J. Webster (ed.), *Wiley Encyclopedia of Electrical and Electronics Engineering*

TIME-DOMAIN ANALYSIS

The study of many linear physical phenomena and processes, including electromagnetic wave radiation, propagation, and scattering, can in principle be done either in the *time domain* or in the *frequency domain*.

In time-domain analysis, the study is done by considering the time as an independent variable. Fields, signals, and systems are considered as explicit functions of time and mathematically described through their natural evolution from the past towards the future. In such time-domain description, time can be treated either as a continuous variable (*continuous time*) or as a discrete variable (*discrete time*) (1,2). Traditionally, continuous-time problems are usually associated with electric circuits, communication, and physics problems, while discrete-time problems are usually associated with time-series analysis, statistical problems, and numerical analysis. However, with the advent of high-speed digital computers, it has been increasingly advantageous to treat phenomena that occur in the time continuum as discrete-time processes through some process of sampling. This has led to an increasing connection between continuous time-domain and discrete time-domain analysis.

Frequency-domain analysis, on the other hand, is done by treating the frequency as an explicit variable. This is because physical processes and signals described in the time domain can often (advantageously) be *spectrally* (Fourier) decomposed, i.e., decomposed into a (possibly infinite) sum of elementary modes or signals, each one with a definite frequency (periodicity). These elementary components are called spectral (Fourier) components. The relative weights of these spectral components in the sum correspond to an alternative (dual) description of the original signal, where now the frequency is taken as an explicit variable. Such spectral decomposition is useful because many signals allow for a much simpler representation in the frequency domain than in the time domain.

Mathematically, the passage from the time domain to the frequency domain is done through the a Fourier transform pair [1,2,3,4,5]. Given a time-domain function $\phi(t)$, its Fourier transform $\tilde{\Phi}(\omega)$ is defined by

$$\tilde{\phi}(\omega) = \mathcal{F}[\phi(t)] = \int_{-\infty}^{+\infty} dt\, e^{i\omega t}\phi(t) \tag{1a}$$

Furthermore, the original time-domain function $\phi(t)$ can be easily reconstructed from $\tilde{\Phi}$ through the use of the *inverse* Fourier transform

$$\phi(t) = \mathcal{F}^{-1}[\tilde{\phi}(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega\, e^{-i\omega t}\tilde{\phi}(\omega) \tag{1b}$$

In the above expressions, we have used the so-called $e^{-i\omega t}$ convention (3,4,5), common in optics and physics. An alternative convention is the $e^{j\omega t}$ (1,2), common in circuit analysis and signal processing. One can go from one convention to the other by simply replacing i with $-j$ in all expressions or vice versa.

2 TIME-DOMAIN ANALYSIS

The Fourier transform pair establishes a one-to-one relationship between the time-domain and frequency-domain descriptions. There are some mathematical conditions to observe on the function $\phi(t)$ for its Fourier transform to exist. These conditions are related to the convergence of the integrals in Eqs. (1a). Discussion of those aspects is beyond the scope of this article. The interested reader may consult, for example, Refs. 1,2,3. Suffice is to say that, for most functions of interest in practice, a Fourier transform pair is well defined.

Many physical processes of interest are described by linear time-invariant (LTI) differential equations. For instance, Maxwell's equations, which govern the behavior of the electromagnetic fields, are written as (4,5)

$$\mu(\boldsymbol{r})\frac{\partial \boldsymbol{H}}{\partial t} = -\nabla \times \boldsymbol{E} \tag{2a}$$

$$\epsilon(\boldsymbol{r})\frac{\partial \boldsymbol{E}}{\partial t} + \boldsymbol{J} = \nabla \times \boldsymbol{H} \tag{2b}$$

$$\nabla \cdot \boldsymbol{E} = \frac{\rho}{\epsilon}(\boldsymbol{r}) \tag{2c}$$

$$\nabla \cdot \boldsymbol{H} = 0 \tag{2d}$$

where $\boldsymbol{E}$ is the electric field, $\boldsymbol{H}$ is the magnetic field, $\boldsymbol{J}$ is the electric current density, and ρ the electric charge density. If $\epsilon(\boldsymbol{r})$ and $\mu(\boldsymbol{r})$ are functions only of position, the above equations constitute a linear, time-invariant system. *Linearity* means that if we take two set of possible field solutions of Eq. (2a), say $\{\boldsymbol{E}_1, \boldsymbol{H}_1\}$ and $\{\boldsymbol{E}_2, \boldsymbol{H}_2\}$, then any linear combination $\{\alpha\boldsymbol{E}_1+\beta\boldsymbol{E}_2, \alpha\boldsymbol{H}_1+\beta\boldsymbol{H}_2\}$ will still be a solution of those equations (principle of superposition). Time invariance means that if we consider a set of excitations (sources) $\{\boldsymbol{J}(t), \rho(t)\}$ and corresponding solutions (responses) $\{\boldsymbol{E}(t), \boldsymbol{H}(t)\}$, and shift the excitations (sources) by an arbitrary amount of time τ, that is, $\{\boldsymbol{J}(t-\tau), \rho(t-\tau)\}$, then the new corresponding solutions are the original solutions shifted by the same amount of time, that is, $\{\boldsymbol{E}(t-\tau), \boldsymbol{H}(t-\tau)\}$.

Differential equations describing LTI systems in the time domain can be easily translated to the frequency domain by replacing the temporal derivative with the algebraic operator $-i\omega$. Time integrals can be similarly replaced by algebraic operators. As a result, the analytical treatment of LTI differential equations (such as Maxwell's equations) or LTI integral equations is significantly simplified when described in the frequency domain. This constitutes a major reason for the popularity for analytical treatments in the frequency domain.

For a discrete-time signal $\phi(n)$, $\infty < n < \infty$, where n is an integer, a Fourier representation is also possible, in the form (2)

$$\phi(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \tilde{\phi}(e^{j\omega})e^{j\omega n}\,d\omega \tag{3a}$$

where we have used the $e^{j\omega t}$ ($e^{j\omega n}$ for the discrete time-domain case) convention mentioned previously, instead. In the above, $\tilde{\Phi}$) is given by

$$\tilde{\phi}(e^{j\omega}) = \sum_{n=-\infty}^{\infty} \phi(n)e^{-j\omega n} \tag{3b}$$

Equations (3a) constitute the *discrete Fourier transform* (*DFT*) pair. In this case, the Fourier transform $\tilde{\Phi}(e^{j\omega})$ is a periodic function with period 2π. The DFT plays an important role in the analysis and design of discrete-time signal-processing algorithms and systems, analogous to the role of the Fourier transform for continuous-time systems. Also important is the fact that very efficient algorithms, collectively known as *fast Fourier transforms* (*FFT*s), exist to compute the DFT (2). It is interesting to observe the correspondence with the continuous-time case: the summation in Eq. (5b) is simply the Fourier *series* of a periodic signal $\tilde{\Phi}(e^{j\omega})$, whereas the integral for the "coefficients" $\phi(n)$ in Eq. (5a) is just the integral that would be used to obtain the coefficients of the Fourier series. Of course, the discrete-time treatment can nevertheless be carried out independently of this correspondence. Another connection between the continuous-time representation and the discrete-time representation is given by the celebrated *Nyquist sampling theorem*, which states that a band-limited signal $\phi(t)$ whose Fourier transform $\tilde{\Phi}(\omega) = 0$ for $\omega > \omega_{\max}$ is uniquely determined by its samples (i.e., a discrete-time representation), $\phi(n) = \phi(n\ \Delta t)$, $-\infty < n < \infty$, if $\omega_s = 2\ \pi/\ \Delta t > 2\omega_{\max}$. The frequency $\omega_{\max}$ is usually referred as the *Nyquist frequency*, and the frequency ω_s as the Nyquist rate. Therefore, a band-limited signal can be reconstructed exactly from discrete samples. It is possible to generalize both the Fourier transform and the DFT to fully exploit the theory of complex variables in the analysis of the time-domain signals. In the case of continuous-time systems, this generalization is given by the *Laplace* transform (1), and for discrete-time systems it is given by the z transform (2).

Apart from its advantages in analytical treatment, the popularity of frequency-domain analysis also stems from the fact that many measurement apparatuses are confined to frequency-domain measurements. With the advent of high-speed digital computers, however, the strong prevalence of frequency-domain analysis in comparison with time-domain analysis has declined. Digital computers altered the way problems can be solved and paved the way to new algorithms and increasing popularity of time-domain analysis (6,7,8,9,10). Moreover, they also promoted dramatic changes in the measurement hardware. Some of the specific reasons for the increasing popularity of time-domain analysis in the digital age are:

(1) Many problems of interest are actually nonlinear or time-variant. In those cases, time-domain analysis provides a more direct and straightforward modeling.
(2) In some cases, fewer arithmetic operations are required in the time domain than in the frequency domain. For instance, for broadband problems, time-domain analysis is more efficient because the problem is localized in the time domain (e.g., short duration times) but not in the frequency domain (large bandwidths).
(3) It is often easier to get frequency-domain results from time-domain data than vice versa.
(4) Time-domain analysis is conceptually closer to our intuition, which develops in the space–time arena.

