OPEN QUESTIONS IN THE CONVERGENCE ANALYSIS OF THE LANCZOS PROCESS FOR THE REAL SYMMETRIC EIGENVALUE PROBLEM

By

Z. Strakos

and

A. Greenbaum

IMA Preprint Series # 934

April 1992

OPEN QUESTIONS IN THE CONVERGENCE ANALYSIS OF THE LANCZOS PROCESS FOR THE REAL SYMMETRIC EIGENVALUE PROBLEM

Z. STRAKOŠ* AND A. GREENBAUM**

Abstract. Some open questions related to convergence of Ritz values and vectors in the Lanczos process are formed and studied. This leads to the feeling that the concept of stabilization of weights in the inner products corresponding to the sequence of related orthogonal polynomials could play an important role in the understanding the behavior of Lanczos process. Stabilization of weights is examined. In the case of exact Lanczos algorithm a simple and almost complete answer to the question is received, while in the finite precision case our attempt to prove the stabilization of weights failed and the question remains open.

1. Introduction. The Lanczos algorithm applied to a given N by N real symmetric matrix A with the initial vector r^0 will be considered. It is given e.g. by the recurrence

$$\begin{aligned} q^1 &= r^0/\|r^0\|, \beta_1 = 0 \\ \alpha_k &= (Aq^k - \beta_k q^{k-1}, q^k) \\ w^k &= Aq^k - \alpha_k q^k - \beta_k q^{k-1} \\ \beta_{k+1} &= \|w_k\| \\ q^{k+1} &= w^k/\beta_{k+1}, \qquad k = 1, 2, \ldots \end{aligned} \tag{1.1}$$

or in the matrix form

$$AQ_k = Q_k T_k + \beta_{k+1} q^{k+1} (e^k)^T \tag{1.2}$$

where Q_k is the N by k matrix with orthonormal columns $q^1, q^2, \ldots, q^k$, $e^k = (0, 0, \ldots, 1)^T$, T_k is symmetric tridiagonal matrix

$$T_k = \begin{pmatrix} \alpha_1 & \beta_2 & & \\ \beta_2 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \beta_k \\ & & \beta_k & \alpha_k \end{pmatrix}, \qquad T_k = Q_k^T A Q_k.$$

*IMA, University of Minnesota, 514 Vincent Hall, 206 Church Street S.E., Minneapolis, MN 55455, USA, and ICIS, Czech Academy of Sciences, Pod vod veží 2, 18207 Praha 8, Czechoslovakia. e-mail: na.strakos@na-net.ornl.gov, strakos@imafs.ima.umn.edu.

**Courant Institute of Mathematical Sciences, 251 Mercer Street, New York, NY 10012, USA. e-mail: greenbaum@nyu.edu.

Let $A = U\Lambda U^T$ and $T_k = S_k\Theta_k S_k^T$ be the eigendecompositions of A and T, $\Lambda = \mathrm{diag}(\lambda_i)$, $\Theta_k = \mathrm{diag}(\theta_j^{(k)})$, U and S_k are matrices with the normalized eigenvectors of A and T_k as columns. The eigenvalues $\theta_j^{(k)}$ of T_k (Ritz values) and the vectors $z_j^{(k)} = Q_k s_j^{(k)}$ are then considered as approximations to the eigenvalues and eigenvectors of the matrix A. Consistently to the notation introduced above we denote the matrix with columns $z_j^{(k)}$ as Z_k. There are many results describing the convergence of Ritz pairs produced by (1.1) to the original eigenpairs, for the comprehensive overview we refer to [Parlett–80] and [Cullum–1985].

In practical computation the orthogonality among Lanczos vectors q^k is lost due to rounding errors and (1.2) must be replaced by the perturbed recurrence

$$AQ_k = Q_k T_k + \beta_{k+1} q^{k+1}(e^k)^T + F_k, \tag{1.3}$$

where Q_k, T_k refer to the really computed values and F_k is the $N \times k$ matrix with columns $f^{(k)}$, $\theta_j^{(k)}$, $s_j^{(k)}$ are the exact eigenpairs of T_k. Fundamental work in the analysis of (1.3) was done by Paige. He proved that $\|F_k\| \le k^{1/2}\|A\|\varepsilon_1$, where ε_1 is a small multiple of the machine roundoff unit ε [Paige–76]. Then

