ELSEVIER
Journal of Computational and Applied Mathematics 147 (2002) 453–469
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

On ε-spectra and stability radii

Laurence Grammont, Alain Largillier *

UPRES-EA 3058 Équipe d'Analyse Numérique de Saint Etienne 23, rue du Dr. P. Michelon, F-42023 Saint Etienne, Cedex 2, France

Received 24 July 2001; received in revised form 3 January 2002

Abstract

Techniques of Krylov subspace iterations play an important role in computing ε-spectra of large matrices. To obtain results about the reliability of this kind of approximations, we propose to compare the position of the ε-spectrum of A with those of its diagonal submatrices. We give theoretical results which are valid for any block decomposition in four blocks, $A_{11}, A_{12}, A_{21}, A_{22}$. We then illustrate our results by numerical experiments. The same kind of problem arises when we compute the stability radius of a large matrix. In that context, we propose a new sufficient condition for the stability of a matrix involving quantities readily computable such as stability radius of small submatrices. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: ε-spectrum; Partitioned matrices; Lyapunov's stability

1. Introduction and preliminaries

Before being a popular tool to study matrices and operators, the notion of ε-spectra appears many times in different forms. Gustafson and Rao [4] quote a paper published in 1962 by Parter [11] which is "an interesting predecessor paper to pseudoeigenvalue". According to Trefethen [15], the notion of pseudoeigenvalues, also called ε-eigenvalues, seems to have been invented independently at least five times: Landau [8], Varah [16], Godunov et al. (Kiriljuk, Kostin, Antonov) [1–3], Trefethen [13], Hinrichsen and Kelb [6], Hinrichsen and Pritchard [7]. As Trefethen explains in [14], the motivation for this concept comes from numerical eigenvalue computations: it is well known that in problems involving nonnormal operators or matrices eigenvalue computations are ill-conditioned. In this context, a generalization of the notion of the spectrum may be useful. Since an eigenvalue can be defined as a pole of the resolvent operator

$$z \to R(z) = (A - zI)^{-1},$$

* Corresponding author.
E-mail addresses: grammont@anum.univ-st-etienne.fr (L. Grammont), larg@anum.univ-st-etienne.fr (A. Largillier).

PII: S0377-0427(02)00480-6

we can focus our interest on the set of complex numbers z for which $\|R(z)\|$ is "large". This leads to the notion of ε-spectrum:

Definition 1.1. Given ε, the ε-spectrum of $A \in \mathbb{C}^{n\times n}$ is the set

$$\Lambda_\varepsilon(A) := \left\{ z \in \rho(A) : \|R(z)\|_2 \geqslant \frac{1}{\varepsilon} \right\} \cup \mathrm{sp}(A),$$

where $\rho(A) := \mathbb{C} \setminus \mathrm{sp}(A)$ is the resolvent set of A, $\mathrm{sp}(A)$ is the spectrum of A and $\|\cdot\|_2$ denotes the subordinated matrix Euclidean norm. It can be defined equivalently by

$$\Lambda_\varepsilon(A) := \{ z \in \rho(A) : \sigma_{\min}(A - zI) \leqslant \varepsilon \} \cup \mathrm{sp}(A),$$

where $\sigma_{\min}$ denotes the smallest singular value.

For a large matrix, Krylov subspace methods (or variations of these methods) are used to approximate pseudospectra (see for example [12]).

Krylov subspace methods compute $A_{11} \in \mathbb{C}^{m\times m}$, $m \ll n$ and $V_1 \in \mathbb{C}^{n\times m}$ such that $AV_1 = V_1 A_{11} + R_1$. Using the following partitioned form

$$A := \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $A_{11} \in \mathbb{C}^{m\times m}$ and $A_{22} \in \mathbb{C}^{(n-m)\times(n-m)}$, the former equality can be written as

$$A = [V_1 \quad V_2] \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix} [V_1 \quad V_2]^* + R, \tag{1}$$

where A_{11}, V_1 and $\|R\|_2 = \|A_{21}\|_2$ are given by the method. The other matrices $A_{22} \in \mathbb{C}^{n-m\times n-m}$ and $V_2 \in \mathbb{C}^{n\times n-m}$ will not be computed.

Section 2 provides a new upper bound for the norm of the resolvent. Then, in order to understand the relationship between $\Lambda_\varepsilon(A)$ and $\Lambda_\varepsilon(A_{11})$, we set up inclusions between $\Lambda_\varepsilon(A)$, $\Lambda_{f(\varepsilon)}(A_{11})$ and $\Lambda_{f(\varepsilon)}(A_{22})$ where f is a function of ε, $\|A_{12}\|_2$ and $\|A_{21}\|_2$. This is done in Section 3.

Our purpose is not to propose a better algorithm to approximate an ε-spectrum. We want to prove a general result which can be helpful to analyse theoretically the approximations obtained in practice. The numerical experiments of Section 4 are presented to illustrate this comparison. Our results about ε-spectra are of theoretical nature and cannot compete with existing tools for computing them. Whereas our results about stability radius are useful in applications.

In stability analysis, the *stability radius* is a crucial notion and it measures the distance of a stable matrix from the set of unstable matrices.

Definition 1.2. A matrix $A \in \mathbb{C}^{n\times n}$ is stable if for all $\lambda \in \mathrm{sp}(A)$, $\mathrm{re}(\lambda) < 0$, where re denotes the real part of a complex number.