In this article, we will describe the basics of time-domain analysis as applied to Maxwell's equations, which govern the behavior of electromagnetic waves. Because of their pervasiveness in electrical engineering, Maxwell's equations will provide a general ground for the discussion of time-domain algorithms. A more detailed discussion of time-domain algorithms in relation to electric circuit and network analysis is presented in the article Time-Domain Network Analysis

Electromagnetic Wave Propagation

The equations governing the electromagnetic field are given by Eq. (2a). If we take a source-free case, $\boldsymbol{J} = \boldsymbol{0}$ and $\rho = 0$, substitute Eq. (2a) into Eq. (2b), and solve for $\boldsymbol{E}$, we obtain

$$\nabla \times \nabla \times \boldsymbol{E} - \mu\epsilon \frac{\partial^2}{\partial t^2} \boldsymbol{E} = \boldsymbol{0} \tag{4}$$

By making use of the identity

$$\nabla \times \nabla \times \boldsymbol{E} = \nabla^2 \boldsymbol{E} - \nabla \nabla \cdot \boldsymbol{E} \tag{5}$$

and of the divergence-free condition for the electric field in a source-free region,

$$\nabla \cdot \boldsymbol{E} = 0 \tag{6}$$

we obtain

$$\nabla^2 \boldsymbol{E} - \mu\epsilon \frac{\partial^2}{\partial t^2} \boldsymbol{E} = \boldsymbol{0} \tag{7}$$

which is know as the *vector wave equation*. The speed of propagation is given by $v = (\mu\epsilon)^{-1/2}$ and depends on the background medium. In vacuum, $v = c = (\mu_0\epsilon_0)^{-1/2} \approx 2.998 \times 10^8$ m/s.

In Cartesian coordinates, the Laplacian operator, ∇^2, is written as

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{8}$$

Furthermore, we can write $\boldsymbol{E}$ in terms of its field components, i.e., $\boldsymbol{E} = \hat{\boldsymbol{x}} E_x + \hat{\boldsymbol{y}} E_y + \hat{\boldsymbol{z}} E_z$, where $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$ are the Cartesian unit vectors. The vector wave equation can therefore be written as three *scalar wave equations*:

$$\nabla^2 E_x - \frac{1}{v^2} \frac{\partial^2}{\partial t^2} E_x = 0 \tag{9}$$

and the same for E_y and E_z. Although the field components appear decoupled in these three equations, they are nevertheless coupled through the divergence-free condition, $\nabla \cdot \boldsymbol{E} = 0$.

The wave-propagating nature of the solutions of Eq. (9) can be easily understood by considering, for simplicity, the one-dimensional case

$$\frac{\partial^2}{\partial x^2} E_x - \frac{1}{v^2} \frac{\partial^2}{\partial t^2} E_x = 0 \tag{10}$$

and considering the *trial* solutions $E_x(x,t) = F^+(x - vt) + F^-(x + vt)$, where $F^\pm$ are twice differentiable functions. It is a simple exercise to verify that this functions satisfy Eq. (10), regardless of the specific functional choice for $F^\pm$. The functions $F^\pm$ are known as propagating wave functions (also called D'Alembert solutions) because they represent a traveling function propagating with speed v in the $+x$ (F^+) and $-x$ (F^-) directions, respectively. The above solutions establishes the propagating-wave nature of the solutions of Maxwell's equations.

The wave equation is a linear second-order partial differential equation: only derivatives up to second order are present. Linear second-order partial differential equations can be classified into three classes, according to the coefficients that multiply the highest-order terms (11). The wave equation is the prototypical *hyperbolic* differential equation. The other classes are the *elliptic* and the *parabolic*. Elliptic equations are associated with steady-state phenomena (boundary-value problems), do not involve time evolution, and therefore will not be

considered here. The prototypical elliptic equation is the Laplace equation, i.e., $\nabla^2\phi = 0$. Parabolic equations are associated with diffusion phenomena, such as heat diffusion (heat equation). They can also occur in the context of electromagnetics, as for instance, in a medium with high conductivity. In that case, Eq. (16b) becomes

$$\frac{\partial^2}{\partial x^2}E_x - \frac{1}{v^2}\frac{\partial^2}{\partial t^2}E_x - \mu\sigma\frac{\partial}{\partial t}E_x = 0 \tag{11a}$$

If the conductivity is sufficiently large so that displacement currents can be neglected, the second term above can be dropped and we have

$$\mu\sigma\frac{\partial}{\partial t}E_x = \frac{\partial^2}{\partial x^2}E_x \tag{11b}$$

which is the diffusion equation, the same parabolic equation that governs heat flow.

Both parabolic and hyperbolic equations are evolutionary equations, which undergo change as a function of time and to which time-domain analysis applies. Parabolic and hyperbolic equations are usually solved through similar techniques, as opposed to elliptic equations.

Note that the passage from the time domain to the frequency domain corresponds to replacing the wave equation (4) by the so-called *Helmholtz equation,*

$$\nabla^2 E_x + \omega^2\mu\epsilon E_x = 0 \tag{12}$$

which is an elliptic equation with no time evolution present.

Time-Domain Differential-Equation Modeling

The time-domain analysis of differential equations using numerical methods on a computer involves a *discretization* procedure, whereby the infinite number of degrees of freedom of the original model is reduced to a finite number, tractable by the computer. For instance, the passage from a continuous-time representation to a discrete-time representation (sampling) mentioned before is an example of a discretization procedure. The spatial discretization of partial differential equations requires a meshing of the space, whereby the spatial domain is replaced by a lattice of discrete points or a set of elementary discrete domains (facets, volumes, etc.). The discretization constitutes the central step in the modeling of differential equations.

When referring to the spatial discretization, three major families of techniques can be identified for differential-equation time-domain modeling. The first are the *finite-difference* techniques, which consists in replacing the derivatives (both temporal and spatial) present in the differential equations with finite-difference approximations. Finite-difference methods are simple to implement and well suited for modeling time-domain problems in moderately complex domains. The second family are the *finite-volume* techniques, where (spatial) local integral relations are derived for the field quantities and discretized through the use of elementary contours, surfaces, and volumes. The third family are the *finite-element* techniques, usually associated with variational formulations and where the fields are projected on the space of some compactly supported basis (interpolatory) functions, usually defined over finite domains (so-called *elements*). Finite elements are better suited for problems involving very complex geometries. They can also provide very accurate error estimates.

There are strong conceptual links between finite-difference, finite-volume, and finite-element techniques (12,13), and in some instances the distinctions between them can become quite blurred. For example, finite-

difference techniques can be recast as particular (point-matched) finite-element techniques with some specific choices for the basis functions. Moreover, the finite-element method can be used as a systematic way to produce more complex finite-difference methods with sharp error estimates. The finite-volume technique can often be also reinterpreted as a finite-difference technique over irregular grids.

Besides these three families of techniques, other families can also be identified. (Pseudo) *spectral methods* (14,15), for example, can be identified as a class of finite methods with global, smooth (for example, Chebyshev or sinusoidal) basis functions, that is, functions defined over the entire spatial domain, instead of compactly supported as in the finite-element method. Pseudospectral methods can be seen as limiting cases of increasing-order finite-difference methods. Another popular time-domain method for Maxwell's equation is the *transmission-line method* (TLM) (16), closely related to the finite-difference method.

When referring to the temporal discretization of hyperbolic equations, the most relevant distinction is between *explicit* and *implicit* techniques. In explicit techniques, the values of the fields at a given instant of time depend only on the previous instant of time. Therefore, they can be written *explicitly* as a function of previously known values in a time-update scheme. In implicit methods, the field values at some instant of time depend not only on field values at previous instants, but also on field values at the same instant of time (and distinct points of space). Therefore, they cannot be written explicitly as a function of previously known field values. Implicit methods require the solution of a linear system at each time step, as opposed to explicit methods. Although for differential-equation modeling the associated matrix is usually sparse, this can be a considerable computational burden. On the other hand, implicit methods are superior in their numerical stability, allowing for larger time discretization steps. This will be discussed in more detail later on.

