driving function for each node is the same except for the sinusoidal variation of amplitude as a function of y (see Eq. (13)). In cases such as this, where the nodes are driven with the same temporal waveform, but are scaled by different amounts, the impulse response for each node cannot be obtained by using unit-amplitude Kronecker delta functions at all the source nodes in the auxiliary simulation. Instead, the amplitudes of the delta functions must be scaled by the same value that scales the source function in the problem at hand. For a two-node source, if one node is given by $p(n\Delta t)$ and the other by $0.5p(n\Delta t)$, then the impulse responses are obtained by driving the first node with the Kronecker delta function $\delta[n]$ and the other with $0.5\delta[n]$.

One must also account for the walls of the guide when obtaining the waveguide impulse responses. To do this, a simulation is performed in which the walls of the guide are present and the nodes in the source screen are driven impulsively (by Kronecker delta functions with appropriate scaling). The way in which the impulse response is recorded is the same as before.

As mentioned previously, in [3],[4] the implementation of a transparent source in which the source function is merely added to the update equation was described. When the source node is an electric field node, this is equivalent to establishing a current at that node and hence the field radiated by such a source is fundamentally different from that radiated by the transparent field source described here. Note that the only difference in implementation is that the transparent current source does not use the convolution that is employed by the transparent field source. Fig. 4 also shows the field at the observation point of the waveguide problem described previously when the source is implemented either as a transparent field screen or as a transparent current screen. The transparent field screen yields precisely the same result as the hard screen and hence excellent agreement with the analytic solution still holds (the transparent screen results are covered by those of the hard-screen). As is expected, the transparent current screen yields a different result. (Note that it is possible to obtain the analytic solution for the fields in the waveguide when the source is a current screen. In that case, the guide transfer function is modified by $\omega\mu/\left(2\sqrt{\omega^2/c^2 - \pi^2/W^2}\right)$.)

To illustrate further the behavior of a transparent field screen, consider the problem shown in the inset of Fig. 5 in which there are two discontinuities in an otherwise homogeneous waveguide. The ultimate goal might be, for example, to determine the field in the vicinity of these discontinuities under a particular pulsed TE_1 excitation. Other than the discontinuities, the guide geometry, the observation point, and the excitation are the same as those that pertained for Fig. 4. In one simulation the discontinuities are illuminated by a hard screen and in the other a transparent field screen. The screens are placed one half of a cut-off wavelength away from the leading edge of the first discontinuity. The transparent screen is backed by a homogeneous section of waveguide that is five cells long. This backing section is then terminated in an eight-cell perfectly-matched layer (PML) [10]. (The standard split-cell PML with central differencing was used. No modifications were made to absorb evanescent energy. However, to test the effects of the ABC, another simulation was performed where the section of the waveguide behind the screen was made long enough to causally isolate the grid termination from the observation point. Those results were virtually indistinguishable from those obtained using the eight-cell PML.) The field at the observation point for the two simulations is shown in Fig. 5. The large ringing seen in the hard-screen result is a consequence of the trapping of energy between the discontinuities and the screen itself, i.e., it is an artifact of the source implementation and not truly indicative of the way in which the discontinuities themselves trap energy. The result obtained using the transparent source, on the other hand, provides an accurate indication of the behavior of the discontinuities in an otherwise uniform segment of waveguide.

5 Conclusions

By convolving the driving function and a grid impulse response, 1-D, 2-D, and 3-D transparent field sources can be created that radiate the same fields as hard sources but that do not scatter energy themselves. Multiple transparent field sources can be used in the same simulation and, if necessary, used in adjacent nodes. This permits the creation of a wide variety of excitations that would be difficult or impossible to achieve otherwise. Transparent field sources also can be used to finely control the excitation of resonant structures without affecting the resonances. Though more expensive to implement than hard sources, the impulse response required to implement a transparent field source must be calculated only once and can be saved for subsequent simulations. The cost of the convolution is typically small compared to other computations for realistic two- and three-dimensional simulations.

