CONVERGENCE OF THE LANCZOS ALGORITHM IN PRESENCE OF CLOSE EIGENVALUES: A NUMERICAL STUDY

GÉRARD MEURANT*

Abstract. This paper is concerned with the behaviour of the Lanczos algorithm in presence of close eigenvalues. We numerically study in great details an example with two close eigenvalues. We show that what happens is very similar to the situation that occurs due to the development of rounding errors when using the Lanczos algorithm in finite precision arithmetic.

1. Introduction. The Lanczos algorithm is one of the most well-known method for computing a few eigenvalues of a symmetric matrix A. In this paper we are interested in the behaviour of the Lanczos algorithm in presence of close eigenvalues. Contrary to the convergence of the Ritz values towards well separated eigenvalues, the convergence to a cluster of close eigenvalues is rather peculiar. For a while it seems that one Ritz value converges to an average value of the eigenvalues belonging to the cluster. This phenomenon, called "misconvergence", is described in B.N. Parlett's book [7]. This problem has also been studied in a paper by A. van der Sluis and H.A. van der Vorst [9]. Here we will consider the problem from another perspective by analyzing an example in great details. Some insights in this phenomenon can be obtained by using the same tools as when studying the Lanczos algorithm in finite precision arithmetic; for a summary on this problem, see [2], [3].

Section 2 briefly recalls some details on the Lanczos algorithm. The example we will study with two close eigenvalues is described in section 3. Numerical results on this example are given in section 4. An explanation of the observed results is provided in section 5 using different tools than in [9]. Section 6 briefly considers finite precision arithmetic. Section 7 provides some numerical experiments with three close eigenvalues. Finally, we give some conclusions.

2. The Lanczos algorithm. Let A be a symmetric matrix of order n. The Lanczos algorithm is the following. Starting from vectors $v^0 = 0$ and $\tilde{v}^1 = v$ for $k = 1, 2, \ldots, m$ it computes

$$\eta_{k-1} = \|\tilde{v}^k\|,$$

$$v^k = \frac{\tilde{v}^k}{\eta_{k-1}},$$

$$\alpha_k = (v^k, Av^k) = (v^k)^T A v^k,$$

$$\tilde{v}^{k+1} = Av^k - \alpha_k v^k - \eta_{k-1} v^{k-1}.$$

This algorithm constructs an orthonormal basis of the Krylov space $\mathcal{K}_m(A, v)$ which is spanned by the vectors $A^j v$ for $j = 0, \ldots, m-1$. The matrix relation for the matrix V_k whose columns are the basis vectors v^j, $j = 1, \ldots, k$ can be written more compactly as

$$AV_k = V_k T_k + \eta_k v^{k+1} (e^k)^T,$$

*(gerard.meurant@gmail.com) started in Prague April 2009, revised May 2010

where e^k is the kth column of the identity matrix of order k. Moreover,

$$V_k^T V_k = I, \quad V_k^T A V_k = T_k,$$

where T_k is a tridiagonal matrix of order k whose diagonal elements are the $\alpha_i, i = 1, \ldots, k$ and the subdiagonal elements are $\eta_i, i = 1, \ldots, k-1$. We have also

$$\eta_{k-1} = (Av^k, v^{k-1}).$$

We will use the Lanczos algorithm with double reorthogonalization of the new basis vector at every iteration. At least for our example, this will give numerical results which are very close to exact arithmetic. Computations done in extended precision using the Matlab Symbolic Toolbox with this example confirm that the double precision results with reorthogonalization well represent exact arithmetic results.

Since we can suppose that $\eta_i > 0$, the tridiagonal matrix T_k has real and simple eigenvalues which we denote by $\theta_j^{(k)}$. They are known as the Ritz values that we order as

$$\theta_1^{(k)} < \theta_2^{(k)} < \cdots < \theta_k^{(k)}.$$

The Ritz values are the approximations of the eigenvalues of A given by the Lanczos algorithm. In exact arithmetic the Lanczos vectors v^k are given by a polynomial in A applied to the initial vector,

$$v^k = p_{1,k}(A)v^1, \quad p_{1,k}(\lambda) = (-1)^{k-1} \frac{\chi_{1,k-1}(\lambda)}{\eta_1 \cdots \eta_{k-1}},$$

where $\chi_{1,k}(\lambda)$ denotes the determinant of $T_k - \lambda I$. The polynomials $p_{1,k}$ of degree $k-1$ are known as the normalized Lanczos polynomials. Their roots are the Ritz values. Of course, they satisfy a three–term recurrence relation

$$\eta_k p_{k+1}(\lambda) = (\lambda - \alpha_k)p_k(\lambda) - \eta_{k-1}p_{k-1}(\lambda), \ k = 1, 2, \ldots$$

with initial conditions, $p_0 \equiv 0$, $p_1 \equiv 1$. These polynomials are orthogonal for the inner product defined by a Riemann-Stieltjes integral with a discrete measure which depends on the eigenvalues of A; see [2].

3. The example. First we start with a Strakoš matrix; see [8]. The matrix of order n is diagonal with eigenvalues

$$\lambda_i = \lambda_1 + \left(\frac{i-1}{n-1}\right)(\lambda_n - \lambda_1)\rho^{n-i}, \ i = 1, \ldots, n.$$

The parameters λ_1 and λ_n are respectively the smallest and largest eigenvalues. The parameter ρ controls the distribution of the eigenvalues. We will use $n = 30$, $\lambda_1 = 0.1$, $\lambda_n = 100$ and $\rho = 0.9$. Let us denote this matrix by S. Since this matrix is diagonal, the matrix of its eigenvectors is the identity matrix. Hence, the projections of the Lanczos vectors on the eigenvectors (which are the important variables to analyze the behavior of the algorithm) are the Lanczos vectors themselves. Starting from a vector v^1 having all its components equal to $1/\sqrt{n}$, the first eigenvalues to be approximated are the largest ones which are well separated.

The matrix A we use as our example is a modification of the matrix S. We change the next to last eigenvalue λ_{n-1} to

$$\lambda_{n-1} = (1 - gap)\lambda_n.$$

Therefore the positive real number *gap* is the relative gap between the two largest eigenvalues and $gap \cdot \lambda_n$ is the absolute gap. The value *gap* is to be chosen small such that we will have a cluster of two close eigenvalues well separated from the other ones.

4. Numerical experiments. As we said before the Lanczos algorithm is used with double reorthogonalization at every iteration. Let us first consider the approximation of the two largest eigenvalues of A. We take $gap = 10^{-10}$ which gives an absolute gap of $1.000000793283107\ 10^{-8}$. Figure 4.1 displays the (log_{10} of the) minimum distances of a Ritz value $\theta_j^{(k)}$ to λ_{30} (solid), λ_{29} (dashed) and $\lambda_M = (\lambda_{30}+\lambda_{29})/2$ (dot-dashed) as a function of the iteration number k. We see that a Ritz value is approaching the cluster in the first 10 iterations. In fact this Ritz value is the largest one $\theta_k^{(k)}$ as it can be seen on figure 4.2 since the distance of $\theta_{k-1}^{(k)}$ to the cluster is large until iteration 20. The largest Ritz value $\theta_k^{(k)}$ becomes larger than λ_{n-1} (that is, enters the cluster) at iteration 12 and it goes almost right away to the average value λ_M of λ_n and λ_{n-1}. As we can see in both figures the distance $|\lambda_n - \theta_k^{(k)}|$ stagnates from iterations 12 to 20. This is related to the fact that the distance of the Ritz value to λ_M is small and close to 10^{-14} at the minimum (see dashed curve in figure 4.1) for a few iterations (roughly between iterations 15 and 20). At iteration 20, $\theta_k^{(k)}$ starts moving again towards λ_n and almost reach the minimum distance at iteration 25. The distance $\lambda_{n-1} - \theta_{k-1}^{(k)}$ becomes small at iteration 19 and then it decreases fast as we can see in figure 4.2.

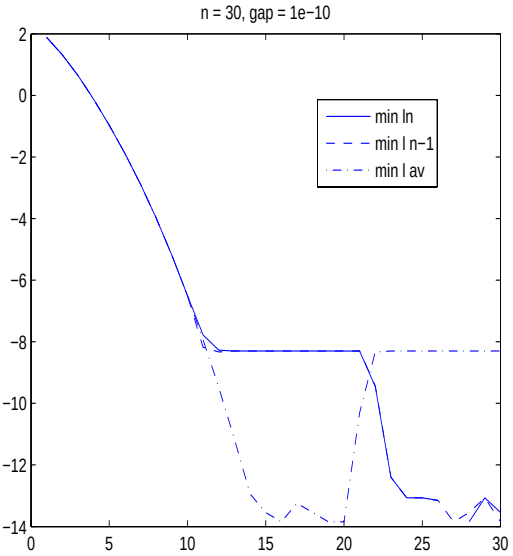

FIG. 4.1. *gap*=10^{-10}, $\log_{10}$ *of minimum distances of a Ritz value to* λ_{30} *(solid),* λ_{29} *(dashed) and* $(\lambda_{30} + \lambda_{29})/2$ *(dot-dashed)*

In fact the third largest eigenvalue $\lambda_{n-2} \approx 75.4383$ has been well approximated by a Ritz value since iteration 5. It is well approximated by $\theta_{k-1}^{(k)}$ until iteration 17 where the distance is $-3.06954\ 10^{-12}$ and then by $\theta_{k-2}^{(k)}$, the distance being $6.2243\ 10^{-12}$ at iteration 18. We see the appearance of a new Ritz value (in between λ_{n-2} and λ_{n-1})

at iteration 18 where the three largest Ritz values are

$$75.43837931033860,\ 85.85112980363365,\ 99.99999999500002$$

At iteration 17 the two largest Ritz values were

$$75.43837931034790,\ 99.99999999499994$$

At iteration 24 the two largest eigenvalues (even though they are quite close) are well approximated by the two largest Ritz values.

This phenomenon has been called "misconvergence" (see [7]) because it seems that for a while the largest Ritz value has converged to the average value λ_M. We will see that the level of stagnation of $\lambda_M - \theta_k^{(k)}$ depends on the size of the gap. When the gap is larger, the level of stagnation is higher and for a "large" gap, the Ritz value cannot be considered as converged but it stagnates for a while.

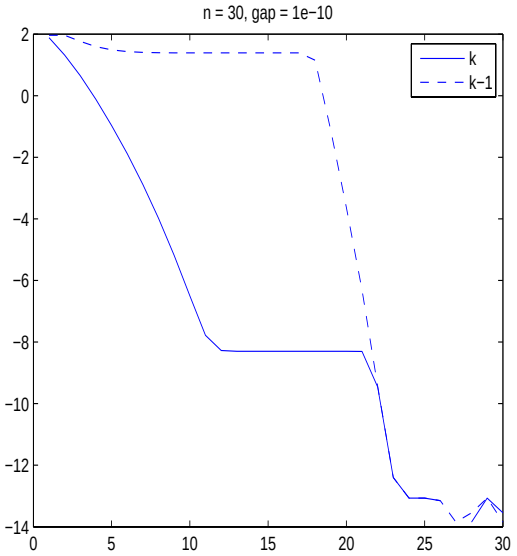

FIG. 4.2. *gap*=10^{-10}, $\log_{10}$ *of distances of Ritz value* $\theta_k^{(k)}$ *to* λ_{30} *(plain), and* $\theta_{k-1}^{(k)}$ *to* λ_{29} *(dashed)*

Normally when a Ritz value converges to an eigenvalue the projection of the Lanczos vector v^k on the corresponding eigenvector must "converge" to zero. This happens because (in our notations when A is diagonal) we have $v_i^k = p_{1,k}(\lambda_i)v_i^1$. If there is a root of $p_{1,k}$ (that is a Ritz value) which converges to λ_i then the corresponding components v_i^k must converge to zero. So considering the two largest eigenvalues, here in our diagonal example, we have to look at the components v_n^k and v_{n-1}^k. This is shown in figure 4.3 where we observe an interesting behavior. The absolute values of these two components are first decreasing until iteration 12, then they increase until iteration 18 for which the value of v_{30}^k is 0.6345647. Note that this is much larger than the initial value v_{30}^1 which was 0.1825741. After iteration 18 the two components decrease up to the end of the computation. In fact, as we can see in figure 4.4 between iterations 13 to 27 the two components have the same absolute values but opposite

signs. This continues until they are at the roundoff level. Both are small (less than 10^{-5}) after iteration 22.

The component v^k_{n-2} does not have the same behavior as we see in figure 4.5, even though there is a small perturbation at iteration 18. There is a change of sign of this component at iteration 18. The logarithms of all the elements of $|V_{30}|$ are displayed in figure 4.6. It turns out that the smallest eigenvalues of A are well approximated only at the last iteration. This is why we do not see a decrease of the corresponding components of the Lanczos vectors.