Finite-Difference Time-Domain Methods. The finite-difference time-domain (FDTD) modeling (8,9, 16) of differential equations replaces derivatives by finite differences. It starts by assuming some field quantity, say $\psi(x,t)$, to exist only at discrete points of space and time separated by fixed intervals Δx and Δt and labeled by the indices m, n that is, $\psi^{m,n} = \psi(m\ \Delta x, n\ \Delta t)$. In general, the intervals do not need to be of equal length, but here they will be assumed so for the sake of simplicity. If $\psi(x,t)$ satisfies a one-dimensional scalar wave equation of the form

$$\frac{1}{v^2}\frac{\partial^2 \psi}{\partial t^2} = \frac{\partial^2 \psi}{\partial x^2} \tag{13}$$

then the spatial derivative can be approximated as

$$\left.\frac{\partial \psi}{\partial x}\right|_{x=m\,\Delta x} \approx \frac{1}{\Delta x}\left(\psi^{m+1,n} - \psi^{m,n}\right) \tag{14a}$$

which corresponds to forward differencing. Alternatively, the spatial derivative can also be approximated as

$$\left.\frac{\partial \psi}{\partial x}\right|_{x=m\,\Delta x} \approx \frac{1}{\Delta x}\left(\psi^{m,n} - \psi^{m-1,n}\right) \tag{14b}$$

or

$$\left.\frac{\partial \psi}{\partial x}\right|_{x=m\,\Delta x} \approx \frac{1}{2\,\Delta x}\left(\psi^{m+1,n} - \psi^{m-1,n}\right) \tag{14c}$$

which correspond to backward and central differencing, respectively. For a second derivative, central differencing can be applied twice,

$$\left.\frac{\partial^2\psi}{\partial x^2}\right|_{x=m\,\Delta x} \approx \frac{1}{\Delta x^2}\left(\psi^{m+1,n} - 2\psi^{m,n} + \psi^{m-1,n}\right) \tag{15}$$

A more systematic way to derive the finite-difference approximation is to expand the functions using a Taylor expansion for $\psi^{m+1,n}$ and $\psi^{m-1,n}$ around $x = m\ \Delta x$, that is, in terms of $\psi^{m,n}$ and its derivatives:

$$\psi^{m+1,n} = \psi^{m,n} + \Delta x\ \frac{\partial\psi}{\partial x}\bigg|^{m,n} + \frac{1}{2}\,\Delta x^2\ \frac{\partial^2\psi}{\partial x^2}\bigg|^{m,n} + \frac{1}{6}\,\Delta x^3\ \frac{\partial^3\psi}{\partial x^3}\bigg|^{m,n} + O(\Delta x^4) \tag{16a}$$

$$\psi^{m-1,n} = \psi^{m,n} - \Delta x\ \frac{\partial\psi}{\partial x}\bigg|^{m,n} + \frac{1}{2}\,\Delta x^2\ \frac{\partial^2\psi}{\partial x^2}\bigg|^{m,n} - \frac{1}{6}\,\Delta x^3\ \frac{\partial^3\psi}{\partial x^3}\bigg|^{m,n} + O(\Delta x^4) \tag{16b}$$

By adding the above equations, we obtain the expression for the second derivative,

$$\left.\frac{\partial^2\psi}{\partial x^2}\right|_{x=m\,\Delta x} = \frac{1}{\Delta x^2}\left(\psi^{m+1,n} - 2\psi^{m,n} + \psi^{m-1,n}\right) + O(\Delta x^2) \tag{17}$$

which shows that the approximation in Eq. (15) is an approximation of second order, that is, with error $O(\Delta x^2)$. On the other hand, backward differencing and forward differencing both have error $O(\Delta x)$.

In a hyperbolic equation, the solution region in the (x,t) domain is open in the coordinate t, so that a solution advances towards positive t from prescribed *initial conditions* at some specified initial time $t = t_0$. On the other hand, in the spatial domain, at any given instant of time, the solution is known for all x and should satisfy prescribed *boundary conditions*. This is illustrated in Fig. 1. Because we are dealing with an initial-value problem in the temporal domain and with a boundary-value problem in the spatial domain, there is a fundamental difference between the time discretization and the spatial discretization.

If we apply the discretization given by Eq. (15) to the time derivative, and substitute in the wave equation in Eq. (13), we arrive at

$$\psi^{m,n+1} = 2\psi^{m,n} - \psi^{m,n-1} + v^2\,\Delta t^2\ \left.\frac{\partial^2\psi}{\partial x^2}\right|_{t=n\,\Delta t} \tag{18}$$

which is a time-stepping formula, that is, given the knowledge of $\psi^{m,n-1}$ and $\psi^{m,n}$, one can find $\psi^{m,n+1}$. In particular, given the values of $\psi^{m,0}$ and $\psi^{m,1}$, the values of $\psi^{m,n}$ at all subsequent time steps can be determined, as long as the time-stepping scheme is stable. The stability of the time-stepping scheme is related to the relative values of Δx, Δt, and v and will be discussed later on. By combining the time and spatial finite differences, the resulting discrete equation becomes

$$\psi^{m,n+1} = 2\psi^{m,n} - \psi^{m,n-1} + v^2\frac{\Delta t^2}{\Delta x^2}\left(\psi^{m+1,n} - 2\psi^{m,n} + \psi^{m-1,n}\right) \tag{19}$$

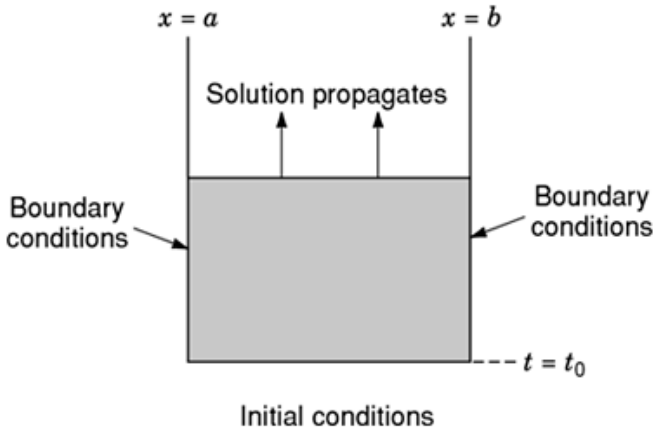

Fig. 1. Time evolution of a hyperbolic partial differential equation in the (x, t) domain, with boundary conditions on x and initial conditions on t.

which can be solved on a computer for some prescribed initial and boundary values.

In the case of Maxwell's equations, we can solve for one of the fields first and discretize the resulting equation, for example the second-order wave equation for $\boldsymbol{E}$, Eq. (12). Otherwise, we can work directly with the first-order curl equations, Eqs. (2a) and (2b), and both fields, $\boldsymbol{E}$ and $\boldsymbol{H}$, simultaneously. Yee's scheme (8,9) is a very popular FDTD scheme to discretize the first-order Maxwell's curl equations in Cartesian coordinates. In terms of components, Eq. (2a) is written as

$$-\mu \frac{\partial H_x}{\partial t} = \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} \tag{20a}$$

$$-\mu \frac{\partial H_y}{\partial t} = \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} \tag{20b}$$

$$-\mu \frac{\partial H_z}{\partial t} = \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} \tag{20c}$$

Yee's FDTD discretization scheme for the above equations starts by spatially staggering the electric and magnetic field components and replacing the spatial derivatives by central differences. The staggering of electric and magnetic fields and the location of each field component are illustrated in Fig. 2. This is equivalent to defining the electric and magnetic field components over different (dual) grids, staggered with respect to each other. A temporal staggering is also used for the electric and magnetic fields in the time evolution. The temporal staggering is such the magnetic field at a time $t = (n + \frac{1}{2})\,\Delta t$ is obtained from the electric field at an instant of time $t = n\,\Delta t$. The electric field at $t = (n + 1)\,\Delta t$ is then obtained from the magnetic field at $t = (n + \frac{1}{2})\,\Delta t$, and the scheme is iterated. Such a time update scheme is usually known as a *leapfrog* scheme. By denoting the field components using $\psi(i\,\Delta x, j\,\Delta y, k\,\Delta z, n\,\Delta t) = \psi^n{}_{i,j,k}$, then the FDTD discretization for Eqs.

(20a) become

$$-\mu\frac{1}{\Delta t}\left(H^{n+\frac{1}{2}}_{x,i,j+\frac{1}{2},k+\frac{1}{2}} - H^{n-\frac{1}{2}}_{x,i,j+\frac{1}{2},k+\frac{1}{2}}\right) = \frac{1}{\Delta y}\left(E^{n}_{z,i,j+1,k+\frac{1}{2}} - E^{n}_{z,i,j,k+\frac{1}{2}}\right) - \frac{1}{\Delta z}\left(E^{n}_{y,i,j+\frac{1}{2},k+1} - E^{n}_{y,i,j+\frac{1}{2},k}\right) \tag{21a}$$

$$-\mu\frac{1}{\Delta t}\left(H^{n+\frac{1}{2}}_{y,i+\frac{1}{2},j,k+\frac{1}{2}} - H^{n-\frac{1}{2}}_{y,i+\frac{1}{2},j,k+\frac{1}{2}}\right) = \frac{1}{\Delta z}\left(E^{n}_{x,i+\frac{1}{2},j,k+1} - E^{n}_{x,i+\frac{1}{2},j,k}\right) - \frac{1}{\Delta x}\left(E^{n}_{z,i+1,j,k+\frac{1}{2}} - E^{n}_{z,i,j,k+\frac{1}{2}}\right) \tag{21b}$$

$$-\mu\frac{1}{\Delta t}\left(H^{n+\frac{1}{2}}_{z,i+\frac{1}{2},j+\frac{1}{2},k} - H^{n-\frac{1}{2}}_{z,i+\frac{1}{2},j+\frac{1}{2},k}\right) = \frac{1}{\Delta x}\left(E^{n}_{y,i+1,j+\frac{1}{2},k} - E^{n}_{y,i,j+\frac{1}{2},k}\right) - \frac{1}{\Delta y}\left(E^{n}_{x,i+\frac{1}{2},j+1,k} - E^{n}_{x,i+\frac{1}{2},j,k}\right) \tag{21c}$$

An analogous discretization applies for the other curl equation (Ampere's law) by duality (4). Once the $\boldsymbol{E}$ field at $t = 0$ and the $\boldsymbol{H}$ field at $t = -\frac{1}{2}\,\Delta t$ are known (initial conditions), the update equations can be used to find the fields at all future time steps.

