INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Spectral portrait for non hermitian large sparse matrices

J.F. Carpraux, J. Erhel, M. Sadkane

N° 2119

Novembre 1993

PROGRAMME 6

Calcul scientifique,
modélisation
et logiciels numériques

apport de recherche

1993

ISSN 0249-6399

INRIA
RENNES

Spectral portrait for non hermitian large sparse matrices

J.F. Carpraux, J. Erhel, M. Sadkane

Programme 6 — Calcul scientifique, modélisation et logiciel numérique
Projet Aladin

Rapport de recherche n°2119 — Novembre 1993 — 14 pages

Abstract: The spectral portrait of a matrix is the picture of its ϵ-spectra for $\epsilon \in [\epsilon_1, \epsilon_2]$, where an ϵ-spectrum of A is the union of all the eigenvalues of all the matrices $A + \Delta$ with $\|\Delta\|_2 \leq \epsilon \|A\|_2$. The spectral portrait is useful to study the stability of various problems, for example, or, as we illustrate in this paper, to visualize the condition number of an eigenvalue. Some methods to estimate the spectral portrait already exist, but only for small matrices. We propose here a new algorithm for non hermitian large sparse matrices.

Key-words: Spectral portrait, ϵ-spectrum, Condition number of an eigenvalue, Davidson's method.

(Résumé : tsvp)

Unité de recherche INRIA Rennes
IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex (France)
Téléphone : (33) 99 84 71 00 – Télécopie : (33) 99 38 38 32

Portrait spectral pour grande matrice creuse non hermitienne

Résumé : Le portrait spectral d'une matrice consiste à dessiner ses ϵ-spectres pour $\epsilon \in [\epsilon_1, \epsilon_2]$, où un ϵ-spectre de A est l'ensemble de toutes les valeurs propres de toutes les matrices $A + \Delta$ où $\|\Delta\|_2 \leq \epsilon \|A\|_2$. Le portrait spectral est, par exemple, très utile pour l'étude de la stabilité de nombreux problèmes ou, comme on illustre dans ce rapport pour visualiser le conditionnement de valeurs propres. Des méthodes de calcul de portraits spectraux existent déjà, mais seulement pour des petites matrices. On propose, ici, un algorithme pour les matrices de grande taille.

Mots-clé : Portrait spectral, ϵ-spectre, Conditionnement d'une valeur propre, Methode de Davidson.

1 Introduction

When computing some eigenvalues of a large square sparse matrix $A \in \mathbb{C}^{n \times n}$, this matrix is quite often the result of a previous computation. Moreover, the backward analysis [12] of algorithms for computing eigenvalues aims at proving that these computed eigenvalues are, in fact, the exact one of a nearby matrix. So, we have to consider that we compute some exact eigenvalues of a matrix $A + \Delta$, $\Delta \in \mathbb{C}^{n \times n}$ such that $\|\Delta\|_2$ is small, and we wish that these eigenvalues are not too far from the eigenvalues of A. In other words, we want to bound the error on an eigenvalue by some constant times the perturbation $\|\Delta\|_2$. This perturbation analysis leads to define and estimate, if it exists, the constant above, which is called the condition number of the eigenvalue [2].

In the hermitian case, it is well-known that the condition number of an eigenvalue is equal to one, so that the error done on the computed eigenvalues is only of the order of $\|\Delta\|_2$. On the contrary, condition numbers can be very large in the non hermitian case. In particular, the condition number of defective eigenvalues is infinite. Several condition number estimators have been designed, see [1] for example. Another approach to study the eigenvalues of perturbed matrices is to draw the spectral portrait of the matrix. It amounts to estimate all the eigenvalues of all the perturbed matrices $A + \Delta$, with $\|\Delta\|_2$ varying in a prescribed range. This spectral portrait provides a lot of information and can be used in various problems [11], for example:

- To measure the distance to a singular matrix :

$$\text{find } \min\{\|E\|_2 \text{ such that } A + E \text{ is singular}\}.$$

- The stability of some problems occurs when the eigenvalues have negative real parts. Thanks to a spectral portrait, we can measure the quantity

$$\min\{\|E\|_2 \text{ such that the problem with } A + E \text{ is not stable}\}.$$

- Study the convergence of linear iterative solvers.