Definition 1.3. The stability radius of a matrix $A \in \mathbb{C}^{n\times n}$ is defined by

$$\mathrm{rs}(A) := \min_{\mathrm{re}(z)=0} \sigma_{\min}(A - zI).$$

It can be easily seen that

$$\mathrm{rs}(A) = \frac{1}{\max_{\mathrm{re}(z)=0} \|R(z)\|_2}.$$

In order to determine the stability radius of a large matrix, one can also use Krylov subspace methods. They lead to decomposition (1).

The upper bound for the norm of the resolvent will lead to some interesting relationships between the stability radii of A and A_{11}. These results will give new sufficient stability conditions that are proposed in Section 5.

2. A bound for resolvent

Let $A \in \mathbb{C}^{n\times n}$ be as follows:

$$A := \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix},$$

where $A_{11} \in \mathbb{C}^{m\times m}$ and $A_{22} \in \mathbb{C}^{(n-m)\times(n-m)}$.

We shall first obtain an upper bound for $\|R(z)\|_2$ in terms of $\|(A_{11}-zI)^{-1}\|_2$ and $\|(A_{22}-zI)^{-1}\|_2$.

We set $R_1(z) := (A_{11}-zI)^{-1}$, $R_2(z) := (A_{22}-zI)^{-1}$, $r_1(z) := \|R_1(z)\|_2$, and $r_2(z) := \|R_2(z)\|_2$.

Lemma 2.1. *Let* $B(z) := -R_1(z)A_{12}$ *and*

$$\Delta := \begin{pmatrix} I & B(z) \\ O & I \end{pmatrix}.$$

Then

$$\|\Delta\|_2^2 = 1 + \frac{\|B(z)\|_2}{2}\left(\|B(z)\|_2 + \sqrt{\|B(z)\|_2^2 + 4}\right).$$

Proof. $\|\Delta\|_2^2 = \rho(\Delta^*\Delta)$.

It is easy to see that

$$\rho(\Delta^*\Delta) \leqslant \rho \begin{bmatrix} 1 & b(z) \\ b(z) & b(z)^2+1 \end{bmatrix}$$

with $b(z) = \|B\|_2$. We have

$$\rho \begin{bmatrix} 1 & b(z) \\ b(z) & b(z)^2+1 \end{bmatrix} = 1 + \frac{b(z)}{2}\left(\sqrt{b(z)^2+4} + b(z)\right).$$

Hence

$$\rho(\Delta^*\Delta) \leqslant 1 + \frac{b(z)}{2}(\sqrt{b(z)^2+4} + b(z)).$$

Let

$$D = \begin{pmatrix} 1 & b(z) \\ 0 & 1 \end{pmatrix},$$

we have

$$\|D\|_2 = 1 + \frac{b(z)}{2}(\sqrt{b(z)^2+4} + b(z)) = d.$$

Let $v = (\alpha, \beta)^{\mathrm{t}}$ a unit vector in $\mathbb{R}^2$ such that $\|Dv\|_2 = d$.

Let $x \in \mathbb{R}^{n-m}$ and $y \in \mathbb{R}^m$ be units vectors such that $B(z)x = b(z)y$.

Then $(\alpha y, \beta x)^{\mathrm{t}} \in \mathbb{R}^n$ is a unit vector and $\|\Delta(\alpha y, \beta x)^{\mathrm{t}}\|_2 = d$. $\square$

Proposition 2.2. *If* $a_{12} := \|A_{12}\|_2$, $r_M(z) := \max\{r_1(z), r_2(z)\}$ *and* $r_m(z) := \min\{r_1(z), r_2(z)\}$ *then the following upper bound holds*:

$$\|R(z)\|_2 \leqslant r_M(z)\sqrt{1 + \frac{a_{12}r_m(z)}{2}(\sqrt{a_{12}^2 r_m(z)^2 + 4} + a_{12}r_m(z))}. \tag{2}$$

Proof. Without loss of generality we may assume $r_m(z) = r_1(z)$.

Since

$$R(z) = \begin{bmatrix} R_1(z) & -R_1(z)A_{12}R_2(z) \\ 0 & R_2(z) \end{bmatrix} = \begin{bmatrix} I & -R_1(z)A_{12} \\ 0 & I \end{bmatrix} \begin{bmatrix} R_1(z) & 0 \\ 0 & R_2(z) \end{bmatrix},$$

it is enough to show that

$$\left\| \begin{bmatrix} I & -R_1(z)A_{12} \\ 0 & I \end{bmatrix} \right\|_2 = \sqrt{1 + \frac{a_{12}r_m(z)}{2}(\sqrt{a_{12}^2 r_m(z)^2 + 4} + a_{12}r_m(z))}.$$

This follows from the Lemma 2.1. $\square$

3. Consequences for ε-spectra

Our aim is to compare the ε-spectrum of

$$A := \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

with those of A_{11} and A_{22} by establishing a suitably fine inclusion. We first treat the case where $A_{21} = 0$ and then we consider the general case.

Proposition 3.1. *Suppose that*

$$A=\begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix}.$$

Then

$$\Lambda_\varepsilon(A_{11}) \cup \Lambda_\varepsilon(A_{22}) \subseteq \Lambda_\varepsilon(A),$$

$$\Lambda_\varepsilon(A) \subseteq \Lambda_{\eta(\varepsilon)}(A_{11}) \cup \Lambda_{\eta(\varepsilon)}(A_{22}), \tag{3}$$

where $\eta(\varepsilon)=\varepsilon\sqrt{1+(a_{12}/\varepsilon)}$ *and* $a_{12}=\|A_{12}\|_2$.