Yee's FDTD scheme is an explicit scheme, and for a given choice of Δx, Δy, and Δz, a basic condition applies to the maximum admissible value of Δt to obtain numerically stable time stepping. This is given by the Courant-Friedrichs-Levy (CFL) criterion (4,8,16),

$$v\,\Delta t \leq \left(\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2} + \frac{1}{\Delta z^2}\right)^{-1/2} \tag{22}$$

For an inhomogeneous medium, v is a function of position, and in the above, v should be taken as its maximum value in the computational domain. The CFL criteria can be derived through a von Neumann stability analysis (8). We note that this criterion is also related to the causality condition. It essentially states that the time step Δt should be smaller than the shortest travel time for waves between the lattice planes, a requirement of causality. However, implicit schemes, which include the same time-step interactions at different spatial grid points, do allow for the relaxation of the CFL condition.

In the case of a parabolic equation such as Eq. (11b), if forward differencing is used for the first-order temporal derivative and central differencing for the second-order spatial derivative (Euler scheme), the resulting scheme will be stable if $2\,\Delta t \leq (\mu\sigma\,\Delta s^2/\,n)$, where n is the dimension of the problem, and Δs is the

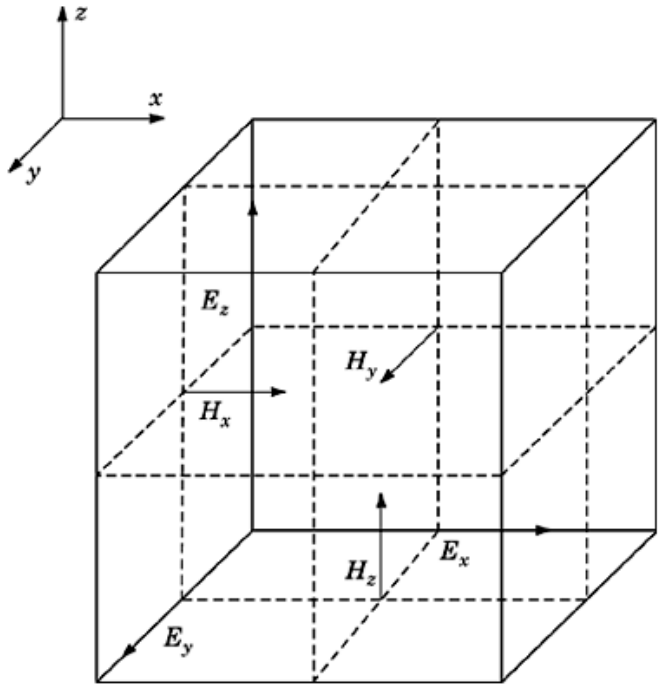

Fig. 2. Yee's elementary staggered FDTD cell, depicting the location of the electric and magnetic field components.

spatial discretization size (assumed uniform) (4). (If central differencing is used for both temporal and spatial derivatives, the update is always unstable.) The right-hand side of this condition is just the time for the field to diffuse between successive planes in the numerical grid. This condition therefore just states that the time step should be smaller than half this diffusion time. In contrast to the discretization of the wave equation, we observe that the computation time for the diffusion equation modeled using FDTD grows as $N^{(n+2)/n}$, where N is the total number of grid points. The computational effort to solve a diffusion problem grows faster with the size of the problem than that in solving the wave equation (the maximum temporal step depends quadratically on the spatial step). This can be physically understood in connection with the difference in how the wavelength (wave equation) and the skin depth (diffusion equation) scale with frequency. One possible way to increase the efficiency of the solution is to use an implicit scheme instead of an explicit one. Another is to add a small wavelike term to the diffusion equation, making it possible to employ a central-differencing scheme for the temporal derivative in which the CFL criterion is always satisfied (unconditionally stable). This must be done without altering much of the physics of the problem (4).

Apart from numerical stability, a second general criterion to be observed in the choice of the discretization size is related to the numerical (or grid) dispersion effects. Numerical dispersion refers to the fact that plane waves do not all propagate with the same phase velocity on the lattice (4,8). Plane waves with different frequencies will have different phase velocities. Moreover, plane waves with the same frequency but different propagation *directions* will also have different phase velocities (due to the anisotropy of the discretized space), this last effect, of course, being present only in two- or three-dimensional problems. As a result of numerical dispersion, a time-domain pulse, which is a linear superposition of plane waves, will be distorted as it propagates through the lattice. To minimize the numerical dispersion error, the spatial discretization size should be chosen small compared to the wavelength (usually between $\lambda/10$ and $\lambda/20$). Alternatively, higher-order finite-difference methods can be employed to reduce the dispersion error (8). Higher-order schemes utilize larger stencils in the

finite-difference approximation for the derivatives. Both alternatives lead to larger computational requirements in memory and CPU time.

In many cases, it is not the numerical dispersion effects but the fine geometrical features of the problem that dictate the maximum values of the spatial discretization steps. In those cases, the use of nonuniform grids, where the spatial discretization size is *locally* reduced to accommodate the fine geometrical details, is often advantageous.

To speed up the time-domain simulation, it is convenient to chose Δt to be as large as possible yet satisfying Eq. (22). As a bonus, it can be shown that, for Yee's scheme, this choice minimizes the numerical dispersion error (8).

Finite-Element Time-Domain Methods. Another popular discretization scheme is the finite-element method (*FEM*) (10,17,18). Although more commonly used for frequency-domain problems (elliptical equations), it has nevertheless been applied with success for time-domain analysis (10). Its major attractiveness is that it is well suited for use with unstructured spatial grids. Finite-element methods can often also be seen as a convenient way to generate complex finite-difference schemes and obtain accurate error estimates. Finite-element schemes are based on general analytical expansion techniques to obtain approximate solutions. They are also usually known as the *Rayleigh–Ritz* method for stationary problems and the method of *weighted residuals* for problems that are posed directly as differential equations (18). Contrary to finite-difference methods, which are based on a finite-point-set approximation to a differential equations, finite-element methods most often utilize piecewise continuous polynomials to expand the unknown functions and generate the discretized version of the equations (in the method of weighted residuals). Hence, instead of discretizing the spatial domain directly, finite-element methods first discretize the function used to represent the solution.

The *point-matched*, or *collocation*, time-domain finite-element method is chosen in the short exposition that follows because of its simplicity and because it already incorporates some of the main features of the time-domain finite-element method. In this scheme, the spatial domain is first subdivided into *finite elements*. In the two-dimensional case, the elements may correspond, for example, to triangles as depicted in Fig. 3. We assume the $\boldsymbol{E}$ and $\boldsymbol{H}$ fields to have the following functional forms:

$$\boldsymbol{E}(\boldsymbol{r}, t) \approx \sum_{i=1}^{M} \phi_i(\boldsymbol{r}) \boldsymbol{E}_i(t) \tag{23a}$$

$$\boldsymbol{H}(\boldsymbol{r}, t) \approx \sum_{j=1}^{N} \psi_j(\boldsymbol{r}) \boldsymbol{H}_j(t) \tag{23b}$$

The functions $\phi_i(\boldsymbol{r})$ and $\psi_j(\boldsymbol{r})$ are called *basis functions* and interpolate the fields within each element using the values of the nodes constituting the element (nodal elements). Alternatively, the degrees of freedom can be associated with edges instead of nodes. In that case, the interpolatory elements are called edge elements. Although nodal elements are simpler to implement, they may produce spurious modes in a finite-element implementation. Proper care should be taken to avoid that when using nodal elements to interpolate $\boldsymbol{H}$ or $\boldsymbol{E}$ (17,18). Edge elements avoid spurious modes because they better mimic the physics of the problem. Using the language of differential forms, $\boldsymbol{H}$ and $\boldsymbol{E}$ are one-forms, while $\boldsymbol{B}$ and $\boldsymbol{D}$ are two-forms. The natural interpolants for one-forms are edge (Whitney) elements. Nodal elements are the natural interpolants for zero-forms (e.g., scalar potentials) only.

