On the Approximation of Pseudospectra of Nonnormal Operators by Discretization, Part I: the First Derivative Operator

A. Harrabi *

CERFACS Technical Report TR/PA/98/37

Abstract

This paper is devoted to the study of the influence of discretization schemes on the pseudospectrum of the first derivative operator. Spectral and finite difference methods are considered. Ample illustrations are provided which show the high sensitivity of the pseudospectrum to the chosen discretization method.

Key words: pseudospectra, first derivative operator, Galerkin approximation, collocation approximation, finite-difference method.

Introduction

Let T be a closed linear operator in a Hilbert space $\mathcal{H}$, and $\epsilon > 0$. The ϵ-pseudospectrum $\sigma_\epsilon(T)$ of T is , in a word, the set of all complex numbers λ such that the resolvent norm $\|(\lambda - T)^{-1}\|$ of T in λ is greater than, or equal to, $1/\epsilon$ or $\lambda \in \sigma(T)$, where $\sigma(T)$ denote the spectrum of T [3, 11]. [1]

For a matrix or an operator of finite dimension, the spectrum consists always of a finite non zero number of eigenvalues. The ϵ-pseudospectrum is a bounded closed neighborhood set of the spectrum. Its consists of a small regions one around each eigenvalue for a very small value of ϵ, and grow when

*University of Sciences Sociales of Toulouse (UT1) and CERFACS, 42 Avenue G. Coriolis, 31057 Toulouse Cedex, France, (harrabi@cerfacs.fr).

[1] See Chatelin and Harrabi [4] for a discussion about the possible definitions in a Banach space.

ϵ increases with a speed depends of the distance and the coupling between the eigenvalues and of theirs ascent. The pseudospectrum is however still confined to a bounded region, it does not escape from an ϵ-neighborhood of the field of values, that is the set of the Rayleigh quotients, u^*Tu with $u^*u = 1$, which is convex bounded set.
In contrast for a closed linear operator on an infinite dimensional space, the spectrum can be anything beginning from the empty set arriving to all the complex plan, as long as it is a closed set. The pseudospectrum is always a closed set with a non empty interior. Its contains an ϵ-neighborhood of the $\sigma(T)$, when $\sigma(T)$ is non empty. When $\sigma(T) = \emptyset$, then $\sigma_\epsilon(T)$ is always unbounded for all $\epsilon > 0$.
This paper is devoted to the study of the sensitivity of the pseudospectrum under discretization schemes for the first derivative operator which is non-normal operator.
Generally, the spectrum and the pseudospectrum of a closed operator T are difficult to calculate. In applications, if T is a differential or an integral operator, then the operator T is discretized and approximated by some discrete operator T_N of finite rank. And the question arises naturally whether the spectra $\sigma(T_N)$ and the ϵ-pseudospectra $\sigma_\epsilon(T_N)$ of T_N can give useful information about the corresponding sets for T for N not astronomically large. More precisely, we would like to know if and how $\sigma_\epsilon(T_N)$ (resp., $\sigma(T_N)$) is converging to $\sigma_\epsilon(T)$ (resp., $\sigma(T)$) as $N \to \infty$?

Here, we will not concentrate on the convergence of the spectra of T_N to T. The reader interested in this subject is referred to Chatelin's book "Spectral approximation of linear operators" [5]. Moreover, we will not give theoretical results. But, we will study the specific example consisting of the first derivative operator $T : u \mapsto u'$ on $[-1, 1]$ with boundary condition $u(1) = 0$. This operator is the simplest nonnormal operator. Its spectrum is empty. But, its ϵ-pseudospectra are unbounded regions, and consist of left-half planes. The operator T is discretized by spectral methods (Galerkin and collocation approximation), and centered finite-difference methods (with Chebyshev and equidistant points). Let us denote by T_N the resulting matrix.

The spectrum and the pseudospectrum of the discrete matrix T_N are computed and plotted for $N = 30$ and 150. These experiments show that the ϵ-pseudospectrum $\sigma_\epsilon(T_N)$ of the matrix T_N becomes much larger when the order of the discretization N increases, and tends to represent $\sigma_\epsilon(T)$. The speed of convergence depends strongly on the discretization scheme. Moreover, the ϵ-pseudospectrum $\sigma_\epsilon(T_N)$ gives a local information about $\sigma_\epsilon(T)$ in the neighborhood of zero, but, for a large complex number $\lambda \in \sigma_\epsilon(T)$ to be in $\sigma_\epsilon(T_N)$ it would require a very large size N of discretization.

The notion of convergence for sets in $\mathbb{C}$ used here is in the Hausdorff sense: it means that for all $\lambda \in \sigma_\epsilon(T)$, there exists a sequence $(\lambda_N)_{N\in\mathbb{N}}$ such that $\lambda_N \in \sigma_\epsilon(T_N)$ for all N, $\lim_{N\to\infty} \lambda_N = \lambda$, and the limit of any convergent subsequence $(\lambda_N)_{N\in\mathbb{N}_1\subset\mathbb{N}}$, such that $\lambda_N \in \sigma_\epsilon(T_N)$ for all $N \in \mathbb{N}_1$, is in $\sigma_\epsilon(T)$.

This paper consists of five sections. The first section is devoted to give some equivalent definitions of the pseudospectrum for closed linear operator in a Hilbert space. The second section explains how to read the figures. The third section summarizes the properties of the first derivative operator T, and its pseudospectra. In the fourth section, we show the spectra and the pseudospectra of the discrete matrix T_N approximating the operator T by a finite difference method. The fifth and last section reports the spectrum and the pseudospectrum of the discrete matrix approximating the first derivative operator by Galerkin and collocation approximations.

1 Pseudospectrum in a Hilbert space

Let $\lambda \in \mathbb{C}$ and $\epsilon > 0$. We say that λ is an ϵ-pseudoeigenvalue of a closed linear operator T in a Hilbert space $\mathcal{H}$ iff λ is in the spectrum $\sigma(T)$ of T or λ is in the resolvent set $\rho(T)$ of T and the resolvent norm $\|(\lambda - T)^{-1}\| \geq 1/\epsilon$ [9]. The ϵ-pseudospectrum of T is defined as the set of all the ϵ-pseudoeigenvalues of T [3, 11]. Equivalently, $\sigma_\epsilon(T)$ is the closed set of all $\lambda \in \sigma(T + E)$ where E is a bounded operator on $\mathcal{H}$ with $\|E\| \leq \epsilon$ with $\|.\|$ is the operator norm on $\mathcal{H}$.

Theorem 1.1 *Let T be a closed linear operator in a Hilbert space $\mathcal{H}$ with dense domain $D(T)$, and $\epsilon > 0$. Then the ϵ-pseudospectrum $\sigma_\epsilon(T)$ satisfies the following equivalent definitions:*

$$\{\lambda \in \mathbb{C};\ \|(\lambda - T)^{-1}\| \geq \frac{1}{\epsilon}\} \cup \sigma(T), \tag{1.1}$$

$$cl(\{\lambda \in \mathbb{C};\ \exists\, x \in D(T),\ \|x\| = 1 \text{ and } \|(\lambda - T)x\| < \epsilon\}) \cup \sigma(T), \tag{1.2}$$

$$cl(\{\lambda \in \mathbb{C};\ \lambda \in \sigma(T + E) \text{ for some } E \in \mathcal{L}(\mathcal{H}) \text{ with } \|E\| \leq \epsilon\}) \tag{1.3}$$

where $cl(S)$ of a set S denotes the closure set of S.

