Journal of Computational Mathematics, Vol.23, No.6, 2005, 657 –672.

A DYNAMICS APPROACH TO THE COMPUTATION OF EIGENVECTORS OF MATRICES *

S. Jiménez
(*Colaborador honorífico, Dept. Matemática Aplicada, Universidad Complutense de Madrid, 28040 Madrid, Spain*)

L. Vázquez
(*Dept. Matemática Aplicada, Facultad de Informática Universidad Complutense de Madrid, 28040 Madrid, Spain; Centro de Astrobiología, CSIC/INTA, 28850 Torrejón de Ardoz, Madrid, Spain*)

Abstract

We construct a family of dynamical systems whose evolution converges to the eigenvectors of a general square matrix, not necessarily symmetric. We analyze the convergence of those systems and perform numerical tests. Some examples and comparisons with the power methods are presented.

Mathematics subject classification: 65F15, 65P99
Key words: Smallest real eigenvalue, Iterative method

1. Introduction

In a previous work [1, 2], a method was proposed to solve systems of linear equations $A\vec{x} = \vec{b}$, by means of considering a dissipative mechanical system associated to the matrix A. This mechanical system evolves under Newton's Second Law towards the solution of the linear system. A numerical simulation was proposed then to calculate the solution in an iterative procedure.

Following a similar point of view, in this paper we construct dynamical systems that have as critical points the eigenvectors of a real square matrix and that evolve towards an eigenvector. In section 2 we present the dynamical systems and their basic properties. In section 3 a numerical scheme is proposed to simulate the evolution and some examples and applications are presented. The main conclusions are summarized in Section 4. Finally, we present the proof of the results in an Appendix.

2. The Dynamical Systems

Let us consider the dynamical system

$$\dot{\vec{x}} = -\frac{A}{\|\vec{x}\|^p}\,\vec{x} + \frac{\vec{x}^{\mathrm{T}} A\vec{x}}{\|\vec{x}\|^{p+2}}\,\vec{x} \tag{1}$$

where $p \in \mathbb{R}$, $\vec{x} \in \mathbb{R}^q$ for some $q \in \mathbb{N}$ and A is a real, $q \times q$ matrix. The norm $\|\ \|$ is the euclidean vectorial norm. If A is symmetric and $p = 2$, the system is equivalent to

$$\dot{\vec{x}} = -\vec{\nabla} U(\vec{x}), \quad U(\vec{x}) = \frac{1}{2}\,\frac{\vec{x}^{\mathrm{T}} A\vec{x}}{\|\vec{x}\|^2}, \tag{2}$$

but we are not assuming any restriction on A.

This system has the following basic properties:

* Received October 16, 2003; final revised February 25, 2005.

1. The fixed points are the eigenvectors of A (and conversely).
2. Conservation law:
$$\frac{\mathrm{d}\|\vec{x}\|^2}{\mathrm{d}t} = \vec{0}. \tag{3}$$
3. Only if A es symmetric and $p = 2$, dissipation law:
$$\frac{\mathrm{d}}{\mathrm{d}t}\left[U(\vec{x})\right] = -\,\dot{\vec{x}}^{\,2}. \tag{4}$$

They are established in a straightforward way: in the first case the equivalence is clear. The conservation and the dissipation laws are obtained taking scalar product with $\vec{x}$ in (**??**) and $\dot{\vec{x}}$ in (2), respectively.

The conservation law (3) supposes that given an initial data $\vec{x}_0$, the corresponding solution lies for all times on the sphere $\|\vec{x}(t)\| = \|\vec{x}_0\|$. By Chillingworth's Theorem (see, for instance, Theorem 1.0.3 in [3]), we know that the solution always exists and is unique provided $\vec{x}_0 \neq \vec{0}$.

By direct calculation, the jacobian $J_p(\vec{x})$ of the dynamical system (**??**) at a given vector $\vec{x}$ is:
$$J_p(\vec{x}) = \frac{1}{\|\vec{x}\|^p}\left(\Big[A - r(\vec{x})I\Big]\Big[pP(\vec{x}) - I\Big] + P(\vec{x})\left[A + A^{\mathrm{T}} - 2r(\vec{x})I\right]\right). \tag{5}$$

Here I stands for the $q \times q$ identity matrix, $P(\vec{x})$ is the orthogonal projector on $\mathrm{span}\{\vec{x}\}$ and $r(\vec{x})$ is the Rayleigh quotient at $\vec{x}$:
$$P(\vec{x}) = \frac{\vec{x}\,\vec{x}^{+}}{\|\vec{x}\|^2}, \qquad r(\vec{x}) = \frac{\vec{x}^{+}A\vec{x}}{\|\vec{x}\|^2},$$
where the superscript $^+$ denotes the transposed, complex conjugate (or in case the vector is real, just the transposed).

Let us consider now the linear stability of the critical points. Let $\vec{u}$ be an eigenvector of A associated to the eigenvalue λ. From Lemma 1 in the Appendix, we have that $J_p(\vec{u})$ is a singular matrix and that the very eigenvector $\vec{u}$ belongs to the kernel:
$$J_p(\vec{u})\vec{u} = \frac{-1}{\|\vec{u}\|^p}\Big[I - P(\vec{u})\Big][A - \lambda I]\vec{u} = \vec{0}. \tag{6}$$

Thus, we see that the jacobian at any critical point has at least one eigenvector with zero real part. This means that in principle nothing can be said on the stability of the critical points from the study of the linear part. If we consider a symmetric A, the conservation law would allow us to conclude that the eigensubspace associated to the smallest eigenvalue is asymptotically stable. In the general case, we have to consider that given an initial data $\vec{x}_0$, the evolution is confined to the surface of the sphere $\|\vec{x}(t)\| = \|\vec{x}_0\|$, thus in order to check linear stability we must restrict ourselves to this manifold. The normal direction to the surface at $\vec{u}$ is given precisely by $\vec{u}$, thus we need to know the local behaviour around $\vec{u}$ in the orthogonal directions to $\vec{u}$. To do this, we compute all the eigenvalues of the jacobian, using the following result:

Theorem 1. *Let $\vec{u}$ be an eigenvector of A associated to the eigenvalue λ. The spectrum of A and that of $J_p(\vec{u})$ are related in the following way:*

1. *eigenvalue λ for A corresponds to eigenvalue 0 for $J_p(\vec{u})$*
2. *eigenvector $\vec{w}$ associated to μ for A corresponds to*
$$\textit{eigenvector} \quad \Big[I - P(\vec{u})\Big]\vec{w} \quad \textit{with eigenvalue} \quad \frac{\lambda - \mu}{\|\vec{u}\|^p} \quad \textit{for } J_p(\vec{u})$$
and this includes the case where $\mu = \lambda$, the case of complex eigenvalues and eigenvectors, as well as the case of generalized eigenvectors.

3. *If the eigensubspace associated to λ has algebraic multiplicity m_{a} and geometric multiplicity m_{g} such that $m_{\mathrm{g}} < m_{\mathrm{a}}$, then the eigensubspace associated to 0 for $J_p(\vec{u})$ has algebraic multiplicity equal to m_{a} and geometric multiplicity equal to either m_{g} or $m_{\mathrm{g}} + 1$. In that later case, a generalized eigenvector of A gives rise to a proper (i.e. not generalized) eigenvector of $J_p(\vec{u})$.*

4. *If $\mu \neq \lambda$, and the eigensubspace associated to μ has algebraic multiplicity m_{a} and geometric multiplicity m_{g}, then the eigensubspace associated to*

$$\frac{\lambda - \mu}{\|\vec{u}\|^p}$$

for $J_p(\vec{u})$ has algebraic multiplicity equal to m_{a} and geometric multiplicity equal to m_{g}.

