

## A New Method of Imposing Boundary Conditions in Pseudospectral Approximations of Hyperbolic Equations\*

By D. Funaro and D. Gottlieb

**Abstract.** A new method to impose boundary conditions for pseudospectral approximations to hyperbolic equations is suggested. This method involves the collocation of the equation at the boundary nodes as well as satisfying boundary conditions. Stability and convergence results are proven for the Chebyshev approximation of linear scalar hyperbolic equations. The eigenvalues of this method applied to parabolic equations are shown to be real and negative.

**Introduction.** The common practice in applying pseudospectral methods to partial differential equations is to satisfy the equation at the interior nodes and to impose the boundary condition at the boundary. This procedure does not take into consideration that the differential equation is satisfied at points arbitrarily close to the boundary. In [4], one of the authors discussed the advantages of imposing a combination of boundary conditions and the equation itself at the boundary nodes, for Chebyshev approximations of the Laplace equation with Neumann conditions. Here we analyze the same idea applied to the linear hyperbolic equation

$$\begin{cases} u_t = u_x, & |x| \leq 1, t > 0, \\ u(x, 0) = f(x), \\ u(1, t) = g(t). \end{cases}$$

We assume that the collocation points are the Gauss-Lobatto Chebyshev quadrature nodes, namely:  $x_j = \cos(\pi j/N)$ ,  $0 \leq j \leq N$ . The stability of the method, with the commonly used boundary treatment, i.e., imposing  $u_N(1, t) = g(t)$ , was analyzed in [10]. Here we show the convergence of the method for the new boundary treatment, namely

$$\frac{\partial u_N}{\partial t}(1, t) - \frac{\partial u_N}{\partial x}(1, t) + \alpha(u_N(1, t) - g(t)) = 0,$$

where  $\alpha$  is positive and large enough. A preliminary theoretical discussion in Section 1 and numerical experiments in the last section show the effectiveness of the method. In Section 3 we use the results obtained for the hyperbolic equations to show that for the heat equation the second-derivative matrices, corresponding to

---

Received July 9, 1987; revised November 20, 1987.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 65N35, 65L10.

*Key words and phrases.* Spectral approximations, hyperbolic equations.

\*This research has been supported in part by AFOSR Grant 85-0303. Also, the research has been supported by the National Aeronautics and Space Administration under NASA Contract No. NAS1-18107 while the authors were in residence at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA 23665.

the Neumann conditions with the new approach, have real and negative eigenvalues. The analogous result for the classical way to impose boundary conditions was previously proven in [8].

The results reported here are only the first step in the implementation of this new method to systems of equations.

**1. Description of the New Method.** In order to illustrate the new method of imposing boundary conditions and to explain what can be gained by this technique, we first treat the following time-independent problem

$$(1.1) \quad \begin{cases} U_x = f, & |x| \leq 1, \\ U(1) = 0, \end{cases}$$

where  $f \in C^s([-1, 1])$  is given ( $s \geq 0$ ).

In the standard pseudospectral Chebyshev method (see for instance [6]), we seek a polynomial of degree  $N$ , say  $v_N$ , such that

$$(1.2) \quad \begin{cases} \text{(a)} & \frac{dv_N}{dx}(x_j) = f(x_j), \quad j = 1, \dots, N; \\ \text{(b)} & v_N(1) = 0, \end{cases}$$

where  $x_j = \cos(\pi j/N)$ ,  $j = 0, 1, \dots, N$ , are the Gauss-Lobatto Chebyshev nodes in  $[-1, 1]$ . In order to determine  $v_N$  from (1.2),  $v_N(x)$  is expressed by its unknown point values  $v_N(x_j)$  using the Lagrange interpolation polynomial

$$v_N(x) = \sum_{k=0}^N v_N(x_k)g_k(x),$$

where

$$g_k(x) = \frac{-(-1)^k(1-x^2)T'_N(x)}{c_k N^2(x-x_k)},$$

with  $T_N(x) \equiv \cos N\theta$ ,  $\cos \theta \equiv x$ , so that

$$T'_N = N \frac{\sin N\theta}{\sin \theta}.$$

Here,  $T_N$  is the  $N$ th-degree Chebyshev polynomial and  $c_j = 1$  if  $1 \leq j \leq N - 1$ , while  $c_0 = c_N = 2$ . Therefore,

$$\frac{dv_N}{dx}(x_j) = \sum_{k=0}^N v_N(x_k) \frac{dg_k}{dx}(x_j), \quad j = 1, \dots, N.$$

Upon substituting the above relations in (1.2), we get a linear system of equations for the point values  $v_N(x_k)$ . We note that in (1.1) the differential equation holds in any arbitrary neighborhood of the boundary, whereas in (1.2) we did not require that the equation also was satisfied at  $x_0 = 1$ . We propose now another procedure that takes into account the differential equation at the boundary as well as the boundary condition.

In our new method, we seek an  $N$ th-degree polynomial  $u_N$  such that

$$(1.3) \quad \begin{cases} \text{(a)} & \frac{du_N}{dx}(x_j) = f(x_j), \quad j = 1, \dots, N; \\ \text{(b)} & \frac{du_N}{dx}(1) - \alpha u_N(1) = f(1), \end{cases}$$

where  $\alpha > 0$  is a suitable constant depending on  $N$ , to be determined later. By writing the equality (1.3)(b) as

$$\frac{1}{\alpha} \left( \frac{du_N}{dx} - f \right) (1) = u_N(1),$$

we note that (1.2) is obtained from (1.3) by letting  $\alpha \rightarrow +\infty$ . We remark that the solution of (1.3) satisfies neither the boundary condition nor the equation at  $x = 1$ ; if the method converges both will be satisfied as  $N \rightarrow +\infty$ .

To show the advantage of the new procedure, we give in Figure 1.1 the plot of the error

$$(1.4) \quad E = E(\alpha) = \left( \frac{\pi}{N} \sum_{j=1}^N (U - u_N)^2(x_j) \frac{1}{c_j} \right)^{1/2}$$

multiplied by  $10^5$  versus  $\alpha$ , for  $f(x) = \sin(x - 1)$  and  $N = 8$ . The point  $x_0$  is not taken into consideration in the sum because the exact solution is known there. It is clear from the figure that  $E(\alpha)$  is not monotone in  $\alpha$  and there exists  $\alpha = \alpha_{\min}$  which minimizes  $E$ . In particular we have  $E(\alpha_{\min}) < E(+\infty)$ . Further experiments indicate that, in terms of  $N$ ,  $\alpha_{\min}$  increases like  $N^2$ .

