Implementation of transparent sources embedded in acoustic finite-difference time-domain grids

John B. Schneider, a) Christopher L. Wagner, and Shira L. Broschat

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

(Received 12 February 1997; accepted for publication 23 September 1997)

The finite-difference time-domain (FDTD) method is a simple but powerful numerical method which has been used to perform a wide variety of complex simulations. One of the considerations in using this method is modeling the source of the incident field. When the physical source of acoustic energy is located within the FDTD grid it has typically been modeled as a “hard” source for which a source function determines the field at a source node. One drawback of a hard source is that it acts itself as a scatterer. Moreover, its scattering cross section is dictated by the spatial step size within the grid; thus, scattering is independent of the underlying physical size of the source. In this paper a scheme is presented that permits the creation of completely transparent sources. These sources, although more expensive to implement than traditional hard sources, radiate the same field as a hard source, but do not scatter energy. © 1998 Acoustical Society of America.

PACS numbers: 43.20.Px, 43.55.Ka [JEG]

INTRODUCTION

The finite-difference time-domain (FDTD) method was first introduced by Yee in 19661 for the study of electromagnetic scattering problems. A counterpart was later developed in acoustics which has been used to study a wide variety of problems, including acoustic propagation in shallow water and in large rooms. The FDTD method permits the use of several different types of insonification. The source can be within the computational grid so that the near-field interaction of the source and its surroundings can be modeled. Alternatively, the source can be located outside the grid (the insonifying field is then introduced using one of the techniques briefly described below). This readily permits the study of the interaction of the source far-field with any scattering object or with the local “environment” (such as a rough surface).

When a source is external to the FDTD grid, the field radiated by that source, i.e., the incident field, is coupled into the grid typically by means of a total-field/scattered-field formulation or a scattered-field formulation. In the total-field/scattered-field formulation (e.g., section 6.5 of Ref. 2) 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 (e.g., section 6.6 of Ref. 2 or 3.4 of Ref. 3), the scattered field radiates directly from any material that differs from the background medium. This latter approach requires calculation of the incident field over all such materials. Both approaches are “mature” and have well-understood advantages and disadvantages.

For many simulations, however, the source of energy must be embedded within the FDTD grid, e.g., for resonators or radiating structures. Most commonly sources are embedded as “hard” sources (e.g., section 6.4 of Ref. 2 or the velocity membrane described in Ref. 4). A hard source is implemented simply by specifying the field at a given node with a source, or “driving,” function. Since the update equation does not apply to this source node, its value is fixed solely by the driving function so that it effectively scatters any field incident upon it. In most instances 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 vanishes 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 to the source. Unfortunately, in many circumstances this requirement is overly restrictive. In this paper the implementation of a source that radiates the same field as a hard source, but that does not scatter, is presented. We call such a source “transparent.”

A node in a FDTD grid that has the same material properties as its neighbors and that is governed by the standard update equation does not, per se, scatter energy. It therefore appears that one may simply implement a transparent 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. Indeed, this approach yields a node that does not act as a scatterer, but the field that radiates from it may bear little resemblance to the actual driving function. In fact the field that radiates from the node, as compared to that from a hard source, is a filtered version of the driving function. The filter transfer function is dictated by the difference between a hard (boundary value) source and the additive (volume) source. (For a discussion of these different sources the reader is directed to section 7.3 of Ref. 5.) Unfortunately,
it does not appear possible to obtain a simple closed-form expression for this transfer function. Thus, there is no way to obtain analytically a function that can be used to drive a simple additive source so that it will radiate the same field as a hard source. However, it is possible to measure, directly from the finite-difference grid, the filter impulse response at the source node and then use it to construct a transparent source that radiates the same field as a hard source. The impulse response is measured at the source node and is fundamentally different from the time-domain Green’s function (which is 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 the Courant limit is used. In two and three dimensions, and in one dimension for Courant numbers other than the limit, the impulse response is infinite in duration. For the remainder of the paper, unless stated otherwise, “transparent source” implies a source that radiates the same field as a hard source (thus this source is distinct from a simple additive source).