Proof. By the interlace theorem for Hermitian matrices, we have

$$\lambda_{\min}((A-zI)^*(A-zI)) \leqslant \lambda_{\min}((A_{11}-zI)^*(A_{11}-zI)),$$

where $\lambda_{\min}$ denotes the smallest eigenvalue. Hence,

$$\sigma_{\min}(A-zI) \leqslant \sigma_{\min}(A_{11}-zI)$$

and

$$\Lambda_\varepsilon(A_{11}) \subseteq \Lambda_\varepsilon(A).$$

As $\sigma_{\min}(A-zI)=\sigma_{\min}(A-zI)^*$, we prove similarly that $\Lambda_\varepsilon(A_{22}) \subseteq \Lambda_\varepsilon(A)$.

In order to prove the second inequality, let $z \in \Lambda_\varepsilon(A)$.

Recall that $\|R(z)\|_2^2 \geqslant 1/\varepsilon^2$. Using the bound (2), we have for $z \in \Lambda_\varepsilon(A)$

$$\frac{1}{\varepsilon^2} - r_M(z)^2\left(1+\frac{r_M(z)^2 a_{12}^2}{2}\right) \leqslant \frac{a_{12} r_M(z)^3}{2}\sqrt{a_{12}^2 r_M(z)^2+4}. \tag{4}$$

There are two cases depending upon whether the first member of the inequality is positive or not.

If

$$\frac{1}{\varepsilon^2} - r_M(z)^2\left(1+\frac{r_M(z)^2 a_{12}^2}{2}\right) \geqslant 0,$$

then (4) implies that

$$\left(1-\frac{a_{12}^2}{\varepsilon^2}\right) r_M(z)^4 - \frac{2}{\varepsilon^2} r_M(z)^2 + \frac{1}{\varepsilon^4} \leqslant 0. \tag{5}$$

If $a_{12} \leqslant \varepsilon$, then

$$\frac{1}{\varepsilon}\sqrt{\frac{1}{1+\frac{a_{12}}{\varepsilon}}} \leqslant r_M(z) \leqslant \frac{1}{\varepsilon}\sqrt{\frac{1}{1-\frac{a_{12}}{\varepsilon}}}.$$

And if $a_{12} \geqslant \varepsilon$, then

$$\frac{1}{\varepsilon}\sqrt{\frac{1}{1+\frac{a_{12}}{\varepsilon}}} \leqslant r_M(z).$$

If

$$\frac{1}{\varepsilon^2} - r_M(z)^2\left(1 + \frac{r_M(z)^2 a_{12}^2}{2}\right) \leqslant 0,$$

then it is easy to prove that we also have

$$r_M(z) \geqslant \frac{1}{\varepsilon}\sqrt{\frac{1}{1 + \frac{a_{12}}{\varepsilon}}}. \qquad \square$$

We give below generalization of Proposition 3.1.

Theorem 3.2. *If*

$$A{:=}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

then

$$\Lambda_\varepsilon(A) \subseteq \Lambda_{\tau(\varepsilon)}(A_{11}) \cup \Lambda_{\tau(\varepsilon)}(A_{22}), \tag{6}$$

where

$$\tau(\varepsilon) = (\varepsilon + a_{21})\sqrt{1 + \frac{a_{12}}{\varepsilon + a_{21}}}.$$

If, moreover, $a_{21} \leqslant \varepsilon$, *we get*

$$\Lambda_{\varepsilon - a_{21}}(A_{11}) \cup \Lambda_{\varepsilon - a_{21}}(A_{22}) \subset \Lambda_\varepsilon(A).$$

Proof. Setting

$$A_0 = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \quad \text{and} \quad \Delta = \begin{bmatrix} 0 & 0 \\ A_{21} & 0 \end{bmatrix},$$

we have $A = A_0 + \Delta$. Thus

$$\sigma_{\min}(A_0 - zI) \leqslant \sigma_{\min}(A - zI) + \sigma_{\max}(\Delta)$$

and

$$\sigma_{\min}(A - zI) \geqslant \sigma_{\min}(A_0 - zI) - a_{21}.$$

Consequently $\Lambda_\varepsilon(A) \subseteq \Lambda_{\varepsilon + a_{21}}(A_0)$ and the first inclusion is obtained by Proposition 3.1.

The second inclusion follows from the inequality

$$\sigma_{\min}(A - zI) \leqslant \sigma_{\min}(A_0 - zI) + \sigma_{\max}(\Delta). \quad \square$$

4. Numerical examples

This section has two parts. First, we compare our inclusion (3) with an inclusion suggested in [9] through an example. Then we give numerical examples illustrating the sharpness of our result. We do not intend to prove that our inclusion is a tool to compute ε-spectra but that our inclusion can be sharp for some matrices.

In [9], Lavallée suggests the following results:

When a matrix is of the form

$$A := \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix},$$

and when $\mathrm{sp}(A_{11}) \cap \mathrm{sp}(A_{22}) = \emptyset$, one can easily eliminate the off-diagonal block A_{12} by the transformation given in the next proposition:

Proposition 4.1. *If* $\mathrm{sp}(A_{11}) \cap \mathrm{sp}(A_{22}) = \emptyset$, *then*

$$A = S \begin{bmatrix} A_{11} & O \\ O & A_{22} \end{bmatrix} S^{-1}$$

with

$$S = \begin{bmatrix} I & -R \\ O & I \end{bmatrix},$$

where R is the solution of the Sylvester equation $A_{11}R - RA_{22} = A_{12}$.