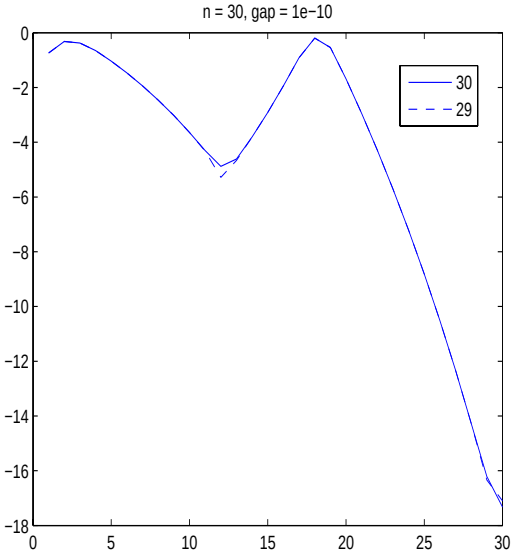

FIG. 4.3. *gap*=10^{-10}, $\log_{10}$ *of* $|v^k_{30}|$ *and* $|v^k_{29}|$

Let us now change the value of the gap between the two largest eigenvalues of A. The results are displayed in figure 4.7 for different values of the relative gap. Remember that the absolute gap is almost a hundred times larger. The iteration number at which $|v^k_{30}|$ starts to increase clearly depends on the size of the gap. Only when the gap is of the order of the machine precision do we not see this phenomenon of oscillation. Then the cluster just appears as one double eigenvalue up to machine precision. Figure 4.8 shows the (log_{10} of the) minimum distances of a Ritz value to λ_{30}, λ_{29} and $(\lambda_{30} + \lambda_{29})/2$ for $gap = 10^{-5}$ (which gives an absolute gap of the order of 10^{-3}). We see that the level of the stagnation of the distances to the last two eigenvalues and the accuracy with which the average is approximated for some time depend on the value of the gap. Moreover, as it can be expected, the two largest eigenvalues of A are well approximated sooner when the gap is larger. The same information is displayed in figure 4.9 for a small gap of 10^{-13}. The stagnation phase is longer (more than 10 iterations).

It is also interesting to look at a comparison with what is obtained with a matrix $\tilde{A}$ for which the two largest eigenvalues are replaced by the average value λ_M and the initial vector is the same as before. Remember that the Lanczos algorithm cannot "see" the double eigenvalue. Everything is just like if we have just one eigenvalue λ_M. The last components of the Lanczos vectors are shown in figure 4.10. We see

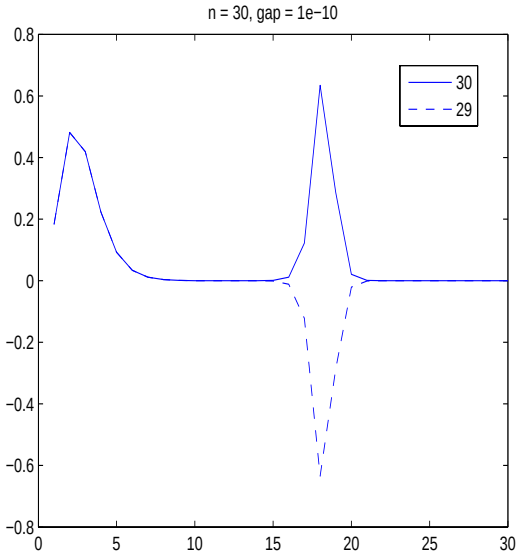

FIG. 4.4. *gap*=10^{-10}, v_{30}^k *and* v_{29}^k

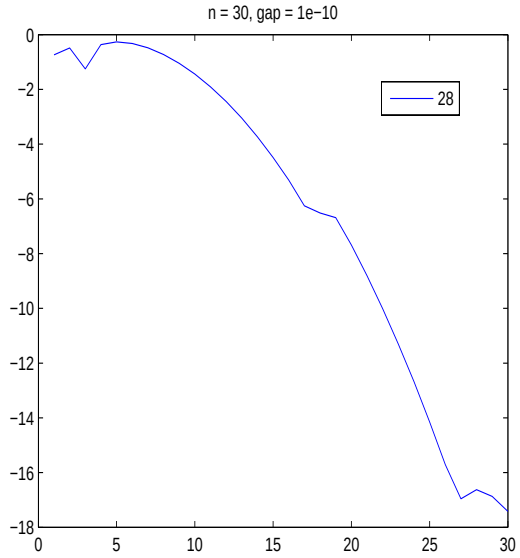

FIG. 4.5. *gap*=10^{-10}, $\log_{10}$ *of* $|v_{28}^k|$

that, up to the point where the curve for A increases, the last components of both computations are very close. In fact, the tridiagonal matrices that are generated are also close. Figure 4.11 shows the diagonal elements of T_k for both computations. They start to differ significantly at iteration 10. Figure 4.12 displays the values of η_k.

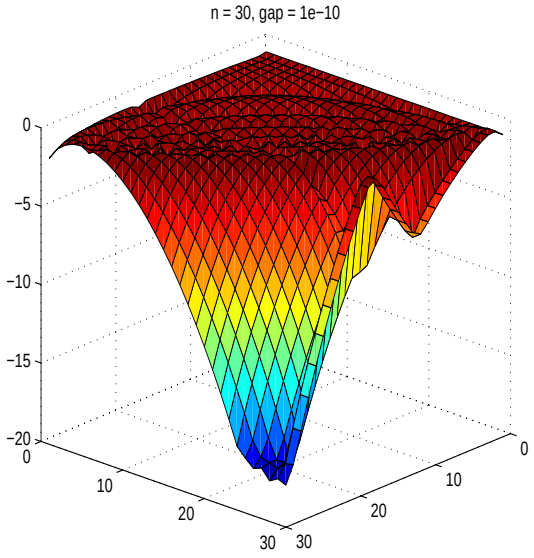

FIG. 4.6. *gap*=10^{-10}, $\log_{10}$ *of all elements of* $|V_{30}|$

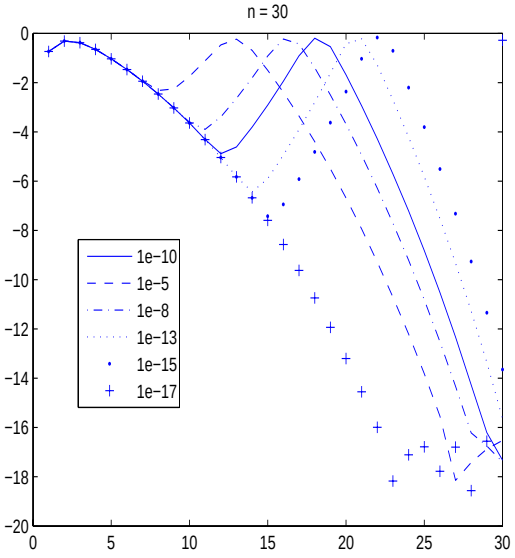

FIG. 4.7. $\log_{10}$ *of* $|v_{30}^k|$, *gap* $= 10^{-10}$ *(solid),* 10^{-5} *(dashed),* 10^{-8} *(dot-dashed),* 10^{-13} *(dotted),* 10^{-15} *(points),* 10^{-17} *(+)*

We have large peaks between iterations 16 to 20. Note that

$$\alpha_k = \sum_{i=1}^{n} \lambda_i (v_i^k)^2, \quad \eta_{k-1} = \sum_{i=1}^{n} \lambda_i v_i^k \, v_i^{k-1}.$$

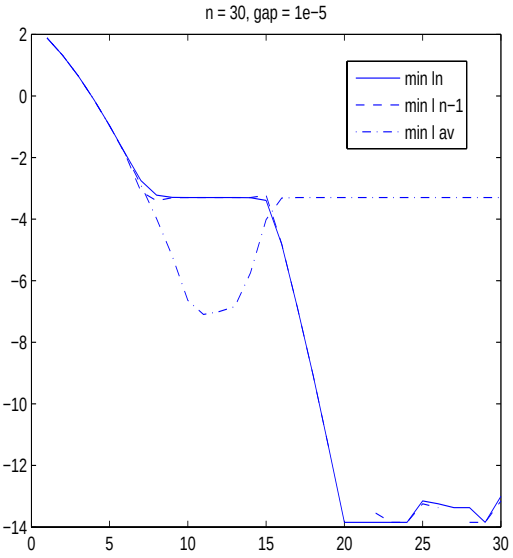

FIG. 4.8. *gap*=10^{-5}, $\log_{10}$ *of minimum distances of a Ritz value to* λ_{30} *(solid),* λ_{29} *(dashed) and* $(\lambda_{30} + \lambda_{29})/2$ *(dot-dashed)*

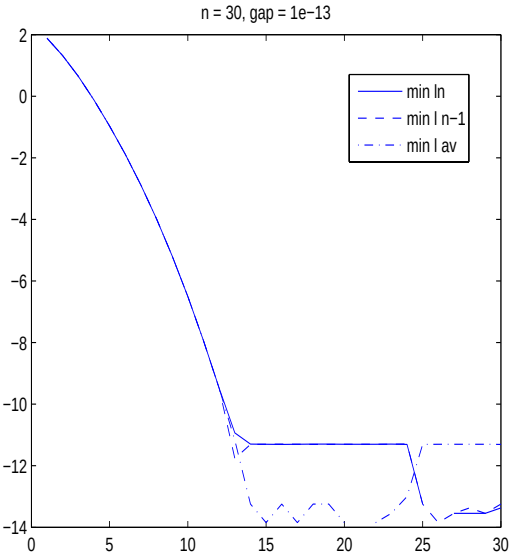

FIG. 4.9. *gap*=10^{-13}, $\log_{10}$ *of minimum distances of a Ritz value to* λ_{30} *(solid),* λ_{29} *(dashed) and* $(\lambda_{30} + \lambda_{29})/2$ *(dot-dashed)*

Therefore, differences in the components of the Lanczos vectors induce differences in the coefficients of the three-term recurrence which in turn give rise to differences in the next Lanczos vectors. However, after the peaks, the values of the two computations are not far away and the rate of decrease is almost the same.

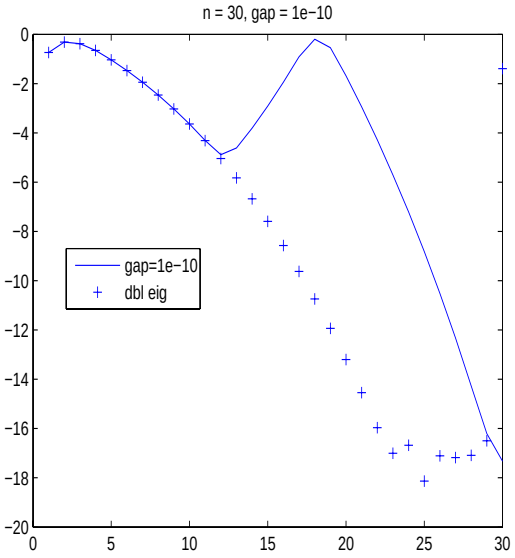

FIG. 4.10. *gap*=10^{-10}, $\log_{10}$ *of* $|v_{30}^k|$ *for* A *(solid) and* $\tilde{A}$ *(+)*

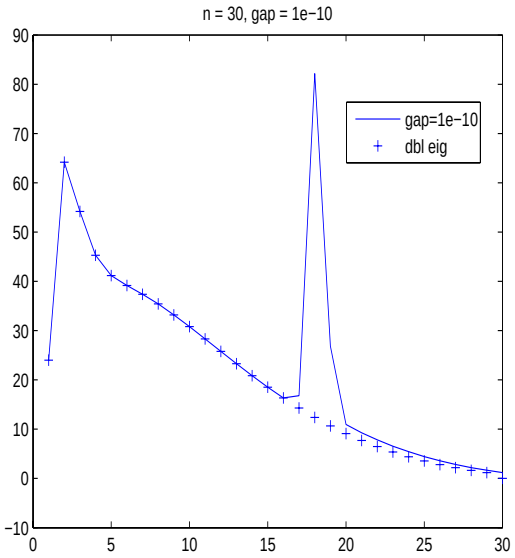

FIG. 4.11. *gap*=10^{-10}, α_k *for* A *(solid) and* $\tilde{A}$ *(+)*

5. A tentative of explanation. In this section we will provide an explanation for the experimental results of the previous section.

5.1. The perturbation problem. Since we assume exact arithmetic and we have a diagonal matrix, the components of the Lanczos vectors are given by a scalar

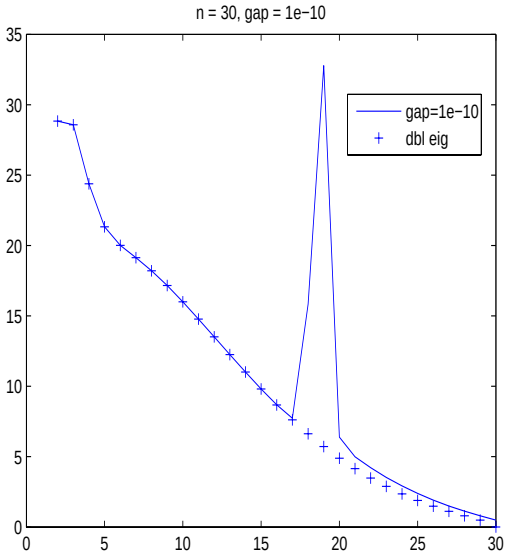

FIG. 4.12. *gap*=10^{-10}, η_k *for* A *(solid) and* $\tilde{A}$ *(+)*

three-term recurrence

$$\eta_k v_i^{k+1} = (\lambda_i - \alpha_k) v_i^k - \eta_{k-1} v_i^{k-1}, \ i = 1, \ldots, n \tag{5.1}$$

whose solution is given by the Lanczos polynomial

$$v_i^{k+1} = p_{1,k+1}(\lambda_i) v_i^1, \quad p_{1,k+1}(\lambda_i) = (-1)^k \frac{\chi_{1,k}(\lambda_i)}{\eta_1 \cdots \eta_k}.$$

The roots of the polynomial $p_{1,k+1}$ of degree k are the Ritz values $\theta_j^{(k)}$ (the eigenvalues of T_k and the roots of the characteristic polynomial $\chi_{1,k}$) at iteration k. Hence

$$p_{1,k+1}(\lambda_i) = \frac{(-1)^k}{\eta_1 \cdots \eta_k} \prod_{l=1}^{k} (\theta_l^{(k)} - \lambda_i). \tag{5.2}$$

Although we assume exact arithmetic, the situation in figures 4.3 and 4.7 reminds of what happens in finite precision arithmetic (without reorthogonalization), see the book [2], chapter 4, page 143. There we see an increase of the perturbation terms due to rounding errors when computing the three-term recurrence in double precision arithmetic. The growth of the perturbations happens when a Ritz value converges to an eigenvalue of A. This situation has been analyzed carefully by Chris Paige [4], [5] and [6]. See also a summary of his results in [2] and [3].

