ERROR BOUNDED SCHEMES FOR TIME-DEPENDENT HYPERBOLIC PROBLEMS

JAN NORDSTRÖM*

Abstract.
In this paper we address the error growth in time for hyperbolic problems on first order form. The energy method is used to study when an error growth or a fixed error bound is obtained. It is shown that the choice of boundary procedure is a crucial point. Numerical experiments corroborate the theoretical findings.

Key words. temporal error, error bound, hyperbolic problem, boundary conditions, energy method, well posed problem, stability, weak formulation, penalty technique

AMS subject classifications. 35M40,35M45,35M50,65M06,65M12,65M15

1. Introduction. Stable approximations of hyperbolic problems on first order form often exhibit a linear (or near linear) error growth in time, see for example [15],[13],[6]. In many of those cases, the wave is restricted or trapped in the domain for long times. Typical examples include periodic problems or problems where the wave is trapped in cavities. However, also cases with a definite bound on the error as time passes can be observed. Typically, in such cases, the wave propagates through the domain for a limited time as in an inflow-outflow problem.

To initiate our investigation we consider the problem

$$u_t + u_x = F(x,t), \quad -\infty < x < \infty, \quad t \geq 0, \tag{1.1}$$

where F is a 1-periodic forcing function in x. The 1-periodic solution which we consider known in this example provide us with initial $(f(x))$ and boundary data $(g(t))$. Assume that we want to compute the numerical solution to (1.1) on the domain $x \in [0,1]$. There are two different ways of imposing the periodic boundary conditions. The first one does it directly by demanding $u(0,t) = u(1,t)$. A second possible choice reads $u(0,t) = g(t)$.

We compute (details of the numerical procedures are discussed later) the two solutions and subtract the exact solution. The error as a function of time for the two types of boundary conditions is shown in Figure 1.1. The computation using the first type of boundary condition (shown by the dotted line) leads to an essentially linear error growth in time while the second type has an error bound.

A hint to why the results are different can be obtained by studying (1.1) on $x \in [0,1]$ with a perturbed forcing function $F + \delta F$ and identical initial

*The Swedish Defence Research Agency, Division of Systems Technology, Department of Computational Physics; and the Department of Information Technology, Scientific Computing, Uppsala University, email: Jan.Nordstrom@foi.se

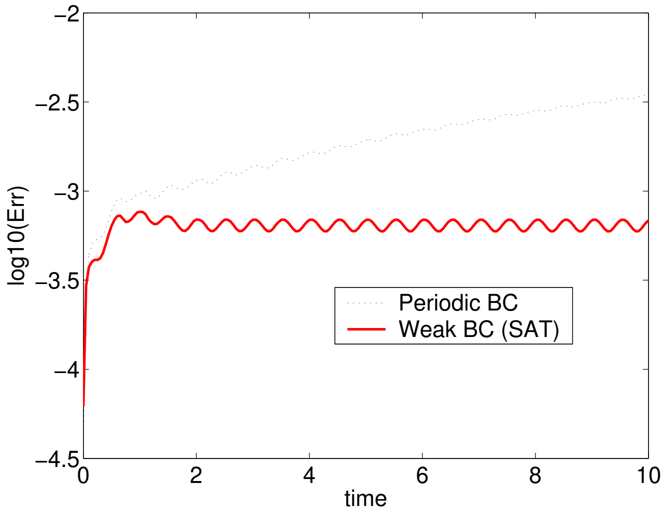

FIGURE 1.1. *The error as a function time.*

and boundary conditions. Denote the solution to the perturbed problem by v and subtract the equations. That yields the error equation,

$$e_t + e_x = \delta F(x,t), \quad 0 \leq x \leq 1, \quad t \geq 0, \tag{1.2}$$

where $e = v - u$ and $e(x,0) = 0$. The two corresponding boundary conditions are $e(0,t) = e(1,t)$ and $e(0,t) = 0$ respectively.

The energy method applied to (1.2) yields

$$||e||_t \leq -\lambda(t)||e|| + ||\delta F||, \tag{1.3}$$

where $||e||^2 = \int_0^1 e^2 dx$. For the boundary condition $e(0,t) = e(1,t)$ we have $\lambda(t) = 0$ which leads to the error growth shown in Figure 1.1. More interesting is the fact that we get $\lambda(t) = (e(1,t)/||e||)^2/2$ by using the boundary condition $e(0,t) = 0$. A constant positive λ leads directly to error boundedness. Actually, as shown in Figure 1.1, $\lambda(t) \geq 0$ suffice for error boundedness.

The error growth in time is particularly problematic for hyperbolic problems. For parabolic problems (add ϵu_{xx} to the right hand side of (1.1)), a temporal error bound is obtained by using Pointcare's inequality $(||e_x||/||e|| \geq \delta > 0)$ even with vanishing boundary terms. However, for vanishing ϵ, as the parabolic problem converges to a hyperbolic problem, the error growth appears since the coefficient $\lambda = \epsilon\delta$ in (1.3) goes to zero.

Some of the phenomena above were discussed in [1],[3],[4]. In this paper we will both extend and simplify that analysis and focus on the error behavior in time for hyperbolic problems. Our theoretical tools will be the energy-method, summation-by-parts operators and penalty procedures for implementing boundary conditions.

The remainder of this paper will proceed as follows. In section 2, we consider different boundary procedures for a general continuous hyperbolic problem and derive an error estimate. The corresponding semi-discrete finite-difference approximation and error estimate is considered in section 3. Numerical experiments are performed in section 4 and finally we draw conclusions in section 5.