This suggests the following inclusion:

Proposition 4.2. *If*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix}$$

and $\mathrm{sp}(A_{11}) \cap \mathrm{sp}(A_{22}) = \emptyset$, *then*

$$\Lambda_\varepsilon(A) \subseteq \Lambda_{f(\varepsilon)}(A_{11}) \cup \Lambda_{f(\varepsilon)}(A_{22}), \tag{7}$$

where

$$f(\varepsilon) = \varepsilon \left(1 + \frac{\|R\|_2}{2} \left(\sqrt{\|R\|_2^2 + 4} + \|R\|_2\right)\right).$$

Proof. As $A = S\,\mathrm{diag}(A_{11}, A_{22}) S^{-1}$, we have $\Lambda_\varepsilon(A) \subseteq \Lambda_{\varepsilon k_2(S)}(A_{11}) \cup \Lambda_{\varepsilon k_2(S)}(A_{22})$ with $k_2(S) = \|S\|_2 \|S^{-1}\|_2$.

$$S = \begin{bmatrix} I & -R \\ O & I \end{bmatrix},$$

so that $k_2(S) = \|S\|_2^2$. Therefore, using Lemma 2.1

$$k_2(S) = 1 + \frac{\|R\|_2}{2}(\sqrt{\|R\|_2^2 + 4} + \|R\|_2). \quad \square$$

Example 1. We compare inclusion (3) and inclusion (7) using the following simple example:

$$A = \begin{bmatrix} \lambda & M \\ 0 & \lambda + \delta \end{bmatrix}.$$

Then we have

$$S = \begin{bmatrix} 1 & l \\ 0 & 1 \end{bmatrix},$$

with $l = |M|/\delta$.

Hence

$$f(\varepsilon) = \varepsilon\left(1 + \frac{|M|}{2\delta}\left(\sqrt{\frac{|M|^2}{\delta^2} + 4} + \frac{|M|}{\delta}\right)\right).$$

Our result (3) implies that

$$\eta(\varepsilon) = \varepsilon\sqrt{1 + \frac{|M|}{\varepsilon}}.$$

In the following figure (Fig. 1) the curve $f(\varepsilon) = \eta(\varepsilon)$ is plotted as a function of δ and M.

Remark. To the left of the curve $f = \eta$ we have $\eta(\varepsilon) < f(\varepsilon)$ so that $\Lambda_{\eta(\varepsilon)}(A_{11}) \cup \Lambda_{\eta(\varepsilon)}(A_{22}) \subset \Lambda_{f(\varepsilon)}(A_{11}) \cup \Lambda_{f(\varepsilon)}(A_{22})$. This means that $\Lambda_{\eta(\varepsilon)}(A_{11}) \cup \Lambda_{\eta(\varepsilon)}(A_{22})$ is closer to $\Lambda_\varepsilon(A)$ than $\Lambda_{f(\varepsilon)}(A_{11}) \cup \Lambda_{f(\varepsilon)}(A_{22})$. In this region we find values of M and δ corresponding to an ill-conditioned matrix A. If $\delta < 2\varepsilon$, then whatever may be the value of M, our bound (3) is better than (7). For a fixed value of $\delta > 2\varepsilon$ and when the departure from normality M is big enough, our bound (3) is better than (7).

For the following examples, the computations were performed in MATLAB 5.1 on a PC Pentium200 and on a SUN SPARCSTATION 4. In order to draw the ε-spectrum, we use the routine pscont.m from the test matrix toolbox [5] or the program given by Trefethen in [15] which permits to treat large matrices.

In order to investigate the tightness of the inclusion (6), we compare the ε-spectrum of A, with the ε-spectrum and $\tau(\varepsilon)$-spectrum of

$$B := \begin{bmatrix} A_{11} & 0 \\ O & A_{22} \end{bmatrix}.$$

Most of the examples are taken from the test matrix toolbox [5].

Example 2. Let A be a Wilkinson matrix $W(n)$ defined as follows (Figs. 2 and 3):

$$w_{i,i} = i, \quad w_{i,i+1} = n, \quad w_{i,j} = 0\ \forall j \neq i,\ j \neq i + 1.$$

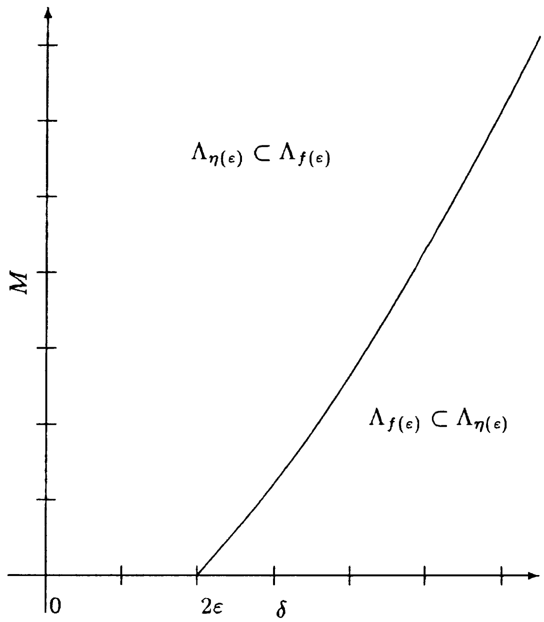

Fig. 1. Comparison of $\eta(\varepsilon)$ with $f(\varepsilon)$.

