Journal of Computational and Applied Mathematics 134 (2001) 243–258

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.com/locate/cam

The linear rational collocation method

Richard Baltensperger, Jean-Paul Berrut *

Département de Mathématiques, Université de Fribourg, Pérolles, CH-1700 Fribourg, Switzerland

Received 13 September 1999; received in revised form 10 April 2000

Abstract

We introduce the collocation method based on linear rational interpolation for solving general hyperbolic problems, prove its stability and its convergence in weighted norms and give numerical examples for its use. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Linear rational interpolation; Time evolution partial differential equations; Spectral method; Conformal point shifts

0. Introduction

Consider the first-order hyperbolic initial boundary value problem

$$\begin{aligned} &u_t(x,t) - u_x(x,t) = 0, \quad x \in\,]-1,1[, \;\; t \in\,]0,T], \\ &u(x,0) = f(x), \quad x \in\,]-1,1[, \\ &u(1,t) = 0, \quad t \in [0,T], \;\; T > 0, \end{aligned} \tag{1}$$

whose exact solution is given by

$$u(x,t) = \begin{cases} f(x+t), & \text{if } x+t \leqslant 1, \\ 0, & \text{otherwise.} \end{cases}$$

Problem (1) can be viewed as a model problem for the more general hyperbolic systems with appropriately specified boundary conditions we will address in Section 2.

* Corresponding author. Tel.: +41-26-300-9196; fax: +41-26-300-9744.

E-mail addresses: richard.baltensperger@unifr.ch (R. Baltensperger), jean-paul.berrut@unifr.ch (J.-P. Berrut).

PII: S0377-0427(00)00552-5

The (spectral) collocation Čebyšev method for solving (1) is based on approximating the solution by a polynomial of degree $\leqslant N$ in the space variable x [7, p. 7],

$$u_N(x,t) := \sum_{k=0}^{N} \tilde{u}(x_k,t)L_k(x) \quad \text{where } L_k(x) := \prod_{j=0, j\neq k}^{N} \frac{x-x_j}{x_k-x_j}, \tag{2}$$

which takes unknown values $\tilde{u}(x_k,t)$ at the Čebyšev extremal points

$$x_k := \cos\frac{k\pi}{N}, \quad k = 0(1)N. \tag{3}$$

$L_k(x)$ in (2) is the Lagrange polynomial associated with x_k. The $\tilde{u}(x_k,t)$ are determined by requiring that the polynomial $u_N(x,t)$ satisfies

$$\begin{aligned} &u_{N,t}(x_j,t) - u_{N,x}(x_j,t) = 0, \quad j = 1(1)N, \ t \in]0,T], \\ &u_N(x,0) = P_N[f](x), \quad x \in]-1,1[, \\ &u_N(1,t) = 0, \quad t \in [0,T], \end{aligned} \tag{4}$$

where $P_N[f](x)$ is the polynomial of degree $\leqslant N$ interpolating f between the Čebyšev points of the second kind (3). In (4) and below, the second index indicates the variable with respect to which differentiation occurs. We can write the first equation of (4) in the matrix form

$$\boldsymbol{u}_{N,t}(t) = \boldsymbol{D}^{(1)}\boldsymbol{u}_N(t), \tag{5}$$

where

$$\boldsymbol{u}_N(t) := [u_N(x_1,t),\ldots,u_N(x_N,t)]^{\mathrm{T}}, \qquad D_{ij}^{(1)} := L_j'(x_i), \quad i,j = 1(1)N.$$

The entries of $\boldsymbol{D}^{(1)}$ can be calculated analytically [6, p. 724, 7, p. 69]. The system of ODEs (5) is then solved by an appropriate time marching technique like a Runge–Kutta or a multistep method.

The main reason for using spectral methods is their higher order which often yields, for a given number of values of the functions appearing in Eq. (1), a far greater accuracy than finite difference or finite element methods: u_N converges spectrally toward u.

Unfortunately, explicit time marching techniques are subject to very tight stability limitations: the time step restriction is $\Delta t = \mathcal{O}(N^{-2})$ for the model problem (1) and $\Delta t = \mathcal{O}(N^{-4})$ for parabolic problems, see [18].

Kosloff and Tal-Ezer have suggested in [13] a transformation approach to overcome these stability restrictions. They have considered the transformation

$$x := g(y,\alpha) = \frac{\arcsin(\alpha y)}{\arcsin(\alpha)}, \quad x,y \in [-1,1], \ \alpha \in (0,1), \tag{6}$$

and the interpolant

$$v_N(x,t) := \sum_{k=0}^{N} \tilde{v}(x_k,t)L_k(g^{-1}(x)) \tag{7}$$

which takes the values $\tilde{v}(x_k,t)$ at the grid points $x_k = g(y_k,\alpha)$, $k = 0(1)N$. The derivatives at the grid points must then be calculated with the chain rule. This results in the computation of sums of products of $N \times N$-matrices. But by a suitable choice of the parameter α in the "stretching function" (6) the stability condition becomes much more favorable.

We present here a new collocation technique, *the linear rational collocation method*. We set the unknown function as a rational interpolant written in its *barycentric form*. The resulting stability condition for the model problem (1) is again weaker than that of the Čebyšev collocation method; moreover, the derivatives of the interpolant may be calculated without the chain rule by using formulas discovered by Schneider and Werner [15].

We describe the method for the hyperbolic problem (1) in the first section. In the second section, we recall results given in [9] about stability and convergence of the polynomial collocation method for general hyperbolic problems like (18). Then, in the third section, we prove the stability and the convergence of the linear rational collocation method for these same problems. Finally, we present numerical results and compare them with those obtained with the classical Čebyšev collocation method.

1. The linear rational collocation method

Let x_k, $k = 0(1)N$, be a set of distinct interpolation points (or nodes). With every vector $\beta = [\beta_0, \dots, \beta_N]$, $\beta_k \neq 0$ for all k, we associate the linear space $\mathscr{R}_N^{(\beta)}$ spanned by the functions

$$L_k^{(\beta)}(x) := \frac{\frac{\beta_k}{x-x_k}}{\sum_{j=0}^N \frac{\beta_j}{x-x_j}}, \quad k = 0(1)N. \tag{8}$$

$L_k^{(\beta)}$ is the Lagrange fundamental rational function with denominator

$$q_N(x) := L(x) \sum_{j=0}^N \frac{\beta_j}{x - x_j}, \quad L(x) := \prod_{j=0}^N (x - x_j), \tag{9}$$

which interpolates the values δ_{jk}, $j = 0(1)N$, at the x_j's. The β_k are called the *weights* of the $L_k^{(\beta)}(x)$. Note that q_N is independent of k.