We prove this in the Appendix. From here we have the following stability result:

Theorem 2. *Let A be such that the eigenvalue with smallest real part, $\lambda_{\min}$, is unique and real. Let be m_{a} and m_{g}, respectively, the algebraic and geometric multiplicites of $\lambda_{\min}$. Let $U_{\min}$ be the set of eigenvectors associated to $\lambda_{\min}$ and $\bar{U}_{\min}$ the set of generalized eigenvectors (in case $m_{\mathrm{a}} > m_{\mathrm{g}}$). Then $U_{\min}$ is a limit set for the system, and:*

- *Solutions inside $U_{\min}$ are fixed points of the system.*
- *The components of the solutions $\vec{x}(t)$ outside $\bar{U}_{\min}$ decay towards an eigenvector $\vec{u} \in U_{\min}$ with asymptotic behaviour:*

$$\|\vec{x}(t) - \vec{u}\| \sim \exp\left[\frac{\lambda_{\min} - \mathcal{R}e(\lambda')}{\|\vec{u}\|^p}\, t\right]$$

where λ' is the eigenvalue with real part nearest to $\lambda_{\min}$.

- *The components of the solutions $\vec{x}(t)$ inside $\bar{U}_{\min}$ decay towards $\vec{u}$ which is the eigenvector such that $\vec{u} = (A - \lambda I)\,\vec{w}_1$, in the notation of Lemma 4 of the Appendix, with asymptotic behaviour:*

$$\|\vec{x}(t) - \vec{u}\| \sim \frac{m_{\mathrm{a}} - m_{\mathrm{g}}}{t}.$$

See also the Appendix for the proof.

We see that the convergence is much faster in the case $m_{\mathrm{a}} = m_{\mathrm{g}}$, when $\lambda_{\min}$ has no generalized eigenvectors. It must be noted that if $\vec{x}_0$ has not a component in $U_{\min}$, the dynamical system cannot reach that eigensubspace. This corresponds to initial values outside $\bar{U}_{\min}$. In that case, we should expect convergence towards an eigenvector associated to the next smaller eigenvalue of A (provided it is real etc.). In fact, when performing numerical simulations, one would expect that small errors may give a component on $U_{\min}$ that gets enhanced, and the numerical solution approaches eventually $U_{\min}$.

Finally, we have a convergence result in the case where $\lambda_{\min}$ is not unique, for instance in the sense that A has also eigenvalues of the form $\lambda_{\min} \pm \mathrm{i}\alpha$ ($\alpha \neq 0$):

Theorem 3. *Let A be real and such that there are several complex eigenvalues with smallest real part, say $\mathcal{R}e_{\min}$. Let U be the direct sum of all the eigensubspaces associated to those eigenvalues. Then U is a limit set for the system and $r(\vec{x})$ converges to $\mathcal{R}e_{\min}$.*

(See the Appendix for the proof.) In this case, there is not convergence towards an eigenvector, but the Rayleigh quotient converges to the real part of the eigenvalues. Checking sepparately the behaviour of $\vec{x}(t)$ and of $r\Big(\vec{x}(t)\Big)$, it is thus possible to identify the case when $\lambda_{\min}$ is not unique.

3. Numerical Simulations

The idea is now to simulate the dynamical system, starting with some initial data $\vec{x}_0$. We chose, for instance, a simple numerical scheme of the form:

$$\frac{\vec{x}_{n+1}-\vec{x}_n}{\tau} = -\frac{A}{\|\vec{x}_n\|^p}\,\vec{x}_n + \frac{\vec{x}_n\cdot A\vec{x}_n}{\|\vec{x}_n\|^{p+2}}\,\vec{x}_n \tag{7}$$

$$\iff \vec{x}_{n+1} = \vec{x}_n - \tau\frac{A}{\|\vec{x}_n\|^p}\,\vec{x}_n + \tau\frac{\vec{x}_n\cdot A\vec{x}_n}{\|\vec{x}_n\|^{p+2}}\,\vec{x}_n. \tag{8}$$

This has the advantage of being explicit. If we take scalar product with $\vec{x}_n$, we get the discrete conservation law

$$\vec{x}_{n+1}\cdot\vec{x}_n = \|\vec{x}_n\|^2 \tag{9}$$

which is the discrete analog to the conservation of the norm. With this numerical scheme the norm is not exactly preserved, but if we consider $\vec{x}_{n+1} = \vec{x}_n + \vec{\delta}_n$, we have $\vec{\delta}_n$ orthogonal to $\vec{x}_n$.

We can understand the scheme as a fixed point iterative method. The specific scheme we have chosen, has the eigenvectors of A as fixed points. So we can consider the numerical scheme as a method in its own and not an approximation. Obviously both methods, discrete and continuous, are related as we will see in what follows.

If the numerical errors are small, we may suppose that the simulations will reproduce the convergence towards an eigenvector and the minimal eigenvalue of A. On the other hand we can study the convergence of the fixed point iteration: the jacobian of the iteration given by (8) is $I + \tau J_p(\vec{x})$ (where J_p is the jacobian of the dynamical system) and its eigenvalues μ satisfy

$$|I + \tau J_p(\vec{x}) - \mu I| = 0 \iff \left| J_p(\vec{x}) - \frac{\mu-1}{\tau}\,I \right| = 0$$

which means that if we denote by γ the eigenvalues of J_p, we have

$$\mu = 1 + \gamma\tau. \tag{10}$$

We can thus ensure the convergence of the numerical scheme if $\max|\mu| < 1$. Near the eigenvector $\vec{u}$ of A, to which the dynamical system converges, this supposes

$$0 < \frac{\tau}{\|\vec{u}\|^p} < \frac{2}{\lambda' - \lambda_{\min}} \tag{11}$$

where λ' is the eigenvalue of A nearest to $\lambda_{\min}$ (and bigger).

This means that the choice of τ is relevant to the convergence and also to the rate of convergence. In fact, numerical simulations show that for a given problem there is an optimal range of values of τ that minimize the number of iterations required to obtain the solution with a given precision. That range depends in general on p but also on the choice of the initial vector, both direction and norm, although the norm that minimizes the number of iterations is usually close to 1. The case $p = 0$ is different in the sense that the number of iterations do not depend on the norm of the initial vector, but only on its direction. On the other hand, if for a given problem we fix the direction of the initial vector, but consider two different choices of p and of the initial norm, the values of τ that optimize employ the same number of iterations in both cases. We may also perform a rescaling on τ defining the new time step as $\tau/\|\vec{u}\|^p$.

For all this, we have chosen to fix $p = 0$ in all our calculations. This leaves us with two objects to consider instead of three: τ and the direction of the initial vector. Of course we do not have a priori indications of which vectors are better suited: if would amount to know beforehand what are the eigenvectors of A. As happens for the dynamical system, if the initial vector has not a component on $U_{\min}$, the iterative method will not converge to that eigensubspace, but to some other eigenvectors, unless numerical errors modify the situation. So finaly we just have

one parameter to consider and that is the value of τ, keeping in mind that we do not need τ to be smaller than 1, since the discrete method can be considered exact and thus without a truncation error.

It is easily seen that the value of τ such that

$$\frac{\tau}{\|\vec{u}\|^p} = \frac{1}{\lambda' - \lambda_{\min}} \tag{12}$$

is optimal in the sense that the component of the solution that belongs to eigenvectors of λ' decays very fast. In that case the next eigenvalue, say λ'', closer to $\lambda_{\min}$ gives the asymptotic behaviour. The optimal τ that minimizes the total number of iterations for a given precision is difficult to establish *a priori*, but it can be done minimizing the product of all eigenvalues of the matrix $I + \tau J_p(\vec{u})$ where $\vec{u}$ is the eigenvector associated to $\lambda_{\min}$ that is the limit of the solution $\vec{x}(t)$. For instance (12) corresponds to minimizing the eigenvalue that comes from λ', minimizing that one and the next gives

$$\frac{\tau}{\|\vec{u}\|^p} = \frac{(\lambda' - \lambda_{\min}) + (\lambda'' - \lambda_{\min})}{2(\lambda' - \lambda_{\min})(\lambda'' - \lambda_{\min})} \tag{13}$$

and so on. There is not much point to this, since these values can only be deduced *a posteriori*. In practice, one should try a value of τ not too small: the discrete time is $t_n = \tau n$ which means that small vallues of τ imply bigger values of the number of iterations n.

Let us present details of the method with some simple examples.

