Computers & Fluids Vol. 14, No. 1, pp. 23–41, 1986
Printed in Great Britain.

0045-7930/86 $3.00 + .00

SPECTRAL AND FINITE DIFFERENCE SOLUTIONS OF THE BURGERS EQUATION

C. BASDEVANT
Laboratoire de Météorologie Dynamique, Ecole Normale Supérieure, 75231 Paris, Cedex 05, France

M. DEVILLE
Unité de Mécanique Appliquée, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

P. HALDENWANG
Département d'Héliophysique, Université de Provence, Centre de St. Jérôme, 13397 Marseille, Cedex 13, France

J. M. LACROIX, J. OUAZZANI, R. PEYRET
Département de Mathématiques, Université de Nice, Parc Valrose, 06000 Nice, France

P. ORLANDI
Dipartimento di Meccanica e Aeronautica, Universita di Roma, Via Eudossiana 18, 00184 Roma, Italy

and

A. T. PATERA
Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

(*Received* 8 *October* 1984)

Abstract—Spectral methods (Fourier Galerkin, Fourier pseudospectral, Chebyshev Tau, Chebyshev collocation, spectral element) and standard finite differences are applied to solve the Burgers equation with small viscosity ($\nu = 1/100\pi$). This equation admits a (nonsingular) thin internal layer that must be resolved if accurate numerical solutions are to be obtained. From the reported computations, it appears that spectral schemes offer the best accuracy, especially if coordinate transformation or elemental subdivision is used to resolve the regions of large variation of the dependent variable.

1. INTRODUCTION

It is well known (Gottlieb and Orszag [1]) that spectral methods applied to problems with smooth solutions offer a spectacular rate of convergence. However, the question remains: If the solution presents sharp variations over a very thin layer, and almost behaves like a discontinuity, do spectral schemes compete with other classical methods such as finite differences? This question was raised during a GAMM workshop held in Louvain-la-Neuve at the end of 1980, and several investigators proposed to solve the nonlinear Burgers equations with a very small viscosity as a test problem. It should be noted that our interest here is in problems admitting solutions with rapid variation, but not discontinuities; the former can obtain exponential convergence with sufficient resolution, the latter require special treatment if spectral accuracy is to be obtained.

In this paper, all the numerical results produced by the participants in this study are reported, discussed and compared. Section 2 describes the equations and the dynamics with a viscosity ν equal to $1/100\pi$. One gives briefly the analytical solution, which is quite difficult to compute. In fact, it is easier to implement a transient algorithm to perform the numerical integration of the Burgers equation than to estimate the exact values. Section 3 presents the numerical methods used by the co-authors. C. Basdevant presents the solution by the Fourier method; M. Deville and P. Haldenwang present a Chebyshev–Tau method, the former using Rosenbrock, the latter Adams–Bashforth Crank–Nicolson (ABCN); J. M. Lacroix, J. Ouazzani and R. Peyret apply an ABCN collocation method; A. T. Patera uses an ABCN scheme with a spectral element spatial

discretization; P. Orlandi computes a finite difference (FD) solution using a second-order approximation and a stretching transformation.

It appears that for the same number of degrees of freedom, spectral methods do yield better accuracy than FD solutions, especially from the point of view of phase errors. However, if the problem involves, as in this case, sharp variations of the velocity profile, spectral methods do not possess an exponential rate of convergence until the region of rapid change is adequately resolved. The exponentially fast convergence is lost because derivatives are evaluated through the region of large variation. We show here that this relatively dismal performance can be improved using a coordinate transformation or element technique which accumulates points in the region of large gradient.

2. THE BURGERS EQUATION—ANALYTICAL SOLUTION

The test problem consists in the Burgers equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = \nu\frac{\partial u^2}{\partial x^2}, \qquad |x| \leqslant 1, \qquad t > 0, \tag{2.1}$$

with boundary conditions

$$u(\pm 1, t) = 0, \tag{2.2}$$

(or equivalently 2-periodicity) and initial condition

$$u(x, 0) = -\sin \pi x. \tag{2.3}$$

For this problem, the viscosity is taken to be $\nu = 10^{-2}/\pi$. For such a small viscosity, the solution develops into a sawtooth wave at the origin from about the time $t = 1/\pi$. The gradient at the origin reaches its maximum value at about $t = 0.5$. This value represents the travel time of the initial peaks of the velocity profile at $x_{\max} = \pm 0.5$ freestreaming with $u = \mp 1$ toward the origin.

The theoretical solution of problem (2.1–2.3) was obtained by Cole [2] and compiled by Benton and Platzmann [3]:

$$u(x, t) = 4\pi\nu[\sum_{n=1}^{\infty} n a_n e^{-n^2\pi^2 t\nu}\sin n\pi x/(a_o + 2\sum_{n=1}^{\infty} a_n e^{-n^2\pi^2 t\nu}\cos n\pi x)], \tag{2.4}$$

where $a_n = (-1)^n I_n(1/2\pi\nu)$ and $I_n(z)$ denotes the modified Bessel functions of first kind. This analytical solution is numerically untractable at small t ($0 \leqslant t \leqslant 2/\pi$) and ν as $I_n(z)$ with z going to infinity behaves asymptotically as $e^z(2\pi z)^{-1/2}$ independent of n. A better form comes from the convolution product in Cole's transformation, i.e.

$$u(x, t) = \left[-\int_{-\infty}^{+\infty} \sin \pi(x - \eta) f(x - \eta)\exp(-\eta^2/4\nu t)\, d\eta \Big/ \int_{-\infty}^{+\infty} f(x - \eta)\exp(-\eta^2/4\nu t)\, d\eta\right], \tag{2.5}$$

with $f(y) = \exp(-\cos \pi y/2\pi\nu)$. Using Hermite integration, a seven-digit accurate result was produced up to $t = 3/\pi$ in order to compute the value of the slope at the origin. The latter is a sensitive quantity and hence a good measure of the accuracy of the numerical methods.

3. THE NUMERICAL METHODS

In this section, we shall describe the various approaches taken for solving the problem (2.1–2.3). They belong to several categories: Fourier spectral schemes, Chebyshev spectral schemes, and finite differences.

3.1 *Fourier spectral method*

With the variables (x, t) changed in $x' = \pi x$, $t' = \pi t$, and dropping the primes, eqns (2.1) and (2.3) become in conservative form

$$\frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial u^2}{\partial x} = \nu' \frac{\partial^2 u}{\partial x^2}, \qquad -\pi \leqslant x \leqslant \pi, \qquad t > 0 \tag{3.1}$$

$$u(x, 0) = -\sin x, \tag{3.2}$$

where $\nu' = \pi\nu = 10^{-2}$. As u is now 2π-periodic in x, one can approximate the solution of (3.1) by a truncated Fourier series

$$u_n(x, t) = \sum_{k=-N}^{N} u_k(t) e^{ikx}. \tag{3.3}$$

Following Gottlieb and Orszag [1], the spectral equations are

$$\frac{du_k}{dt} = -\frac{ik}{2} N_k(\mathbf{u}, \mathbf{u}) - \nu' k^2 u_k, \qquad -N \leqslant k \leqslant N, \tag{3.4}$$

with the initial conditions

$$\begin{aligned} u_k(0) &= 0 \qquad \text{for} \qquad k \neq 1, \\ u_1(0) &= i/2, \qquad u_{-1}(0) = -i/2. \end{aligned} \tag{3.5}$$

In (3.4), $\mathbf{u} = [u_k]$ and $N_k(\mathbf{u}, \mathbf{u})$ denotes the nonlinear term which is provided by a convolution product. The standard Galerkin and pseudo-spectral methods differ in practice in the nonlinear term evaluation.

Let us assume $M \geqslant N$ and define

$$\begin{aligned} \tilde{u}_k &= u_k, \qquad |k| \leqslant N, \\ \tilde{u}_k &= 0, \qquad N < |k| \leqslant M. \end{aligned} \tag{3.6}$$

The algorithm proceeds as follows:

(i) Perform an inverse Discrete Fourier Transform (DFT) using Fast Fourier algorithms (FFT) of length M on $\tilde{u}_k$ to obtain $u(x_j)$, $x_j = 2\pi j/2M$, $0 \leqslant j \leqslant 2M - 1$.
(ii) Compute $w(x_j) = u(x_j)^2$, $0 \leqslant j \leqslant 2M - 1$.
(iii) Perform the direct DFT of length M on the $w(x_j)$ values to produce $u*u(k) = w_k$, $-M \leqslant k \leqslant M$.