2. The continuous problem. We will consider a general three dimensional hyperbolic problem in a half space of the form,

$$\begin{aligned} Su_t + Au_x + Bu_y + Cu_z = \tilde{F}, &\quad x \geq 0, \quad t \geq 0, \\ Lu = \tilde{G}, &\quad x = 0, \quad t \geq 0, \\ u = \tilde{H}, &\quad x \geq 0, \quad t = 0. \end{aligned} \tag{2.1}$$

We assume that all matrices (S, A, B, C) are symmetric (or simultaneously symmetrizable, see [2]) and that S is a diagonal strictly positive matrix. In terms of applications this means that we consider aerodynamic, aeroacoustic (the Euler equations) and electromagnetic (the Maxwell's equations) problems. We also assume that u, $\tilde{F}, \tilde{H}$ have compact support in $x \geq 0$.

Definition 2.1. *(2.1) is said to be strongly well posed if an unique solution exists and the estimate*

$$\|u\|^2 + \int_0^t \|u\|_\Gamma^2 d\tau \leq K_c e^{\eta_c t} \left(\|\tilde{H}\|^2 + \int_0^t (\|\tilde{F}\|^2 + \|\tilde{G}\|_\Gamma^2) d\tau \right) \tag{2.2}$$

holds. K_c and η_c are constants and do not depend on $\tilde{F}, \tilde{G}$ or $\tilde{H}$. $\|\cdot\|_\Gamma$ and $\|\cdot\|$ are suitable continuous norms.

Remark: If one can estimate the solution for zero boundary data only, the problem is called well posed, see [7] for more details.

2.1. Well-posedness. The energy-method applied to (2.1) yields

$$((u, u)_S)_t = \int_{x=0} u^T A u \; dydz + 2(u, \tilde{F})_I, \tag{2.3}$$

where I is the identity matrix and we have introduced the scalar product

$$(u, v)_T = \int_{x \geq 0} u^T T v \; dxdydz. \tag{2.4}$$

T is a symmetric positive definite matrix and $(u, u)_T = ||u||_T^2$ defines a norm. In Definition 2.1 above,

$$||u||_I^2 = ||u||^2 \quad \text{and} \quad \int_{x=0} u^T u \; dydz = \|u\|_\Gamma^2,$$

were used.

We have the splitting,

$$u^T A u = (X^T u)^T (\Lambda^+ + \Lambda^-)(X^T u) = u^T A^+ u + u^T A^- u. \tag{2.5}$$

The superscripts $+, -$ indicate the positive and negative portions of the eigenvalue matrix Λ and A. The (characteristic) boundary condition

$$Lu = (X^T u)^+ = \tilde{G} = (X^T \tilde{g})^+, \quad \text{or} \quad A^+ u = A^+ \tilde{g} \tag{2.6}$$

yields the estimate,

$$((u,u)_S)_t = \int_{x=0} \tilde{g}^T A^+ \tilde{g} + u^T A^- u \; dydz + 2(u, \tilde{F})_I. \tag{2.7}$$

A direct integration of (2.7) yields an estimate of the form (2.2) which proves the following proposition.

PROPOSITION 2.2. *(2.1) augmented with (2.6) is strongly well posed.*

2.2. The continuous error equation. Consider (2.1) with a perturbed forcing function $\tilde{F} + \delta\tilde{F}$ and identical initial and boundary conditions. Denote the solution to the perturbed problem by v and subtract the equations. That yields the error equation,

$$\begin{aligned} Se_t + Ae_x + Be_y + Ce_z = \delta\tilde{F}, \quad x \geq 0, \quad t \geq 0 \\ Le = 0, \quad x = 0, \quad t \geq 0 \\ e = 0, \quad x \geq 0, \quad t = 0. \end{aligned} \tag{2.8}$$

The energy method applied to (2.8) and the relations,

$$((e,e)_S)_t = 2||e||_S(||e||_S)_t, \quad (e, \delta\tilde{F})_I \leq ||e||_I ||\delta\tilde{F}||_I, \tag{2.9}$$

yields,

$$(||e||_S)_t \leq C_0 ||e||_S + C_1 ||\delta\tilde{F}||_I, \tag{2.10}$$

where

$$C_0 = \frac{\int_{x=0} e^T A e \; dydz}{2||e||_S^2}, \quad C_1 = \frac{||e||_I}{||e||_S}. \tag{2.11}$$

Remark: We call the boundary procedure dissipative if $C_0 < 0$. Dissipative boundary treatments, for example using (2.6), lead to strong stability.

The problem (2.10) is of the form

$$u_t \leq -\lambda(t) u + F(t), \quad u(0) = f, \quad \lambda \geq 0, \tag{2.12}$$

where F has a maximum value $|F|_{max}$. We will need the following lemma.

LEMMA 2.3. *The solution u to (2.12) is bounded for $t \geq 0$ provided that λ is non-zero "almost" everywhere.*

Proof: A direct calculation yields

$$u(t) \leq e^{-\theta(0,t)}f + \int_0^t e^{-\theta(\xi,t)}F(\xi)d\xi, \quad \theta(\xi,t) = \int_\xi^t \lambda(\eta)d\eta. \tag{2.13}$$

Clearly $\theta(\xi,t)$ is a positive monotone function of time. Given the assumption that λ is non-zero "almost" everywhere, we have the estimate

$$\theta(\xi,t) \geq \delta_0(t-\xi), \quad \delta_0 > 0. \tag{2.14}$$

The estimate (2.14) and (2.13) leads to

$$|u| \leq e^{-\delta_0 t}f + |F|_{max}(1 - e^{-\delta_0 t})/\delta_0, \tag{2.15}$$

which means that u remains bounded as time increases. ■

We can now state one of the main results of this paper.