$$\|Az_j^{(k)} - \theta_j^{(k)} z_j^{(k)}\| = \|q^{k+1}\beta_{k+1}(e^k, s_j^{(k)}) + F_k s_j^{(k)}\| \le \delta_{kj} + k^{1/2}\|A\|\varepsilon_1 \tag{1.4}$$

where $\delta_{kj} = \beta_{k+1}|(e^k, s_j^{(k)})| = \beta_{k+1}|s_{kj}^{(k)}|$, $s_{kj}^{(k)}$ is the bottom element of the j-th eigenvector of the matrix T_k. For $\|z_j^{(k)}\| \sim 1$ this gives a bound for the convergence of $\theta_j^{(k)}$ because

$$\min_i |\lambda_i - \theta_j^{(k)}| \le \frac{\|Az_j^{(k)} - \theta_j^{(k)} z_j^{(k)}\|}{\|z_j^{(k)}\|} \le \frac{1}{\|z_j^{(k)}\|}(\delta_{kj} + k^{1/2}\|A\|\varepsilon_1). \tag{1.5}$$

Paige analyzed the case $\|z_j^{(k)}\| \ll o(1)$ and proved in fact the next theorem (cf. [Paige–80], p. 249).

THEOREM 1.1 (Paige). *For any Ritz value $\theta_j^{(k)}$ determined at the k-th step of the Lanczos process (1.3)*

$$\min_i |\lambda_i - \theta_j^{(k)}| \le \max\{2.5(\delta_{kj} + k^{1/2}\|A\|\varepsilon_1), (k+1)^3\|A\|\varepsilon_2\}, \tag{1.6}$$

$$|(z_j^{(k)}, q^{k+1})| = \frac{|\varepsilon_{jj}^{(k)}|}{\delta_{kj}}, \quad |\varepsilon_{jj}^{(k)}| \le k\|A\|\varepsilon_2 \equiv \varepsilon_3. \tag{1.7}$$

Thus, $\delta_{kj} \ll o(1)$ implies convergence of $\theta_j^{(k)}$ to some eigenvalue λ_i regardless the value $\|z_j^{(k)}\|$ is small or not (ε_2 is a small multiple of the machine precision unit ε, see [Paige–80]).

Moreover, if the orthogonality among q^{k+1} and $z_j^{(k)}$ is lost, then $\theta_j^{(k)}$ is convergent to some eigenvalue λ_i. The reverse statement cannot be simply stated from (1.6) and (1.7). First, it is not necessarily true that $|\lambda_i - \theta_j^{(k)}| \ll o(1)$ implies $\delta_{kj} \ll o(1)$. Second, even if δ_{kj} is small then it can still be $|\varepsilon_{jj}^{(k)}| \ll \delta_{kj}$ and $(z_j^{(k)}, q^{k+1})$ can be small. We will return to this point later.

We have not mentioned yet any assumption concerning the multiplicity of the matrix A eigenvalues. In exact arithmetic this question has no real impact to the analysis, because the initial vector q^1 projected to the invariant subspace corresponding to some eigenvalue λ_i determine a unique eigenvector from the subspace participating in the process. Exact Lanczos applied to A with a given q^1 can thus give no information about the multiplicity of eigenvalues. In the finite precision Lanczos run the multiplicity can be effectively tested [Paige–72, Cullum–85]. Although we touch this point in next sections, there is not our aim to discuss it in detail here. Therefore from now on the eigenvalues of A are supposed to be simple and ordered such that

$$\lambda_1 < \lambda_2 < \cdots < \lambda_N.$$

Similarly, Ritz values are supposed to be ordered in a similar way,

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

We assume moreover that the projection of initial vector q^1 to any eigenvector of the matrix A is nonzero, i.e. $(q^1, u_i) \neq 0$, $i = 1, 2, \ldots, N$.