To any function $u(x,t)$ we can then associate its interpolant in $\mathscr{R}_N^{(\beta)}$ given by

$$r_N(x,t) := \sum_{k=0}^N \tilde{u}(x_k,t) L_k^{(\beta)}(x) = \frac{\sum_{k=0}^N \frac{\beta_k}{x-x_k} \tilde{r}(x_k,t)}{\sum_{k=0}^N \frac{\beta_k}{x-x_k}}, \tag{10}$$

which takes the values $\tilde{r}(x_k,t)$ at the points x_k, $k = 0(1)N$.

If we assume that q_N does not have any zero in the interval of collocation (which implies that the β_k alternate signs [5,15]), we can use the formulas given in [15] for differentiating the rational function $r_N(x,t)$ with respect to the x-variable. This defines the linear rational collocation method for the model problem (1) as

$$\begin{aligned} &r_{N,t}(x_j,t) - r_{N,x}(x_j,t) = 0, \quad j = 1(1)N, \ t \in]0,T], \\ &r_N(x,0) = R_N[f](x), \quad x \in]-1,1[, \\ &r_N(1,t) = 0, \quad t \in [0,T], \end{aligned} \tag{11}$$

where $R_N[f](x)$ is the rational function with denominator (9) interpolating f between the x_k's.

As in the polynomial collocation case, we can write the system (11) in the matrix form

$$\boldsymbol{r}_{N,t}(t) = \boldsymbol{D}^{(1)} \boldsymbol{r}_N(t), \tag{12}$$

where

$$\boldsymbol{r}_N(t) := [r_N(x_1,t),\ldots,r_N(x_N,t)]^{\mathrm{T}}, \qquad D_{ij}^{(1)} := (L_j^{(\beta)})'(x_i), \quad i,j = 1(1)N.$$

The entries of $\boldsymbol{D}^{(1)}$ are given by (see [3])

$$D_{ij}^{(1)} = \begin{cases} \dfrac{\beta_j}{\beta_i}\dfrac{1}{x_i - x_j}, & \text{if } i \neq j, \\ -\displaystyle\sum_{k=0,k\neq i}^{N} \dfrac{\beta_k}{\beta_i}\dfrac{1}{x_i - x_k}, & \text{if } i = j. \end{cases} \tag{13}$$

The numerical effort needed for solving (12) or (5) is the same.

To see the connection between the rational function we employ for the rational collocation method, on one side, and the polynomial used in the classical Čebyšev collocation method on the other side, we first rewrite the polynomial $u_N(x,t)$ in its barycentric form [12,14]

$$u_N(x,t) = \frac{\sum_{k=0}^{N''} \frac{(-1)^k}{x-x_k}\tilde{u}(x_k,t)}{\sum_{k=0}^{N''} \frac{(-1)^k}{x-x_k}}, \tag{14}$$

where the $''$ means that the first and the last terms of the sum are to be halved.

It has been shown in [3] that, if u is analytic (in the spatial variable x) in an ellipse and if the collocation points x_k are shifted by a conformal map – with the same weights, the resulting interpolant still converges spectrally (at the same rate as the polynomial interpolating between the Čebyšev points).

(14) is then a rational function (10) with $\beta_0 := 1/2$, $\beta_k := (-1)^k$, $k = 1(1)N-1$ and $\beta_N := (-1)^N/2$; it does not have any pole in the interval of interpolation [4].

We will now briefly reconstruct the rational function $r_N(x,t)$ as the quotient of two interpolants, as in [3]. Let $x \in I$, $y \in J$, where I, J are two intervals in $\mathbb{R}$, let g be a conformal map from a relatively compact domain $\mathscr{D}_1$ in $\mathbb{C}$ containing J to another relatively compact domain $\mathscr{D}_2$ containing I and such that $g(J) = I$. Without loss of generality we set $J := [-1,1]$. Finally, let u be a complex-valued function defined on the interval I (in the space variable x).

We define $x_k := g(y_k)$, where the y_k, $k = 0(1)N$, are the Čebyšev points of the second kind in J, and we study the rational interpolant r_N of the function $u : I \times [0,T] \to \mathbb{C}$ between the x_k's.

Let $s : \mathscr{D}_1 \times \mathscr{D}_1 \to \mathbb{C}$ be the analytic function of two variables

$$s(y,z) := \frac{y - z}{g(y) - g(z)}.$$

In order to interpolate the function $\tilde{u}(y,t) := u(g(y),t)$ on the interval J, we write it as

$$u(g(y),t) = \frac{u(g(y),t)s(y,z)}{s(y,z)} =: \frac{v(y,z,t)}{s(y,z)}, \tag{15}$$

where $(y,z) \in J \times J$. We freeze the variable z and we construct the polynomial interpolating $v(y,z,t)$ between the $N+1$ Čebyšev points of the second kind $y_k = \cos(k\pi/N)$,

$$p_N[v](y,z,t) := \sum_{k=0}^{N} v(y_k,z,t)L_k(y) = \frac{2^{N-1}}{N}L(y)\sum_{k=0}^{N''}\frac{(-1)^k}{y - y_k}v(y_k,z,t), \tag{16}$$

where

$$L_k(y) := \prod_{j=0, j\neq k}^{N} \frac{y - y_j}{y_k - y_j} \quad \text{and} \quad L(y) := \prod_{k=0}^{N}(y - y_k).$$

Repeating the same process, we interpolate the denominator of (15) by

$$q_N[s](y,z) := \frac{2^{N-1}}{N} L(y) \sum_{k=0}^{N''} \frac{(-1)^k}{y - y_k} s(y_k, z). \tag{17}$$

In the special case $z = y \neq y_k$, if we form the quotient of the two functions (16) and (17) and if we set $x := g(y)$ and $x_k := g(y_k)$, $k = 0(1)N$, the result is precisely the linear rational interpolant (14) of $u(x,t)$ between the (conformally) transformed Čebyšev points x_k:

$$r_N[u] \equiv \frac{p_N[v]}{q_N[s]}.$$

2. Stability and convergence results for the polynomial collocation method applied to hyperbolic problems

In [9], Canuto and Quarteroni have given stability results and error estimates for spectral and (spectral) collocation approximations of hyperbolic equations. We briefly recall them here for the collocation case.