The expansion in Eq. (23a) is required to be complete, that is, it can represent any function up to the order of approximation, but the basis functions are not required to be orthogonal. By substituting the above

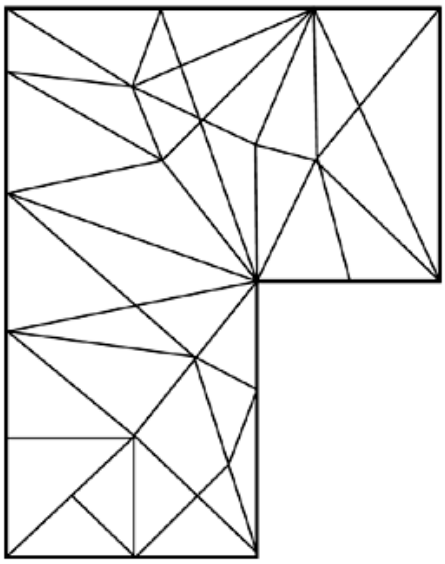

Fig. 3. Typical finite-element meshing of an L-shaped domain using triangular elements.

expressions into Maxwell's equations, we obtain

$$\sum_{j=1}^{N} \psi_j(\boldsymbol{r}) \frac{d\boldsymbol{H}_j(t)}{dt} = -\frac{1}{\mu} \sum_{i=1}^{M} \nabla\phi_i(\boldsymbol{r}) \times \boldsymbol{E}_i(t) \tag{24a}$$

$$\sum_{i=1}^{M} \phi_i(\boldsymbol{r}) \frac{d\boldsymbol{E}_i(t)}{dt} = \frac{1}{\epsilon} \sum_{j=1}^{N} \nabla\psi_j(\boldsymbol{r}) \times \boldsymbol{H}_j(t) \tag{24b}$$

Because both $\phi_i(\boldsymbol{r})$ and $\psi_j(\boldsymbol{r})$ are known functions, the only unknowns in Eqs. (24a) are the time-dependent nodal values of the electric and magnetic fields $\boldsymbol{E}_i(t)$ and $\boldsymbol{H}_j(t)$ at the nodal points $\boldsymbol{r}_i$. By conveniently normalizing the basis functions so that $\phi_i(\boldsymbol{r}_j) = \psi_i(\boldsymbol{r}_j) = \delta_{ij}$, and enforcing Eqs. (24a) at each nodal point (point matching), we obtain

$$\frac{d\boldsymbol{H}_j(t)}{dt} = -\frac{1}{\mu} \sum_{i=1}^{M} \nabla\phi_i(\boldsymbol{r}_j) \times \boldsymbol{E}_i(t) \tag{25a}$$

$$\frac{d\boldsymbol{E}_i(t)}{dt} = \frac{1}{\epsilon} \sum_{j=1}^{N} \nabla\psi_j(\boldsymbol{r}_i) \times \boldsymbol{H}_j(t) \tag{25b}$$

This corresponds to the semidiscrete system of the finite-element method (i.e., discretization in space only). Because both $\psi_j(\boldsymbol{r})$ and $\psi_j(\boldsymbol{r})$ have finite support in space, only a few terms in the summation in Eq. (25a)

will contribute, and, because of the point matching, an explicit scheme for the time update is obtained. The leapfrog scheme can then be used similarly to the FDTD method. In this manner we have

$$\boldsymbol{H}_j^{n+\frac{1}{2}} = \boldsymbol{H}_j^{n-\frac{1}{2}} - \frac{\Delta t}{\mu} \sum_{i=1}^{M} \nabla \phi_i(\boldsymbol{r}_j) \times \boldsymbol{E}_i^n \tag{26a}$$

$$\boldsymbol{E}_i^{n+1} = \boldsymbol{E}_i^n + \frac{\Delta t}{\epsilon} \sum_{j=1}^{N} \nabla \psi_j(\boldsymbol{r}_i) \times \boldsymbol{H}_j^{n+\frac{1}{2}} \tag{26b}$$

Other strategies are possible to obtain the expansion coefficients in Eq. (23a). In general, instead of a point matching, testing functions (spanning the *test* space) are utilized in conjunction with the basis functions (spanning the *trial* space), and the resultant semidiscrete system produces a implicit scheme. These are known as *Galerkin* methods and are perhaps the most popular. Point matching can be shown to be equivalent to a special case of a Galerkin method with Dirac delta test functions (distributions) associated with the electric and magnetic field nodes. Another strategy is the *least-squares* method of weighted residuals, where the square of residuals is integrated over the domain of the problem and the expansion coefficients in Eq. (23a) are obtained by minimizing the resulting integral. Note that in contrast to the FDTD scheme described previously, all three components of the electric field are placed at the same node for Eqs. (26a). The same is true for the magnetic field components.

There are many important issues connected with the finite-element method that are beyond the scope of this article, such as variational formulations, choice of meshing and elements utilized, bases and test functions, nodal versus edge elements, and mass lumping (a procedure used to convert an implicit time-update scheme to an explicit one). For a detailed discussion of those issues, the reader is referred to Refs. 17,18.

Finite-Volume Time-Domain Methods. The term *finite-volume* has a loose meaning in the literature, and it often can refer to different discretization methods. For instance, it sometimes synonymous with three-dimensional FDTD schemes on irregular, unstructured grids. Here we will use the classification employed in Ref. 19. The finite-volume time-domain method starts by subdividing the computational domain into elementary volumes (most often irregular). Upon integration of Maxwell's curl equations in each elementary volume, we obtain

$$-\mu \int_V \frac{\partial \boldsymbol{H}}{\partial t} dv = \int_V \nabla \times \boldsymbol{E} dv = \int_{\partial V} \hat{\boldsymbol{n}} \times \boldsymbol{E} ds \tag{27a}$$

$$\epsilon \int_V \frac{\partial \boldsymbol{E}}{\partial t} dv = \int_V \nabla \times \boldsymbol{H} dv = \int_{\partial V} \hat{\boldsymbol{n}} \times \boldsymbol{H} ds \tag{27b}$$

Many time-domain algorithms can be obtained through different choices of elementary volumes and surfaces. As with the finite-difference and finite-element methods, the discretization scale should resolve the wavelength well to minimize dispersion error. The grid may consist of cubes, distorted cubes, tetrahedrons, prisms, or their combination. By using a leapfrog scheme similar to the previously discussed finite-difference

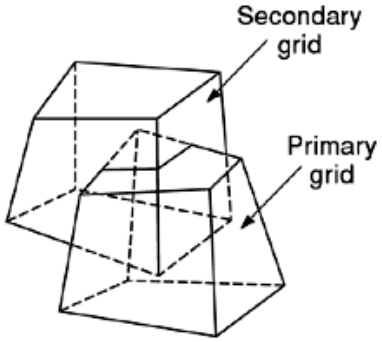

Fig. 4. Illustration of the dual grid construction used in the finite-volume method with cubical elements.

and finite-element discretizations, the time update equations for the above become

$$-\frac{\mu}{\Delta t}\int_V \left(\boldsymbol{H}^{n+\frac{1}{2}} - \boldsymbol{H}^{n-\frac{1}{2}}\right) dv = \int_{\partial V} \hat{\boldsymbol{n}} \times \boldsymbol{E}\, ds \tag{28a}$$

$$\frac{\epsilon}{\Delta t}\int_V \left(\boldsymbol{E}^{n+1} - \boldsymbol{E}^{n}\right) dv = \int_{\partial V} \hat{\boldsymbol{n}} \times \boldsymbol{H}\, ds \tag{28b}$$

Similarly to Yee's FDTD case, a dual grid is often utilized as illustrated in Fig. 4, where the electric field components are located on the primary grid and the magnetic field components are located on the secondary grid. At each time step a double interpolation should be used to interpolate field values from one grid to another. There is a close association of the finite-volume time-domain method with the FDTD method, in that it can be shown that the FDTD can be derived through application of the integral relations

$$-\mu \int_A \frac{\partial \boldsymbol{H}}{\partial t} \cdot \hat{\boldsymbol{n}} dS = \int_{\partial A} \boldsymbol{E} \cdot d\boldsymbol{l} \tag{29a}$$

$$\epsilon \int_A \frac{\partial \boldsymbol{E}}{\partial t} \cdot \hat{\boldsymbol{n}} dS = \int_{\partial A} \boldsymbol{H} \cdot d\boldsymbol{l} \tag{29b}$$

Therefore, while the finite-volume time-domain method is derived by using the curl of the electric field over a surface to find the magnetic vector at the center of the enclosed volume as in Eq. (27a) [and vice versa on the dual grid, Eq. (27b)], the FDTD can be derived through using the circulation of the electric field around a contour to update the magnetic field over the enclosed area as in Eq. (29a) [and vice versa on the dual-grid, Eq. (29b)]. For an additional discussion of the finite-volume time-domain method, the reader may consult Ref. 19.

Time Integration Schemes

In the previous sections, we have illustrated the main features of different discretization methods using leapfrog schemes. The leapfrog scheme is convenient for a system of two first-order equations such as Maxwell's curl equations, because it achieves overall second-order accuracy in time with only first-order time differencing. However, many other time-stepping schemes are possible. In this section, we will illustrate some of those.