Where does the perturbation come from in our exact arithmetic problem? The last two equations in (5.1) can be written (somehow artificially) as

$$\eta_k v_{n-1}^{k+1} = (\lambda_M - \alpha_k) v_{n-1}^k - \eta_{k-1} v_{n-1}^{k-1} + (\lambda_{n-1} - \lambda_M) v_{n-1}^k, \tag{5.3}$$

$$\eta_k v_n^{k+1} = (\lambda_M - \alpha_k) v_n^k - \eta_{k-1} v_n^{k-1} + (\lambda_n - \lambda_M) v_n^k, \tag{5.4}$$

with $\lambda_M = (\lambda_{n-1} + \lambda_n)/2$ if we assume that $v_i^1 = 1/\sqrt{n}\ \forall i$. In a more general case

$$\lambda_M = \frac{\lambda_n (v_n^1)^2 + \lambda_{n-1}(v_{n-1}^1)^2}{(v_n^1)^2 + (v_{n-1}^1)^2}.$$

So what we obtain with equation (5.1) for $i = 1, \ldots, n-2$ and equations (5.3) and (5.4) is that everything seems to look like if we were computing the last two equations with perturbations (or in variable precision arithmetic) for a problem with a double eigenvalue at λ_M .

Let g be the absolute gap, $g = \lambda_n - \lambda_{n-1}$. The perturbation terms are $-\frac{g}{2} v_{n-1}^k$ in equation (5.3) and $\frac{g}{2} v_n^k$ in equation (5.4). Of course, they change at each iteration with the values of v_{n-1}^k and v_n^k. If the last components decrease for a while then the perturbation terms in equations (5.3) and (5.4) also decrease.

The matrix equation related to the perturbed system is

$$\tilde{A} V_k = V_k T_k + \eta_k v^{k+1} (e^k)^T + D V_k, \tag{5.5}$$

where $\tilde{A}$ is diagonal with $\lambda_1, \ldots, \lambda_{n-2}, \lambda_M, \lambda_M$ on the diagonal and D is diagonal with a zero diagonal except for the last two terms which are $-g/2$ and $g/2$.

The solution of perturbed three-term recurrences like (5.3) and (5.4) is known in terms of polynomials, see [2], theorem 4.3, page 150.

THEOREM 5.1. *Let $h_{n-1} = -g/2$ and $h_n = g/2$. The solution of recurrences (5.3) and (5.4) is*

$$v_i^{k+1} = p_{1,k+1}(\lambda_M) v_i^1 + h_i \sum_{l=1}^{k} p_{l+1,k+1}(\lambda_M) \frac{v_i^l}{\eta_l},\ i = n-1, n, \tag{5.6}$$

where

$$p_{j,k+1}(\lambda) = \frac{(-1)^{k-j+1}}{\eta_j \cdots \eta_k} \prod_{m=1}^{k-j+1} (\theta_m^{(j,k)} - \lambda), \tag{5.7}$$

where $\theta_m^{(j,k)}$ are the eigenvalues of the trailing submatrix $T_{j,k}$ which is obtained from T_k by discarding the $j-1$ first rows and columns.

The polynomial $p_{1,k}$ satisfies the unperturbed recurrence

$$\eta_k p_{1,k+1}(\lambda) = (\lambda - \alpha_k) p_{1,k}(\lambda) - \eta_{k-1} p_{1,k-1}(\lambda),\ k = 1, 2, \ldots \tag{5.8}$$

with $p_{1,0} \equiv 0$, $p_{1,1} \equiv 1$. The $p_{j,k+1}$ are known as the associated polynomials. They satisfy the following three-term recurrence for a given j,

$$\eta_k p_{j,k+1}(\lambda) = (\lambda - \alpha_k) p_{j,k}(\lambda) - \eta_{k-1} p_{j,k-1}(\lambda),\ k = j, j+1, \ldots \tag{5.9}$$

with initial conditions $p_{j,j-1} \equiv 0, \quad p_{j,j} \equiv 1$. Note that the perturbation terms in equation (5.6) depend on the previous v_i^l, $i = n$ or $n-1$.

Since the gap is small we could also have used a Taylor expansion of the polynomial $p_{1,k+1}$ at $\lambda = \lambda_M$ by writing

$$\begin{aligned} p_{1,k+1}(\lambda_n) &= p_{1,k+1}(\lambda_n - \lambda_M + \lambda_M) \\ &= p_{1,k+1}(\lambda_M + g) \\ &= p_{1,k+1}(\lambda_M) + g\, p'_{1,k+1}(\lambda_M) + \frac{g^2}{2} p''_{1,k+1}(\lambda_M) + \ldots . \end{aligned}$$

Note that since we have a polynomial of degree k, this expansion is finite. In fact we can identify the powers of g with the expression in equation (5.6) using the associated polynomials. The derivatives of $p_{1,k+1}$ satisfy also recurrence relations. However we will see that using only the first derivative is not enough to obtain the correct behavior of $p_{1,k}(\lambda_n)$.

Our claim is that the absolute value of the first term $|p_{1,k+1}(\lambda_M)|$ decreases (at least until iteration 17 in the example with a gap of 10^{-10}) and that the absolute values of the perturbation terms are increasing because the Ritz value $\theta_k^{(k)}$ converges to the cluster. At some point (iteration 13 in the example) the perturbation terms become larger than the first term $|p_{1,k+1}(\lambda_M)|$ and the (absolute value of the) components are increasing. However, $|v_i^k|$ has to be smaller than 1. So the increase has to be stopped and this implies the appearance of an eigenvalue close to λ_M for $T_{j,k}$, $j \geq 2$, through increases of α_k and η_k. This happens for $k = 18$. After this, some already existing perturbation terms are decreasing for a while and new ones are increasing as we will see.

The questions which are of interest are the following:

-1. What is the behavior of the polynomials $p_{1,k}$ and $p_{j,k}, j \geq 2$ as a function of k at $\lambda = \lambda_M$?
-2. Why is $\theta_k^{(k)}$ first "converging" to λ_M?
-3. What is the duration of the stagnation phase?
-4. When and why does $\theta_k^{(k)}$ move towards λ_n?

These questions were partially answered in [9].

5.2. Question 1. What is the behavior of the polynomials $p_{1,k}$ and $p_{j,k}$, $j \geq 2$ as a function of k for $\lambda = \lambda_M$?

This question was studied in chapter 4 of [2] (for roundoff perturbations) where it is proved that when $|p_{1,k}|$ decreases then $|p_{j,k}|$ decreases. Let us first compute the polynomials $p_{1,k}$ and $p_{j,k}$ at λ_M. The coefficients α_k and η_k are given by our Lanczos computation with double reorthogonalization. Since when computing the values of the polynomials using three-term recurrence we cannot use anything like orthogonalization we have to be careful of not introducing too large rounding errors. Therefore we computed everything in variable precision using the Matlab Symbolic Toolbox. The following computations for the polynomials are done with 32 decimal digits. We use a relative gap of 10^{-10}.

Figure 5.1 displays the polynomial $p_{1,k}$ (in fact the logarithm of $|p_{1,k}(\lambda_M)v_n^1|$) at λ_M as a function of k. The dashed curve is the last component of the Lanczos vectors. We see that the values at λ_M decrease until iteration 17. This happens because the largest Ritz value is approaching the cluster. The values at the beginning (up to iteration 11) are close to those of v_n^k because the perturbation terms are small. The values at λ_M decrease to a value of $6.88518\ 10^{-11}$, and they more or less stagnate until iteration 20. Finally they increase rapidly up to 10^8. This happens because the largest Ritz value (which is a root of the polynomial) moves away from λ_M. In fact, as we can see in figure 5.2 the values of the polynomial $p_{1,k}(\lambda_M)$ are positive until iteration 16 (that is when the largest Ritz value is still smaller than λ_M) and then negative when the absolute values increase. We have large negative values after iteration 24.

Let us now look at the first associated polynomials $p_{j,k}(\lambda_M)$ for j small. We have to multiply their values by $(g\, v_n^{j-1})/(2\eta_{j-1})$. This is shown in figure 5.3 for

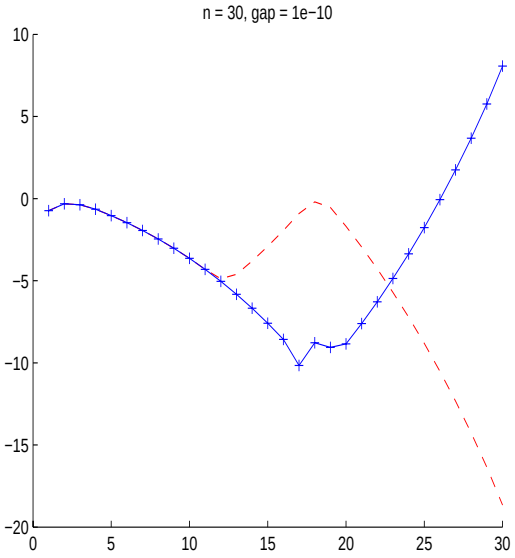

FIG. 5.1. *gap*=10^{-10}, $\log_{10}$ *of* $|p_{1,k}(\lambda_M)v_n^1|$ *(solid) and* $|v_n^k|$ *(dashed)*

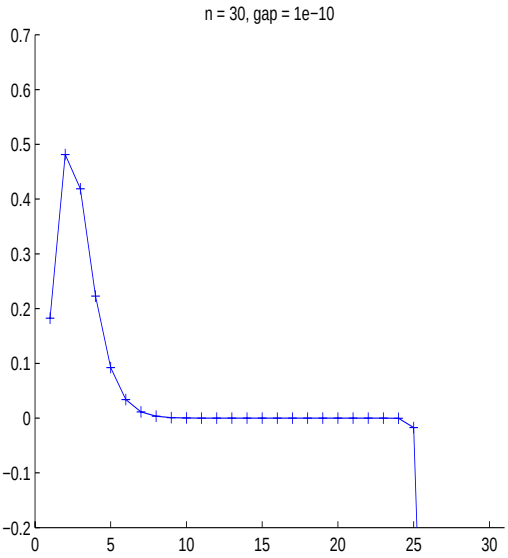

FIG. 5.2. *gap*=10^{-10}, $p_{1,k}(\lambda_M)v_n^1$

$j = 2, \ldots, 5$. The symbol + gives the sum

$$p_{1,k+1}(\lambda_M)v_n^1 + \frac{g}{2}\sum_{j=2}^{5} p_{j,k+1}(\lambda_M)\frac{v_n^{j-1}}{\eta_{j-1}}$$

The first associated polynomials increase from the beginning up to iteration 18 and

then decrease up to iteration 25 or 26 and finally they increase again. The sum nicely approximate the dashed curve which is $|v_n^k|$ until iteration 22. Most of the values of the associated polynomials $j = 2, \ldots, 5$ are positive but not large enough to compensate for the large negative values of $p_{1,k}(\lambda_M)$. We see that in log scale the increase of the perturbation terms is almost symmetric to the decrease of the first polynomial until iteration 16. This phenomenon was studied in [2].

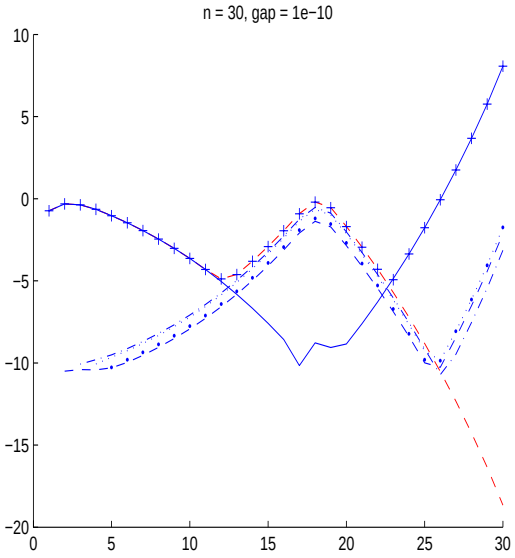

FIG. 5.3. *gap*=10^{-10}, $\log_{10}$ *of* $|p_{1,k}(\lambda_M)v_n^1|$ *(solid),* $|v_n^k|$ *(dashed),* $|g\, p_{j,k}(\lambda_M)v_n^{j-1}/(2\eta_{j-1})|$, $j = 2:5$ *(different symbols) and sum of the values (+)*

If the first perturbation terms do not explain the behavior of v_n^k, let us consider more terms. Figure 5.4 displays all the perturbation terms. Even though this figure is a little messy, we clearly see that the sum of all the terms (+) gives the value of v_n^k (dashed curve) as it should. There are some interesting features in this figure. In the bottom left we see that the starting values of the first perturbations terms decrease because they are given by values of v_n^k which is decreasing at this moment. After iteration 12 the perturbation terms are larger than the first term (polynomial $p_{1,k}$), therefore v_n^k is now increasing and consequently the starting values of the perturbation terms increase too.