Some methods to estimate a spectral portrait already exist, but only for small matrices [8, 11]. These methods are based on the Singular Value Decomposition algorithm, see for example [7], but this algorithm cannot be

applied to large matrices because of the expense of storage requirements.

This paper is organized as follows : in section 2 we recall the link between the condition number and the spectral portrait, in sections 3,4 we propose an algorithm for computing the spectral portrait of large matrices based upon the computation of the smallest singular value by a modification of the Davidson's method [10], in section 5 we give some numerical examples.

2 Condition number and spectral portrait

2.1 Condition number of the eigenvalue problem

The condition number of an eigenvalue consists in a measure of the variation of this eigenvalue through a matrix perturbation Δ [2].
If the eigenvalue λ of A is not defective, the error $|\Delta\lambda| = |\lambda' - \lambda|$, where λ' is an eigenvalue of $A + \Delta$, can be bounded by $|\Delta\lambda| \leq C_\lambda \|\Delta\|_2$, where C_λ is the condition number of λ.

2.2 Spectral portrait in the complex plane

The spectral portrait of a matrix is the picture of its ϵ-spectra for $\epsilon \in [\epsilon_1, \epsilon_2]$, where the ϵ-spectrum of A, denoted by $\Lambda_\epsilon(A)$, is, for fixed ϵ, the union of all the eigenvalues of all the matrices $A+\Delta$ with $\|\Delta\|_2 \leq \epsilon \|A\|_2$.

Definition 1 *Let $\mu \in \mathbb{C}$, then $\mu \in \Lambda_\epsilon(A)$ if there exists a matrix Δ ($\|\Delta\|_2 \leq \epsilon \|A\|_2$) such that $det(A + \Delta - \mu I) = 0$.*

This definition 1 is equivalent to [8]:

Definition 2

$$\mu \in \Lambda_\epsilon(A) \quad if \quad \|(A - \mu I)^{-1}\|_2 \geq \frac{1}{\epsilon \|A\|_2}.$$

2.3 Relation between condition number and spectral portrait

The spectral portrait can be used to give an estimation of the condition number of the eigenvalue problem. Indeed, for fixed ϵ, $\Lambda_\epsilon(A)$ is the union of

patches around clusters of eigenvalues of A. For $\epsilon = 0$, $\Lambda_0(A) = \Lambda(A)$: the patches are reduced to points (the different eigenvalues of A), and there is a value of ϵ for which $\Lambda_\epsilon(A)$ contains all the eigenvalues of A ($\epsilon = 2\,\|A\|_2$). So, let λ be a non defective eigenvalue of A, then we can consider some $\epsilon \in]\,0, \epsilon_\lambda]$ for which the cluster of eigenvalues of the patch , $_\epsilon$ around λ is reduced to λ. For these values of ϵ, we can estimate the condition number by :

$$C_\lambda \approx \frac{diam(,\,_\epsilon)}{\epsilon\,\|A\|_2} \qquad \text{with} \quad diam(,\,_\epsilon) = \max_{\mu_1,\mu_2 \in \Gamma_\epsilon} |\mu_1 - \mu_2|$$

The computation of the spectral portrait consists in the computation of $\epsilon(\mu, A) = \|(A - \mu I)^{-1}\|_2^{-1}$ for μ describing a grid of the complex plane.

Remarks:

1. $\mu \in \Lambda_\epsilon(A)$ for all $\epsilon \geq \dfrac{\epsilon(\mu, A)}{\|A\|_2}$;
2. $\epsilon(\mu, A) = \|(A - \mu I)^{-1}\|_2^{-1} = \sigma_{min}(A - \mu I)$ where $\sigma_{min}(A - \mu I)$ is the smallest singular value of $A - \mu I$.