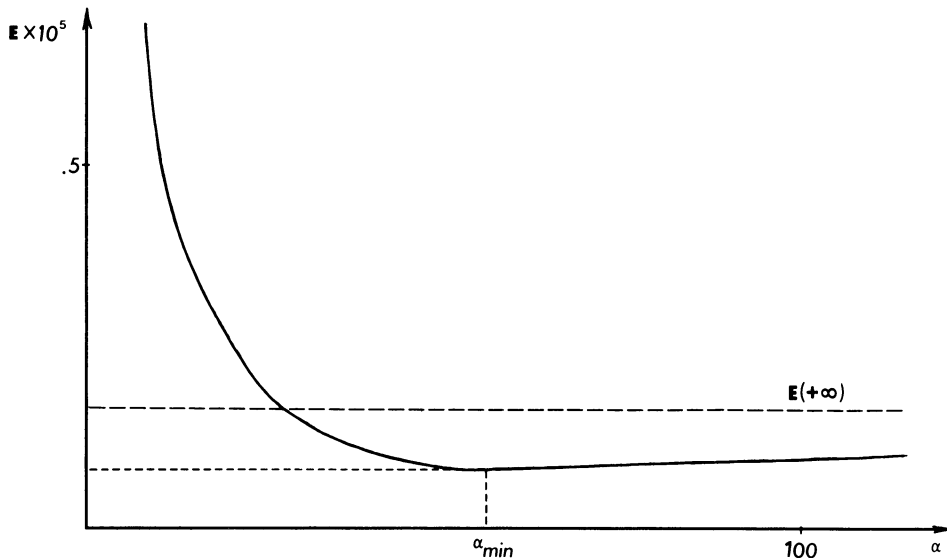


FIGURE 1.1  
*Behavior of the error versus  $\alpha$ .*

We would like to explain why the procedure (1.3) should be, in general, better than (1.2). We start by noting that if  $f$  is a polynomial of degree  $N - 1$  at most, then both (1.2)(a) and (1.3)(a) hold, not only at the grid points  $x_j$ , but for every  $x$  since both sides of the equations are polynomials of degree  $N - 1$ . In particular  $(du_N/dx)(1) = f(1)$ , thus by (1.3)(b) we get  $u_N(1) = 0$ , leading to the conclusion that  $u_N(x) = v_N(x), \forall x$ . Suppose now that  $f$  is a polynomial of degree  $N$ . We can assume, because of the linearity, that

$$(1.5) \quad f(x) = \frac{(1+x)T'_N(x)}{2N^2}.$$

Hence,  $f(x_j) = 0$ ,  $j = 1, \dots, N$ , and  $f(1) = 1$ . Any other polynomial, up to a constant factor, can be obtained from (1.5) by adding some suitable polynomial of lower degree. In this case it is easily verified that the solution  $U$  of (1.1) is given by

$$(1.6) \quad U(x) = \frac{1}{2N^2} \left[ \frac{1}{2} \frac{N}{N+1} T_{N+1}(x) + T_N(x) + \frac{1}{2} \frac{N}{N-1} T_{N-1}(x) - \frac{2N^2 - 1}{N^2 - 1} \right].$$

It is clear that the solution of (1.2) is

$$(1.7) \quad v_N(x) = 0, \quad \forall x.$$

On the other hand, the solution of (1.3)(a) is a constant and from (1.3)(b) we get

$$(1.8) \quad u_N(x) = -\frac{1}{\alpha}, \quad \forall x.$$

With  $1/\alpha \equiv \beta$ , the error is given by

$$(1.9) \quad E = \left( \frac{\pi}{N} \sum_{j=1}^N (U(x_j) + \beta)^2 \frac{1}{c_j} \right)^{1/2}.$$

To minimize  $E$  one has to choose  $\beta$  as the negative mean of  $U$ , namely

$$\beta_{\min} = -\frac{\pi}{N} \sum_{j=1}^N U(x_j) \frac{1}{c_j} \bigg/ \frac{\pi}{N} \sum_{j=1}^N \frac{1}{c_j} = -\sum_{j=0}^N U(x_j) \frac{1}{c_j} \bigg/ \sum_{j=1}^N \frac{1}{c_j},$$

and an easy calculation shows that

$$(1.10) \quad \alpha_{\min} = \frac{1}{\beta_{\min}} = \frac{2N^2 - N}{2N^2 - 1} (N^2 - 1) \sim N^2.$$

This explains the behavior of  $\alpha_{\min}$  as a function of  $N$ .

In Table 1.1, we summarize the results of another experiment. This time we chose  $f(x) = -\frac{3}{2}(1-x)^{1/2}$ , with the boundary condition  $U(1) = 1$ , so that the solution was  $U(x) = (1-x)^{3/2} + 1$ . We have tried the two different ways of imposing boundary conditions, i.e.,

$$(1.11) \quad v_N(1) = 1,$$

$$(1.12) \quad \frac{du_N}{dx} - \alpha_{\min}(u_N(1) - 1) = f(1) \quad (\text{where } \alpha_{\min} \text{ is given by (1.10)}),$$

and we varied the number of grid points  $N$ .

TABLE 1.1

*Comparison of the errors between the two ways of imposing boundary conditions.*

$N$	Condition (1.11)	Condition (1.12)
2	0.281837	0.243315
4	0.338991E-01	0.183894E-01
6	0.995738E-02	0.478991E-02
8	0.418410E-02	0.190129E-02

**2. The Time-Dependent Problem.** In this section we show how to apply the new procedure of setting the boundary conditions, described in the previous section, to a scalar hyperbolic equation. In a future paper, we will discuss the case

of a system of hyperbolic equations. An analysis of the convergence of this method will be carried out for Chebyshev approximations.