At iteration 18 a new (large) Ritz value appears in T_k (in between λ_{n-2} and λ_{n-1}) as well as in all the trailing submatrices $T_{j,k}$, $j = 2, \ldots, 18$ as we will see. This happens because $|v_n^k|$ is now large and this induces a dramatic change of values for α_k and η_k. This stops the increase of the perturbation terms. However, the perturbations which start after iteration 18 have large initial values which give the fast increase that we see on the upper right of the figure. As we see in figure 5.5 these last perturbation terms are large and positive and their sum cancels out the large negative values of $p_{1,k}(\lambda_M)$. So we do need all the terms in the sum from 2 to 30 in equation (5.6) to obtain the correct behavior of the last component of the Lanczos vectors.

Does the value of the derivative $p'_{1,k}$ at λ_M explain the behavior of v_n^k? Figure 5.6 shows the values of the second term in the Taylor expansion $g\, p'_{j,k}(\lambda_M)v_n^1$ as the dashed

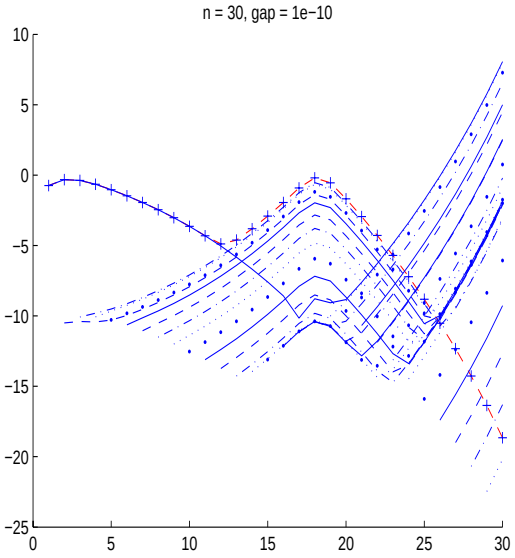

FIG. 5.4. *gap*=10^{-10}, $\log_{10}$ *of* $|p_{1,k}(\lambda_M)v_n^1|$ *(solid),* $|v_n^k|$ *(dashed) and* $|g\,p_{j,k}(\lambda_M)v_n^{j-1}/(2\eta_{j-1})|$, $j = 2:30$ *(different symbols), sum of the values (+)*

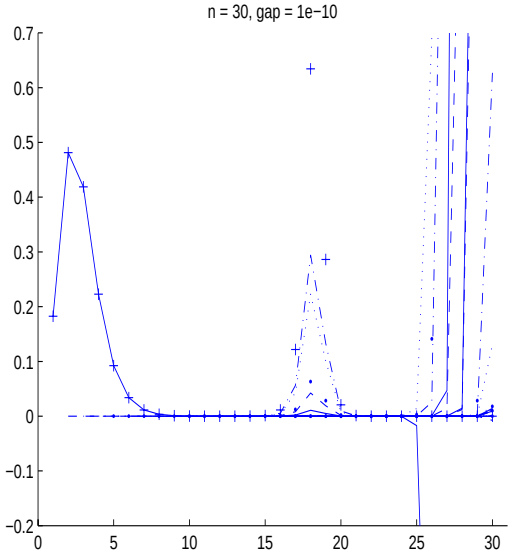

FIG. 5.5. *gap*=10^{-10}, $p_{1,k}(\lambda_M)v_n^1$ *(solid),* $g\,p_{j,k}(\lambda_M)v_n^{j-1}/(2\eta_{j-1})$, $j = 2:30$ *(different symbols) and sum of the values (+)*

curve. As before with the first associated polynomials it gives the right behavior of v_n^k when it increases but after some decrease the first derivative is going up again and does not explain what happens in the last iterations.

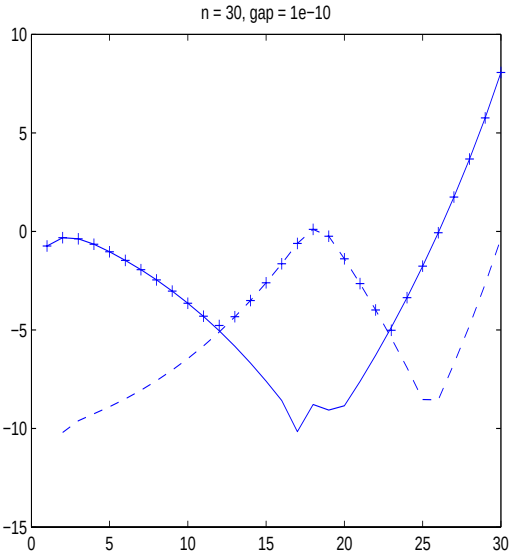

FIG. 5.6. *gap*=10^{-10}, $\log_{10}$ *of* $|p_{1,k}(\lambda_M)v_n^1|$ *(solid),* $|g\,p'_{j,k}(\lambda_M)v_n^1|$ *(dashed) and sum of the values (+)*

How can we prove (or at least understand) that the polynomial $p_{1,k}$ is first decreasing at λ_M and the polynomials $p_{j,k}$ increasing?

Since

$$p_{1,k+1}(\lambda) = \frac{(-1)^k}{\eta_1 \cdots \eta_k} \prod_{l=1}^{k} (\theta_l^{(k)} - \lambda),$$

let us first look at the Ritz values and the eigenvalues of $T_{j,k}$, $j \geq 2$. As long as $\theta_k^{(k)} \leq \lambda_M$ all the terms $\theta_l^{(k)} - \lambda_M$ are negative and due to the factor $(-1)^k$ the polynomial is positive at λ_M.

Figure 5.7 displays the Ritz values, from those of T_1 at the top to those of T_{30} at the bottom. The vertical (green) dotted lines represent the eigenvalues of A which are also shown as large (red) dots on the x-axis. The largest Riz value goes very fast to the cluster. We also see the appearance of a Ritz value between 80 and 90 at iteration 18. Then it moves close to the cluster. Figure 5.8 is a zoom on the cluster. The largest Ritz value jumps right to the middle of the cluster and the stagnation ends when $\theta_{k-1}^{(k)}$ arrives quite close to the cluster. Figure 5.9 shows the eigenvalues of the trailing matrix $T_{2,k}$. An eigenvalue close to 100 appears at iteration 19. This happens for all the trailing matrices up to $T_{17,k}$ but not for $T_{j,k}, j \geq 18$ as we can see in figure 5.10 for $j = 18$.

When $p_{1,k}(\lambda_M) \neq 0$, the ratio $p_{1,k+1}(\lambda_M)/p_{1,k}(\lambda_M)$ can be written as

$$\frac{p_{1,k+1}(\lambda_M)}{p_{1,k}(\lambda_M)} = -\frac{\theta_k^{(k)} - \lambda_M}{\eta_k} \prod_{l=1}^{k-1} \frac{\theta_l^{(k)} - \lambda_M}{\theta_l^{(k-1)} - \lambda_M}. \tag{5.10}$$

By interlacing properties of the Ritz values all the ratios in the product are larger

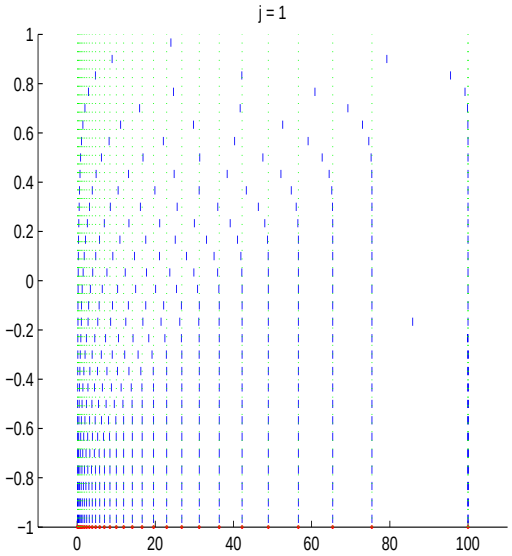

FIG. 5.7. *gap*=10^{-10}, *eigenvalues of* T_k *(Ritz values),* $k = 1$ *(top) to* $k = 30$ *(bottom)*

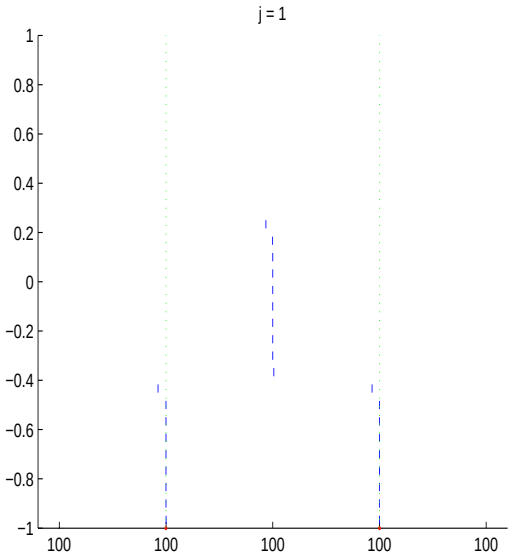

FIG. 5.8. *gap*=10^{-10}, *eigenvalues of* T_k, $k = 1$ *(top) to* $k = 30$ *(bottom) - Zoom to the cluster*

than 1. Hence for $|p_{1,k+1}(\lambda_M)/p_{1,k}(\lambda_M)|$ to be smaller than 1 we need

$$\frac{|\theta_k^{(k)} - \lambda_M|}{\eta_k} < 1,$$

but this is only a necessary condition. More generally this ratio has to be small

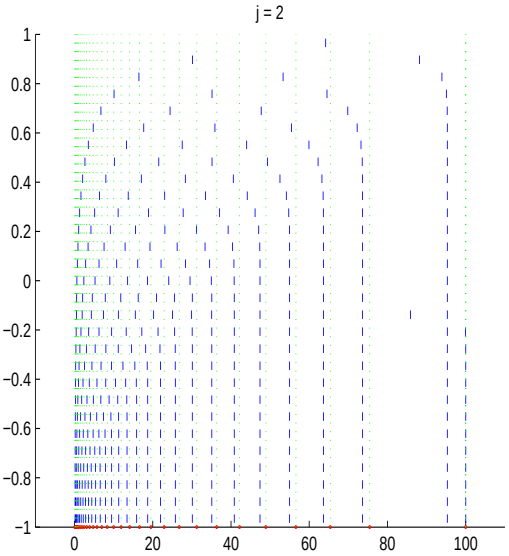

FIG. 5.9. *gap*=10^{-10}, *eigenvalues of* $T_{2,k}$, $k = 2$ *(top) to* $k = 30$ *(bottom)*

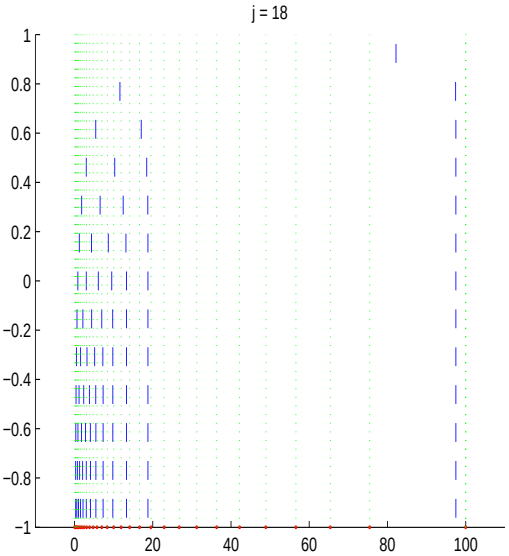

FIG. 5.10. *gap*=10^{-10}, *eigenvalues of* $T_{18,k}$, $k = 18$ *(top) to* $k = 30$ *(bottom)*

for the polynomial values to decrease rapidly. The values of this ratio are given in figure 5.11. Except for the first iteration it is much smaller than 1 but it is increasing after iteration 20 which is the end of the stagnation phase. Although computing these quantities in double precision is not really reliable, figure 5.12 displays $|\theta_k^{(k)} - \lambda_M|/\eta_k$ and $\prod_{l=1}^{k-1} |\theta_l^{(k)} - \lambda_M|/|\theta_l^{(k-1)} - \lambda_M|$. These two curves are almost the mirror of each

other. Their product which is the ratio of the polynomial values is shown in figure 5.13 up to iteration 14 because after that point there are some numerical problems to compute it.

