Journal of Computational Physics **160,** 67–87 (2000)
doi:10.1006/jcph.2000.6421, available online at http://www.idealibrary.com on IDEAL®

Strict Stability of High-Order Compact Implicit Finite-Difference Schemes: The Role of Boundary Conditions for Hyperbolic PDEs, II

Saul S. Abarbanel,*,† Alina E. Chertock,*,‡ and Amir Yefet§

*Department of Applied Mathematics, School of Mathematical Sciences, Tel-Aviv University, Tel-Aviv, Israel; and §Department of Mathematical Sciences, New Jersey Institute of Technology, University Heights, Newark, New Jersey

E-mail: †saul@math.tau.ac.il, ‡cheral@math.tau.ac.il and alina@math.lbl.gov, and §yefet@atlas.njit.edu

Received March 2, 1999; revised November 30, 1999

This paper deals with the problem of systems of hyperbolic PDEs in one and two space dimensions, using the theory of part I [7]. © 2000 Academic Press

Key Words: hyperbolic PDEs; boundary conditions; stability; accuracy; error bounds.

1. INTRODUCTION

In this Part II of this series we shall continue the discussion about numerical methods for hyperbolic initial boundary value problems (IBVPs). This part is devoted to solving one- and two-dimensional hyperbolic *systems*. In Section 2, the theory and methodology presented in Part I [7] are modified to accommodate partially reflecting or absorbing boundary conditions and to solve the one-dimensional hyperbolic system. As was mentioned in Part I, time stability in the scalar case does not imply time stability for systems; see [1, 2]. Despite the fact that for hyperbolic systems we succeeded in proving the time stability only for some special cases, numerical examples show that the method is effective and provide time stability even when a theoretical foundation is lacking. As in the scalar case, the fourth- and sixth-order schemes are used for solving model problems. The formal accuracy of each scheme is determined by doing a grid refinement study. The numerical results show that the convergence rate of the schemes used here agrees well with theory. In order to investigate numerically whether the schemes are time stable we compute the error for long time integrations and additionally determine the eigenvalue spectrum of the semidiscrete system. In all cases, no eigenvalues with a positive real part are found which indicate the time stability of the schemes.

0021-9991/00 $35.00

As an application where high-order accurate approximations are needed we consider in Section 3 the two-dimensional Maxwell's equations in free space. The SAT method used for the diagonalized system in 1-D is adopted to solve the two-dimensional system, which cannot be digonalized. The problem is solved using both the fourth- and the sixth-order schemes. Numerical results are compared with those obtained by Turkel and Yefet in [5, 6]. They solved the same problem by using the Ty(2,4) scheme, which is a fourth-order compact implicit difference scheme on staggered meshes.

Section 2 of Part II was written by the authors of Part I. Section 3 of Part II was written in collaboration with A. Yefet.

2. 1-D HYPERBOLIC SYSTEMS

2.1. *General Theory and Description of the Method*

Consider a first-order hyperbolic system of partial differential equations

$$\frac{\partial \mathbf{u}}{\partial t} + A\frac{\partial \mathbf{u}}{\partial x} = 0, \qquad 0 \le x \le 1,\ t \ge 0, \tag{2.1}$$

where without loss of generality $\mathbf{u}(x,t) = (u^1(x,t), \dots, u^r(x,t))^T$ and A is a diagonal matrix with constant entries:

$$A = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_k & & & \\ & & & \lambda_{k+1} & & \\ & & & & \ddots & \\ & & & & & \lambda_r \end{pmatrix}, \qquad \begin{aligned} &\lambda_1 > \lambda_2 > \cdots, \lambda_k > 0, \\ &\lambda_r < \cdots < \lambda_{k+2} < \lambda_{k+1} < 0. \end{aligned}$$

The solution of (2.1) is uniquely determined if we prescribe initial values

$$\mathbf{u}(x,0) = \mathbf{f}(x), \qquad 0 \le x \le 1, \tag{2.2}$$

and boundary conditions

$$\begin{aligned} \mathbf{u}^{\mathrm{I}}(0,t) &= L\mathbf{u}^{\mathrm{II}}(0,t) + g^{\mathrm{I}}(t) \\ \mathbf{u}^{\mathrm{II}}(1,t) &= R\mathbf{u}^{\mathrm{I}}(1,t) + g^{\mathrm{II}}(t), \qquad t \ge 0, \end{aligned} \tag{2.3}$$

where L and R are fixed matrices of orders $k \times (r-k)$ and $(r-k) \times k$, respectively, $g^{\mathrm{I}}(t)$ is a given k-vector, $g^{\mathrm{II}}(t)$ is a given $(r-k)$-vector, and

$$\mathbf{u}^{\mathrm{I}} = (u^1, \dots, u^k)^T, \qquad \mathbf{u}^{\mathrm{II}} = (u^{k+1}, \dots, u^r)^T \tag{2.4}$$

is a partition of $\mathbf{u}$ into its outflow and inflow components, respectively, corresponding to the partition of A.

It is well known that (2.3) is well posed for any L and R, but in order to assure that the solution of (2.1) is bounded in time (when $g^{\mathrm{I}}(t)$ and $g^{\mathrm{II}}(t)$ are bounded in time), it is

sufficient to assume that

$$\|L\| \cdot \|R\| \le 1, \tag{2.5}$$

where the nonsquare matrix norm is defined by

$$\|L\| = \rho(L^T L)^{1/2} \tag{2.6}$$

and $\rho(L^T L)$ is the spectral radius of $L^T L$.

In order to solve the initial-boundary value problem (2.1) by a finite-difference approximation, we introduce, as in the scalar case, a mesh size h and denote by $\mathbf{u}^i = (u_0^i, u_1^i, \dots, u_N^i)^T$, $i = 1, \dots, r$, vectors of unknowns corresponding to the grid points $x_0, \dots, x_N$ ($N = 1/h$) and by $\mathbf{v}^i$ the numerical approximation to $\mathbf{u}^i$. Assuming that we have the same matrices P, Q, $\tilde{P}$, $\tilde{Q}$ and the the same vectors $\vec{S}_0$, $\vec{S}_N$ as in the scalar case—see Part I—we approximate the (2.1) by the scheme

$$\begin{aligned} P\frac{d\mathbf{v}^i}{dt} &= -\lambda_i Q\mathbf{v}^i + \lambda_i \vec{S}_0\big(\mathbf{v}_0^i - (L\mathbf{v}^{\mathrm{II}} + g^{\mathrm{I}})_0^i\big), \qquad 1 \le i \le k \\ \tilde{P}\frac{d\mathbf{v}^i}{dt} &= -\lambda_i \tilde{Q}\mathbf{v}^i + \lambda_i \vec{S}_N\big(\mathbf{v}_N^i - (R\mathbf{v}^{\mathrm{I}} + g^{\mathrm{II}})_N^i\big), \qquad k+1 \le i \le r. \end{aligned} \tag{2.7}$$

To prove the convergence of the scheme (2.7) we will derive an equation for the error function $\mathcal{E}$ and show that its discrete norm (to be defined later) is bounded by a function $F(t, h, u)$, where t, h, and u are the time, the mesh size, and the exact solution, respectively. It will be shown that $F(t, h, u)$ is bounded in time by a linear growth and tends to zero with mesh refinement.

Since $\mathbf{u}_0^i - (L\mathbf{u}^{\mathrm{II}} + g^{\mathrm{I}})_0^i = 0$ for $1 \le i \le k$ and $\mathbf{u}_N^i - (Ru^{\mathrm{I}} + g^{\mathrm{II}})_N^i = 0$ for $k+1 \le i \le r$, we may write for the vectors $\mathbf{u}^i$

$$\begin{aligned} P\frac{d\mathbf{u}^i}{dt} &= -\lambda_i Q\mathbf{u}^i + \lambda_i \vec{S}_0\big(\mathbf{u}_0^i - (L\mathbf{u}^{\mathrm{II}} + g^{\mathrm{I}})_0^i\big) + P\mathbf{T}^i, \qquad 1 \le i \le k \\ \tilde{P}\frac{d\mathbf{u}^i}{dt} &= -\lambda_i \tilde{Q}\mathbf{u}^i + \lambda_i \vec{S}_N\big(\mathbf{u}_N^i - (R\mathbf{u}^{\mathrm{I}} + g^{\mathrm{II}})_N^i\big) + \tilde{P}\mathbf{T}^i, \qquad k+1 \le i \le r, \end{aligned} \tag{2.8}$$

where $\mathbf{T} = (\mathbf{T^0}, \dots, \mathbf{T^k}, \mathbf{T^{k+1}}, \dots, \mathbf{T^r})$ is the $r \times N$ long vector of the truncation errors due to numerical differencing.

Denote by $\varepsilon^i = \mathbf{u}^i - \mathbf{v}^i$ $(1 \le i \le r)$ the solution error vectors and subtract (2.7) from (2.8) to get

$$\begin{aligned} P\frac{d\varepsilon^i}{dt} &= -\lambda_i Q\varepsilon^i + \lambda_i \vec{S}_0\big(\varepsilon_0^i - (L\varepsilon^{\mathrm{II}})_0^i\big) + P\mathbf{T}^i, \qquad 1 \le i \le k \\ \tilde{P}\frac{d\varepsilon^i}{dt} &= -\lambda_i \tilde{Q}\varepsilon^i + \lambda_i \vec{S}_N\big(\varepsilon_N^i - (R\varepsilon^{\mathrm{I}})_N^i\big) + \tilde{P}\mathbf{T}^i, \qquad k+1 \le i \le r. \end{aligned} \tag{2.9}$$

We define now the scalar product

$$(\varepsilon^i, \varepsilon^j) = \sum_{m=0}^{N} \varepsilon_m^i \varepsilon_m^j \tag{2.10}$$

and the discrete norms

$$\|\varepsilon^{\mathrm{I}}\|^2 = \sum_{i=1}^{k} \frac{\|R\|}{\lambda_i}(\varepsilon^i, \varepsilon^i), \qquad \|\varepsilon^{\mathrm{II}}\|^2 = \sum_{i=k+1}^{r} \frac{\|L\|}{|\lambda_i|}(\varepsilon^i, \varepsilon^i) \tag{2.11}$$

and

$$\|\mathcal{E}\|^2 = \|\varepsilon^{\mathrm{I}}\|^2 + \|\varepsilon^{\mathrm{II}}\|^2, \tag{2.12}$$

where $\mathcal{E}$ is the $r \times N$ long error vector whose first $k \times N$ entries are the entires of ε^{I} and the other $(r-k) \times N$ entries are the ones of $\varepsilon^{\mathrm{II}}$.