Consider the equation

$$(2.1) \quad \begin{cases} U_t = U_x, & |x| \leq 1, t > 0, \\ U(1, t) = g(t), \\ U(x, 0) = f(x). \end{cases}$$

The pseudospectral semidiscrete approximation to (2.1) suggested in this paper involves seeking a polynomial  $u_N$  of degree at most  $N$  such that

$$(2.2) \quad \begin{cases} \frac{\partial u_N}{\partial t} = \frac{\partial u_N}{\partial x} & \text{at } x = x_j, j = 1, \dots, N, \forall t > 0, \\ \frac{\partial u_N}{\partial t}(1, t) = \frac{\partial u_N}{\partial x}(1, t) - \alpha(u_N(1, t) - g(t)), \\ u_N(x_j, 0) = f(x_j), & j = 0, \dots, N. \end{cases}$$

The choice of the nodes  $\{x_j\}$  determines the particular spectral method. For example, the points

$$(2.3) \quad x_j = \cos \frac{\pi j}{N}, \quad j = 0, 1, \dots, N,$$

determine the usual pseudospectral Chebyshev method, whereas the points

$$(2.4) \quad x_j = \cos \frac{\pi j}{N+1}, \quad j = 0, 1, \dots, N,$$

determine a different version (see [7]). The pseudospectral Legendre method is defined by choosing  $x_j$  to be the extrema of the  $N$ th-degree Legendre polynomial.

We would like to show here the convergence of the solution  $u_N(x, t)$  of (2.2) to  $U(x, t)$  defined in (2.1) when  $N \rightarrow +\infty$ , in the case of the Chebyshev method defined by (2.3). The stability proof for the Chebyshev method (2.2) and (2.3) for  $\alpha = \infty$  is discussed in [10]. We use the same basic ideas to get directly a convergence proof, for  $\alpha \neq \infty$ . The proof here is presented in detail since it will serve as a reference for our future work discussing systems. We start with the following preliminary results.

LEMMA 2.1. *Let  $u_N(x, t)$  be the solution of (2.2) when  $x_j$  are given by (2.3); then*

$$(2.5) \quad \frac{\partial u_N}{\partial t} = \frac{\partial u_N}{\partial x} + \tau \frac{(1+x)T'_N(x)}{2N^2},$$

where  $\tau = -\alpha(u_N(1, t) - g(t))$ .

*Proof.* It is sufficient to note that (2.5) exactly coincides with (2.2), when evaluated at the collocation nodes.  $\square$

We define now  $P_M U$  as the polynomial interpolating  $U$  at the points  $\cos(\pi j/M)$ ,  $j = 0, 1, \dots, M$ . Note in particular that  $(P_M U)(1, t) = U(1, t)$ , for any  $M$ . We are ready to write the error equation.

LEMMA 2.2. *Let  $\varepsilon_N(x, t) = u_N(x, t) - P_{N-3}U(x, t)$ ; then*

$$(2.6) \quad \begin{cases} \frac{\partial \varepsilon_N}{\partial t} = \frac{\partial \varepsilon_N}{\partial x} - \alpha \frac{(1+x)T'_N(x)}{2N^2} \varepsilon_N(1, t) + Q(x, t), \\ \varepsilon_N(x, 0) = P_N f - P_{N-3} f, \end{cases}$$

where  $Q(x, t)$  is a polynomial of degree  $N - 3$  in  $x$ , given by

$$Q(x, t) = \frac{\partial}{\partial x}(P_{N-3}U) - P_{N-3} \left( \frac{\partial}{\partial x}U \right).$$

*Proof.* We apply  $P_{N-3}$  to (2.1) to get

$$(2.7) \quad \frac{\partial(P_{N-3}U)}{\partial t} = \frac{\partial(P_{N-3}U)}{\partial x} - Q(x, t) - \alpha[(P_{N-3}U)(1, t) - g(t)] \frac{(1+x)T'_N(x)}{2N^2},$$

with the initial condition  $[P_{N-3}U]_{t=0} = P_{N-3}f$ . In fact, note that, since  $(P_{N-3}U)(1, t) = g(t)$ , the last term that was introduced in (2.7) is zero. Hence, (2.6) follows from (2.7) and (2.5).  $\square$

Next we will show that  $\varepsilon_N(x, t)$  tends to zero as  $N$  increases. The proof will be based on a careful energy estimate for (2.6). For this, we need the following lemmas.

LEMMA 2.3. *Let  $w(x) = \sum_{k=0}^{4N-1} b_k T_k(x)$ ; then*

$$(2.8) \quad \frac{\pi}{N} \sum_{j=0}^N \frac{w(x_j)}{c_j} = \int_{-1}^1 \frac{w(x)}{\sqrt{1-x^2}} dx + \pi b_{2N},$$

where  $c_0 = c_N = 2$  and  $c_k = 1$  for  $0 < k < N$ .

*Proof.* We test (2.8) for  $T_k$ ,  $k = 0, \dots, 4N - 1$ . If  $0 \leq k \leq 2N - 1$ , (2.8) is a well-known quadrature formula (see [9, p. 50]); if  $k = 2N$ , it is a trivial result by noticing that  $T_{2N}(x_j) = 1$ ; if  $2N + 1 \leq k \leq 4N - 1$ , then by writing  $T_{2N+m} = 2T_N T_m - T_{2N-m}$ , (2.8) follows easily from the orthogonality of the Chebyshev polynomials.  $\square$

LEMMA 2.4. *Let  $v(x) = \sum_{k=0}^N a_k T_k(x)$ ; then*

$$(2.9) \quad \begin{aligned} & \frac{\pi}{N} \sum_{j=0}^N \frac{1}{c_j} (1+x_j)(1-\beta x_j)v(x_j)v_x(x_j) \\ &= \int_{-1}^1 \frac{(1+x)(1-\beta x)vv_x}{\sqrt{1-x^2}} dx + \frac{\pi}{2} \left[ (1-\beta)Na_N^2 - \beta \frac{2N-1}{2} a_N a_{N-1} \right], \end{aligned}$$

for any  $\beta$  real.