Depending on the respective values of M and N, the scheme will remove or leave the aliasing errors.

If $M > \frac{3}{2}N$, the scheme is the *Galerkin method* and

$$N_k(\mathbf{u}, \mathbf{u}) = \sum_{p=-N}^{N} u_p u_{k-p}, \tag{3.7}$$

while for $\frac{3}{2}N \geqslant M \geqslant N$, aliasing errors pollute the solution and the scheme is a pseudo-spectral approximation.

$$\begin{aligned} N_k(\mathbf{u}, \mathbf{u}) &= \sum u_p u_q, \\ &\quad |p|, |q| \leqslant N \\ &\quad k = p + q + 2Me, \end{aligned} \tag{3.8}$$

where e may be 0 or ± 1. Aliasing terms come from the inability for the $2M$ grid points x_j in the physical space to separate a wave from its higher-order harmonics.

The time marching technique applied to (3.4) uses the leapfrog scheme for the nonlinear term and an "exact" integration for the linear part. Rewriting (3.4) for a single Fourier mode as

$$\frac{\mathrm{d}u_k}{\mathrm{d}t} = \mathcal{N}_k(\mathbf{u}, \mathbf{u}) + \lambda u_k, \qquad \mathcal{N}_k = -\frac{ik}{2} N_k,$$

$$\lambda = -\nu' k^2, \tag{3.9}$$

Eq. (3.9) may be time integrated

$$u_k(t + \Delta t) = u_k(t - \Delta t)\mathrm{e}^{2\lambda\Delta t} + \int_{t-\Delta t}^{t+\Delta t} \mathcal{N}_k(\mathbf{u}, \mathbf{u})(\tau)\mathrm{e}^{\lambda(t+\Delta t-\tau)}\,\mathrm{d}\tau.$$

Using the trapezoidal rule, the previous relationship leads to the explicit scheme

$$u_k(t + \Delta t) = u_k(t - \Delta t)\mathrm{e}^{2\lambda\Delta t} + 2\Delta t \mathrm{e}^{\lambda\Delta t}\mathcal{N}_k(\mathbf{u}, \mathbf{u})(t), \tag{3.10}$$

which is second-order accurate in time and unconditionally stable for the viscous terms.

3.2 *Chebyshev spectral methods*

Let u_N denote the Chebyshev approximation of the dependent variable u. Therefore,

$$u_N(x, t) = \sum_{n=0}^{N} a_n(t) T_n(x), \tag{3.11}$$

where $T_n(x)$ is the nth order Chebyshev polynomial defined by $T_n(x) = \cos(n \cos^{-1} x)$.

3.2.1. *Methods in spectral space (TAU).* Using the *Tau projection method* for the linear part of (2.1) and the pseudo-spectral scheme for the nonlinear term, the semi-discrete spectral equations are [1]

$$c_n \frac{\mathrm{d}a_n}{\mathrm{d}t} = \mathcal{N}_n(\mathbf{a}, \mathbf{a}) + \mathcal{L}_n(\mathbf{a}), \qquad 0 \leqslant n \leqslant N - 2$$

$$c_0 = 2, \qquad c_n = 1, \qquad \text{for} \qquad n > 0, \tag{3.12}$$

where $\mathcal{N}_n$ and $\mathcal{L}_n$ represent the nonlinear and linear terms, respectively. Denoting by $\mathbf{a} = [a_n]$ the set of Chebyshev coefficients, the operators on the right-hand side of (3.12) may be defined as

$$\mathcal{L}_n(\mathbf{a}) = \nu \sum_{\substack{m=n+2 \\ m+n \text{ even}}}^{N} m(m^2 - n^2) a_m,$$

$$\mathcal{N}_n(\mathbf{a}, \mathbf{a}) = -2 \sum_{\substack{m=-N \\ m+p \geqslant n+1 \\ m+p+n \text{ odd}}}^{N} \sum_{p=-N}^{N} c_{|m|} c_{|p|} p a_{|p|} a_{|m|}$$

$$-2 \sum_{\substack{m=-N \\ m+p \geqslant 2N-n+1 \\ m+p+n \text{ odd}}}^{N} \sum_{p=-N}^{N} c_{|m|} c_{|p|} p a_{|p|} a_{|m|}. \tag{3.13}$$

The system (3.12) is closed by the boundary conditions

$$\sum_{n=0}^{N} (-1)^n a_n = 0; \qquad \sum_{n=0}^{N} a_n = 0.$$

The nonlinear terms is expensive to evaluate directly from (3.13) and is obtained in practice by pseudo-spectral calculation. From the coefficients a_n, it is possible to compute the Chebyshev coefficients of the first-order derivative by standard recurrence relationships and the nonlinear term is evaluated in physical space. From this evaluation, we get $\mathcal{N}_n(\mathbf{a}, \mathbf{a})$. To go from Chebyshev to physical space, and vice versa, can be performed easily by discrete Chebyshev transforms or by matrix multiplications on a vector machine. Two time schemes are considered to approximate (3.12).

The first method, denoted ABCN, uses the Adams–Bashforth scheme for the convective part, and the Crank–Nicolson scheme for the viscous term.

$$c_n \frac{a_n(t + \Delta t) - a_n(t)}{\Delta t} = \frac{3}{2} \mathcal{N}_n(t) - \frac{1}{2} \mathcal{N}_n(t - \Delta t) + \frac{1}{2} [\mathcal{L}_n(t) + \mathcal{L}_n(t + \Delta t)], \quad (3.14)$$

where $\mathcal{N}_n(t)$ is a short notation for $\mathcal{N}_n[\mathbf{a}(t), \mathbf{a}(t)]$. It is shown in Deville *et al.* [4] that this scheme of second-order accuracy in time is sufficient to achieve a moderate accuracy level at a reasonable computational cost. Note that (3.14) implies the solution of a linear algebraic system with constant coefficients at each time step.

The second method is a semi-implicit Runge–Kutta scheme of Rosenbrock type. It is due to Michelsen and Villadsen [5] and is third-order accurate in time. If we set

$$\Lambda(\cdot) = 1 - c\Delta t[\mathcal{N}(\mathbf{a}, \cdot) + \mathcal{N}(\cdot, \mathbf{a}) + \mathcal{L}(\cdot)],$$

where $\mathbf{a}$ is evaluated at time t, the Rosenbrock algorithm is

$$\begin{aligned} \Lambda(\mathbf{k}_1) &= \Delta t \mathcal{F}(\mathbf{a}), \\ \Lambda(\mathbf{k}_2) &= \Delta t \mathcal{F}(\mathbf{a} + b_2 \mathbf{k}_1), \\ \Lambda(\mathbf{k}_3) &= \Delta t \mathcal{F}(b_{31}\mathbf{k}_1 + b_{32}\mathbf{k}_2) \\ \mathbf{a}(t + \Delta t) &= \mathbf{a}(t) + R_1\mathbf{k}_1 + R_2\mathbf{k}_2 + \mathbf{k}_3, \end{aligned} \quad (3.15)$$

where $\mathcal{F}(\cdot) = \mathcal{N}(\cdot, \cdot) + \mathcal{L}(\cdot)$, and $c = 0.43586652$, $b_2 = 0.75$, $b_{31} = -8\,(c^2 - 2c + 1)/6c$, $b_{32} = 2\,(6c^2 - 6c + 1)/9c$, $R_1 = 11/27 - b_{31}$, $R_2 = 16/27 - b_{32}$.

The algorithm is strongly stable and is designed to integrate stiff equations. The method is rather expensive as it requires factorization of a linear system at each time cycle. Furthermore, the construction of the associated matrix implies $O(N^3)$ operations.

3.2.2. *Methods in physical space (Collocation).* This kind of procedure was proposed, for example, by Lanczos [6] and Finlayson [7]. Here, Chebyshev polynomials are used in a collocation pseudo-spectral approximation with the ABCN scheme (Ouazzani and Peyret [8]).