Differentiating the scalar products $(P\varepsilon^i, \varepsilon^i)$ and $(\tilde{P}\varepsilon^i, \varepsilon^i)$ and using Eq. (2.9) yields

$$\begin{aligned}
\frac{d}{dt}(P\varepsilon^i, \varepsilon^i) &= -\lambda_i(Q\varepsilon^i, \varepsilon^i) + \lambda_i(\vec{S}_0, \varepsilon^i)\big(\varepsilon_0^i - (L\varepsilon^{\mathrm{II}})_0^i\big) + (P\mathbf{T}^i, \varepsilon^i), \qquad 1 \le i \le k \\
\frac{d}{dt}(\tilde{P}\varepsilon^i, \varepsilon^i) &= -\lambda_i(\tilde{Q}\varepsilon^i, \varepsilon^i) + \lambda_i(\vec{S}_N, \varepsilon^i)\big(\varepsilon_N^i - (R\varepsilon^{\mathrm{I}})_N^i\big) + (\tilde{P}\mathbf{T}^i, \varepsilon^i), \qquad k+1 \le i \le r.
\end{aligned} \tag{2.13}$$

We now use the definitions of $\vec{S}_0$ and $\vec{S}_N$, the properties of Q and $\tilde{Q}$ (from assumption 3 and remarks from Part I), and the fact that the λ_i are positive for $1 \le i \le k$ and negative for $k+1 \le i \le r$ to get

$$\begin{aligned}
\frac{d}{dt}(P\varepsilon^i, \varepsilon^i) &= \lambda_i(\tau - 1)q_{00}\big(\varepsilon_0^i\big)^2 - \lambda_i q_{11}\big(\varepsilon_1^i\big)^2 - \lambda_i \tau q_{00}(L\varepsilon^{\mathrm{II}})_0^i \varepsilon_0^i \\
&\quad - \lambda_i(q_{01} + q_{10})\varepsilon_1^i (L\varepsilon^{\mathrm{II}})_0^i - \lambda_i\big[q_{NN}\big(\varepsilon_N^i\big)^2 + (q_{N-1N} + q_{NN-1})\varepsilon_{N-1}^i \varepsilon_N^i \\
&\quad + q_{N-1N-1}\big(\varepsilon_{N-1}^i\big)^2\big] + (P\mathbf{T}^i, \varepsilon^i), \qquad 1 \le i \le k \\
\frac{d}{dt}(\tilde{P}\varepsilon^i, \varepsilon^i) &= |\lambda_i|(\tau - 1)q_{00}\big(\varepsilon_N^i\big)^2 - |\lambda_i| q_{11}\big(\varepsilon_{N-1}^i\big)^2 - |\lambda_i| \tau q_{00}(R\varepsilon^{\mathrm{I}})_N^i \varepsilon_N^i \\
&\quad - |\lambda_i|(q_{01} + q_{10})\varepsilon_{N-1}^i (R\varepsilon^{\mathrm{I}})_N^i - |\lambda_i|\big[q_{NN}\big(\varepsilon_0^i\big)^2 + (q_{N-1N} + q_{NN-1})\varepsilon_0^i \varepsilon_1^i \\
&\quad + q_{N-1N-1}\big(\varepsilon_1^i\big)^2\big] + (\tilde{P}\mathbf{T}^i, \varepsilon^i), \qquad k+1 \le i \le r.
\end{aligned} \tag{2.14}$$

We multiply the first equation of (2.14) by $\|R\|/\lambda_i$ and sum up from $i = 0$ to k, and we multiply the second equation by $\|L\|/|\lambda_i|$ and sum up from $i = k+1$ to r. We then add these two sums and, assuming that q_{N-1N-1} is positive, the resulting expression may be written thusly:

$$\begin{aligned}
&\frac{d}{dt}\sum_{i=0}^{k} \frac{\|R\|}{\lambda_i}(P\varepsilon^i, \varepsilon^i) + \frac{d}{dt}\sum_{i=k+1}^{r} \frac{\|L\|}{|\lambda_i|}(\tilde{P}\varepsilon^i, \varepsilon^i) \\
&= \sum_{i=0}^{k} \Bigg[\|R\|(\tau - 1)q_{00}\big(\varepsilon_0^i\big)^2 - \|R\| q_{11}\big(\varepsilon_1^i\big)^2 - \|R\| \tau q_{00}(L\varepsilon^{\mathrm{II}})_0^i \varepsilon_0^i \\
&\quad - \|R\|(q_{01} + q_{10})\varepsilon_1^i (L\varepsilon^{\mathrm{II}})_0^i - \|R\| \left(\frac{q_{N-1N} + q_{NN-1}}{2\sqrt{q_{N-1N-1}}} \varepsilon_N^i + \sqrt{q_{N-1N-1}}\, \varepsilon_{N-1}^i \right)^2
\end{aligned}$$

$$
\begin{aligned}
&- \|R\|\left(q_{NN} - \frac{(q_{N-1N} + q_{NN-1})^2}{4q_{N-1N-1}}\right)\left(\varepsilon_N^i\right)^2\Bigg] + \sum_{i=0}^{k}\frac{\|R\|}{\lambda_i}(P\mathbf{T}^i, \varepsilon^i)\\
&+ \sum_{i=k+1}^{r}\Bigg[\|L\|(\tau-1)q_{00}\left(\varepsilon_N^i\right)^2 - \|L\|q_{11}\left(\varepsilon_{N-1}^i\right)^2 - \|L\|\tau q_{00}(R\varepsilon^{\mathrm{I}})_N^i\varepsilon_N^i\\
&- \|L\|(q_{01}+q_{10})\varepsilon_{N-1}^i(R\varepsilon^{\mathrm{I}})_N^i - \|L\|\left(\frac{q_{N-1N}+q_{NN-1}}{2\sqrt{q_{N-1N-1}}}\varepsilon_0^i + \sqrt{q_{N-1N-1}}\varepsilon_1^i\right)^2\\
&- \|L\|\left(q_{NN} - \frac{(q_{N-1N}+q_{NN-1})^2}{4q_{N-1N-1}}\right)\left(\varepsilon_0^i\right)^2\Bigg] + \sum_{i=k+1}^{r}\frac{\|L\|}{|\lambda_i|}(\tilde{P}\mathbf{T}^i, \varepsilon^i).
\end{aligned}
$$

Again, as in Part I, we require the expression $q_{NN}\varepsilon_0^2 + (q_{N-1N} + q_{NN-1})\varepsilon_0\varepsilon_1 + q_{NN}\varepsilon_1^2$ to be positive for all $\varepsilon_0, \varepsilon_1 \in \mathbf{R}$. This implies

$$
q_{N-1N-1} > 0, \qquad q_{NN} - \frac{(q_{N-1N}+q_{NN-1})^2}{4q_{N-1N-1}} > 0. \tag{2.15}
$$

We next define new discrete scalar products (note the difference from (2.10):

$$
\begin{aligned}
[\varepsilon^{\mathrm{I}}, \varepsilon^{\mathrm{I}}]_m &= \sum_{i=1}^{k}\varepsilon_m^i\varepsilon_m^i\\
[\varepsilon^{\mathrm{II}}, \varepsilon^{\mathrm{II}}]_m &= \sum_{i=k+1}^{r}\varepsilon_m^i\varepsilon_m^i.
\end{aligned} \tag{2.16}
$$

Replacing the sums in the last equation with these vector operations and using the properties of the matrices P and $\tilde{P}$ we get an estimate for the discrete norm $\|\mathcal{E}\|$,

$$
\begin{aligned}
\frac{1}{2}c_0\frac{d}{dt}\|\mathcal{E}\|^2 \le{}& (\tau-1)q_{00}\|R\|[\varepsilon^{\mathrm{I}}, \varepsilon^{\mathrm{I}}]_0 - \|R\|q_{11}[\varepsilon^{\mathrm{I}}, \varepsilon^{\mathrm{I}}]_1 - \|R\|\tau q_{00}[L\varepsilon^{\mathrm{II}}, \varepsilon^{\mathrm{I}}]_0\\
&- \beta\|R\|[\varepsilon^{\mathrm{I}}, \varepsilon^{\mathrm{I}}]_N + (\tau-1)q_{00}\|L\|[\varepsilon^{\mathrm{II}}, \varepsilon^{\mathrm{II}}]_N - \|L\|q_{11}[\varepsilon^{\mathrm{II}}, \varepsilon^{\mathrm{II}}]_{N-1}\\
&- \|L\|\tau q_{00}[R\varepsilon^{\mathrm{I}}, \varepsilon^{\mathrm{II}}]_N - \beta\|L\|[\varepsilon^{\mathrm{II}}, \varepsilon^{\mathrm{II}}]_0 - 2\|R\|\mathbf{q}_{01}\sum_{i=1}^{k}(\varepsilon^{\mathrm{I}})_1^i(L\varepsilon^{\mathrm{II}})_0^i\\
&- 2\|L\|\mathbf{q}_{01}\sum_{i=k+1}^{r}(\varepsilon^{\mathrm{II}})_{N-1}^i(R\varepsilon^{\mathrm{I}})_N^i + \sum_{i=0}^{k}\frac{\|R\|}{\lambda_i}(P\mathbf{T}^i, \varepsilon^i) + \sum_{i=k+1}^{r}\frac{\|L\|}{|\lambda_i|}(\tilde{P}\mathbf{T}^i, \varepsilon^i),
\end{aligned}
$$

where

$$
\mathbf{q}_{01} = \frac{1}{2}(q_{01}+q_{10}), \qquad \beta = q_{NN} - \frac{(q_{N-1N}+q_{NN-1})^2}{4q_{N-1N-1}} > 0.
$$

