Progress In Electromagnetics Research, PIER 28, 231–252, 2000

A DETAILED EXAMINATION OF THE FINITE-VOLUME, TIME-DOMAIN METHOD FOR MAXWELL'S EQUATIONS

J. L. Young and R. O. Nelson

Microelectronics Research Center
University of Idaho
Moscow, ID 83844-1024, USA

D. V. Gaitonde

Air Force Research Laboratory
Wright Patterson Air Force Base
OH 45433-7521, USA

1. INTRODUCTION

The finite-volume, time-domain (FVTD) method has been considered in recent years as a viable tool for the simulation of the electromagnetic field [1–3]. Its popularity among some researchers stems from its overwhelming success in predicting vortices, shocks and the like in complicated high-speed fluid flows [4]. The idea behind the conversion of a computational fluid dynamics (CFD) solver into a computational

electromagnetics (CEM) solver has its roots in the mathematical similarities of the equations of Maxwell and the equations of Euler. Since both systems are hyperbolic, the argument is made that similar discretization procedures for both systems may be invoked as well.

The previous argument is indeed valid if the task is to model a plane wave in an unbounded medium. For this situation, the FVTD scheme can be shown to be both consistent and conditionally stable, which implies from Lax's theorem that the scheme's data will converge to the correct answer for infinitesimal cell sizes and time steps [5]. However, the role of the solver is not to model such simplistic cases. The purpose of the solver is to model the electromagnetic event as created by finite sources and disturbed by physical boundaries and/or media. Herein, then, lies the question of the role of the nodally collocated FVTD solver in computational electromagnetics.

It is not the purpose of this paper to editorialize on the FVTD role in CEM, but to demonstrate certain key features, attributes and deficiencies of the FVTD solver. We do so by first laying a precise mathematical foundation on which the scheme is built. This foundation, which seems to be well-understood in the CFD community for conservation-based equations, has not been adequately articulated to the CEM community in the context of Maxwell's equations. From this foundation, the role of the grid metrics and the primitive vector is clearly explained, both for the interior scheme as well as the boundary scheme.

The paper concludes with some case studies that demonstrate key features of the FVTD scheme. These case studies include one-dimensional wave propagation in a free-space cavity and in a partially filled dielectric cavity, three-dimensional wave propagation in a circular waveguide, and three-dimensional scattering from a post in an unbounded medium.

2. FORMULATION

It is well known that the mathematical description of the electromagnetic field is given as a solution to the two curl and two divergence equations of Maxwell [6]. In the ensuing subsections, a short description of these equations is provided and methods for their discretization are examined in detail.

2.1 Point Vector Form

Letting $\mathbf{E}$ and $\mathbf{H}$ denote the vector symbols of the electric and magnetic fields, respectively, Maxwell postulated that

$$\mu \frac{\partial \mathbf{H}}{\partial t} + \nabla \times \mathbf{E} = 0, \tag{1}$$

$$\epsilon \frac{\partial \mathbf{E}}{\partial t} - \nabla \times \mathbf{H} = 0, \tag{2}$$

$$\nabla \cdot \mathbf{E} = 0, \tag{3}$$

and

$$\nabla \cdot \mathbf{H} = 0. \tag{4}$$

The previous description assumes that the space under consideration is devoid of sources. In addition, the assumptions that the permittivity ϵ and the permeability μ are scalars and of constant value are made. For purposes of numerical computation, the two curl equations are discretized and advanced in time; the two divergence equations are additional constraints imposed upon the fields.

An alternate form of Maxwell's equations that has significance to certain discretization operators is obtained by adding and subtracting like terms within the two curl equations:

$$\mu \frac{\partial \mathbf{H}}{\partial t} + \nabla \times \mathbf{F}_e^+ + \nabla \times \mathbf{F}_e^- = 0 \tag{5}$$

$$\epsilon \frac{\partial \mathbf{E}}{\partial t} - \nabla \times \mathbf{F}_m^+ - \nabla \times \mathbf{F}_m^- = 0, \tag{6}$$

where

$$\mathbf{F}_e^+ = (\mathbf{E} - \eta \mathbf{n} \times \mathbf{H})/2, \tag{7}$$

$$\mathbf{F}_e^- = (\mathbf{E} + \eta \mathbf{n} \times \mathbf{H})/2, \tag{8}$$

$$\mathbf{F}_m^+ = \frac{1}{2}\left(\mathbf{H} + \frac{\mathbf{n} \times \mathbf{E}}{\eta}\right), \tag{9}$$

and

$$\mathbf{F}_m^- = \frac{1}{2}\left(\mathbf{H} - \frac{\mathbf{n} \times \mathbf{E}}{\eta}\right). \tag{10}$$

In Eqns. (7)–(10), $\mathbf{n}$ is an arbitrary unit vector yet to be specified. Obviously, $\mathbf{F}_e^+ + \mathbf{F}_e^- = \mathbf{E}$ and $\mathbf{F}_m^+ + \mathbf{F}_m^- = \mathbf{H}$, as required. More

importantly, for a plane wave traveling in the $\mathbf{n}$ direction, $\mathbf{F}_e^-$ and $\mathbf{F}_m^-$ are identically zero, whereas $\mathbf{F}_e^+$ and $\mathbf{F}_m^+$ are not. Likewise, for a plane wave traveling in the negative $\mathbf{n}$ direction, $\mathbf{F}_e^+$, and $\mathbf{F}_m^+$ are identically zero, whereas $\mathbf{F}_e^-$ and $\mathbf{F}_m^-$ are not. Thus, it can be argued that the role of $\mathbf{F}_e^+$, $\mathbf{F}_m^+$, $\mathbf{F}_e^-$, and $\mathbf{F}_m^-$ is to discriminate between a plane wave traveling in the positive or negative $\mathbf{n}$ directions.

2.2 Global Vector Form

The global form of the governing equations can be obtained by integrating the point-form equations over some volume V. The volume is enclosed by the surface S, which is defined by the outward pointing normal $\mathbf{n}$. Integrating over V and invoking the vector curl theorem, we obtain

$$\epsilon \frac{\partial \mathbf{E}^a}{\partial t} = \frac{1}{V} \int_S \mathbf{n} \times \mathbf{H} \, dS \tag{11}$$

and

$$\mu \frac{\partial \mathbf{H}^a}{\partial t} = \frac{1}{V} \int_S \mathbf{E} \times \mathbf{n} \, dS. \tag{12}$$

Here $\mathbf{E}^a$ and $\mathbf{H}^a$ are the average values of $\mathbf{E}$ and $\mathbf{H}$, respectively, within the volume V. That is,

$$\mathbf{E}^a = \frac{1}{V} \int_V \mathbf{E} \, dV \tag{13}$$

and

$$\mathbf{H}^a = \frac{1}{V} \int_V \mathbf{H} \, dV. \tag{14}$$

A chief attribute of both Eqns. (11) and (12) is the exclusive manifestation of tangential field components on the volume surface. This attribute is compatible with the boundary conditions for $\mathbf{E}$ and $\mathbf{H}$ [6].