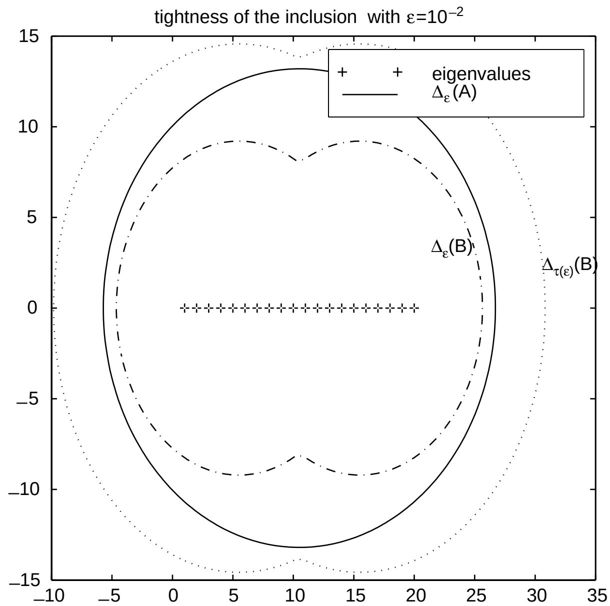

Fig. 2. $n = 20,\ m = 10$.

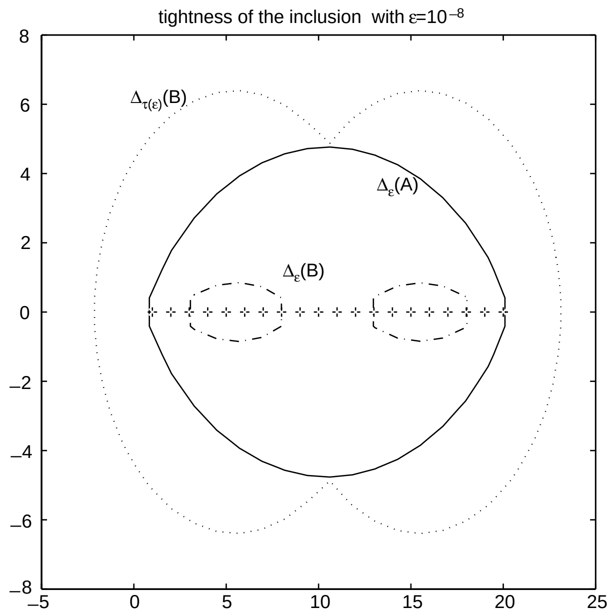

Fig. 3. $n = 20,\ m = 10$.

Remark. We observe that when ε decreases, $\Delta_{\tau(\varepsilon)}(B)$ becomes a better approximation of $\Delta_\varepsilon(A)$ than $\Delta_\varepsilon(B)$. For $\varepsilon = 10^{-12}$, we cannot even see the shape of $\Delta_\varepsilon(B)$.

Example 3. Let

$$A = \begin{bmatrix} \mathrm{Frank}(m) & \alpha I \\ O & \mathrm{Frank}(n-m) \end{bmatrix},$$

where Frank(n) is defined in [5], $\alpha \in \mathbb{R}$ and $I = \mathrm{eye}(m, n-m)$ (Figs. 4 and 5).
$\Delta_{\tau(\varepsilon)}(B)$ is a better approximation of $\Delta_\varepsilon(A)$ than $\Delta_\varepsilon(B)$ when $\alpha = a_{12}$ increases.

Example 4. Let A be the matrix

$$\begin{bmatrix} \mathrm{Frank}(m) & \mathrm{ones} \\ O & \mathrm{Frank}(m) \end{bmatrix},$$

where Frank(n) is defined in [5] and $m = 200$ (Fig. 6).

Remark. Large sizes amplify the differences between $\Delta_\varepsilon(B)$ and $\Delta_{\tau(\varepsilon)}(B)$. $\Delta_{\tau(\varepsilon)}(B)$ is a good approximation of $\Delta_\varepsilon(A)$. On the other hand $\Delta_\varepsilon(B)$ is close to the eigenvalues and far from $\Delta_\varepsilon(A)$.

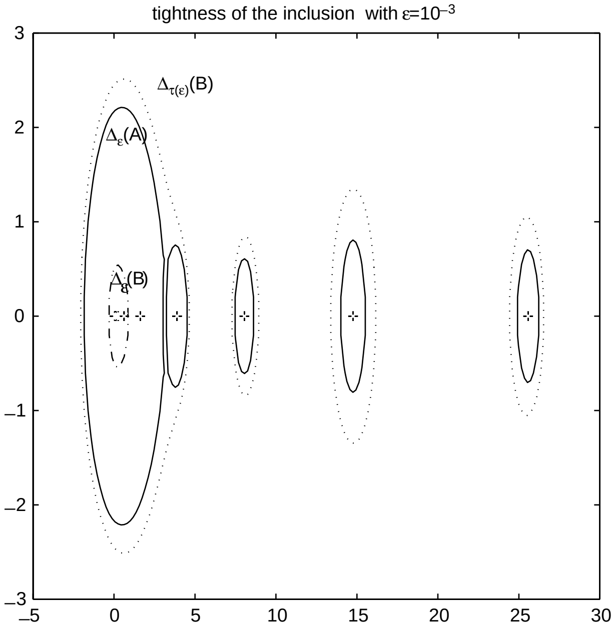

Fig. 4. $n = 20$, $m = 10$, $\alpha = 200$.