Single-Step Methods. Starting from a problem already discretized in the spatial domain, the resulting semidiscrete problem can be written as

$$\frac{d\boldsymbol{v}(t)}{dt} = \boldsymbol{F}(\boldsymbol{v}, t) \tag{30}$$

where $\boldsymbol{v}$ and $\boldsymbol{F}$ are vectors, and $\boldsymbol{v}$ is subject to some initial condition, $\boldsymbol{v}(t_0) = \boldsymbol{v}_0$. In general $\boldsymbol{F}$ can be a nonlinear function. The algorithms for solving first-order equations such as Eq. (30) immediately generalize to systems of first-order equations, and because a higher-order equation can be cast as a system of first-order equations, the methods discussed here apply to differential equations of arbitrary order. In the case of Maxwell's equations, $\boldsymbol{F}$ is linear and Eq. (30) can be decomposed as

$$\frac{d\boldsymbol{v}_1(t)}{dt} = \boldsymbol{F}_1(\boldsymbol{v}_2) \tag{31a}$$

$$\frac{d\boldsymbol{v}_2(t)}{dt} = \boldsymbol{F}_2(\boldsymbol{v}_1) \tag{31b}$$

In this case, $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ represent the field components at all grid points, while $\boldsymbol{F}_1(\boldsymbol{v}_2) = \boldsymbol{F}_1 \cdot \boldsymbol{v}_2$ and $\boldsymbol{F}_2(\boldsymbol{v}_1) = \boldsymbol{F}_2 \cdot \boldsymbol{v}_1$ are (sparse) matrices representing the curl operators on the grid.

A time-step update scheme for the prototypical Eq. (30) can be written as

$$\boldsymbol{v}^{n+1} = \boldsymbol{v}^n + \Delta t \left[\theta \boldsymbol{F}(\boldsymbol{v}^n, t^n) + (1-\theta)\boldsymbol{F}(\boldsymbol{v}^{n+1}, t^{n+1})\right] \tag{32}$$

with $t^n = n\ \Delta t$, and $0 \le \theta \le 1$. This is called a *theta method*. If $\theta = 0$, we arrive at *Euler's method* (explicit), and if $\theta = \frac{1}{2}$, we arrive at the *trapezoidal rule.*

We can interpret Eq. (32) geometrically, by assuming the slope of the solution to be piecewise constant and given by a linear combination, weighted through the variable θ, of the derivatives at the endpoint of each discretization interval. Alternatively, we can also use a Taylor expansion as in Eqs. (16a) to arrive at the above approximation. By doing so, it can be shown that the theta approximation is of first order, except for $\theta = 1/2$, when it is of second order (trapezoidal rule). Moreover, if $\theta = 0$ the method is explicit, and otherwise implicit.

Multistep Methods. The theta method and the leapfrog scheme described before are examples of *step-by-step* methods, that is, once the solution for $\boldsymbol{v}^{n+1}$ is obtained, the value of $\boldsymbol{v}^n$ is discarded for future updates. However, past values at more than one time step can be used to compute the value at next time step. This gives rise to *multistep* methods, which can be written in general form as

$$\sum_{m=0}^{s} a_m \boldsymbol{v}_{n+m} = \Delta t \sum_{m=0}^{s} b_m \boldsymbol{F}(\boldsymbol{v}^{n+m}, t^{n+m}) \tag{33}$$

If $b_s = 0$, the method is explicit; otherwise, it is implicit. Of course, the efficacy of the method depends on the values of the coefficients a_m and b_m. The method above has $2s + 1$ degrees of freedom (coefficients), and in principle an optimal scheme could be obtained by adjusting the maximum number of coefficients. However, it can be shown that an implicit method of order $2s$ does not converge for $s \geq 3$. It is also possible to show that the maximum order for a convergent s-step method, as in Eq. (33), is equal to s for explicit schemes and to 2 int$[(s + 2)/2]$ for implicit schemes. This result is known as the *Dahlquist first barrier* (11).

Convergence is a most important characteristic of a time-stepping scheme. A scheme is said to be convergent if, for every equation of the form in Eq. (30), $|\boldsymbol{v}^n{}_{\Delta t} - \boldsymbol{u}(t^n)| \to 0$ as $\Delta t \to 0$ for all n, where $\boldsymbol{u}$ denotes the actual solution (i.e., exact) in the case of an ordinary-differential-equation problem or a semidiscrete solution in the case of a partial-differential-equation problem, and $\boldsymbol{v}^n{}_{\Delta t}$ denotes the numerical solution obtained by using a time step of size Δt. The numerical solution of a convergent scheme tends to the analytical or semidiscrete solution as the discretization step size approaches zero. Note that, because the semidiscrete solution is itself an approximation to the exact solution, a time-stepping convergent scheme may not converge to the exact solution of the problem if the spatial discretization is not consistent (13). In that case, to study the convergence properties of the overall numerical scheme, we must consider the discretization with respect to time and space conjointly.

A popular multistep scheme is the two-step *predictor–corrector* scheme. We again consider Eq. (30). In the first step (predictor), we predict the function value at the middle of the discretization interval:

$$\boldsymbol{v}^{n+\frac{1}{2}} = \boldsymbol{v}^n + \frac{\Delta t}{2}\boldsymbol{F}(\boldsymbol{v}^n, t^n) \tag{34}$$

and this is followed by a corrector step, where the next step value is estimated through

$$\boldsymbol{v}^{n+1} = \boldsymbol{v}^n + \Delta t\, \boldsymbol{F}(\boldsymbol{v}^{n+\frac{1}{2}}, t^{n+\frac{1}{2}}) \tag{35}$$

The predictor–corrector scheme can be related to the popular *Lax–Wendroff* scheme when $\boldsymbol{F}$ is a linear time-invariant function. The Lax–Wendroff scheme uses the following identity from a Taylor expansion:

$$\boldsymbol{v}^{n+1} = \boldsymbol{v}^n + \Delta t \left.\frac{d\boldsymbol{v}}{dt}\right|^n + \frac{\Delta t^2}{2}\left.\frac{d^2\boldsymbol{v}}{dt^2}\right|^n + O(\Delta t^3) \tag{36}$$

to estimate the derivative at t^n as

$$\left.\frac{d\boldsymbol{v}}{dt}\right|^n \approx \frac{1}{\Delta t}\left(\boldsymbol{v}^{n+1} - \boldsymbol{v}^n\right) - \frac{\Delta t}{2}\left.\frac{d^2\boldsymbol{v}}{dt^2}\right|^n \tag{37}$$

The second derivative of $\boldsymbol{v}$ in Eq. (37) can be obtained from Eq. (30):

$$\left.\frac{d^2\boldsymbol{v}}{dt^2}\right|^n = \left.\frac{d\boldsymbol{F}(\boldsymbol{v}, t)}{dt}\right|^n = \left.\frac{d\boldsymbol{F}(\boldsymbol{v}, t)}{d\boldsymbol{v}} \cdot \boldsymbol{F}(\boldsymbol{v}, t)\right|^n + \left.\frac{\partial \boldsymbol{F}(\boldsymbol{v}, t)}{\partial t}\right|^n \tag{38}$$

and, if $\boldsymbol{F}(\boldsymbol{v},t) = \boldsymbol{F} \cdot \boldsymbol{v}$ (linear, time-invariant), Eq. (36) then becomes

$$\boldsymbol{v}^{n+1} = \boldsymbol{v}^n + \Delta t\, \boldsymbol{F} \cdot \boldsymbol{v}^n + \frac{\Delta t^2}{2} \boldsymbol{F}^2 \cdot \boldsymbol{v}^n \tag{39}$$

which is a second-order scheme.

Runge–Kutta (*RK*) methods form another popular class of time-stepping schemes. They are essentially based on (estimated) numerical quadrature rules to approximate the integral

$$\boldsymbol{v}^{n+1} = \boldsymbol{v}^n + \int_{t^n}^{t^{n+1}} \boldsymbol{F}\left(\boldsymbol{v}(t'), t'\right) dt' \tag{40}$$

through a sum like

$$\boldsymbol{v}^{n+1} \approx \boldsymbol{v}^n + \Delta t \sum_{k=1}^{K} c_k \boldsymbol{F}\left(\boldsymbol{v}(t_k), t_k\right) \tag{41}$$

where $t_k \in (t^n, t^{n+1})$, $k = 1, \ldots, K$. Because the values of $\boldsymbol{F}\left(\boldsymbol{v}(t_k), t_k\right)$ are not known *a priori*, they also need to be approximated. The idea is to use an estimate $\boldsymbol{F}\left(\boldsymbol{v}\left(t_l\right), t_l\right) \approx \zeta_l$, given as a linear combination of the previous estimates ζ_k, $k = 0, \ldots, l-1$, namely,

$$\begin{aligned}
\zeta_1 &= \boldsymbol{v}^n \\
\zeta_2 &= \boldsymbol{v}^n + \Delta t\, A_{21} \boldsymbol{F}\left(\zeta_1, t^n\right) \\
&\vdots \\
\zeta_l &= \boldsymbol{v}^n + \Delta t \sum_{k=1}^{l-1} A_{lk} \boldsymbol{F}\left(\zeta_k, t^n + \alpha_k\, \Delta t\right) \\
\boldsymbol{v}^{n+1} &= \boldsymbol{v}^n + \Delta t \sum_{k=1}^{K} c_k \boldsymbol{F}\left(\zeta_k, t^n + \alpha_k\, \Delta t\right)
\end{aligned} \tag{42}$$

The matrix (A_{lk}) is called the RK matrix, c_k are the RK weights, and α_k are the RK nodes. The number K denotes the stage of the RK scheme (and usually, but not always, also corresponds to its order). Popular third-order schemes are the classical RK scheme, with $K = 3$ and $\alpha_1 = 0$, $\alpha_2 = \frac{1}{2}$, $\alpha_3 = 1$, $c_1 = \frac{1}{6}$, $c_2 = \frac{2}{3}$, $c_3 = \frac{1}{6}$, $A_{21} = \frac{1}{2}$, $A_{31} = -1$, $A_{32} = 2$, and $A_{kl} = 0$ for $k \leq l$; and the Nystrom RK scheme, also with $K = 3$, and $\alpha_1 = 0$, $\alpha_2 = \alpha_3 = \frac{2}{3}$, $c_1 = \frac{1}{4}$, $c_2 = c_3 = \frac{3}{8}$, $A_{21} = A_{32} = \frac{2}{3}$, and all other $A_{kl} = 0$.