PROPOSITION 2.4. *If the boundary procedure in (2.1) is dissipative, i.e. $C_0 < 0$ "almost" everywhere in (2.10), then the error is bounded in time. It is bounded by the size of the forcing function.*

Proof: Lemma 2.3 applied to (2.10) leads to the result. ■

Remark: We include a non-homogeneous initial condition in (2.12) even though (2.10) has zero initial data. The reason being that C_0 can be zero initially since $e(0) = 0$. If so, the error grows linearly for a short time and form the non-zero initial error.

An immediate result of Proposition 2.2 and 2.4 is the following corollary.

COROLLARY 2.5. *Strong well posedness lead to a temporal error bound.*

3. The semi-discrete problem.

Let $\boldsymbol{u}, \boldsymbol{u_x}$ be the numerical approximation of the scalar quantities u, u_x respectively. The approximation of the first derivative $\boldsymbol{u_x}$ is introduced as

$$\boldsymbol{u_x} = P^{-1}Q\boldsymbol{u}, \tag{3.1}$$

where P and Q are matrices. If a spatial operator is of the form (3.1) and the conditions (i) and (ii) below are full-filled, the operator is referred to as a Summation By Parts operator (SBP) (see [11]).

(i) The matrix P is symmetric, positive definite and bounded, $\Delta x p I \leq P \leq \Delta x q I$, where $p > 0$ and q are bounded independent of $1/\Delta x$.

(ii) The matrix Q is almost skew-symmetric. $Q + Q^T = diag(-1, 0, .., 0, 1)$.

The boundary and interface conditions will be imposed by the weak penalty technique referred to as the SAT (Simultaneous Approximation Term) procedure. For details on penalty procedures for finite difference, finite volume and spectral methods, see [5], [14] and [8] respectively.

The unknowns are organized in a vector $\boldsymbol{u}$ where $\boldsymbol{u}_{i,j,k}$ corresponds to the grid point (x_i, y_j, z_k). The semi-discrete version of problem (2.1) on the finite difference form and using the SBP operators above can be written

$$\begin{aligned}\Big(\boldsymbol{I}_x \otimes \boldsymbol{I}_y \otimes \boldsymbol{I}_z \otimes S\Big)\boldsymbol{u}_t + \Big(\boldsymbol{P}_x^{-1}\boldsymbol{Q}_x \otimes \boldsymbol{I}_y \otimes \boldsymbol{I}_z \otimes A\Big)\boldsymbol{u}&\\ \Big(\boldsymbol{I}_x \otimes \boldsymbol{P}_y^{-1}\boldsymbol{Q}_y \otimes \boldsymbol{I}_z \otimes B\Big)\boldsymbol{u} + \Big(\boldsymbol{I}_x \otimes \boldsymbol{I}_y \otimes \boldsymbol{P}_z^{-1}\boldsymbol{Q}_z \otimes C\Big)\boldsymbol{u}&\\ = \Big(\boldsymbol{P}_x^{-1}B_x \otimes \boldsymbol{I}_y \otimes \boldsymbol{I}_z \otimes \Sigma\Big)(\boldsymbol{u} - \tilde{\boldsymbol{g}}) + \tilde{\boldsymbol{F}}&\\ \boldsymbol{u}(0) = \tilde{\boldsymbol{H}}&\end{aligned} \tag{3.2}$$

The symbol $\otimes$ denotes the Kronecker product, see appendix A. In (3.2), $\boldsymbol{I}_x, \boldsymbol{I}_y, \boldsymbol{I}_z$ are identity matrices of appropriate size, B_x is identically zero except that $(B_x)_{11} = 1$ and Σ is an undetermined matrix. The first term on the right hand side imposes the boundary conditions at $x = 0, i = 0$ weakly using the SAT technique. The characteristic boundary condition (2.6) can be imposed in this way by a special choice of Σ.

Definition 3.1. *(3.2) is said to be strongly stable if, for a sufficiently fine mesh $\Delta x, \Delta y, \Delta z < h$, there is an unique solution that satisfies*

$$\|\boldsymbol{u}\|^2 + \int_0^t \|\boldsymbol{u}\|_\Gamma^2 d\tau \le K_d e^{\eta_d t}\left(\|\tilde{\boldsymbol{H}}\|^2 + \int_0^t (\|\tilde{\boldsymbol{F}}\|^2 + \|\tilde{\boldsymbol{G}}\|_\Gamma^2) d\tau\right). \tag{3.3}$$

K_d and η_d are constants and do not depend on $\tilde{\boldsymbol{F}}, \tilde{\boldsymbol{H}}$ or $\tilde{\boldsymbol{G}}$. $\|\cdot\|_\Gamma$ and $\|\cdot\|$ are suitable discrete norms.

Remark: If one can estimate the solution for zero boundary data only, the problem is called stable, see [7] for more details.