3 Computation of $\sigma_{min}(A_\mu)$ where $A_\mu = (A - \mu I)$

The following algorithm is an adaptation of the modified Davidson's algorithm which computes the smallest singular value ν and the associated right singular vector x of a large sparse matrix [10]. In fact, it computes the smallest eigenvalue and the associated eigenvector of the hermitian matrix $A_\mu^H A_\mu$ (where A_μ^H is the conjugate transpose of A_μ).

The idea behind this algorithm is to build gradually a dense hermitian matrix $H_k := V_k^H\, A_\mu^H\, A_\mu\, V_k$ (steps 1-3) using projection techniques, then we compute the smallest eigenpair of the projected matrix H_k (step 4). If this eigenpair is a good approximation of the smallest eigenpair of $A_\mu^H A_\mu$ then we stop (step 6). Otherwise, we increase the subspace V (step 8) by incorporating a new direction (step 7) to the previous subspace. This algorithm is an algorithm with restart, this means that if the size of the basis V is greater than a fixed size m, we restart the algorithm with the last Ritz vector x_k and the corresponding direction t_k (step 8).

C_k stands for a $n \times n$ preconditioning matrix whose choice is discussed in [10]. Here, C_k is an approximation of $(A_\mu^H A_\mu)^{-1}$ (in fact, we realize an incomplete LU decomposition of A_μ). MGS stands for Modified Gram Schmidt Procedure.

Algorithm 1 :

Choose m and tol

Choose an initial vector $V_1 \in \mathbb{C}^{n \times 1}$, such that $\|V_1\|_2 = 1$

for $k = 1, \ldots$ **do**

1. Compute the matrix $U_k := A_\mu V_k$
2. Compute the matrix $W_k := A_\mu^H U_k$
3. Compute the Rayleigh matrix $H_k := V_k^H W_k$
4. Compute the smallest eigenpair (ν_k^2, y_k) of H_k
5. Compute the Ritz vector $x_k := V_k y_k$
6. Compute the residual $r_k := W_k y_k - \nu_k^2 x_k$
   if $\|r_k\| \leq tol$ **then** exit
7. Compute the new direction $t_k := C_k r_k$
8. **if** $\dim(V_k) \leq m - 1$
   then $V_{k+1} := [V_k, \dfrac{t_k}{\|t_k\|_2}]$ where $t_k := (I - V_k V_k^H) t_k$
   else $V_{k+1} := MGS(x_k, t_k)$
   end if

end for

At convergence, ν_k and x_k approximate the sought singular value and singular vector.
An important characteristic of Algorithm 1 is that the matrix $A_\mu^H A_\mu$ is not required explicitly. We only need two subroutines that compute $A_\mu u$ and $A_\mu^H v$ for given u and v. At step k, the basis V_{k+1} is obtained from V_k by incorporating the vector $t_k := C_k r_k$ after orthonormalization. The subspace spanned by V_k is not a Krylov subspace, and since the matrice C_k is not diagonal, a linear system must be solved at each iteration. The hope is to reach the convergence very quickly with a small value of m, thus rewarding the extra cost involved by this system resolution. A detailed convergence analysis of Algorithm 1 can be found in [3], and a simplified convergence

result for the smallest singular value in [10].

4 Computation of the spectral portrait

We cannot compute the spectral portrait of a matrix A "everywhere" (more precisely in the disk centered at 0 of radius $\|A\|_2$), because this computation would be too expensive. So, we assume that we only want to know the spectral portrait in the neighbourhood of a complex value, to check for example if this value is a good approximation for an eigenvalue of A. So, we define a grid of the complex plane on which we want to draw the spectral portrait. For this, we give two points of the complex plane, (x_1, y_1) (bottom left point of the grid) and (x_2, y_2) (upper right one), and the number of points in the two directions, nx and ny.

The spectral portrait computation consists merely in the computation of $\epsilon(\mu, A) = \sigma_{min}(A - \mu I)$ for each $\mu = x + iy$ of the grid.