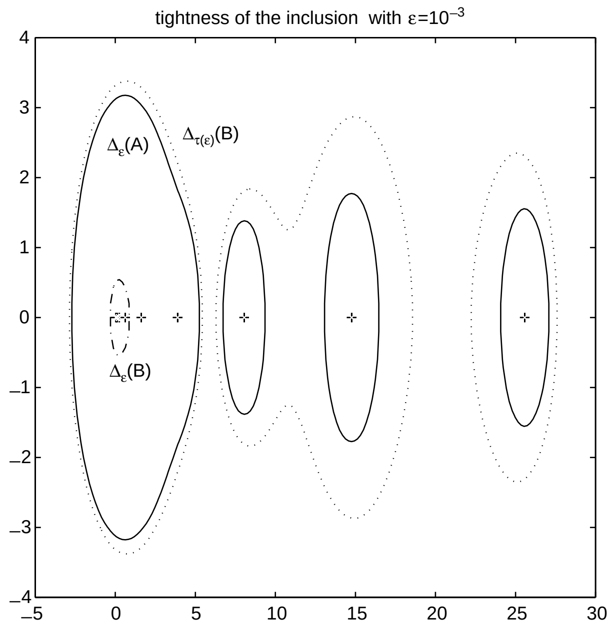

Fig. 5. $n = 20$, $m = 10$, $\alpha = 1000$.

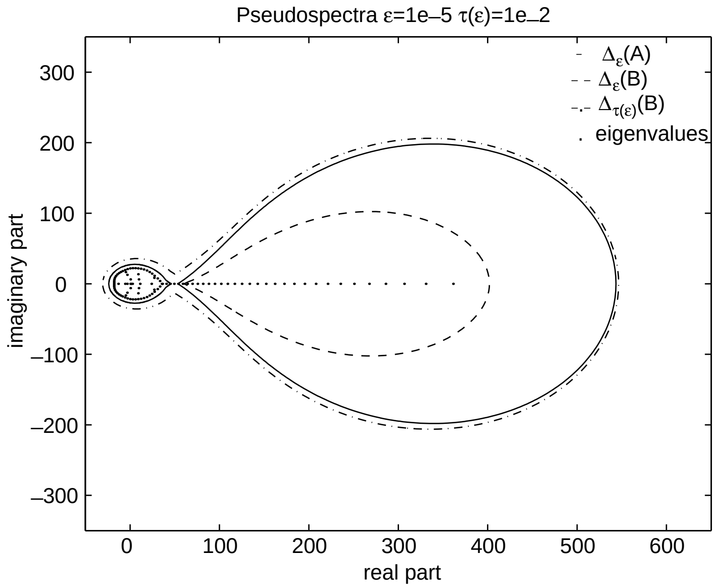

Fig. 6. $m = 200$.

5. Application to matrix stability theory

Systems of ordinary differential equations are often used as a mathematical model in Physics, Biology and Economics. The equilibrium properties or asymptotic behaviour of such systems are of main interest. A useful tool for this analysis is matrix stability analysis (see Definition 1.3).

When dealing with practical applications, we are led to handle large matrices for which there is obviously no way to compute all its eigenvalues. For such a case, Malyshev and Sadkane propose to combine Krylov methods and standard methods in the case of Lyapunov stability. More precisely, $A \in \mathbb{C}^{n\times n}$ can be partitioned in the following way:

$$A = [V_1 \quad V_2] \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix} [V_1 \quad V_2]^* + R, \tag{8}$$

where $A_{11} \in \mathbb{C}^{m\times m}$, $m \ll n$, $V_1 \in \mathbb{C}^{n\times m}$ and $\|R\|$ are given by a Krylov method followed by a refinement scheme such as Arnoldi plus inverse iterations or polynomial acceleration. The other matrices $A_{22} \in \mathbb{C}^{n-m\times n-m}$ and $V_2 \in \mathbb{C}^{n\times n-m}$ will not be computed. The question is: if A_{11} is stable according to the standard stability analysis, can we be sure that A is also stable?

It is interesting to give a lower bound for this radius that guarantees the stability of the matrix A. Malyshev and Sadkane propose a lower bound of $\mathrm{rs}(A)$ involving $\mathrm{rs}(A_{11})$ (computed by a bisection method for example) and the smallest eigenvalue of $-(A_{22} + A_{22}^*)/2$ (in [10]).

Let

$$A_0 = [V_1 \quad V_2] \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix} [V_1 \quad V_2]^*,$$

$$R = [V_1 \quad V_2] \begin{bmatrix} O & O \\ A_{21} & O \end{bmatrix} [V_1 \quad V_2]^*,$$

$$\mu = \lambda_{\min}\left(-\frac{A_{22}+A_{22}^*}{2}\right),$$

$$s_1 = \mathrm{rs}(A_{11}),$$

$$a_{21} = \|R\|_2.$$

The following proposition recalls the bound obtained in [10]:

Proposition 5.1. *Let*

$$t_1 = \max\left(\frac{1}{\mu}, \frac{1}{r_1} + \frac{a_{12}}{\mu r_1}\right),$$

$$t_2 = \max\left(\frac{1}{r_1}, \frac{1}{\mu}, + \frac{a_{12}}{\mu r_1}\right),$$

$$t_3 = \sqrt{\left(\frac{1}{r_1^2} + \frac{1}{\mu^2} + \frac{a_{12}^2}{\mu^2 r_1^2}\right)}.$$

If A_{11} is stable, $(A_{22}+A_{22}^*)/2$ is negative definite, and

$$a_{21} < \frac{1}{\min(\sqrt{t_1 t_2}, t_3)}$$

then A is stable and

$$\mathrm{rs}(A) \geqslant \frac{1}{\min(\sqrt{t_1 t_2}, t_3)} - a_{21}. \tag{9}$$

In this section, we propose another lower bound and a sufficient condition for a matrix to be stable.