Both the exact and finite precision Lanczos processes generate a sequence of real symmetric tridiagonal matrices with positive subdiagonals (Jacobi matrices). In the exact arithmetic the process is terminated by $\beta_{N+1} = 0$. In the finite precision arithmetic one can hardly meet vanishing β_{k+1}; usually the process can be run for an unlimited number of steps with $\beta_{k+1} \gg 0$. Using the notation introduced above we briefly recall some properties of Jacobi matrices.

a) The eigenvalues of T_k must be distinct, i.e. in previous notation $\theta_1^{(k)} < \theta_2^{(k)} < \cdots < \theta_k^{(k)}$.

b) If $k < t$, then for any Jacobi matrix T_t with T_k as its leading principal submatrix there is an eigenvalue $\theta_r^{(t)}$ in any open interval $(\theta_j^{(k)}, \theta_{j+1}^{(k)})$, $(-\infty, \theta_1^{(k)})$, $(\theta_k^{(k)}, +\infty)$.

c) Let $\psi_k(\theta) = \det(\theta I - T_k)$ be the characteristic polynomial of T_k, $\psi_{2,k}$ be the characteristic polynomial of $T_{2,k}$ (submatrix of T_k obtained by cutting off the first row and column). Then

$$(1.8) \qquad s_{1j}^{(k)} s_{kj}^{(k)} \psi_k'(\theta_j^{(k)}) = \prod_{r=2}^{k} \beta_k$$

$$(1.9) \qquad (s_{1j}^{(k)})^2 \psi_k'(\theta_j^{(k)}) = \psi_{2,k}(\theta_j^{(k)}), \quad \psi_k'(\theta_j^{(k)}) = \prod_{r \neq j} (\theta_j^{(k)} - \theta_r^{(k)})$$

$$(1.10) \qquad (s_{kj}^{(k)})^2 \psi_k'(\theta_j^{(k)}) = \psi_{k-1}(\theta_j^{(k)})$$

$$(1.11) \qquad (\theta_j^{(k)} - \theta_r^{(k-1)})(s_j^{(k)})^T \begin{pmatrix} s_r^{(k-1)} \\ 0 \end{pmatrix} = \beta_k s_{k-1,r}^{(k-1)} s_{kj}^{(k)}$$

where $s_{1j}^{(k)}$ is the top element of the j-th normalized eigenvector of the matrix T_k. Proof for a), c) can be found in [Parlett–80], b) is well known property of orthogonal polynomials and in the language of Jacobi matrices it is presented e.g. in [van der Sluis–87]. From (1.8) we see that $s_{1j}^{(k)} \neq 0, s_{kj}^{(k)} \neq 0$.

Let now T_N is a Jacobi matrix with $\theta_j^{(N)} = \lambda_j, j = 1, 2, \ldots N$. Then from (1.2) we can easily derive that T_N is the Lanczos matrix produced in the step N of the Lanczos algorithm applied to the matrix $A = U \Lambda U^T$, $\Lambda = \mathrm{diag}(\lambda_i)$, U is arbitrary orthonormal matrix with the initial vector q^1 defined as

$$(1.12) \qquad q^1 = U(s_{11}^{(N)}, s_{12}^{(N)}, \ldots, s_{1N}^{(N)})^T = U s^0, \quad s^0 = (s_{11}^{(N)}, \ldots, s_{1N}^{(N)})^T.$$

Moreover, the Lanczos run (exact or finite precision) applied to A, q^1 generates in the steps 1 thru n a sequence of Jacobi matrices $T_1, T_2, \ldots, T_n$ which are equivalent to the matrices produced by the (exact) Lanczos algorithm applied to T_n, e^1. This suggests that theoretical properties of Jacobi matrices provide essential tool for the analysis of the Lanczos algorithm and its behavior in finite precision arithmetic (cf. the basic result in [Greenbaum–1989]).

The correspondence between Jacobi matrices and orthonormal polynomials (which is well known since the time of classical development) can be formulated in the following way. Let $\lambda_1 < \lambda_2 < \cdots < \lambda_N$ be N distinct real points, and $m_1, m_2, \ldots m_N$ be N positive values,

$$\sum_{i=1}^{N} m_i = 1.$$

The innerproduct of polynomials φ, ψ with respect to the discrete points $\lambda_1, \ldots, \lambda_N$ and corresponding weights $m_1, \ldots, m_N$ is defined by the formula

$$(1.13) \qquad (\varphi, \psi) = \sum_{i=1}^{N} m_i \varphi(\lambda_i) \psi(\lambda_i).$$