Algorithm 2 :

1. $\mu = x_1 + i\, y_1, \qquad step_x = \dfrac{x_2 - x_1}{nx - 1}, \qquad step_y = \dfrac{y_2 - y_1}{ny - 1}$

for $j = 1, nx$ **do**

  for $k = 1, ny$ **do**

  2. Compute $\sigma_{min}(A - \mu I)$ (by Algorithm 1)

  3. $\mu = \mu + i\, step_y$ (next μ in the current column)

end for

4. $\mu = \mu + step_x$ (next column)

5. $step_y = -step_y$ (we change the direction of going through the column)

end for

We see that we have to compute $\sigma_{min}(A - \mu I)$ for nearby μ. For that reason, it is interesting that, in Algorithm 1, we compute not only the smallest singular value but also the associated vector. Indeed, in the first step of Algorithm 1, we have to choose an initial basis V which is a vector. Now, during the computation of $\sigma_{min}(A - \mu I)$ for one μ, we increase the basis V by adding at each iteration a new direction, and the last one provides convergence. So, if we take it as an initial vector for the computation of $\sigma_{min}(A - \mu' I)$ where μ' is near from μ, then it should improve the convergence since closed matrices have closed singular values and often closed singular vectors.

For this reason, we go through the grid as described in Algorithm 2 (steps 3,4,5). We sweep the grid column by column and alternatively from bottom to top and from top to bottom. This means that we always deal with closed values μ.

Now we define a color map by subdividing the range $[\min\{\epsilon(\mu, A)\}, \max\{\epsilon(\mu, A)\}]$ into intervals of equal length and by assigning a color to each interval. Therefore each point μ of the grid will be affected to a color according to the value $\epsilon(\mu, A)$. Points of the same color correspond to an ϵ-spectrum.

In fact, it would be very interesting to find an effective method, to compute only the ϵ-spectrum for given ϵ. Indeed, we can imagine, that the user knows the error done on the computation of the matrix A, and with this error he wants only to know the corresponding ϵ-spectrum. To realize this, we can follow the level lines [8].

5 Examples of spectral portraits

5.1 Comparison between the algorithm based on SVD and Algorithm 2

We would like to show here the validity of Algorithm 2. We choose as an example a matrix with two ill-conditioned eigenvalues, hence with a large ϵ-spectrum even for small ϵ. The spectral portraits computed on one hand with the SVD algorithm and on the other hand with Algorithm 2, are depicted in Figures 1, 2 for the following triangular matrix A [6]:

$$A = \begin{pmatrix} -2 & 25 & 0 & 0 & 0 & 0 & 0 \\ 0 & -3 & 10 & 3 & 3 & 3 & 0 \\ 0 & 0 & 2 & 15 & 3 & 3 & 0 \\ 0 & 0 & 0 & 0 & 15 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 & 10 & 0 \\ 0 & 0 & 0 & 0 & 0 & -2 & 25 \\ 0 & 0 & 0 & 0 & 0 & 0 & -3 \end{pmatrix}$$

Clearly, the eigenvalues are $\{-3, -2, 0, 2, 3\}$ where -3 and -2 are eigenvalues of multiplicity 2, which are ill-conditioned as can be seen on the spectral portrait.
Even in this difficult case where very small singular values are attained at points of the grid, Figures 1 and 2 are nearly the same.

The grid used in Figures 1, 2 and 3 is defined by:

$$(x_1, y_1) = (-4, -1), \ \ (x_2, y_2) = (4, 1), \ \ nx = ny = 100$$

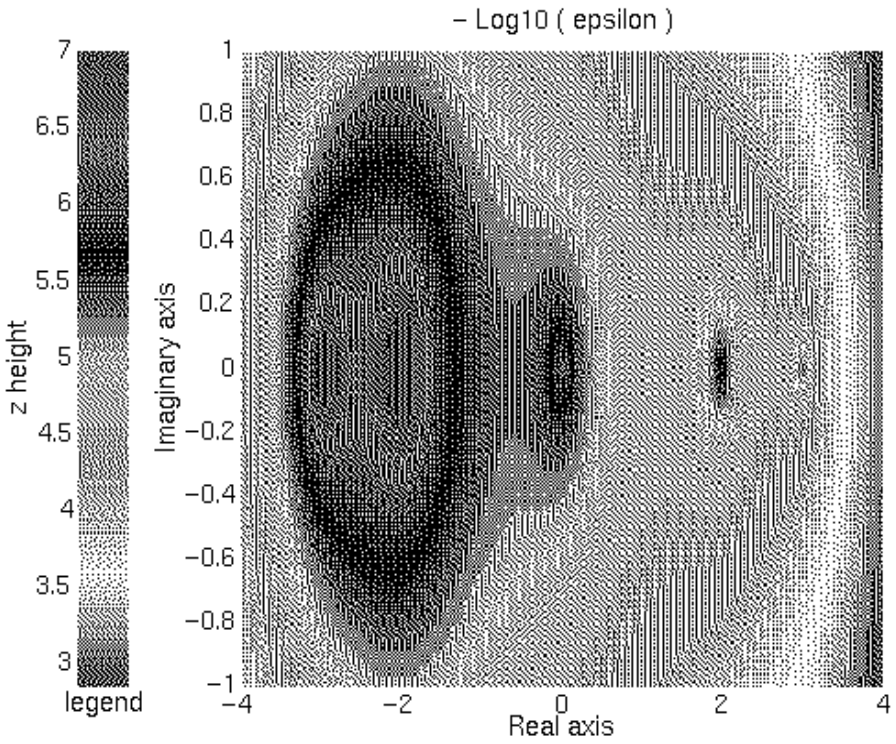

Figure 1: using Singular Value Decomposition

To estimate the condition number C_λ of an eigenvalue λ of a matrix A thanks to a spectral portrait (see section 2.3), we consider the diameter of

a patch , $_\epsilon$ around λ, and the value of ϵ corresponding to , $_\epsilon$. To estimate ϵ, we look at the value corresponding to the color of , $_\epsilon$ in the legend : this is the value of $-Log10(\epsilon)$, so we infer the value of ϵ. An estimation of C_λ is then given by the ratio between the diameter of , $_\epsilon$ and ϵ.