Let us consider the interval $I :=]-1, 1[\subset \mathbb{R}$ and its boundary $\Gamma := \partial I = \{-1, 1\}$. For a given weight function $w(x)$ over I we define

$$L_w^2(I) := \{\phi : I \to \mathbb{R} \mid \phi \text{ is measurable and } (\phi,\phi)_w < \infty\},$$

with $(\phi,\psi)_w := \int_I \phi(x)\psi(x)w(x)\,\mathrm{d}x$ and $\|\phi\|_{0,w} := (\phi,\phi)_w^{1/2}$. For any integer $k > 0$, we consider the weighted Sobolev space

$$H_w^k(I) := \{\phi \in L_w^2(I) \mid \mathrm{d}^m\phi/\mathrm{d}x^m \in L_w^2(I),\ 0 \leqslant m \leqslant k\},$$

with the norm $\|\phi\|_{k,w}^2 := \sum_{m=0}^{k} \|\mathrm{d}^m\phi/\mathrm{d}x^m\|_{0,w}^2$. Note that $H_w^0(I) \equiv L_w^2(I)$. In the following C will be a generic positive constant independent of the discretization parameter N and of $u(x,t)$, the solution of the problem.

Suppose we are given two functions b and b_0 such that b, $b_x := \partial b/\partial x$ and b_0 all belong to $L^\infty(I)$, and set $\Gamma^- := \{x \in \Gamma \mid xb(x) < 0\}$, $\Gamma^+ := \Gamma \backslash \Gamma^-$. Let $f = f(x,t)$ and $u^0(x)$ be two assigned functions, and consider the following general hyperbolic problem ($T > 0$):

$$\begin{aligned} &u_t(x,t) + (bu)_x(x,t) + b_0(x)u(x,t) = f(x,t), \quad x \in I,\ t \in]0,T], \\ &u(x,0) = u^0(x), \quad x \in I, \\ &u(x,t) = 0, \quad x \in \Gamma^-,\ t \in]0,T]. \end{aligned} \tag{18}$$

We assume that $u \in L^2(H_w^1)$ and $u_t \in L^2(L_w^2)$ for suitable weight functions to be defined later; here

$$L^p(H_w^k) := \left\{\phi : [0,T] \to H_w^k \mid \phi \text{ is measurable and } \int_0^T \|\phi(t)\|_{k,w}^p \,\mathrm{d}t < \infty\right\}.$$

Moreover, we assume that b and b_0 satisfy the following condition (which can always be arranged by a change of unknown $u(x,t) \to e^{\lambda t}u(x,t)$):

$$\tfrac{1}{2}\, b_x + b_0 - \tfrac{1}{2}\, bw_x/w \geqslant 0 \quad \text{in } I. \tag{19}$$

In what follows $w(x)$ will be the Jacobi weight function defined as

$$w(x) := (1+x)^{e^-}(1-x)^{e^+}\hat{w}(x),$$

$$e^{\pm} := \begin{cases} 0 & \text{if } \pm 1 \in \Gamma^-, \\ 1 & \text{if } \pm 1 \in \Gamma^+, \end{cases}$$

$$\hat{w}(x) := (1-x^2)^{-1/2}. \tag{20}$$

For the (spectral) collocation Jacobi approximation, we define the space

$$V_N := \{\phi \in \mathscr{P}_N(I) \mid \phi(x) = 0 \,\forall x \in \Gamma^-\},$$

where $\mathscr{P}_N(I)$ denotes the linear space of the polynomials of degree $\leqslant N$ on I. Let $\{\hat{x}_j, \hat{w}_j\}$, $j{=}0(1)N$, denote the nodes and the weights of the Gaussian integration formula relatively to the weight $\hat{w}$ and with the points of Γ^- as preassigned nodes. Depending on Γ^-, we are dealing with the following formulae:

- $\Gamma^- = \emptyset$: (Čebyšev–)Gauss integration formula;
- $\Gamma^- = \Gamma$: (Čebyšev–)Gauss–Lobatto integration formula;
- $\Gamma^- = \{-1\}$: (Čebyšev–)Gauss–Radau integration formula with $x_0 = -1$;
- $\Gamma^- = \{1\}$: (Čebyšev–)Gauss–Radau integration formula with $x_N = 1$.

Hereafter we assume $b_0 \in C^0(\bar{I})$, $b \in C^1(\bar{I})$, $f \in C^0([0,T] \times \bar{I})$ and $u^0 \in C^0(\bar{I})$. The collocation approximation to (18) is the solution of the following problem, which takes advantage of the skew-symmetric decomposition of the operator: find $u_N \in C^1(V_N)$ such that

$$u_{N,t}(\hat{x}_j,t) + \tfrac{1}{2}\,[bu_{N,x} + P_{N,x}[bu_N]](\hat{x}_j,t) + (\tfrac{1}{2}\, b_x + b_0)u_N(\hat{x}_j,t) = f(\hat{x}_j,t), \tag{21}$$

$$j = 0(1)N,\ \ \hat{x}_j \notin \Gamma^-,\ \ t \in \,]0,T],$$

$$u_N(\hat{x}_j,0) = u^0(\hat{x}_j), \quad j = 0(1)N.$$

We set $w_j := (1+\hat{x}_j)^{e^-}(1-\hat{x}_j)^{e^+}\hat{w}_j$, $j = 0(1)N$,

$$(\phi,\psi)_{N,w} := \sum_{j=0}^{N} \phi(\hat{x}_j)\psi(\hat{x}_j)w_j, \quad \forall \phi,\psi \in C^0(\bar{I}),$$

and the discrete norm $\|\phi\|_{N,w} := (\phi,\phi)_{N,w}^{1/2}$.

With these norms we have the following theorems [9, pp. 635–636].

Theorem 1. *The following stability result holds*:

$$\|u_N\|_{L^\infty(L^2_w)} \leqslant C\left(\|u^0\|_{N,w} + \left(\int_0^T \|f(t)\|^2_{N,w}\,dt\right)^{1/2}\right). \tag{22}$$

Theorem 2. *Assume that* $u \in L^\infty(H^\sigma_{\hat{w}})$ *and that* $b \in H^\sigma_{\hat{w}}$ *for* $\sigma > 2$. *Then*

$$\|u - u_N\|_{L^\infty(L^2_{\hat{w}})} \leqslant CN^{2-\sigma}\|u\|_{L^\infty(H^\sigma_{\hat{w}})}. \tag{23}$$

We see that the order of convergence of u_N toward u depends only on the regularity of the latter.