Substituting the estimates

$$
\begin{aligned}
[L\varepsilon^{\mathrm{II}}, \varepsilon^{\mathrm{I}}]_0 &\le \|L\| \cdot \|\varepsilon^{\mathrm{II}}\|_0 \cdot \|\varepsilon^{\mathrm{I}}\|_0,\\
[R\varepsilon^{\mathrm{I}}, \varepsilon^{\mathrm{II}}]_N &\le \|R\| \cdot \|\varepsilon^{\mathrm{I}}\|_N \cdot \|\varepsilon^{\mathrm{II}}\|_N,
\end{aligned}
$$

$$\sum_{i=1}^{k}(\varepsilon^{\mathrm{I}})_1^i(L\varepsilon^{\mathrm{II}})_0^i \le \sqrt{\sum_{i=1}^{k}\left[(\varepsilon^{\mathrm{I}})_1^i\right]^2\cdot\sum_{i=1}^{k}\left[(L\varepsilon^{\mathrm{II}})_0^i\right]^2} \le \|L\|\cdot\|\varepsilon^{\mathrm{I}}\|_1\cdot\|\varepsilon^{\mathrm{II}}\|_0,$$

$$\sum_{i=k+1}^{r}(\varepsilon^{\mathrm{II}})_{N-1}^i(R\varepsilon^{\mathrm{I}})_N^i \le \sqrt{\sum_{i=k+1}^{r}\left[(\varepsilon^{\mathrm{II}})_{N-1}^i\right]^2\cdot\sum_{i=k+1}^{r}\left[(R\varepsilon^{\mathrm{I}})_N^i\right]^2} \le \|R\|\cdot\|\varepsilon^{\mathrm{I}}\|_N\cdot\|\varepsilon^{\mathrm{II}}\|_{N-1}$$

where

$$\|\varepsilon^{\mathrm{I}}\|_m = \sqrt{[\varepsilon^{\mathrm{I}},\varepsilon^{\mathrm{I}}]_m},$$
$$\|\varepsilon^{\mathrm{II}}\|_m = \sqrt{[\varepsilon^{\mathrm{II}},\varepsilon^{\mathrm{II}}]_m}, \qquad m = 0,1,N-1,N,$$

into the last inequality for $\|\mathcal{E}\|$ and collecting like terms yields

$$\begin{aligned}\frac{1}{2}c_0\frac{d}{dt}\|\mathcal{E}\|^2 \le{}& \big\{(\tau-1)q_{00}\cdot\|R\|\cdot\|\varepsilon^{\mathrm{I}}\|_0^2 - \|R\|q_{11}\|\varepsilon^{\mathrm{I}}\|_1^2\\ &+|\tau q_{00}|\cdot\|R\|\cdot\|L\|\cdot\|\varepsilon^{\mathrm{I}}\|_0\cdot\|\varepsilon^{\mathrm{II}}\|_0 + 2|\mathbf{q}_{01}|\cdot\|R\|\cdot\|L\|\cdot\|\varepsilon^{\mathrm{I}}\|_1\cdot\|\varepsilon^{\mathrm{II}}\|_0\\ &-\beta\|L\|\cdot\|\varepsilon^{\mathrm{II}}\|_0^2\big\} + \big\{(\tau-1)q_{00}\cdot\|R\|\cdot\|\varepsilon^{\mathrm{II}}\|_N^2 - \|R\|q_{11}\|\varepsilon^{\mathrm{I}}\|_{N-1}^2\\ &+|\tau q_{00}|\cdot\|R\|\cdot\|L\|\cdot\|\varepsilon^{\mathrm{I}}\|_N\cdot\|\varepsilon^{\mathrm{II}}\|_N + 2|\mathbf{q}_{01}|\cdot\|R\|\cdot\|L\|\cdot\|\varepsilon^{\mathrm{I}}\|_N\cdot\|\varepsilon^{\mathrm{II}}\|_{N-1}\\ &-\beta\|L\|\cdot\|\varepsilon^{\mathrm{I}}\|_N^2\big\} + \sum_{i=0}^{k}\frac{\|R\|}{\lambda_i}(P\mathbf{T}^i,\varepsilon^i) + \sum_{i=k+1}^{r}\frac{\|L\|}{|\lambda_i|}(\tilde{P}\mathbf{T}^i,\varepsilon^i). \end{aligned}\tag{2.17}$$

We require now each curly bracket to be nonpositive. Thus we need

$$\begin{aligned}&(\tau-1)q_{00}\cdot\|R\|\cdot\|\varepsilon^{\mathrm{I}}\|_0^2 - \|R\|q_{11}\|\varepsilon^{\mathrm{I}}\|_1^2 + |\tau q_{00}|\cdot\|R\|\cdot\|L\|\cdot\|\varepsilon^{\mathrm{I}}\|_0\cdot\|\varepsilon^{\mathrm{II}}\|_0\\ &\quad+2|\mathbf{q}_{01}|\cdot\|R\|\cdot\|L\|\cdot\|\varepsilon^{\mathrm{I}}\|_1\cdot\|\varepsilon^{\mathrm{II}}\|_0 - \beta\|L\|\cdot\|\varepsilon^{\mathrm{II}}\|_0^2 \le 0\end{aligned}\tag{2.18}$$

and also

$$\begin{aligned}&(\tau-1)q_{00}\cdot\|R\|\cdot\|\varepsilon^{\mathrm{II}}\|_N^2 - \|R\|q_{11}\|\varepsilon^{\mathrm{I}}\|_{N-1}^2 + |\tau q_{00}|\cdot\|R\|\cdot\|L\|\cdot\|\varepsilon^{\mathrm{I}}\|_N\cdot\|\varepsilon^{\mathrm{II}}\|_N\\ &\quad+2|\mathbf{q}_{01}|\cdot\|R\|\cdot\|L\|\cdot\|\varepsilon^{\mathrm{I}}\|_N\cdot\|\varepsilon^{\mathrm{II}}\|_{N-1} - \beta\|L\|\cdot\|\varepsilon^{\mathrm{I}}\|_N^2 \le 0\end{aligned}\tag{2.19}$$

for all $\varepsilon^{\mathrm{I}},\varepsilon^{\mathrm{II}}\in\mathbf{R}$.

It is possible to show that both inequalities are satisfied (and hence the algorithm is time stable) if

$$\begin{aligned}&q_{11}>0, \qquad q_{N-1N-1}>0, \qquad (\tau-1)q_{00}<0,\\ &\beta = q_{NN} - \frac{(q_{N-1N}+q_{NN-1})^2}{4q_{N-1N-1}} > 0,\\ &\frac{1}{4}\tau^2 q_{11}q_{00}^2\|R\|\cdot\|L\| + (\tau-1)q_{00}\big(\beta q_{11} - \mathbf{q}_{01}^2\|R\|\cdot\|L\|\big) < 0.\end{aligned}\tag{2.20}$$

Assuming for the moment that these inequalities hold we can write

$$\frac{1}{2}c_0\frac{d}{dt}\|\mathcal{E}\|^2 \le \sum_{i=0}^{k}\frac{\|R\|}{\lambda_i}(P\mathbf{T}^i,\varepsilon^i) + \sum_{i=k+1}^{r}\frac{\|L\|}{|\lambda_i|}(\tilde{P}\mathbf{T}^i,\varepsilon^i).\tag{2.21}$$

Using (2.9) from Part I and the definition (2.11) of the discrete norms we get

$$\frac{1}{2}c_0\frac{d}{dt}\|\mathcal{E}\|^2 \le c_1\|\mathbf{T}\|\|\mathcal{E}\| \tag{2.22}$$

and after dividing by $\|\mathcal{E}\|$,

$$\frac{d}{dt}\|\mathcal{E}\| \le \frac{c_1}{c_0}\|\mathbf{T}\|, \tag{2.23}$$

leading to

$$\|\mathcal{E}\| \le \frac{c_1}{c_0}\sup_{0\le\tau\le t}\|\mathbf{T}\|\,t. \tag{2.24}$$

We are ready now to formulate the theorem:

THEOREM 2.1. *Let the method defined by Eq.* (2.7) *satisfy* (2.20), *for the discretization of the hyperbolic system* (2.1) *with initial and boundary conditions* (2.2), (2.3). *Then it is stable and leads to an error whose norm is growing linearly in time.*

Remark. We recall that in order to solve the hyperbolic system numerically we use the same matrices P, Q, $\tilde{P}$, $\tilde{Q}$ and the same vectors $\vec{S}_0$, $\vec{S}_N$ as in the scalar case, i.e.,

$$q_{00} = -\frac{2}{3}, \qquad q_{11} = \frac{1}{6} > 0, \qquad \mathbf{q}_{01} = \frac{1}{3},$$
$$\beta = q_{NN} - \frac{(q_{N-1N} + q_{NN-1})^2}{4q_{N-1N-1}} = \frac{1}{6} > 0.$$

With this choice of the matrix Q the inequalities (2.20) hold if

$$\frac{1 - \|R\|\cdot\|L\| - \sqrt{D}}{2\|R\|\cdot\|L\|} \le \tau \le \frac{1 - \|R\|\cdot\|L\| + \sqrt{D}}{2\|R\|\cdot\|L\|}, \tag{2.25}$$

where

$$D = (1 - \|R\|\cdot\|L\|)(1 - 5\|R\|\cdot\|L\|).$$

We can choose τ, which satisfies (2.25), if $D \ge 0$. This happens if $\|R\|\cdot\|L\| \le 1/5$. But this is only a sufficient condition, because numerical experiments (see the discussion in the next subsection) show that the numerical solution converges to the analytical solution for all $t < \infty$ even if $1/5 < \|R\|\cdot\|L\| \le 1$.