3.1. Stability. By applying the energy method to (3.2) (we multiply from the left by $\boldsymbol{u}^T(\boldsymbol{P}_x \otimes \boldsymbol{P}_y \otimes \boldsymbol{P}_z \otimes I)$ and add the transpose) we obtain

$$\begin{aligned}((\boldsymbol{u}, \boldsymbol{u})_{P,S})_t = \boldsymbol{u}_0^T\Big(\boldsymbol{P}_y \otimes \boldsymbol{P}_z \otimes A + \Sigma + \Sigma^T\Big)\boldsymbol{u}_0 - 2\boldsymbol{u}_0^T\Big(\boldsymbol{P}_y \otimes \boldsymbol{P}_z \otimes \Sigma\Big)\tilde{\boldsymbol{g}}&\\ + 2(\boldsymbol{u}, \tilde{\boldsymbol{F}})_{P,I},&\end{aligned} \tag{3.4}$$

where $\boldsymbol{u}_0$ denotes the restriction of $\boldsymbol{u}$ to the boundary at $x = 0, i = 0$ and

$$(\boldsymbol{u}, \boldsymbol{v})_{P,T} = \boldsymbol{u}^T\Big(\boldsymbol{P}_x \otimes \boldsymbol{P}_y \otimes \boldsymbol{P}_z \otimes T\Big)\boldsymbol{v}, \tag{3.5}$$

is the discrete analog to the continuous scalar product (2.4). T is a symmetric positive definite matrix and $(\boldsymbol{u}, \boldsymbol{u})_{P,T} = ||\boldsymbol{u}||_{P,T}^2$ defines a norm. In

Definition 3.1 above,

$$||\boldsymbol{u}||^2_{P,I} = ||\boldsymbol{u}||^2 \quad \text{and} \quad \boldsymbol{u}_0^T\Big(\boldsymbol{P}_y \otimes \boldsymbol{P_z} \otimes I\Big)\boldsymbol{u}_0 = \|\boldsymbol{u}\|^2_\Gamma,$$

were used.

Clearly, we have a bounded growth if $\boldsymbol{\phi}^T(A + \Sigma + \Sigma^T)\boldsymbol{\phi} = 0$ for a general $\boldsymbol{\phi}$ and $\tilde{g} = 0$. We have proved the following proposition.

PROPOSITION 3.2. *The approximation (3.2) with Σ chosen such that*

$$\boldsymbol{\phi}^T(A + \Sigma + \Sigma^T)\boldsymbol{\phi} = 0,$$

for a general $\boldsymbol{\phi}$ is stable.

A particularly useful and strong concept of stability is obtained in the following way. By introducing the previous notation $A = A^+ + A^-$ we find that the boundary terms in (3.4) can be written

$$\begin{aligned} BT = {} & \boldsymbol{u}_0^T\Big(\boldsymbol{P}_y \otimes \boldsymbol{P_z} \otimes A^+ + \Sigma + \Sigma^T\Big)\boldsymbol{u}_0 \\ & + \boldsymbol{u}_0^T\Big(\boldsymbol{P}_y \otimes \boldsymbol{P_z} \otimes A^-\Big)\boldsymbol{u}_o - 2\boldsymbol{u}_0^T\Big(\boldsymbol{P}_y \otimes \boldsymbol{P_z} \otimes \Sigma\Big)\tilde{\boldsymbol{g}}. \end{aligned} \tag{3.6}$$

The choice $\Sigma = \sigma A^+$ and $\sigma \leq -1/2$ means that all the quadratic term in $\boldsymbol{u}_0$ give a negative contribution. The special choice $\sigma = -1$, (3.4) and (3.6) yields the estimate,

$$\begin{aligned} ((\boldsymbol{u}, \boldsymbol{u})_{P,S})_t = {} & \tilde{\boldsymbol{g}}^T\Big(\boldsymbol{P}_y \otimes \boldsymbol{P}_z \otimes A^+\Big)\tilde{\boldsymbol{g}} + \boldsymbol{u}_0^T\Big(\boldsymbol{P}_y \otimes \boldsymbol{P}_z \otimes A^-\Big)\boldsymbol{u}_0 \\ & + 2(\boldsymbol{u}, \tilde{\boldsymbol{F}})_{P,I} - Rest, \end{aligned} \tag{3.7}$$

where

$$Rest = \Big(\boldsymbol{u}_0 - \tilde{\boldsymbol{g}}\Big)^T\Big(\boldsymbol{P}_y \otimes \boldsymbol{P}_z \otimes A^+\Big)\Big(\boldsymbol{u}_0 - \tilde{\boldsymbol{g}}\Big). \tag{3.8}$$

The discrete estimate (3.7) is completely analogous to the continuous estimate (2.7) except for the small negative definite rest term $-Rest$.

A direct integration of (3.7) yields an estimate of the form (3.3) which proves the following proposition.

PROPOSITION 3.3. *The approximation (3.2) with $\Sigma = \sigma A^+$ and $\sigma \leq -1/2$ is strongly stable.*

Remark: Note that the energy method combined with the penalty technique more or less automatically leads to the right number of boundary conditions (equal to the number of positive eigenvalues) and the most appropriate form (the characteristic boundary conditions (2.6)). Note also that the rest term is proportional to the accuracy at the boundary since

$$Rest = \Big((X^T\boldsymbol{u})_0 - \tilde{\boldsymbol{G}}\Big)^T\Big(\boldsymbol{P}_y \otimes \boldsymbol{P}_z \otimes \Lambda^+\Big)\Big((X^T\boldsymbol{u})_0 - \tilde{\boldsymbol{G}}\Big). \tag{3.9}$$