3.1 Some examples in dimension 3

3.1.1 Choice of parameters and rate of convergence

Let be

$$A = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \qquad \vec{x}_0 = \frac{1}{\sqrt{3}}(1,1,1)^{\mathrm{T}}. \tag{14}$$

We have chosen the initial data such that it has the same component on the three eigensubspaces. We have fixed the relative precision of the solution to be less than 10^{-10} in λ. We can compute the values τ_1 and τ_2 according to (12) and (13), respectively. In this case ($\lambda_{\min} = 1$, $\lambda' = 4$, $\lambda'' = 6$), they are:

$$\tau_1 = \frac{1}{3}, \quad \tau_2 = \frac{4}{15}. \tag{15}$$

In figure 1 we compare the convergence for these values of τ. We have chosen different initial vectors (normalized to 1):

$$\vec{v}_1 = (1,1,1)^{\mathrm{T}}, \quad \vec{v}_2 = (1,0,1)^{\mathrm{T}}, \quad \vec{v}_3 = (1,1,0)^{\mathrm{T}}. \tag{16}$$

The first one has components along all three eigensubspaces, the second one has components only along the eigenvectors associated to $\lambda_{\min}$ and λ', and the third one only along the eigenvectors associated to $\lambda_{\min}$ and λ''. We see that the decay with $\tau = \tau_1$ is governed by the asymptotic behaviour of the eigenvectors associated to λ'' (first and fourth curves) and that the eigenvectors associated to λ' decay superlinearly to zero (second curve). Finally τ_2 (third curve) is the optimal choice in the general case when the initial data has components along all three eigenspaces.

In this example, we may wish to compute all three eigenvalues and the corresponding eigenvectors. The mimimum is computed using the method. The maximum can be obtained using the method on the matrix $-A$, in which case we have to change the sign of the eigenvalue. As for the intermediate value, we can obtain it using the method on A but with an initial data with no components on $U_{\min}$. In this case, and once the eigenvector associated to $\lambda_{\min}$ is known, it

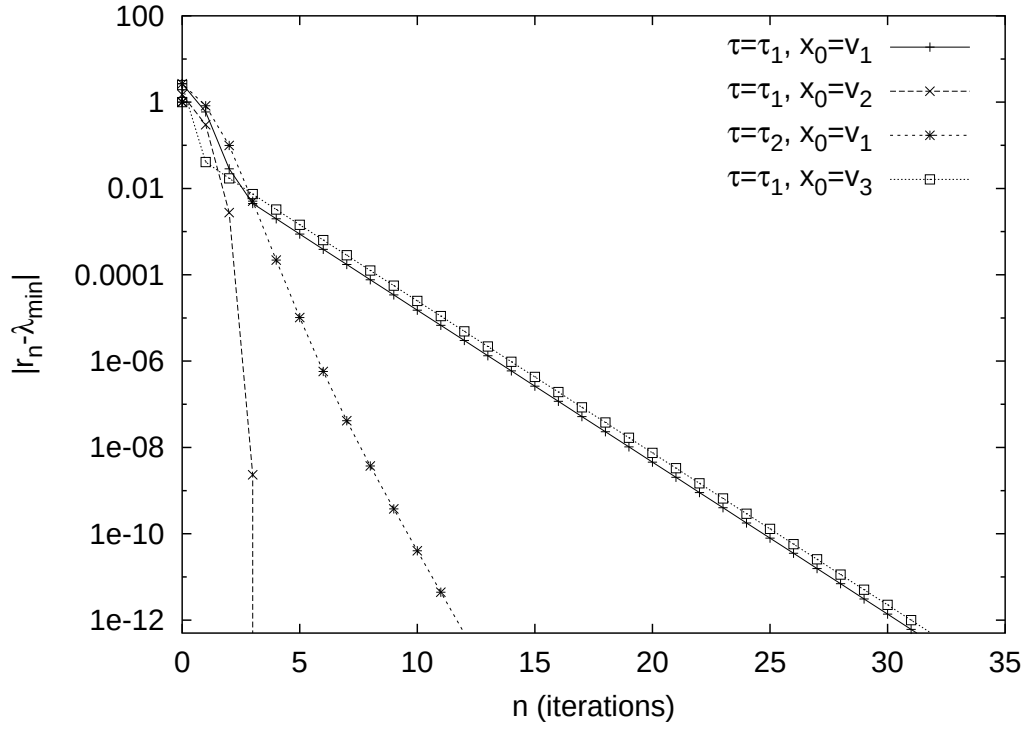

Figure 1: $|r(\vec{x}_n) - \lambda_{\min}|$ versus number of iterations n for matrix A with different values of τ and of the initial vector $\vec{x}_0$ according to (15) and (16). The vertical segment in the second curve means that at the following iteration step (number 4) the value is exactly $\lambda_{\min}$.

is very simple. In a more general situation, it is possible to compute all eigenvalues and their eigenvectors: the details will be presented elsewhere.

3.1.2 Real eigenvectors

Let us compare now the cases of diagonalizable versus non-diagonalizable: we will consider four cases, given by matrices

$$A_1 = A = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \quad A_2 = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$A_3 = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}; \quad A_4 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

The first case corresponds to a diagonalizable problem with $\dim U_{\min} = 1$. In the second one $\dim U_{\min} = 2$, with algebraic and geometric dimensions equal to 2. The third case is nondiagonalizable with algebraic dimension 2 and geometric dimension 1. Finally, the fourth case has algebraic dimension 3 and geometric dimension 1. In all cases $\lambda_{\min} = 1$. We have represented in figure 2 the two dimensional projection of trajectories on a semi-sphere.

In case 1, we have plotted the projection on the z-plane of the attracting basin of eigenvector $\vec{u} = (0, 0, 1)^{\mathrm{T}}$, which corresponds to vectors with positive z component, and trajectories of solutions from different initial values, all of them with unit norm. The arrows give the indication of movement along the solutions as time increases. The eigenvalues of the jacobian $J_0(\vec{u})$ are -2 with eigenvector $\vec{v}_1 = (0, 1, 0)^{\mathrm{T}}$ and -5 with eigenvector $\vec{v}_2 = (1, 0, 0)^{\mathrm{T}}$. As we see, near the