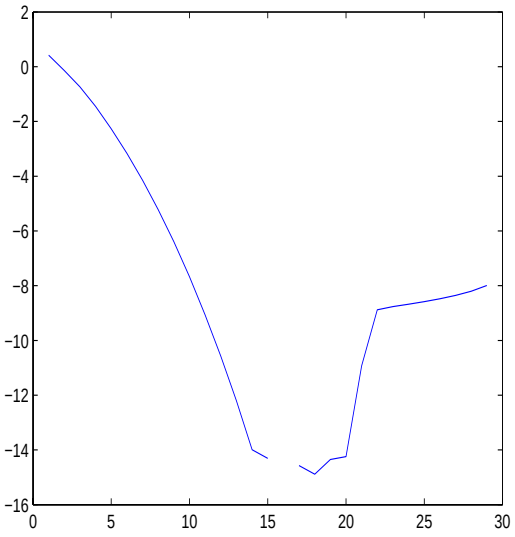

FIG. 5.11. *gap*=10^{-10}, $\log_{10}$ *of* $|\theta_k^{(k)} - \lambda_M|/\eta_k$

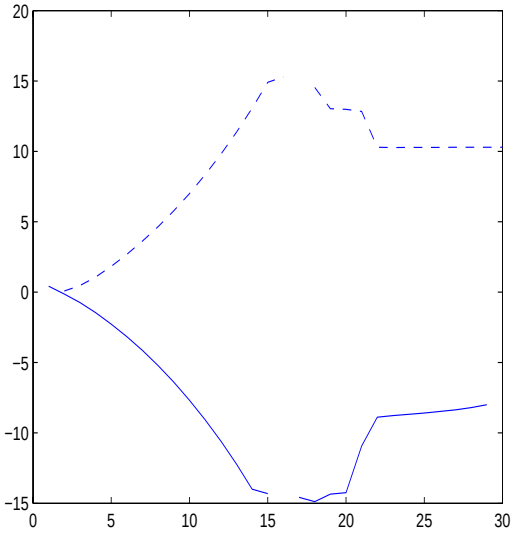

FIG. 5.12. *gap*=10^{-10}, $\log_{10}$ *of* $|\theta_k^{(k)} - \lambda_M|/\eta_k$ *(solid) and* $\prod_{l=1}^{k-1} |\theta_l^{(k)} - \lambda_M|/|\theta_l^{(k-1)} - \lambda_M|$ *(dashed)*

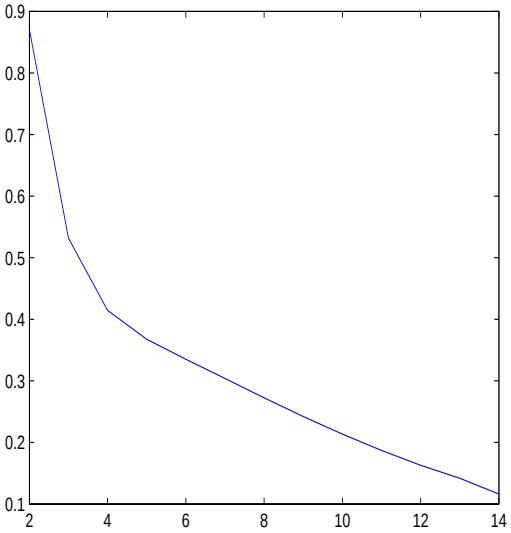

FIG. 5.13. *gap*=10^{-10}, $|p_{1,k+1}(\lambda_M)/p_{1,k}(\lambda_M)|$

Another approach to this problem is to use a result in Belmehdi [1].

PROPOSITION 5.2. *Let $j > 1$, $k \geq j$,*

$$p_{j,k+1}p_{1,k} - p_{j,k}p_{1,k+1} = \frac{\eta_{j-1}}{\eta_k} p_{1,j-1}.$$

Dividing by $p_{1,k}p_{j,k}$ that we suppose to be nonzero, we obtain

$$p_{1,k}p_{j,k} = \frac{\eta_{j-1}}{\eta_k} \frac{p_{1,j-1}}{\frac{p_{j,k+1}}{p_{j,k}} - \frac{p_{1,k+1}}{p_{1,k}}}.$$

Hence, if

$$\eta_k \left(\frac{p_{j,k+1}}{p_{j,k}} - \frac{p_{1,k+1}}{p_{1,k}} \right)$$

does not depend too much on k, then the product $p_{1,k}p_{j,k}$ is independent of k and $p_{j,k}(\lambda)$ is increasing when $p_{1,k}(\lambda)$ is decreasing as a function of k for a given λ. Figure 5.14 shows that this quantity is indeed almost constant (at least in log scale) at λ_M up to iteration 16 and then from iterations 20 to 25.

A third way of looking at the decrease or increase of these polynomials is to consider the three-term recurrence. If we write $P_k(\lambda) = (p_1(\lambda) \cdots p_k(\lambda))^T$ with $p_1 \equiv 1$ and $p_j \equiv p_{1,j}$ we have

$$(T_k - \lambda I)P_k(\lambda) + \eta_k p_{k+1}(\lambda)e^k = 0,$$

where e^k is the last column of the identity matrix of order k. The Ritz values $\theta_j^{(k)}$ are the eigenvalues of T_k and the roots of p_{k+1}. Generally $\lambda_M \neq \theta_j^{(k)}$ even though it can

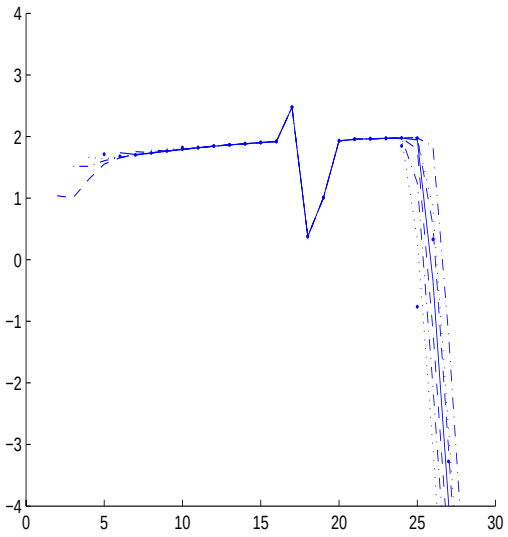

FIG. 5.14. *gap*=10^{-10}, $\log_{10}$ *of* $|\eta_k(p_{j,k+1}/p_{j,k} - p_{1,k+1}/p_{1,k})|$ *at* λ_M, $j = 2, \ldots, 10$

be close to $\theta_k^{(k)}$. Therefore $T_k - \lambda_M I$ is nonsingular. Hence

$$P_k(\lambda_M) = (\lambda_M I - T_k)^{-1}\eta_k p_{k+1}(\lambda_M)e^k.$$

Considering the last component of this vector equation leads to

$$p_k(\lambda_M) = [(\lambda_M I - T_k)^{-1}]_{k,k}\eta_k p_{k+1}(\lambda_M).$$

The ratio of the polynomial values at λ_M is given by

$$\frac{p_{k+1}(\lambda_M)}{p_k(\lambda_M)} = \frac{1}{\eta_k[(\lambda_M I - T_k)^{-1}]_{k,k}}.$$

The k, k element of the inverse is given by the first pivot function $\delta_k(\lambda)$, see [2], page 11. We have

$$[(\lambda I - T_k)^{-1}]_{k,k} = -\frac{1}{\delta_k(\lambda)}.$$

The value $\delta_k(\lambda_M)$ is the last element of the diagonal of a Cholesky-like factorization of $T_k - \lambda_M I$. It is given by the non linear recurrence

$$\delta_1(\lambda) = \alpha_1 - \lambda, \quad \delta_l(\lambda) = \alpha_l - \lambda - \frac{\eta_{l-1}^2}{\delta_{l-1}(\lambda)}, \ l = 2, \ldots, k.$$

Hence, we have $|p_{k+1}(\lambda_M)| < |p_k(\lambda_M)|$ if and only if

$$|\delta_k(\lambda_M)| < \eta_k.$$

The function δ_k is a rational function. It has zeros at $\theta_j^{(k)}$, $j = 1, \ldots, k$ and poles at $\theta_j^{(k-1)}, j = 1, \ldots, k-1$. It is decreasing between the poles. Here the function of interest is a scaled version of δ_k. Let

$$\gamma_1(\lambda) = \frac{\alpha_1 - \lambda}{\eta_1}, \quad \gamma_l(\lambda) = \frac{\alpha_l - \lambda}{\eta_l} - \frac{\eta_{l-1}}{\eta_l} \frac{1}{\gamma_{l-1}(\lambda)}, \; l = 2, \ldots, k.$$

We would like to know when $|\gamma_k(\lambda_M)| < 1$. Figures 5.15 to 5.21 show parts of the function γ_k for different values of the iteration number k. The star where the curve crosses the x-axis is the root $\theta_k^{(k)}$. The other star (located to the right for the first iterations) is the value of λ_M and the circle is the value of the function $\gamma_k(\lambda_M)$ that we wish to have an absolute value smaller than 1. Let us denote $\theta_0^{(k-1)} = -\infty$ and $\theta_k^{(k-1)} = \infty$. Since the function γ_k is continuous and decreasing in between the poles and have a zero in each of these intervals, we know that there exists an interval in every $[\theta_j^{(k-1)}, \theta_{j+1}^{(k-1)}]$, $j = 0, \ldots, k-1$ for which the values of the function are such that $|\gamma_k(\lambda)| < 1$. However, this interval can eventually be very small.

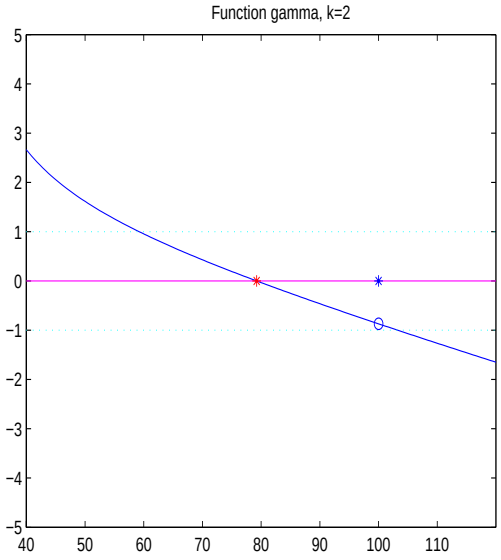

FIG. 5.15. *gap*=10^{-10}, *function* γ_k, $k = 2$

For our example, we are interested in the last interval $[\theta_{k-1}^{(k-1)}, \infty[$. There exist two values $\underline{\lambda}$ and $\overline{\lambda}$ (which depend on k) defined implicitly by

$$\alpha_k - \underline{\lambda} - \frac{\eta_{k-1}}{\gamma_{k-1}(\underline{\lambda})} = -\eta_k, \quad \alpha_k - \overline{\lambda} - \frac{\eta_{k-1}}{\gamma_{k-1}(\overline{\lambda})} = \eta_k,$$

such that

$$\theta_{k-1}^{(k-1)} < \underline{\lambda} < \overline{\lambda},$$

and $\gamma_k(\underline{\lambda}) = 1$, $\gamma_k(\overline{\lambda}) = -1$. The question is: when does λ_M belong to the interval $[\underline{\lambda}, \overline{\lambda}]$? It is not true at the first iteration but become true for iterations 2 to 16. Since

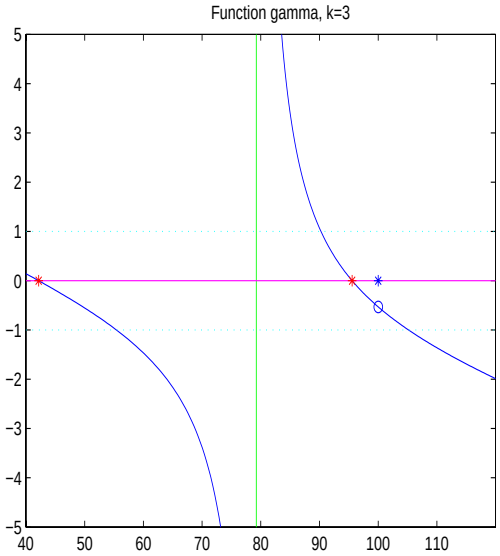

FIG. 5.16. *gap*=10^{-10}, *function* γ_k, $k = 3$

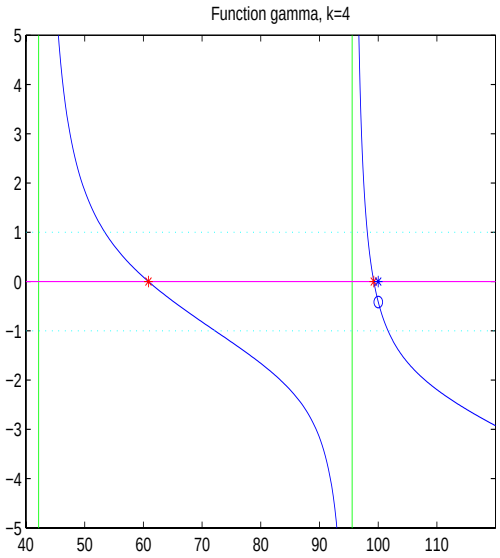

FIG. 5.17. *gap*=10^{-10}, *function* γ_k, $k = 4$

the largest Ritz value $\theta_k^{(k)}$ moves towards λ_M the function γ_k becomes steeper (or more and more "vertical"). Therefore the interval $[\underline{\lambda}, \overline{\lambda}]$ shrinks. This steepening of the functions δ_k and γ_k arise because the Ritz value $\theta_k^{(k)}$ is approaching an eigenvalue of A. Denoting by Z_k the orthonormal matrix of the eigenvectors of T_k, we have