3.2. The semi-discrete error equation. By inserting the continuous solution u (and using the correct boundary data, forcing function and initial data) into the scheme (3.2) and subtracting the discrete solution we obtain the error equation

$$\begin{aligned} \left(\boldsymbol{I}_x \otimes \boldsymbol{I}_y \otimes \boldsymbol{I}_z \otimes S\right)\boldsymbol{e}_t + \left(\boldsymbol{P}_x^{-1}\boldsymbol{Q}_x \otimes \boldsymbol{I}_y \otimes \boldsymbol{I}_z \otimes A\right)\boldsymbol{e} \\ \left(\boldsymbol{I}_x \otimes \boldsymbol{P}_y^{-1}\boldsymbol{Q}_y \otimes \boldsymbol{I}_z \otimes B\right)\boldsymbol{e} + \left(\boldsymbol{I}_x \otimes \boldsymbol{I}_y \otimes \boldsymbol{P}_z^{-1}\boldsymbol{Q}_z \otimes C\right)\boldsymbol{e} \\ = \left(\boldsymbol{P}_x^{-1}B_x \otimes \boldsymbol{I}_y \otimes \boldsymbol{I}_z \otimes \Sigma\right)\boldsymbol{e} + \boldsymbol{T}\boldsymbol{e} \\ \boldsymbol{e}(0) = 0. \end{aligned} \tag{3.10}$$

Clearly the only source of error is the truncation error $\boldsymbol{T}\boldsymbol{e}$.

By following the recipe outlined above (the energy method) we obtain

$$((\boldsymbol{e},\boldsymbol{e})_{P,S})_t = \boldsymbol{e}_0^T\left(\boldsymbol{P}_y \otimes \boldsymbol{P}_{\boldsymbol{z}} \otimes A + \Sigma + \Sigma^T\right)\boldsymbol{e}_0 + 2(\boldsymbol{u},\boldsymbol{T}\boldsymbol{e})_{P,I}, \tag{3.11}$$

where $\boldsymbol{e}_0$ denotes the restriction of $\boldsymbol{e}$ to the boundary at $x = 0, i = 0$.

Remark: The error equation (3.11) completely determines the accuracy of the solution for long times. With a given truncation error (given by the scheme) we obtain a more or less favorable error development depending on how we choose our boundary conditions (choice of Σ).

The relations

$$((\boldsymbol{e},\boldsymbol{e})_{P,S})_t = 2||\boldsymbol{e}||_{P,S}(||\boldsymbol{e}||_{P,S})_t, \quad (\boldsymbol{e},\boldsymbol{T}\boldsymbol{e})_{P,I} \leq ||\boldsymbol{e}||_{P,I}||\boldsymbol{T}\boldsymbol{e}||_{P,I}, \tag{3.12}$$

inserted into (3.11) yields

$$(||\boldsymbol{e}||_{P,S})_t \leq C_0||\boldsymbol{e}||_{P,S} + C_1||\boldsymbol{T}\boldsymbol{e}||_{P,I}, \tag{3.13}$$

where

$$C_0 = \frac{\boldsymbol{e}_0^T\left(\boldsymbol{P}_y \otimes \boldsymbol{P}_z \otimes (A + \Sigma + \Sigma^T)\right)\boldsymbol{e}_0}{2||\boldsymbol{e}||_{P,S}^2}, \quad C_1 = \frac{||\boldsymbol{e}||_{P,I}}{||\boldsymbol{e}||_{P,S}}. \tag{3.14}$$

Remark: If C_0 vanish, we have the usual (essentially) linear error growth in time for hyperbolic problems (see for example [15],[13],[6]).

We can now state the other main result of this paper.

PROPOSITION 3.4. *If the boundary procedure in (3.2) is dissipative, i.e. $C_0 < 0$ "almost" everywhere in (3.13), then the error is bounded in time. It is bounded by the size of the truncation error.*

Proof: Lemma 2.3 applied to (3.13) leads to the result. ∎

Proposition 3.3 and 3.4 immediately leads to the following corollary.

COROLLARY 3.5. *Strong stability lead to a temporal error bound.*

4. Examples and numerical experiments. By performing numerical calculation we will check the validity of Proposition 3.4.

4.1. The one-way wave equation. The numerical approximation of (1.1) on the domain $[0, 1]$ including the penalty terms for the boundary condition has the general form given in (3.2). The one-dimensional approximation is given by inserting $\boldsymbol{P}_y = \boldsymbol{P}_z = \boldsymbol{I}_y = \boldsymbol{I}_z = 1$, $B = C = R = 0$ and $A = 1$. We get,

$$\boldsymbol{u}_t + P^{-1}Q\boldsymbol{u} = P^{-1}\sigma(u_0 - g(t))\boldsymbol{a_0} + \boldsymbol{F}, \quad t \geq 0, \tag{4.1}$$

where $\boldsymbol{a_0} = (1, 0, 0....0)^T$. Recall that F is a 1-periodic forcing function. The exact 1-periodic solution which we consider known in this example provide us with the boundary data $(g(t))$.

The energy method on the corresponding error equation yields directly

$$(||\boldsymbol{e}||_P)_t \leq -\lambda(t)||\boldsymbol{e}||_P + ||\boldsymbol{Te}||_P, \tag{4.2}$$

where $||\boldsymbol{e}||_P^2 = e^T Pe$ and $\lambda = (-(1 + 2\sigma)e_0^2 + e_N^2)/2||\boldsymbol{e}||_P^2$. By the choice $1 + 2\sigma \leq 0$ we get an error bound in time, see Proposition 3.4. To compare with the general formulation, note that $A^+ = 1, \Sigma = \sigma$ and $A^+ + \Sigma + \Sigma^T = 1 + 2\sigma$. The calculation using the scheme (4.1) and the scheme where the periodicity is imposed directly (details not shown here) is shown in Figure 1.1 and 4.1. The amplitude of the error is proportional to Δx^4 since we use a fourth order method.