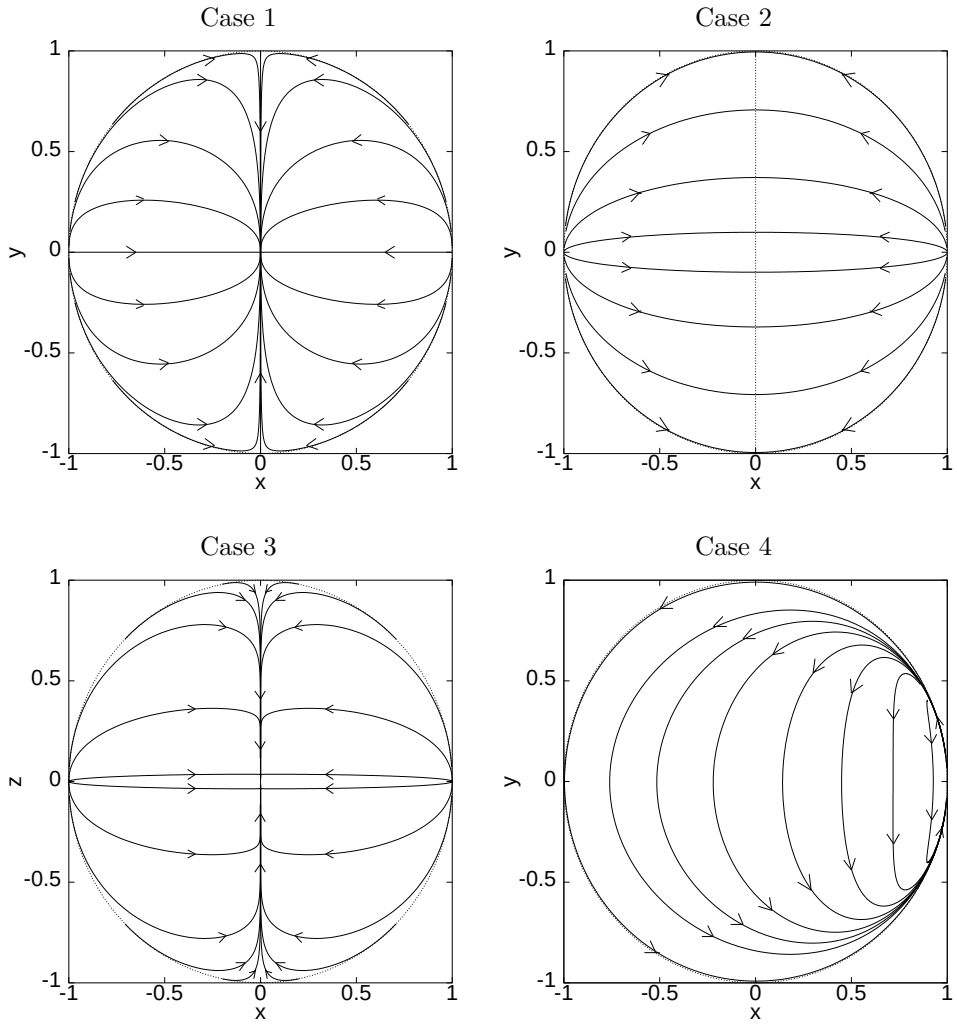

Figure 2: Projection of solutions in the cases 1 to 4: matrices with real eigenvalues

equilibrium point (that coresponds to the origin of the plot) the y component of the solutions decays faster than the x component, which agrees with the linear approximation. The picture is similar to that of a node in a planar system.

In case 2, $U_{\min} = \text{span}\{\vec{u}_1, \vec{u}_2\}$ with $\vec{u}_1 = (0,0,1)^{\mathrm{T}}$, $\vec{u}_2 = (0,1,0)^{\mathrm{T}}$, and all the points of the sphere with $x = 0$ correspond to eigenvectors of $\lambda_{\min}$. It is represented in the plot by a dotted line. Solutions tend towards an eigenvector, in fact following a geodesic on the sphere.

In case 3, $\vec{u} = (0,1,0)^{\mathrm{T}}$, the other eigenvector (both of A and of $J_0(\vec{u})$) being $\vec{v} = (1,0,0)^{\mathrm{T}}$. The third direction corresponds to a generalized eigenvector of $\lambda_{\min}$: $\vec{w} = (0,0,1)^{\mathrm{T}}$. We have plotted the projections on the y-plane. Simulations show that the x component decays rapidly and, as we see in the plot, the trajectories approach the generalized eigensubspace of $\lambda_{\min}$, that is $\text{span}\{\vec{u}, \vec{w}\}$, and eventually they reach the equilibrium point $\vec{u}$. The general picture is similar to that of case 1.

Case 4 is clearly different. Here we have only one eigenvector: $\vec{u} = (1,0,0)^{\mathrm{T}}$, and two gener-

alized eigenvectors $\vec{w}_1 = (0,1,0)^{\mathrm{T}}$, $\vec{w}_2 = (0,0,1)^{\mathrm{T}}$, such that $A\vec{w}_1 = \vec{u} + \vec{w}_1$ and $A\vec{w}_2 = \vec{w}_1 + \vec{w}_2$. We have projected the trajectories with positive z component on the z-plane. The behaviour is similar to that of a parabolic sector in a planar system. As we see, near $\vec{u}$ the direction along $\vec{w}_1$ is unstable and that along $-\vec{w}_1$ stable, such that eventually all trajectories approach the equilibrium point. The behaviour on the other semisphere is similar and approaches $-\vec{u}$.

3.1.3 Complex eigenvectors

Although our hypothesis is that $\lambda_{\min}$ is real, we present in case 5 an example where the (generalized) eigenvectors of a second eigenvalue are complex. We also consider case 6 where $\lambda_{\min}$ is unique but imaginary and there is another real eigenvalue with bigger real part. Finally we consider case 7 where $\lambda_{\min}$ is not unique, but there are two eigenvalues with same real part, one real and the other imaginary. In this two last cases, the convergence is not guaranteed since we do not fulfill the fundamental hypothesis of $\lambda_{\min}$ being unique and real. The matrices we are considering in these examples are:

$$A_5 = \begin{pmatrix} 6 & -2 & 0 \\ 2 & 6 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \quad A_6 = \begin{pmatrix} 1 & -2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}; \quad A_7 = \begin{pmatrix} 1 & -2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The corresponding plots are in figure 3

In case 5, $\lambda_{\min} = 1$ associated to $(0,0,1)^{\mathrm{T}}$ and we have two complex conjugate eigenvalues: $6 \pm \mathrm{i}2$ associated to $(1,0,0)^{\mathrm{T}}$ and$(0,1,0)^{\mathrm{T}}$. We have plotted the projection of trajectories with $z > 0$ on the z-plane. The eigenvector is an asymptotically stable focus.

In case 6, we have that the eigenvalues with minimal real part are $1 \pm \mathrm{i}2$, associated to $\vec{u}_1 = (1,0,0)^{\mathrm{T}}$ and $\vec{u}_2 = (0,1,0)^{\mathrm{T}}$. Besides, there is a real eigenvalue, 3, associated to $(0,0,1)^{\mathrm{T}}$. This last eigenvector behaves as an unstable focus and the trajectories tend to span$\{\vec{u}_1, \vec{u}_2\}$, and describe a circular motion. Although there is no convergence towards any vector, if we compute $r(\vec{x})$, we see that it converges towards the real part of the complex eigenvalues.

Finally, in case 7, all eigenvalues have the same real part. The complex eigenvalues are $1 \pm \mathrm{i}2$, associated to $\vec{u}_1 = (1,0,0)^{\mathrm{T}}$ and $\vec{u}_2 = (0,1,0)^{\mathrm{T}}$, and the real eigenvalue is 1, associated to $(0,0,1)^{\mathrm{T}}$. This eigenvector behaves as a center and the trajectories describe a circular motion. As in the previous case, there is no convergence towards any vector, but if we compute $r(\vec{x})$, we see that it is always equal to 1.

3.2 Examples in dimension 5: comparison with the Power Methods

We will now compare the performance of the method and that of similar iterative methods such as the Direct Power Method (DPM) and the Inverse Power Method with Seed (IPMS)[4].

We define different matrices with a specific spectrum and use them as a test. Let us consider matrices

$$M_1 = \begin{pmatrix} -6 & 6 & -4 & 4 & 1 \\ -23 & 21 & -14 & 14 & 5 \\ -23 & 20 & -14 & 15 & 6 \\ 22 & -18 & 13 & -12 & -5 \\ -68 & 56 & -42 & 42 & 17 \end{pmatrix}, \; M_2 = \begin{pmatrix} -7 & 8 & -5 & 6 & 2 \\ -24 & 23 & -15 & 16 & 6 \\ -23 & 20 & -14 & 15 & 6 \\ 13 & -10 & 8 & -8 & -4 \\ -46 & 37 & -30 & 33 & 15 \end{pmatrix}.$$

These matrices have canonical Jordan forms Λ_i, such that $P\Lambda_i P^{-1} = M_i$ with:

$$\Lambda_1 = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}, \quad \Lambda_2 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix},$$