The integration of Eqns. (5) and (6) throughout the volume V is accomplished in the exact same manner as in the previous section. The result is

$$\epsilon \frac{\partial \mathbf{E}^a}{\partial t} = \frac{1}{V} \int_S \mathbf{n} \times \mathbf{F}_m^+ \, dS + \frac{1}{V} \int_S \mathbf{n} \times \mathbf{F}_m^- \, dS \tag{15}$$

and

$$\mu \frac{\partial \mathbf{H}^a}{\partial t} = \frac{1}{V} \int_S \mathbf{F}_e^+ \times \mathbf{n} \, dS + \frac{1}{V} \int_S \mathbf{F}_e^- \times \mathbf{n} \, dS. \tag{16}$$

The unit vector $\mathbf{n}$ in Eqns. (7)–(10) is now defined to be the vector normal to the surface S.

2.3 Global Curvilinear Form

Given the similarity of Eqns. (11), (12), (15), and (16), we consider the following canonical integral equation as the basis for the ensuing mathematical development:

$$\frac{\partial \mathbf{U}^a}{\partial t} = \frac{1}{V} \int_S \mathbf{n} \times \mathbf{F} \, dS \tag{17}$$

where $\mathbf{U}^a$ denotes either $\epsilon \mathbf{E}^a$ or $\mu \mathbf{H}^a$; $\mathbf{F}$ denotes either $-\mathbf{E}$, $\mathbf{H}$, $-\mathbf{F}_e^+ - \mathbf{F}_e^-$ or $\mathbf{F}_m^+ + \mathbf{F}_m^-$, depending upon the context. Also, in anticipation of the numerical procedure, the volume V is assumed to be comprised of six faces S_1, S_2, $\cdots$, S_6; the sum of these faces is the closed surface S. For a six-sided volume, Eqn. (17) reduces to

$$\frac{\partial \mathbf{U}^a}{\partial t} = \frac{1}{V} \sum_{l=1}^{6} \mathbf{I}_l \tag{18}$$

where

$$\mathbf{I}_l = \int_{S_l} \mathbf{n}_l \times \mathbf{F} \, dS; \tag{19}$$

here $\mathbf{n}_l$ is the outward-pointing normal to S_l.

The evaluation of the integrals in Eqn. (18) is facilitated by establishing a curvilinear transformation that maps V into V', the transformed region of integration. The mapping is accomplished through the following equations: $\xi = \xi(x, y, z)$, $\eta = \eta(x, y, z)$, and $\zeta = \zeta(x, y, z)$. By invoking the Jacobian J, we compute the average value via

$$\mathbf{U}^a = \frac{1}{V} \int_{V'} \mathbf{U}' J \, d\xi d\eta d\zeta; \tag{20}$$

here $\mathbf{U}'(\xi, \eta, \zeta, t) = \mathbf{U}(x(\xi, \eta, \zeta), y(\xi, \eta, \zeta), z(\xi, \eta, \zeta), t)$.

Next consider the integration over the six faces. The surfaces S_1, S_2, S_3, S_4, S_5, and S_6 are defined by $\xi = \xi_1$ (ξ_1 is a constant), $\xi = \xi_2$, $\eta = \eta_1$, $\eta = \eta_2$, $\zeta = \zeta_1$, and $\zeta = \zeta_2$, respectively. One of these surfaces, say S_2 and its integral $\mathbf{I}_2$, is the focus of the following development; the remaining five surface integrals are evaluated in a

similar fashion. Considering Eqn. (19) for $l = 2$, we obtain, through a change of variables the equation,

$$\mathbf{I}_2 = \int_{S_2'} \mathbf{n}_2 \times \mathbf{F}'(\xi_2, \eta, \zeta, t) J |\nabla \xi| d\eta d\zeta, \tag{21}$$

where $\mathbf{F}'(\xi, \eta, \zeta, t) = \mathbf{F}(x(\xi, \eta, \zeta),\, y(\xi, \eta, \zeta),\, z(\xi, \eta, \zeta), t)$ and S_2' is the transformed region associated with S_2. Since $\nabla \xi$ is perpendicular to surfaces of constant ξ, $\mathbf{n}_2 |\nabla \xi| = \nabla \xi$ and

$$\mathbf{I}_2 = \int_{S_2'} \nabla \xi \times \mathbf{F}'(\xi_2, \eta, \zeta, t) J d\eta d\zeta. \tag{22}$$

One approach towards evaluating I_2 from known cell average values consists of introducing a primitive vector $\mathbf{W}$, whose derivative is I_2 [7]. Stated mathematically,

$$\frac{\partial \mathbf{W}_2}{\partial \xi} \equiv \mathbf{I}_2 \tag{23}$$

Primitive vectors for the other surface integrals are defined in a similar manner. It is apparent from the above equation that value of the surface integral hinges upon the mechanism for computing the primitive vector and its derivative.

Values for the primitive vector are deduced by considering a more general definition for its derivative than the one provided by Eqn. (23). Let

$$\frac{\partial \mathbf{W}(\xi)}{\partial \xi} = \int_{S'(\xi)} \nabla \xi \times \mathbf{F}'(\xi, \eta, \zeta) J d\eta d\zeta. \tag{24}$$

From this definition, it follows that $S'(\xi_2) = S_2'$ and

$$\left. \frac{\partial \mathbf{W}(\xi)}{\partial \xi} \right|_{\xi = \xi_2} = \frac{\partial \mathbf{W}_2}{\partial \xi}. \tag{25}$$

Then, by means of Riemann integration of Eqn. (24),

$$\mathbf{W}(\xi) = \mathbf{W}(\xi_1) + \int_{\xi_1}^{\xi} \int_{S'(\tau)} \nabla \xi \times \mathbf{F}'(\tau, \eta, \zeta) J d\eta d\zeta d\tau. \tag{26}$$

The evaluation of this integral at $\xi = \xi_2$ yields

$$\mathbf{W}(\xi_2) = \mathbf{W}(\xi_1) + \int_{\xi_1}^{\xi_2} \int_{S'(\tau)} \nabla\xi \times \mathbf{F}'(\tau, \eta, \zeta) J d\eta d\zeta d\tau. \tag{27}$$

By definition the integration term is proportional to the average value of $\nabla\xi \times \mathbf{F}$ over the volume V and hence,

$$\mathbf{W}(\xi_2) = \mathbf{W}(\xi_1) + V(\nabla\xi \times \mathbf{F})^a. \tag{28}$$

The antecedent equation fundamentally states that the *exact* value of $\mathbf{W}$ at domain walls can be determined from the *average* value of $\nabla\xi \times \mathbf{F}$ within that domain.