Similarly, in the case of the fourth-order scheme,

$$q_{00} = -\frac{5}{8}, \qquad q_{11} = \frac{1}{8} > 0, \qquad \mathbf{q}_{01} = \frac{1}{4},$$
$$\beta = q_{NN} - \frac{(q_{N-1N} + q_{NN-1})^2}{4q_{N-1N-1}} = \frac{1}{4} > 0,$$

leading to

$$\frac{4 - 2\|R\|\cdot\|L\| - 2\sqrt{D}}{5\|R\|\cdot\|L\|} \le \tau \le \frac{4 - 2\|R\|\cdot\|L\| + 2\sqrt{D}}{5\|R\|\cdot\|L\|}, \tag{2.26}$$

where

$$D = (2 - \|R\| \cdot \|L\|)(2 - 6\|R\| \cdot \|L\|).$$

We can find τ, which satisfies (2.26), if $\|R\| \cdot \|L\| \le 1/3$. Numerical experiments performed in the next section show that the fourth-order scheme is time stable even if $1/3 < \|R\| \cdot \|L\| \le 1$.

However, if $\|R\| \cdot \|L\| \le 1/5$ in the case of the sixth-order scheme (or $\|R\| \cdot \|L\| \le 1/3$ in the case of the fourth-order scheme) then (2.18) and (2.19) are strictly negative and (2.24) is replaced, just as in Part I, by a constant bound.

2.2. *Numerical Experiments*

Consider the hyperbolic system

$$\frac{\partial \mathbf{u}}{\partial t} + A\frac{\partial \mathbf{u}}{\partial x} = 0, \qquad 0 \le x \le 1,\ t \ge 0, \tag{2.27}$$

where

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \mathbf{u} = \begin{pmatrix} u \\ v \end{pmatrix}, \tag{2.28}$$

with initial data

$$u(x, 0) = \sin 2\pi x, \quad v(x, 0) = -\sin 2\pi x, \qquad 0 \le x \le 1, \tag{2.29}$$

and boundary conditions

$$u(0, t) = v(0, t), \quad v(1, t) = u(1, t), \qquad t \ge 0. \tag{2.30}$$

The exact solution is

$$\begin{aligned} u(x, t) &= \sin 2\pi(x - t), \\ v(x, t) &= -\sin 2\pi(x + t), \qquad 0 \le x \le 1,\ t \ge 0. \end{aligned} \tag{2.31}$$

Note that due to (2.30), $\|R\| \cdot \|L\| = 1$ and thus we test the most severe reflection case.

As in the scalar case of Part I we solve the problem (2.27)–(2.30) numerically using two different schemes: fourth-order compact with third-order boundary closure and sixth-order compact with fifth-order boundary closure. And again we compare two methods for implementation of the boundary conditions: (i) conventional, which implies the overwriting of the value of the solution at the boundary point with the analytic boundary condition after each Runge–Kutta stage, and (ii) the SAT method described in the previous subsection. In all cases, the standard fourth-order Runge–Kutta method is used for time integration, with a suitable Δt such that the desired overall accuracy is maintained.[1]

[1] We did not a use sixth-order Runge–Kutta integrator because we are not aware of any *stable* sixth-order Runge–Kutta method suitable for a system of ODEs. However, in the scalar case of Part I we did use a sixth-order Runge–Kutta method.

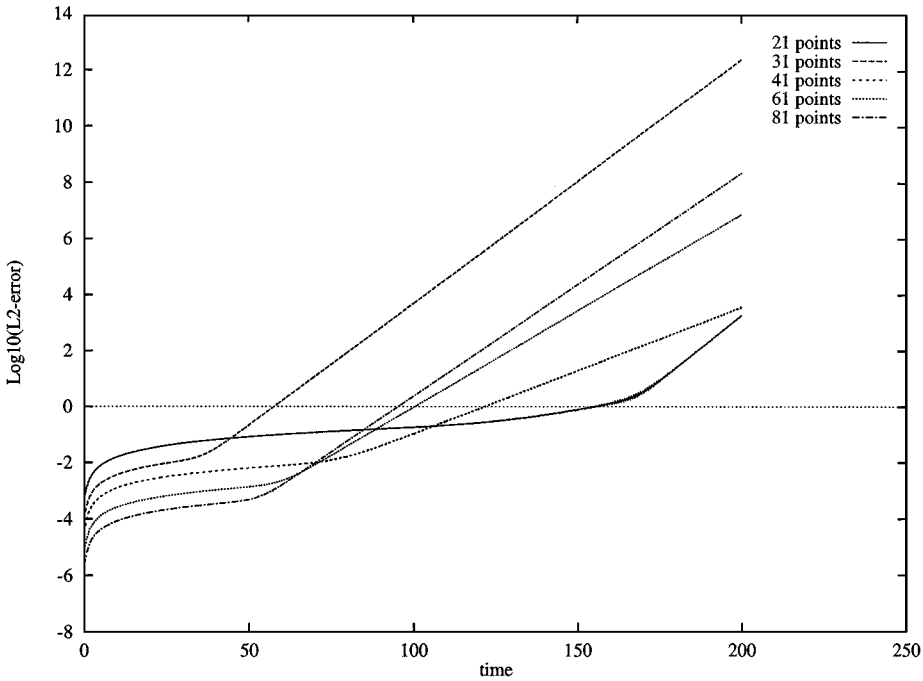

FIG. 1. The L_2-error as a function of time for the fourth-order approximation using conventional implementation of boundary conditions with CFL $= 0.5$.

Conventional boundary conditions. In Part I it was shown that for the scalar case the fourth-order scheme is time stable while the sixth-order scheme is not when using conventional implementation of boundary conditions. Using these schemes to solve the test problem (2.27)–(2.30) we found that neither scheme was time stable when applied to a system of equations. Figures 1 and 2 show L_2-error as a function of time for the fourth-order compact scheme and sixth-order compact scheme, respectively, for different grids. As one can see, results diverge exponentially from the analytic solution.

On the other hand, we shall show that SAT procedure ensures time stability (only a sublinear temporal growth) for the hyperbolic system, for both the fourth- and the sixth-order schemes.

SAT boundary conditions. First we verify that SAT implementation of boundary conditions retains the formal accuracy of the spatial operator. Results of the grid convergence study of the spatial operators with SAT parameter $\tau = 2$ for both orders of accuracy are presented in Table I. The entries are the absolute error $\log_{10}(L_2)$ at a fixed time $t = T$ and the convergence rate. The convergence rate is computed as

$$\log_{10}\left(\frac{\|\mathbf{u}-\mathbf{u}^{h_1}\|_2}{\|\mathbf{u}-\mathbf{u}^{h_2}\|_2}\right)\Big/\log_{10}\left(\frac{h_1}{h_2}\right), \tag{2.32}$$

where $\mathbf{u} = (\mathbf{u}(x_0, t), \mathbf{u}(x_2, t), \ldots, \mathbf{u}(x_N, t))^T$ is the exact solution, $\mathbf{u}^h$ is the numerical solution with mesh width h, and $\|\mathbf{u}-\mathbf{u}^h\|_2$ is the discrete L_2 norm of the absolute error. The data in this table indicate that the convergence rate asymptotically approaches the theoretical value of 4 for the fourth-order operator and 6 for the sixth-order operator. Figures 3 and 4 show the error as a function of time for long time integration using the fourth-order

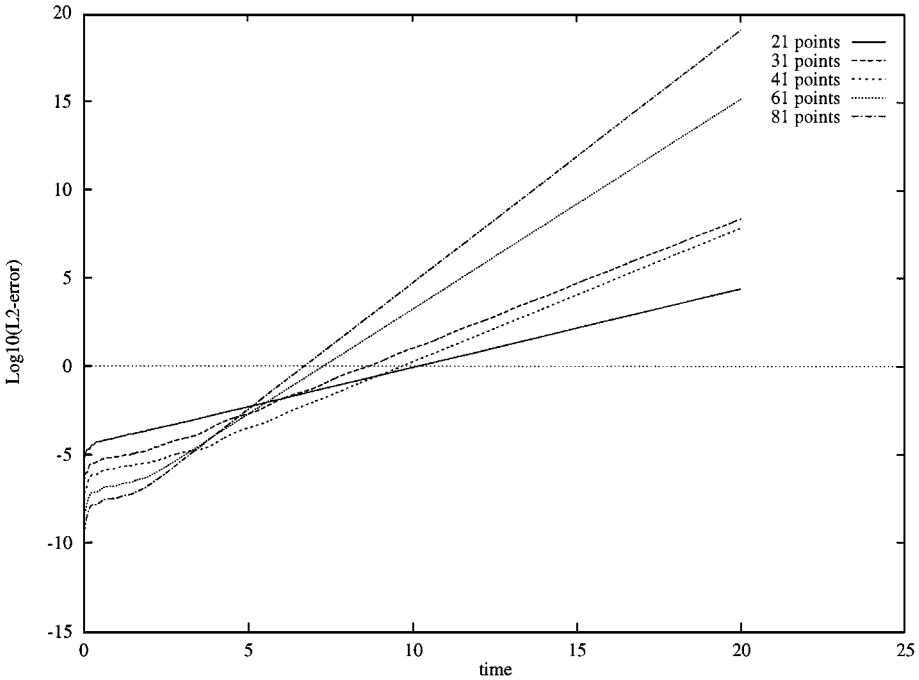

FIG. 2. The L_2-error as a function of time for the sixth-order approximation using conventional implementation of boundary conditions with CFL = 0.1.

and the sixth-order difference operators, respectively, for different grids. No exponential growth exists, and both schemes are found to be strictly stable. In Figs. 5 and 6 the eigenvalue spectrum for both schemes for different grids is shown. One can see that there are no eigenvalues with a positive real part.