The second-order derivative at each interior collocation point $x_i = \cos \pi i/N$; $i = 1, \ldots, N - 1$; is expressed in terms of the values u_k at all points $k = 0, \ldots, N$ by the formula

$$\frac{\partial^2 u_N}{\partial x^2}(x_i, t) = \sum_{k=0}^{N} L_{ik} u_N(x_k, t), \quad (3.16)$$

where L_{ik} is given by

$$L_{i,k} = \frac{(-1)^{i+k+1}}{\bar{c}_k(1 - x_i^2)(x_i - x_k)^2}(2 - x_i x_k - x_i^2), \qquad 1 \leqslant i \leqslant N - 1$$

$$0 \leqslant k \leqslant N$$

$$(\bar{c}_0 = \bar{c}_N = 2,\ \bar{c}_i = 1 \text{ for } 1 \leqslant i \leqslant N - 1) \qquad i \neq k$$

$$L_{i,i} = -\frac{1}{3(1 - x_i^2)}\left[\frac{3}{1 - x_i^2} + N^2 - 1\right] \qquad 1 \leqslant i \leqslant N - 1.$$

Therefore, the resulting algebraic system may be written as

$$u_N(x_i, t + \Delta t) - \frac{\nu \Delta t}{2} \frac{\partial^2 u_N}{\partial x^2}(x_i, t + \Delta t) = u_N(x_i, t) - \Delta t \left[\frac{3}{2} \left(u_N \frac{\partial u_N}{\partial x} \right)(x_i, t) - \frac{1}{2} \left(u_N \frac{\partial u_N}{\partial x} \right)(x_i, t - \Delta t) - \frac{\nu}{2} \frac{\partial^2 u_N}{\partial x^2}(x_i, t) \right], \quad 1 \leqslant i \leqslant N - 1$$

$$u_N(-1, t + \Delta t) = 0 \qquad u_N(1, t + \Delta t) = 0.$$

The nonlinear term is computed through the pseudospectral technique.

As already noted, a region of large gradient develops near the origin. The thickness of this layer is approximately 0.042 and the number of collocation points in this region, $-0.021 \leqslant x \leqslant 0.021$, is not sufficient if the cutoff value N is too small. This number is 7 for $N = 512$, 3 for $N = 256$ and only 1 if $N \leqslant 128$. Table 1 presents the abscissa $x_{N/2-1}$ of the first collocation point adjacent to the origin. In order to obtain more collocation points in this region, we use a coordinate transformation

$$x = g(\zeta), \tag{3.17}$$

where $g(\zeta)$ is an odd function satisfying $g(\pm 1) = \pm 1$. The chosen function g is

$$x = g(\zeta) = (1 - \alpha)\zeta^3 + \alpha\zeta, \qquad \alpha > 0. \tag{3.18}$$

For $\alpha(=g'(0))$ small, the collocation points in physical (x) coordinate are clustered near the origin. The transformed eqn (2.1) yields

$$\frac{\partial u}{\partial t} + \frac{u}{g'} \frac{\partial u}{\partial \zeta} = \frac{\nu}{g'^2} \left(\frac{\partial u^2}{\partial \zeta^2} - \frac{g''}{g'} \frac{\partial u}{\partial \zeta} \right), \quad |\zeta| \leqslant 1, \qquad t > 0, \tag{3.19}$$

and is solved by the foregoing ABCN collocation scheme. The coefficients of the expansion of the first-order derivatives appearing in (3.19) may be found in Ouazzani and Peyret [8].

3.2.3. *Spectral element methods.* The spectral element decomposition [9, 10] breaks the computational domain up into a series of macro-elements of length L^k, $k = 1, \ldots, M$, and expands the unknown in each element as a Lagrangian interpolant through the Gauss–Lobatto Chebyshev collocation points,

$$u_N^k = \sum_{i=0}^{N} u_i^k h_i(r), \tag{3.20a}$$

$$h_i(r) = \frac{2}{N} \sum_{n=0}^{N} \frac{1}{\bar{c}_i \bar{c}_n} T_n(r_i) T_n(r), \tag{3.20b}$$

where r is the local (elemental) coordinate system $-1 \leqslant r \leqslant 1$.

To solve the Burgers equation, we treat the nonlinear term explicitly (3rd order Adams–Bashforth) with a modified collocation scheme, while the viscous term is handled implicitly with a variational formulation [9]. The elemental matrices are defined as

Table 1. Abscissa of the first adjacent point to the origin with respect to the cutoff value ($-1 \leqslant x \leqslant 1$)

N	32	64	128	256	512
$x_{N/2-1}$	0.0980	0.0491	0.0245	0.0123	0.0061

$$A_{ij}^k = -\frac{2}{L^k}\int_{-1}^{1} \frac{dh_i}{dr}\frac{dh_j}{dr}\,dr \tag{3.21a}$$

$$B_{ij}^k = \frac{L^k}{2}\int_{-1}^{1} h_i h_j\,dr \tag{3.21b}$$

$$C_{ij}^k = A_{ij}^k - \frac{2}{\nu\Delta t}B_{ij}^k, \tag{3.21c}$$

where A_{ij}^k, B_{ij}^k and C_{ij}^k are essentially the Laplacian, mass matrix and Helmholtz operator, respectively. We denote direct stiffness summation [9] as

$$[C] = \sum_k{}' C_{ij}^k, \tag{3.22}$$

where $[\cdot]$ indicates a global matrix or vector. Periodic boundary conditions are applied by identifying the same global node number to the first and last node.

In terms of these matrices, the explicit step is then

$$[B][\hat{u}^{n+1} - u^n] = -\frac{\Delta t}{2}\sum_{q=0}^{2}\alpha_q\left\{\sum_k{}'\left(\sum_j B_{ij}^k \sum_p (u_p^k u_p^k)^{n-q}\frac{dh_p}{dr}(r_j)\right)\right\} \tag{3.23a}$$

where the α_q are the Adams–Bashforth coefficients,

$$\alpha_0 = \tfrac{23}{12}, \qquad \alpha_1 = -\tfrac{4}{3}, \qquad \alpha_2 = \tfrac{5}{12}, \tag{3.23b}$$

where the superscript n refers to the time level $n\Delta t$.

Note multiplication by the mass matrix B_{ij}^k following collocation and differentation recovers certain desirable characteristics of the Galerkin approach, in particular conservation. The viscous step is then constructed from the variational formulation for the Helmholtz equation giving

$$[C][u^{n+1} + u^n] = -\frac{2}{\nu\Delta t}[B][u^n + \hat{u}^{n+1}]. \tag{3.24}$$

Since it is $[B][\hat{u}^{n+1}]$ that is required in (3.24), no (matrix inversion) penalty is associated with use of the consistent mass matrix in eqn (3.23a).

This technique, which has been successfully applied to the Navier–Stokes equations [9, 10] is made efficient by use of collocation for the nonlinear terms, and parallel static condensation [10] for eqn (3.24). Thus for both efficiency and convenience considerations [no patching is required following the variational formulation (3.24)], the method borrows from both finite element and spectral practice. Note the technique is not a true Chebyshev method, as the unity rather than $(1 - z^2)^{-1/2}$ weighting is used in the elemental matrices. Inasmuch, for the elliptic contributions, the scheme reduces to a particular implementation of p-type finite element ideas [11].

3.3. *Finite difference approximation*

Classical centered finite differences are used in conjunction with a coordinate transformation which accumulates points in the large gradient region, and places a limited number of grid points where the solution does not change rapidly. The coordinate transformation is given by the analytical expression

$$x = g(\zeta) = \begin{cases} 1 + \dfrac{\tanh 2S(\zeta - 1)}{\tanh S}, & 0.5 \leqslant \zeta < 1 \\ -1 + \dfrac{\tanh 2S\zeta}{\tanh S}, & 0 < \zeta < 0.5 \end{cases} \tag{3.25}$$

where S is a stretching parameter and here is set to 4. In the transformed space, second-order derivatives are evaluated in the form

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{g'}\frac{\partial}{\partial \zeta}\left(\frac{u'}{g'}\right). \tag{3.26}$$

Time discretization is performed by a linearized noniterative scheme which can produce any of the schemes: Euler implicit, Crank–Nicolson and three-point backward. This implicit noniterative method is due to Briley and McDonald [12], and full details about this finite difference scheme may be found in Orlandi and Briscolini [13].