The definition of the pseudospectra of a closed linear operator T can be extended to Banach spaces with the use of the following hypothesis: the resolvent norm $\|(\lambda - T)^{-1}\|$ of the operator T is not constant in an open set of the resolvent set of T. This hypothesis is satisfied in the case of Hilbert spaces and in Banach spaces of finite dimension [1, 6]. The question remains open for Banach spaces with infinite dimension. See also Chatelin and Harrabi [4].

2 Figures

Figure 1 illustrates the approximations $\sigma_{m\epsilon}(T)$ and $\sigma_{s\epsilon}(T)$ for the pseudospectrum of the first derivative operator. The other figures illustrate the spectrum and the pseudospectrum of some matrices obtained by various discretization schemes of T. Let us denote these discretized matrices by T_N as above.

The figures show the pseudospectra of the discrete matrices T_N by representing the level curves $\|(z - T_N)^{-1}\| = 10^{\epsilon}$ for all $\epsilon = 1, 2, \dots, 7$, that is, the boundary of the ϵ-pseudospectrum $\sigma_\epsilon(T_N)$ for $\epsilon = 10^{-1}, 10^{-2}, \dots, 10^{-7}$. The smallest singular value $\sigma(z - T_N)$ is computed on a 50×50 grid in some chosen rectangle $[x_1,\, x_2] \times [y_1,\, y_2]$. The resulting array is then fed to the contour plot routine of Matlab. The eigenvalues of the matrices T_N are denoted by $+$ signs. [2]

The resolvent norm of first derivative operators is difficult to be calculate, and the ϵ-pseudospectra $\sigma_\epsilon(T)$ can not be known exactly. The best thing which can be done is to give some approximation of the pseudospectra of T. In section 3, we propose two sets, $\sigma_{m\epsilon}(T)$ and $\sigma_{s\epsilon}(T)$ such that $\sigma_{m\epsilon}(T) \subseteq \sigma_\epsilon(T) \subseteq \sigma_{s\epsilon}(T)$. In Figure 1, we plot the two sets $\sigma_{m\epsilon}(T)$ (dashed curves) and $\sigma_{s\epsilon}(T)$ (solid curves). We note that these two sets are very close to each other and therefore give a very good visual approximation to $\sigma_\epsilon(T)$. Moreover, we can say that $\sigma_\epsilon(T)$ is graphically equal to one of these two sets since, generally, the figures showing the pseudospectra of a matrix do not give the exact pseudospectra but are accurate by two millimeters or three. This is the case indeed for $\sigma_{m\epsilon}(T)$ or $\sigma_{s\epsilon}(T)$ since the difference between them do not exceed three millimeters.

3 The first derivative operator

Let T be the first derivative operator in the Hilbert space $L^2[-1, 1]$.

$$Tu = u', \quad u(1) = 0. \tag{3.1}$$

[2]Figure 6 appeared in [11, 12].

The domain of definition of the operator T is $D(T) = \{u \in L^2[-1,\, 1];\ u' \in L^2[-1,\, 1] \text{ and } u(1) = 0\}$ which is a dense subspace of $L^2[-1,\, 1]$.

We note that the operator T is *nonnormal.* In fact, the adjoint operator of T is the operator T^* acting in $L^2[-1,\, 1]$ and defined by $T^*u = -u'$ with boundary condition $u(-1) = 0$. Then $D(T) \neq D(T^*)$, and $D(TT^*) \neq D(T^*T)$. Therefore T is not normal (in spite of the fact that $T^* = -T$ in $D(T) \cap D(T^*)$) [5, 8].

The spectrum $\sigma(T)$ of T is empty. Moreover, for all $\lambda \in \mathbb{C}$, and all functions $v \in L^2[-1,\, 1]$

$$(\lambda - T)^{-1}v(y) \quad = \quad \int_y^1 e^{-\lambda(x-y)}v(x)\,dx \quad = \quad \int_{-1}^1 h_\lambda(y,x)v(x)\,dx \qquad (3.2)$$

where $(\lambda - T)^{-1}$ is the resolvent of the operator T on λ, and

$$h_\lambda(y,x) \quad = \quad \begin{cases} e^{-\lambda(x-y)} & \text{if } \ -1 \leq y \leq x \leq 1 \\ 0 & \text{if } \ -1 \leq x \leq y \leq 1, \end{cases} \qquad (3.3)$$

The following theorem summarizes the features of the resolvent norm $\|(\lambda - T)^{-1}\|$ of T in λ, which are necessary to study the pseudospectrum of T.

Theorem 3.1 *Let $\lambda = a + ib$ a complex number, where $a,\, b \in \mathbb{R}$ and $i = \sqrt{-1}$. Then*

1. *the resolvent norm $\|(\lambda - T)^{-1}\|$ of T in λ is independent of the imaginary part of λ: $\|(\lambda - T)^{-1}\| = \|(a - T)^{-1}\|$ for all $b \in \mathbb{R}$.*

2. *The resolvent norm $\|(\lambda - T)^{-1}\|$ of T in λ is a decreasing function of the real part $a = Re(\lambda)$ of λ. Equivalently, if $a',\, a \in \mathbb{R}$ such that $a' \leq a$ then $\|(a' - T)^{-1}\| \geq \|(a - T)^{-1}\|$.*

3. *$\lim_{a\to-\infty} \|(a - T)^{-1}\| = +\infty$, and $\lim_{a\to+\infty} \|(a - T)^{-1}\| = 0$.*

Proof

1. Let us first prove that the resolvent norm is constant over all straight line parallel to the imaginary axis. For this end, let $\lambda = a + ib \in \mathbb{C}$ where $a,\, b \in \mathbb{R}$ and $i = \sqrt{-1}$. We will show that $\|(\lambda - T)^{-1}\| = \|(a - T)^{-1}\|$ for all $b \in \mathbb{R}$. Let $v_n \in L^2[-1,\, 1]$ and $\|v_n\| = 1$ such that $\|(\lambda - T)^{-1}v_n\| \to \|(\lambda - T)^{-1}\|$ when $n \to \infty$ (the existence of such a sequence of function $(v_n)_{n\in\mathbb{N}}$ is just a feature of the definition of the resolvent norm as the maximum of $\|(\lambda - T)^{-1}v\|$ for all $v \in D(T)$ with