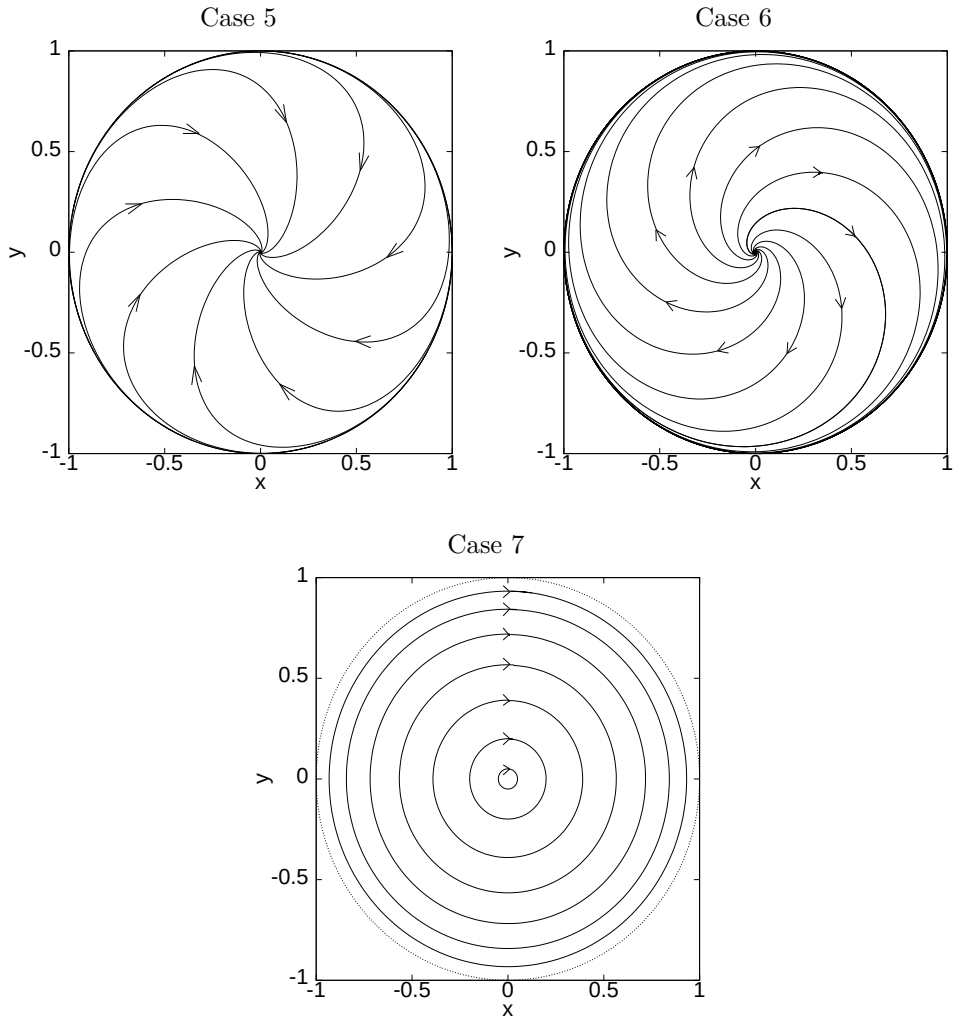

Figure 3: Projection of solutions in the cases 5 to 7: matrices with complex eigenvalues

$$P = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & -2 & 0 & 1 \\ -1 & 1 & 0 & 1 & -1 \\ 3 & 0 & -1 & -2 & 2 \end{pmatrix}.$$

Matrix P has been chosen assigning arbitrarily the values 0, ± 1, ± 2 to its elements (an element has been changed to 3 so that the resulting matrices M_i have all the elements with integer values, for sake of simplicity). In this way, the eigenvectors are not mutually orthogonal.

In figure 4 we compare the results of simulating the dynamical system (DS) with different values of τ and of the DPM for matrix M_1 in order to obtain $\lambda_{\max} \equiv 3$. In order to have DS converging towards the maximum eigenvalue, we have taken $-M_1$. In all computations the initial vector is $(0, 0, 1, 1, 1)^{\mathrm{T}}$, normalized. As we see, the number of iterations is similar for both methods, provided we chose a reasonable value of τ for the DS.

In figure 5 we compare the results of simulating the dynamical system (DS) with different

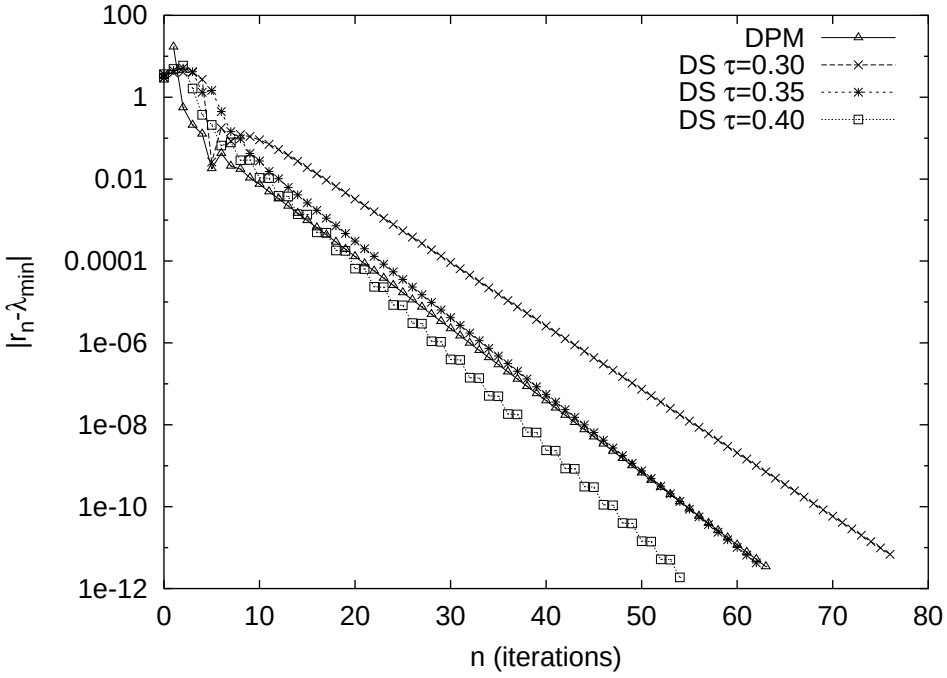

Figure 4: Comparison of $|r(\vec{x}_n) - \lambda_{\max}|$ versus number of iterations n for matrix M_1 with the Power Method and the dynamical system with three different values of τ.

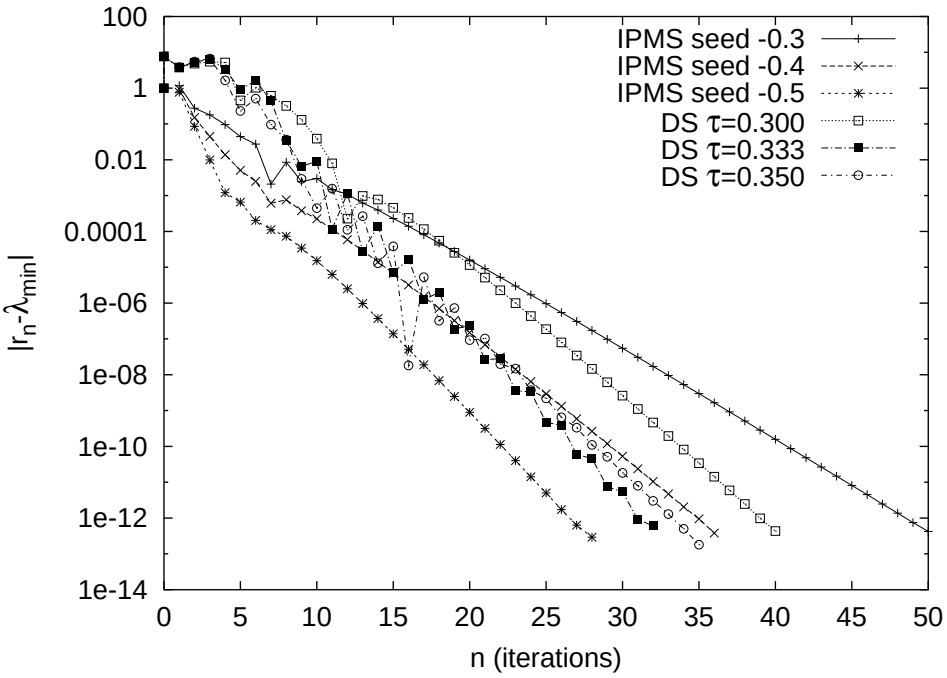

Figure 5: Comparison of $|r(\vec{x}_n) - \lambda_{\min}|$ versus number of iterations n for matrix M_1 with the Inverse Power Method with three seeds and the dynamical system with three different values of τ.