4. NUMERICAL RESULTS

This section presents the numerical results generated by the proposed algorithms.

4.1. *Fourier method*

An accurate solution was obtained with a Galerkin procedure using 1364 degrees of freedom ($M = 1024$, $N = 682$) and a time step $\Delta t = 5.10^{-4}/\pi$. Figure 1 shows this solution at times $t = 1/\pi, 2/\pi, 3/\pi$. The previous numerical solution is considered as the "reference" solution for error estimates in Figs. 2 and 3. Typical behaviors of the L_2 error appear on Fig. 2. In Fig. 2(a) we present results for a fine resolution ($N = 682$) which shows clearly a $(\Delta t)^{-2}$ decrease of the error at time $t = 1/\pi$ and $2/\pi$. For a coarser resolution ($N = 170$) (Fig. 2b) the error is still of second order at time $t = 1/\pi$ while at time $t = 2/\pi$, the error decays approximately as $(\Delta t)^{-1}$ between $\Delta t = 5.10^{-2}/\pi$ and $\Delta t = 10^{-3}/\pi$. For smaller Δt values, the error remains constant. For very coarse resolutions, the results are still worse. The accuracy of the time integration is destroyed by the spatial errors. Any time-step reduction does not help to cure the problem.

Figure 3 presents the L_2 error with respect to the number of the Fourier modes in the Galerkin schemes. The time step is fixed at $\Delta t = 10^{-3}/\pi$. The rate of convergence is $O(N^{-4})$ at least for the finite N used here and for t close to $2/\pi$. For high N values, the error is constant due to the limited precision of the computer used (AP 120, 38 bits

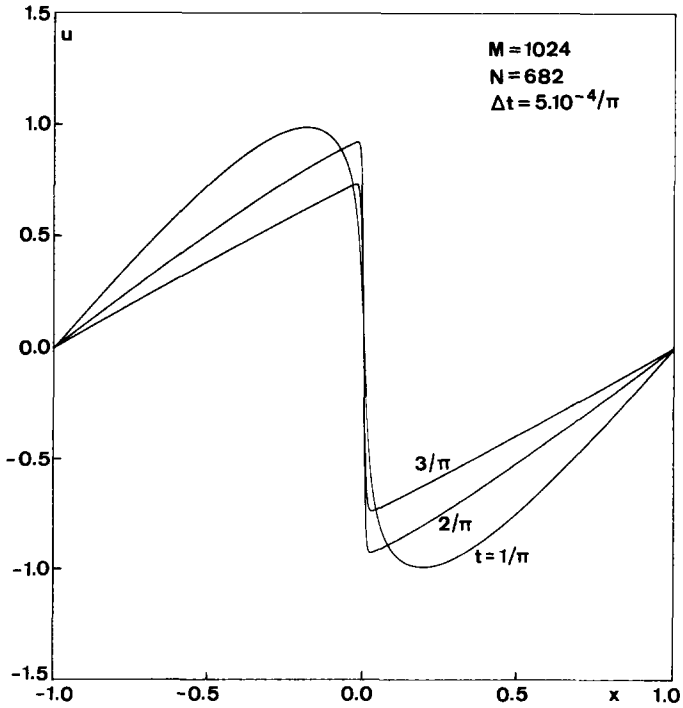

Fig. 1. Reference solution of the Fourier method for $t = 1/\pi, 2/\pi, 3/\pi$ with the Galerkin procedure ($N = 682$, $\Delta t = 5.10^{-4}/\pi$).

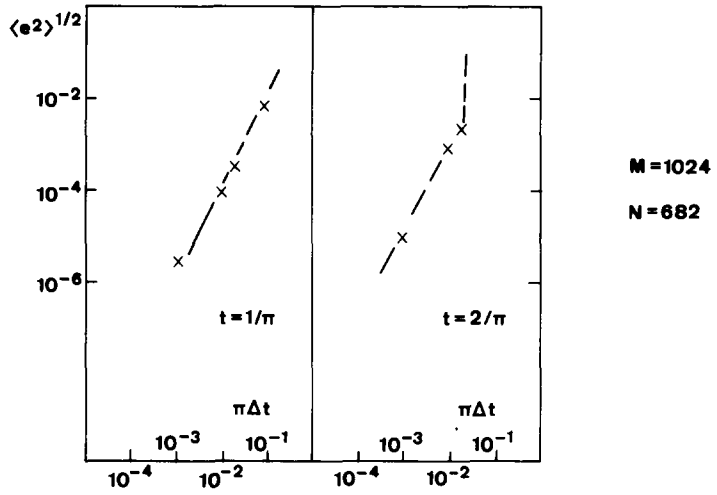

Fig. 2. L_2 error with respect to the time step Δt, at time $t = 1/\pi$ and $2/\pi$. (a). Galerkin method $M = 1024$, $N = 682$.

words). In pseudo-spectral schemes, aliasing errors are harmless provided the cutoff N is larger than 128.

When the resolution is too low to approximate properly the narrow layer developed by the solution, small scales of the numerical solution are unduly excited and therefore, energy which would have been dissipated at small scales, accumulates in the high wave numbers. To avoid this effect, a modified dissipativity term is added to the right-hand side of (3.4), namely

$$-\nu_\beta k^{2\beta} u_k.$$

This procedure is well known in two-dimensional turbulence (Basdevant *et al.* [14]). This damping term is *local* in physical *and* Fourier space for large β values ($\beta = 2, \ldots, 8$). It will act only on small scales and mainly in the vicinity of the origin. Figure 4 displays the effect of this modified dissipativity on a fully aliased collocation method at low resolution ($N = 64$). On this coarse mesh, this extra dissipativity localizes (near the origin) and decreases the error.

4.2. *Chebyshev methods*

4.2.1. *Methods in spectral space (TAU).* Here, because of the oddness of the initial condition which is preserved by eqn (2.1), the integration domain is reduced to [0, 1] in

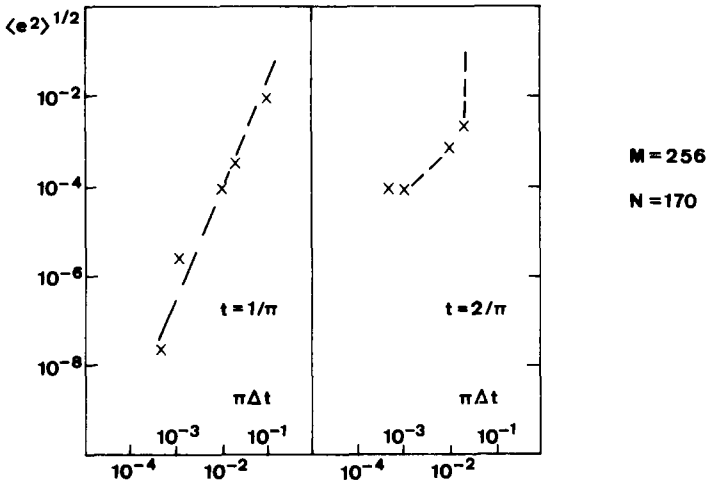

Fig. 2(b). Galerkin method $M = 256$, $N = 170$.

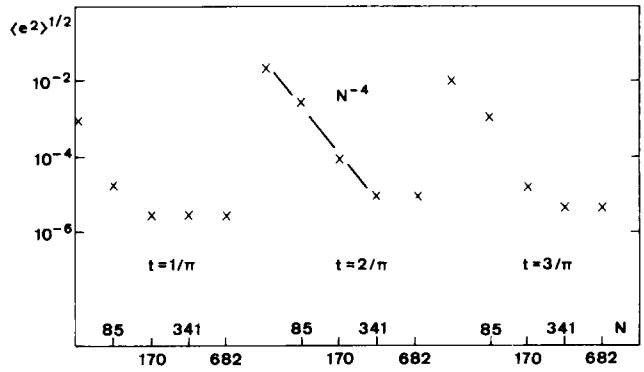

Fig. 3. L_2 error with respect to the cutoff N for Galerkin scheme. The time step is fixed at $\Delta t = 10^{-3}/\pi$.

order to take advantage of the Chebyshev efficiency near the boundaries. At the origin, we impose $u(0, t) = 0$. The solution in the interval $[-1, 1]$ with a large gradient at $x = 0$ would necessitate a very large number of terms as it will be shown in the next section in the case of the ABCN collocation scheme. We include these results to indicate the drastically different convergence properties of Chebyshev methods for internal and boundary layers.