The best-known fourth-order Runge–Kutta method has $K = 4$, and $\alpha_1 = 0$, $\alpha_2 = \alpha_3 = \frac{1}{2}$, $\alpha_4 = 1$, $c_1 = c_4 = \frac{1}{6}$, $c_2 = c_3 = \frac{1}{3}$, $A_{21} = A_{32} = \frac{1}{2}$, $A_{43} = 1$, and all other $A_{kl=0}$.

One of the reasons for the popularity of the fourth-order RK scheme is that, to obtain a fifth-order scheme, we need at least six stages. All these RK schemes correspond to *explicit* schemes, but *implicit* RK schemes can also be derived (11). Implicit schemes require the solution of a linear system at each time step, but present superior stability properties.

Convolution and Recursive Functions

In (time-)dispersive media, the relationship between the electric field vectors $\boldsymbol{D}$ and $\boldsymbol{E}$ in the time domain (constitutive relation) is given by a convolutional operator of the form

$$\boldsymbol{D}(t) = \int_0^t dt'\, \boldsymbol{E}(t-t')\epsilon(t') = \epsilon_\infty \epsilon_0 \boldsymbol{E}(t) + \epsilon_0 \int_0^t dt'\, \boldsymbol{E}(t-t')\chi(t') \tag{43}$$

Translated to the frequency domain, the above relation reads

$$\boldsymbol{D}(\omega) = \epsilon(\omega)\boldsymbol{E}(\omega) \tag{44}$$

where $\epsilon(\omega)$ is the (frequency-dependent) permittivity, ϵ_0 is the free-space permittivity, ϵ_∞ is the infinite-frequency permittivity (instantaneous response), and $\chi(t)$ is the time-domain susceptibility. In Eq. (43) causality was invoked by having $\epsilon(t) = 0$ for $t < 0$.

A naive implementation of the convolution in Eq. (43) for a time-stepping scheme requires the storage of the whole previous time history of $\boldsymbol{E}(t)$ in order to obtain $\boldsymbol{D}(t)$ at each successive time step. However, this can be avoided through the use of a *recursive convolution* algorithm (20,21), as described below.

Letting $t^n = n\ \Delta t$, we have

$$\boldsymbol{D}(t) \approx \boldsymbol{D}(t^n) = \boldsymbol{D}^n = \epsilon_\infty \epsilon_0 \boldsymbol{E}^n + \epsilon_0 \int_0^{n\,\Delta t} dt'\, \boldsymbol{E}(n\,\Delta t - t')\chi(t') \tag{45}$$

If the time interval Δt is sufficiently small, we may approximate the field quantities as constants over each interval (piecewise-constant approximation) so that the above integration becomes a summation of the form

$$\boldsymbol{D}^n = \epsilon_\infty \epsilon_0 \boldsymbol{E}^n + \epsilon_0 \sum_{m=0}^{n-1} \boldsymbol{E}^{n-m} \int_{m\,\Delta t}^{(m+1)\,\Delta t} dt'\ \chi(t') \tag{46}$$

The objective is to incorporate the summation above in a explicit time-domain update at minimal computational cost. A discretized form of the constitutive relation can be written as

$$\boldsymbol{D}^{n+1} - \boldsymbol{D}^n = \epsilon_\infty \epsilon_0 (\boldsymbol{E}^{n+1} - \boldsymbol{E}^n) + \epsilon_0 (\boldsymbol{P}^{n+1} - \boldsymbol{P}^n) \tag{47}$$

with

$$\boldsymbol{P}^n = \sum_{m=0}^{n-1} \boldsymbol{E}^{n-m} \int_{m\,\Delta t}^{(m+1)\,\Delta t} dt'\, \chi(t') = \sum_{m=0}^{n-1} \boldsymbol{E}^{n-m} \tilde{\chi}^m \tag{48}$$

where we define

$$\tilde{\chi}^m = \int_{m\,\Delta t}^{(m+1)\,\Delta t} dt'\ \chi(t') \tag{49}$$

We assume the susceptibility function $\chi(t)$ to be an exponential function in time. As we will see later, this assumption is not as restrictive as it may seem.

At first, it appears that the summation in Eq. (48) (convolution) requires the storage of all the past values of $\boldsymbol{E}^n$. In order to see how, on assuming an exponential dependence for $\chi(t')$, this storage is actually not necessary, we use Eq. (48) and rewrite the last term in Eq. (47) as

$$\boldsymbol{P}^{n+1} - \boldsymbol{P}^n = \boldsymbol{E}^{n+1}\tilde{\chi}^0 + \sum_{m=0}^{n-1} \boldsymbol{E}^{n-m}\left(\tilde{\chi}^{m+1} - \tilde{\chi}^m\right) \tag{50}$$

We also define

$$\Delta\tilde{\chi}^n = \left(\tilde{\chi}^{n+1} - \tilde{\chi}^n\right) \tag{51}$$

and the recursive accumulator

$$\boldsymbol{Q}^n = \sum_{m=0}^{n-1} \boldsymbol{E}^{n-m}\Delta\tilde{\chi}^n \tag{52}$$

For a susceptibility function of the form

$$\chi(t) = A_0 e^{-t/\tau_0} u(t) \tag{53}$$

for $t > 0$ and zero otherwise, Eq. (52) can then be written recursively as

$$\boldsymbol{Q}^n = \boldsymbol{E}^n\tilde{\chi}^0 + e^{-\Delta t/\tau_0}\boldsymbol{Q}^{n-1} \tag{54}$$

with $\boldsymbol{Q}^0 = \boldsymbol{Q}^1 = 0$. From the above, we observe that only the previous value of the recursive accumulator is needed.

The recursive convolution algorithm can be easily extended to the case where the susceptibility function $\chi(t)$ displays sinusoidal or damped sinusoidal behavior, by equating it to the real part of a complex exponential susceptibility function (22). Moreover, the restriction of χ to a exponential function above is not problematic, because the susceptibility function from a more arbitrary frequency-dependent material $\epsilon(\omega)$ can be modeled as a sum of exponentials, for example, via Prony's method.

Other variants of the recursive convolution algorithm exist. For instance, instead of approximating the electric field as a constant at each time interval in Eq. (46), one can approximate it as a linear function over each interval. This gives rise to the so-called *piecewise-linear recursive convolution* algorithm (23).

Moreover, it is also possible to efficiently incorporate the dispersive behavior of Eq. (43) into update equations through entirely different approaches, such as the z-transform method or the auxiliary differential equation (*ADE*) method. In the ADE approach, for instance, the frequency-dependent dispersion relation is transformed into a ordinary differential equation in the time variable, which relates $\boldsymbol{E}$ and $\boldsymbol{D}$. To illustrate the approach, we take $\epsilon(\omega)$ having the form (Debye dispersion)

$$\epsilon(\omega) = \frac{\alpha}{1 - i\omega\tau} \tag{55}$$

If we substitute this expression into Eq. (44), we arrive at

$$\left(1 - i\omega\tau\right) \boldsymbol{D} = \alpha \boldsymbol{E} \tag{56}$$

In the time domain, the above becomes

$$\left(1 + \tau \frac{\partial}{\partial t}\right) \boldsymbol{D} = \alpha \boldsymbol{E} \tag{57}$$

In a update scheme, the above equation can be easily updated concomitantly with Maxwell's curl equations. A more detailed description of the ADE and z-transform approaches is beyond the scope of this article, however. The interested reader may consult Refs. 24,25,26,27.

Time-Domain Integral-Equation Methods

Time-domain methods can be used not only in conjunction with differential equations, but also in conjunction with integral equations (time-domain integral equations). Models based on time-domain integral equations provide, in general, a more efficient formulation of, for example, surface-scattering phenomena or, in general, of problems where a Green's function is available (6,7,16,28).