3. 2-D HYPERBOLIC SYSTEMS

3.1. *Application to Maxwell's Equations*

As an application where high-order accurate approximation are needed we consider Maxwell's equations. In free space they are given by

TABLE I

Grid Convergence of Two High-Order Schemes for $u_t + Au_x = 0$, Using the SAT Implementation of Boundary Conditions with the SAT Parameter $\tau = 2$ and CFL = 0.5 for the Fourth-Order Scheme, CFL = 0.1 for the Sixth-Order Scheme ($T = 10$)

	Fourth-order compact		Sixth-order compact	
Grid	$\log_{10}(L_2)$	Rate	$\log_{10}(L_2)$	Rate
21	−2.657		−4.371	
31	−3.332	3.83	−5.462	6.19
41	−3.817	3.89	−6.231	6.15
61	−4.506	3.91	−7.299	6.07
81	−4.998	3.94	−8.041	5.97

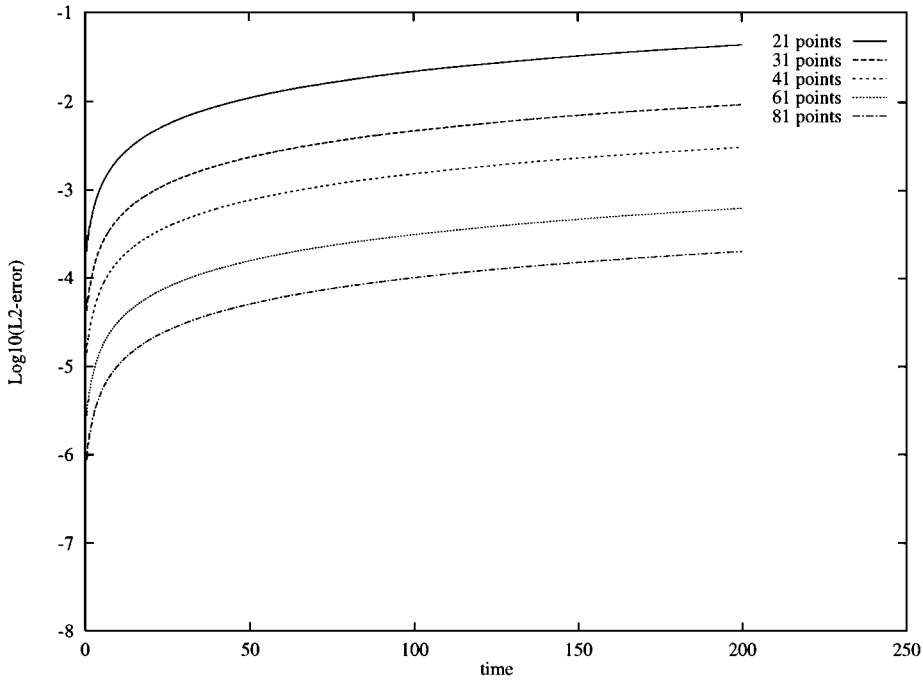

FIG. 3. The L_2-error as a function of time for the fourth-order approximation using SAT method for implementation of boundary conditions with $\tau = 2$, CFL $= 0.5$.

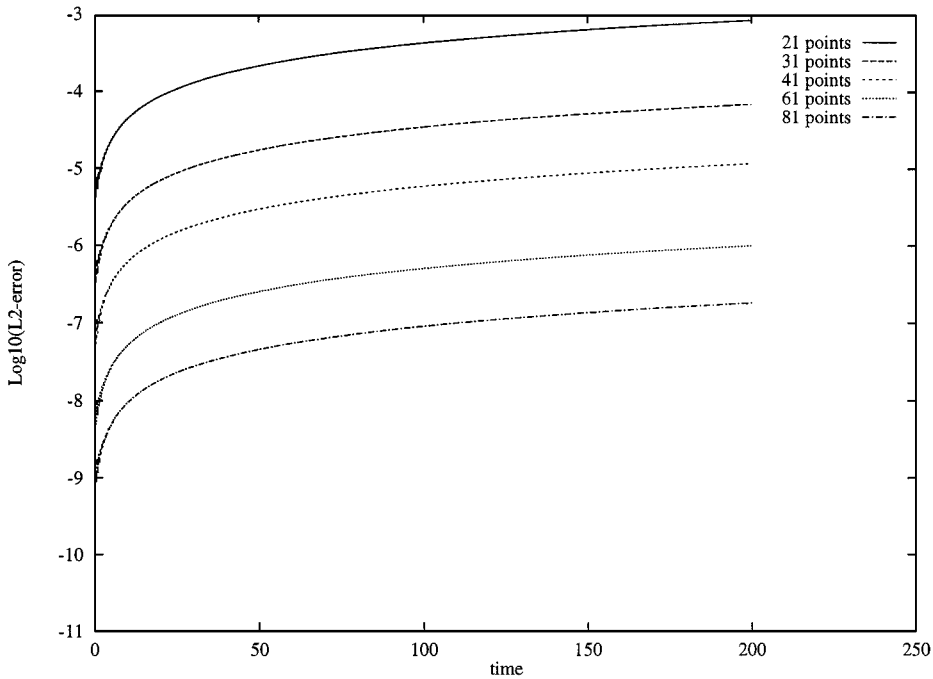

FIG. 4. The L_2-error as a function of time for the sixth-order approximation using SAT method for implementation of boundary conditions with $\tau = 2$, CFL $= 0.1$.

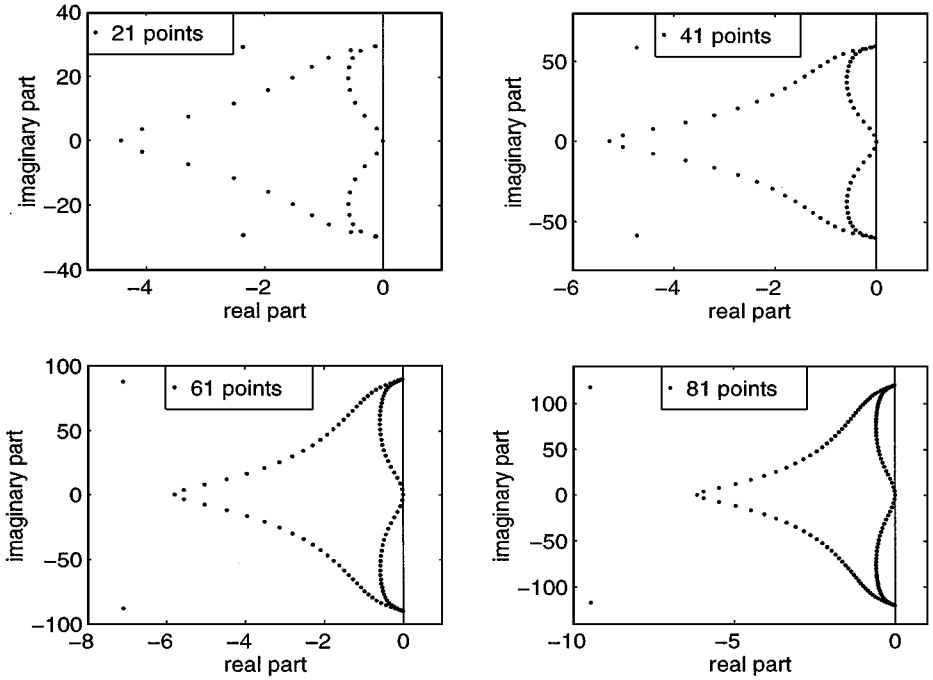

FIG. 5. Semidiscrete eigenvalue spectrum for the fourth-order approximation using SAT method for implementation of boundary conditions with $\tau = 2$.

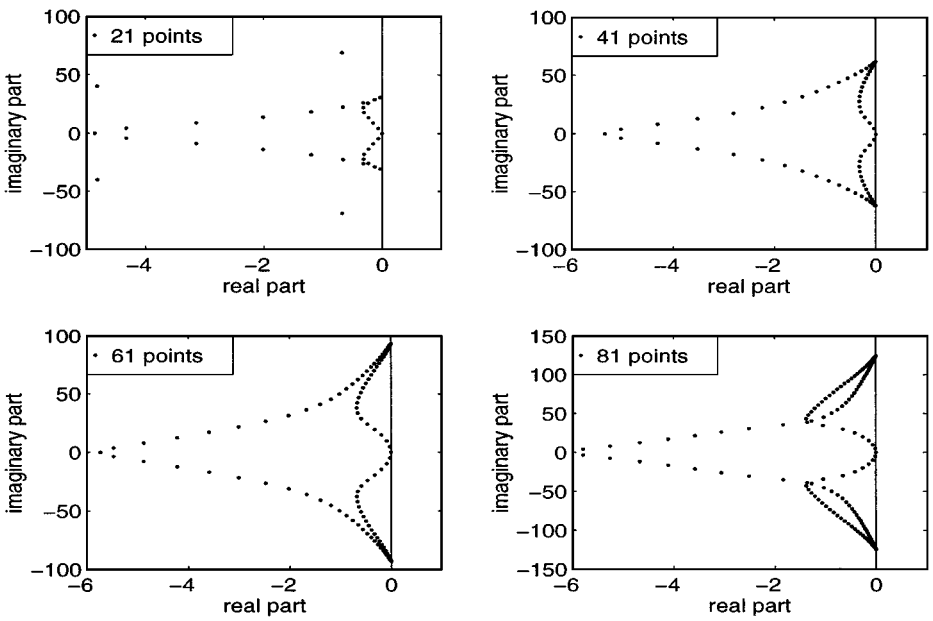

FIG. 6. Magnification of semidiscrete eigenvalue spectrum close to the imaginary axis for the sixth-order approximation using the SAT method for implementation of boundary conditions with $\tau = 2$.

$$\begin{aligned}
\frac{\partial B}{\partial \tilde{t}} + \nabla \times \mathrm{E} &= 0 \qquad \text{(Faraday's law)},\\
\frac{\partial D}{\partial \tilde{t}} - \nabla \times \mathrm{H} &= \mathrm{J} \qquad \text{(Ampere's law)},\\
\mathrm{B} &= \mu \mathrm{H},\\
\mathrm{D} &= \epsilon \mathrm{E},
\end{aligned} \tag{3.1}$$

coupled with Gauss's law

$$\begin{aligned}
\nabla \cdot \mathrm{B} &= 0,\\
\nabla \cdot \mathrm{D} &= 0.
\end{aligned} \tag{3.2}$$

If we assume perfectly conducting conditions at the outer edge of the domain then the boundary conditions are

$$\begin{aligned}
\vec{n} \times \mathrm{E} &= 0,\\
\vec{n} \cdot \mathrm{H} &= 0,
\end{aligned} \tag{3.3}$$

where $\vec{n}$ is a normal vector to the surface of the domain.