$$T_k - \lambda I = Z_k(\Theta_k - I)Z_k^T.$$

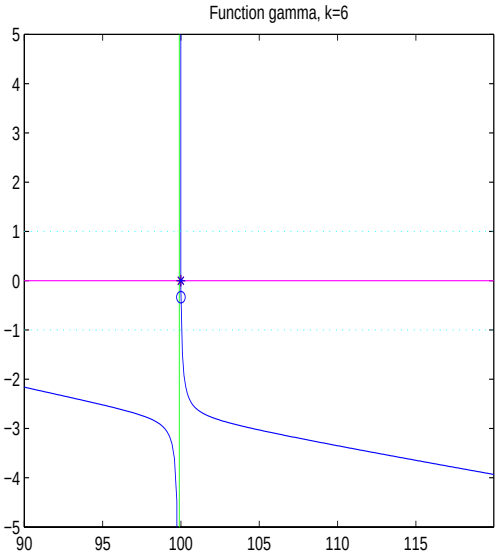

FIG. 5.18. *gap*=10^{-10}, *function* γ_k, $k = 6$

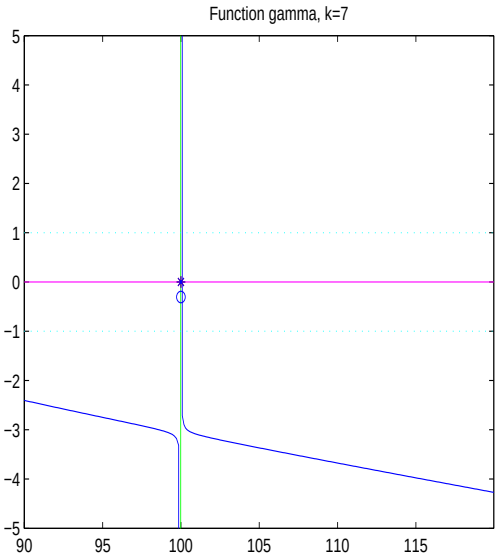

FIG. 5.19. *gap*=10^{-10}, *function* γ_k, $k = 7$

Therefore,

$$\delta_k(\lambda) = \frac{1}{(e^k)^T Z_k(\Theta_k - \lambda I)^{-1} Z_k^T e^k}.$$

The vector $Z_k^T e^k$ has its components equal the last elements of the eigenvectors.

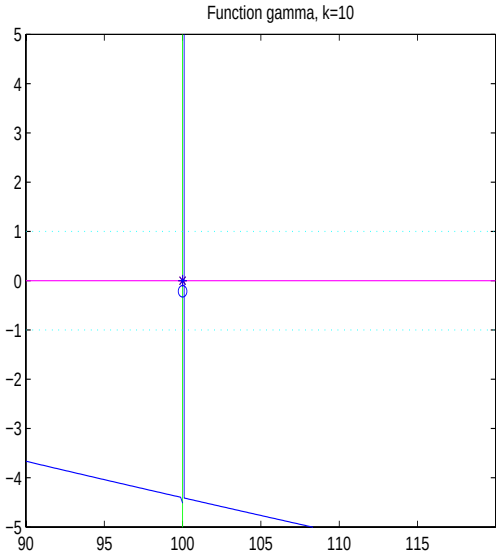

FIG. 5.20. *gap*=10^{-10}, *function* γ_k, $k = 10$

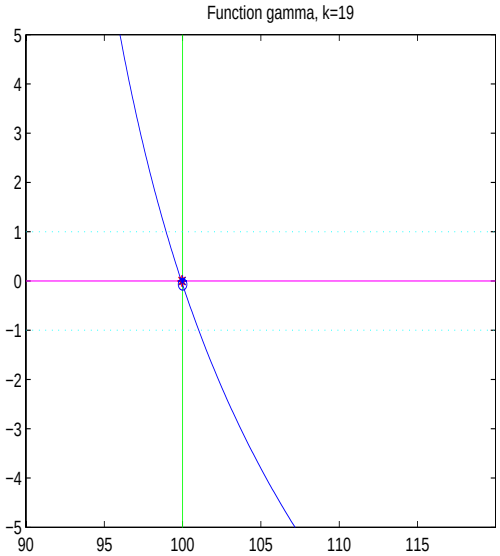

FIG. 5.21. *gap*=10^{-10}, *function* γ_k, $k = 19$

Therefore,

$$\delta_k = \frac{1}{\sum_{j=1}^{k} \frac{(z_k^j)^2}{\theta_j^{(k)} - \lambda}} = \frac{\prod_{j=1}^{k} (\theta_j^{(k)} - \lambda)}{\sum_{j=1}^{k} (z_k^j)^2 \prod_{\substack{m=1 \\ m \neq j}}^{k} (\theta_m^{(k)} - \lambda)}.$$

When a Ritz value "converges" the last component of the corresponding eigenvector goes to zero. This explains the behavior of the function δ_k as k increases. At some point, $\theta_k^{(k)}$ becomes larger than λ_M and the value of the function γ_k is now computed on the next to last branch which is almost vertical close to the pole. Hence, we obtain a negative value whose absolute value is large.

5.3. Question 2. Why is $\theta_k^{(k)}$ first "converging" to λ_M?

It is instructive to consider how the Ritz values at iteration $k+1$ are obtained from the Ritz values at the previous iteration. The eigenvalues of T_{k+1} are solutions of the so-called "secular equation" for θ

$$\alpha_{k+1} - \eta_k^2 \sum_{j=1}^{k} \frac{\xi_j^2}{\theta_j^{(k)} - \theta} = \theta,$$

where $\xi_j = z_k^j = (Z_k^T e^k)_j$ and Z_k is the orthogonal matrix of the eigenvectors of T_k which means the ξ_js are the last components of the eigenvectors of T_k. Solving the secular equation can be interpreted as finding the intersections of the rational function

$$g(\theta) = \eta_k^2 \sum_{j=1}^{k} \frac{\xi_j^2}{\theta_j^{(k)} - \theta},$$

and the straight line of equation $\alpha_{k+1} - \theta$ with slope -1. The function g (not to be confused with the gap) is increasing in the intervals between the poles $\theta_j^{(k)}$ which are the Ritz values at the previous iteration. However, the function g tends to be flat in between the poles when the corresponding value ξ_j is small. This corresponds to the "convergence" of a Ritz value. In this case the function is almost vertical close to the poles and almost horizontal in between, see figure 5.27.

Figure 5.22 displays the situation at the first iteration. The only Ritz value (the pole) is $\alpha_1 = 23.9886$. At iteration 2, the value of α has moved to $\alpha_2 = 64.1879$ because

$$\alpha_k = \sum_{i=1}^{n} \lambda_i (v_i^k)^2,$$

and the components v_{n-1}^2 and v_n^2 have increased. Hence, the largest Ritz value is moving rapidly to the right during the first iterations. This "convergence" gives a decrease of the components v_{n-1}^k and v_n^k and consequently α_k then moves to the left. At iteration 7 (figures 5.24 and 5.25) we have a good approximation of λ_{n-2} and the largest Ritz value is close to the cluster. This is why we have this shape of the function g in the last two intervals. The function is almost flat between the poles and almost vertical close to the poles. It happens because the last elements of the corresponding eigenvectors of T_k are small. As long as we have this situation the two largest Ritz values cannot move by much because they are found on the vertical portions of g since α_k is far away. At iteration 17, α suddenly moves to the right because we have large values of v_{n-1}^k and v_n^k, see figure 5.26. This allows to have a Ritz value in between λ_{n-2} and λ_{n-1}. At iteration 18, α has moved back to the left but the second largest Ritz value is now moving fast towards the cluster. It is also interesting to look at iteration 22 (figures 5.27 and 5.28) for which we obtain Ritz values (at iteration 23) which are close to λ_{n-1} and λ_n.

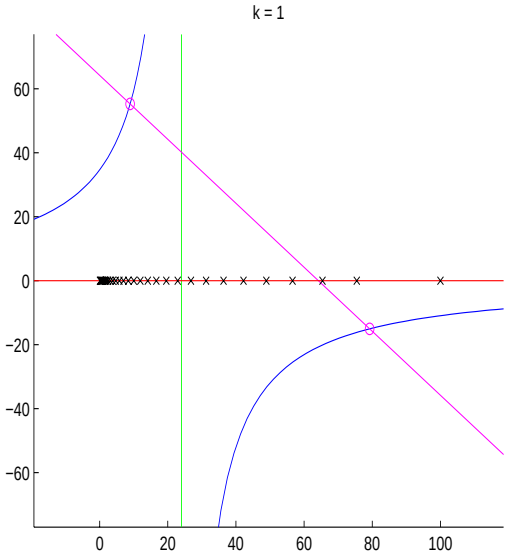

FIG. 5.22. *gap*=10^{-10}, *solving the secular equation* $k = 1$

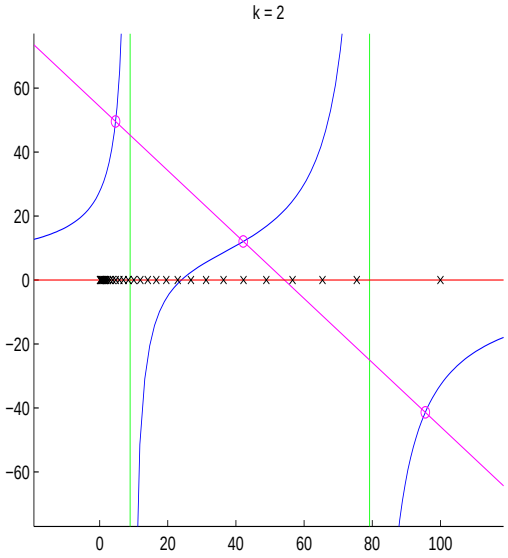

FIG. 5.23. *gap*=10^{-10}, *solving the secular equation* $k = 2$

So we understand why, at the beginning, the Ritz values do not "see" that we have a cluster of close eigenvalues. In fact the growth of $|v_{n-1}^k|$ and $|v_n^k|$ after iteration 12 is a blessing since it allows the appearance of a new Ritz value that can move to the cluster. It remains to see why in the example the largest Ritz value goes to the middle of the cluster and stays there for a few iterations.

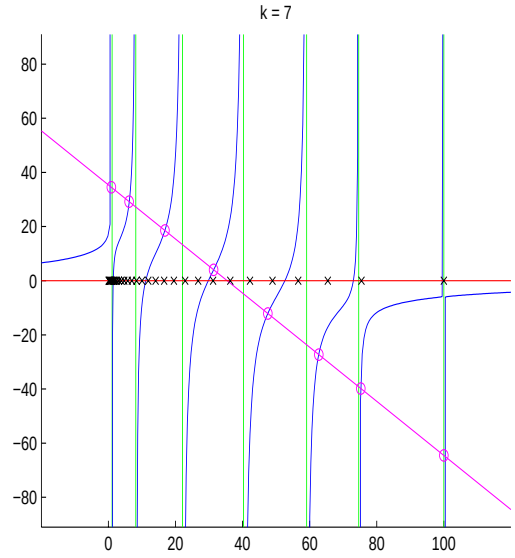

FIG. 5.24. *gap*=10^{-10}, *solving the secular equation* $k = 7$

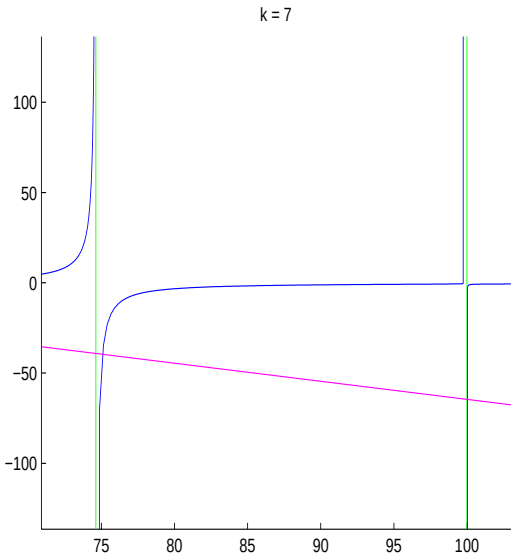

FIG. 5.25. *gap*=10^{-10}, *solving the secular equation* $k = 7$ *(zoom)*

Let us assume that $\theta_{k-1}^{(k)}$ is far from the cluster and $\theta_k^{(k)}$ is within the cluster. The Lanczos polynomial can be written as

$$p_{1,k+1}(\lambda) = C_k(\lambda)(\theta_{k-1}^{(k)} - \lambda)(\theta_k^{(k)} - \lambda).$$

The values of $C_k(\lambda)$ for $\lambda \in [\lambda_{n-1}\, \lambda_n]$ are not much different. Therefore to look for