The previous equation is exact, provided that the metric information associated with $\nabla\xi$ is known exactly. To compute the metrics, a mathematical description of the ξ is required. Consider the description $\xi = \xi_l + \xi_n$, where ξ_l and ξ_n represent the linear and non-linear terms of ξ, respectively. Since $\nabla\xi_l$ is constant, it follows from Eqn. (28) that

$$\mathbf{W}(\xi_2) = \mathbf{W}(\xi_1) + V\nabla\xi_l \times \mathbf{F}^a + V(\nabla\xi_n \times \mathbf{F})^a. \tag{29}$$

If $\nabla\xi_n$ is regarded as sufficiently small throughout the volume, then

$$\mathbf{W}(\xi_2) \approx \mathbf{W}(\xi_1) + V\nabla\xi_l \times \mathbf{F}^a, \tag{30}$$

which is a second-order approximation. The error of this approximation ε is obtained by the mean value theorem of integral calculus:

$$\varepsilon = \mathbf{F}'(\xi_\circ, \eta_\circ, \zeta_\circ) \int_{V'} \nabla\xi_n d\xi d\eta d\zeta, \tag{31}$$

where $\xi_\circ, \eta_\circ, \zeta_\circ$ is some point in the volume V'. For volumes that are parallelepipeds, $\varepsilon = 0$.

A value for $\nabla\xi_l$ can be obtained from the geometry of the volume. In the Appendix, we show that $\nabla\xi \approx \delta_\xi \mathbf{A}_\xi^a / V$, where δ_ξ is the distance between S_1 and S_2 and $\mathbf{A}_\xi^a$ is the average directed surface area in the ξ direction within V. By definition,

$$\mathbf{A}_\xi^a = \frac{1}{\delta_\xi} \int_{\xi_1}^{\xi_2} \mathbf{A}(\xi) d\xi. \tag{32}$$

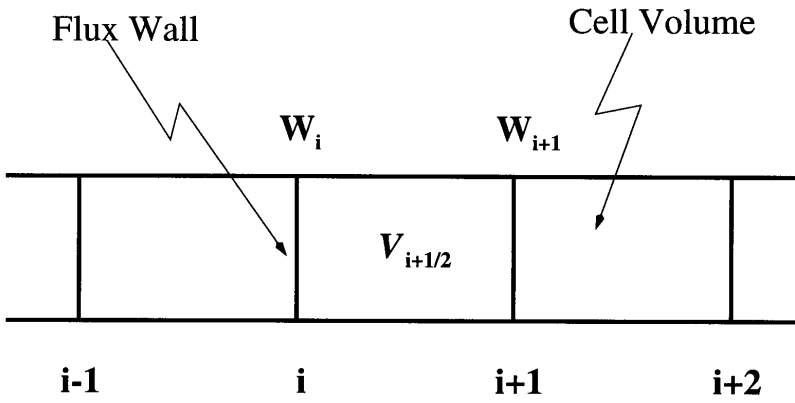

Figure 1. A pictorial representation of the indexing scheme used in the flux reconstruction.

Therefore, Eqn. (30) reduces to

$$\mathbf{W}(\xi_2) \approx \mathbf{W}(\xi_1) + \delta_\xi \mathbf{A}^a_\xi \times \mathbf{F}^a, \tag{33}$$

Due to the simplicity of the Eqn. (33), the code development for the primitive vector is trivial.

To obtain an approximation for the derivative of the primitive vector, the domain V is sub-divided into many sub-volumes, each denoted as $V_{i-1/2,j-1/2,k-1/2}$; Eqn. (18) is now an equation over each sub-domain. The six surfaces defining each sub-volume are still denoted locally as S_1 through S_6. Globally, the surfaces S_1 and S_2 are defined by the equations $\xi = \xi_{i-1}$ and $\xi = \xi_i$, respectively. Likewise, $\eta = \eta_{j-1}$, $\eta = \eta_j$, $\zeta = \zeta_{k-1}$, and $\zeta = \zeta_k$ globally define S_3, S_4, S_5, and S_6, respectively. Using this notation, we cast Eqn. (28) as follows:

$$\mathbf{W}_i = \mathbf{W}_{i-1} + V_{i-1/2}(\nabla\xi \times \mathbf{F})^a_{i-1/2}, \qquad 1 \le i \le I_{\text{pts}} \tag{34}$$

where $(\nabla\xi \times \mathbf{F})^a_{i-1/2}$ is the average value of $\nabla\xi \times \mathbf{F}$ in the volume $V_{i-1/2}$ defined by the grid lines $i-1$ and i. (The indices j, k are dropped for clarity; see Figure 1 for a one-dimensional depiction.) Since $\mathbf{W}$ is to be differentiated, $\mathbf{W}_0$ may be set to zero, or to any other value.

To estimate the derivative of $\mathbf{W}$ at the cell surface i (i.e., the value of $\mathbf{I}_2$), we simply employ standard difference stencils. For example, a second-order central approximation is given by

$$\left.\frac{\partial \mathbf{W}}{\partial \xi}\right|_i \approx \left(\frac{\mathbf{W}_{i+1} - \mathbf{W}_{i-1}}{2\delta_\xi}\right). \tag{35}$$

Substituting Eqn. (34) in Eqn. (35), we obtain

$$\left.\frac{\partial \mathbf{W}}{\partial \xi}\right|_i \approx \frac{V_{i+1/2}(\nabla\xi \times \mathbf{F})^a_{i+1/2} + V_{i-1/2}(\nabla\xi \times \mathbf{F})^a_{i-1/2}}{2\delta_\xi} \tag{36}$$

Under the assumption that the metrics are computed to second-order,

$$\left.\frac{\partial \mathbf{W}}{\partial \xi}\right|_i \approx \left(\frac{(\mathbf{A}^a_\xi \times \mathbf{F}^a)_{i+1/2} + (\mathbf{A}^a_\xi \times \mathbf{F}^a)_{i-1/2}}{2}\right), \tag{37}$$

where $(\mathbf{A}^a_\xi \times \mathbf{F}^a)_{i\pm1/2}$ is the average value of $\mathbf{A}_\xi$ in volume $i \pm 1/2$ crossed into the average value of $\mathbf{F}$ in volume $i \pm 1/2$. Eqn. (37) is valid for cells of unequal size or for stretched grids. For uniform grids, it seems reasonable to replace Eqn. (37) with

$$\left.\frac{\partial \mathbf{W}}{\partial \xi}\right|_i \approx \mathbf{A}_i \times \left(\frac{\mathbf{F}^a_{i+1/2} + \mathbf{F}^a_{i-1/2}}{2}\right). \tag{38}$$

Per Eqn. (23), the right hand side of Eqn. (38) is an estimated value of the surface integral $\mathbf{I}_2$. The equation states that the value of the surface integral is *reconstructed* from average values of $\mathbf{F}$.