Remark: The result in Figure 4.1 is truly remarkable. By mesh refinement, we can obtain an *arbitrarily* high accuracy at *any future time.*

4.2. The Maxwell's equations. The distribution of electromagnetic fields are described by the Maxwell's equations

$$\mu\frac{\partial H}{\partial t} = -\nabla \times E, \quad \epsilon\frac{\partial E}{\partial t} = \nabla \times H - J, \tag{4.3}$$

$$\nabla \cdot \epsilon E = \rho, \quad \nabla \cdot \mu H = 0, \tag{4.4}$$

combined with the equation of continuity $\rho_t + \nabla \cdot J = 0$. Here E is the electric field, H the magnetic field, J is the electric current density and ρ is the charge density. ϵ and μ are permittivity and permeability respectively.

With $J = 0$ we can formulate the two-dimensional version of (4.3) in TE (transverse electric) mode as

$$Su_t + Au_x + Bu_y = 0, \tag{4.5}$$

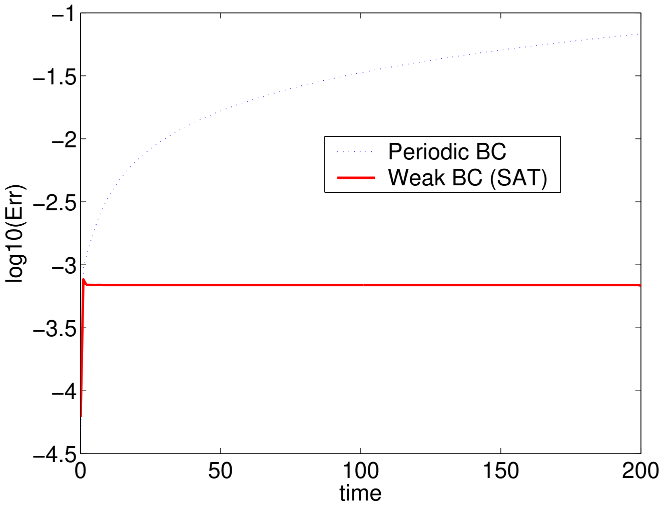

FIGURE 4.1. *The error for a very long time.*

where we consider the domain $(x, y) \in [-1, 1] \times [0, 1]$ and

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, B = -\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, S = \begin{bmatrix} \mu & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{bmatrix}, u = \begin{bmatrix} H_z \\ E_x \\ E_y \end{bmatrix}$$

4.2.1. The continuous problem. By using the energy method on eq.(4.5) and using that the matrices A and B are symmetric, leads to

$$\frac{d}{dt}\|u\|_S^2 = -\int_0^1 u^T A u|_{x=-1}^{x=1} dy - \int_{-1}^1 u^T B u|_{y=0}^{y=1} dx \tag{4.6}$$

where $\|u\|_S^2 = u^T S u$. The problem (4.5) augmented with initial condition is well-posed if we specify the E_y component at $x = -1$ and $x = 1$ and the E_x component at $y = 0$ and $y = 1$.

Remark: In terms of an error equation where the boundary data and forcing function is zero this means that the time derivative of the norm of the error is zero. With a non-zero forcing function included, we would get the classical linear error growth.

The problem (4.5) augmented with initial condition is strongly well-posed if we specify the ingoing characteristic variable at each boundary.

This yield the estimate,

$$\frac{d}{dt}\|u\|_S^2 = \int_0^1 (u^T A^- u + g^T A^+ g)_{x=-1} dy - \int_0^1 (u^T A^+ u + g^T A^- g)_{x=1} dy$$
$$+ \int_{-1}^1 (u^T B^- u + g^T B^+ g)_{y=0} dx - \int_{-1}^1 (u^T B^+ u + g^T B^- g)_{y=1} dx,$$

where $A^{+,-} = X_A \Lambda_A^{+,-} X_A^T$ and $B^{+,-} = X_B \Lambda_B^{+,-} X_B^T$ and

$$X_A = \begin{bmatrix} -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix}, \quad \Lambda_A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$X_B = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \Lambda_B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Remark: In terms of an error equation where the boundary data and forcing function is zero this means the time derivative of the norm of the error is negative semi-definite. With a non-zero forcing function included, we will get an error bound, see Proposition 2.4 above.

4.2.2. Calculations. The numerical approximation of (4.5) including the penalty terms for boundary conditions has the general form in (3.2). The more restricted two-dimensional approximation is obtained with $\boldsymbol{P}_z = \boldsymbol{I}_z = 1$ and $C = 0$. The different types of boundary conditions are obtained by varying the form of the penalty matrix Σ. For the time integration we use the classical fourth order Runge-Kutta method with a very small time-step.

We will compute the wave propagation over a material discontinuity. All details (especially concerning the treatment of the discontinuity itself) for this calculation is given in [15]. Here we will focus on the effect of the outer boundary conditions. We have an exact solution which is a plane wave that propagates over the discontinuity. That solution provides us with boundary data and we impose the two different sets of boundary conditions mentioned above.