*Proof.* The result is an application of the previous lemma (see also [10]).  $\square$

LEMMA 2.5. *Let  $\varepsilon_N(x, t)$  be defined by (2.6). Suppose that*

$$\varepsilon_N(x, t) = \sum_{k=0}^N a_k T_k(x);$$

then

$$(2.10) \quad \frac{d}{dt}(2a_N - a_{N-1})^2 = -4N(2a_N^2 - a_N a_{N-1}).$$

*Proof.* We can argue as in [10] using (2.6) and the fact that  $Q(x, t)$  is a polynomial of degree  $N - 3$ .  $\square$

LEMMA 2.6. Let  $\varepsilon_N(x, t)$  be defined by (2.6); then

$$\begin{aligned}
 (2.11) \quad & \frac{\pi}{N} \sum_{j=0}^N \frac{1}{c_j} (1+x_j)(1-\beta x_j) \varepsilon_N(x_j, t) \frac{\partial \varepsilon_N}{\partial x}(x_j, t) \\
 & = \int_{-1}^1 \frac{(1+x)(1-\beta x)}{\sqrt{1-x^2}} \varepsilon_N \frac{\partial \varepsilon_N}{\partial x} dx - \pi \beta \frac{2N-1}{16N} \frac{d}{dt} (2a_N - a_{N-1})^2 \\
 & \quad - \frac{\pi}{2} N \left( 3\beta - 1 - \frac{\beta}{N} \right) a_N^2.
 \end{aligned}$$

*Proof.* Combine the results of Lemma 2.4 and Lemma 2.5.  $\square$

THEOREM 2.1. Define

$$\begin{aligned}
 (2.12) \quad \|\varepsilon_N\|^2 & = \frac{\pi}{N} \sum_{j=0}^N \frac{1}{c_j} (1+x_j) \left( 1 - \frac{1}{2}x_j \right) \varepsilon_N^2(x_j, t) \\
 & \quad + \pi \frac{2N-1}{16N} (2a_N - a_{N-1})^2,
 \end{aligned}$$

and let

$$\gamma_K = \frac{1}{2K} \sum_{m=1}^K \frac{1-z_m+z_m^2}{1-z_m},$$

where  $z_m$  are the zeros of  $T_K$  and  $K$  is chosen such that  $K \geq N+1$ . Then we have

$$\begin{aligned}
 (2.13) \quad & \frac{1}{2} \frac{d}{dt} \|\varepsilon_N\|^2 + \frac{\pi}{2} \left( \frac{\alpha}{N} - \gamma_K \right) \varepsilon_N^2(1, t) \\
 & \leq 2 \int_{-1}^1 (1+x) \left( 1 - \frac{x}{2} \right)^2 Q^2(x, t) \sqrt{1-x^2} dx, \quad \forall t > 0.
 \end{aligned}$$

*Proof.* We evaluate the equation in (2.6) at the points  $x_j$ , then multiply by

$$\frac{\pi}{N c_j} (1+x_j) \left( 1 - \frac{1}{2}x_j \right) \varepsilon_N(x_j, t)$$

and sum up over  $j = 0, \dots, N$  to get

$$\begin{aligned}
 (2.14) \quad & \frac{\pi}{N} \sum_{j=0}^N \frac{1}{c_j} (1+x_j) \left( 1 - \frac{1}{2}x_j \right) \varepsilon_N(x_j, t) \frac{\partial}{\partial t} \varepsilon_N(x_j, t) \\
 & = \frac{\pi}{N} \sum_{j=0}^N \frac{1}{c_j} (1+x_j) \left( 1 - \frac{1}{2}x_j \right) \varepsilon_N(x_j, t) \frac{\partial}{\partial x} \varepsilon_N(x_j, t) \\
 & \quad - \frac{\pi \alpha}{2N} \varepsilon_N^2(1, t) + \frac{\pi}{N} \sum_{j=0}^N \frac{1}{c_j} (1+x_j) \left( 1 - \frac{1}{2}x_j \right) \varepsilon_N(x_j, t) Q(x_j, t).
 \end{aligned}$$

The right-hand side of (2.14) is composed of three terms. We start by estimating the last term. First, we realize that the polynomial  $(1+x)(1-\frac{1}{2}x)\varepsilon_N Q$  is of degree  $2N-1$  and therefore, by Lemma 2.3, we have

$$\begin{aligned}
 & \frac{\pi}{N} \sum_{j=0}^N \frac{1}{c_j} (1+x_j) \left( 1 - \frac{1}{2}x_j \right) \varepsilon_N(x_j, t) Q(x_j, t) \\
 & = \int_{-1}^1 (1+x) \left( 1 - \frac{1}{2}x \right) \varepsilon_N(x, t) Q(x, t) \frac{dx}{\sqrt{1-x^2}}.
 \end{aligned}$$

Upon using the Gauss quadrature formula based on  $z_m, m = 1, \dots, K$ , one gets

$$\begin{aligned}
 & \left| \int_{-1}^1 (1+x) \left(1 - \frac{1}{2}x\right) \varepsilon_N Q \frac{dx}{\sqrt{1-x^2}} \right| \\
 &= \left| \frac{\pi}{K} \sum_{m=1}^K (1+z_m) \left(1 - \frac{1}{2}z_m\right) \varepsilon_N(z_m, t) Q(z_m, t) \right| \\
 &\leq \frac{\pi}{2K} \sum_{m=1}^K \frac{1-z_m+z_m^2}{2(1-z_m)} \varepsilon_N^2(z_m, t) \\
 &\quad + \frac{\pi}{2K} \sum_{m=1}^K \frac{2(1-z_m)}{1-z_m+z_m^2} (1+z_m)^2 \left(1 - \frac{1}{2}z_m\right)^2 Q^2(z_m, t) \\
 &\leq \frac{\pi}{2K} \sum_{m=1}^K \frac{1-z_m+z_m^2}{2(1-z_m)} \varepsilon_N^2(z_m, t) \\
 &\quad + \frac{2\pi}{K} \sum_{m=1}^K (1-z_m)(1+z_m)^2 \left(1 - \frac{1}{2}z_m\right)^2 Q^2(z_m, t) \\
 &= \frac{\pi}{2K} \sum_{m=1}^K \frac{1-z_m+z_m^2}{2(1-z_m)} \varepsilon_N^2(z_m, t) \\
 &\quad + 2 \int_{-1}^1 (1+x) \left(1 - \frac{1}{2}x\right)^2 Q^2(x, t) \sqrt{1-x^2} dx.
 \end{aligned}$$