The previous equation is recognized as being nothing more than a central average reconstruction and is one of many reconstruction formulas. As with the primitive vector, its computation is trivial. Within this equation, there exists two approximations. The first is associated with the discretization of the derivative of the primitive vector. The second is associated with the accuracy of the metric computation. Both of these approximations influence the overall accuracy of the reconstructed value. Instead of using a second-order central derivative, a fourth-order central derivative could be employed. If however, the metric is computed to second-order, the reconstructed value is strictly second-order. For cells that are near parallelepipeds, the nonlinear metric terms are negligible and the accuracy of the reconstructed

value is strongly influenced by the accuracy of the derivative calculation. Thus, even if the scheme is not strictly high-order, it still may be highly accurate.

One difficulty with the employment of central differences for the calculation of the primitive vector derivative (in conjunction with nodally collocated schemes) is the problem of grid decoupling [8]. Since a central operator has holes in its stencil, the unknown and its derivative at a node are decoupled from each other. For point-source excitation, the result of this effect is an interlacement of zeros in the data [9]. To overcome this problem, we consider a split flux treatment, which discretizes $\mathbf{F}^+$ and $\mathbf{F}^-$, rather than $\mathbf{F}$ [10]. Since $\mathbf{F} = \mathbf{F}^+ + \mathbf{F}^-$, it follows from Eqn. (34) that $\mathbf{W} = \mathbf{W}^+ + \mathbf{W}^-$, where

$$\mathbf{W}_i^+ = \mathbf{W}_{i-1}^+ + (\nabla\xi \times \mathbf{F}^+)_{i-1/2}^a, \qquad 1 \le i \le I_{\text{pts}} \tag{39}$$

and

$$\mathbf{W}_i^- = \mathbf{W}_{i-1}^- + (\nabla\xi \times \mathbf{F}^-)_{i-1/2}^a, \qquad 1 \le i \le I_{\text{pts}} \tag{40}$$

The differentiation of $\mathbf{W}^+$ and $\mathbf{W}^-$ is accomplished by establishing difference stencils within the domain of influence of these two vectors. That is, for the positive-going waves (i.e., $\mathbf{F}^+$), backward stencils are invoked; for the negative-going waves (i.e., $\mathbf{F}^-$), forward stencils are invoked. For example, a second-order windward approximation yields

$$\left.\frac{\partial \mathbf{W}^+}{\partial \xi}\right|_i \approx \left(\frac{3\mathbf{W}_i^+ - 4\mathbf{W}_{i-1}^+ + \mathbf{W}_{i-2}^+}{2\delta_\xi}\right) \tag{41}$$

and

$$\left.\frac{\partial \mathbf{W}^-}{\partial \xi}\right|_i \approx \left(\frac{-3\mathbf{W}_i^- + 4\mathbf{W}_{i+1}^- - \mathbf{W}_{i+2}^-}{2\delta_\xi}\right). \tag{42}$$

For second-order metrics, the final result, as obtained from the previous equations and expressed in terms of $\mathbf{F}^a$, is

$$\left.\frac{\partial \mathbf{W}^+}{\partial \xi}\right|_i \approx \mathbf{A}_i \times \left(\frac{3(\mathbf{F}^+)_{i-1/2}^a - (\mathbf{F}^+)_{i-3/2}^a}{2}\right), \tag{43}$$

and

$$\left.\frac{\partial \mathbf{W}^-}{\partial \xi}\right|_i \approx \mathbf{A}_i \times \left(\frac{3(\mathbf{F}^-)_{i+1/2}^a - (\mathbf{F}^-)_{i+3/2}^a}{2}\right), \tag{44}$$

The sum of these two derivatives is the derivative of the total primitive vector (or, the value of the surface integral under investigation.)

A table of commonly used stencils that may be invoked for the derivative of $\mathbf{W}$, $\mathbf{W}^+$ or $\mathbf{W}^-$ may be found in many texts (e.g., [11]). Other stencils that are neither central or strictly forward (or backward) can also be constructed via standard Taylor series analysis. For such stencils, we denote them as windward biased. In the numerical examples section, we choose to adopt the third-order windward biased stencil (3WB1) as the stencil under investigation. That is, let

$$\frac{\partial \mathbf{W}^+}{\partial \xi}\Big|_i \approx \left(\frac{2\mathbf{W}^+_{i+1} + 3\mathbf{W}^+_i - 6\mathbf{W}^+_{i-1} + \mathbf{W}^+_{i-2}}{6\delta_\xi} \right) \tag{45}$$

and

$$\frac{\partial \mathbf{W}^-}{\partial \xi}\Big|_i \approx \left(\frac{-2\mathbf{W}^-_{i-1} - 3\mathbf{W}^-_i + 6\mathbf{W}^-_{i+1} - \mathbf{W}^-_{i+2}}{6\delta_\xi} \right). \tag{46}$$

Based upon stability arguments and four-stage Runge-Kutta integration, central differences may be used for either $\mathbf{W}$, $\mathbf{W}^+$ or $\mathbf{W}^-$. Windward and windward biased stencils are typically used for only $\mathbf{W}^+$ or $\mathbf{W}^-$.

This section is closed by noting that all of the difference stencils discussed so far are of the explicit type. An alternative to explicit differencing is implicit differencing, or compact differencing. For example, consider the following central compact operator [12, 13]:

$$\alpha \mathbf{Q}_{i+1} + \mathbf{Q}_i + \alpha \mathbf{Q}_{i-1} = \beta(\mathbf{W}_{i+1} - \mathbf{W}_{i-1}). \tag{47}$$

If $\alpha = 1/4$ and $\beta = 3/4$, then $\mathbf{Q}_i$ is approximately the derivative of $\mathbf{W}$ at i; the degree of approximation is fourth-order. Unlike the fourth-order explicit central operator, which has a stencil that spans five cells, the fourth-order compact operator spans only three cells. This attribute is often quoted in the context of boundary operator implementation. Although the above equation represents a tri-diagonal matrix equation, the solution of that equation is easily obtained by employing the Thomas algorithm [14]. The number of operations associated with the computation of $\mathbf{Q}_i$ is five.

2.4 Boundary Conditions

In addition to the equations that govern propagation, Maxwell's equations in global form require that tangential $\mathbf{E}$ and $\mathbf{H}$ be con-

tinuous across a dielectric interface. At a perfect electrical conductor defined by the normal $\mathbf{n}$ there exist two conditions of interest:

$$\mathbf{n} \times \mathbf{E} = 0 \qquad \mathbf{n} \times \nabla \times \mathbf{H} = 0. \tag{48}$$

These equations are equivalent to

$$\mathbf{E}_t = 0 \qquad \frac{\partial \mathbf{H}_t}{\partial n} = 0, \tag{49}$$

where the subscript t denotes the tangential component. Although other boundary conditions may be derived, uniqueness of solution is assured when the tangential fields are specified on the boundary. Moreover, the boundary conditions are compatible with the integrand of the surface integrals, which highlights an important feature of the nodally collocated finite-volume procedure. The algorithmic implementation of these boundary conditions is considered next.