The results obtained by the ABCN Tau method and the Rosenbrock algorithm are virtually indistinguishable. Figure 5 shows the solution obtained at various times by the ABCN scheme with $N = 64$ and $\Delta t = 1/300\pi$. In this computation, the accuracy is about $5 \cdot 10^{-5}$. Figures 6, 7 and 8 present the time evolution of three sensitive quantities, namely the slope $|\partial u/\partial x|$ at the origin, the maximum velocity $u_M = \max_x |u|$ and the abscissa $x_{\max}$ of u_M. The Rosenbrock algorithm needs no stability restriction and the choice of time step comes only from accuracy considerations.

4.2.2. *Method in physical space (Collocation).* Problem (2.1) is first solved on $[-1, 1]$ and Fig. 9 presents plots for different values of N. (On these plots, linear interpolation is used between collocation points.) The time step is always $\Delta t = 10^{-2}/\pi$, except for the case $N = 512$ where $\Delta t = 10^{-2}/2\pi$ to avoid oscillations. It is quite striking to notice that a very large number of polynomials is needed to obtain nice and smooth profiles (no

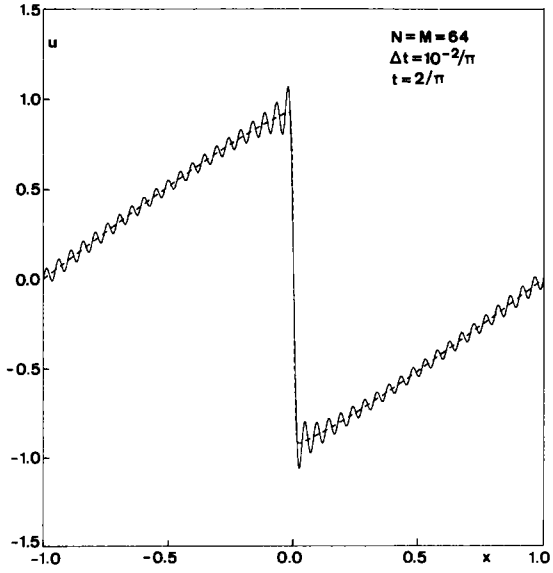

Fig. 4. Effect of the modified dissipativities at low resolution for a fully aliased collocation method. $N = M = 64$, $t = 2/\pi$, $\Delta t = 10^{-2}/\pi$. (a). Comparison of the collocation solution (solid line) without any modelization term with the reference solution (dashed line).

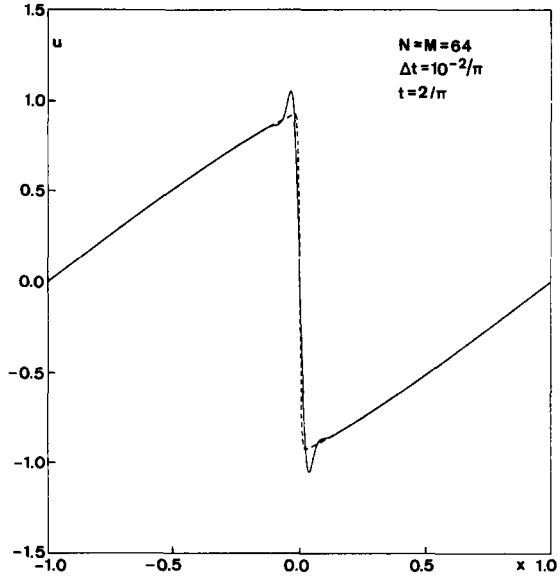

Fig. 4(b). Comparison of the collocation solution with $-\nu_2\,[-(\partial^2)/(\partial x^2)]^2$, $\nu_2 = 1.6\ 10^{-16}$, (solid line) with the reference solution (dashed line).

filtering is used). This effect comes from the fact that at $t = 2/\pi$, the region of large gradient near the origin is very narrow. To get the right value of this gradient, a sufficient number of collocation points is necessary, and from Table 1, one can conclude that N should be at least equal to 256 to get enough spatial resolution. With the coordinate transformation (3.18) and (3.19), results change drastically. Figure 10, obtained with Chebyshev interpolation, displays the numerical solution for different N values ($\nu = 1/100\pi$, $\Delta t = 1/100\pi$) and various α values. On Fig. 11, one finds the evolution of the variable u in the transformed plane. The dilatation produced by the coordinate transformation gives less steep slopes at the origin, which are well represented by Chebyshev approximations with lower cutoff values ($N = 64$). Table 2 gives the values of $|\partial u/\partial x|_{\max}$ at the origin with respect to α. We will return to this point later in section 5.

The collocation scheme was also applied to the interval [0, 1]. In this case, plots are indiscernible from those obtained with the Chebyshev–Tau schemes.

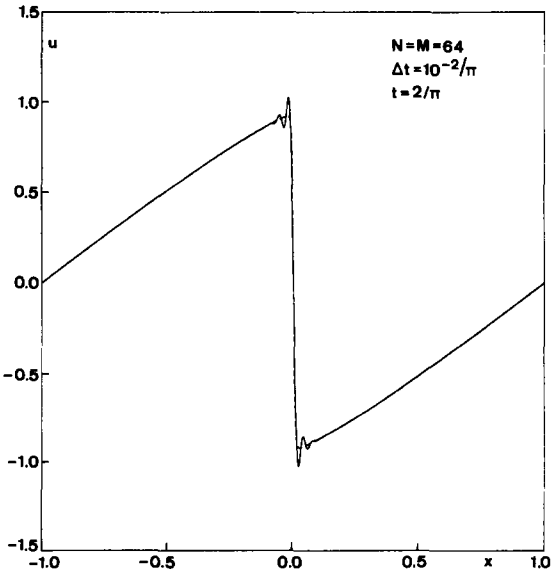

Fig. 4(c). Comparison of the collocation solution with $-\nu_4\,[-(\partial^2)/(\partial x^2)]^4$, $\nu_4 = 3.3\ 10^{-16}$, (solid line) with the reference solution (dashed line).

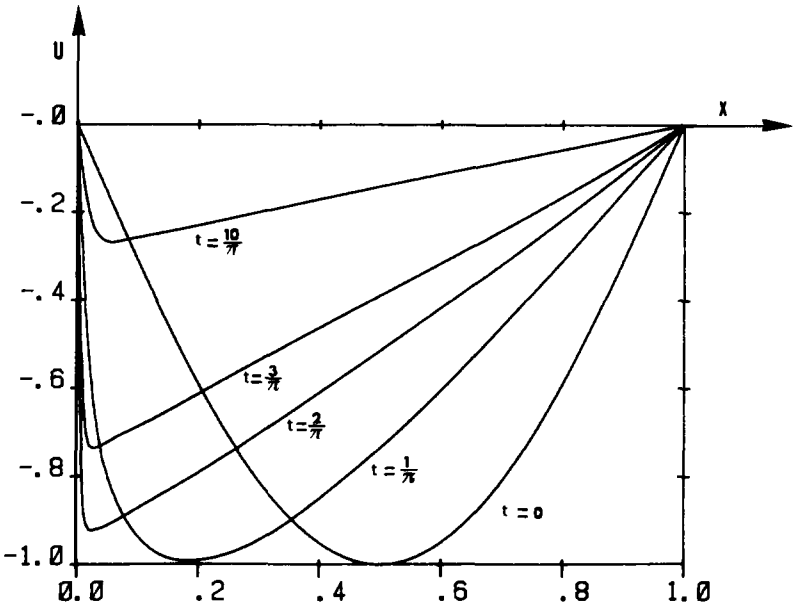

Fig. 5. Solution of the Chebyshev spectral ABCN scheme ($N = 64$, $\Delta t = 1/300\pi$).