For the first term in the right-hand side of (2.14) we use the result in Lemma 2.6 with  $\beta = \frac{1}{2}$ . Therefore, by (2.12) and the previous estimate, we get

$$\begin{aligned}
 \frac{1}{2} \frac{d}{dt} \|\varepsilon_N\|^2 &\leq \frac{1}{2} \int_{-1}^1 (1+x) \left(1 - \frac{1}{2}x\right) \frac{\partial}{\partial x} [\varepsilon_N^2(x, t) - \varepsilon_N^2(1, t)] \frac{dx}{\sqrt{1-x^2}} \\
 &\quad - \frac{\pi}{2} N \left(\frac{1}{2} - \frac{1}{2N}\right) a_N^2 - \frac{\pi\alpha}{2N} \varepsilon_N^2(1, t) \\
 (2.15) \quad &\quad + \frac{\pi}{2K} \sum_{m=1}^K \frac{1-z_m+z_m^2}{2(1-z_m)} \varepsilon_N^2(z_m, t) \\
 &\quad + 2 \int_{-1}^1 (1+x) \left(1 - \frac{1}{2}x\right)^2 Q^2(x, t) \frac{dx}{\sqrt{1-x^2}}.
 \end{aligned}$$

Integration by parts for the first term in the right-hand side of (2.15) yields

$$\begin{aligned}
 & \frac{1}{2} \int_{-1}^1 (1+x) \left(1 - \frac{1}{2}x\right) \frac{\partial}{\partial x} [\varepsilon_N^2(x, t) - \varepsilon_N^2(1, t)] \frac{dx}{\sqrt{1-x^2}} \\
 &= -\frac{1}{2} \int_{-1}^1 \frac{1-x+x^2}{2(1-x)} [\varepsilon_N^2(x, t) - \varepsilon_N^2(1, t)] \frac{dx}{\sqrt{1-x^2}} \\
 &= -\frac{1}{2} \left\{ \frac{\pi}{K} \sum_{m=1}^K \frac{1-z_m+z_m^2}{2(1-z_m)} \varepsilon_N^2(z_m, t) - \gamma_K \pi \varepsilon_N^2(1, t) \right\},
 \end{aligned}$$

where we noted that the last integrand is a polynomial of degree  $2N + 1 \leq 2K - 1$  and therefore the Gauss quadrature formula is exact. Going back to (2.15), one finally gets (2.13).  $\square$



*Remark 2.1.* It can be shown that  $\|\cdot\|$  defined in (2.12) is actually a norm. In fact, it is possible to find a positive constant  $c$ , independent of  $N$ , such that

$$\|\varepsilon_N\|^2 \geq c \int \varepsilon_N^2(1+x) \frac{dx}{\sqrt{1-x^2}},$$

for every polynomial  $\varepsilon_N$  of degree at most  $N$ .  $\square$

Finally, by integrating (2.13), we get the main result of this section.

**THEOREM 2.2.** *Let  $\alpha$  be such that  $\pi(\alpha/N - \gamma_{N+1}) > C^*$ , where  $C^*$  does not depend on  $N$ ; then we have*

$$\begin{aligned} \|\varepsilon_N(\cdot, t)\|^2 + C^* \int_0^t \varepsilon_N^2(1, \tau) d\tau \\ (2.16) \quad \leq \|P_N f - P_{N-3} f\|^2 \\ + 4 \int_0^t \int_{-1}^1 (1+x) \left(1 - \frac{x}{2}\right)^2 Q^2(x, \tau) \sqrt{1-x^2} dx d\tau. \quad \square \end{aligned}$$

The previous theorem is a convergence result by noting that the right-hand side of (2.16) goes to zero in a spectral way (see for instance [1]).

*Remark 2.2.* One can check that  $\gamma_{N+1}/N$  converges to  $\frac{1}{2}$  when  $N$  goes to  $+\infty$ . This means that, by taking  $\alpha$  proportional to  $N^2$ , the hypothesis of Theorem 2.2 is satisfied. This assumption is similar to that made for the time-independent problem (see (1.10)).  $\square$

**3. Boundary Conditions for Elliptic Equations.** A theoretical analysis of the convergence for pseudospectral approximations of the solution of Neumann problems, with a modified approach to treat the boundary conditions similar to that examined in the previous sections, has been developed in [4]. Here we shall prove that the matrices relative to such approximations have real and strictly negative eigenvalues (note that, in the Chebyshev case, these matrices are not symmetric). For this purpose, we consider the parabolic equation

$$(3.1) \quad U_t = U_{xx}, \quad |x| \leq 1,$$

with the Neumann boundary conditions

$$(3.2) \quad U_x(\pm 1) = 0.$$

The solution is determined up to a constant. The Chebyshev method with the new boundary treatment involves seeking an  $N$ th-degree polynomial  $u_N$  such that

$$(3.3) \quad \frac{\partial u_N}{\partial t} = \frac{\partial^2 u_N}{\partial x^2} \quad \text{at } x_j = \cos \frac{j\pi}{N}, \quad j = 1, \dots, N-1,$$

and

$$(3.4) \quad \begin{cases} \frac{\partial u_N}{\partial t} - \frac{\partial^2 u_N}{\partial x^2} + \alpha \frac{\partial u_N}{\partial x} = 0 & \text{at } x = 1, \\ \frac{\partial u_N}{\partial t} - \frac{\partial^2 u_N}{\partial x^2} - \alpha \frac{\partial u_N}{\partial x} = 0 & \text{at } x = -1, \end{cases}$$

where  $\alpha$  is a positive constant to be determined later on. The eigenvalue problems associated with (3.3)–(3.4) consist of finding a nonvanishing polynomial  $v$ , of degree at most  $N$ , such that

$$(3.5) \quad \lambda v = v_{xx} \quad \text{at } x = x_j, \quad j = 1, \dots, N-1,$$

and

$$(3.6) \quad \begin{cases} \lambda v - v_{xx} + \alpha v_x = 0 & \text{at } x = 1, \\ \lambda v - v_{xx} - \alpha v_x = 0 & \text{at } x = -1. \end{cases}$$

The problem (3.5) admits the trivial solution  $\lambda = 0$ . We will show that the other eigenvalues are real and strictly negative. We begin by noticing that one can explicitly derive the characteristic polynomial of (3.5)–(3.6). In fact, (3.5) can be written as follows:

$$(3.7) \quad \lambda v = v_{xx} + aR + bS, \quad a, b \in \mathbf{R},$$

where

$$R(x) = \frac{xT'_N(x)}{N^2}, \quad S(x) = \frac{T'_N(x)}{N^2}.$$

Therefore, following [8], we have the next result.