3. Stability and convergence results for the linear rational collocation method

For the linear rational approximation, we define the space

$$W_N := \{\phi \in \mathscr{R}_N^{(\beta)} \mid \phi(x) = 0 \ \forall x \in \Gamma^-\},$$

where $\mathscr{R}_N^{(\beta)}$ denotes the space of the rational functions spanned by the functions $L_k^{(\beta)}(x)$ with $\beta := [1/2, -1, 1, \ldots, (-1)^j, \ldots, (-1)^N/2]$.

Again we assume $b_0 \in C^0(\bar{I})$, $b \in C^1(\bar{I})$, $f \in C^0([0,T] \times \bar{I})$ and $u^0 \in C^0(\bar{I})$. The linear rational (spectral) collocation approximation to the hyperbolic problem (18) is defined here as the solution of the following problem: find $r_N \in C^1(W_N)$ such that

$$\begin{aligned} & r_{N,t}(\hat{x}_j, t) + \tfrac{1}{2}[br_{N,x} + R_{N,x}[br_N]](\hat{x}_j, t) + (\tfrac{1}{2} b_x + b_0) r_N(\hat{x}_j, t) = f(\hat{x}_j, t), \\ & \quad j = 0(1)N, \ \ \hat{x}_j \notin \Gamma^-, \ \ t \in]0, T], \\ & r_N(\hat{x}_j, 0) = u^0(\hat{x}_j), \quad j = 0(1)N, \end{aligned} \tag{24}$$

where $\hat{x}_j := g(\hat{y}_j)$ and where the $\hat{y}_j$ are the Gauss points depending on the problem to be solved. $R_N[br_N](x)$ is the rational function with denominator (9) interpolating br_N. Note that the $\hat{x}_j$'s are in I and the $\hat{y}_j$'s are in $J :=]-1, 1[$. In the following, $\Gamma := \partial J = \{-1, 1\}$; moreover, we will assume that the conformal map g is such that $g(-1) = -1$ and $g(1) = 1$ and therefore $g' > 0$ on J.

To study stability and convergence, we will show that the rational collocation scheme is equivalent to a polynomial collocation scheme applied to an *associate* problem in a transformed space. Then we will prove that the latter scheme is stable and convergent.

Theorem 3. *The* (*skew-symmetric*) *rational collocation method* (24) *is equivalent to a polynomial collocation method in the transformed space.*

Proof. We rewrite the solution in the transformed space using the variable transformation $x := g(y)$. The solution $u(x,t)$ of problem (18) becomes $\tilde{u}(y,t) := u(g(y), t)$, which we multiply by the quotient

$$\frac{q_N[s](y,z)}{q_N[s](y,z)}.$$

For sufficiently large N one has $q_N[s](y,z) \neq 0 \ \forall y, z$ ($q_N[s]$ is a convergent approximation of the function $s \neq 0$, see [3] — we conjecture that $q_N[s]$ does not vanish for any $N > 0$). Defining $v(y,z,t) := \tilde{u}(y,t) \cdot q_N[s](y,z)$, $\tilde{b}(y) := b(g(y))$, $\tilde{b}_0(y) := b_0(g(y))$ and $\tilde{f}(y,t) := f(g(y),t)$, we can then rewrite (18) as an *associate* hyperbolic problem in the transformed space:

$$v_t + \frac{1}{g'}(\tilde{b}v)_y + \left(\tilde{b}_0 - \frac{\tilde{b}}{g'}\frac{q_{N,y}[s]}{q_N[s]}\right) v = \tilde{f} q_N[s], \quad y, z \in J, \ \ t \in]0, T],$$

$$v(y,z,0) = v^0(y,z) := u^0(g(y))q_N[s](y,z), \quad y,z \in J,$$

$$v(y,z,t) = 0, \quad y \in \Gamma^-, \; z \in J, \; t \in]0,T]. \tag{25}$$

Applying the skew-symmetric decomposition to the equation, we find

$$v_t + \frac{1}{2g'}[\tilde{b}v_y + (\tilde{b}v)_y] + \left[\frac{1}{2g'}\tilde{b}_y + \bar{b}_0\right] v = k, \quad y,z \in J, \; t \in]0,T], \tag{26}$$

where we have set

$$\bar{b}_0 := \tilde{b}_0 - \frac{\tilde{b}}{g'}\frac{q_{N,y}[s]}{q_N[s]} \quad \text{and} \quad k := \tilde{f}q_N[s].$$

In what follows, $V_N := \{\phi \in \mathscr{P}_N(J) \mid \phi(y) = 0 \; \forall y \in \Gamma^-\}$. The collocation scheme we apply to problem (26) is the following: find $v_N \in C^1(V_N)$ such that

$$v_{N,t}(\hat{y}_j,z,t) + \frac{1}{2g'(\hat{y}_j)}[\tilde{b}v_{N,y} + P_{N,y}[\tilde{b}v_N]](\hat{y}_j,z,t) + \left(\frac{1}{2g'}\tilde{b}_y + \bar{b}_0\right) v_N(\hat{y}_j,z,t) = k(\hat{y}_j,z,t), \tag{27}$$
$$j = 0(1)N, \; \hat{y}_j \notin \Gamma^-, \; z \in J, \; t \in]0,T],$$

$$v_N(\hat{y}_j,z,0) = v^0(\hat{y}_j,z), \quad j = 0(1)N, \; z \in J.$$

Since P_N interpolates, one has at every collocation point $\hat{y}_j$

$$\bar{b}_0 v_N = \tilde{b}_0 v_N - \frac{1}{2g'}\frac{q_{N,y}[s]\tilde{b}}{q_N[s]}v_N - \frac{q_{N,y}[s]}{2g'}\frac{P_N[\tilde{b}v_N]}{q_N[s]}.$$

When $z = \hat{y}_j$, we obtain after few manipulations

$$\frac{v_{N,t}}{q_N[s]}(\hat{y}_j,z,t) + \frac{1}{2g'(\hat{y}_j)}\left[\tilde{b}\left(\frac{v_N}{q_N[s]}\right)_y + \left(\frac{P_N[\tilde{b}v_N]}{q_N[s]}\right)_y\right](\hat{y}_j,z,t)$$
$$+ \left(\frac{1}{2g'}\tilde{b}_y + \tilde{b}_0\right)\frac{v_N}{q_N[s]}(\hat{y}_j,z,t) = \tilde{f}(\hat{y}_j,z,t), \quad j = 0(1)N, \; \hat{y}_j \notin \Gamma^-, \; z \in J, \; t \in]0,T],$$

$$v_N(\hat{y}_j,z,0) = v^0(\hat{y}_j,z), \quad j = 0(1)N, \; z \in J.$$

When $z = y$, this is precisely the skew-symmetric rational collocation method (24) in the space variable y applied to (18). □

Remark

1. We merely use the change of variable to prove stability and convergence in the transformed "y-space". Unlike [13,16] we solve the problem in the original "x-space".
2. The collocation scheme introduced in (27) is different from the classical collocation scheme (21) applied to the hyperbolic equation (25), so Theorems 1 and 2 cannot be directly applied. Nevertheless, the fact that the derivative of the transformation g differs from zero will enable us to prove stability and convergence of our scheme.