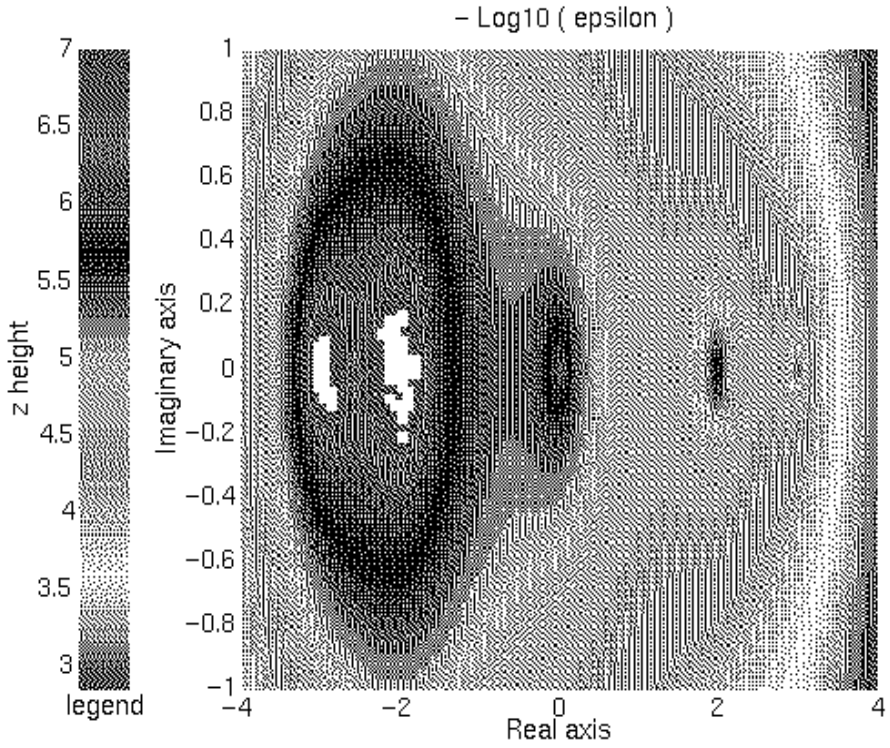

Figure 2: using Algorithm 2 with $tol = 10^{-3}$ and $m = 2$

As we can see in Figure 2 computed with Algorithm 2, some values are missing. These values correspond to very small singular values of the order $10^{-7}, 10^{-8}$, that is to say to small eigenvalues of $A_\mu^H A_\mu$ of the order $10^{-14}, 10^{-16}$. It is well-known that $\|\nu_k^2 - \nu^2\| \leq C_{\nu^2}\|r_k\|^2$ (see [9] for example). Hence the convergence threshold *tol* must be very small to estimate small eigenvalues.

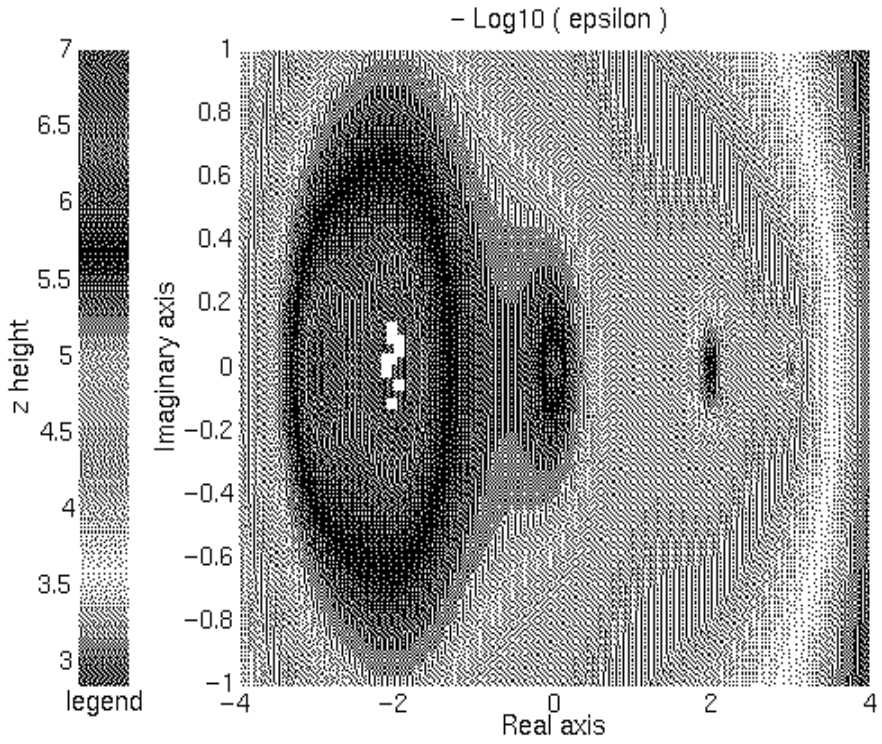

Figure 3: using Algorithm 2 with $tol = 10^{-14}$ and $m = 2$

We plot in Figure 3 the same spectral portrait as in Figure 2 with $tol = 10^{-14}$ instead of 10^{-3} and observe now only a few missing values, as expected. But the computation time is very high. For example, in the previous example, with $tol = 10^{-14}$ the computation time is at least five times greater than with $tol = 10^{-3}$. So, we prefer to keep tol not too small, because we have enough information with this value and the time of computation is reduced.

5.2 Spectral portrait of a few large matrices

We give here examples of spectral portraits for three non hermitian large sparse matrices coming from the Harwell-Boeing set of test matrices [5]. We can see on the three pictures that the eigenvalues are well conditioned, since the values of $-Log10(\epsilon)$ are not too high (see legend).