LEMMA 3.1. *The solution  $v$  of (3.7) is given by*

$$(3.8) \quad v(x) = ap(x, \mu) + bq(x, \mu),$$

where  $\mu = 1/\lambda$  and

$$(3.9) \quad \begin{cases} p(x, \mu) = \sum_{k=0}^{\infty} R^{(2k)}(x) \mu^{k+1}, \\ q(x, \mu) = \sum_{k=0}^{\infty} S^{(2k)}(x) \mu^{k+1}. \end{cases}$$

*Proof.* We first note that  $p$  and  $q$  are polynomials in  $x$ . Then, it is easily verified that

$$(3.10) \quad \begin{cases} \lambda p - p_{xx} = R & \text{in } \mathbf{R}, \\ \lambda q - q_{xx} = S & \text{in } \mathbf{R}, \end{cases}$$

and therefore  $v$  defined in (3.8) is the solution of (3.7). This completes the proof.  $\square$

To get the characteristic polynomial of the second derivative operator, we need to substitute (3.8) into (3.6) and make use of (3.10) to get

$$(3.11) \quad \begin{cases} a \left[ R(1) + \alpha \frac{\partial p}{\partial x}(1, \mu) \right] + b \left[ S(1) + \alpha \frac{\partial q}{\partial x}(1, \mu) \right] = 0, \\ a \left[ R(-1) - \alpha \frac{\partial p}{\partial x}(-1, \mu) \right] + b \left[ S(-1) - \alpha \frac{\partial q}{\partial x}(-1, \mu) \right] = 0. \end{cases}$$

From now on we suppose that  $N$  is even (for  $N$  odd, similar arguments can be applied), so that we have  $R(1) = S(1) = R(-1) = -S(-1) = 1$  and  $p(x, \mu) = p(-x, \mu)$ ,  $q(x, \mu) = -q(-x, \mu)$ . Hence, we can state

THEOREM 3.1. *The complex number  $\lambda \neq 0$  is an eigenvalue of (3.6) if and only if  $\mu = 1/\lambda$  satisfies*

$$(3.12) \quad 2 \left[ 1 + \alpha \frac{\partial p}{\partial x}(1, \mu) \right] \left[ 1 + \alpha \frac{\partial q}{\partial x}(1, \mu) \right] = 0.$$

*Proof.* The left-hand side of (3.12) is the determinant of (3.11). Since we are looking for a nontrivial solution of (3.6), this determinant must vanish.  $\square$

Now define

$$(3.13) \quad \begin{cases} g(\mu) = 1 + \alpha \frac{\partial p}{\partial x}(1, \mu), \\ h(\mu) = 1 + \alpha \frac{\partial q}{\partial x}(1, \mu). \end{cases}$$

It is not difficult to check that  $g$  and  $h$  are polynomials in  $\mu$  of degree  $N/2$ . In order to show that the roots of  $g(\mu)$  and  $h(\mu)$  are real negative and distinct, we use the notion of a positive pair (see [5] and [8]). Two polynomials form a positive pair if their roots are real negative and interlaced. We shall prove, for instance, that  $g(\mu)$  and  $p(1, \mu)/\mu$  form a positive pair. To show that, we first need the following result.

LEMMA 3.2. *Let*

$$(3.14) \quad f(\mu) = g(\mu^2) + \alpha\mu \left[ \frac{p(1, \mu^2)}{\mu^2} \right],$$

where  $g$  is defined in (3.13) and  $p$  in (3.9). Then  $f$  is a Hurwitz polynomial (i.e., all its roots lie in the left side of the imaginary axis) provided  $\alpha$  is sufficiently large.

*Proof.* By the definitions (3.9) and (3.13), one easily gets

$$(3.15) \quad \begin{aligned} f(\mu) &= 1 + \alpha \sum_{k=0}^{\infty} R^{(2k+1)}(1)\mu^{2k+2} + \alpha\mu \sum_{k=0}^{\infty} R^{(2k)}(1)\mu^{2k} \\ &= 1 + \alpha \sum_{m=0}^{\infty} R^{(m)}(1)\mu^{m+1}. \end{aligned}$$

We show that  $f$  is the characteristic polynomial relative to the pseudospectral approximation of a hyperbolic problem. In fact, define

$$\rho(x, \mu) = \sum_{m=0}^{\infty} R^{(m)}(x)\mu^{m+1};$$

then it is readily verified that

$$(3.16) \quad \frac{1}{\mu} \rho(x, \mu) = \rho_x(x, \mu) - \alpha \rho(1, \mu) R(x),$$

and that the roots of  $f(\mu) = 1 + \alpha \rho(1, \mu) = 0$  give the corresponding eigenvalues. Now, (3.16) actually is the eigenvalue problem associated with the hyperbolic equation

$$(3.17) \quad \frac{\partial w_N}{\partial t} = \frac{\partial w_N}{\partial x} - \alpha w_N(1, t) R(x).$$

With a proof similar to that of Theorem 2.1, where  $w_N$  plays the role of  $\varepsilon_N$  with  $g \equiv 0$  and  $Q \equiv 0$ , it is possible to show that, for some norm  $\|\cdot\|$ , we have  $d\|w_N\|^2/dt < 0$  if  $\alpha$  is suitably large. This implies that  $f$  is Hurwitz.  $\square$

As an immediate result of Lemma 3.2, we have the next theorem.

**THEOREM 3.2.** *If  $\alpha$  is sufficiently large, then the roots  $\mu$  of the polynomial  $g$  defined in (3.13) are real negative and distinct.*

*Proof.* The theorem is a consequence of  $f$  being a Hurwitz polynomial. In fact, this is a necessary and sufficient condition for  $g(\mu)$  and  $\rho(1, \mu)/\mu$  to form a positive pair (see [5, p. 228]). In particular, the roots of  $g$  are real and negative.  $\square$

In the same way, we can also prove

**THEOREM 3.3.** *If  $\alpha$  is sufficiently large, then the roots  $\mu$  of the polynomial  $h$  defined in (3.13) are real negative and distinct.*

*Proof.* It can be verified that the polynomials  $h(\mu)$  and  $q(1, \mu)/\mu$  form a positive pair by showing that  $h(\mu^2) + \alpha\mu[q(1, \mu^2)/\mu^2]$  is a Hurwitz polynomial.  $\square$

Finally, by Theorems 3.1, 3.2, and 3.3, we can conclude with the following result.