We will now prove the spatial stability (in the transformed space) of the equivalent polynomial collocation scheme. For that purpose, we shall partly apply to the rational case ideas borrowed form [9].

Theorem 4. *For sufficiently large N the following stability result holds for all z*:

$$\|v_N\|_{L^\infty(L^2_w)} \leqslant C \left(\|v^0\|_{N,w} + \left(\int_0^T \|k(t)\|^2_{N,w}\, \mathrm{d}t \right)^{1/2} \right), \tag{28}$$

with k as in the proof of Theorem 3; *therefore, for $z=y$, the rational collocation scheme is stable in the transformed space.*

Proof. Assume that

$$\tfrac{1}{2}\tilde{b}_y + g'\bar{b}_0 - \tfrac{1}{2}\tilde{b}w_y/w \geqslant 0 \quad \text{in } J \times J. \tag{29}$$

This condition can always be fulfilled, see [9]. If we first multiply (27) by g', the problem can be equivalently written in variational form: find $v_N \in C^1(V_N)$ such that

$$(g'v_{N,t}, \phi)_{N,w} + \tfrac{1}{2}((\tilde{b}v_{N,y} + P_{N,y}[\tilde{b}v_N]), \phi)_{N,w} + ((\tfrac{1}{2}\tilde{b}_y + g'\bar{b}_0)v_N, \phi)_{N,w} = (g'k, \phi)_{N,w},$$

$$\forall \phi \in V_N, \ \forall z \in J,$$

$$v_N(0) = P_N[(u^0 \circ g)q_N[s]], \quad \forall z \in J. \tag{30}$$

Repeating the same calculations as in [9] we get for all $\phi, \psi \in V_N$

$$(\tilde{b}\phi_y + P_{N,y}[\tilde{b}\phi], \psi)_{N,w} = -(\phi, P_{N,y}[\tilde{b}\psi] + \tilde{b}\psi_y)_{N,w} - T(\psi,\phi) - T(\phi,\psi),$$

where $T(\phi,\psi) := \int_{-1}^{1} P_{N,y}[\tilde{b}\phi]\psi\, w_y\, \mathrm{d}y$. Then it follows that

$$(\tilde{b}v_{N,y} + P_{N,y}[\tilde{b}v_N], v_N)_{N,w} = -T(v_N, v_N). \tag{31}$$

From the definition of w, we obtain

$$T(v_N,v_N) = \int_{-1}^{1} P_{N,y}[\tilde{b}v_N]v_N \frac{w_y}{\hat{w}}\hat{w}\, \mathrm{d}y = \sum_{j=0}^{N} \tilde{b}(\hat{y}_j)v_N^2(\hat{y}_j)\frac{w_y(\hat{y}_j)}{w(\hat{y}_j)}w_j. \tag{32}$$

Let $\phi = v_N$ in (30); from (31) and (32), we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\sum_{j=0}^{N} g'(\hat{y}_j)v_N^2(\hat{y}_j)w_j + \sum_{j=0}^{N}\left(\frac{1}{2}\tilde{b}_y + g'\bar{b}_0 - \frac{1}{2}\tilde{b}\frac{w_y}{w}\right)(\hat{y}_j)v_N^2(\hat{y}_j)w_j = (g'k, v_N)_{N,w}.$$

Using (29), we obtain the inequality

$$(g'v_{N,t}, v_N)_{N,w} \leqslant (g'k, v_N)_{N,w}.$$

Therefore we find (since $g' > 0$) for all z

$$\frac{\mathrm{d}}{\mathrm{d}t}\min_{y\in[-1,1]} g'(y)\sum_{j=0}^{N} v_N^2(\hat{y}_j,z,t)w_j \leqslant 2\sum_{j=0}^{N} g'(\hat{y}_j)k(\hat{y}_j,z,t)v_N(\hat{y}_j,z,t)w_j$$

$$\leqslant C_1\|k\|^2_{N,w} + C_2\|v_N\|^2_{N,w}.$$

The Gronwall lemma completes the proof. □

This stability result will be useful in the following theorem for proving the spatial convergence of the rational collocation method in the transformed space (for all z).

Theorem 5. *Assume that $v \in L^\infty(H^\sigma_{\hat{w}})$ and that $\tilde{b} \in H^\sigma_{\hat{w}}$ for $\sigma > 2$. Then for all z*

$$\|v - v_N\|_{L^\infty(L^2_w)} \leqslant CN^{2-\sigma}\|v\|_{L^\infty(H^\sigma_{\hat{w}})}, \tag{33}$$

ensuring that for $z = y$ the rational collocation scheme is convergent in the transformed space.

Proof. Write $\bar{v} := P_N[v] \in L^\infty(V_N)$. By (18) and by the interpolation property of P_N ($P_N[u](\hat{y}_j) = u(\hat{y}_j)$, $0 \leqslant j \leqslant N$, for $u \in C^0(\bar{J})$), $\bar{v}$ satisfies

$$(g'\bar{v}_t, \phi)_{N,w} + \tfrac{1}{2}(\tilde{b}\bar{v}_y + P_{N,y}[\tilde{b}\bar{v}], \phi)_{N,w} + ((\tfrac{1}{2}\tilde{b}_y + g'\bar{b}_0)\bar{v}, \phi)_{N,w}$$

$$= (g'k, \phi)_{N,w} - \tfrac{1}{2}(\tilde{b}(v - \bar{v})_y + (\tilde{b}v)_y - P_{N,y}[\tilde{b}v], \phi)_{N,w}, \quad \forall \phi \in V_N, \ \forall z \in J,$$