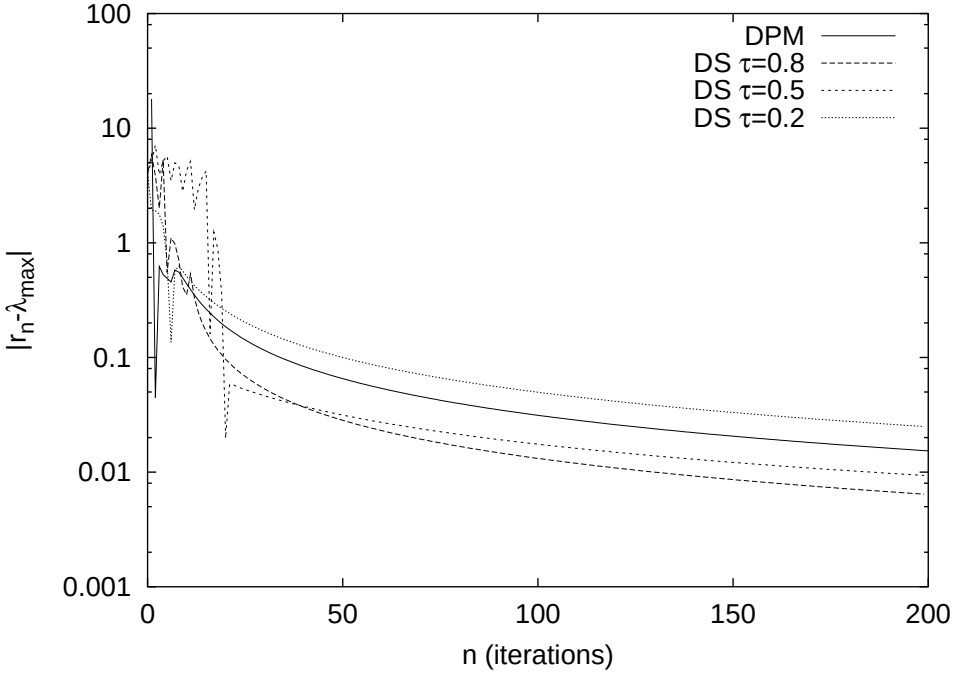

Figure 6: Comparison of $|r(\vec{x}_n) - \lambda_{\max}|$ versus number of iterations n for matrix M_2 with the Power Method and the dynamical system with three different values of τ. Due to the big number of iterations, the data are represented by lines rather that points

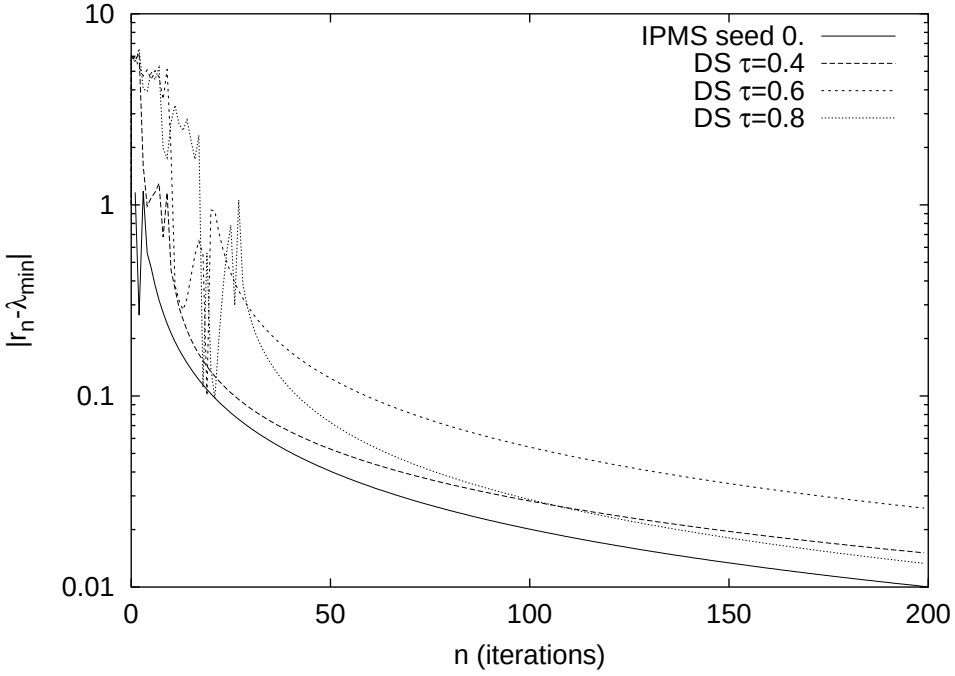

Figure 7: Comparison of $|r(\vec{x}_n) - \lambda_{\min}|$ versus number of iterations n for matrix M_2 with the Inverse Power Method and the dynamical system with three different values of τ.

values of τ and of the IDPS for matrix M_1 in order to obtain $\lambda_{\min} = -1$. We cannot start the IPMS with seed 0, since 1 is also an eigenvalue and the method do not converge. As we see, the number of iterations is similar for both methods, provided we chose a reasonable value of τ for the DS and of the seed for the IPMS.

In figure 6 we compare the results of simulating the dynamical system (DS) with different values of τ and of the DPM for matrix M_2 in order to obtain $\lambda_{\max} \equiv 3$. In order to have DS converging towards the maximum eigenvalue, we have taken $-M_2$. In all computations the initial vector is $(0, 0, 1, 1, 1)^{\mathrm{T}}$, normalized. As we see, the precition of both methods is similar for a number of iterations fixed. We are in the case of generalized eigenvectors where the decay towards the eigenvector is not exponential.

In figure 7 we compare the results of simulating the dynamical system (DS) with different values of τ and of the IDPS for matrix M_2 in order to obtain $\lambda_{\min} = 1$. We start the IPMS with seed 0. The precition of both methods is similar for a number of iterations fixed.

4. Conclusions

We have presented a new family of methods to obtain the eigenvectors of the given matrix A. The different kinds of behaviour allow in pratice to know whether the associated eigenvalue is real and unique or if there are others with the same real part. We also can deduce whether there are generalized eigenvectors.

From the numerical point of view, its performance is similar to that of the Power Methods, with the difference of the conditions of applicability: this method can be used in situations where the Power Methods do not converge because there are two eigenvalues with same absolute value or equally distant from the seed.

Acknowledgements. The authors thank the support of the Centro de Ciências Matemáticas (Funchal, Portugal) under grant POCTI/MAT/40706/2001 of the FCT (Portugal). We also thank the partial support received from the projects: NATO Science Programme Linkage Grant PST.CLG.978177 and grant BFM2002-02359 of the Ministerio de Ciencia y Tecnología (Spain). Finally, we are indebted to the referee for his valuable comments and corrections.

Appendix: proof of the theorems

In this Appendix, we give proofs of the three theorems. Each section corresponding to one of the results.

A.1 Proof of Theorem 1

In order to build the proof, we start with some preliminary Lemmas.

Lemma 1. *Let $\vec{u}$ be an eigenvector of A associated to eigenvalue λ. We have*

$$J_p(\vec{u}) = \frac{-1}{\|\vec{u}\|^p}\Big[I - P(\vec{u})\Big][A - \lambda I]$$

Proof. we just consider the hypothesis and perform the computations. The result is obtained from (5) using that $AP(\vec{u}) = \lambda P(\vec{u}) = P(\vec{u})A^{\mathrm{T}}$.

Lemma 2. *Let $\vec{u}$ be an eigenvector of A associated to eigenvalue λ. Let μ be an arbitray value. We have*

$$\left(J_p(\vec{u}) - \frac{\lambda - \mu}{\|\vec{u}\|^p} I\right)\Big[I - P(\vec{u})\Big] = \frac{-1}{\|\vec{u}\|^p}\Big[I - P(\vec{u})\Big][A - \mu I].$$

Proof. The result is obtained by direct calculations and Lemma 1.

Lemma 3. *Let $\vec{u}$ and $\vec{v}$ be mutually linearly independent eigenvectors of A, associated respectively to eigenvalues λ and μ. Then*

$$\Big[I - P(\vec{u})\Big]\vec{v} \quad \textit{is an eigenvector of } J_p(\vec{u}) \textit{ with eigenvalue } \quad \frac{\lambda - \mu}{\|\vec{u}\|^p}.$$

Furthermore: that eigenvector of $J_p(\vec{u})$ is orthogonal to $\vec{u}$.