$\|v\| = 1$). Let $\tilde{v}_n \in L^2[-1,\, 1]$ defined by $\tilde{v}_n(x) = e^{-ibx}v_n(x)$. Then, $\|\tilde{v}_n\| = 1$, and

$$\begin{aligned}
\|(\lambda - T)^{-1}v_n\|^2 &= \int_{-1}^{1} |e^{\lambda y} \int_y^1 e^{-ax}e^{-ibx}v_n(x)\, dx|^2\, dy, \\
&= \int_{-1}^{1} e^{2ay}| \int_y^1 e^{-ax}\tilde{v}_n(x)\, dx|^2\, dy, \\
&= \|(a-T)^{-1}\tilde{v}_n\|^2, \\
&\leq \|(a-T)^{-1}\|.
\end{aligned}$$

Then $\lim_{n\to\infty} \|(\lambda - T)^{-1}v_n\| \leq \|(a-T)^{-1}\|$. Therefore, $\|(\lambda - T)^{-1}\|$, which is equal to the limit of $\|(\lambda - T)^{-1}v_n\|$, is smaller than $\|(a - T)^{-1}\|$. In the same way, we can prove that $\|(a-T)^{-1}\| \leq \|(\lambda - T)^{-1}\|$. Therefore $\|(\lambda - T)^{-1}\| = \|(a-T)^{-1}\|$ for all $\lambda = a + ib$ where $a,\, b \in \mathbb{R}$. This proves that the norm of the resolvent of T at λ is independent of the imaginary part of λ.

2. Let us now prove that the resolvent norm $\|(a-T)^{-1}\|$ of T in $a \in \mathbb{R}$ is decreasing function of the real number a. We notice that the function $h_a(x,y)$ defined by (3.3) is positive for all real a. For all $a',\, a \in \mathbb{R}$ such that $a' \leq a$, we have $h_{a'}(y,x) \geq h_a(y,x)$. Then, for all positive function v in $L^2[-1,\, 1]$, we have $h_{a'}(y,x)v(x) \geq h_a(y,x)v(x)$ for all $x, y \in [-1,\, 1]$, and $\|(a' - T)^{-1}v\| \geq \|(a-T)^{-1}v\|$. The fact that $\|(a-T)^{-1}\| = \sup_{\|v\|=1; v\geq 0} \|(a-T)^{-1}v\|$ since $h_a(y,x) \geq 0$ proves the assumption that $\|(a' - T)^{-1}\| \geq \|(a-T)^{-1}\|$.

3. Let us now prove the third and last point of the theorem.

   (a) We recall that $\|(a-T)^{-1}\| \leq \|h_a\|$, then to show that $\lim_{a\to+\infty} \|(a-T)^{-1}\| = 0$, it is enough to show that $\lim_{a\to+\infty} \|h_a\| = 0$. For all real data a, we have

   $$\|h_a\| = \frac{(e^{-4a} + 4a - 1)^{\frac{1}{2}}}{2|a|}. \tag{3.4}$$

   It easy to see that $\lim_{a\to+\infty} \|h_a\| = 0$, since $\|h_a\| \simeq a^{-1/2}$ for large positive real data a.

   (b) We note that $\|(a-T)^{-1}\| \geq \|(a-T)^{-1}v\|/\|v\|$ for all $v \in L^2[-1,\, 1]$. Let $\mathbb{1}(x) = 1$ for all $x \in [-1,\, 1]$. If $\lim_{a\to-\infty} \|(a-T)^{-1}\mathbb{1}\|/\|\mathbb{1}\| = +\infty$ then $\lim_{a\to-\infty} \|(a-T)^{-1}\| = +\infty$. A simple calculation gives

   $$\frac{\|(a-T)^{-1}\mathbb{1}\|}{\|\mathbb{1}\|} = \left(\frac{-e^{-4a} + 4e^{-2a} + 4a - 3}{4a^3}\right)^{1/2}.$$

Then, we have $\|\|(a-T)^{-1}\mathbb{1}\|/\|\mathbb{1}\| \simeq e^{-2a}/\sqrt{-4a^3}$, for all large real negative number a, which goes to plus infinity when $a \to -\infty$. This therefore proves the assumption.

$\square$

The following theorem gives the expression of the ϵ-pseudospectra $\sigma_\epsilon(T)$ of the operator T for all $\epsilon > 0$.

Theorem 3.2 *The ϵ-pseudospectrum of T is a left half plan, for all $\epsilon > 0$. Precisely, for all $\epsilon > 0$, we have*

$$\sigma_\epsilon(T) = \{\lambda \in \mathbb{C} : Re\lambda \leq a_\epsilon\},$$

where a_ϵ is a real number such that $\|(a_\epsilon - T)^{-1}\| = 1/\epsilon$, and $Re\lambda$ is the real part of λ.

Proof

The last theorem is a direct consequence of the fact that the resolvent norm $\|(\lambda - T)^{-1}\|$ of T on λ is a decreasing continuous function from $+\infty$ to 0 of the real part of λ, independently of the imaginary part of λ (see Theorem 3.1). The continuity of the resolvent norm is due to the continuity of the resolvent. $\square$

Any explicit formula to the resolvent norm of the operator T is hard to derive, and therefore the ϵ-pseudospectrum $\sigma_\epsilon(T)$ is not known explicitly. Therefore, the only thing that can be done is to approximate $\sigma_\epsilon(T)$. An approximation of $\sigma_\epsilon(T)$ can be find via some approximation of $\|(a-T)^{-1}\|$ over all real data a and exploiting the feature of $\sigma_\epsilon(T)$ (which is a left half-plane). We give two approximate sets $\sigma_{m\epsilon}(T)$ and $\sigma_{s\epsilon}(T)$ such that $\sigma_{m\epsilon}(T) \subseteq \sigma_\epsilon(T) \subseteq \sigma_{s\epsilon}(T)$. These two sets ($\sigma_{m\epsilon}(T)$ and $\sigma_{s\epsilon}(T)$) are given by bounding the resolvent norm $\|(a-T)^{-1}\|$ of T for all real data a. We note that $\|(a-T)^{-1}v\|/\|v\| \leq \|(a-T)^{-1}\| \leq \|h_a\|$ for all non zero function $v \in L^2[-1, 1]$ and all real data a. Let $v_a(x) = e^{-ax}$, then $v_a \in L^2[-1, 1]$, and

$$\frac{\|(a-T)^{-1}v_a\|}{\|v_a\|} = \left(\frac{e^{-4a} + 8a - e^{4a}}{4a^2(1-e^{4a})}\right)^{1/2}.$$

The choice of the function v_a as above is governed by the fact that the value $\|(a-T)^{-1}v_a\|/\|v_a\|$ is very close to $\|h_a\|$ for all real data a. Let $\epsilon > 0$, and