Appendix

The Ricker wavelet is used for the driving function $p(t)$ in all the simulations presented here. The Ricker wavelet is equivalent to the second derivative of a Gaussian; it is simple to implement; it has no dc component; and, its spectral content is fixed by a single parameter. The Ricker wavelet is typically written

$$p(t) = \left(1 - 2\pi^2 f_R^2 \left(t - t_R\right)^2\right) \exp\left[-\pi^2 f_R^2 \left(t - t_R\right)^2\right] \tag{21}$$

where f_R is the peak frequency and t_R is the temporal delay. The "peak frequency" is the frequency with the greatest spectral content. The delay can be set to any desired amount, but it is convenient to express it as a multiple of $1/f_R$. Here we use $t_R = 1/f_R$. The FDTD simulation is assumed to start at $t = 0$, but $p(t)$ is not zero for $t < 0$. However, with a delay of $t_R = 1/f_R$, $|p(t < 0)|$ is bound by 0.001, which is small compared to the peak value of unity. Thus, the transient caused by "switching on" $p(t)$ at $t = 0$ is relatively small. (For applications that demand a smoother transition, the bound on $|p(t < 0)|$ can be made arbitrarily small by increasing t_R.) The Fourier transform of (21) is

$$P(\omega) = \frac{2}{f_R\sqrt{\pi}} \left(\frac{\omega}{2\pi f_R}\right)^2 \exp\left[-jt_R\omega - \left(\frac{\omega}{2\pi f_R}\right)^2\right] \tag{22}$$

The functions $p(t)$ and $|P(\omega)|$ are shown in Fig. 6. For the sake of illustration, we have arbitrarily chosen f_R to be 1 Hz. Different values of f_R change the horizontal scale but they they do not change the general shape of the curve. To obtain unit amplitude at the peak frequency, $P(\omega)$ has been scaled by $f_R\pi e/2$.

In an FDTD simulation the discrete function $p[n] = p(n\Delta t)$ is used instead of the continuous one. This function can be expressed in terms of dimensionless quantities as follows. First, assume a uniform spatial step of δ and let the Courant number $c\Delta t/\delta$ be s. Further assume that the spatial step size is such that there are N_R spatial steps at the wavelength λ_R corresponding to the peak frequency of the Ricker wavelet, i.e., $\lambda_R = N_R\delta$. While the value of s is dictated by the stability limit, N_R is *chosen* to ensure that the peak frequency is sampled at N_R points per wavelength. Writing $\delta = c\Delta t/s$, $\lambda_R = c/f_R = ct_R$, and $\lambda = N_R\delta$, one obtains

$$t_R = \frac{1}{f_R} = N_R\Delta t/s \tag{23}$$

Letting time t be $n\Delta t$ and expressing t_R and f_R as in (23), the discrete form of (21) can be written as

$$p[n] = \left(1 - 2\pi^2 \left(\frac{sn}{N_R} - 1\right)^2\right) \exp\left[-\pi^2 \left(\frac{sn}{N_R} - 1\right)^2\right] \tag{24}$$

The parameters that specify $p[n]$ are the Courant number s and the points per wavelength at the peak frequency N_R. This function appears to be independent of the temporal and spatial step sizes, but it does depend on their ratio via the Courant number.

For a time-domain simulation in which the total number of time steps is N_T, the discrete spectrum $P[m]$ is given by $P(\omega)$ with ω replaced by $2\pi m\Delta f$ in (22). Making this substitution and using (23) yields

$$P[m] = \frac{2}{f_R\sqrt{\pi}} \left(\frac{mN_R}{sN_T}\right)^2 \exp\left[-j2\pi\frac{mN_R}{sN_T} - \left(\frac{mN_R}{sN_T}\right)^2\right] \tag{25}$$

Again, this is independent of the spatial and temporal step sizes (except via their ratio in the Courant number), but it does depend on the additional parameter N_T. As was the case for (18), (25) is based on the assumption that N_T is large enough so that no frequency aliasing occurs. However, this number is not tied to the number of time steps in any actual FDTD simulation. It merely has to be sufficiently large for purposes of obtaining the unaliased analytic solution.

Acknowledgments

The authors would like to thank Dr. Kurt L. Shlager of Lockheed Martin, Sunnyvale, CA for helpful comments in the preparation of this paper.

This work was supported by the Office of Naval Research, Code 321OA.

References

[1] A. Taflove, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Boston, MA: Artech House, 1995.

[2] K. S. Kunz and R. J. Luebbers, *The Finite Difference Time Domain Method for Electromagnetics*. Boca Raton, FL: CRC Press, 1993.

[3] A. P. Zhao, A. V. Räisänen, and S. R. Cvetkovic, "A fast and efficient FDTD algorithm for the analysis of planar microstrip discontinuities by using a simple source excitation scheme," *IEEE Microwave Guided Wave Lett.*, vol. 5, no. 10, pp. 341–343, 1995.