For a given $\lambda_i, m_i, i = 1, \ldots, N$, (1.13) defines the unique set of orthonormal polynomials $1, \varphi_1, \ldots, \varphi_N$, where $\varphi_k(\lambda)$ is of degree k with a positive coefficient corresponding to λ^k. It is not difficult to see that $\varphi_1, \ldots, \varphi_N$ are the normalized Lanczos polynomials $\psi_1, \ldots, \psi_N$ generated by the Lanczos algorithm applied to the matrix $A = U\Lambda U^T$, $\Lambda = \mathrm{diag}(\lambda_i)$, U is arbitrary orthonormal matrix, with the initial vector q^1 given by

$$q^1 = U(m_1^{1/2}, m_2^{1/2}, \ldots, m_N^{1/2})^T. \tag{1.14}$$

Comparing (1.12) and (1.14) gives

$$m_i = (s_{1i}^N)^2 = (u_i, q^1)^2 \tag{1.15}$$

Lanczos polynomials $\psi_k(\lambda)$ can be expressed as

$$\begin{aligned} \psi_k(\lambda) &= \prod_{j=2}^{k+1} \beta_j \quad \varphi_k(\lambda), \quad j = 1, 2, \ldots, N-1, \\ \psi_N(\lambda) &= \prod_{j=2}^{N} \beta_j \quad \varphi_N(\lambda). \end{aligned} \tag{1.16}$$

Moreover, each $\psi_k, k = 1, 2, \ldots, N$, is determined uniquely by the minimizing property

$$\|\psi_k\| = \min\{\|\psi\|, \psi \text{ is monic polynomial of degree at most } k\ \}, \tag{1.17}$$

where $\|\ \ \|$ is a norm induced by the innerproduct (1.13). From (1.16) immediately follows

$$\begin{aligned} q^{k+1} &= \psi_k(A) q^1 / \prod_{j=2}^{k+1} \beta_j \\ e^{k+1} &= \psi_j(T_N) e^1 / \prod_{j=2}^{k+1} \beta_j, \quad k = 1, 2, \ldots N-1. \end{aligned} \tag{1.18}$$

For proofs see [Parlett–1980]. Classical orthogonal polynomials are related to the results reviewed here via the Gauss quadrature rule [Golub–1969]. For a simple description of the relation to the continued fractions see [Gragg–1984], also [Strakoš–1991b].

Thus, the Lanczos run (exact or finite precision) applied to A, q^1 generates in steps 1 thru n a sequence of monic polynomials

$$1, \psi_1, \ldots, \psi_n$$

exactly orthogonal with respect to the innerproduct

$$(\varphi, \psi) = \sum_{j=1}^{n} m_j^{(n)} \varphi(\theta_j^{(n)}) \psi(\theta_j^{(n)}), \tag{1.19}$$

where $\theta_1^{(n)} < \theta_2^{(n)} < \cdots < \theta_n^{(n)}$ are the roots of ψ_n (eigenvalues of T_n) and the weights $m_j^{(n)}$ are given by

$$m_j^{(n)} = (s_{1j}^{(n)})^2, \quad j = 1, 2, \ldots, n. \tag{1.20}$$

Each ψ_k, $k = 1, 2, \ldots, n$ is determined uniquely by the minimizing property (1.18), where norm $\| \ \|$ is induced by the innerproduct (1.19). In the next step $\{\theta_j^{(n)}, m_j^{(n)}\}_{j=1}^{n}$ is replaced by $\{\theta_j^{(n+1)}, m_j^{(n+1)}\}_{j=1}^{n+1}$ in such a way, that

$$1, \psi_1, \ldots, \psi_n$$

which were exactly orthogonal with respect to the innerproduct defined by n-points $\theta_j^{(n)}$ with weights $m_j^{(n)}$, $j = 1, 2, \ldots, n$, are now *exactly* orthogonal with respect to the $n+1$ dimensional innerproduct defined by $n+1$ points $\theta_j^{(n+1)}$ with weights $m_j^{(n+1)}$, $j = 1, 2, \ldots,$ $n+1$, and $\psi_{n+1}(\theta_j^{(n+1)}) = 0$. The real Lanczos run evaluates in this manner a sequence of monic polynomials orthogonal in this special way with respect to a sequence of innerproducts. The innerproducts are changing with increasing dimension n while preserving the orthogonality among the previously determined polynomials. One can expect that the innerproducts tend to approach in some sense the original innerproduct (1.13), (1.15). This raises a question about the convergence of Ritz values and the stabilization of weights (1.20).