Proof. Using Lemma 2 we have

$$\left(J_p(\vec{u}) - \frac{\lambda-\mu}{\|\vec{u}\|^p} I\right)\Big[I - P(\vec{u})\Big]\vec{v} = \frac{-1}{\|\vec{u}\|^p}\Big[I - P(\vec{u})\Big][A-\mu I]\vec{v}.$$

This is null since $[A-\mu I]\vec{v} = \vec{0}$. On the other hand, $\vec{u}$ and $\vec{v}$ being linearly independent, we have $\Big[I - P(\vec{u})\Big]\vec{v} \neq \vec{0}$. Thus we prove the existence of the eigenvector and eigenvalue of $J_p(\vec{u})$.

Besides, since $I - P(\vec{u})$ is the orthogonal projector on $\text{span}\{\vec{x}\}^{\perp}$, the eigenvector is by construction orthogonal to $\vec{u}$.

What happens if A is not diagonalizable is dealt with in the next lemmas:

Lemma 4. *Let $\vec{u}$ be an eigenvector of A associated to λ. We suppose that λ has an algebraic multiplicity m_{a} and a geometric multiplicity $m_{\text{g}} < m_{\text{a}}$. Let $\vec{w}_k$ be a generalized eigenvector associated to λ such that:*

$$\begin{cases} (A-\lambda I)\vec{w}_k \neq \vec{0}, \\ \cdots \\ (A-\lambda I)^k\,\vec{w}_k \neq \vec{0}, \\ (A-\lambda I)^{k+1}\,\vec{w}_k = \vec{0}, \end{cases} \qquad \textit{with} \quad 1 \leq k \leq m_{\text{g}}.$$

We have two possibilities:

a) If $(A-\lambda I)\vec{w}_1 \in \text{span}\{\vec{u}\}$ then

$$\begin{cases} J_p(\vec{u})\vec{w}_k \neq \vec{0}, \\ \cdots \\ J_p^{k-1}(\vec{u})\vec{w} \neq \vec{0}, \\ J_p^{k}(\vec{u})\vec{w} = \vec{0}. \end{cases}$$

b) If $(A-\lambda I)\vec{w}_1 \notin \text{span}\{\vec{u}\}$ then

$$\begin{cases} J_p(\vec{u})\vec{w}_k \neq \vec{0}, \\ \cdots \\ J_p^{k}(\vec{u})\vec{w} \neq \vec{0}, \\ J_p^{k+1}(\vec{u})\vec{w} = \vec{0}. \end{cases}$$

Proof. From Lemma 1 we know that for any vector $\vec{w}$

$$J_p(\vec{u})\vec{w} = \frac{-1}{\|\vec{u}\|^p}\Big[I - P(\vec{u})\Big][A-\lambda I]\vec{w}$$

and thus, to the power ℓ,

$$J_p^{\ell}(\vec{u})\vec{w} = \frac{(-1)^{\ell}}{\|\vec{u}\|^{\ell p}}\left(\Big[I - P(\vec{u})\Big][A-\lambda I]\right)^{\ell}\vec{w} = \frac{(-1)^{\ell}}{\|\vec{u}\|^{\ell p}}\Big[I - P(\vec{u})\Big][A-\lambda I]^{\ell}\vec{w}$$

since:

$$[A-\lambda I]\Big[I - P(\vec{u})\Big] = [A-\lambda I].$$

Applying this to $\vec{w} = \vec{w}_k$ with $\ell = 1, \ldots, k+1$ proves case b). On the other hand, if

$$(A-\lambda I)\vec{w}_1 \in \text{span}\{\vec{u}\}$$

we have that

$$\Big[I - P(\vec{u})\Big][A-\lambda I]\vec{w}_1 = \vec{0}$$

and then

$$J_p(\vec{u})\vec{w}_1 = \vec{0},$$

which is the difference needed to prove case a).

Case a) always happens if $m_{\text{g}} = m_{\text{a}} - 1$. It also happens by chance if, among all the proper eigenvectors associated to λ, we choose our vector $\vec{u}$ belonging to $\text{span}\{(A-\lambda I)\vec{w}_1\}$.

Now we study the case when the generalized eigenvectors are associated to an eigenvalue μ of A such that $\mu \neq \lambda$:

Lemma 5. *Let $\vec{u}$ be an eigenvector of A associated to λ. Let be $\mu \neq \lambda$ and $\vec{w}_k$ a generalized eigenvector associated to μ, such that:*

$$\begin{cases} (A-\mu I)\vec{w}_k \neq \vec{0}, \\ \cdots \\ (A-\mu I)^k \vec{w}_k \neq \vec{0}, \\ (A-\mu I)^{k+1} \vec{w}_k = \vec{0}. \end{cases}$$

We have:

$$\begin{cases} \left(J_p(\vec{u}) - \frac{\lambda-\mu}{\|\vec{u}\|^p} I\right)\Big[I - P(\vec{u})\Big]\vec{w}_k \neq \vec{0}, \\ \cdots \\ \left(J_p(\vec{u}) - \frac{\lambda-\mu}{\|\vec{u}\|^p} I\right)^k \Big[I - P(\vec{u})\Big]\vec{w}_k \neq \vec{0}, \\ \left(J_p(\vec{u}) - \frac{\lambda-\mu}{\|\vec{u}\|^p} I\right)^{k+1} \Big[I - P(\vec{u})\Big]\vec{w}_k = \vec{0}. \end{cases}$$

Proof. We use Lemma 2, repetedly, on any vector $\vec{w}$:

$$\left(J_p(\vec{u}) - \frac{\lambda-\mu}{\|\vec{u}\|^p} I\right)^\ell \Big[I - P(\vec{u})\Big]\vec{w} = \frac{(-1)^\ell}{\|\vec{u}\|^{\ell p}}\Big[I - P(\vec{u})\Big][A-\mu I]^\ell \vec{w}$$

and apply this to $\vec{w} = \vec{w}_k$, $\ell = 1, \ldots, k+1$. Now, contrarily to what happened in the case a) of Lemma 4, $[A-\mu I]^k \vec{w}$ never belongs to $\mathrm{span}\{\vec{u}\}$ since $\lambda \neq \mu$, and only when $\ell = k+1$ can this be null.

We are now in a position to prove Theorem 1: the first point is proven using Lemma 3 with $\lambda = \mu$. The second point is proven using Lemma 3 in the case of proper eigenvectors, and Lemmas 4 and 5 in the case of generalized eigenvectors. It is easily seen that the complex case is also fulfilled. The third point is given directly by Lemma 4. Finally, the fourth point is proven by Lemma 5, which completes the proof.

A.2 Proof of Theorem 2

As for the proof of Theorem 2, if the solution belongs to $U_{\min}$ it is an eigenvector and thus a fixed point. Even if we have several eigenvectors linearly independent (say $\{\vec{u}_i\}_{i=1}^q$, where $q = m_{\mathrm{g}}$) associated to the eigenvalue, any solution of the form $\vec{x}(t) = \sum_{i=1}^q a_i(t)\vec{u}_i$ is a constant: substituting in the equation we have

$$\begin{aligned} &\sum_{i=1}^q \dot{a}_i(t)\vec{u}_i + \frac{\lambda_{\min}}{\|\vec{x}\|^p}\sum_{i=1}^q a_i(t)\vec{u}_i - \frac{\lambda_{\min}}{\|\vec{x}\|^p}\sum_{i=1}^q a_i(t)\vec{u}_i = \vec{0} \\ \iff \qquad &\sum_{i=1}^q \dot{a}_i(t)\vec{u}_i = \vec{0} \iff \forall i,\ \dot{a}_i = 0. \end{aligned} \tag{17}$$

Thus, any point of $U_{\min}$ is a fixed point.

We suppose thus that a general solution has components outside $U_{\min}$. Those can be of two types: outside $\bar{U}_{\min}$ and inside $\bar{U}_{\min} - U_{\min}$.