To simplify the notation we shall consider the two dimensional case with ϵ, μ constants and $J = 0$. We nondimensionalize the variables, $t = c\tilde{t}/L$, $x = \tilde{x}/L$, $y = \tilde{y}/L$, $E = E$, $H = \sqrt{(\epsilon/\mu)}\mathrm{H}$, where ϵ and μ are the permittivity and permeability coefficients, in free space, respectively, c is the speed of light, and L is a length of the domain. The 2-D version of system (3.1), (3.2) decouples into two independent sets of equations. We shall consider the TM (transverse magnetic) system in a square domain $\Omega = \{(x, y) \in \mathbf{R}^2 \mid 0 \le x \le 1, 0 \le y \le 1\}$. The TM equations then become

$$\begin{aligned}
\frac{\partial E_z}{\partial t} &= \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \qquad (x, y) \in \Omega,\ t \ge 0\\
\frac{\partial H_x}{\partial t} &= -\frac{\partial E_z}{\partial y}\\
\frac{\partial H_y}{\partial t} &= \frac{\partial E_z}{\partial x}
\end{aligned} \tag{3.4}$$

with the boundary conditions

$$\begin{aligned}
E_z(0, y, t) = E_z(1, y, t) &= 0, \qquad t \ge 0,\\
E_z(x, 0, t) = E_z(x, 1, t) &= 0.
\end{aligned} \tag{3.5}$$

We take as initial conditions,

$$\begin{aligned}
E_z(x, y, 0) &= \sin(\omega_1 x)\sin(\omega_2 y), \qquad (x, y) \in \Omega,\\
H_x(x, y, 0) &= 0,\\
H_y(x, y, 0) &= 0,
\end{aligned} \tag{3.6}$$

where $\omega_1 = \pi n$ and $\omega_2 = \pi m$ ($n, m = \pm 1, \pm 2, \pm 3, \ldots$).

The exact solution is

$$\begin{aligned} E_z(x,y,t) &= \sin(\omega_1 x)\sin(\omega_2 y)\cos(\omega t),\\ H_x(x,y,t) &= -\frac{\omega_2}{\omega}\sin(\omega_1 x)\cos(\omega_2 y)\sin(\omega t),\\ H_y(x,y,t) &= \frac{\omega_1}{\omega}\cos(\omega_1 x)\sin(\omega_2 y)\sin(\omega t), \end{aligned} \tag{3.7}$$

where $\omega = \sqrt{\omega_1^2 + \omega_2^2}$.

The matrix form of the equations (3.4) is

$$\begin{aligned} \frac{\partial}{\partial t}\begin{pmatrix} E_z \\ H_x \\ H_y \end{pmatrix} &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \frac{\partial}{\partial x}\begin{pmatrix} E_z \\ H_x \\ H_y \end{pmatrix} + \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \frac{\partial}{\partial y}\begin{pmatrix} E_z \\ H_x \\ H_y \end{pmatrix} \\ &= A_1 \frac{\partial}{\partial x}\begin{pmatrix} E_z \\ H_x \\ H_y \end{pmatrix} + A_2 \frac{\partial}{\partial y}\begin{pmatrix} E_z \\ H_x \\ H_y \end{pmatrix}, \end{aligned} \tag{3.8}$$

where

$$A_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The SAT method for implementation of boundary conditions is used for diagonalized systems in one dimension. We encounter a problem when dealing with this two-dimensional problem, because it is impossible to diagonalize the two matrices A_1 and A_2 simultaneously. To overcome this problem of how to state the boundary conditions we consider the two-dimensional Maxwell's equations (3.4) in each space dimension independently. We decompose (3.8) into the one-dimensional Maxwell's equations[2]

$$\frac{\partial}{\partial t}\begin{pmatrix} E_z \\ H_y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{\partial}{\partial x}\begin{pmatrix} E_z \\ H_y \end{pmatrix}, \tag{3.9}$$

$$\frac{\partial}{\partial t}\begin{pmatrix} E_z \\ H_x \end{pmatrix} = -\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \frac{\partial}{\partial y}\begin{pmatrix} E_z \\ H_x \end{pmatrix}, \tag{3.10}$$

with $E_z = 0$ at the boundaries (see (3.5)), and we denote

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

We shall limit our detailed discussion only to Eq. (3.9). The treatment of the equation (3.10) is similar.

[2] This decomposition is not, of course, equivalent to the original system (3.8). It is done for lack of a 2-D characteristic theory. This practice follows what has been done previously in the context of 2-D gas dynamics; see [4].

We diagonalize the matrix A and change the variables. Let M be a diagonalizing matrix of A and let Λ be a diagonal matrix having the eigenvalues of A, i.e.,

$$M^{-1}AM = \Lambda = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.11}$$

and

$$M = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}, \qquad M^{-1} = \frac{1}{2}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}. \tag{3.12}$$

Equation (3.9) is transformed into

$$\frac{\partial}{\partial t}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\frac{\partial}{\partial x}\begin{pmatrix} u \\ v \end{pmatrix}, \tag{3.13}$$

where

$$\begin{pmatrix} u \\ v \end{pmatrix} = M^{-1}\begin{pmatrix} E_z \\ H_x \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -E_z + H_y \\ E_z + H_y \end{pmatrix}.$$

The boundary conditions can be written as

$$u(0, y, t) = v(0, y, t), \qquad v(1, y, t) = u(1, y, t). \tag{3.14}$$

This is equivalent to the requirement of $E_z = 0$ on the boundaries. Note also that (3.14) is in the form (2.3) with $g^{\mathrm{I}}(t) = g^{\mathrm{II}}(t) = 0$ and $R = L = 1$.

We add to the system (3.13) an artificial zero term which is similar to the SAT term for a one-dimensional hyperbolic system and rewrite it as

$$\frac{\partial}{\partial t}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\frac{\partial}{\partial x}\begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} \alpha[u(0, y, t) - v(0, y, t)] \\ \beta[v(1, y, t) - u(1, y, t)] \end{pmatrix}, \tag{3.15}$$

where α and β are some constants.

When we return to the original variables, i.e., E_z, H_y, we get

$$\begin{aligned} \frac{\partial}{\partial t}\begin{pmatrix} E_z \\ H_y \end{pmatrix} &= A\frac{\partial}{\partial x}\begin{pmatrix} E_z \\ H_y \end{pmatrix} + M\begin{pmatrix} \alpha[u(0, y, t) - v(0, y, t)] \\ \beta[v(1, y, t) - u(1, y, t)] \end{pmatrix} \\ &= A\frac{\partial}{\partial x}\begin{pmatrix} E_z \\ H_y \end{pmatrix} + \begin{pmatrix} -\alpha[u(0, y, t) - v(0, y, t)] + \beta[v(1, y, t) - u(1, y, t)] \\ \alpha[u(0, y, t) - v(0, y, t)] + \beta[v(1, y, t) - u(1, y, t)] \end{pmatrix}. \end{aligned} \tag{3.16}$$

Using the fact that

$$\begin{pmatrix} E_z \\ H_y \end{pmatrix} = M\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -u + v \\ u + v \end{pmatrix}$$

we replace the boundary terms $u(0,y,t)-v(0,y,t)$, $v(1,y,t)-u(1,y,t)$ in (3.16) by the original variables $E_z(0,y,t)$, $E_z(1,y,t)$.

Thus (3.16) becomes

$$\frac{\partial}{\partial t}\begin{pmatrix}E_z\\H_y\end{pmatrix}=\begin{pmatrix}0&1\\1&0\end{pmatrix}\frac{\partial}{\partial x}\begin{pmatrix}E_z\\H_y\end{pmatrix}+\begin{pmatrix}\alpha E_z(0,y,t)+\beta E_z(1,y,t)\\-\alpha E_z(0,y,t)+\beta E_z(1,y,t)\end{pmatrix}. \tag{3.17}$$

We now call attention to the fact that the systems (3.9) and (3.17) are equivalent (see (3.5)). In a similar fashion we get for E_z, H_x a system which is equivalent to (3.10):

$$\frac{\partial}{\partial t}\begin{pmatrix}E_z\\H_x\end{pmatrix}=-\begin{pmatrix}0&1\\1&0\end{pmatrix}\frac{\partial}{\partial y}\begin{pmatrix}E_z\\H_x\end{pmatrix}+\begin{pmatrix}+\alpha E_z(x,0,t)+\beta E_z(x,1,t)\\\alpha E_z(x,0,t)-\beta E_z(x,1,t)\end{pmatrix}. \tag{3.18}$$

When we approximate the nondiagonalized equations (3.17) and (3.18) numerically by using the SAT method for implementation of boundary conditions we shall add SAT boundary terms for both directions, which resemble the artificial zero terms that appear in the equations (3.17), (3.18). Let Δx and Δy be mesh widths in the x- and y-directions, and divide the axes into subintervals of length Δx and Δy, respectively. For $i=0,\dots,N_1$ and $j=0,\dots,N_2$ we use the notation

$$E_{z_{ij}}(t)=E_z(x_i,y_j,t),\qquad H_{x_{ij}}(t)=H_x(x_i,y_j,t),\qquad H_{y_{ij}}(t)=H_y(x_i,y_j,t),$$
$$x_i=i\Delta x,\qquad y_j=j\Delta y,$$
$$N_1\Delta x=1,\qquad N_2\Delta y=1,$$

where $E_{z_{ij}}(t)$, $H_{x_{ij}}(t)$, and $H_{y_{ij}}(t)$ are vector grid functions. We denote by $e_{z_{ij}}$, $h_{x_{ij}}$, and $h_{y_{ij}}$ the numerical approximations to the projections $E_{z_{ij}}(t)$, $H_{x_{ij}}(t)$, and $H_{y_{ij}}(t)$, respectively. Without loss of generality we take $N=N_1=N_2$, i.e., $\Delta x=\Delta y$.

Before proceeding to the semidiscrete problem let us define

$$D_x=\tilde{P}^{-1}\tilde{Q},\qquad D_y=P^{-1}Q, \tag{3.19}$$

where $(N+1)\times(N+1)$ matrices P, Q and $\tilde{P}$, $\tilde{Q}$ are the same matrices used to solve the hyperbolic system in the one-dimensional case and described in detail in Part I and in [3]. We note that in practice P^{-1} and $\tilde{P}^{-1}$ are never evaluated. Rather, the decomposition $P=LU$ and $\tilde{P}=\tilde{L}\tilde{U}$ is calculated once for each matrix. L and U ($\tilde{L}$ and $\tilde{U}$) are bidiagonal matrices with one of them having "ones" along the diagonal. Hence, the inverse of L and U ($\tilde{L}$ and $\tilde{U}$) is very cheap (two additions and three multiples per point).