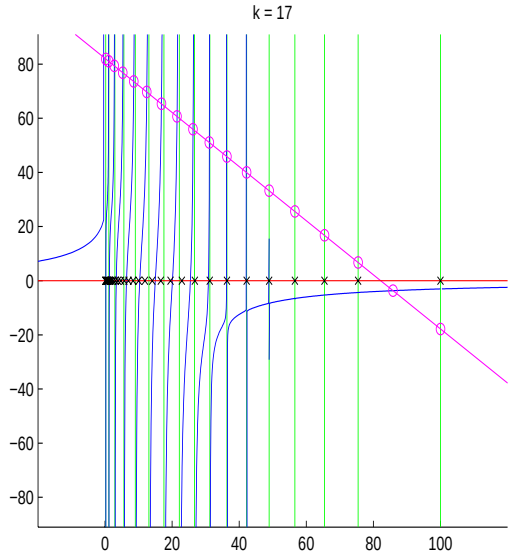

FIG. 5.26. *gap*=10^{-10}, *solving the secular equation* $k = 17$

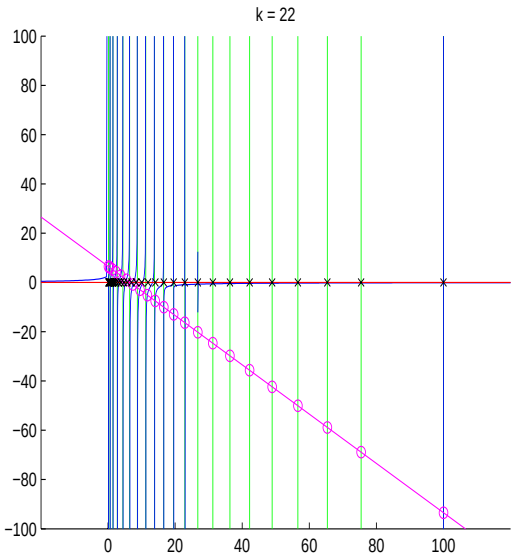

FIG. 5.27. *gap*=10^{-10}, *solving the secular equation* $k = 22$

the root located in the cluster we are interested in

$$t(\lambda) = (\theta_{k-1}^{(k)} - \lambda)(\theta_k^{(k)} - \lambda) = \lambda^2 - (\theta_{k-1}^{(k)} + \theta_k^{(k)})\lambda + \theta_{k-1}^{(k)}\,\theta_k^{(k)} = 0.$$

For simplicity of notations, let us use $\theta_1 = \theta_{k-1}^{(k)}$ and $\theta_2 = \theta_k^{(k)}$. So θ_1 is far from the cluster and θ_2 (that we are looking for) is within the cluster. It is not too difficult

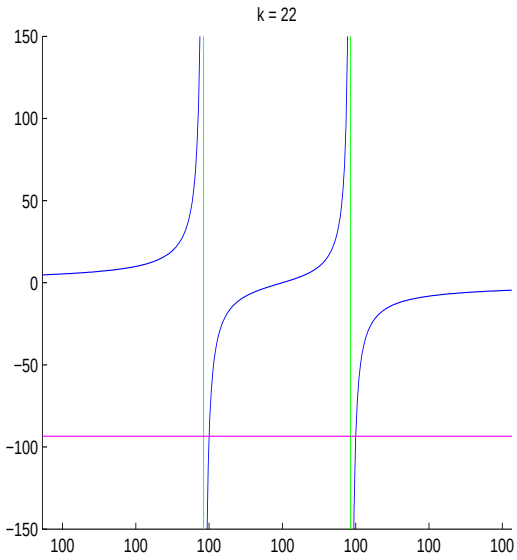

FIG. 5.28. *gap*=10^{-10}, *solving the secular equation* $k = 22$ *(zoom on the cluster)*

to see that the derivative $p'_{1,k+1}$ is negative at θ_1. Hence going from θ_1 to λ_n the polynomial is first decreasing and then increasing with a root θ_2 in the cluster. We are interested in the shape of $t(\lambda)$ within the cluster. To see more clearly the issues, let us do a change of variable from $[\lambda_{n-1}\, \lambda_n]$ to $[-1\, 1]$. Let

$$\mu = \frac{2}{g}\lambda - \frac{\lambda_{n-1} + \lambda_n}{g} = \frac{2}{g}(\lambda - \lambda_M).$$

As a function of μ, t is written as

$$t(\lambda(\mu)) = \left(\frac{g^2}{4}\right)\mu^2 + (g\lambda_M - \frac{g}{2}(\theta_1 + \theta_2))\mu + \lambda_M^2 - \lambda_M(\theta_1 + \theta_2) + \theta_1\theta_2.$$

Since g is small and $|\mu| \leq 1$ the function t must behave as a straight line of equation

$$g(\lambda_M - \frac{\theta_1 + \theta_2}{2})\mu + \lambda_M^2 - \lambda_M(\theta_1 + \theta_2) + \theta_1\theta_2.$$

Figure 5.29 shows the Lanczos polynomial $p_{1,k+1}$ for $k = 15$ for $\lambda \in [\lambda_{n-2}\, \lambda_n]$. Figure 5.30 displays the polynomial in the cluster interval $[\lambda_{n-1}\, \lambda_n]$ at the same iteration. It is a linear function to a good approximation. Now, let us denote by y_n (resp. y_{n-1}) the value of the polynomial at λ_n (resp. λ_{n-1}). The value of the root θ_2 within the cluster is approximately

$$\theta_2 \approx \frac{y_n\lambda_{n-1} - y_{n-1}\lambda_n}{y_n - y_{n-1}}.$$

In our example when the perturbations are dominant we have $v_{n-1}^k \approx -v_n^k$ since the initial perturbations have opposite signs. Therefore $y_{n-1} \approx -y_n$ and

$$\theta_2 \approx \frac{\lambda_{n-1} + \lambda_n}{2} = \lambda_M.$$

During this phase of the computation we have the remarkable property

$$\left|\frac{p_{1,k}(\lambda_{n-1})}{p_{1,k}(\lambda_n)}\right| \approx \frac{(v_n^1)^2}{(v_{n-1}^1)^2}.$$

As long as the polynomial is almost linear within the cluster the root $\theta_k^{(k)}$ is at the middle of the cluster (when the initial weights are the same).

As we can observe in figure 5.31 the polynomial is not linear anymore within the cluster when the second largest Ritz value becomes close to the cluster.

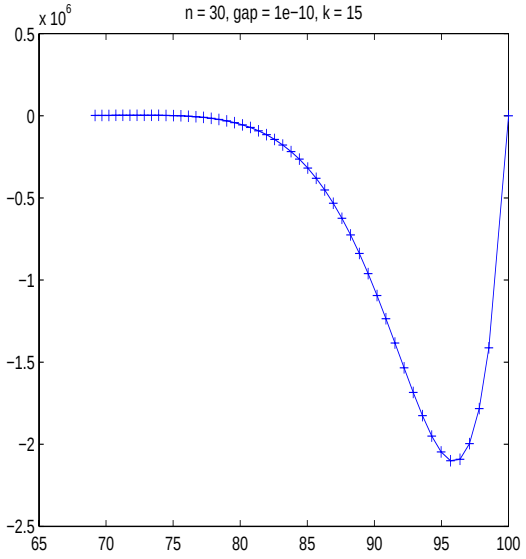

FIG. 5.29. *gap*=10^{-10}, *values of the Lanczos polynomial* $k = 15$

5.4. Question 3. What is the duration of the stagnation phase?

We have seen that the stagnation starts almost immediately when $\theta_k^{(k)}$ enters the cluster. It ends soon after we reach the peak of v_n^k and obtain a large value of α_k because another Ritz value is moving towards the cluster. In our example we have the distance of $\theta_k^{(k)}$ to λ_{n-1} and to λ_n which must be the same but with opposite signs during and after the stagnation phase. We have (see [2] page 29)

$$\lambda_i - \theta_k^{(k)} = \eta_k \frac{z_k^k v_i^{k+1}}{\sum_{l=1}^k z_l^k v_i^l}.$$

Therefore in our case we must have (approximately)

$$\frac{z_k^k v_n^{k+1}}{\sum_{l=1}^k z_l^k v_n^l} = -\frac{z_k^k v_{n-1}^{k+1}}{\sum_{l=1}^k z_l^k v_{n-1}^l}.$$

This gives the relation

$$\sum_{l=1}^k z_l^k (v_n^{k+1} v_{n-1}^l - v_{n-1}^{k+1} v_n^l).$$

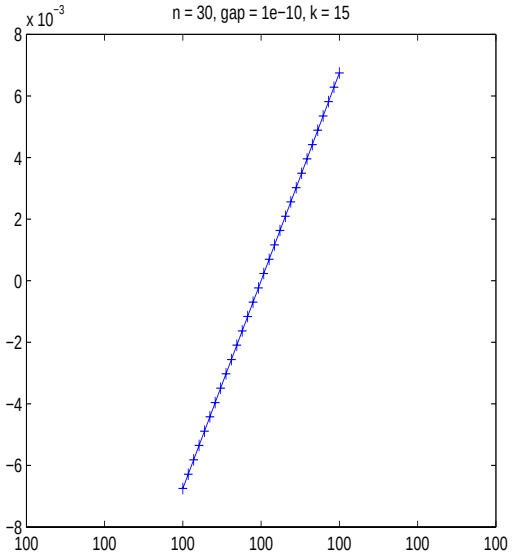

FIG. 5.30. *gap*=10^{-10}, *values of the Lanczos polynomial* $k = 15$ *(zoom on the cluster)*

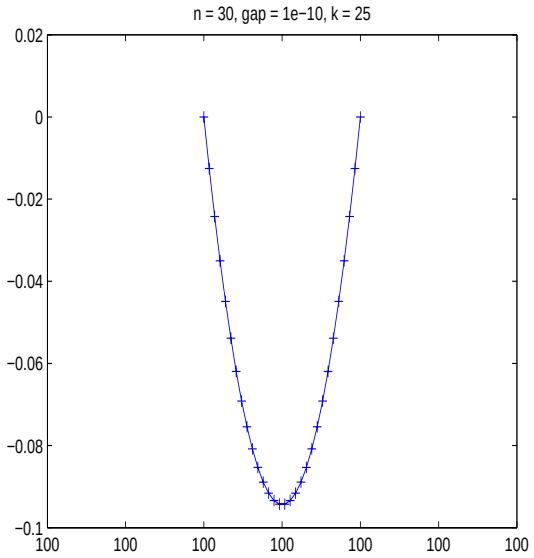

FIG. 5.31. *gap*=10^{-10}, *values of the Lanczos polynomial* $k = 25$ *(zoom on the cluster)*

After the beginning of the stagnation phase we have $v_n^{k+1} = -v_{n-1}^{k+1}$. Hence we must have

$$\sum_{l=1}^{k} z_l^k (v_{n-1}^l - v_n^l) \approx 0. \tag{5.11}$$

At the beginning of the computation (with equal weights) the difference between v^l_{n-1} and v^l_n is smaller than 10^{-6} until iteration 12. Then the difference grows up to iteration 18. Figure 5.32 shows the components of the last eigenvector of T_k for $k = 1, \ldots, 30$ and figure 5.33 displays the sum in equation (5.11). The sum is reasonably small until iteration 20 for which the value is $3.04118\ 10^{-5}$.

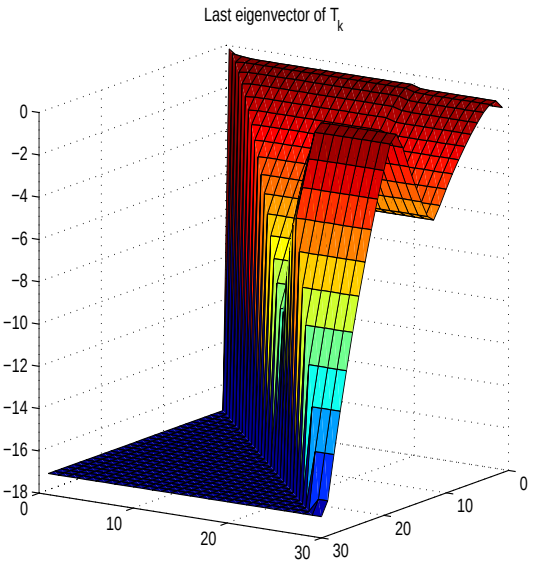

FIG. 5.32. *gap*=10^{-10}, *components of the last eigenvector of* T_k

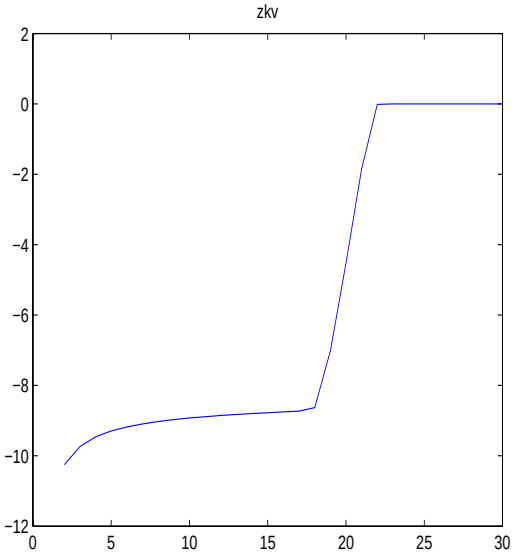

FIG. 5.33. *gap*=10^{-10}, *sum in equation (5.11)*

5.5. Question 4. When and why does $\theta_k^{(k)}$ move towards λ_n?

We have almost already answered to that question. The largest Ritz value starts moving again when the second largest Ritz value come close to the cluster. Then the polynomial $p_{1,k}$ is no longer almost linear within the cluster. So somehow $\theta_k^{(k)}$ is "pushed" by $\theta_{k-1}^{(k)}$. How fast this happens depends on the values of α_k.