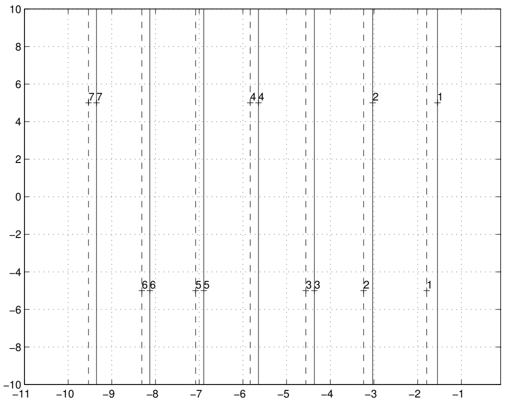

Figure 1: *The ϵ-pseudospectra $\sigma_\epsilon(T)$ of the first derivative operator T with boundary condition $u(1) = 0$ are left half-planes. The dashed curves (resp., the solid lines) are the boundaries of $\sigma_{m\epsilon}(T)$ (resp., $\sigma_{s\epsilon}(T)$) defined by (3.6) (resp., (3.5)) for $\epsilon = 10^{-1}, 10^{-2}, \dots, 10^{-7}$. Note that $\sigma_{m\epsilon}(T) \subseteq \sigma_\epsilon(T) \subseteq \sigma_{s\epsilon}(T)$. The operator T has an empty spectrum.*

let $\sigma_{s\epsilon}(T)$ and $\sigma_{m\epsilon}(T)$ defined by follows

$$\begin{aligned}
\sigma_{s\epsilon}(T) &= \{\lambda = a + ib \in \mathbb{C};\ \|h_a\| \geq \frac{1}{\epsilon}\},\\
&= \{\lambda = a + ib \in \mathbb{C};\ (\frac{e^{-4a} + 4a - 1}{4a^2})^{1/2} \geq \frac{1}{\epsilon}\}. \qquad (3.5)\\
\sigma_{m\epsilon}(T) &= \{\lambda = a + ib \in \mathbb{C};\ \|(a - T)^{-1} v_a\| / \|v_a\| \geq \frac{1}{\epsilon}\},\\
&= \{\lambda = a + ib \in \mathbb{C};\ (\frac{e^{-4a} + 8a - e^{4a}}{4a^2(1 - e^{4a})})^{1/2} \geq \frac{1}{\epsilon}\}. \qquad (3.6)
\end{aligned}$$

The two sets form left half planes like $\sigma_\epsilon(T)$ since the two functions $\|h_a\|$ and $\|(a - T)^{-1} v_a\| / \|v_a\|$ are decreasing continuous functions of a, and satisfy $\lim_{a\to-\infty} \|h_a\| = +\infty$ (resp., $\lim_{a\to-\infty} \|(a - T)^{-1} v_a\| / \|v_a\| = +\infty$) and $\lim_{a\to+\infty} \|h_a\| = 0$ (resp., $\lim_{a\to+\infty} \|(a-T)^{-1} v_a\| / \|v_a\| = 0$). Moreover, these two sets are close together. The difference between the two sets do not exceed two millimeters or three on Figure 1 and therefore they give an excellent approximation to $\sigma_\epsilon(T)$. Therefore, we will assume that $\sigma_\epsilon(T)$ is "graphically" equal to $\sigma_{s\epsilon}(T)$.

Now, it is time to explain how Figure 1 has been constructed. Figure 1 shows the two sets $\sigma_{m\epsilon}(T)$ (dashed curves) defined by (3.6) and $\sigma_{s\epsilon}(T)$ (continuous curves) defined by (3.5) for the seven values of ϵ: 10^{-1}, 10^{-2}, ..., 10^{-7}. The values of a_ϵ such that $\|(a_\epsilon - T)^{-1}v_{a_\epsilon}\|/\|v_{a_\epsilon}\| = 1/\epsilon$ (resp., a_ϵ such that $\|h_{a_\epsilon}\| = 1/\epsilon$) are calculated explicitly for all $\epsilon = 10^{-1}, 10^{-2}, \ldots, 10^{-7}$. Therefore, the vertical straight line passing through a_ϵ is plotted for all values of a_ϵ.

4 Finite difference approximations

We recall that the finite difference method consists in choosing N points $(x_i)_{0\le i\le N}$ in the interval of interest (here $|x| \le 1$, such that $x_0 = -1 < x_1 < \cdots < x_{n-1} < x_N = 1$) and to approximate the derivative of u by some convergent numerical differentiation formulas. In the case of interest, the first derivative u' will be approximated by

$$u'(x_i) \;=\; \sum_{-j_1}^{j_2} \alpha_j u_{i+j}, \tag{4.1}$$

where u_{i+j} represents the value of u at x_{i+j}.

Two things are of capital importance in a finite difference method: the choice of the points (hence the mesh size $|x_{i+1} - x_i|$), and the choice of j_1, j_2 and the coefficients α_j for all $-j_1 \le j \le j_2$ giving the degree of accuracy of the approximation.

Here, we will use a centered finite difference scheme to approximate u' with $j_1 = j_2 = 1$ at all points x_i, $1 \le i \le N-1$, and a forward finite difference on $x_0 = -1$. So we obtain:

$$u'(x_0 = -1) \;=\; \frac{u(x_1) - u(x_0)}{x_1 - x_0}, \quad u'(x_i) \;=\; \frac{u(x_{i+1}) - u(x_{i-1})}{x_{i+1} - x_{i-1}}, \ \forall\, i \ge 1. \tag{4.2}$$

The centered finite difference approximation of T yields the matrix of order N with entries $(t_{ij})_{1\le i,j\le N}$ defined as follow

$$t_{12} \;=\; -t_{11} \;=\; \frac{1}{x_1 - x_0}, \qquad t_{ii\pm 1} \;=\; \pm\frac{1}{x_{i+1} - x_{i-1}}, \ \forall\, 2 \le i \le N. \tag{4.3}$$

In the following two subsections, we have chosen two different types of partition: the first one with the uses equidistant points, the second one uses Chebyshev points. The spectrum and the pseudospectrum of the approximated matrices are reported in both cases. We denote by T_N^{fed} (resp., T_N^{fch}) the matrix obtained with the finite difference method of the first derivative operator T on equidistant points (resp., Chebyshev points).

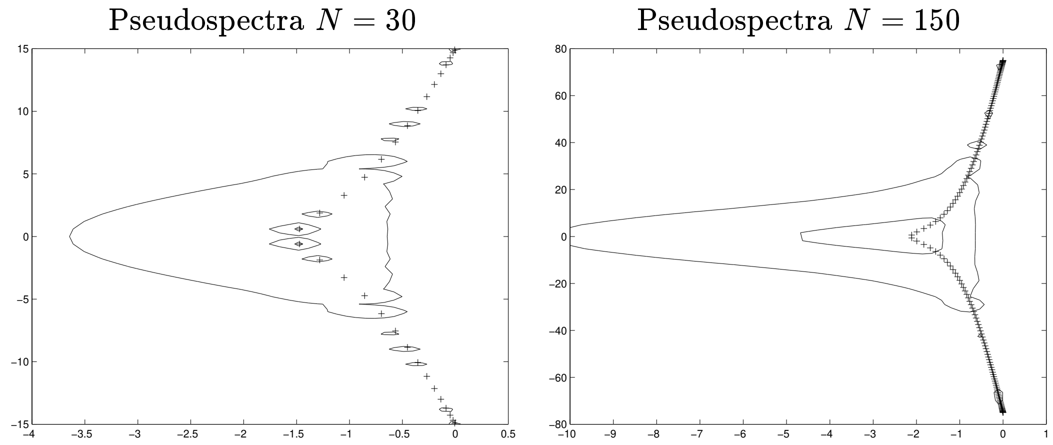

Figure 2: *The centered finite difference approximation matrix* T_N^{fed} *of order* N *with equidistant points.* $\|T_{30}^{fed}\| \simeq 22.7$, *and* $\|T_{150}^{fed}\| \simeq 113.4$.