[4] A. P. Zhao and A. V. Räisänen, "Application of a simple and efficient source excitation technique to the FDTD analysis of waveguide and microstrip circuits," *IEEE Trans. Microwave Theory Tech.*, vol. 44, no. 9, pp. 1535–1539, 1996.

[5] D. N. Buechler, D. H. Roper, C. H. Durney, and D. A. Christensen, "Modeling sources in the FDTD formulation and their use in quantifying source and boundary condition errors," *IEEE Trans. Microwave Theory Tech.*, vol. 43, no. 4, pp. 810–814, 1995.

[6] Z. Wu and J. Fang, "Numerical implementation and performance of perfectly matched layer boundary condition for waveguide structures," *IEEE Trans. Microwave Theory Tech.*, vol. 43, no. 12, pp. 2676–2683, 1995.

[7] O. M. Ramahi, "Complementary boundary operators for wave propagation problems," *J. Comput. Phys.*, vol. 133, pp. 113–128, 1997.

[8] M. Celuch-Marcysiak, A. Kozak, and W. K. Gwarek, “A new efficient excitation scheme for the FDTD method based on the field and impedance template,” in *IEEE Antennas and Propagat. Soc. Int. Symp.*, vol. 2, (Baltimore, MD), pp. 1296–1299, July 1996.

[9] M. Celuch-Marcysiak and W. K. Gwarek, “A novel variable source impedance technique for emulating unperturbed eigenmodes in the FD-TD method,” in *Proc. 3rd IEE Conf. on Computation in Electromagnetics*, (Bath, England), pp. 277–282, Apr. 1996.

[10] J.-P. Berenger, “A perfectly matched layer for the absorption of electromagnetic waves,” *J. Comput. Phys.*, vol. 114, no. 1, pp. 185–200, 1994.

	$E_z(i_{\text{src}}-2)$	$H_y(i_{\text{src}}-2)$	$E_z(i_{\text{src}}-1)$	$H_y(i_{\text{src}}-1)$	$E_z(i_{\text{src}})$	$H_y(i_{\text{src}})$	$E_z(i_{\text{src}}+1)$	$H_y(i_{\text{src}}+1)$	$E_z(i_{\text{src}}+2)$
$n=0$					f^0				
$n=\frac{1}{2}$				$\frac{1}{Z}f^0$		$-\frac{1}{Z}f^0$			
$n=1$			f^0		f^1		f^0		
$n=\frac{3}{2}$		$\frac{1}{Z}f^0$		$\frac{1}{Z}f^1$		$-\frac{1}{Z}f^1$		$-\frac{1}{Z}f^0$	
$n=2$	f^0		f^1		f^2		f^1		f^0

Figure 1: Values of E_z and H_y in a one-dimensional grid when the node $E_z(i_{\text{src}})$ is implemented as a hard source. The values assume the Courant number s is unity. Node location is given along the top and the time step is indicated along the left. A blank indicates the field is zero.

	$E_z(i_{\text{src}}-2)$	$H_y(i_{\text{src}}-2)$	$E_z(i_{\text{src}}-1)$	$H_y(i_{\text{src}}-1)$	$E_z(i_{\text{src}})$	$H_y(i_{\text{src}})$	$E_z(i_{\text{src}}+1)$	$H_y(i_{\text{src}}+1)$	$E_z(i_{\text{src}}+2)$
$n=0$					f^0				
$n=\frac{1}{2}$				$\frac{1}{Z}f^0$		$-\frac{1}{Z}f^0$			
$n=1$			f^0		f^1-f^0		f^0		
$n=\frac{3}{2}$		$\frac{1}{Z}f^0$		$\frac{1}{Z}(f^1-f^0)$		$-\frac{1}{Z}(f^1-f^0)$		$-\frac{1}{Z}f^0$	
$n=2$	f^0		f^1-f^0		$f^2-f^1+f^0$		f^1-f^0		f^0

Figure 2: Values of E_z and H_y in a one-dimensional grid when the node $E_z(i_{\text{src}})$ is given by the sum of the usual update equation and the driving function f^n. The values assume the Courant number s is unity.