4.2.3 *Spectral element method.* This technique shares the rapid convergence of spectral methods for problems with smooth solutions, but has the important advantage of being able to distribute points where they are required for a particular problem. Although this is of limited use for solutions with singularities, for solutions with nonsingular inner layers, this can significantly reduce the number of grid points required. This is demonstrated in Table 3, where we have chosen $M = 4$ elements, $[-1, -0.05]$, $[-0.05, 0]$, $[0, 0.05]$, $[0.05, 1.0]$ and $N = 16$. Due to the point crowding near $x = 0$, good accuracy is obtained with limited spatial resolution. An inevitable by-product of the good interior resolution is a small time step; however the inefficiency associated with this can be greatly reduced by using subcycling [15], in which several (inexpensive) convective steps are taken for each implicit viscous step.

4.3. *Finite difference methods*

Finite difference calculations provide values of the slope at the origin which are close to the curve shown on Fig. 6. The finite difference results are in good agreement with

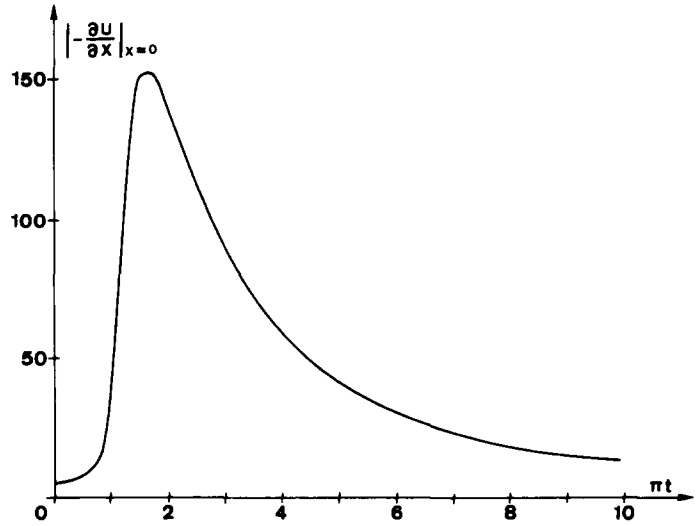

Fig. 6. Evolution of the absolute value of the slope at the origin versus time.

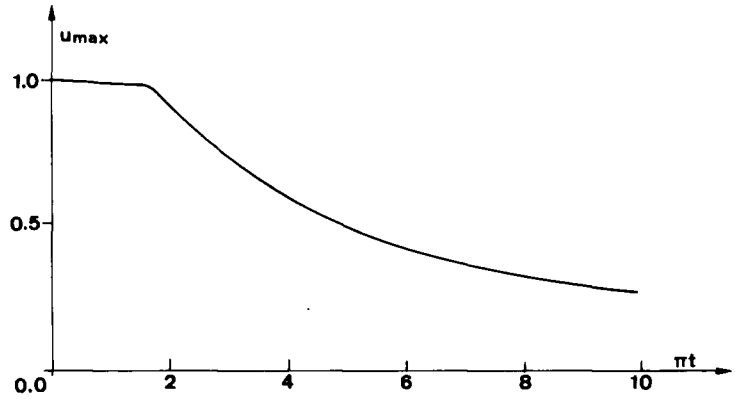

Fig. 7. Evolution of $u_{max} = \max_x |u(x, t)|$.

these produced by the ABCN spectral scheme. The time integration used the three-point backward discretization. The numerical results are presented in Table 3 and will be discussed below.

5. COMPARISON AND CONCLUSIONS

As the computations reported in this paper were performed at different places on different computers it is extremely difficult to relate the computing times of the various methods. Consequently it is impossible to carry out any precise comparison between algorithms as regards efficiency. However, a few comments may be done.

The ABCN Tau scheme requires the solution of a linear system. Several choices are available:

(i) LU factorization in a preprocessing stage followed by the solution of two triangular systems at each time step.

(ii) Inversion of the matrix at the beginning of the time integration followed by a matrix multiplication at each time step. This procedure can be very efficient on a vector computer.

(iii) Decoupling of even and odd modes and solution at each step of two quasi-tridiagonal systems.

The collocation ABCN scheme may either use (i) or (ii). The Rosenbrock algorithm is more expensive because the matrix buildup has to be performed at each step together

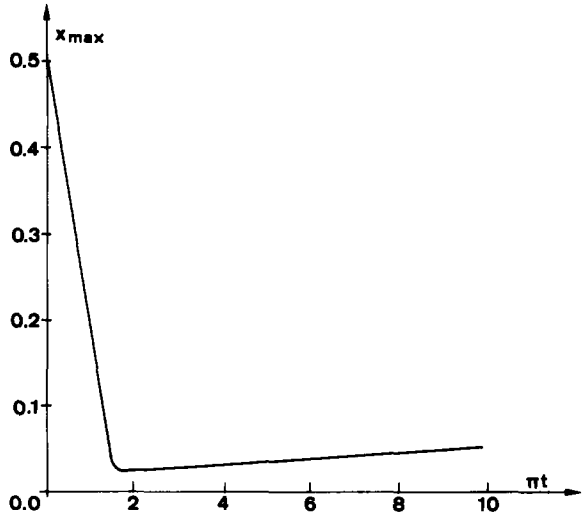

Fig. 8. Evolution of x_{max}, the abscissa of u_{max}.

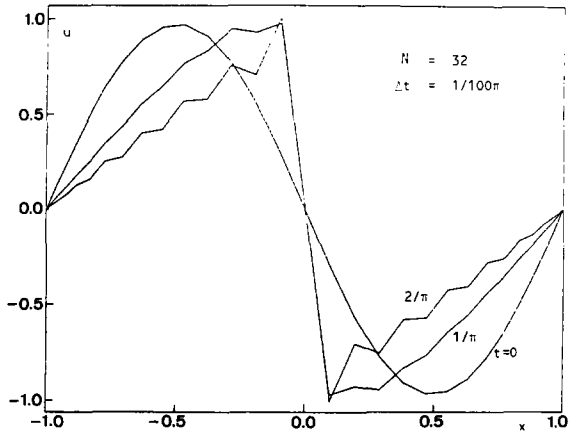

Fig. 9. Solution at different times for various values of N ($\nu = 10^{-2}/\pi$ $-1 \leqslant x \leqslant 1$) by the Chebyshev collocation ABCN scheme.

Fig. 9(a).

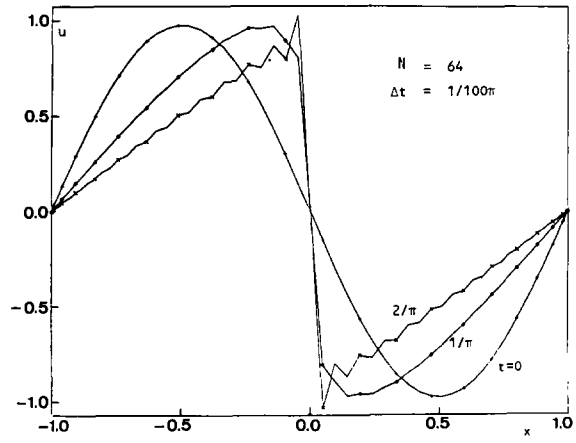

Fig. 9(b).

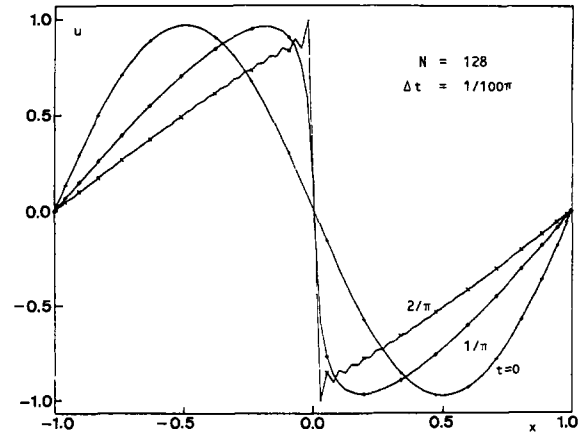

Fig. 9(c).

with the solution of the linear system. As for the other Chebyshev techniques, the major computational expense for the spectral element method is solution of the implicit equations associated with the diffusion terms. Use of static condensation allows this inversion to be done relatively efficiently, as this procedure reduces all coupling between

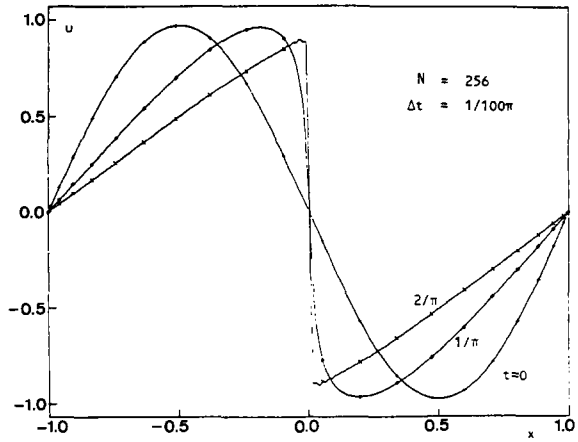

Fig. 9(d).