4.1 Equidistant points

In this subsection, the finite difference mentioned above is used with equidistant points. Let $N > 1$; the mesh size is $h = \frac{2}{N}$, and the equidistant points are $x_j = -1 + jh,\ j = 0, 1, 2, \ldots, N$. So, the first derivative is approximated as follows

$$u'(x_0 = -1) \approx \frac{u_1 - u_0}{h}, \qquad u'(x_j) \approx \frac{u_{j+1} - u_{j-1}}{2h}.$$

The first derivative operator T with boundary condition $u(1) = 0$ is approximated by tridiagonal matrix T_N^{fed} of order N and entries

$$t_{11} = -\frac{1}{h};\ t_{12} = \frac{1}{h};\ t_{NN-1} = \frac{1}{2h};\ t_{ii\pm 1} = \pm\frac{1}{2h} \text{ for all } 2 \le i \le N-1.$$

The matrix T_N^{fed} differs from a tridiagonal Toeplitz matrix by its first two entries t_{11} and t_{12}.

In Figure 2, we report the pseudospectrum of the matrix T_{30}^{fed} of order 30 in the left, and its of T_{150}^{fed} of order 150 in the right part. The spectrum of the matrices T_N^{fed} is in the left half plane (corresponding to negative real parts). The ϵ-pseudospectrum approximates poorly, in this case, $\sigma_\epsilon(T)$ (compare with Figure 1):

- we note that, for example, $\sigma_{10^{-1}}(T_N^{fed})$ does not contains point z with real part less than $-N/10$ ($\sigma_{10^{-1}}(T_{30}^{fed}) \subset \{z \in \mathbb{C} : \mathrm{Re} z \geq 3.5 \simeq 30/10\}$, respectively $\sigma_{10^{-1}}(T_{150}^{fed}) \subset \{z \in \mathbb{C} : \mathrm{Re} z \geq 15 \simeq 150/10\})$. We note that this fact is also observed for large order of discretization.

- the right part of the level curve of $\sigma_\epsilon(T_N^{fed})$ are far from being vertical segments, and do not approximate very well the vertical lines bounding the ϵ-pseudospectra of the operator T.

We note that $\sigma_\epsilon(T_N^{fed})$ *"converges"* to $\sigma_\epsilon(T)$ which is also observed for larger dimensions N of T_N^{fed}: in fact, for example, $\sigma_{10^{-1}}(T_N^{fed})$ becomes larger and larger as N increases (see the difference of scaling between the pseudospectra for $N = 30$, and $N = 150$) and remains in the left half plane like $\sigma_{10^{-1}}(T)$. But, this *"convergence"* is not very good, in the sense, that $\sigma_\epsilon(T_N^{fed})$ does not extend very well in side the left half-plane as explained previously.

4.2 Chebyshev points

The points $x_i = -\cos(i\pi/N)$, for $i = 0, 1, \ldots, N$, are the zeros of the Chebyshev polynomial of order N and are called the Chebyshev points. The centered finite difference method based on these points yields a tridiagonal matrix T_N^{fch} of order N with entries $(t_{i\,j})_{1\leq i,j\leq N}$ indicated by the formula (4.3).

In the left plot of Figure 3, the pseudospectra of the matrix T_{30}^{fch} of order 30 is given, those of T_{150}^{fch} of order 150 is reported in in the right plot.

The ϵ-pseudospectrum of T_N^{fch} approximates more accurately $\sigma_\epsilon(T)$ than does $\sigma_\epsilon(T_N^{fed})$ (compare Figure 2 and Figure 3). Moreover, we note that $\sigma_\epsilon(T_N^{fch})$ converges more quickly to $\sigma_\epsilon(T)$ than $\sigma_\epsilon(T_N^{fed})$. In the case of Chebyshev points, $\sigma_{10^{-1}}(T_N^{fch})$, for example, contains pints with module of the order of N (see Figure 3). Locally in the neighborhood of zero, we have, also, a clear improvement. But, unfortunately the right part of the level curve of $\sigma_\epsilon(T_N^{fed})$ approximates with difficulty a vertical segments, and does not approximate very well the vertical lines bounding the ϵ-pseudospectra of the operator T. This disadvantage reinforced by the fact that $\sigma_\epsilon(T_N^{fch})$ does not extend very well in side the left half-plane remains an important drawback of the finite difference discretizations.

5 Spectral methods

In this section, we consider a spectral method to approximate T. We will use first the Galerkin approximation with the basis $\phi_k = t_k - t_0$ where t_k is

Pseudospectra $N = 30$ Pseudospectra $N = 150$

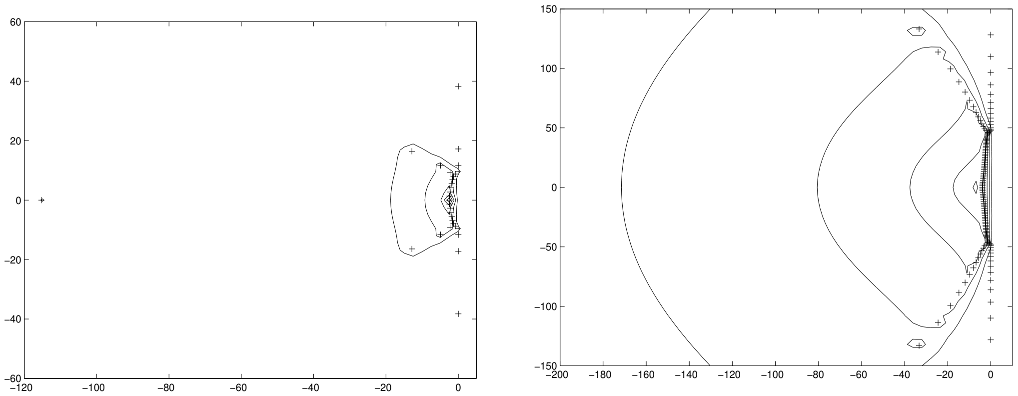

Figure 3: *The centered finite difference approximation matrix* T_N^{fch} *of order* N *with Chebyshev points.* $\|T_{30}^{fch}\| \simeq 260$, *and* $\|T_{150}^{fch}\| \simeq 6.5 \times 10^3$.

the Chebyshev polynomial of degree k. Second, we will use the collocation approximation, where the collocation points are the Chebyshev points.

In both cases, the operator T is approximated by a matrix of order N. These matrices are denoted by T_N^{GC} in the case of Galerkin approximation, and by T_N^{CC} in the case of collocation approximation. The spectra and pseudospectra of T_N^{CC} and T_N^{GC} will be presented and compared to those of the first derivative operator T as we have done in the case of finite difference method in the last section.