$$\bar{v}(0) = P_N[(u^0 \circ g)q_N[s]], \quad \forall z \in J.$$

Thus the error $e := v_N - \bar{v} \in L^\infty(V_N)$ solves the equation

$$(g'e_t, \phi)_{N,w} + \tfrac{1}{2}(\tilde{b}e_y + P_{N,y}[\tilde{b}e], \phi)_{N,w} + ((\tfrac{1}{2}\tilde{b}_y + \bar{g}'b_0)e, \phi)_{N,w} = \frac{1}{2}(G, \phi)_{N,w},$$

$$\forall \phi \in V_N, \ \forall z \in J,$$

$$e(0) = 0, \quad \forall z \in J,$$

where $G := \tilde{b}(v - \bar{v})_y + ((\mathscr{I} - P_N)(\tilde{b}v))_y$, $\mathscr{I}$ denoting the identity operator. We want to establish a bound for $\|G\|_{N,w}$. To this end, note that since $(1+y)^{e^-}(1-y)^{e^+} \leqslant 2$ on J, one has (for all z)

$$\|G\|_{N,w} \leqslant \sqrt{2}\|G\|_{N,\hat{w}} \leqslant \sqrt{2}(\|\tilde{b}(v - \bar{v})_y\|_{N,\hat{w}} + \|((\mathscr{I} - P_N)(\tilde{b}v))_y\|_{N,\hat{w}}).$$

As demonstrated in [9], we can bound the terms of the right-hand side by

$$\|\tilde{b}(v - \bar{v})_y\|_{N,\hat{w}} \leqslant CN^{2-\sigma}\|v\|_{\sigma,\hat{w}}$$

and

$$\|((\mathscr{I} - P_N)(\tilde{b}v))_y\|_{N,\hat{w}} \leqslant CN^{2-\sigma}\|\tilde{b}\|_{\sigma,\hat{w}}\|v\|_{\sigma,\hat{w}}.$$

Thus $\|G\|_{N,w} \leqslant CN^{2-\sigma}\|v\|_{\sigma,\hat{w}}$. Applying the stability result (28) to e, we get (for all z)

$$\|e\|_{L^\infty(L^2_w)} \leqslant CN^{2-\sigma}\|v\|_{L^2(H^\sigma_{\hat{w}})}.$$

We conclude with the triangle inequality that (for every z)

$$\|v - v_N\|_{0,w} \leqslant \|v - \bar{v}\|_{0,w} + \|e\|_{0,w} \leqslant \sqrt{2}\|v - \bar{v}\|_{0,\hat{w}} + \|e\|_{0,w}$$

and the result follows from the approximation property

$$\|u - P_N[u]\|_{\mu,\hat{w}} \leqslant CN^{2\mu-\sigma}\|u\|_{\sigma,\hat{w}}, \quad 0 \leqslant \mu \leqslant \sigma, \ \forall u \in H^\sigma_{\hat{w}}(J), \ \sigma > \tfrac{1}{2}$$

(see [8]). □

As in the polynomial case, the convergence depends only on the smoothness of the solution, and is faster than any negative power of N if the solution is C^∞; thus in this case the linear rational collocation method is spectrally accurate.

Note that in this section we have not taken into account the discretisation in time. The latter leads to time step restrictions (see [11] for more details on this concept) which will be discussed in the next section.

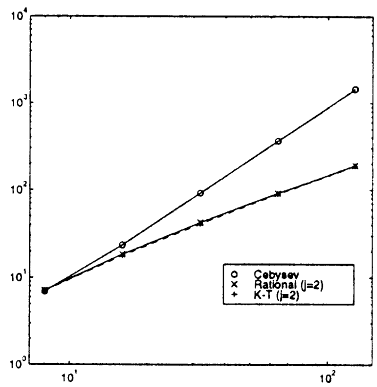

Fig. 1. Spectral radius (logarithmic plot).

4. Numerical examples

In this section we discuss some numerical computations. We first display the spectral radius and the condition number of the first derivative matrix for the model problem (1). Then we give the results of solving three hyperbolic problems and one parabolic problem using the routine ODE45 (in MATLAB 4) in time with a tolerance of 10^{-6}. All computations were performed on a DEC AlphaServer 2100A 5/300.

(6) is chosen as the mapping function as in [13,3].

In Fig. 1 we present the spectral radius ρ of the differentiation matrices $\boldsymbol{D}^{(1)}$ for the model problem (1) when $\boldsymbol{D}^{(1)}$ is calculated using the technique proposed in [2], for several values of N ($N = 8, 16, 32, 64, 128$) and α ($\alpha = \cos(j\pi/N)$, $j = N/2, 2$ – for $j = N/2$, one has $\alpha = 0$ and ρ is the spectral radius of the polynomial Čebyšev collocation differentiation operator). We also give the spectral radius of the Kosloff and Tal-Ezer differentiation matrix $\tilde{\boldsymbol{D}}^{(1)} := \boldsymbol{A}\boldsymbol{D}^{(1)}$ [13], where $\boldsymbol{A}$ is the diagonal matrix whose nonzero entries are given by $A_{ii} := 1/g'(y_i, \alpha)$ and where $\boldsymbol{D}^{(1)}$ is the Čebyšev (spectral) collocation differentiation matrix.

From Fig. 1 we see, e.g., that for $N = 128$ the spectral radius is almost eight times smaller for the rational differentiation operator than for the Čebyšev differentiation operator. It is of the same size as the Kosloff and Tal-Ezer differentiation operator.

A useful measure of the normality of the differentiation matrix is the size of the condition number $\kappa(K) = \|K\| \cdot \|K^{-1}\|$ of the matrix K whose columns are the normalized eigenvectors of $\boldsymbol{D}^{(1)}$, resp. $\tilde{\boldsymbol{D}}^{(1)}$.

The condition numbers are plotted in Fig. 2 for several values of N ($N = 8, 16, 32, 64, 128$) and α ($\alpha = \cos(j\pi/N)$, $j = N/2, 1, 2, 3$). We see that the Čebyšev differentiation operator (for the model problem (1)) is strongly non-normal, as observed in [18] (see the circles in Fig. 2). In contrast, the condition number of the rational differentiation operator does not increase too rapidly with N (see Fig. 2). In dashed, we display the condition numbers of the Kosloff and Tal-Ezer differentiation operator (for the same values of α). They are about half those of the rational differentiation operator.