6. Computations in finite precision arithmetic. When we use the Lanczos algorithm without reorthogonalization we have also to deal with perturbations arising from rounding errors. However in our example the gap is large enough such that we first see the perturbation from the cluster of the two largest eigenvalues. Now we can do more than 30 iterations. Figure 6.1 displays the logarithm of $|v_n^k|$. There are two types of oscillations. The ones going down to almost 10^{-5} are coming from the cluster and the ones going down to the square root of the machine precision 10^{-8} arise from the rounding errors.

We obtain also multiple copies of the eigenvalues in the cluster. For instance, at iteration 50 we have two copies of each of the two eigenvalues in the cluster.

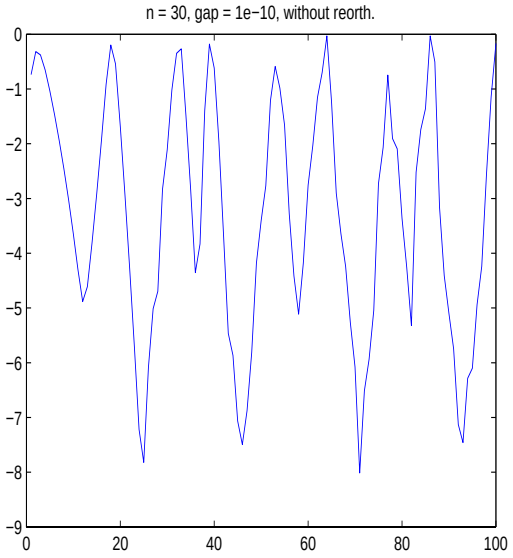

FIG. 6.1. *gap*=10^{-10}, $\log_{10}$ *of* $|v_{30}^k|$ *without reorthogonalization*

7. A cluster with 3 eigenvalues. In this example the largest eigenvalue is still 100 and we have $\lambda_{n-1} = (1 - gap)\lambda_n$ and $\lambda_{n-2} = (1 - 2\,gap)\lambda_n$. The interesting elements of the Lanczos vectors in a computation using reorthogonalization are displayed in figure 7.1. The component v_{n-1}^k has a different behavior because in this example λ_{n-1} is very close to the first stagnation point of $\theta_k^{(k)}$ as explained in the next paragraph.

The largest Ritz value $\theta_k^{(k)}$ stagnates at a point which is the average of the three largest eigenvalues from iteration 11 to iteration 19. This is close to λ_{n-1}. Then when $\theta_{k-1}^{(k)}$ arrives $\theta_k^{(k)}$ moves towards λ_n but soon after stagnates again until iteration 27. The second largest Ritz value $\theta_{k-1}^{(k)}$ stagnates in the cluster from iteration 21 to

iteration 27. The distance $\theta_{k-1}^{(k)} - \lambda_{n-2}$ is the same as the distance $\lambda_n - \theta_k^{(k)}$ during this second stagnation phase. It ends when $\theta_{k-2}^{(k)}$ becomes close to the cluster.

Figure 7.2 shows the values of the different polynomials at the average of the three largest eigenvalues.

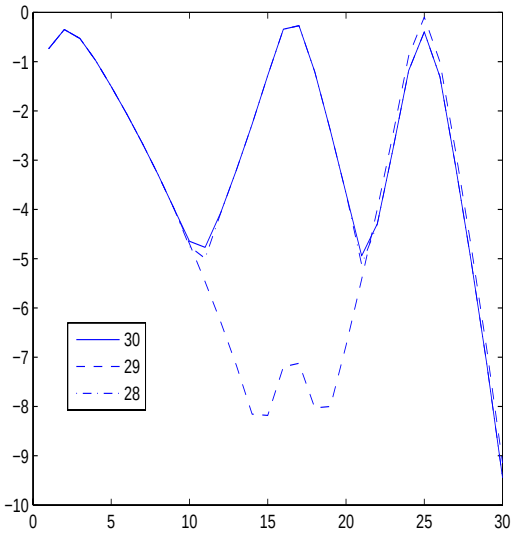

FIG. 7.1. *gap*=10^{-10}, $\log_{10}$ *of* $|v_{30}^k|$, $|v_{29}^k|$ *and* $|v_{28}^k|$

8. Computations with the Arnoldi algorithm. Even though the matrix A is symmetric since it is diagonal we can use the Arnoldi algorithm instead of the Lanczos algorithm. Since the Arnoldi process explicitly orthogonalize against all the previous basis vectors we do not use any extra reorthogonalization. Figure 8.1 displays the logarithm of $|v_n^k|$ for the modified Gram-Schmidt (MGS) implementation of Arnoldi. The results are comparable with those of the Lanczos algorithm with double reorthogonalization except at the end when we see a small increase for MGS. The relative differences can be observed in figure 8.2.

Figure 8.3 shows the results when Arnoldi is implemented using Householder reflections. The result is better at the end even though the last iteration is clearly different from the Lanczos result. Relative differences are given in figure 8.4. Surprisingly they are a little bit higher than with MGS except of course for the last iterations even though global orthogonality is better preserved in the Householder implementation.

9. Conclusions. In this paper we have done a detailed numerical study of the convergence of the Ritz values when using the Lanczos algorithm for a matrix with two close eigenvalues. We have shown that the behavior of the Ritz values can be explained (to some extent) using the same tools as for the Lanczos algorithm in finite precision arithmetic. We hope this study can contribute to a better understanding of the Lanczos algorithm in presence of close eigenvalues.

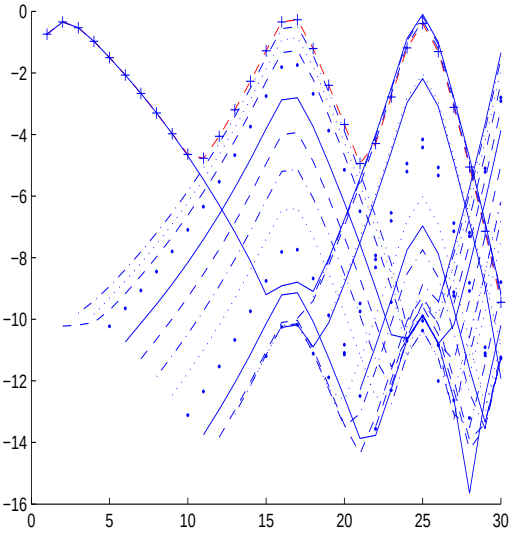

FIG. 7.2. *gap*=10^{-10}, *cluster of 3 eigs,* $\log_{10}$ *of* $|p_{1,k}(\lambda_M)v_n^1|$ *(solid),* $|v_n^k|$ *(dashed) and* $|g\, p_{j,k}(\lambda_M)v_n^{j-1}/(2\eta_{j-1})|$, $j = 2:30$ *(different symbols), sum of the values (+)*

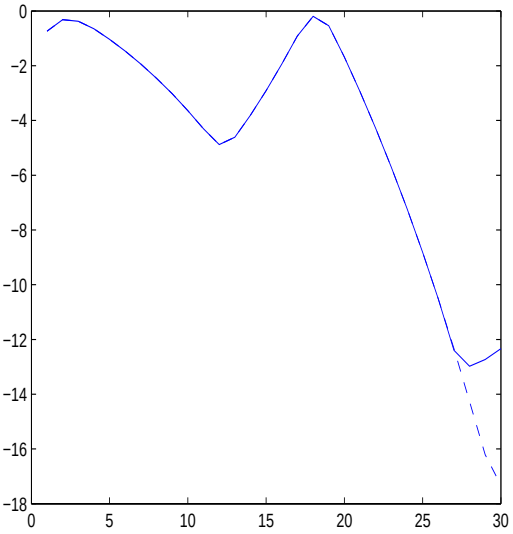

FIG. 8.1. *gap*=10^{-10}, $\log_{10}$ *of* $|v_{30}^k|$, *MGS-Arnoldi (solid) and Lanczos with reorthogonalization (dashed)*

REFERENCES

[1] S. BELMEHDI, *On the associated orthogonal polynomials*, J. Comput. Appl. Math., v 32, (1990), pp 311-319.

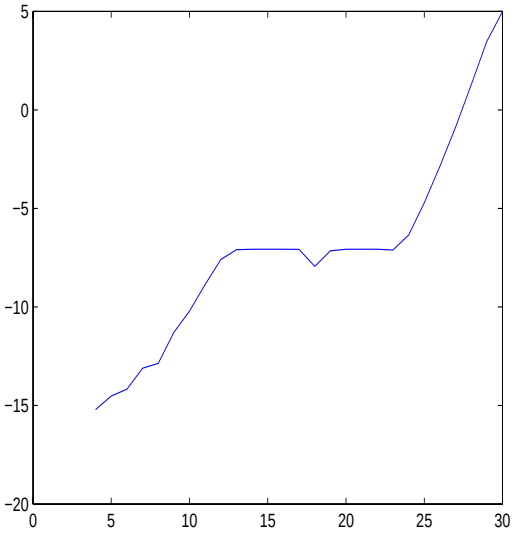

FIG. 8.2. *gap*=10^{-10}, $\log_{10}$ *of the relative difference on* $|v_n^k|$ *between MGS-Arnoldi and Lanczos with reorthogonalization*

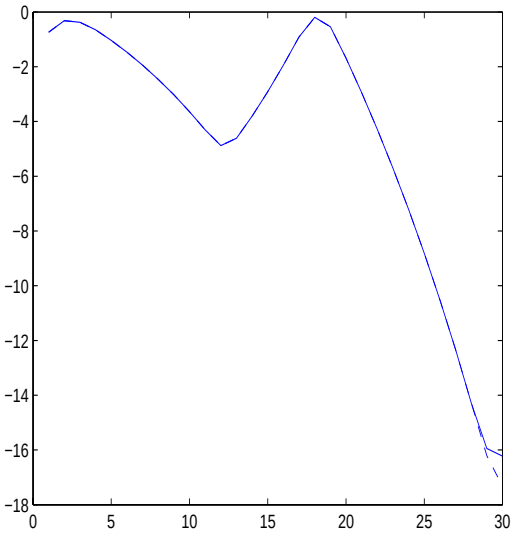

FIG. 8.3. *gap*=10^{-10}, $\log_{10}$ *of* $|v_{30}^k|$, *Householder-Arnoldi (solid) and Lanczos with reorthogonalization (dashed)*

[2] G. MEURANT, *The Lanczos and Conjugate Gradient algorithms, from theory to finite precision computations*, SIAM, (2006).
[3] G. MEURANT AND Z. STRAKOŠ, *The Lanczos and conjugate gradient algorithms in finite precision arithmetic*, Acta Numerica, v 15, (2006), pp 471–542.

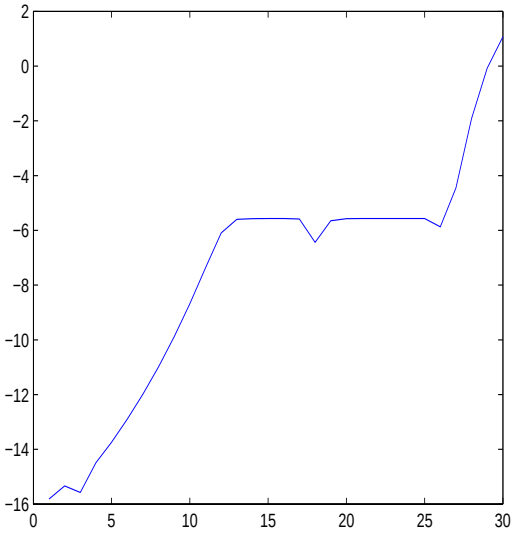

FIG. 8.4. *gap*=10^{-10}, $\log_{10}$ *of the relative difference on* $|v_n^k|$ *between Householder-Arnoldi and Lanczos with reorthogonalization*

[4] C.C. PAIGE, *The computation of eigenvalues and eigenvectors of very large sparse matrices*, Ph.D. thesis, University of London, (1971).
[5] C.C. PAIGE, *Error analysis of the Lanczos algorithm for tridiagonalizing a symmetric matrix*, J. Inst. Maths Applics., v 18, (1976), pp 341–349.
[6] C.C. PAIGE, *Accuracy and effectiveness of the Lanczos algorithm for the symmetric eigenproblem*, Linear Algebra Appl., v 34, (1980), pp 235–258.
[7] B.N. PARLETT, *The symmetric eigenvalue problem*, Prentice Hall, (1980). Reprinted by SIAM in the series Classics in Applied Mathematics (1998).
[8] Z. STRAKOŠ, *On the real convergence rate of the conjugate gradient method*, Linear Alg. Appl., v 154–156 (1991), pp 535–549.
[9] A. VAN DER SLUIS AND H.A. VAN DER VORST, *The convergence behavior of Ritz values in the presence of close eigenvalues*, Linear Algebra Appl., v 88, (1987), pp 651–694.