5.1 Galerkin approximation

In this subsection, we use the Galerkin approximation of the operator T. The chosen basis is $\phi_i(x) = t_i(x) - t_0(x)$ where t_i is the Chebyshev polynomial of order i for all $i \in \mathbb{N}$: $t_i(x) = \cos(i \arccos x)$. Observe that ϕ_i satisfies the boundary condition $\phi_i(1) = 0$ for all $i \geq 1$ since $t_i(1) = 1$ for all $i \in \mathbb{N}$.

A simple calculation show that the Chebyshev polynomials satisfy the following properties:

$$(t_i,t_j)=\int_{-1}^{1} t_k(x)t_l(x)\,dx \quad = \quad \begin{cases} 0 & \text{if } \; i+j: \; \text{odd} \\ \frac{1}{1-(i+j)^2}+\frac{1}{1-(i-j)^2} & \text{if } \; i+j: \; \text{even} \end{cases} \tag{5.1}$$

$$(t_i,{t'}_j)=\int_{-1}^{1} t_k(x){t'}_l(x)\,dx \quad = \quad \begin{cases} \frac{2j^2}{j^2-i^2} & \text{if } \; i+j: \; \text{odd} \\ 0 & \text{if } \; i+j: \; \text{even} \end{cases} \tag{5.2}$$

It implies that:

$$(\phi_i,\phi_j) \quad = \quad \begin{cases} 2+\frac{1}{1-(i+j)^2}+\frac{1}{1-(i-j)^2} & \text{if } j \text{ odd}, \quad i \text{ odd} \\ 2+\frac{2}{1-i^2} & \text{if } j \text{ odd}, \quad i \text{ even} \\ 2+\frac{2}{1-j^2} & \text{if } j \text{ even}, \quad i \text{ odd} \\ 2+\frac{2}{1-j^2}+\frac{2}{1-i^2}+\frac{1}{1-(i+j)^2}+\frac{1}{1-(i-j)^2} & \text{if } j \text{ even}, \quad i \text{ even} \end{cases}$$

$$(\phi_i,{\phi'}_j) \quad = \quad \begin{cases} -2 & \text{if } j \text{ odd}, \quad i \text{ odd} \\ -2+\frac{2j^2}{j^2-i^2} & \text{if } j \text{ odd}, \quad i \text{ even} \\ \frac{2j^2}{j^2-i^2} & \text{if } j \text{ even}, \quad i \text{ odd} \\ 0 & \text{if } j \text{ even}, \quad i \text{ even} \end{cases}$$

The first derivative operator T is approximated by the Galerkin projection $T_N^{GC} = P_N T P_N$ where P_N is the orthogonal projection on the polynomial space $I\!P_N$ of degree less than, or equal to, N vanishing at $x = 1$. Let $u(x) = \sum_{k=1}^{\infty} u_k\phi_k(x)$, then

$$T_N^{GC}u \;=\; P_N T P_N u \;=\; \sum_{l=1}^{N}(\sum_{k=1}^{N} t_{l\,k}u_k)\phi_l \quad \text{where} \quad \phi'_k = \sum_{l=1}^{\infty} t_{l\,k}\phi_l.$$

The approximate operator T_N^{GC} can be viewed as a matrix of order N in $\mathbb{C}^{N\times N}$ denoted also by T_N^{GC}. Its entries $(t_{ij})_{1\le i,j\le N}$ satisfy $\phi'_j = \sum_{i=1}^{\infty} t_{ij}\phi_i$. Let us proceed to find the explicit formula of the entries of T_N^{GC}.
Some simple calculations yield $T_N^{GC} = A^{-1}B$ where $A \in \mathbb{C}^{N\times N}$ with entries $a_{ij} = (\phi_i,\phi_j)$, and $B \in \mathbb{C}^{N\times N}$ with entries $b_{ij} = (\phi_i,\phi'_j)$ for all $1 \le i,j \le N$.

In Figure 4, we report the pseudospectra of the discretization matrix T_{30}^{GC} of order 30 in the left part, and those of T_{150}^{GC} of order 150 in the right part.

Globally, $\sigma_\epsilon(T_N^{GC})$ converges very well and rapidly to $\sigma_\epsilon(T)$. Moreover, they extend very well in left half-plane. Not that, for example, $\sigma_{10^{-1}}(T_N^{GC})$ contains points z with modules much larger than the order of discretization N,

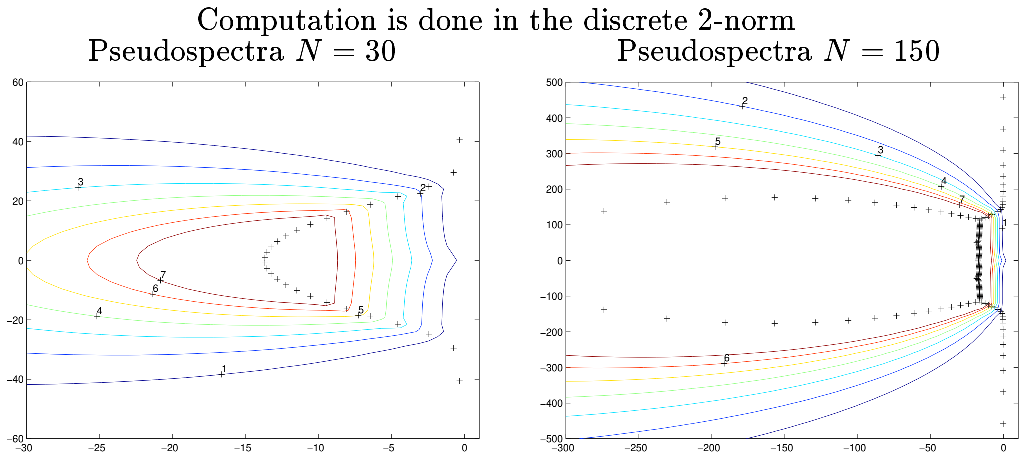

Figure 4: *The Galerkin approximation matrix* T_N^{GC} *of order* N *of the first derivative operator with boundary condition* $u(1) = 0$ *with use of a weighted norms.* $\|T_{30}^{GC}\| \simeq 435$, $\|T_{150}^{GC}\| = 10^4$.

and this feature become more and more when N increase (compare the scaling between $N = 30$ and $N = 150$ in Figure 4). But, in the neighborhood of zero, $\sigma_\epsilon(T_N^{GC})$ has always a tip in the right part near the real axis, which appears for large values of N. This disadvantage is an important drawback of this discretization.