The following picture is the spectral portrait of the matrix HOR131. It arises in the flow network problem. It is square of order 434 with 4710 nonzero elements. The grid used is defined by:

$$(x_1, y_1) = (-0.5, -0.2),\ \ (x_2, y_2) = (1, 0.2),\ \ nx = 1000,\ \ ny = 10$$

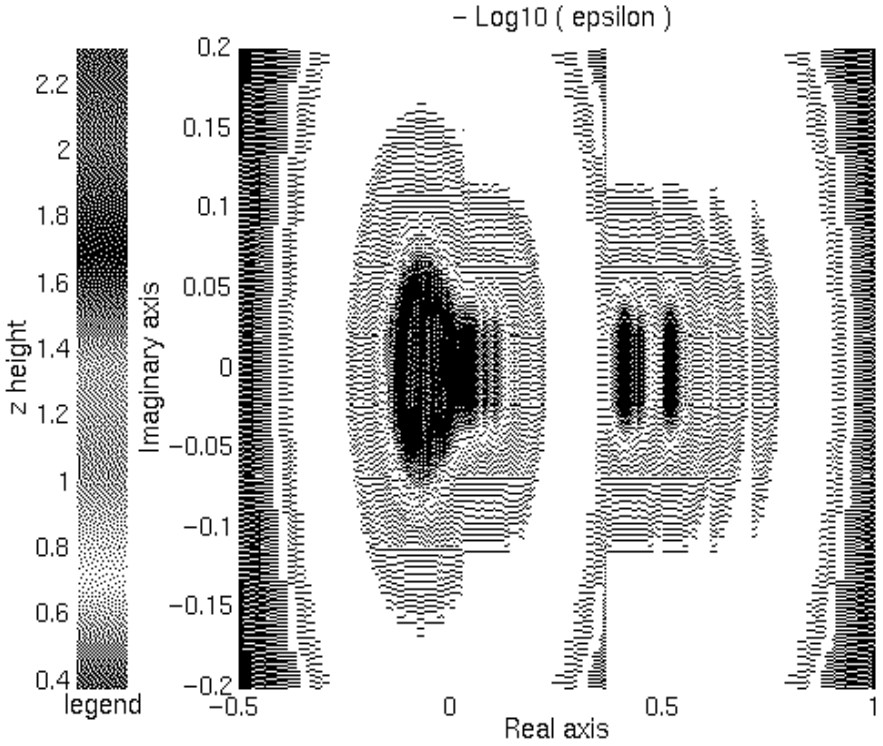

Figure 4: Matrix HOR131, $tol = 10^{-3}$, $m = 40$

In Figure 4, we see that the eigenvalues are close from each other, but well conditioned.

The next picture was realized with the matrix PORES3. It arises from reservoir simulation. It is square of order 532 and has 3474 nonzero elements. The grid used is defined by:

$$(x_1, y_1) = (-6150, -3.5), \quad (x_2, y_2) = (-6100, 35), \quad nx = 50, \quad ny = 100$$