Perhaps the simplest way to implement a transparent source, and the one used in this study, is first to run a 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 (symmetry can be exploited to reduce the size of the simulation as discussed later). The transparent source is then realized, in part, by convolving the impulse response with the driving function. Once found, the impulse response can be used for all subsequent simulations that have the same number of dimensions and the same Courant number, i.e., it is independent of any other aspect of the problem under consideration. This approach is only appropriate for a single-node source or a collection of noninteracting source nodes (as would be used in a “classic” phased array). Nevertheless, it can be extended to permit the construction of multiple, interacting source nodes. Thus, it is possible to construct transparent source “screens” that excite a system exactly as would a hard boundary-valued source. These transparent screens can be used, for example, to excite waveguides and, unlike hard screens, a transparent screen can be placed immediately adjacent to any waveguide discontinuity. (A full discussion of transparent screens is presented in Ref. 6.)

In Sec. I implementation of transparent sources in one dimension is described, and in Sec. II implementation in two and three dimensions is described.

I. ONE-DIMENSIONAL TRANSPARENT SOURCES

Although a transparent source in one dimension is of little practical use, it is helpful first to consider implementation in one dimension (the extension to two and three dimensions is trivial). The differential equations governing acoustic propagation are

\[ \frac{\partial \mathbf{v}}{\partial t} = -\frac{1}{\rho} \nabla \rho, \]  
\[ \frac{\partial \rho}{\partial t} = -c^2 \rho \nabla \cdot \mathbf{v}, \]

where \( c \) is the speed of sound, \( \rho \) is the density, \( p \) is the pressure, and \( \mathbf{v} \) is the velocity. We reduce the problem to one dimension so there is no variation in the \( y \) and \( z \) directions. By employing finite differences to approximate the derivatives and offsetting the pressure and velocity evaluation points, both temporally and spatially, one can solve for future field values in terms of current and past field values (see, for example, Ref. 4 or 7). This leads to the standard update equations for the FDTD method in one dimension:

\[ v_{x}^{n+1/2}(i) = v_{x}^{n-1/2}(i) - \frac{1}{c \rho \Delta x} \left[ p_{x}^{n+1}(i) - p_{x}^{n}(i) \right], \]  
\[ p_{x}^{n+1}(i) = p_{x}^{n}(i) - c \rho \left( \frac{c \Delta t}{\Delta x} \left[ v_{x}^{n+1/2}(i) - v_{x}^{n+1/2}(i-1) \right] \right), \]

where \( \Delta x \) and \( \Delta t \) are the spatial and temporal step sizes, respectively, the superscript indicates the time step, and the argument indicates the spatial location so that \( p_{x}^{n}(i) = p(i \Delta x, n \Delta t) \) and \( v_{x}^{n+1/2}(i) = v_{x}(i + 1/2) \Delta x, [n + 1/2] \Delta t \). For brevity, the spatial offset between the \( p \) and \( v \) nodes is suppressed in the arguments of the discrete forms. The term \( c \Delta t / \Delta x \), hereafter identified as \( s \), is the Courant number which must be chosen to satisfy the stability requirements:

\[ s \leq 1/\sqrt{N}, \]

where \( N \) is the number of spatial dimensions. The maximum value of \( s \) yields the minimum amount of numerical dispersion and the longest simulation duration for a given number of time steps. Although the maximum value of \( s \) should be used whenever possible, the Courant limit cannot be used throughout the computational domain for simulations of inhomogeneous regions since, for stability, the limit must hold in the fastest medium as discussed below. Thus it is important to consider source implementations that do not restrict the Courant number. Note that in one dimension with an \( s \) of unity the FDTD algorithm yields an exact solution for propagation in a homogeneous medium.