This problem is due to the fact that the basis $(\phi_i)_{1\leq i\leq N}$ is not orthonormal. In fact the discretization matrix T_N^{GC} is calculated with respect to the continuous 2-norm $\|\phi\|_2 = (\int_{-1}^{1} |\phi|^2\, dx)^{1/2}$, in contrast the calculation of its pseudospectrum (more precisely of its resolvent norm) is done in the discrete 2-norm $\|\,.\,\|$ in $\mathbb{C}^N$. To avoid this, it makes more sense to compute the resolvent norm of T_N^{GC} with respect to the continuous 2-norm $\|\,.\,\|_2$, or used an orthonormal basis instead of $(\phi_i)_{1\leq i\leq N}$. Let the weighted matrix E be defined such that $\|\phi\|_2 = \|Eu\|$ where $\phi = \sum_{i=1}^{N} u_i\phi_i$, $u = (u_1, u_2, \ldots, u_N)^T$. In fact, we have

$$\begin{aligned}
\|\phi\|_2^2 &= \sum_{i,j=1}^{N} \overline{u_i} u_j \int_{-1}^{1} \phi_i(x)\phi_j(x)\, dx, \\
&= u' A u, \\
&= \|Eu\|^2,
\end{aligned}$$

where the last equality is done since the matrix A is symmetric and can be decomposed in the form $A = E'E$ where E' denote the adjoint matrix of E.

It is easy to see that

$$\|(z - T_N^{GC})^{-1}\|_2 \;=\; \|E(z - T_N^{GC})^{-1}E^{-1}\| \;=\; \|(z - ET_N^{GC}E^{-1})^{-1}\|. \quad (5.3)$$

The reader is referred to Reddy [10, pp. 44-55] for more details about the last equality.

In Figure 5, we report the pseudospectra of T_{30}^{GC} of order 30 in the left part, and those of T_{150}^{GC} of order 150 in the right part where the computation of the resolvent norm is now done with respect to the continuous 2-norm $\| \,.\, \|_2$ instead of the discrete one used in Figure 4.

We note that globally the ϵ-pseudospectrum of the discretization matrix T_N^{GC} have a similar shape and width with use of the continuous or the discrete 2-norm. But, now, when the continuous 2-norm is used, we note a good local improvement in the neighborhood of zero: note the presence of the vertical segments in the pseudospectra of T_N^{GC} which are close to the vertical line bounding the pseudospectra of T instead of the tip in the right part when the discrete 2-norm is used (compare Figure 5 and 4).

In this subsection, all the experiments are done with the use of the Chebyshev basis. We know in advance that this basis is a very more accurate and convenient basis. We hope to discus the influence of the choice of basis on the accuracy of the approximation of the pseudospectrum in a future work.

Remark: computing the pseudospectrum of the discretization matrix T_N^{GC} by the continuous 2-norm of the space of polynomial $I\!P_N$ is equivalent, in some sense, to use an orthonormal basis of $I\!P_N$ instead of $(\phi_i)_{1\leq i\leq N}$.

To show that, we note first that the weighted matrix E used in the formula (5.3) to compute the pseudospectra of T_N^{GC} is chosen such that $E'E = A$ where $A = ((\phi_i, \phi_j))_{1\leq i,j\leq N}$. Let $(\psi_i)_{1\leq i\leq N}$ be an orthonormal basis of $I\!P_N$ such that $\psi_i(1) = 0$. (Such a basis exists, we can use any Gram-Schmidt orthogonalization of $(\phi_i)_{1\leq 1\leq N}$.) Let us denote by $(x_{ji})_{1\leq j\leq N}$ the coordinate of ϕ_i in the orthogonal basis $(\psi_j)_{1\leq j\leq N}$ for all i such that $1 \leq i \leq N$, and let $X = (x_{ij})_{1\leq i,j\leq N}$. Then, a simple calculation show that $A = ((\phi_i, \phi_j))_{1\leq i,j\leq N} = X'X$, since $(\psi_i)_{1\leq i\leq N}$ is orthogonal. Then, we have $E'E = X'X$, which implies that $EX^{-1} = W$ is an unitary matrix. On the other hand, the theorem of the *change of basis* show that the representation of the operator P_nTP_n in the basis $(\psi_i)_{1\leq i\leq N}$ is the matrix $XT_N^{GC}X^{-1}$. The

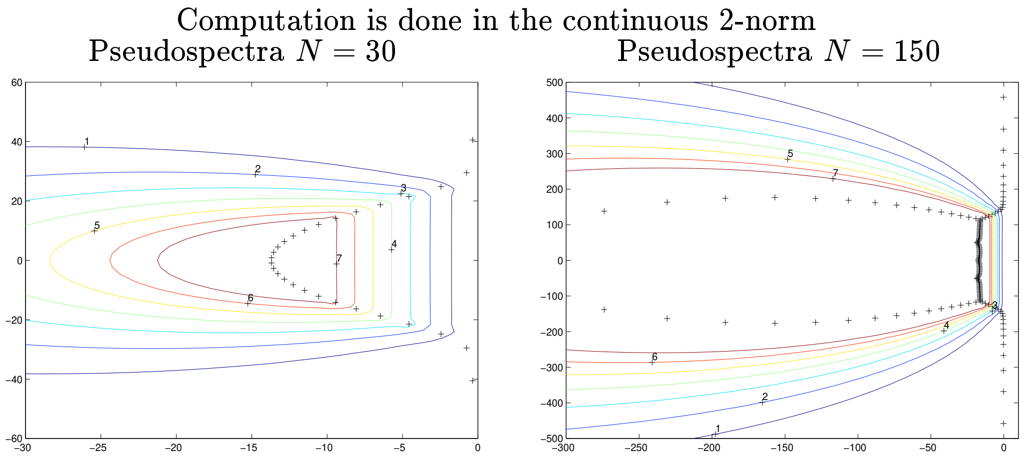

Figure 5: *The Galerkin approximation matrix T_N^{GC} of order N of the first derivative operator with boundary condition $u(1) = 0$ with use of a weighted norms.* $\|T_{30}^{GC}\| \simeq 435$, $\|T_{150}^{GC}\| = 10^4$.

substitution of $XT_N^{GC}X^{-1}$ in the formula (5.3) gives:

$$\begin{aligned} \|(z - T_N^{GC})^{-1}\|_2 &= \|EX^{-1}(z - XT_N^{GC}X^{-1})^{-1}XE^{-1}\|, \\ &= \|(z - XT_N^{GC}X^{-1})^{-1}\| \end{aligned} \tag{5.4}$$

where the last equality is due to the fact that EX^{-1} is unitary. In fact the formula (5.4) shows that the calculation of the resolvent norm in the $\|\,.\,\|_2$ continuous 2-norm is equivalent to calculate the resolvent norm of T_N^{GC} in the discrete 2-norm when now the discretization operator is represented in an orthogonal basis. In fact, in formula (5.3) the discretization operator P_NTP_N is represented in $(\widetilde{\psi}_i)_{1\leq i\leq N}$ where the matrix of passage from the basis $(\phi_i)_{1\leq i\leq N}$ to $(\widetilde{\psi}_i)_{1\leq i\leq N}$ is now E^T. A simple calculation show that the basis $(\widetilde{\psi}_i)_{1\leq i\leq N}$ is, also, orthogonal.

5.2 Collocation approximation

In this subsection, we use the collocation approximation to approximate the operator T defined by (3.1). The points chosen here are the Chebyshev points which are the zeros of the Chebyshev polynomial of degree N equal to $x_j = -\cos(j\pi/N)$, for all $j = 0, 1, \ldots, N$. Also let $l_j(x) = \prod_{i=0,\, i\neq j}^{N} \frac{x - x_i}{x_j - x_i}$ for all $j = 0, 1, \ldots, N$ be the Lagrange polynomials.