**THEOREM 3.4.** *If  $\alpha$  is sufficiently large, then the eigenvalues  $\lambda \neq 0$  of the second-derivative Chebyshev matrix with the boundary conditions (3.6) are real and negative.  $\square$*

It is easily verified that  $\alpha$  turns out to be proportional to  $N^2$ , as is also pointed out in [4], where an explicit formula for  $\alpha$  is given.

**4. Analysis of the Eigenvalues and Numerical Experiments.** In this section we analyze the behavior of the eigenvalues of the  $(N + 1) \times (N + 1)$  matrix associated with the scheme (1.3). Applying the same proof of Theorem 2.1 in Section 2 to the equation (2.5) with  $g \equiv 0$ , we get  $d\|u_N\|^2/dt < 0$ .

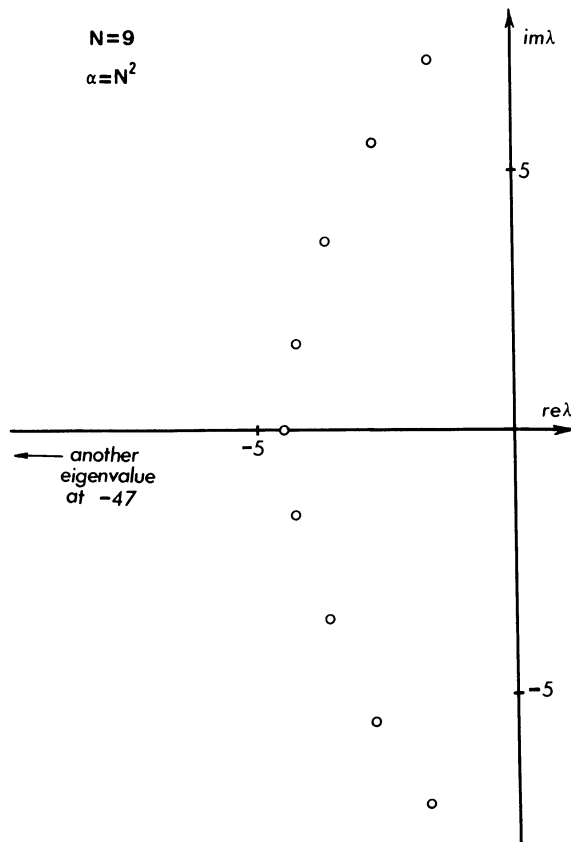


FIGURE 4.1  
*Eigenvalues in the complex plane using scheme (1.3).*

This implies that all the eigenvalues have negative real parts. In Figure 4.1 they are plotted for  $N = 9$  and  $\alpha = N^2$ . The distribution in the complex plane is similar to that of the eigenvalues corresponding to the  $N \times N$  matrix associated with the system (1.2). The extra eigenvalue coming from (1.3) is real negative and its magnitude is proportional to  $N^2$ . If  $R_N(\lambda)$  is the  $N$ th-degree characteristic polynomial related to (1.2) (see [2] for the explicit expression of the coefficients), it is easily verified that the eigenvalues of (1.3) are the  $N + 1$  roots of the equation

$$(4.1) \quad \lambda^{N+1} + \alpha R_N(\lambda) = 0.$$

Following [6], it can be shown that the eigenfunction corresponding to the root  $\lambda$  of (4.1) (up to a normalizing constant) takes the following form

$$(4.2) \quad u(x) = \sum_{k=0}^N h^{(k)}(x)\lambda^{N-k}, \quad \text{where } h(x) = T'_N(x)(1+x).$$

To discretize in time (2.2), we can use the second-order Runge-Kutta method. The analysis of the stability of the method, based on the knowledge of the eigenvalues of (1.3), gives an upper bound on the time step  $\Delta t$ . By choosing  $\alpha$  proportional to  $N^2$ , the restriction on  $\Delta t$  is given by the formula

$$(4.3) \quad \Delta t \leq \frac{2}{\alpha - .4N^2}.$$

Therefore, by taking  $\alpha = N^2$ , condition (4.3) says that  $\Delta t \leq 3.3/N^2$ . This restriction is slightly more severe than that obtained by exactly imposing the boundary condition in  $x = 1$ . Indeed, in this last case, we had  $\Delta t \leq 17/N^2$  (see [2]). The more restrictive condition on  $\Delta t$  is due to the presence of the real eigenvalue with the largest magnitude. One could think that this result negatively influences time discretization for scheme (2.2). Nevertheless, we argue that this is not the case. In fact, consider problem (2.1), when the initial guess is  $f(x) = 1 - \cos(x - 1)$  and  $g \equiv 0$ . We discretize the equation by collocation at the Chebyshev nodes  $x_j$ ,  $j = 1, \dots, N$ . Two different conditions are tested in  $x = 1$ , namely

$$(4.4) \quad \begin{aligned} \text{a) } & u_N(1, t) = 0, \\ \text{b) } & \frac{\partial u_N}{\partial t}(1, t) = \frac{\partial u_N}{\partial x}(1, t) - \alpha u_N(1, t). \end{aligned}$$

We take  $N = 8, \alpha = N^2$  and  $t \in [0, T]$  with  $T = 1$ , and we evaluate the error  $E$  as in (1.4) using both the schemes, respectively obtained by imposing conditions a or b in (4.4). Second-order Runge-Kutta is used for time discretization. Figure 4.2 shows the behavior of the error versus  $\Delta t$ . As the analysis of the eigenvalues pointed out, by increasing  $\Delta t$  using condition b, instability occurs earlier than using condition a.