This paper is organized as follows. Next section specify the problems which motivated the work. Section 3 collects some basic theoretical results. Most of them are known but they are rather dispersed in the literature, some of them are thought to be new. Because many answers remain incomplete, we present some preliminary results too. Hopefully they can be used in the further work. Section 4 describes our numerical experiments. In sections 5-8 theoretical results are applied to the questions formulated in Section 2.

2. Questions. In this section we try to explain the motivation of our work on this paper.

Question 1. Approximation bounds based on Ritz values at several consecutive steps.

The value of δ_{kj} plays a key role in the approximation bounds (1.4), (1.6) developed by Paige. Can the value δ_{kj} be bounded using only information about Ritz values in several consecutive steps? In Section 5 we show that in many interesting cases the answer is positive. If, e.g., Ritz values $\theta_j^{(k)}$, $\theta_l^{(k+1)}$ at steps k, $k+1$ are close to one another and both of them are well separated from the other Ritz values at that steps, then the quantities δ_{kj}, $\delta_{k+1,l}$ are small. From the relation

$$\delta_{kj} \geq \min_r |\theta_j^{(k)} - \theta_r^{(t)}|, \quad t > k, 1 \leq r \leq t, \tag{2.1}$$

(cf. [Paige–80, Wilkinson–65]) this implies that there must be Ritz value $\theta_r^{(t)}$ close to $\theta_j^{(k)}$ in any subsequent step and from Theorem 1.1 these Ritz values are close to some eigenvalue λ_i of A. Moreover, we develop bounds for δ_{kj} (based on Ritz values in two or three consecutive steps) which can be applied to the outer eigenvalues in a cluster, such as might be generated by a finite precision computation. They lead to a practical test for convergence which does not require computation of eigenvector elements of the tridiagonal matrix. Unfortunately the formulas very rarely give approximation bounds of order less than $\varepsilon^{1/2}$ and they do not establish small value of δ_{kj} for interior Ritz values in the cluster.

Question 2. Clustered Ritz values and eigenvector approximations in the finite precision Lanczos computation.

For the clusters of Ritz values generated in the finite precision Lanczos computation (cf. [Parlett–80], paragraphs 13.3 and 13.6) our experiments suggest much more. It was observed that for *any* Ritz value $\theta_j^{(k)}$ in the tight well separated cluster the value δ_{kj} is small and consequently the corresponding Ritz vector (if nonvanishing) is a good approximation to some eigenvector of the matrix A. Theoretical affirmation of this empirical observation could significantly simplify the error analysis of the Lanczos process given in [Greenbaum–89]. We do not have a proof. Our partial results are discussed in Section 6.

Question 3. Stabilization of weights.

As it was recalled in the Introduction, Lanczos process (exact or finite precision) for the matrix A and initial vector q^1 generates a sequence of innerproducts (1.19). It is natural to ask in which sense the innerproducts (1.19) approximate the original innerproduct (1.13), (1.15). Let e.g. Ritz values $\theta_{j_0}^{(k)}, \theta_{j_1}^{(k+1)}, \ldots, \theta_{j_l}^{(k+l)}$ computed in steps $k, k+1, \ldots, k+l$, approximate eigenvalue λ_i of A. Does it imply that the weights $m_{j_0}^{(k)}, m_{j_1}^{(k+1)}, \ldots, m_{j_l}^{(k+l)}$ are approximately equal to the value m_i? For the exact Lanczos process, assuming that Ritz values $\theta_{j_0}^{(k)}, \theta_{j_1}^{(k+1)}, \ldots, \theta_{j_l}^{(k+1)}$ are well separated from the other Ritz values, we answer this question positively in Section 8. For a finite precision Lanczos process and well separated Ritz values a partial answer is given. For a finite precision Lanczos process and a cluster of Ritz values the question is reformulated as follows: do Ritz values in a tight well separated cluster approximating eigenvalue λ_i of A share the weight m_i? Despite the preliminary results of Section 8, this question remains practically open.

Question 4: Identification of spurious (ghost) eigenvalues.