We quote the following lemma from [10].

Lemma 5.2.

$$s_2 \geqslant \lambda_{\min}\left(\frac{-A_{22}-A_{22}^*}{2}\right).$$

It will be used in the proof of the following theorem:

Theorem 5.3. *Let set $s = \max(s_1, \mu)$. If A_{11} is unstable, then A_0 is unstable and*

$$\mathrm{rs}(A) \leqslant a_{21}.$$

If A_{11} is stable and $(A_{22}+A_{22}^*)/2$ is negative definite, then

$$\mathrm{rs}(A_0) \geqslant \min\{s_1,\mu\} \frac{s}{\sqrt{s^2+\frac{a_{12}^2}{2}+\frac{a_{12}}{2}\sqrt{4s^2+a_{12}^2}}}. \tag{10}$$

If moreover

$$a_{21} < \min\{s_1,\mu\} \frac{s}{\sqrt{s^2+\frac{a_{12}^2}{2}+\frac{a_{12}}{2}\sqrt{4s^2+a_{12}^2}}},$$

then A is stable and

$$\mathrm{rs}(A) \geqslant \min\{s_1,\mu\} \frac{s}{\sqrt{s^2+\frac{a_{12}^2}{2}+\frac{a_{12}}{2}\sqrt{4s^2+a_{12}^2}}} - a_{21}. \tag{11}$$

Proof. We have $\sigma_{\min}(A_0 - zI) \leqslant \sigma_{\min}(A_{11} - zI)$. If $\min_{\mathrm{re}(z)=0} \sigma_{\min}(A_{11} - zI) \leqslant 0$ then $\min_{\mathrm{re}(z)=0} \sigma_{\min}(A_0 - zI) \leqslant 0$ and then A_0 is unstable. Since $A = A_0 + R$, $\sigma_{\min}(A - zI) = \sigma_{\min}(A_0 - zI + R) \leqslant \sigma_{\min}(A_0 - zI) + \sigma_{\max}(R)$. But $\sigma_{\max}(R) = \|R\|_2 = a_{21}$, so $\sigma_{\min}(A - zI) \leqslant \sigma_{\min}(A_0 - zI) + a_{21}$ and $\min_{\mathrm{re}(z)=0} \sigma_{\min}(A - zI) \leqslant \min_{\mathrm{re}(z)=0} \sigma_{\min}(A_0 - zI) + a_{21} \leqslant a_{21}$. Hence, $\mathrm{rs}(A) \leqslant a_{21}$.

Let us assume that A_{11} is stable and that $(A_{22}+A_{22}^*)/2$ is negative definite. According to Definition 1.3,

$$\frac{1}{\mathrm{rs}(A_0)} = \max_{\mathrm{re}(z)\geqslant 0} \|(A_0 - zI)^{-1}\|_2.$$

But $V = [V_1 \quad V_2]$ is unitary, so

$$\|(A_0 - zI)^{-1}\|_2 = \left\| \left(\begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix} - zI \right)^{-1} \right\|_2 .$$

Taking the maximum of the set $\mathrm{re}(z) \geqslant 0$ and taking into account

$$\max_{\mathrm{re}(z)\geqslant 0} r_1(z) = \frac{1}{\mathrm{rs}(A_{11})} = \frac{1}{s_1}$$

and

$$\max_{\mathrm{re}(z)\geqslant 0} r_2(z) = \frac{1}{\mathrm{rs}(A_{22})} = \frac{1}{s_2},$$

we have

$$\frac{1}{\mathrm{rs}(A_0)} \leqslant \frac{1}{\min(s_1,s_2)} \sqrt{1 + \frac{a_{12}}{2\max(s_1,s_2)} \left(\frac{a_{12}}{\max(s_1,s_2)} + \sqrt{\frac{a_{12}^2}{\max(s_1,s_2)^2} + 4} \right)}.$$

By the previous lemma we have $\max(s_1,s_2) \geqslant s$ and $\min(s_1,s_2) \geqslant \min(s_1,\mu)$.

These relations lead to the inequality (10).

Table 1

Quantity	Computation
μ	$4.91 \cdot 10^{-2}$
s_1	$2.71 \cdot 10^{-3}$
a_{12}	0.65
t_1	$1.01 \cdot 10^{4}$
t_2	$9.83 \cdot 10^{3}$
t_3	$9.81 \cdot 10^{3}$
s	$4.91 \cdot 10^{-2}$

From following

$$\sigma_{\min}(A_0 - zI) \leqslant \sigma_{\min}(A - zI) + a_{21},$$

we have inequality (11) and the sufficient condition on the stability of A. $\square$

5.1. Numerical experiment

We illustrate the theoretical results of Theorem 5.3 on a practical test problem taken from [10] in order to compare our results with the results obtained in [10].