Suppose that $i = 0$ (locally defined surface S_1) resides on a perfect electric conductor (PEC) and the domain is specified by $i \geq 0$. If the integrand of $\mathbf{I}_1$ is a function of tangential $\mathbf{E}$, then $\mathbf{I}_1 \equiv 0$. If the integrand of $\mathbf{I}_1$ is a function of tangential $\mathbf{H}$, then from the identity

$$\frac{\partial \mathbf{W}_1}{\partial \xi} = \frac{1}{V} \int_{S_1} \mathbf{n}_1 \times \mathbf{H} \, dS, \tag{50}$$

it follows that a Neumann condition is implemented by setting

$$\frac{\partial^2 \mathbf{W}_1}{\partial \xi^2} = 0 \qquad \text{on } S_1. \tag{51}$$

To second-order, this condition is satisfied by employing an image term $\mathbf{W}_1(-1)$ and by setting that image term to the value of $2\mathbf{W}_1(0) - \mathbf{W}(1)$. With this image term, second-order central (2C) or 3WB1 operators can be invoked on the boundary; second-order windward (2W) or 3WB1 operators can be invoked one cell in from the boundary. Note: For the case when 3WB1 operators are used in the interior and 2C operators are used on the boundary, the global spatial accuracy is still third-order [15].

For the situation when S_1 resides on a dielectric interface, at least two possibilities arise. The fact that the tangential fields are continuous across the interface (but a discontinuous first derivative) allows one

to reconstruct at that interface with second-order central averages; the reconstruction is in non-split form. In split form, the process of reconstruction is more complicated. The complication arises from the discontinuity of $\mathbf{F}$, due to the η term. To handle this discontinuity, the primitive vectors associated with $\mathbf{E}$ and $\mathbf{H}$ are built as before. However, two values of $\mathbf{F}$ (i.e., two values of $\mathbf{F}^+$ and two values of $\mathbf{F}^-$) must be deduced, depending upon the cell under consideration. To the left of S_1, $\mathbf{F}$ is a function of η_a; this $\mathbf{F}$ is used in the time advancement of $\mathbf{U}^a_{i-1/2}$. To the right of S_1, $\mathbf{F}$ is a function of η_b; this $\mathbf{F}$ is used in the time advancement of $\mathbf{U}^a_{i+1/2}$.

The truncation of the domain or the development of an absorbing boundary condition (ABC) can be accomplished by several means. The perfectly matched layer (PML) is one such possibility [16]. The development of such a PML for high-order FVTD schemes is the subject of future work. In this work, characteristic theory is adopted for simplicity. In split flux form, we have already noted that $\mathbf{F}^+$ and $\mathbf{F}^-$ only act on positive- and negative-going waves, respectively. Thus a suitable ABC can be devised by setting either $\mathbf{F}^+$ or $\mathbf{F}^-$ to zero, depending upon which vector represents an out-going wave.

2.5 Integration

Upon completion of the reconstruction phase, the semi-discrete equations take on the form

$$\frac{\partial \mathbf{u}}{\partial t} = \mathbf{A}\mathbf{u}, \tag{52}$$

which constitutes a constant coefficient, linear system of ordinary differential equations. Here $\mathbf{u}$ represents every value of $\mathbf{U}^a$ throughout the computational domain and $\mathbf{A}$ is the connectivity matrix. Since our desire is to obtain high-order accuracy, the classical four-stage scheme is invoked and labeled as RK4. In vector form and in terms of the auxiliary vectors $\mathbf{K}_1 \cdots \mathbf{K}_4$, the standard RK4 scheme is as follows [17]:

$$\begin{aligned}
\mathbf{K}_1 &= \delta \mathbf{A}\mathbf{u}^n \\
\mathbf{K}_2 &= \delta \mathbf{A}\left(\mathbf{u}^n + \mathbf{K}_1/2\right) \\
\mathbf{K}_3 &= \delta \mathbf{A}\left(\mathbf{u}^n + \mathbf{K}_2/2\right) \\
\mathbf{K}_4 &= \delta \mathbf{A}(\mathbf{u}^n + \mathbf{K}_3) \\
\mathbf{u}^{n+1} &= \mathbf{u}^n + (\mathbf{K}_1 + 2\mathbf{K}_2 + 2\mathbf{K}_3 + \mathbf{K}_4)/6.
\end{aligned} \tag{53}$$

Other techniques for low storage RK integrators are reported in the literature [18, 19]. Upon the incorporation of all four stages into the equation for $\mathbf{u}^{n+1}$, the final result is

$$\mathbf{u}^{n+1} = \left(1 + \delta_t \mathbf{A} + \frac{\delta_t^2}{2}\mathbf{A}^2 + \frac{\delta_t^3}{6}\mathbf{A}^3 + \frac{\delta_t^4}{24}\mathbf{A}^4\right)\mathbf{u}^n. \tag{54}$$

Furthermore, by substituting Eqn. (52) into Eqn. (54), we obtain

$$\mathbf{u}^{n+1} = \left(1 + \delta_t \frac{\partial}{\partial t} + \frac{\delta_t^2}{2}\frac{\partial^2}{\partial t^2} + \frac{\delta_t^3}{6}\frac{\partial^3}{\partial t^3} + \frac{\delta_t^4}{24}\frac{\partial^4}{\partial t^4}\right)\mathbf{u}^n. \tag{55}$$

which is the expected truncated Taylor series. These two equations reveal the main sources of error: 1) Spatial discretization errors incorporated in the approximation of $\partial\mathbf{u}/\partial t$ by $\mathbf{Au}$ and 2) temporal errors as a result of the truncated Taylor series. The manifestations of these errors, usually couched in terms of dissipation and dispersion errors, are best characterized in the frequency domain. Since such an analysis has been provided in many sources, we only note that the predominant error mechanism is dissipation [12].

One final note: Even though upwind schemes are well known to prevent grid decoupling, grid decoupling can be induced via the Runge-Kutta integrator. Since the Runge-Kutta integrator requires that both field components be advanced simultaneously, the excitation of one field component is not made manifest to the other field component at the instant the excitation is applied. This effect is most noticeable when modeling volumetric and sheet sources; the effect is observed near the source, where the data extremely alternates between positive and negative values.

3. NUMERICAL EXAMPLES

In the examples that follow, 3WB1 reconstruction formulas are invoked. We choose this stencil for the following reasons: 1) The combination of 3WB1 interior stencils with 2C boundary operators yields a globally third-order accurate spatial scheme and 2) Windward-based stencils for nodally collocated schemes partially inhibit grid-decoupled solutions. (Recall that central-based stencils induce grid-decoupled solutions). As discussed next, this stencil choice may or may not always yield satisfactory results.