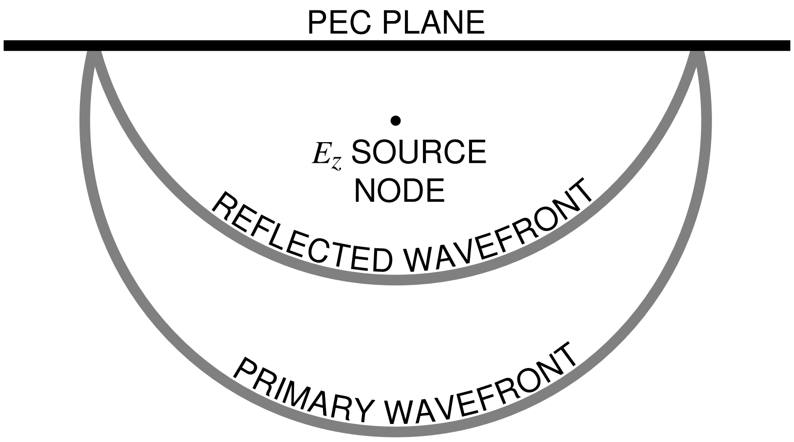

Figure 3: Electric field about an E_z source node in the vicinity of a PEC surface after 220 time steps. (a) Sketch of problem geometry. (b) Hard source. (c) Transparent source.

Figure 3(b):

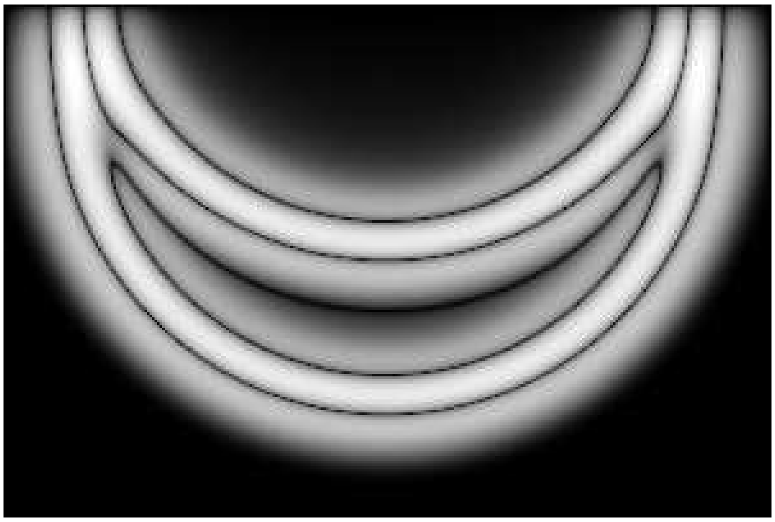

Figure 3(c):