We test the bounds on the Orr–Sommerfeld operator defined by

$$\frac{1}{\alpha R} L^2 y - i(ULy - U'' y) - \lambda L y = 0,$$

where α and R are positive parameters, λ is a spectral parameter number, $U = 1 - x^2$, y is a function defined on $[-1, 1]$ with $y(\pm 1) = y'(\pm 1) = 0$, $L = (d^2/\mathrm{d}x^2) - \alpha^2$. The discretization

$$x_i = -1 + ih, \quad h = \frac{2}{n+1},$$

$$L_h = \frac{1}{h^2}\,\mathrm{Tridiag}(-1, -2 - \alpha^2 h^2, 1),$$

$$U_h = \mathrm{diag}(1 - x_1^2, \ldots, 1 - x_n^2)$$

with $\alpha = 1$, $R = 1000$, $n = 400$ yields to the eigenvalue problem

$$Au = \lambda u, \quad A = \frac{1}{\alpha R} L_h - i L_h^{-1}(U_h L_h + 2I_n).$$

The block decomposition (1) of A is performed by Arnoldi method combined with complex Chebyshev acceleration (see [10]). Arnoldi iterations are stopped at $m = 10$. The results of the computations of the quantities involved in the inequalities (11) and (9) are taken from [10] for the sake of comparison.

Using the quantities given in the above Table 1, we obtain from inequality (9)

$$\mathrm{rs}(A) \geqslant 1.01 \cdot 10^{-4} - a_{21}$$

and from inequality (11)

$$\mathrm{rs}(A) \geqslant 2.24 \cdot 10^{-4} - a_{21}.$$

For this example, our lower bound of $\mathrm{rs}(A_0)$ is more than twice as the one of [10].

6. Conclusion

We give a new upper bound for the norm of the resolvent. From it, we deduce an approximation of $\Delta_\varepsilon(A)$ which contains $\Delta_\varepsilon(A)$. As, in general, the ε-spectrum is approximated from below, our result is interesting. Also the only approximation of that kind which we know seems to be worse than ours in the case of ill-conditioned matrices (see Fig. 1). As the numerical examples show, we should not neglect the upper off-diagonal block of a matrix to compute its ε-spectrum (see Fig. 5). If we want to cancel this term, we have to increase the value of ε.

As far as the stability radius of a partitioned matrix is concerned, we set up a lower bound. We test this bound on the Orr–Sommerfeld operator. Our bound is twice as good as the one established in [10].

Acknowledgements

The authors wish to thank Miloud Sadkane whose ideas motivated this paper.

References

[1] S.K. Godunov, Spectral portraits of matrices and criteria of spectrum dichotomy, in: J. Herzberger, L. Atanassova (Eds.), International Symposium on Computer Arithmetic and Scientific Computation, Oldenburg, Germany, North-Holland, Amsterdam, 1991.
[2] S.K. Godunov, A. Antonov, O.P. Kiriljuk, V.I Kostin, Guaranteed Accuracy in Numerical Linear Algebra, Kluwer Academic Publishers, Dordrecht, 1993.
[3] S.K. Godunov, O.P. Kiriljuk, V.I Kostin, Spectral portraits of matrices, Technical Report 3, Institute of Mathematics, Academy of Science URSS, Novosibirsk, 1990.
[4] K.E. Gustafson, D.K.M. Rao, Numerical Range, The Field of Values of Linear Operators and Matrices, Springer, New York, 1996.
[5] N.J. Higham, The test matrix toolbox for matlab (version 3.0), Technical Report, Manchester Centre for Computational Mathematics, 1995.
[6] D. Hinrichsen, B. Kelb, Spectral value sets: a graphical tool for robustness analysis, Systems Control Lett. 21 (1993) 127–136.
[7] D. Hinrichsen, A.J. Pritchard, On spectral variations under bounded real matrix perturbations, Numer. Math. 60 (1992) 509–524.
[8] H.J. Landau, The notion of approximate eigenvalues applied to an integral equation of laser theory, Quart. Appl. Math. 35 (1977) 165–172.
[9] P.F. Lavallée, Nouvelles approches de calcul du ε-spectre de matrices et de faiscaux de matrices, Ph.D. Thesis No. 1876, Octobre 1997.
[10] A.N. Malyshev, M. Sadkane, On the stability of large matrices, J. Comput. Appl. Math. 102 (1999) 303–313.
[11] S. Parter, Stability, convergence and pseudo-stability of finite difference equations for an over-determined problem, Numer. Math. 4 (1962) 277–292.

[12] K.C. Toh, L.N. Trefethen, Calculation of pseudospectra by the Arnoldi iteration, SIAM J. Sci. Comput. 17 (1996) 1–15.
[13] L.N. Trefethen, Approximation theory and numerical linear algebra, in: J. Mason, M. Cox (Eds.), Algorithms and Approximation II, 1991, Chapman & Hall, London, 1990, pp. 336–360.
[14] L.N. Trefethen, Pseudospectra of matrices, in: D.F. Griffiths, G.A. Watson (Eds.), Numerical Analysis, Longman, Harlow, 1992, pp. 234–266.
[15] L.N. Trefethen, Computation of pseudospectra, Acta Numerica 8 (1999) 247–295.
[16] J. Varah, On the separation of two matrices, SIAM J. Numer. Anal. 16 (2) (1979) 216–222.