Let $[e_z]$, $[h_x]$, and $[h_y]$ be the $(N+1)\times(N+1)$ matrices with the elements $e_{z_{ij}}$, $h_{x_{ij}}$, and $h_{y_{ij}}$, respectively, and denote by $[e_z]_j^R$, $[h_x]_j^R$, and $[h_y]_j^R$ the jth row of each of these matrices and by $[e_z]_i^C$, $[h_x]_i^C$, and $[h_y]_i^C$ the ith column of each of these matrices.

We now write the semidiscrete approximation to (3.17) as

$$\begin{aligned}\frac{d}{dt}[e_z]_j^R&=D_x[h_y]_j^R-\tilde{P}^{-1}\big(\vec{S}_0e_{z_{0j}}+\vec{S}_Ne_{z_{Nj}}\big),\\\frac{d}{dt}[h_y]_j^R&=D_x[e_z]_j^R-\tilde{P}^{-1}\big(-\vec{S}_0e_{z_{0j}}+\vec{S}_Ne_{z_{Nj}}\big),\end{aligned} \tag{3.20}$$

and the semidiscrete approximation to (3.18) as

$$\begin{aligned}\frac{d}{dt}[e_z]_i^C &= -[h_x]_i^C D_y^T + P^{-1}\big(\vec{S}_0 e_{z_{i0}} + \vec{S}_N e_{z_{iN}}\big),\\ \frac{d}{dt}[h_x]_i^C &= -[e_z]_i^C D_y^T + P^{-1}\big(\vec{S}_0 e_{z_{i0}} - \vec{S}_N e_{z_{iN}}\big),\end{aligned} \tag{3.21}$$

where the $(N+1)$ long vectors $\vec{S}_N$ and $\vec{S}_0$ are exactly the same vectors as in the one-dimensional case, i.e.,

$$\vec{S}_0 = \begin{pmatrix} \tau q_{00} \\ (q_{01}+q_{10}) \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \vec{S}_N = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -(q_{01}+q_{10}) \\ -\tau q_{00} \end{pmatrix}. \tag{3.22}$$

We now compose the two one-dimensional systems into the two-dimensional set and approximate the equations (3.8) in the following way:

$$\begin{aligned}\frac{d}{dt}[e_z] &= D_x[h_y] - \big([e_z]_0^C \vec{S}_0^T + [e_z]_N^C \vec{S}_N^T\big)\,\tilde{P}^{-1} - [h_x]D_y^T + P^{-1}\big(\vec{S}_0[e_z]_0^R + \vec{S}_N[e_z]_N^R\big)\\ \frac{d}{dt}[h_x] &= -[e_z]D_y^T + P^{-1}\big(\vec{S}_0[e_z]_0^R - \vec{S}_N[e_z]_N^R\big)\\ \frac{d}{dt}[h_y] &= D_x[e_z] - \big(-[e_z]_0^C \vec{S}_0^T + [e_z]_N^C \vec{S}_N^T\big)\,\tilde{P}^{-1}.\end{aligned} \tag{3.23}$$

3.2. *Maxwell's Equations: Numerical Simulations*

The problem (3.4), (3.5), (3.7) was solved using both the fourth-order scheme and the sixth-order scheme. The boundary conditions are imposed using the SAT algorithm described above. In all cases, the temporal advance is via the standard fourth-order Runge–Kutta method. The time step is chosen small enough to ensure the local stability of the Runge–Kutta method and retain the desired overall accuracy (see footnote 1). The simulations were all run to equivalent times $T = 100$ for both the fourth- and the sixth-order schemes and different grids ($N = N_1 = N_2 = 20, 40, 80$). We chose CFL $= 1/10$, $\tau = 2$ for the fourth-order scheme and CFL $= 1/15$, $\tau = 2$ for the sixth-order scheme. In Figs. 7–9 the $\log_{10}$ of the L_2 error is computed for both schemes and different grids. As one can see, the error grows linearly in time; no exponential growth exists, indicating temporal stability of the schemes. Figure 11 shows the e_z component of the numerical solution at time $T = 2$ obtained by using the sixth-order scheme with $N = 80$, $\tau = 2$.

Unlike previous sections, where we compared two procedures for imposing of boundary conditions (the conventional procedure and the SAT procedure), in this section we shall compare our results with the results obtained by Turkel and Yefet; see [5, 6]. They solved the same problem by using the Ty(2,4) scheme, which is a fourth-order compact implicit difference scheme on staggered meshes. For time integration they used the staggered leapfrog method. The Ty(2,4) algorithm was run for $N = 20$, CFL $= 1/18$; $N = 40, 80$, CFL $= 1/44$.

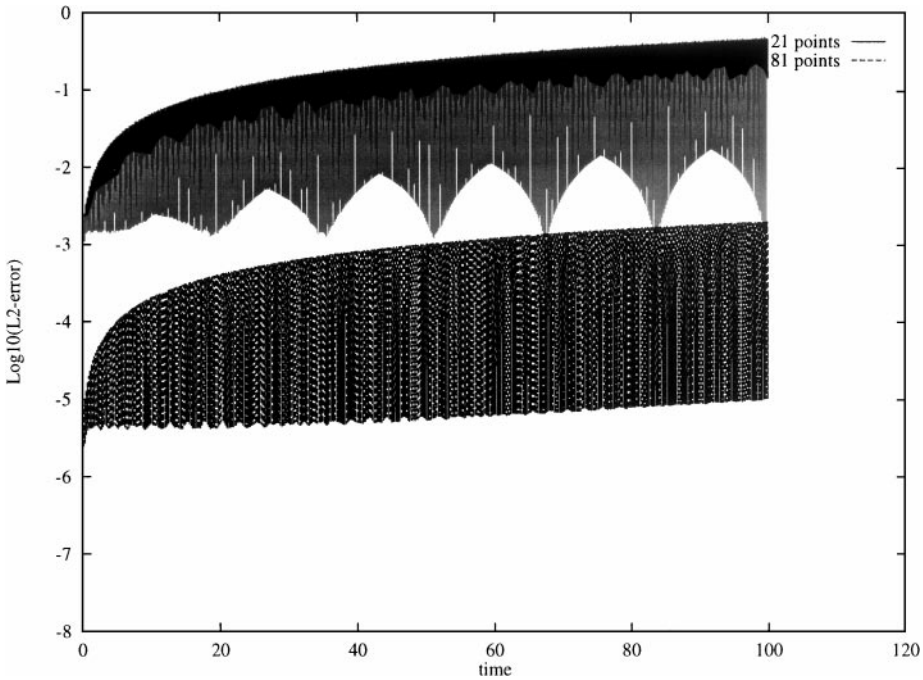

FIG. 7. The L_2-error as a function of time for the SAT fourth-order approximation with CLF $= 0.1$, $\tau = 2$, $\omega_1 = 3\pi$, $\omega_2 = 4\pi$, $\omega = 5\pi$. $N = 20, 80$.

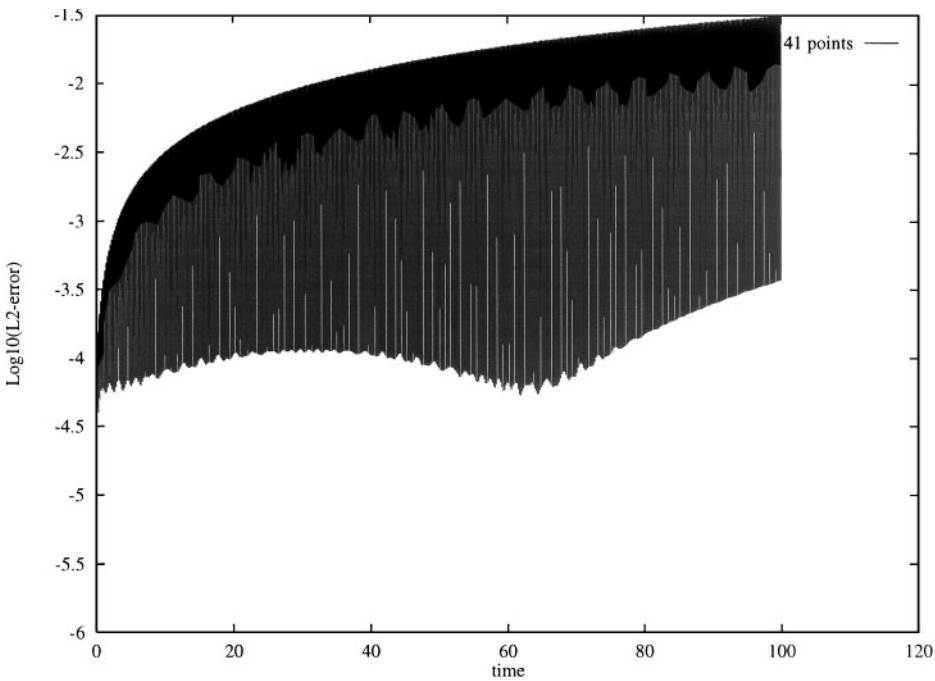

FIG. 8. The L_2-error as a function of time for the SAT fourth-order approximation with CFL $= 0.1$, $\tau = 2$, $\omega_1 = 3\pi$, $\omega_2 = 4\pi$, $\omega = 5\pi$. $N = 40$.