Consider a one-dimensional (1-D) computational domain in which the source is a pressure node at \( i_{0} \). A hard source is realized by setting the source node equal to a given driving function \( f(n \Delta t) = f^{0} \). The pressure at the source node is then \( p^{n}(i_{0}) = f^{0} \), 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 \( p \) and \( v \) 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 pressure at node \( i \) and time step \( n \) is given by

\[ p^{n}(i) = f^{n-|i-i_{0}|}. \]  

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. Although we explicitly consider only pressure sources here, the transparent-source implementation presented applies equally to pressure and velocity sources.
With the goal of creating a transparent 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

\[ p^{n+1}(i_{\text{src}}) = p^n(i_{\text{src}}) - c \rho s [v_x^{n + 1/2}(i_{\text{src}}) - v_x^{n + 1/2}(i_{\text{src}} - 1)] + f^{n+1}. \]

Figure 2 shows the values of \( p \) and \( v_x \) 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 pressure at an arbitrary node is given by

\[ p^i = \sum_{m=0}^{n-|i-i_{\text{src}}|} (-1)^{m+n-|i-i_{\text{src}}|} f^m. \]

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." Unfortunately, the field that is radiated by the source node may not resemble the driving function as desired. To illustrate this, 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 radiated field 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 (5), 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 source that radiates the same field as the hard source can be achieved using

\[ p^{n+1}(i_{\text{src}}) = p^n(i_{\text{src}}) - c \rho s [v_x^{n + 1/2}(i_{\text{src}}) - v_x^{n + 1/2}(i_{\text{src}} - 1)] + f^{n+1} + f^n. \]

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

One-dimensional FDTD simulations performed using Courant numbers other than the limit do not permit such a simple implementation of a transparent 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. This is a consequence of the inherent numerical dispersion.

Numerical dispersion is a function of the number of points per wavelength for a given spectral component of the signal. Thus, the spatial step size, \( \Delta x \), typically is determined by the amount of dispersion that can be tolerated and then, assuming one dimension, the Courant number is dictated by the stability requirement \( \Delta t = \Delta x / c_{\text{fastest}} \), where \( c_{\text{fastest}} \) is the speed of sound in the fastest material within the grid. For the source implementation considered here, numeri-
cal dispersion is not an issue. What is significant, as shown shortly, is that a Courant number less than unity results in echo terms that depend on the entire history of the driving function.

To facilitate the construction of transparent 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 used at the source node and the value returned is recorded as part of the impulse response. However, the value of the source node is not set to this value.) Thus, the impulse response is calculated from the previous value of the source node and its surrounding velocity 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 $p^n(i_{src}) = \delta[n]$ while the impulse response is

$$I^n = p^{n-1}(i_{src}) - c \rho s[v_x^{n-1/2}(i_{src}) - v_x^{n-1/2}(i_{src} - 1)].$$  (9)

Using the Courant limit, $s = 1$, the impulse response is $I^n = -\delta[n-1]$. For Courant numbers less than unity, the impulse response is infinite in duration. 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

$$I^0 = 0,$$
$$I^1 = 1 - 2s^2,$$
$$I^2 = -2s^2 + 2s^4,$$
$$I^3 = -2s^2 + 6s^4 - 4s^6,$$
$$I^4 = -2s^2 + 12s^4 - 20s^6 + 10s^8,$$
$$I^5 = -2s^2 + 20s^4 - 60s^6 + 70s^8 - 28s^{10}.$$

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 a simple 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 the simulation, a hard source is driven impulsively and the impulse response is recorded using (9). Alternatively, symmetry can be exploited, since the velocity $v_x^{n-1/2}(i_{src} - 1)$ is the negative of $v_x^{n-1/2}(i_{src})$, so that only half the 1-D space is needed. The impulse response can then be found using

$$I^n = p^{n-1}(i_{src}) = 2c \rho s v_x^{n-1/2}(i_{src}),$$  (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 $p^0 = f^0$, the source node at the next time step is $p^1 = f^1 + I^1 p^0$; at the next it is $p^2 = f^2 + I^1 f^1 + I^2 p^0$; 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 $p^n = f^n$, $p^1 = f^1$, $p^2 = f^2$, etc. The cancellation 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. Stated 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

$$p^{n+1}(i_{src}) = p^n(i_{src}) - c \rho s [v_x^{n+1/2}(i_{src}) - v_x^{n+1/2}(i_{src} - 1)]$$
$$+ f^{n+1} - \sum_{m=0}^{n} I^{n-m+1} f^m. \quad (11)$$