By using the tangential boundary condition at the boundaries, i.e the E_y component at $x = -1$ and $x = 1$ and the E_x component at $y = 0$ and $y = 1$ we obtain a stable approximation according to Proposition 3.2. For example, at the boundary $x = -1$ that means that we have

$$\Sigma = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \text{and consequently} \quad \boldsymbol{A} + \Sigma + \Sigma^T = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

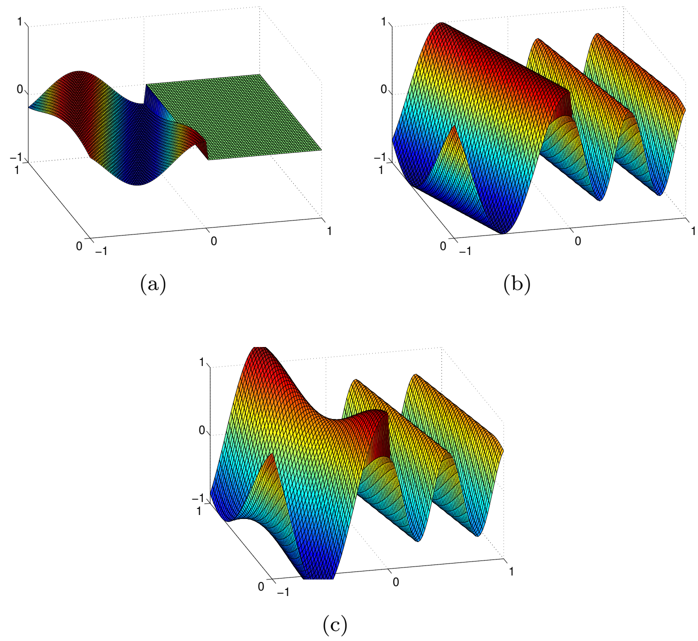

FIGURE 4.2. *(a) the wave propagating to the left, (b) the wave propagating to the right, (c) the total wave,* $\Theta_i = \frac{\pi}{6}$.

which yields $C_0 = 0$. An error bounded scheme cannot be expected, see Proposition 3.4.

By specifying the ingoing characteristic variables at the boundaries, i.e the $(E_y + H_z)/\sqrt{2}$ at $x = -1$ and $(E_y - H_z)/\sqrt{2}$ at $x = 1$ and $(E_x - H_z)/\sqrt{2}$ at $y = 0$ and $(E_x + H_z)/\sqrt{2}$ at $y = 1$ we obtain a strongly stable approximation according to Proposition 3.3. For example, at the boundary $x = -1$ we have (by choosing $\sigma = -1$)

$$\Sigma = -\boldsymbol{A}^+ = -\frac{1}{2}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{A} + \Sigma + \Sigma^T = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

which yields $C_0 < 0$. An error bounded scheme can be expected, see Proposition 3.4.

As an example we show what happens when the wave hits the discontinuity with an angle of incidence $\Theta_i = \frac{\pi}{6}$. A significant reflection at the interface between the two medium occurs, see figure 4.2. In [15] it was shown that the accuracy in space is of design order.

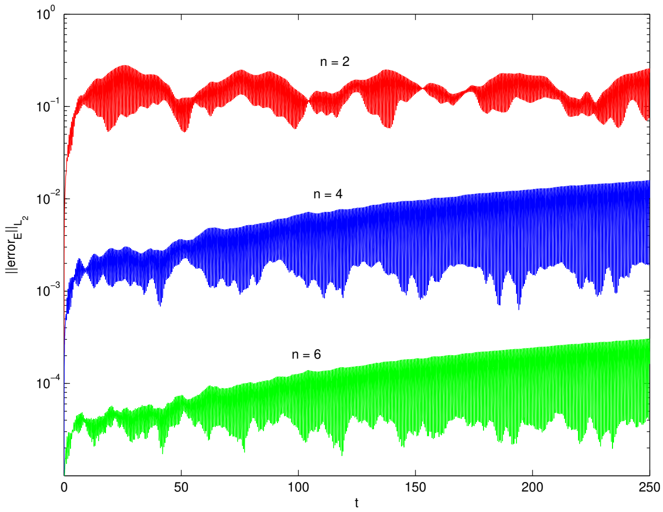

FIGURE 4.3. *The error using non characteristic boundary conditions.*

Next we study the error development in time. If we specify the tangential electrical fields at the boundaries and measure the error as a function of time we obtain the classical linear error growth shown in Figure 4.3. By specifying the characteristic variables we obtain the error bounded result shown in Figure 4.4. The results in Figure 4.3 and Figure 4.4 are computed with 2,4,6 orders of accuracy. The error levels using the characteristic boundary conditions are significantly lower than the one's obtained for non characteristic boundary conditions.

Remark: The result in Figure 4.4 is remarkable. By mesh refinement, we can obtain an *arbitrarily* high accuracy at *any future time*.

Remark: The error boundedness in this paper is due to the boundary treatment only, the internal discretization is not important. The same result is obtained using the spectral element code USEMe, see [9],[10].

5. Conclusions. We have shown that an error bound can be obtained for a general time-dependent hyperbolic problem if a sufficiently dissipative (or sufficiently stable) boundary procedure can be constructed.

Two examples, the one-dimensional one-way wave equation and the two-dimensional Maxwell's equations in transverse mode were studied and numerical calculations were performed.

The results of the calculations using the errorbounded schemes are remarkable. With accurate boundary data and mesh refinement, we can obtain an *arbitrarily* high accuracy at *any future time*.