TABLE 4.1  
*Comparison of the errors for different  $\Delta t$  and  $N$ .*

	$\Delta t = .01$		$\Delta t = .001$		$\Delta t = .0001$	
$N$	Condition a	Condition b	Condition a	Condition b	Condition a	Condition b
8	.1649E-02	.1124E-02	.1644E-02	.1122E-02	.1644E-02	.1122E-02
16	.2076E-03	.2039E-03	.1996E-03	.1962E-03	.1995E-03	.1961E-03
32	.6837E-04	OVERFLOW	.3846E-04	.3653E-04	.3842E-04	.3649E-04

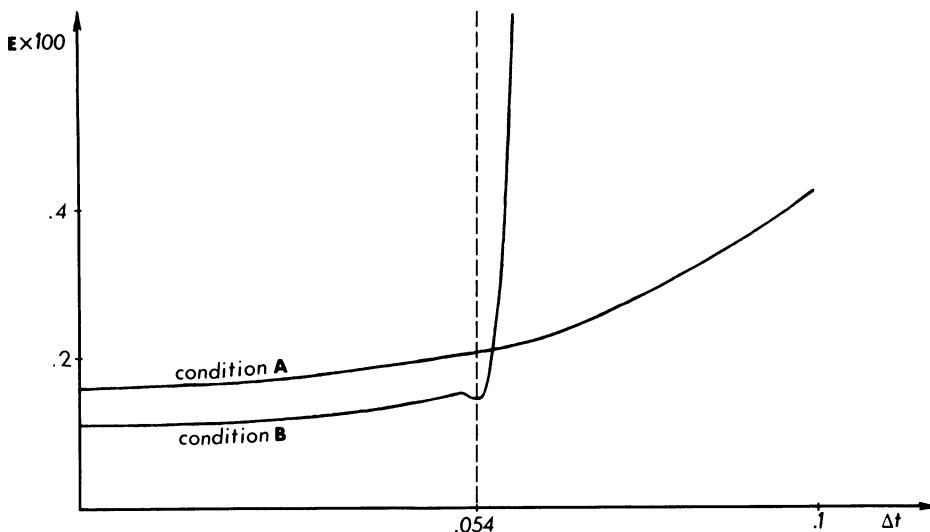


FIGURE 4.2

*Comparison of the errors versus  $\Delta t$  using different boundary conditions.*

However, the error using method b is significantly smaller than that of method a. Moreover, the choice of larger  $\Delta t$  for method a causes deterioration of the accuracy.

For the same example, Table 4.1 shows the error when  $T = 1$  for various choices of  $N$ . Similarly, Table 4.2 shows the error when different values of  $T$  are used and  $\Delta t = .001$ . In almost all the cases, the use of condition b is preferred, especially when large values of  $T$  are considered. Similar results can be obtained when time-dependent boundary conditions are considered.

TABLE 4.2

*Comparison of the errors for different  $T$  and  $N$ .*

	$T = .5$		$T = 2$		$T = 10$	
$N$	Condition a	Condition b	Condition a	Condition b	Condition a	Condition b
8	.1107E-02	.1200E-02	.8904E-03	.7875E-03	.2010E-07	.4098E-09
16	.2594E-03	.2540E-03	.9567E-04	.8506E-04	.8312E-10	.2718E-13
32	.4168E-04	.3904E-04	.1187E-04	.1174E-04	.2231E-14	.9872E-19

We conclude this section by discussing preconditioning for the matrix corresponding to (1.3). For the matrix resulting from scheme (1.2) an efficient preconditioner was proposed in [3]. Such preconditioner can be written as a product of two  $N \times N$  matrices  $Z$  and  $D$ , where  $D$  is the upwind finite-differences matrix at the collocation nodes and  $Z$  is a shift in the space of polynomials of degree  $N - 1$ , from the values at the staggered grid points to the values at the initial grid. The eigenvalues after preconditioning are real positive and between 1 and  $\pi/2$ . An analogous result holds for the  $(N + 1) \times (N + 1)$  matrix corresponding to the scheme (1.3). As preconditioner for this matrix we take  $\hat{Z}\hat{D}$ , where  $\hat{Z}$  and  $\hat{D}$  have respectively

the form

$$\hat{Z} = \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \boxed{Z} \\ \vdots & \\ 0 & \end{pmatrix}, \quad \hat{D} = \begin{pmatrix} -\alpha & 0 \cdots 0 \\ \frac{1}{x_0 - x_1} & \boxed{D} \\ 0 & \\ \vdots & \\ 0 & \end{pmatrix}.$$

The preconditioned eigenvalues can be explicitly computed also in this case. They are

$$\lambda_0 = 1; \quad \lambda_m = \frac{m \sin(\pi/2N)}{\sin(m\pi/2N)}, \quad m = 1, \dots, N.$$

In particular,  $1 \leq \lambda_m < \pi/2$ . The corresponding eigenfunctions, up to a multiplicative constant, are

$$u_m(x) = T_m(x) - 1 + \frac{m^2}{\alpha(\lambda_m + 1)}, \quad m = 0, 1, \dots, N.$$

The preconditioner presented above is particularly suggested when steady-state solutions of problem (2.2) have to be computed.

Dipartimento di Matematica  
Università di Pavia  
Strada Nuova 65  
27100 Pavia, Italy

Division of Applied Mathematics  
Brown University  
Providence, Rhode Island 02912

1. C. CANUTO & A. QUARTERONI, "Error estimates for spectral and pseudospectral approximation of hyperbolic equations," *SIAM J. Numer. Anal.*, v. 19, 1982, pp. 629-642.
2. D. FUNARO, "Some results about the spectrum of the Chebyshev differencing operator," in *Numerical Approximation of P.D.E.*, Mathematical Studies, No. 133, Part III (E. L. Ortiz, ed.), North-Holland, Amsterdam, 1987, pp. 271-284.
3. D. FUNARO, "A preconditioning matrix for the Chebyshev differencing operator," *SIAM J. Numer. Anal.*, v. 24, 1987, pp. 1024-1031.
4. D. FUNARO, "Domain decomposition methods for pseudo spectral approximations-Part One: Second order equations in one dimension," *Numer. Math.*, v. 52, 1988, pp. 329-344.
5. F. R. GANTMACHER, *The Theory of Matrices*, Chelsea, New York, 1974.
6. D. GOTTLIEB & S. A. ORSZAG, *Numerical Analysis of Spectral Methods: Theory and Applications*, CBMS Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, Pa., 1977.
7. D. GOTTLIEB, "The stability of pseudospectral Chebyshev methods," *Math. Comp.*, v. 36, 1981, pp. 107-118.
8. D. GOTTLIEB & L. LUSTMAN, "The spectrum of the Chebyshev collocation operator for the heat equation," *SIAM J. Numer. Anal.*, v. 20, 1983, pp. 909-921.
9. T. J. RIVLIN, *The Chebyshev Polynomials*, Wiley, New York, 1974.
10. A. SOLOMONOFF & E. TURKEL, *Global Collocation Methods for Approximation and the Solution of Partial Differential Equations*, ICASE Report No. 86-60, submitted to *J. Comput. Phys.*