Let us start considering the first case. It is the only possibility, for instance, if $m_{\mathrm{a}} = m_{\mathrm{g}}$. From point 2 in theorem 1, we have that all eigenvectors of $J_p(\vec{u})$ but $\vec{u}$ (either true or generalized) belong to $\mathrm{span}\{\vec{u}\}^\perp$. Thus the local behaviour around $\vec{u}$ is given by those other eigenvectors. Let be $\vec{u}_{\min} \in U_{\min}$ such that it lies on the surface of the sphere $\|\vec{x}\| = \|\vec{x}_0\|$. If the geometric multiplicity m_{g} of $\lambda_{\min}$ is 1 and if we force the solutions to lie on that same sphere, it is clear that $\vec{u}_{\min}$ is asymptotically stable. We only have to show that something similar is also true if $m_{\mathrm{g}} > 1$. Let us denote by $\bar{U}_{\min}$ the set of generalized eigenvectors associated to $\lambda_{\min}$. We have that on one hand $J_p(\vec{u}_{\min})$ has no eigenvalues with positive real part, and that all other

eigenvectors associated to $\lambda_{\min}$ are eigenvectors of $J_p(\vec{u}_{\min})$ with zero real part. On the other hand, for any other eigenvector $\vec{v}$ of A (either true or generalized), any eigenvector of A that belongs to $U_{\min}$ give rise to eigenvectors of $J_p(\vec{v})$ with negative real part. Thus any trajectory outside $\bar{U}_{\min}$ decays towards $\bar{U}_{\min}$, and its behaviour is governed by the smallest eigenvalue of $J_p(\vec{u})$ which is

$$\frac{\lambda_{\min} - \lambda'}{\|\vec{u}\|^p},$$

hence the asymptotic behaviour.

We study now the second case, and consider trajectories evolving inside $\bar{U}_{\min}$. We will see that they tend to some true eigenvector of $\lambda_{\min}$. In order to simplify the computations we choose $\|\vec{x}(0)\| = 1$, but this is not necessary. We use a notation similar to that of Lemma 4: let be $\mathcal{B} = \left\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_q\right\}$ an orthonormal basis of $U_{\min}$ and $\mathcal{B}' = \left\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_K\right\}$ a basis of $\bar{U}_{\min} - U_{\min}$ (the subspace of generalized but not true eigenvectors of $\lambda_{\min}$) such that:

$$\begin{cases} (A - \lambda_{\min} I)\vec{w}_j = \vec{w}_{j-1},\ j = 2, 3, \ldots, K, \\ (A - \lambda_{\min} I)\vec{w}_1 = \vec{u}_q. \end{cases} \tag{18}$$

($q = m_{\mathrm{g}}$ and $K = m_{\mathrm{a}} - m_{\mathrm{g}}$, but we have chosen this in order to simplify the notation).

Let us now consider a more general solution of the form

$$\vec{x}(t) = \sum_{i=1}^{q} a_i(t)\vec{u}_i + \sum_{j=1}^{K} b_j(t)\vec{w}_j. \tag{19}$$

Using (18) we have

$$A\vec{x}(t) = \lambda_{\min}\vec{x}(t) + b_1(t)\vec{u}_q + \sum_{j=1}^{K-1} b_{j+1}(t)\vec{w}_j \tag{20}$$

and from here

$$\vec{x}(t)^{\mathrm{T}} A\vec{x}(t) = \lambda_{\min}\|\vec{x}(t)\|^2 + b_1(t)\vec{x}(t)^{\mathrm{T}}\vec{u}_q + \sum_{j=1}^{K-1} b_{j+1}(t)\vec{x}(t)^{\mathrm{T}}\vec{w}_j. \tag{21}$$

Let us, for the time being, call

$$h(t) \equiv b_1(t)\vec{x}(t)^{\mathrm{T}}\vec{u}_q + \sum_{j=1}^{K-1} b_{j+1}(t)\vec{x}(t)^{\mathrm{T}}\vec{w}_j. \tag{22}$$

Substituting in the dynamical system we have

$$\begin{aligned} \dot{\vec{x}}(t) &= -\lambda_{\min}\vec{x}(t) - b_1(t)\vec{u}_q - \sum_{j=1}^{K-1} b_{j+1}(t)\vec{w}_j + \lambda_{\min}\vec{x}(t) + h(t)\vec{x} \\ &= -b_1(t)\vec{u}_q - \sum_{j=1}^{K-1} b_{j+1}(t)\vec{w}_j + \lambda_{\min}\vec{x}(t) + h(t)\vec{x}, \end{aligned} \tag{23}$$

while by direct and differentiation of the solution:

$$\dot{\vec{x}}(t) = \sum_{i=1}^{q} \dot{a}_i(t)\vec{u}_i + \sum_{j=1}^{K} \dot{b}_j(t)\vec{w}_j. \tag{24}$$

Putting all this together, and using the fact that all the $\vec{u}$'s and $\vec{w}$'s are linearly independent, we get the following set of equations for the coefficients:

$$\begin{cases} \dot{a}_i = h(t)a_i,\ i = 1, \ldots,\ q-1; \\ \dot{a}_q = h(t)a_q - b_1; \\ \dot{b}_j = h(t)b_j - b_{j+1},\ j = 1, \ldots,\ K-1; \\ \dot{b}_K = h(t)b_K. \end{cases} \tag{25}$$

The solution of this system is of the form

$$\begin{cases} a_i(t) = \pm\alpha_i \Big[\mathcal{R}(t,2K)\Big]^{-2},\ i = 1, \ldots,\ q-1; \\ a_q(t) = \pm\mathcal{P}(t,K)\Big[\mathcal{R}(t,2K)\Big]^{-2}; \\ b_j(t) = \pm\mathcal{P}(t,K-j)\Big[\mathcal{R}(t,2K)\Big]^{-2},\ j = 1, \ldots,\ K-1; \\ b_K(t) = \pm\Big[\mathcal{R}(t,2K)\Big]^{-2}; \end{cases} \tag{26}$$

where $\alpha_i = a_i(0)/b_K(0)$, and $\mathcal{P}(t,K)$ and $\mathcal{R}(t,2K)$ are polynomial with leading terms of the form, respectively,

$$\frac{(-1)^K t^K}{K!} \qquad \text{and} \qquad \frac{t^{2K}}{(K!)^2}. \tag{27}$$

A few words about this solution: first of all, none of these coefficients can become singular, since $\vec{x}$ exists for all times and is bounded. Secondly, none of them can be equal to zero, unless they were zero at the initial time, and in that case remain zero for all times. This also means that the sign ($\pm$) we should consider is just that of $b_K(0)$. Finally, due to the form of the leading terms (27), we see that all coefficients but $a_q(t)$, tend to zero as t goes to infinity, and that $a_q(t)$ tends to ± 1, as should be expected. The coefficient that goes to zero more slowly is $b_1(t)$, its asymptotic behaviour is

$$|b_1(t)| \sim \frac{1}{(K-1)!} t^{K-1} \left(\frac{t^{2K}}{(K!)^2}\right)^{-2} = \frac{K}{t} = \frac{m_{\mathrm{a}} - m_{\mathrm{g}}}{t}. \tag{28}$$

Thus we see that any trajectory inside $\bar{U}_{\min}$ decays towards the proper eigenvector $\vec{u}_q$.

A.3 Proof of Theorem 3

We finally present the proof of Theorem 3: in this case there is not in principle convergence to a vector, but as in the previous theorem, all solutions decay towards the eigenvectors of A associated to the eigenvalues with smallest real part. This means the space U. On the other hand, it is easy to check that for any vector belonging to U, the Rayleigh quotient is just $\mathcal{R}e_{\min}$.

References

[1] L. Vázquez and J.L. Vázquez-Poletti, A new Approach to Solve Systems of Linear Equations, *Journal of Computational Mathematics*, **19**:4 (2001) 445-448.

[2] L. Vázquez and S. Jiménez, Analysis of a Mechanical Solver for Linear Systems of Equations, *Journal of Computational Mathematics*, **19**:1 (2001), 9-14.

[3] J. Guckenheimer and P. Holmes, "Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vectors Fields", Springer Verlag, New York, 1983.

[4] F. Chatelin, "Eigenvalues of Matrices", John Wiley & Sons, New York, 1995.