The error boundedness in this paper is due to the boundary treat-

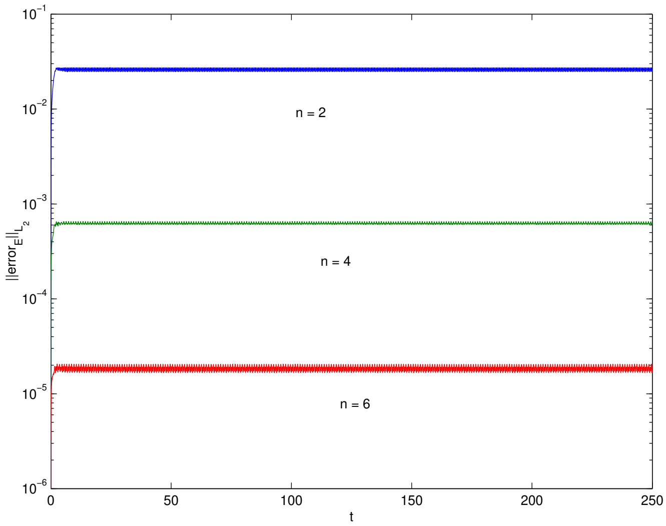

FIGURE 4.4. *The error using characteristic boundary conditions.*

ment only, the internal discretization (finite differences was used) is not important. The same result is obtained using a spectral element code.

Appendix A. The Kronecker product.

DEFINITION A.1. *Let $\boldsymbol{A}$ be a $p \times q$ matrix and let $\boldsymbol{B}$ be an $m \times n$ matrix, then*

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{pmatrix} a_{0,0}\boldsymbol{B} & \cdots & a_{0,q-1}\boldsymbol{B} \\ \vdots & \ddots & \vdots \\ a_{p-1,0}\boldsymbol{B} & \cdots & a_{P-1,0}\boldsymbol{B} \end{pmatrix}.$$

The $p \times q$ block matrix $\boldsymbol{A} \otimes \boldsymbol{B}$ is called a Kronecker product. Let $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{D}$ be matrices of arbitrary sizes, such that the specified operations are defined, then (see [12]),

$$\begin{aligned} (\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) &= (\boldsymbol{AC}) \otimes (\boldsymbol{BD}) \\ (\boldsymbol{A} + \boldsymbol{B}) \otimes \boldsymbol{C} &= \boldsymbol{A} \otimes \boldsymbol{C} + \boldsymbol{B} \otimes \boldsymbol{C} \\ (\boldsymbol{A} \otimes \boldsymbol{B})^T &= \boldsymbol{A}^T \otimes \boldsymbol{B}^T \\ (\boldsymbol{A} \otimes \boldsymbol{B})^{-1} &= \boldsymbol{A}^{-1} \otimes \boldsymbol{B}^{-1}. \end{aligned}$$

REFERENCES

[1] S. Abarbanel, A. Ditkowski, and B. Gustafsson, *On error bounds of finite difference approximations to partial differential equations - temporal behavior and rate of convergence*, Journal of Scientific Computing, 15 (2000), pp. 79–116.

[2] S. Abarbanel and D. Gottlieb, *Optimal time splitting for two and three dimensional Navier-Stokes equations with mixed derrivatives*, Journal of Computational Physics, 41 (1981), pp. 1–33.

[3] S. Abrabanel and A. Chertock, *Strict stability of high-order compact implicit finite-difference schemes: The role of boundary conditions for hyperbolic pdes, I*, Journal of Computational Physics, 160 (2000), pp. 42–66.

[4] S. Abrabanel, A. Chertock, and A. Yefet, *Strict stability of high-order compact implicit finite-difference schemes: The role of boundary conditions for hyperbolic pdes, II*, Journal of Computational Physics, 160 (2000), pp. 67–87.

[5] M.H. Carpenter, D. Gottlieb, and S. Abarbanel, *Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: Methodology and application to high-order compact schemes*, Journal of Computational Physics, 129 (1994).

[6] A. Ditkowski, K. Dridi, and J. S. Hesthaven, *Convergent cartesian grid methods for Maxwell's equations in complex geometries*, Journal of Computational Physics, 170 (2001), pp. 39–80.

[7] B. Gustafsson, H.-O. Kreiss, and J. Oliger, *Time dependent problems and difference methods*, John Wiley & Sons, Inc., 1995.

[8] J.S. Hestahaven and D. Gottlieb, *A stable penalty method for the compressible Navier-Stokes equations, i. Open boundary conditions*, SIAM Journal of Scientific Computing, 17 (1996).

[9] J.S. Hesthaven and T. Warburton, *High-order nodal methods on unstructured grids. i. time-domain solution of Maxwell's equations*, Journal of Computational Physics, 181 (2002), pp. 186–221.

[10] ———, *High-order nodal methods on discontinuous galerkin methods for the Maxwell's equations*, Royal Soc. London Ser A 362, (2004), pp. 493–524.

[11] H.-O. Kreiss and G. Scherer, *Finite element and finite difference methods for hyperbolic partial differential equations, in: C. De Boor (Ed.), Mathematical Aspects of Finite Elements in Partial Differential Equation*, Academic Press, New York, 1974.

[12] C. Van Loan, *Computational Frameworks for the Fast Fourier Transform*, SIAM, 1992.

[13] J. Nordström and M. H. Carpenter, *High-order finite difference methods, multidimensional linear problems and curvilinear coordinates*, Journal of Computational Physics, 173 (2001), pp. 149–174.

[14] J. Nordström, K. Forsberg, C. Adamsson, and P. Eliasson, *Finite volume methods, unstructured meshes and strict stability*, Applied Numerical Mathematics, 45 (2003), pp. 453–473.

[15] J. Nordström and R. Gustafsson, *High order finite difference approximations of electromagnetic wave propagation close to material discontinuities*, Journal of Scientific Computing, 18 (2003), pp. 215–234.