A typical time-domain integral equation can be written as

$$g(\boldsymbol{r}, t) = \int_D d\boldsymbol{r}' \int dt' \; f(\boldsymbol{r}', t') K(\boldsymbol{r}, \boldsymbol{r}', t - t') \tag{58}$$

where $g(\boldsymbol{r},t)$ is the excitation function, $K(\boldsymbol{r},\boldsymbol{r}',t)$ is the kernel of the integral equation (often the Green's function of an associated differential equation), $f(\boldsymbol{r},t)$ is the unknown solution, for $\boldsymbol{r} \in D$, and D is some integration volume or surface. To solve Eq. (58) numerically, the unknown is represented in terms of suitably chosen spatiotemporal basis functions:

$$f(\boldsymbol{r}, t) = \sum_{n=1}^{N_s} \sum_{i=0}^{N_t} a_{n,i} b_n(\boldsymbol{r}) T_i(t) \tag{59}$$

where $a_{n,i}$ are the unknown coefficients. Spatial and the temporal basis functions having local support are often employed. Upon substituting Eq. (59) into Eq. (58) and taking the inner product (also called a *testing procedure*) with each of the spatial basis functions $b_m(\boldsymbol{r})$ at discrete times $t = t_j = j\,\Delta t$, the following matrix equation is obtained:

$$\bar{\boldsymbol{Z}}_0 \cdot \bar{\boldsymbol{A}}_j = \bar{\boldsymbol{G}}_j - \sum_{l=1}^{j-1} \bar{\boldsymbol{Z}}_l \cdot \bar{\boldsymbol{A}}_{j-l} \tag{60}$$

where the elements of the vectors $\bar{\boldsymbol{A}}_j$ and $\bar{\boldsymbol{G}}_j$ are given by

$$[\bar{\boldsymbol{A}}_i]_m = a_{m,i} \tag{61}$$

and

$$[\bar{\boldsymbol{G}}_i]_m = \int_D d\boldsymbol{r}\, b_m(\boldsymbol{r}) g(\boldsymbol{r}, t_i) \tag{62}$$

The elements of the (sparse) interaction matrix are given by

$$[\bar{\boldsymbol{Z}}_i]_{mn} = \int_D d\boldsymbol{r}\, b_m(\boldsymbol{r}) \int_D d\boldsymbol{r}'\; b_n(\boldsymbol{r}') \int dt'\, K(\boldsymbol{r}, \boldsymbol{r}', t - t') T_{j-i}(t')|_{t=t_j} \tag{63}$$

Equation (60) relates the expansion coefficients at the jth time step to the excitation and the expansion coefficients at previous time steps. This algorithm is often termed the marching-on-time (*MOT*) scheme (28). Numerical methods based on integral equations are often termed *boundary element methods* because, from the knowledge of the Green's function of the problem, the discretization needs to be performed only on the boundary of the domain. In electromagnetics, for historical reasons, the general procedure of projecting an integral equation into a matrix equation is often termed *method of moments* (29).

Current Issues

Time-domain algorithms continue to be a topic of active research. Because of the need to model large-scale broadband and nonlinear systems and with the ever increasing advance in the computational resources available, time-domain techniques have become an indispensable tool. Among the topics of current research in time-domain analysis are (a) higher-order methods to reduce the inherent dispersion error caused in finite-difference, finite-element, and finite-volume methods (30), (b) the use of irregular grids (structured or unstructured) to better model irregular geometries and to avoid the staircasing approximation of regular grids (31), (c) the use of (pseudo)spectral methods to reduce the number of nodes per wavelength necessary for a given accuracy (15), and (d) multigrid (32), adaptative meshing (33), or subgridding (34) schemes. In case of time-domain integral equations, the use of time-domain fast multipole methods (28) to reduce the computational complexity of the algorithms is also a topic under current investigation. Moreover, asymptotic techniques, such as the geometric theory of diffraction (*GTD*), originally developed in the frequency domain, have also been translated in recent years to the time domain (35). For a more detailed discussion of modern aspects of time-domain analysis, the reader is referred to Refs. 6,7,8 and references therein.

BIBLIOGRAPHY

1. A. V. Oppenheim,A. W. Willsky,I. T. Young, *Signals and Systems*, Englewood Cliffs, NJ: Prentice-Hall, 1983.
2. A. V. OppenheimR. W. Schafer, *Discrete-Time Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1989.
3. E. Butkov, *Mathematical Physics*, New York: Addison-Wesley, 1968.
4. W. C. Chew, *Waves and Fields in Inhomogeneous Media*, Piscataway, NJ: IEEE Press, 1995.
5. J. A. Kong, *Electromagnetic Wave Theory*, Cambridge, MA: EMW Publishing, 1999.
6. E. K. Miller, Time domain modeling in electromagnetics, *J. Electromag. Waves Appl.*, **8** (9): 1125–1172, 1994.
7. S. M. Rao (ed.), *Time Domain Electromagnetics*, Academic Press, 1999.
8. A. Taflove, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, Boston: Artech House, 1995.
9. K. S. Yee, Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, *IEEE Trans. Antennas Propag.*, **14**: 302–307, 1966.

10. A. C. Cangellaris,C.-C. Li,K. K. Mei, "Point-matched time domain finite element methods for electromagnetic radiation and scattering," IEEE Trans. Antennas Propag., **35**: 1160–1173, 1987.
11. A. Iserles, *A First Course in the Numerical Analysis of Differential Equations*, Cambridge: Cambridge University Press, 1996.
12. C. Matiussi, An analysis of finite volume, finite element, and finite difference methods using concepts form algebraic topology, *J. Comput. Phys.*, **33** (2): 289–309, 1997.
13. F. L. TeixeiraW. C. Chew, Lattice electromagnetic theory from a topological viewpoint, *J. Math. Phys.*, **40** (1): 169–187, 1999.
14. B. Fornberg, *A Practical Guide to Pseudospectral Methods*, Cambridge: Cambridge University Press, 1998.
15. Q. H. Liu, Large scale simulations of electromagnetic and acoustic measurements using the pseudospectral time-domain (PSTD), *IEEE Trans. Geosci. Remote Sens.*, **37**: 917–926, 1999.
16. M. Sadiku, *Numerical Techniques in Electromagnetics*, Boca Raton, FL: CRC Press, 1992.
17. J. M. Jin, *The Finite Element Method in Electromagnetics*, New York: Wiley, 1993.
18. P. P. SilvesterR. L. Ferrari, *Finite Elements for Electrical Engineers*, 3rd ed., Cambridge: Cambridge University Press, 1996.
19. K. S. YeeJ. S. Chen, The finite-difference (FDTD) and the finite-volume time-domain (FVTD) methods in solving Maxwell's equations, *IEEE Trans. Antennas Propag.*, **45**: 354–363, 1997.
20. R. J. Luebbers, *et al.* A frequency-dependent finite-difference time-domain formulation for dispersive materials, *IEEE Trans. Electromagn. Compat.*, **32**: 222–227, 1990.
21. R. SiushansianJ. LoVetri, Efficient evaluation of convolution integrals arising in FDTD formulations of electromagnetic dispersive media, *J. Electromagn. Waves Appl.*, **11**: 101–117, 1997.
22. R. J. LuebbersF. Hunsberger, FDTD for *n*-th order dispersive media, *IEEE Trans. Antennas Propag.*, **40**: 1297–1301, 1992.
23. D. F. KelleyR. J. Luebbers, Piecewise linear recursive convolution for dispersive media using FDTD, *IEEE Trans. Antennas Propag.*, **44**: 792–797, 1996.
24. D. M. Sullivan, *Z*-transform theory and the FDTD method, *IEEE Trans. Antennas Propag.*, **44**: 28–34, 1996.
25. D. M. Sullivan, Frequency-dependent FDTD methods using *Z* transforms, *IEEE Trans. Antennas Propag.*, **40**: 1223–1230, 1992.
26. T. KashiwaI. Fukai, A treatment by the FDTD method of the dispersive characteristics associated with electronic polarization, *Microw. Opt. Technol. Lett.*, **3** (6): 203–205, 1990.
27. W. H. WeedonC. M. Rappaport, A general method for FDTD modeling of wave propagation in arbitrary frequency dispersive media, *IEEE Trans. Antennas Propag.*, **45**: 401–410, 1997.
28. A. A. Ergin,B. Shanker,E. Michielssen, "Fast evaluation of three-dimensional transient wave fields using diagonal translation operators," *J. Comput. Phys.*, **146** (1): 157–180, 1998.
29. R. F. Harrington, *Field Computations by Moment Methods*, Piscataway, NJ: IEEE Press, 1993.
30. R. D. Gralia,D. R. Wilton,A. F. Peterson, Higher-order interpolatory vector bases for computational electromagnetics, *IEEE Trans. Antennas Propag.*, **45**: 329–342, 1997.
31. E. A. Navarro, *et al.* Some considerations about the finite difference time domain method in general curvilinear coordinates, *IEEE Microw. Guided Wave Lett.*, **4**: 396–398, 1994.
32. M. J. WhiteM. F. Iskander, Development of a multigrid FDTD code for three-dimensional applications, *IEEE Trans. Antennas Propag.*, **45**: 1512–1517, 1997.
33. I. S. KimW. J. R. Hoefer, A local mesh refinement algorithm for the time domain finite difference method using Maxwell's curl equations, *IEEE Trans. Microw. Theory Tech.*, **38**: 812–815, 1990.
34. P. Monk, Sub-gridding FDTD schemes, *Appl. Comput. Electromagn. Soc. J.*, **11**: 37–46, 1996.
35. T. W. Veruttipong,Time domain version of the uniform GTD, IEEE Trans. Antennas Propag., **38**: 1757–1764, 1990.

F. L. TEIXEIRA
Massachusetts Institute of Technology