We begin this section with a simple problem of a guassian pulse in a one-dimensional perfectly conducting cavity. The form of the pulse is $e^{-w^2t^2}$, where $w = 4.14 \times 10^{10}\,1/\mathrm{s}$; this value for w yields a frequency spread in excess of $20\,\mathrm{GHz}$. For the simulation parameters, let $\delta_t = 1.179\,\mathrm{ps}$ and $\delta_x = 0.0005\,\mathrm{m}$. The pulse is assumed to be traveling at the speed of light. To ascertain both the dissipative and dispersive effects on the pulse, we marched the pulse 11,314 time steps, which is the required time for the pulse to attenuate by ten percent from its initial value; see Figure 2. The slight ripple effect seen in that plot of Figure 2 is to be expected due to dissipation, dispersion and boundary related errors. However, since the pulse has traveled twenty round trips, such effects are certainly anticipated.

To validate the dielectric boundary condition methodology, the problem of a guassian pulse within a partially filled dielectric cavity is considered; the dielectric has a relative permittivity value of 2.3 and fills the first fifty cells of the 200 cell cavity. The pertinent parameters are the same as in the previous example. The initial pulse is centered at cell 100 and is traveling to the left as time advances. Representative results after 230 time steps are provided in Figure 3. As seen directly from the figure, the theoretical transmission coefficient of -0.795 and the theoretical reflection coefficient of -0.205 are predicted.

The one-dimensional problems clearly demonstrate the validity of the FVTD theory. The aforementioned examples demonstrate both the long-term stability of the simulation and the validity of the designed boundary operators for perfectly conducting and dielectric interfaces. For three-dimensional curvilinear domains, the results are not as promising, as discussed next.

The propagation of an electromagnetic wave in a cylindrical waveguide of radius x is the subject of the following discussion; the waveguide is filled with free-space. The domain spans $10 \times 20 \times 100$ cells in ρ, ϕ, and z, respectively. The domain is assumed to be periodic in z and the electromagnetic wave is initialized with the TE_{11} time-harmonic mode of frequency $\omega = 2\pi\,\mathrm{r/s}$. In the radial direction, the signal is sampled at twenty points per wavelength. For this study, let $c = 1\,\mathrm{m/s}$, $\delta_\rho = 0.1\,\mathrm{m}$, $\delta_\phi = 2\pi/20\,\mathrm{r}$, $\delta_z = 5\lambda_z/100 = 0.05229\,\mathrm{m}$, and $\delta_t = 0.01269\,\mathrm{s}$. Figure 4 shows the electric field distribution among various planes within the guide. This figure serves as a qualitative check by showing that the mode has been properly set up within the guide. Next consider Figures 5 and 6, which show all six components of the

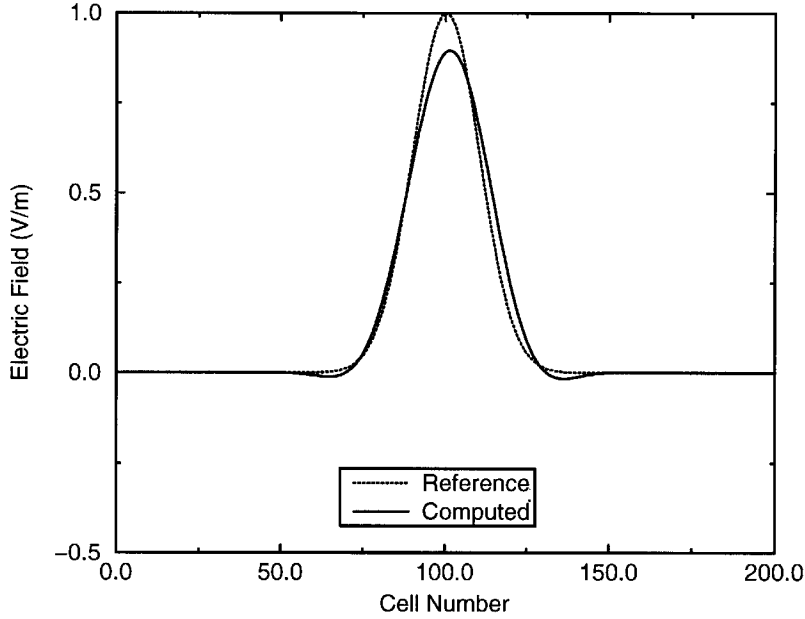

Figure 2. Electric field inside a perfectly conducting cavity.

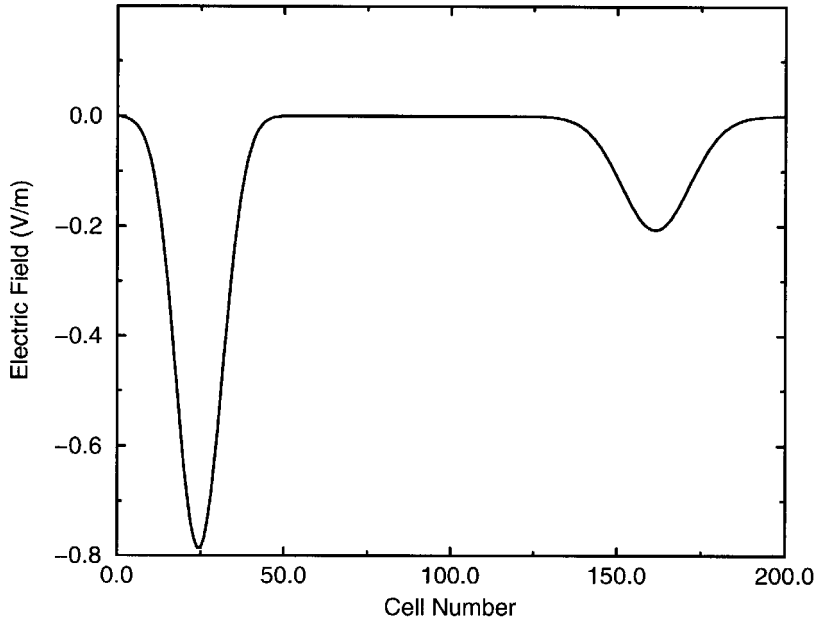

Figure 3. Electric field inside a partially filled dielectric cavity.