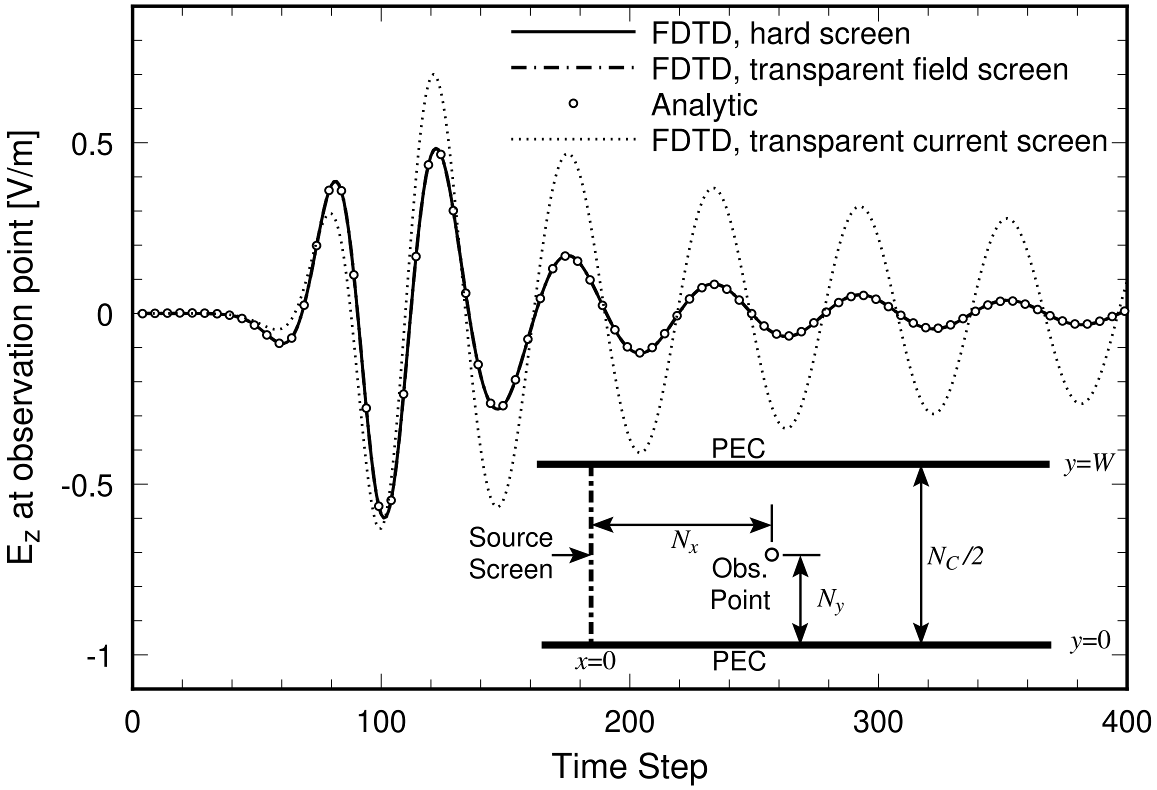

Figure 4: Field at an observation point nearly mid-way between the plates and one-half of a cut-off wavelength in front of the source screen. The cut-off wavelength λ_C is equal to the wavelength at the peak frequency of the pulse λ_R. The analytic solution (for a field source) was obtained using a 2048-point DFT. The sampling in the FDTD simulation was such that there were 40 cells per wavelength at the peak frequency of the Ricker wavelet. The source is implemented either as a transparent field screen, a hard field screen, or a transparent current screen. The results for the transparent field screen are identical to those of the hard screen and the two results appear as a single curve passing through the analytic results.