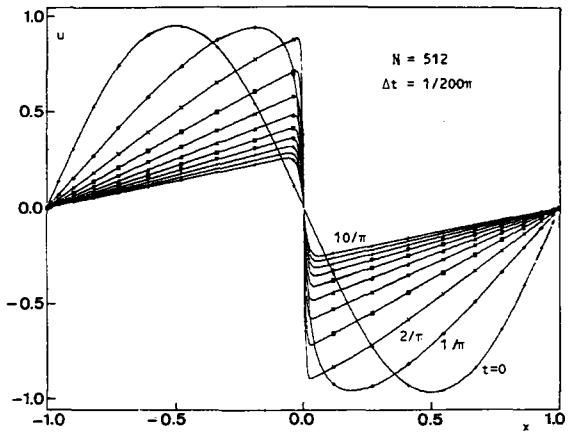

Fig. 9(e).

elements to elemental boundary points, with the remainder of the elemental calculations being performed independently. This procedure is particularly effective on a parallel processor, where operation counts commensurate with substructured low-order techniques can be obtained [10]. The finite difference method involves the solution of a tridiagonal system at each step. Finally, the explicit Fourier method is probably the most efficient per time as it involves only FFT's.

If one treats the full domain $[-1, 1]$, the numerical experiments show that the presence of the very sharp sawtooth profile requires many grid points or expansion functions. The spectacular rate of convergence of spectral schemes which is of "infinite order" [1] for infinitely differentiable functions is only obtained, if ever, for N unmoderately large. All spectral schemes (Fourier Galerkin, Fourier pseudospectral, ABCN collocation) perform poorly for $\Delta t = 10^{-2}/\pi$ even with large cutoffs ($N/M = 682/1024$ or $N = 512$). However, better results are obtained if the time step is reduced to small values like $\Delta t = 5.10^{-4}/\pi$. Well-chosen coordinate transformations [(3.18) and (3.25)] and elemental decomposition alleviate the difficulty and help to recover acceptable accuracy (see ABCN collocation with coordinate transformation and FD).

For half the domain $[0, 1]$, the accumulation of Chebyshev points near the origin is favorable for the achievable precision. This allows the user to decrease the number of grid points to $N = 64$. The results obtained on $[0, 1]$ are roughly speaking of the same order of precision as those coming from the previous methods working on $[-1, 1]$ with a coordinate transformation or element method; however, the latter are more flexible.

From Table 3, we can notice that the FD scheme yields two significant digits for $|\partial u/\partial x|_{\max}$ at the origin, while the ABCN spectral and collocation techniques (ABCN and

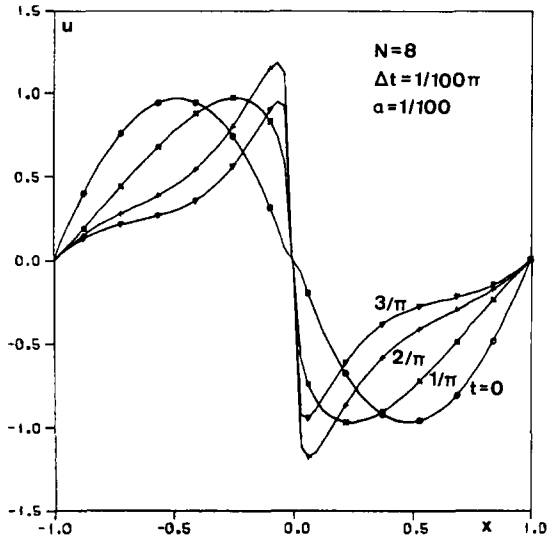

Fig. 10. Solution at different times for various values of N by the Chebyshev collocation ABCN scheme ($-1 \leqslant x \leqslant 1$) using a mapping $x = g(\zeta)$.

Fig. 10(a).

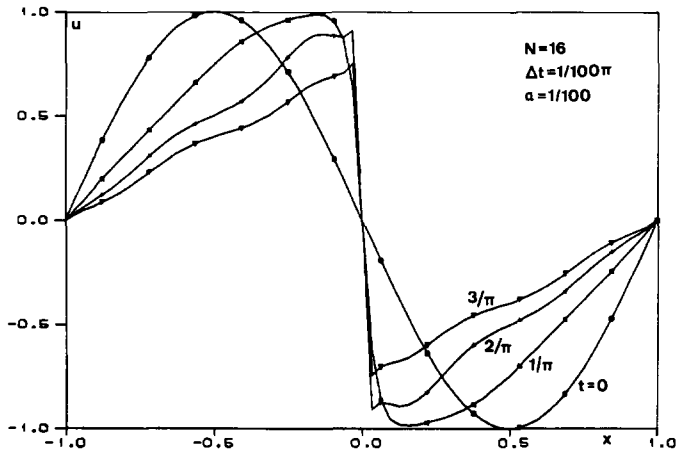

Fig. 10(b).

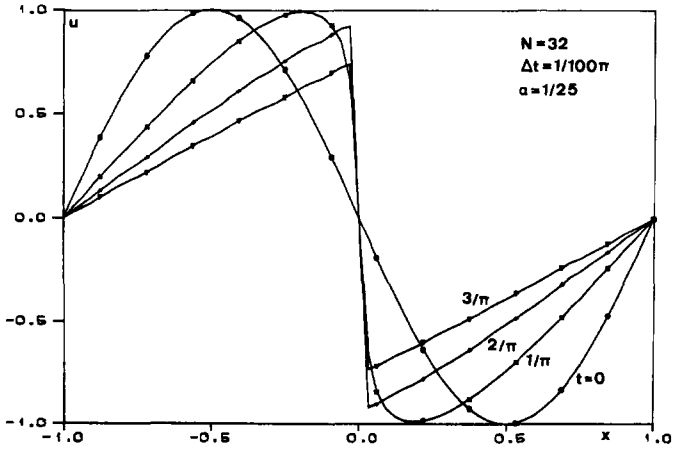

Fig. 10(c).

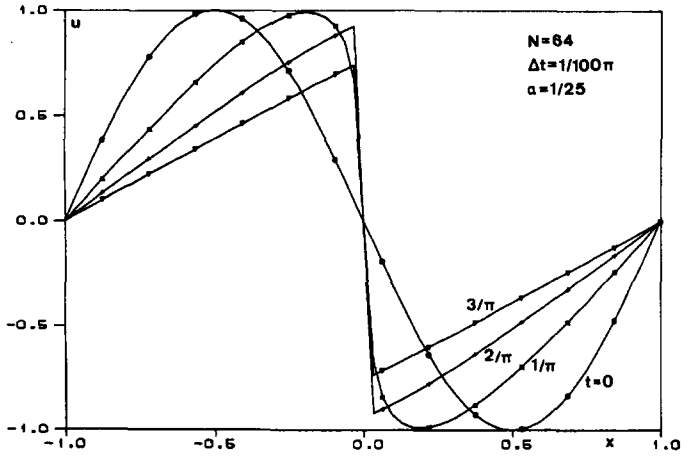

Fig. 10(d).

spectral element) produce four significant digits. This last accuracy is also achieved by the Fourier–Galerkin scheme with $\Delta t = 5.10^{-4}/\pi$. The spectral Rosenbrock algorithm which is third order in time gives five significant digits with $N = 64$ and still retains two significant digits when $N = 32$.