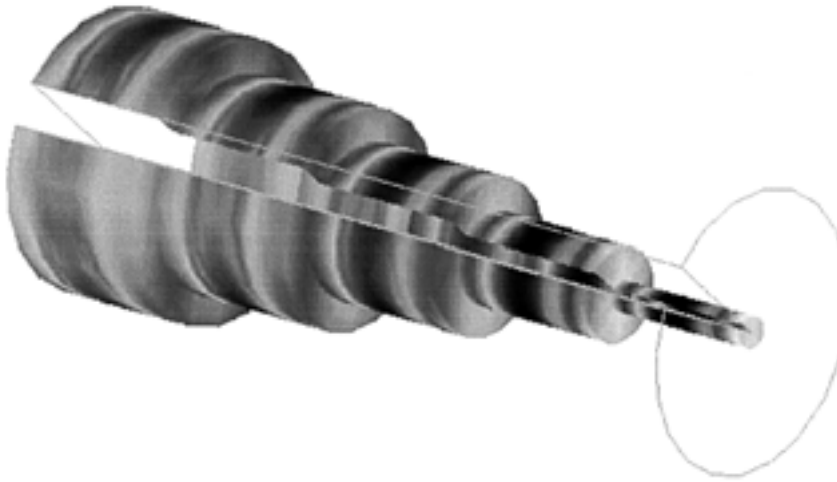

Figure 4. The electric field within the cylindrical waveguide.

electromagnetic signal as a function of radial cell distance at time steps 200 and 2000, respectively. Clearly, the solver is able to predict the basic field phenomenology within the guide. Upon closer examination of the data, however, one notes the small effect of numerical dissipation in the early time and large dissipative effects in the late time. Note: After 2000 times steps, the total energy in the electromagnetic wave has decreased by a factor of three. The highly dissipative nature of the computed solution is consistent with the theoretical performance of windward schemes. A second effect is also made manifest in the data: A non-zero longitudinal electric field. The longitudinal component is indeed small relative to its transverse counterparts, which implies that solver has set up a quasi-TE mode rather than a pure TE mode. The generation of this non-physical component is due to the grid and its corresponding metrics. Since each cell is not a parallelepiped, the numerically computed metrics are only second-order accurate (for parallelepiped cells, the metrics are computed exactly.) Hence, the nonlinear terms of the metrics are not captured and the corresponding solution is corrupted by metric induced errors.

To demonstrate the effect of grid decoupling due to the four-stage Runge-Kutta integrator, we provide Figure 7. In this case, a perfectly conducting post is placed between two perfectly conducting plates

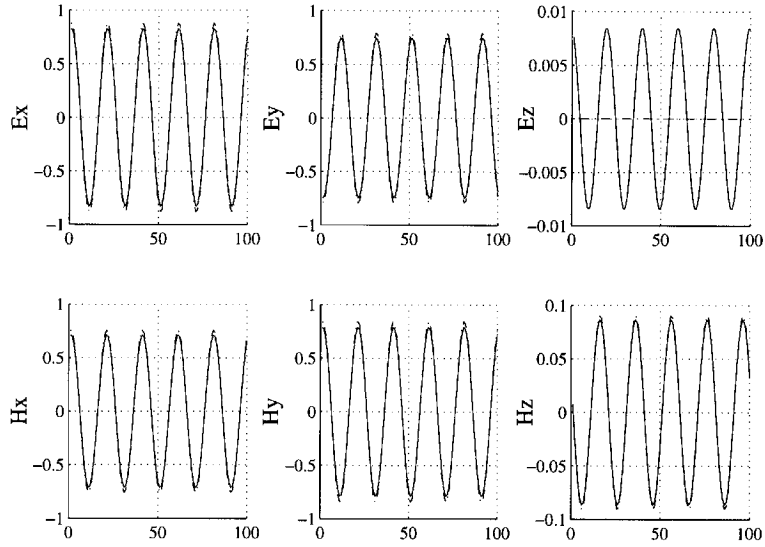

Figure 5. The field components of **E** and **H** within the cylindrical waveguide at $t = 200\delta_t$. The dotted lines are the analytical solution; the solid lines are the computed solution.

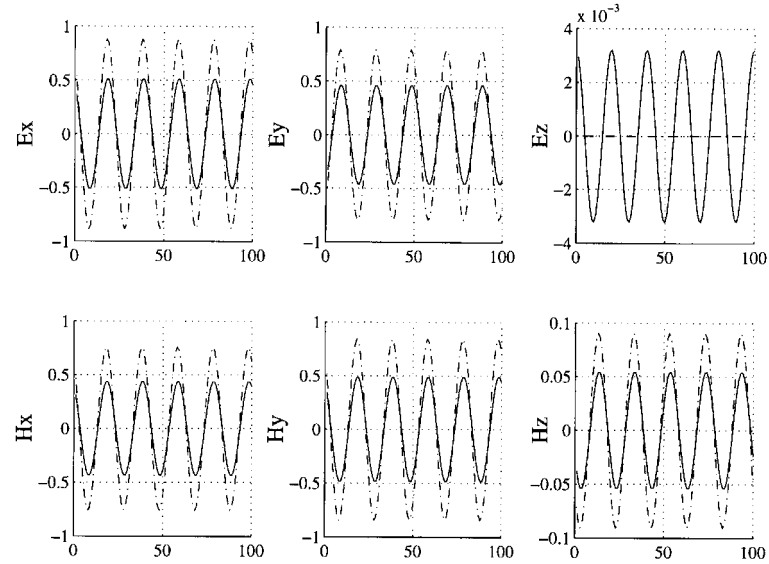

Figure 6. The field components of **E** and **H** within the cylindrical waveguide at $t = 2,000\delta_t$. The dotted lines are the analytical solution; the solid lines are the computed solution.

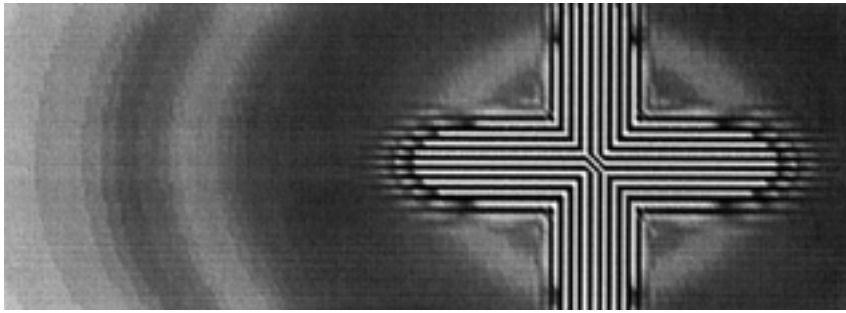

Figure 7. Near-field electromagnetic field data about a post. The striping of the data suggests strong grid decoupling.

of height $0.4\,\mathrm{mm}$. (Even though the solver is simulating a three-dimensional geometry, the plates image this post continuously to infinity, thus reducing the dimensionality of the solution to two.) The post occupies one cell in its cross-section and is excited by a uniform plane wave traveling at the speed of light. A scattered field formulation is used for this scenario in which $\delta_x = 0.2\,\mathrm{cm}$, $\delta_y = 0.2\,\mathrm{cm}$, $\delta_z = 0.0667\,\mathrm{cm}$, and $\delta_t = .1742\,\mathrm{ps}$. Per Figure 7, the grid decoupling effect is obviously present in the near-field data. In the far-field, the solver predicts the anticipated cylindrical wave.

4. CONCLUDING REMARKS

This paper has provided a rigorous treatment of the nodally collocated FVTD methodology. The treatment clearly demonstrates how one can design the reconstruction formulas for both low and high-order, accurate solutions – the accuracy is partially contingent upon the discrete approximation of the derivative of the primitive vector. Moreover, from this same primitive vector, suitable boundary operators can be equally designed. Finally, for curvilinear grids, the analysis herein clearly shows the role of the metric calculation on solution accuracy. Since numerous works have discussed stability, dispersion errors and dissipation errors for this scheme (e.g., see [20]), such a discussion has not been provided.