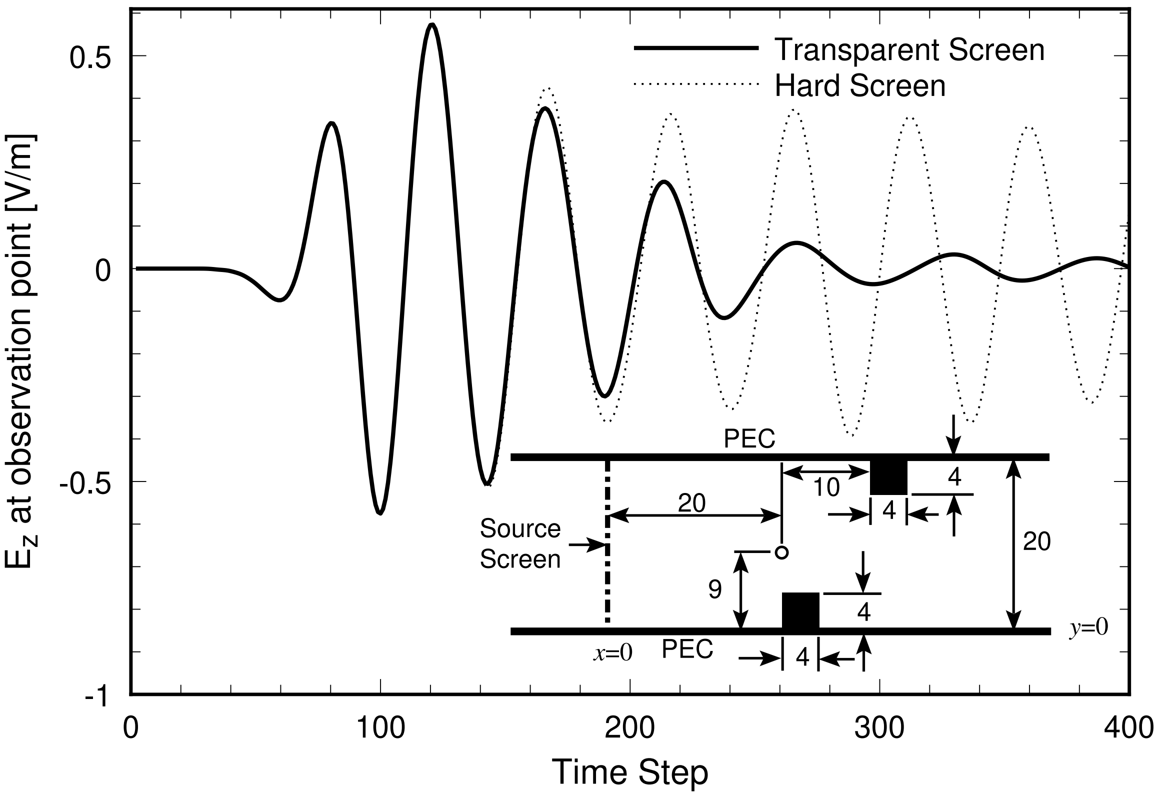

Figure 5: Field at an observation point in a guide with two discontinuities. The inset figure shows the parallel-plate waveguide geometry. Dimensions are in number of cells. The small circle 20 cells in front of the screen indicates the location of the observation point. The excitation is such that there are 40 cells per wavelength at the peak frequency. The source screen is implemented either as hard or transparent yielding the dotted line and the solid line, respectively.

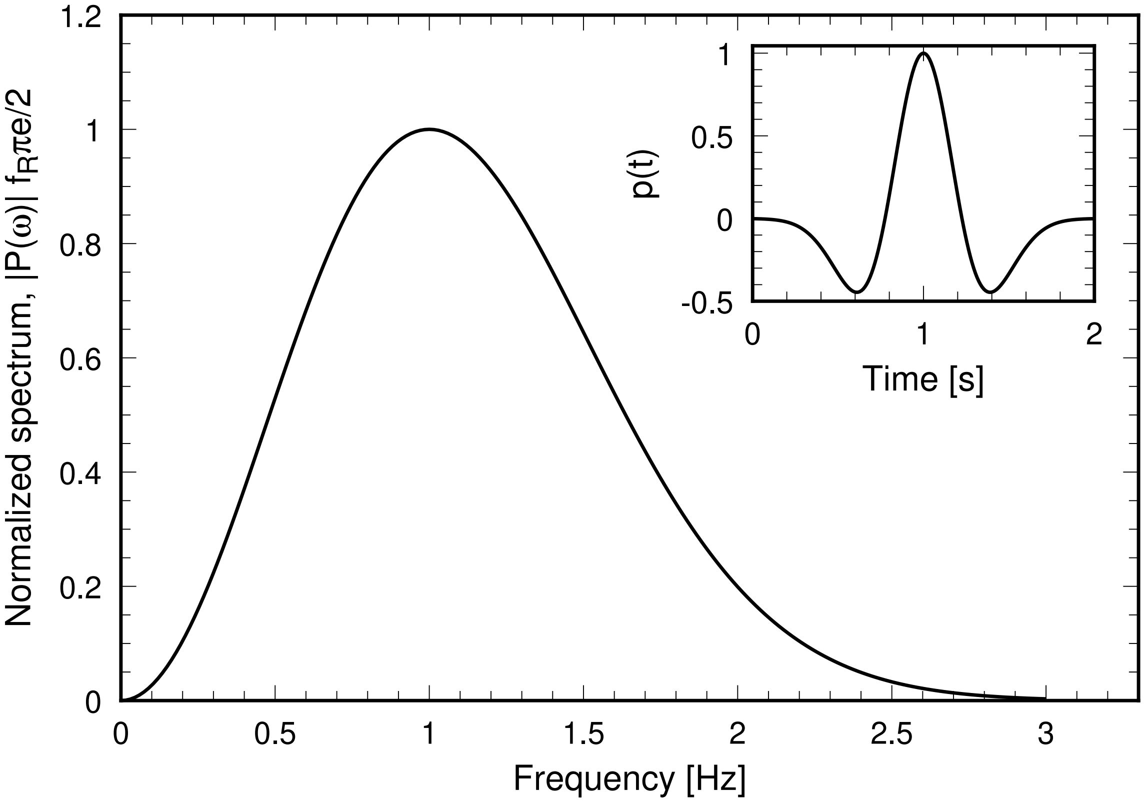

Figure 6: Normalized spectrum of the Ricker wavelet with $f_R = 1$ Hz. The corresponding temporal form $p(t)$ is shown in the inset box. For other values of f_R, the horizontal axis in the time domain is scaled by $1/f_R$. For example, if f_R were 1 MHz, the peak would occur at 1 μs rather than at 1 s. In the spectral domain, the horizontal axis is directly scaled by f_R so that if f_R were 1 MHz, the peak would occur at 1 MHz.