Although the discussion has been for the implementation of a single pressure source, the same approach can be used when the source is a velocity node, and multiple (non-interacting) sources can exist in the same computational domain. In one dimension, the impulse response for a velocity node is the same as for a pressure node. In two and three dimensions, the impulse responses for pressure and velocity nodes differ, but the basic source implementation, which is described in the next section, is the same.

II. TRANSPARENT SOURCES IN TWO AND THREE DIMENSIONS

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

$$p^{n+1}(r_{src}) = (N - D \text{ update equation}) + f^{n+1}$$
$$- \sum_{m=0}^{n} I^{n-m+1} f^m, \quad (12)$$

where $N$ is the number of dimensions, $r_{src}$ is the source location, “$N$-D update equation” is the update equation for a pressure node for the given number of dimensions, and $I_x$ 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. (Note that the goal is to obtain transparent sources that radiate the same fields as a hard source in an unbounded medium. It is also possible to obtain the impulse response that would correspond to a hard source in a bounded medium or with a scatterer present. In these cases the impulse response must be measured with any scatterers present. This distinction is important, for example, when using a transparent screen to excite a waveguide.)

As was the case for one dimension, symmetry can be exploited when measuring the impulse response in two and three dimensions. For a pressure source located at the origin, the horizontal and vertical components of velocity satisfy
\[ v_x(x, y) = -v_x(-x, y), \]  
\[ v_y(x, y) = -v_y(x, -y), \]  
\[ v_z(x, y) = v_z(y, x). \]  

With (13) the computational domain can be divided in half and only the “right” half retained; with (14) the remaining computational domain can be divided in half and only the “top” half retained; and, finally, with (15) the remaining computational domain can be cut in half along a diagonal and only the lower right half retained. In this way the impulse response can be found using a computational space that is only one-eighth the size of a full two-dimensional space. This is important when the impulse response must be determined over a long duration. In addition, the reduction of the grid size is of practical importance since the impulse response should be calculated in an “unbounded” homogeneous grid. Because the driving function is an impulse, termination of the grid with artificial absorbing boundary conditions (ABC’s) cannot provide an adequate model of an unbounded medium. Instead, the computational domain should be large enough so that reflections from the termination of the computational domain are insignificant over the duration of interest. This obviates the need for ABC’s and the edge of the computational domain can be left perfectly reflecting.

In three dimensions, for a pressure node at the origin, the following hold:
\[ v_x(x, y, z) = -v_x(-x, y, z), \]  
\[ v_y(x, y, z) = -v_y(x, -y, z), \]  
\[ v_z(x, y, z) = -v_z(x, y, -z), \]  
\[ v_z(x, y, z) = v_z(y, x, z), \]  
\[ v_z(x, y, z) = v_z(x, z, y), \]  
\[ v_z(x, y, z) = v_z(z, y, x). \]  

These symmetry relations can be exploited to reduce the size of the computational domain needed to obtain the impulse response to \( \frac{1}{8} \) the size of a full three-dimensional computational domain.