Instead of adjusting the numerical parameters in such a way to yield *excellent* results or to select examples that only show only the positive attributes of the scheme, we have chosen to use typical parameters and

common geometries to demonstrate certain deficiencies of the scheme. These deficiencies include strong dissipation errors of windward based stencils, metric errors associated with non-parallelepiped cells and integrator induced grid decoupled solutions.

Some solutions exist that minimize or eliminate the aforementioned shortcomings. For example, in the work of Gaitonde, Shang and Young [7], it was demonstrated that high-order, central implicit stencils in conjunction with spatial filters can significantly reduce dissipation effects, as compared with 3WB1 stencils. The filters serve two purposes: 1) To stabilize the boundary operators and 2) inhibit grid decoupled solutions. (It is not clear at this time, however, if decoupled solutions due to the RK4 integrator are eliminated with these same filters.) Unfortunately, these filters increase computational overhead and some question remains whether the efficiency of the final scheme has been compromised.

For curvilinear grids, it appears at this time that second-order metric computations will not suffice. Although, a high-order metric computation is admissible in theory, the practical implications of such a computation still needs to be studied.

APPENDIX

Consider some arbitrary surface $S(\xi)$ and its corresponding directed area $\mathbf{A}(\xi)$, which is defined as

$$\mathbf{A}(\xi) = \int_{S(\xi)} \nabla\xi J d\eta d\zeta \tag{56}$$

Decomposing ξ into its linear and non-linear parts ξ_l and ξ_n, respectively, assuming that the non-linear part is negligible, and recognizing that $\nabla\xi_l$ is constant, we obtain

$$\mathbf{A}(\xi) \approx \nabla\xi_l \int_{S(\xi)} J d\eta d\zeta \tag{57}$$

The integration of this area from ξ_1 to ξ_2 yields,

$$\int_{\xi_1}^{\xi_2} \mathbf{A}(\xi) d\xi \approx \int_{\xi_1}^{\xi_2} \nabla\xi_l \int_{S(\xi)} J d\eta d\zeta d\xi = V \nabla\xi_l. \tag{58}$$

Hence,

$$\nabla\xi_l \approx \frac{\delta_\xi \mathbf{A}_\xi^a}{V} \tag{59}$$

where

$$\mathbf{A}_{\xi}^{a} = \frac{1}{\delta_{\xi}} \int_{\xi_1}^{\xi_2} \mathbf{A}(\xi) d\xi. \tag{60}$$

Here $\mathbf{A}_{\xi}^{a}$ is the average directed surface area in the direction of ξ within the volume V. We close by noting that if the grid is a stretched Cartesian grid, the stretching is automatically incorporated into the definition of the area's average value.

ACKNOWLEDGMENT

This work was funded by the United States Department of Defense Experimental Program to Stimulate Competitive Research, grant number F49620-96-1-0469. The authors are grateful to Dr. Arje Nachman of the Air Force Office of Scientific Research and Dr. Joseph S. Shang of the Air Force Research Laboratory for their technical support and encouragement of this research effort. Supercomputer resources were provided by the Department of Defense's NAVOCEANO Major Share Resource Center.

REFERENCES

1. Shang, J. S., and R. M. Fithen, "A comparative study of characteristic-based algorithms for the Maxwell equations," *J. Comp. Phys.*, Vol. 125, 378–394, 1996.
2. Palaniswamy, S., W. F. Hall, and V. Shankar, "Numerical solution to Maxwell's equations in the time domain on nonuniform grids," *Radio Sci.,* Vol. 31, No. 4, 1996.
3. Madsen, N. K., and R. W. Ziolkowski, "A three-dimensional modified finite volume technique for Maxwell's equations," *Electromagnetics*, Vol. 10, 147–161, 1990.
4. Anderson, W. K., J. L. Thomas, and B. van Leer, "A comparison of finite volume flux vector splittings for the Euler equations," *AIAA 23rd Aerospace Sciences Meeting*, AIAA-85-0122, Reno, NV, Jan. 1985.
5. Richtmyer, R., and K. Morton, *Difference Methods for Initial-Value Problems,* Wiley, New York, 1967.
6. Harrington, R. F., *Time-Harmonic Electromagnetic Fields*, McGraw-Hill, New York, 1961.
7. Gaitonde, D., J. Shang, and J. L. Young, "Practical aspects of higher-order numerical schemes for wave propagation phenomena," *Intl. J. Num. Methods in Engr.*, Vol. 45, 1849–1869.

8. Anderson, D. A., J. C. Tannehill, and R. H. Pletcher, *Computational Fluid and Mechanics and Heat Transfer*, Taylor & Francis, Bristol, PA, 1984.
9. Ray, S. L., "Grid decoupling in finite element solutions of Maxwell's equations," *IEEE Trans. Ant. Propagat.,* Vol. 40, No. 4, 1992.
10. Shang, J. S., and D. Gaitonde, "Characteristic-based, time-dependent maxwell equation solvers on a general curvilinear frame," *AIAA Journal*, Vol. 33, No. 3, 491–498, 1995.
11. Al-Khafaji, A. W., *Numerical Methods in Engineering Practice*, Holt, Rinehart and Winstron, New York, 1986.
12. Gaitonde, D., and J. S. Shang, "Optimized compact-difference-based finite-volume schemes for linear wave phenomena," *J. Comp. Phys.*, Vol. 138, 617–643, 1997.
13. Gottlieb, D., and B. Yang, "Comparisons of staggered and non-staggered schemes for Maxwell's equations," *12 Annual Rev. of Progress in Appl. Comp. Electromagn.,* 1122–1131, Monterrey, CA, 1996.
14. Thomas, L. H., "Elliptic problems in linear difference equations over a network," *Watson Sci. Comput. Lab. Rept.*, Columbia University, New York, 1949.
15. Gustafsson, B., "The convergence rate for difference approximations to mixed initial boundary value problems," *Math. Comp.*, Vol. 29, 396–406, 1975.
16. Zhao, L., and A. C. Cangellaris, "GT-PML: Generalized theory of perfectly matched layers and its application to the reflectionless truncation of finite-difference time-domain grids," *IEEE Trans. Microwave Theory Tech.,* Vol. 44, No. 12, 1996.
17. Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes*, Cambridge University Press, Cambridge, 1986.
18. Fyfe, D. J., "Economical evaluation of Runge-Kutta formulae, *Math. Comput.,* Vol. 20, 392–398, 1966.
19. Williamson, J. H., "Low-storage Runge-Kutta schemes," *J. Comp. Physics.,* Vol. 35, 48–56, 1980.
20. Shang, J., "A perspective of computational electromagnetics in the time domain," *28th AIAA Plasmadynamics & Lasers Conference*, AIAA 97-2356, Atlanta, GA, June 1997.