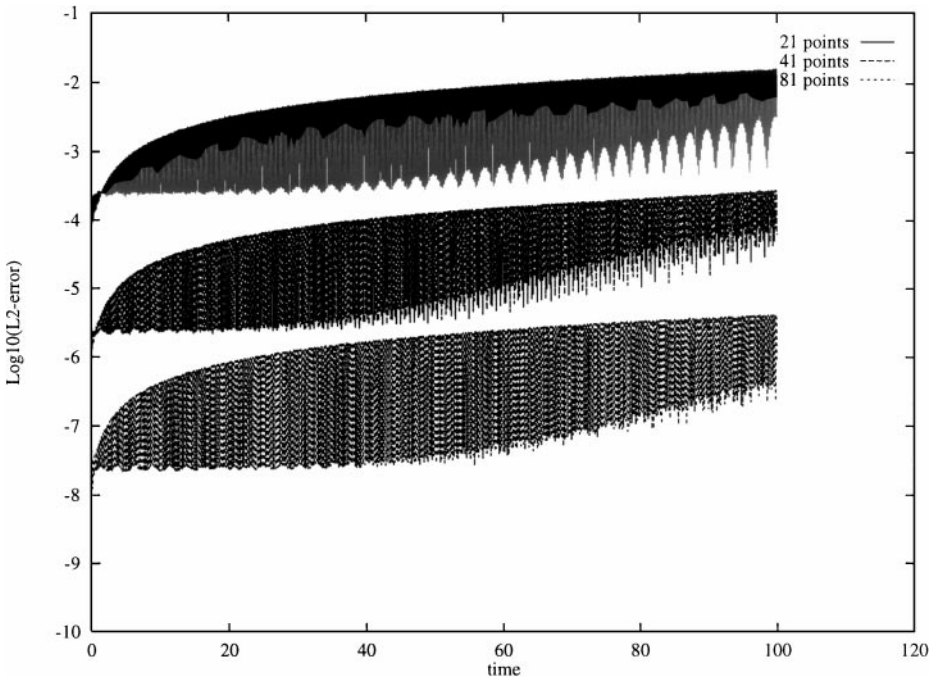

FIG. 9. The L_2-error as a function of time for the SAT sixth-order approximation with CFL $= 1/15$, $\tau = 2$, $\omega_1 = 3\pi$, $\omega_2 = 4\pi$, $\omega = 5\pi$. $N = 20$, 40, 80.

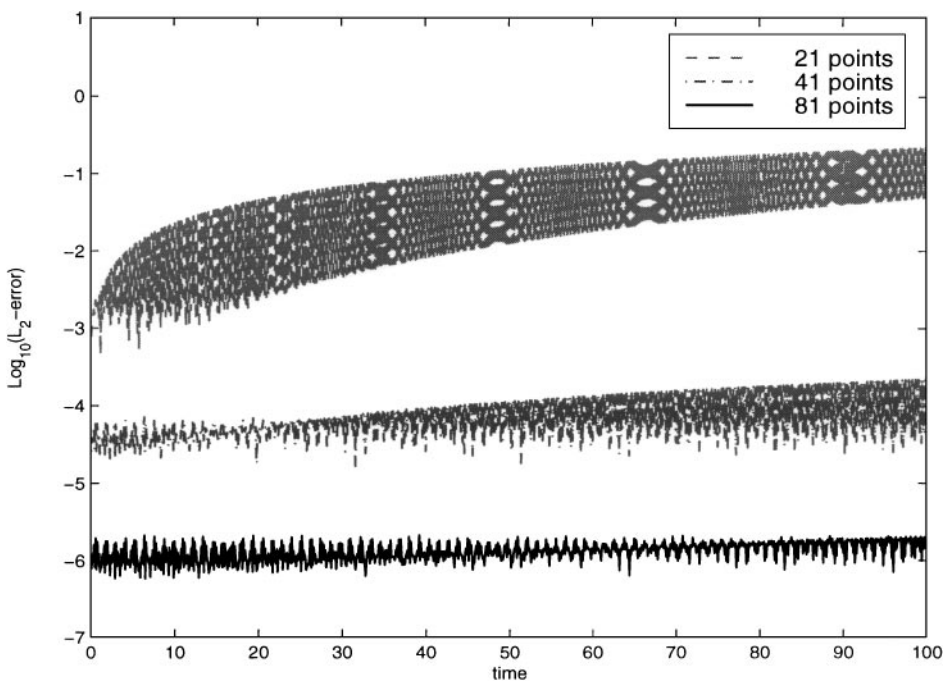

FIG. 10. The L_2-error as a function of time for the Ty(2,4) fourth-order approximation for $N = 20$: CFL $= 1/18$; for $N = 40$, 80: CFL $= 1/44$. $\omega_1 = 3\pi$, $\omega_2 = 4\pi$, $\omega = 5\pi$.

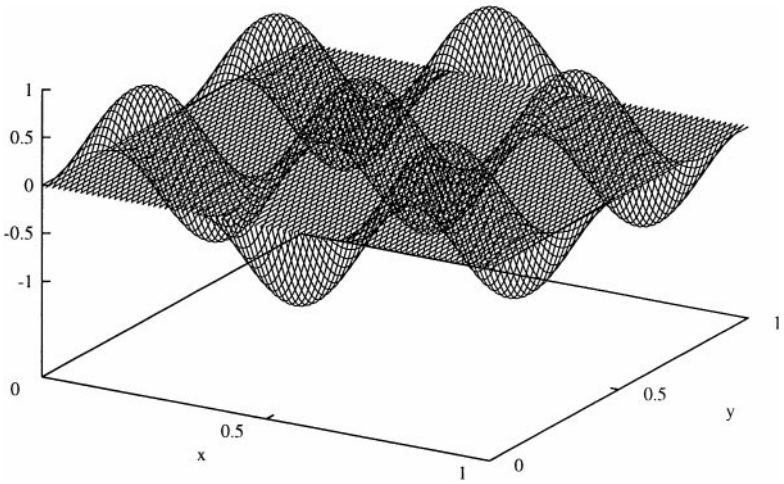

FIG. 11. e_z component of the numerical solution at $T = 2$ obtained using the SAT sixth-order approximation with $N = 80$, CFL $= 1/15$, $\tau = 2$, $\omega_1 = 3\pi$, $\omega_2 = 4\pi$, $\omega = 5\pi$.

The $\log_{10}$ of the L_2 error, obtained by using the Ty(2,4) fourth-order scheme, is plotted in Fig. 10. Note that the Ty algorithm was run with a time step, Δt, almost 2 times smaller for $N = 20$ and almost $4\frac{1}{2}$ times smaller for $N = 40$, 80 than one used for the fourth-order SAT scheme. It should also be observed that the results obtained by using the SAT schemes and presented in Figs. 7–9 are printed every Δt step while the results obtained by using the Ty(2,4) scheme and presented in Fig. 10 are printed every $1/(10\Delta t)$ steps (i.e., only 1000 points are printed, in contrast to about 20,000–80,000 points for our printout graphs).

In order to check on the order of accuracy, the runs were repeated for ($N = N_1 = N_2 = 20$, 40, 80). Table II shows a grid refinement study for all three spatial operators. The absolute error $\log_{10}(L_2)$ at a fixed time $t = T$ and the convergence rate between two grids are plotted. The results in this table agree very well with the predicted ones for fourth and sixth order. We note than the error obtained by using the Ty(2,4) fourth-order scheme is smaller than the error in SAT fourth-order scheme, but the SAT sixth-order scheme outperforms both.

TABLE II

Grid Convergence of Schemes for the Two-Dimensional Maxwell Equations

	Ty(2,4) fourth-order		SAT fourth-order		SAT sixth-order	
Grid	$\log_{10}(L_2)$	Rate	$\log_{10}(L_2)$	Rate	$\log_{10}(L_2)$	Rate
21	−2.677		−2.644		−3.580	
41	−4.234	5.17	−4.089	4.80	−5.416	6.10
81	−5.751	5.03	−5.326	4.11	−7.261	6.13

Note. $T = 10$, $\omega_1 = 3\pi$, $\omega_2 = 4\pi$, $\omega = 5\pi$. Here CFL $= 1/10$ for the SAT fourth-order scheme and CFL $= 1/15$ for the SAT sixth-order scheme. For Ty(2,4): $N = 20$, CFL $= 1/18$; $N = 40$, 80, CFL $= 1/44$.

4. CONCLUDING REMARKS

In this Part II of this series, the methodology presented in Part I was used to solve one- and two-dimensional hyperbolic systems. Analytical proof of time stability for one-dimensional hyperbolic systems was obtained for a restricted class of problems, namely when $\|L\| \cdot \|R\| \leq 1/5$ for the sixth-order accurate scheme and $\|L\| \cdot \|R\| \leq 1/3$ for the fourth-order scheme. However, it has been numerically verified, by both measuring the error for long time integrations and determining the eigenvalue spectrum of the semidiscrete system, that the method was effective and provided time stability even when a theoretical foundation was lacking. We have shown application in the most severe case of $\|L\| \cdot \|R\| = 1$.

The numerical experiments were concluded by solving the two-dimensional Maxwell's equations in free space. The SAT method used for solving diagonalized systems in one dimension was adopted to solve a nondiagonalizable two-dimensional system. Numerical results obtained by using both fourth- and sixth-order SAT schemes were compared with the results yielded by the fourth-order Ty(2,4) scheme derived by Turkel and Yefet in [5, 6].

REFERENCES

1. M. H. Carpenter, D. Gottlieb, and S. Abarbanel, The stability of numerical boundary treatments for compact high-order finite-difference schemes, *J. Comput. Phys.* **108**, 272 (1993).
2. M. H. Carpenter, D. Gottlieb, and S. Abarbanel, Time-stable boundary conditions for finite difference schemes solving hyperbolic systems: Methodology and applications to high-order compact schemes, *J. Comput. Phys.* **111**, 220 (1994).
3. A. Chertock, *Strict Stability of High-Order Compact Implicit Finite-Difference Schemes—The Role of Boundary Conditions for Hyperbolic PDEs*, Ph.D. thesis, Tel-Aviv University, Tel-Aviv, Israel, November 1998.
4. B. Gustafsson, H.-O. Kreiss, and J. Oliger, *Time Dependent Problems and Difference Methods* (Wiley, New York, 1995).
5. A. Yefet and E. Turkel, Fourth order compact implicit method for the Maxwell equations with discontinuous coefficients, *Appl. Num. Math.*, to appear.
6. E. Turkel, High-order methods, in *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, edited by A. Taflove (Artech House, Boston, 1998), p. 63.
7. S. S. Abarbanel and A. E. Chertock, Strict stability of high-order compact implicit finite-difference schemes: The role of boundary conditions for for hyperbolic PDEs, I, *J. Comput. Phys.* **158**, 1–25 (2000).