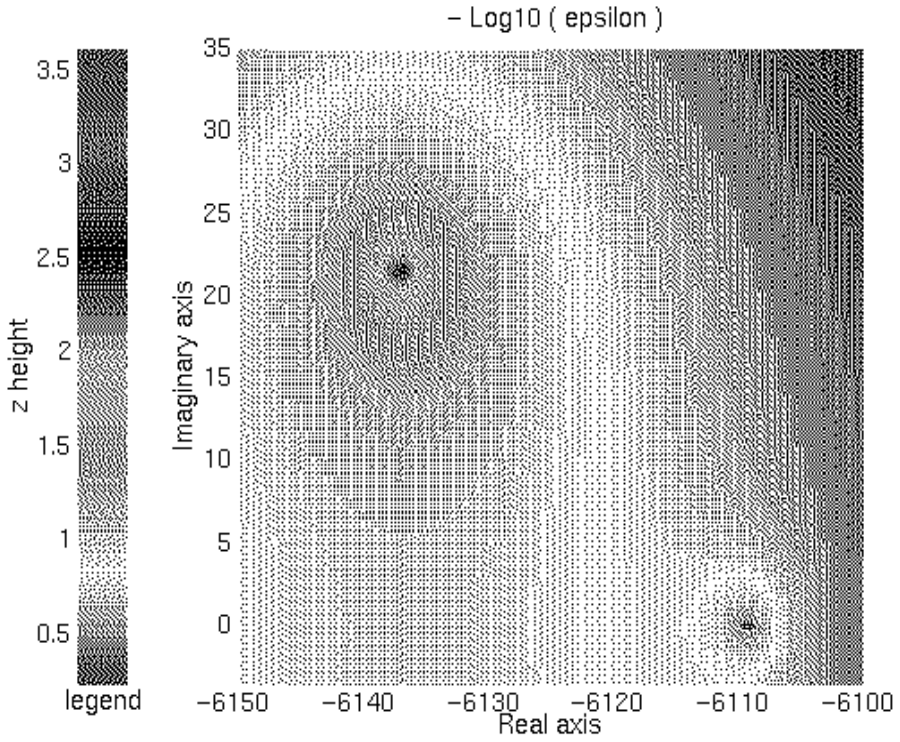

Figure 5: Matrix PORES3, $tol = 10^{-4}$, $m = 40$

In Figure 5, we can see that the condition number of the eigenvalue in the top left corner is not so good than for the other eigenvalue, because the patches around the first eigenvalue are bigger.

6 Conclusion

We have proposed an algorithm for estimating the spectral portrait of non hermitian large sparse matrices. The algorithm retains the advantage of Davidson's procedure in that the matrix A (resp. A^H) is accessed only in the form of matrix vector products. We would like to conclude with the following remarks :

- The algorithm we proposed can easily be parallelized. It involves sparse matrix-vector products and BLAS primitives.
- The choice of the preconditioner (step 7 of Algorithm 1) is crucial for the success of the method and it can be improved.
- We cannot easily reduce the number of grid points, because we must be sure to capture all the eigenvalues.

References

[1] Z. Bai, J. Demmel, A. McKenney, *On the conditioning of the nonsymmetric eigenproblem: Theory and software.* LAPACK Working Note 13, 1989.

[2] F. Chatelin, *Valeurs propres de matrices.* Collection mathématiques appliquées pour la maîtrise, Masson, 1988.

[3] M. Crouzeix, B. Philippe, M. Sadkane, *The Davidson method.* Tech. Rep., INRIA Report No. 1353, 1990. To appear in SIAM J. Sci. Stat. Comput.

[4] E.R. Davidson, *The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices.* Comp. Phys., 1975, pp. 87-94.

[5] I.S. Duff, R.G. Grimes, J.G. Lewis, *Sparse matrix test problems.* ACM Trans. Math. Softw. 15, 1-14, 1989.

[6] S.K. Godunov, *Spectral Portraits of Matrices and Criteria of Spectrum Dichotomy.* Research Report, Novosibirsk.

[7] G.H. Golub, C.F. Van Loan, *Matrix Computations.* The Johns Hopkins University Press, Baltimore and London, 1989.

[8] V.I. Kostin, *On definition of matrices' spectra.* High Performance Computing II, 1991.

[9] B.N. Parlett, *The Symmetric Eigenvalue Problem.* Prentice Hall, Englewood Cliffs, 1980.

[10] B. Philippe, M. Sadkane, *Computation of the singular subspace associated with the smallest singular values of large matrices.* IRISA Report No 754, 1993.

[11] L.N. Trefethen, *Non-Normal Matrices and Pseudo-Eigenvalues.* Book to appear.

[12] J.H. Wilkinson, *Rounding errors in the algebraic processes* Englewoods Cliffs, N.J : Prentice-Hall, 1963.

Unité de recherche INRIA Lorraine, Technôpole de Nancy-Brabois, Campus scientifique,
615 rue de Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105,
78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS
Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
(France)
ISSN 0249-6399