Next we solve the model problem (1) with $f(x) = \cos^2(\pi x/2)$ as the initial condition. We use the polynomial and rational (spectral) collocation methods in the space variable x, and the embedded

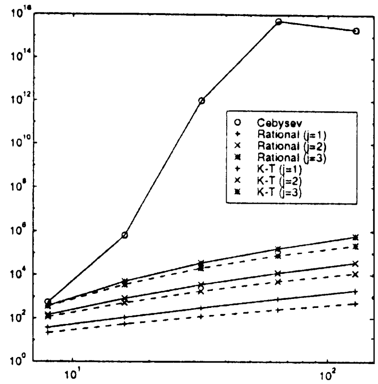

Fig. 2. Condition numbers of some matrices K (logarithmic plot).

Table 1
Absolute error of the numerical solution at time $t=1$ for Problem (1)

	Čebyšev		Rational ($j=2$)		Rational ($\alpha=0.9$)	
N	nsteps	E_{abs}	nsteps	E_{abs}	nsteps	E_{abs}
8	22	9.521×10^{-3}	22	1.232×10^{-2}	22	2.762×10^{-2}
16	33	1.240×10^{-3}	32	1.785×10^{-3}	32	1.245×10^{-3}
32	48	3.280×10^{-4}	49	3.511×10^{-4}	48	1.951×10^{-4}
64	124	7.831×10^{-5}	74	9.132×10^{-5}	74	5.154×10^{-5}
128	477	1.915×10^{-5}	115	4.568×10^{-5}	263	1.183×10^{-5}

Table 2
Absolute error of the numerical solution at time $t=1$ for Problem (1) with f as in (34)

	Čebyšev		Rational ($j=2$)		Rational ($\alpha=0.9$)	
N	nsteps	E_{abs}	nsteps	E_{abs}	nsteps	E_{abs}
8	34	2.280×10^{-1}	39	9.755×10^{-1}	37	4.026×10^{0}
16	84	2.284×10^{0}	80	4.656×10^{0}	79	5.991×10^{0}
32	129	1.741×10^{0}	145	9.883×10^{-1}	137	1.442×10^{1}
64	183	1.351×10^{-1}	188	1.116×10^{-3}	187	1.563×10^{-2}
128	473	4.153×10^{-5}	189	1.376×10^{-4}	259	3.070×10^{-5}

Runge–Kutta method [10] of the MATLAB 4 routine ODE45 in time with a tolerance of 10^{-6}. The method underlying the latter is a Runge–Kutta–Fehlberg method of order 4(5).

In Tables 1–4 we list the number of steps (nsteps) needed to compute the solution at time $t=1$ and the maximum error (E_{abs}) between the exact and the approximated solutions.

In Table 1, we see that for $N=128$ the number of steps needed to compute the solution at time $t=1$ with the rational collocation method (and $j=2$) is four times smaller than with the "classical"

Table 3
Absolute error of the numerical solution at time $t = 1$ for Problem (35)

	Čebyšev		Rational ($j = 2$)		Rational ($\alpha = 0.9$)	
N	nsteps	E_{abs}	nsteps	E_{abs}	nsteps	E_{abs}
8	18	1.795×10^{-5}	19	2.817×10^{-5}	18	9.314×10^{-4}
16	19	5.862×10^{-8}	19	1.859×10^{-5}	19	7.074×10^{-6}
32	19	5.861×10^{-8}	21	5.151×10^{-6}	19	5.799×10^{-8}
64	29	1.328×10^{-8}	29	1.396×10^{-6}	19	1.004×10^{-7}
128	102	3.046×10^{-7}	52	5.592×10^{-7}	52	1.694×10^{-8}

Table 4
Absolute error of the numerical solution at time $t = 1$ for Problem (36)

	Čebyšev		Rational ($j = 2$)		Rational ($\alpha = 0.9$)	
N	nsteps	E_{abs}	nsteps	E_{abs}	nsteps	E_{abs}
8	59	2.234×10^{-8}	45	3.196×10^{-6}	36	5.908×10^{-5}
16	854	6.656×10^{-9}	267	5.734×10^{-7}	309	2.068×10^{-7}
32	13546	7.250×10^{-8}	1337	9.886×10^{-8}	4206	7.761×10^{-9}

collocation method. Even for fixed α, we have better results, also for the maximum error.

In Table 2, we list the results for the model problem with the same function f as in [13],

$$f(x) = [x\exp((-x-1)^2)\cos(6\pi x) - 1]^4. \tag{34}$$

The solution oscillates a great deal and for small N the approximation is poor with all methods. We see again that, for $N = 128$, the number of steps needed to approximate the solution is halved for $\alpha = 0.9$ and reduced by a factor of almost 4 for $j = 2$. The solution computed with $\alpha = 0.9$ is even better than that given by the polynomial collocation method.

Next, we solve the problem

$$u_t + xu_x = 0, \quad x \in I, \ \ t \in]0, T],$$

$$u(x, 0) = f(x), \quad x \in I. \tag{35}$$

The function f is given (as in the first problem) by $f(x) := \cos^2(\pi x/2)$, the exact solution is $u(x, t) = f(x\exp(-t))$.

The results given in Table 3, are again better with the rational collocation method. The number of steps required to compute the solution is halved for $\alpha = 0.9$ and $j = 2$.

In order to shed more light on the stability improvement, we plot in Figs. 3 and 4 the boundary of the ε-pseudospectrum [17]

$$\Lambda_\varepsilon(A) := \{z \in \mathbb{C}: \|(zI - A)^{-1}\| \geqslant \varepsilon^{-1}\}$$

of the differential operator $-x\mathrm{d}/\mathrm{d}x$ for $N = 32$ and different ε ($10^{-3}, 10^{-5}$ and 10^{-7}). We see that in the case of polynomial collocation the eigenvalues are very sensitive to roundoff, as was first

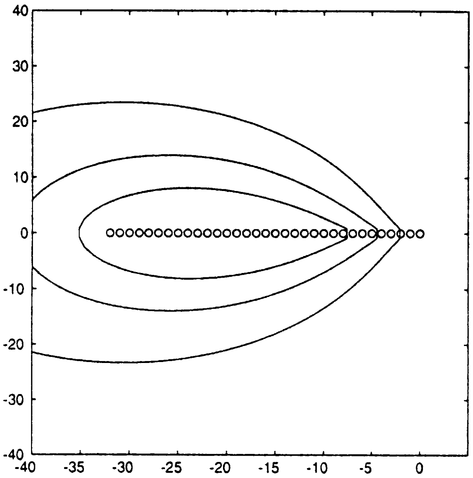

Fig. 3. Boundaries of the pseudospectra $\Lambda_\varepsilon(A)$ for $\varepsilon = 10^{-3}$, 10^{-5} and 10^{-7} of the polynomial collocation matrix A of the operator $-x\mathrm{d}/\mathrm{d}x$, together with the eigenvalues of A (circles).