If we inspect now the value of the time when $|\partial u/\partial x|_{\max}$ is attained, the time marching schemes with spectral (Fourier or Chebyshev) approximation are within 1% from the exact value. In contrast, the finite difference scheme presents a large time lag as the maximum value appears at later times than the analytical solution. This is due to lagging phase errors which characterize very often the second-order finite difference schemes. Let us note that for equivalent spatial approximations (for example, Chebyshev schemes on [0, 1]), the higher the accuracy in time, the better the resulting value. As a conclusion, with spectral discretization, higher-order time schemes should be used.

The coordinate transformation (3.18) deserves a comment. The first collocation point in the neighborhood of the origin in the physical plane is $x \sim \pm\alpha\pi/N$. From Table 2, an optimum grid is obtained for $\alpha = \frac{1}{23}$. At smaller α values, the scheme degrades and the placing of grid points closer to the origin does not improve the numerical results. A possible explanation comes from the error associated to the matrix computation and its inverse. In the matrix elements, α^2 terms appear at the denominator, which might grow unduly for small α values. As for the coordinate transformation, the spectral element decomposition allows placement of points where they are required, with a corresponding

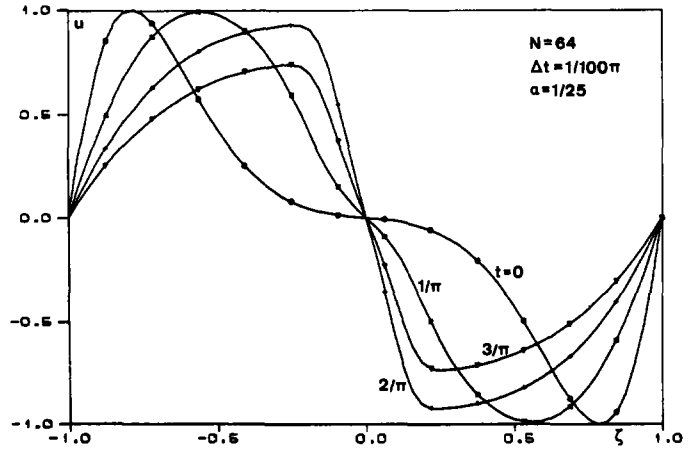

Fig. 11. Solution at different times by the Chebyshev collocation ABCN scheme in the transformed plane ζ.

Table 2. Maximum value of the slope at the origin and corresponding time value for $N = 64$, $\Delta t = 1/100\pi$ with respect to the parameter α of the coordinate transformation (3.18)

α	$\lvert\partial u/\partial x\rvert_{max}$	∂t_{max}
1	37.3959	1.54
1/10	150.7737	1.60
1/20	151.7396	1.60
1/25	152.1225	1.60
1/40	153.3561	1.60
1/50	153.3257	1.60
1/60	152.6868	1.59
1/100	146.6897	1.59

Table 3. Maximum absolute value of the slope at the origin and time t_{max} at which this maximum value occurs for all the numerical algorithms and for the analytical solution.

Method	Interval	$\left\lvert\frac{\partial u}{\partial x}\right\rvert_{max}$	$\pi \cdot t_{max}$	N/M	$\Delta t \cdot \pi$
- Fourier Galerkin	[-1,1]	151.942	1.6035	682/1024	5.10^{-4}
		142.665	1.60	682/1024	10^{-2}
		148.975	1.603	170/256	5.10^{-4}
		142.313	1.60	170/256	10^{-2}
-Fourier Pseudospectral	[-1,1]	142.606	1.60	256/256	10^{-2}
		144.237	1.60	128/128	10^{-2}
-ABCN collocation	[-1,1]	145.877	1.60	512	5.10^{-3}
+coordinate transform.	[-1,1]	152.123	1.60	64	10^{-2}
-Spectral Element	[-1,1]	152.04	1.6033	16 x 4	$10^{-2}/6$
- FD	[-1,1]	150.1	1.63	81	10^{-2}
- Chebyshev					
ABCN spectral	[0,1]	152.05	1.60	64	1/300
Rosenbrock spectral	[0,1]	151.998	1.60	64	10^{-2}
	[0,1]	150.144	1.60	32	10^{-2}
ABCN collocation	[0,1]	152.126	1.60	64	10^{-2}
- Analytical		152.00516	1.6037		

increase in accuracy (see Table 3). Unlike the transformation method, the variations in resolution are local, not global, and would appear to be more general for multidimensional problems.

From this investigation, we conclude that spectral methods are not well suited to the calculation of thin inner layers. If a sufficient number of modes is used and if a good coordinate transformation or domain decomposition is designed, the accuracy in the sharp region is quite acceptable. Standard numerical schemes such as finite differences resort also to coordinate transformation, but suffer from numerical dispersion. Spectral methods are better from this point of view and call for higher-order time marching schemes.

Acknowledgements—The Nice team performed all calculations on the Cray-1 computer of Centre de Calcul Vectoriel pour la Recherche. Louvain-la-Neuve and M.I.T. acknowledge the support of NATO, Grant No. SA. 5-2-05 RG (035/84) for international collaboration.

REFERENCES

1. D. Gottlieb and S. A. Orszag, *Numerical Analysis of Spectral Methods: Theory and Applications. SIAM Monograph no.* 26, SIAM, Philadelphia, U.S.A. (1977).

2. J. D. Cole, On a quasi-linear parabolic equation occurring in aerodynamics. *Quart. Appl. Math.* **9**, 225–236 (1951).
3. E. R. Benton and G. N. Platzmann, A table of solutions of one-dimensional Burgers equation. *Quart. Appl. Math.* **29**, 195–212 (1972).
4. M. Deville, P. Haldenwang and G. Labrosse, Comparison of time integration (finite difference and spectral) for the nonlinear Burgers equation. *Notes on Numerical Fluid Mechanics* (Edited by H. Viviand), Vol. 5, pp. 64–76. Vieweg Verlag, Braunschweig (1982).
5. J. Villadsen and M. L. Michelsen, *Solution of Differential Equation Models by Polynomial Approximation.* Prentice-Hall, Englewood Cliffs, New Jersey (1978).
6. C. Lanczos, *Applied Analysis.* Prentice-Hall, Englewood Cliffs, New Jersey (1956).
7. B. A. Finlayson, *The Method of Weighted Residuals and Variational Principles.* Academic Press, New York (1972).
8. J. Ouazzani and R. Peyret, A pseudo-spectral solution of binary gas mixture flows. *Notes on Numerical Fluid Mechanics* (Edited by Pandolfi/Piva), Vol. 7, pp. 275–282. Vieweg Verlag, Braunschweig (1984).
9. A. T. Patera, A spectral element method for fluid dynamics: Laminar flow in a channel expansion. *J. Comput. Phys.* **54**, 468–488 (1984).
10. K. Z. Korczak and A. T. Patera, A spectral element method applied to unsteady flows at moderate Reynolds number. *Proc. 9th Int. Conf. on Num. Meth. Fluid Dyn.*, Paris, France, Springer, New York (1984).
11. I. Babuska and M. R. Dorr, Error estimates for the combined h and p versions of the finite element method. *Numer. Math.* **37**, 257–277 (1981).
12. W. R. Briley and H. McDonald, Solution of three-dimensional compressible Navier–Stokes equations by an implicit technique. *Proceed. 4th Internat. Conf. Numer. Meth. Fluid Dynamics,* Boulder, CO, Springer, New York (1975).
13. P. Orlandi and M. Briscolini, Direct simulation of burgulence. *Proceed.* 1983 *Internat. Conf. on Computational Techniques and Applications* (Edited by J. Noye), North-Holland, Amsterdam (1984).
14. C. Basdevant, B. Legras, R. Sadourny, and M. Beland, A study of barotropic model flows; Intermittency, waves and predictability. *J. Atmos. Sci.* **38**, 2305–2326 (1981).
15. P. M. Gresho, S. T. Chan, R. L. Lee, and C. D. Upson, A modified Finite Element Method for solving the time-dependent, incompressible Navier–Stokes equations. Part 1: Theory. *Int. J. for Num. Meth. in Fluids* **4**, 557–597 (1984).