To illustrate how the impulse response is obtained in two dimensions, consider a problem which must be run for 16 000 time steps and for which the Courant number “seen” by the source is the Courant limit. To obtain an impulse response that is completely free of any boundary artifacts over 16 000 time steps, a simulation could be done using an 8000 by 8000 cell grid with the source at the center. Since there are three fields per cell, a pressure node and two velocity nodes, this requires a simulation with 192 million unknowns. By exploiting symmetry, however, this number is reduced to 24 million which, unlike the original number of unknowns, is handled easily on inexpensive workstations. In practice, 24 million unknowns, which guarantees the absence of boundary artifacts in the impulse response, is excessive. In two dimensions, the magnitude of the leading edge of the impulse is reduced by the square of the Courant number as it travels to successive neighbors. Thus, after the leading edge of the impulse has traveled from the source to the edge of the computational domain and back it is reduced by \( s^{2\cdot000} \). (This reduction is, however, purely theoretical. In practice, for both two and three dimensions, this theoretical reduction is smaller than the smallest number that can be represented using double-precision numbers.) In two dimensions the maximum value of \( s \), i.e., the Courant limit, is \( 1/\sqrt{2} \). In three dimensions, the Courant limit is \( 1/\sqrt{3} \) and the leading edge of the impulse dies out even more quickly. Therefore, the simulation domain may be reduced further in size without incurring significant errors in the impulse response.

Since an impulse response only needs to be calculated once and then can be recalled for use in any simulation that has the same dimensions and the same Courant number, it is anticipated that, by exploiting symmetry and using a powerful workstation, impulse responses of sufficient duration can be obtained to satisfy the requirements of the vast majority of problems. Thus, for example, if an impulse response of 16 000 time steps has been obtained, it subsequently can be used for any problem that uses 16 000, or fewer, times steps. However, some problems may require an exceedingly large number of time steps for which a completely “clean” impulse response does not exist. For these cases is there a way to obtain the impulse response in a reasonable fashion? There appear to be simple solutions in one and three dimensions, but not in two dimensions. In one dimension, the impulse response will decay to zero and the rate at which it approaches zero is a function of the Courant number. The closer the Courant number is to the stability limit, the more rapidly the response approaches zero. Therefore, after a sufficient number of time steps, the impulse response can be approximated by zero. The actual number of steps beyond which this approximation can be employed is determined by inspection of the impulse response for a given Courant number.

In three dimensions, the impulse response approaches a nonzero constant that depends on the Courant number. The “envelope” of the deviation from this constant is inversely proportional to the time step. Hence, the longer the simulation, the closer the impulse response is bound to this constant. Figure 3 shows the 3-D impulse responses over 280 time steps that correspond to Courant numbers of \( 1/\sqrt{3} \) (i.e., the limit), 0.40, 0.25, and 0.10. The inset shows the case for the Courant number \( 1/\sqrt{3} \) with the vertical scale expanded by a factor of approximately 2500. Clearly the fluctuations are small and decreasing. Though not shown, expanded views for the other Courant numbers exhibit similar behavior. Thus, after a sufficient number of time steps the impulse response can be approximated by this constant.

Figure 4 shows the 2-D impulse response over 750 time steps for four different Courant numbers. Unlike in one and three dimensions, the response does not quickly converge to a constant value, but rather decays very slowly to zero. 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 sources, consider a two-dimensional point source (i.e., a line source in three dimensions) near a planar pressure-release surface as shown in Fig. 5(a). The source is a pressure node and the driving function is a Ricker wavelet. 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 to yield the Courant limit. Figure 5(b) and (c) shows the pressure in the vicinity of a hard and a transparent 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 pressure found at the corresponding node. The hard source, Fig. 5(b), while radiating the same primary field as the transparent source, scatters the reflected field as evidenced by the nonblack region between the reflected wave and the pressure-release surface. For the transparent source, Fig. 5(c), the primary wave is identical to that of the hard source, but the source does not interfere with the field reflected by the pressure-release surface.

III. CONCLUSIONS

By convolving the driving function and the grid impulse response, 1-D, 2-D, and 3-D transparent sources can be created that radiate the same fields as hard sources but that do not scatter energy themselves. Multiple transparent sources can be used in the same simulation and, if necessary, used in adjacent nodes. This permits the creation of a wide variety of insonifications that would be difficult or impossible to achieve otherwise. Transparent 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 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.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research, Code 321OA.
5P. M. Morse and H. Feshbach, Methods of Theoretical Physics (McGraw–Hill, New York, 1953).