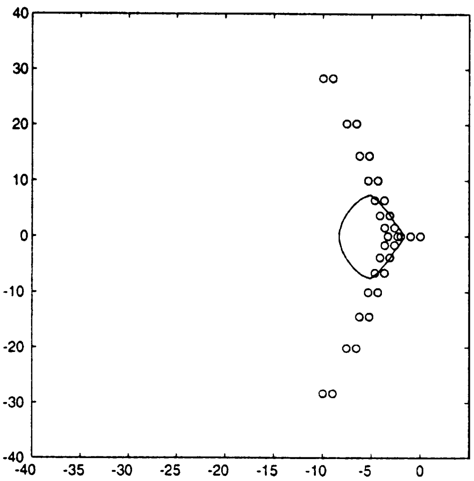

Fig. 4. Same as Fig. 3, but for the rational collocation operator.

noticed in [18]. For rational collocation (with $j = 2$) the eigenvalues are much more stable. The ε-pseudospectrum deviates from the spectrum only for $\varepsilon \leqslant 10^{-3}$.

Finally we present results for the parabolic problem

$$\begin{aligned} &u_t = u_{xx}, \quad x \in I, \ \ t \in \,]0,T], \\ &u(x,0) = \sin((\pi/2)(x+1)), \quad x \in I, \\ &u(-1,t) = u(1,t) = 0, \quad t \in [0,T], \end{aligned} \tag{36}$$

whose exact solution is given by $u(x,t) = \sin((\pi/2)(x+1)\exp((-\pi^2/4)t))$.

Table 4 shows that the rational collocation method is once again more efficient than the polynomial method. For $\alpha = 0.9$ and $N = 32$, the number of steps needed to approximate the solution is divided by about 3. For $j = 2$ and $N = 32$ the number of steps is 12 times smaller.

The results displayed in Tables 1–4 behave the same way if the tolerance is reduced or increased. The linear rational collocation method still remains better than the polynomial Čebyšev method.

We have also performed some tests with implicit methods. The improvements documented in Tables 1–4 can still be observed, though somewhat less pronounced.

5. Conclusion

We have introduced the linear rational collocation method and we have proved stability and convergence results for its application to hyperbolic problems. We have given numerical examples showing that the time step restriction is consistently much weaker with the rational than with the polynomial collocation method.

Comparisons performed with Kosloff and Tal-Ezer's modified Čebyšev method [13] consistently give extremely close results, with equal exponents of E_{abs} for both methods. The linear rational method seems nevertheless to have some advantages like, e.g., the fact that the underlying interpolant is a projection as well as the simplicity of Schneider and Werner's formula for its derivatives.

This method has already been applied to the two dimensional wave equation, see [1] for details. We can also apply the rational method to more complicated (nonlinear) problems and we can again expect weaker stability restrictions than with its polynomial counterpart.

In a future paper, we plan to give some stability and convergence results for general parabolic problems.

Acknowledgements

The authors wish to thank the unknown referees for their comments which have improved the present work.

References

[1] R. Baltensperger, Résolution numérique de l'équation des ondes en dimension deux par la méthode de collocation rationnelle linéaire, Ph.D. Thesis, Institute of Mathematics, University of Fribourg, Switzerland, 1999.
[2] R. Baltensperger, J.-P. Berrut, The errors in calculating the pseudospectral differentiation matrices for Čebyšev–Gauss–Lobatto points, Comput. Math. Appl. 37 (1999) 41–48. Errata: Comput. Math. Appl. 38 (1999) 119.
[3] R. Baltensperger, J.-P. Berrut, B. Noël, Exponential convergence of a linear rational interpolant between transformed Čebyšev points, Math. Comp. 68 (1999) 1109–1120.
[4] J.-P. Berrut, Rational functions for guaranteed and experimentally well-conditioned global interpolation, Comput. Math. Appl. 15 (1988) 1–16.
[5] J.-P. Berrut, H.D. Mittelmann, Lebesgue constant minimizing linear rational interpolation of continuous functions over the interval, Comput. Math. Appl. 33 (1997) 77–86.
[6] J.P. Boyd, Chebyshev and Fourier Spectral Methods, Springer, Berlin, 1989.
[7] C. Canuto, M.Y. Hussaini, A. Quarteroni, T.A. Zang, Spectral Methods in Fluid Dynamics, Springer, New York, 1988.

[8] C. Canuto, A. Quarteroni, Approximation results for orthogonal polynomials in Sobolev spaces, Math. Comp. 38 (1982) 67–86.
[9] C. Canuto, A. Quarteroni, Error estimates for spectral and pseudospectral approximations of hyperbolic equations, SIAM J. Numer. Anal. 19 (1982) 629–642.
[10] E. Hairer, S.P. Nørsett, G. Wanner, Solving Ordinary Differential Equations I. Nonstiff Problems, Springer, Berlin, 1987.
[11] E. Hairer, G. Wanner, Solving Ordinary Differential Equations II, Springer, Berlin, 1991.
[12] P. Henrici, Essentials of Numerical Analysis, Wiley, New York, 1982.
[13] D. Kosloff, H. Tal-Ezer, A modified Chebyshev pseudospectral method with an $\mathcal{O}(N^{-1})$ time step restriction, J. Comput. Phys. 104 (1993) 457–469.
[14] H.E. Salzer, Lagrangian interpolation at the Chebyshev points $x_{n,v} = \cos(v\pi/n), v = 0(1)n$; some unnoted advantages, Comput. J. 15 (1972) 156–159.
[15] C. Schneider, W. Werner, Some new aspects of rational interpolation, Math. Comp. 47 (1986) 285–299.
[16] T. Tang, M. Trummer, Boundary layer resolving pseudospectral methods for singular perturbation problems, SIAM J. Sci. Comp. 17 (1996) 430–438.
[17] L.N. Trefethen, Pseudospectra of matrices, in: D.F. Griffiths, G.A. Watson (Eds.), Numerical Analysis 1991, Longman Scientific and Technical, New York, 1992, pp. 234–266.
[18] L.N. Trefethen, M.R. Trummer, An instability phenomenon in spectral methods, SIAM J. Numer. Anal. 24 (1987) 1008–1023.