The collocation approximation of the operator T defined by (3.1) with the use of Lagrange polynomials as a basis and the Chebyshev points as the collocation points yields the matrix T_N^{CC} of order N with entries t_{ij} equal to $l'_j(x_i)$ for all $1 \leq i, j \leq N$. This matrix is called the Chebyshev spectral differentiation matrix with the boundary condition $u(1) = 0$. Its entries can be expressed by explicit formulas [11, 2], or can be derived by FFT (see [7]).

We report, in Figure 6, the spectrum and the pseudospectrum of the approximate matrix T_{30}^{CC} of order 30 in the upper part. Those of T_{150}^{CC} are in the middle part. Two zooms of the ϵ-pseudospectrum of T_{150}^{CC} are shown in the lower part.

We notice here the presence of vertical segments in the pseudospectrum of the matrices T_N^{CC} which are very close to the vertical lines limiting the pseudospectrum of T which are a left half planes. We note that $\sigma_\epsilon(T_N^{CC})$ approximates more accurately $\sigma_\epsilon(T)$ than any other discretization. Moreover, $\sigma_\epsilon(T_N^{CC})$ give a very good fit locally in the neighborhood of zero. As for the other discretizations, a complex number $\lambda \in \sigma_\epsilon(T)$ with a large modulus, particularly with a large imaginary part requires a very large order N to be in $\sigma_\epsilon(T_N^{CC})$.

6 Conclusion

Experimentally, we have seen that the ϵ-pseudospectrum of the discretized matrix converges to the ϵ-pseudospectrum of the first derivative operator. The speed of the convergence and the sensitivity of the ϵ-pseudospectra of the discretization matrices differ strongly according to the discretization scheme. We note that the classification given in this paper is specific to the first derivative operator and can differ from an operator to another. We note also that the ϵ-pseudospectrum of discrete matrices or operators of finite rank can give at most a local information on the pseudospectra of the operator it approximates. And, generally a large order of discretization is required to have a good local approximation.

References

[1] A. Böttcher. Pseudospectra and singular values of large convolution operators. *J. Integral Equations Applics*, 6:267–301, 1994.

[2] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral Methods in Fluid Dynamics.* Springer-Verlag, New York, 1987.

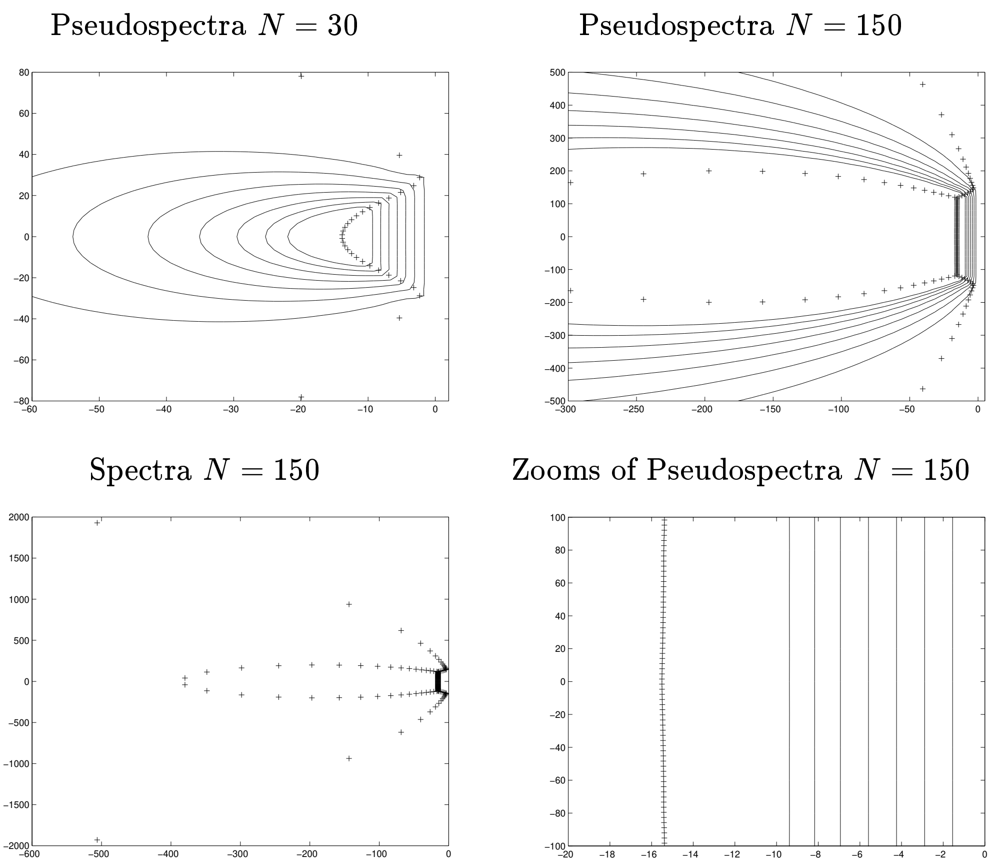

Figure 6: *The collocation approximation matrix* T_N^{CC} *of order* N *of the first derivative operator* T *with boundary condition* $u(1) = 0$. *The collocation points are these of Chebyshev.* $\|T_{30}^{CC}\| \simeq 495$, $\|T_{150}^{CC}\| \simeq 1.2 \times 10^4$.

[3] F. Chaitin-Chatelin and V. Frayssé. *Lectures on Finite Precision Computations.* SIAM, Philadelpha, 1996.

[4] F. Chaitin-Chatelin and A. Harrabi. About definitions of pseudospectra of closed operators in banach spaces. Tech. Rep. TR/PA/98/08, CERFACS, France, 1998.

[5] F. Chatelin. *Spectral Approximation of Linear Operators.* Academic Press, New York, 1983.

[6] A. Harrabi. Pseudospectre d'une suite d'opérateurs bornés. Tech. Rep. TR/PA/97/48, CERFACS, France, 1997. To appear in Mathematical Modelling and Numerical Analysis, Vol. 6, 1998.

[7] N. Higham. The test matrix for matlab (version 3.0). Numerical Analysis Report 276, Manchester Center for Computational Mathematics, 1995.

[8] T. Kato. *Perturbation Theory for Linear Operators.* Springer, New York, 1976.

[9] H. J. Landau. On Szegö's eigenvalue distribution theorem and non-hermitian kernels. *J. Analyse Math*, 28:335–357, 1975.

[10] S. C. Reddy. *Pseudospectra of Operators and Discretization Matrix and an Application to Stability of the Methods of Lines.* PhD thesis, Massachusetts Institute of Technology, Combridge, Massachusetts 02139, August 1991.

[11] L. N. Trefethen. Pseudospectra of matrices. In Numerical Analysis 1991, editor, *Longman, Harlow.* D. F. Griffiths and G. A. Watson, 1992.

[12] L. N. Trefethen. Pseudospectra of linear operators. *SIAM Rev.*, 39(3):383–406, September 1997.