The approach based on the identification of *spurious* Ritz values was developed by Cullum and Willoughby, e.g. [Cullum-85]. It is a feasible alternative to the approach determining convergent Ritz values based on Paige's work. Spurious eigenvalues are those Ritz values computed in the finite precision Lanczos run, which are the result of the loss of orthogonality and can never be found among the Ritz values (even slightly perturbed) of the exact process for the same data. We will focus on spurious Ritz values which are well separated from theirs neighbors (we are not going to discuss the question of multiplicity of the original eigenvalues).

If we consider for simplicity only well separated Ritz values, then spurious Ritz values computed at the step k are in [Cullum-85] identified with those eigenvalues $\theta_j^{(k)}$ of T_k which are also (numerically) eigenvalues of $T_{2,k}$. The justification for this test is based on creating the associated conjugate gradient procedure and proving its asymptotic convergence (cf. [Cullum–80]). A more general consideration shows in Section 7 that spurious eigenvalues may appear in the early stage of computation when the corresponding CG process is far from convergence and in a part of the spectra which is unconverged. Moreover, there are spurious Ritz values which are not (numerically) eigenvalues of $T_{2,k}$. The discussion in Section 7 shows that any general procedure that reliably identifies *any* spurious eigenvalue must be very complicated.

3. Theoretical properties. Some theoretical properties of Jacobi matrices and of the Lanczos algorithm will be presented followed by some results taking into account the effect of roundoff in real Lanczos runs. First, we need to describe the relation between Ritz values and weights (1.20) in the sequence of innerproducts (1.19). This is related to the inverse eigenvalue problem for Jacobi matrices. In the language of orthogonal polynomials and Lanczos method the situation is simple.

Let $\theta_j^{(n)}, j = 1, \ldots, n$ and $\theta_j^{(2,n)}, j = 1, \ldots, n-1$ be any sets of real numbers with the interlacing property

$$\theta_1^{(n)} < \theta_1^{(2,n)} < \theta_2^{(n)} < \cdots < \theta_{n-1}^{(2,n)} < \theta_n^{(n)}.$$

Then there is unique Jacobi matrix T_n with $\theta_j^{(n)}, j = 1, \ldots, n$, as its eigenvalues and $\theta_j^{(2,n)}, j = 1, 2 \ldots, n-1$, as eigenvalues of $T_{2,n}$. Clearly, let $\psi_n(\theta) = \prod_{j=1}^{n} (\theta - \theta_j^{(n)})$, $\psi_{2,n}(\theta) = \prod_{j=1}^{n-1} (\theta - \theta_j^{(2,n)})$. Then the Lanczos algorithm applied to the diagonal matrix $D_n = \mathrm{diag}(\theta_j^{(n)})$ with the initial vector $q^1 = s^0 = (s_{11}^{(n)}, \ldots, s_{1n}^{(n)})$, where $s_{1j}^{(n)}$ is determined by (1.9), gives the result. Applying the previous argument to the matrix T_n transposed according the skew-diagonal (or using the Lanczos in the reverse ordering, see [Parlett–80]) the result is modified as follows. Let $\theta_j^{(n)}, j = 1, \ldots, n$ and $\theta_j^{(n-1)}, j = 1, \ldots, n-1$ be any sets of real numbers with the interlacing property

$$\theta_1^{(n)} < \theta_1^{(n-1)} < \theta_2^{(n)} < \cdots < \theta_{n-1}^{(n-1)} < \theta_n^{(n)},$$

then there is unique Jacobi matrix T_n with the eigenvalues $\theta_j^{(n)}, j = 1, \ldots, n$, and $\theta_j^{(n-1)}, j = 1, \ldots, n-1$, as eigenvalues of T_{n-1}. Similar or related results were published by many authors, cf. [Gantmakher–50], [Atkinson–64], [Wendroff–61], [Hochstadt–74], [Hald–76], [deBoor–78], [Hochstadt–79], [Deift-84], [Gragg–84] and [deBoor–86]. Scott used the result to find the initial vector for which the orthogonality among Lanczos vectors is well preserved in the real Lanczos run and gave thus a nice application of the Paige's theory.

The correspondence of Jacobi matrices and orthogonal polynomials was exploited by de Boor and Golub, who gave first (to our knowledge) the simple and elegant proof of these results and pointed out the connection to the Lanczos method [de Boor–78], [Golub–83]. We recall some other results presented in these works.

The vector of weights

$$m^{(n)} = (m_1^{(n)}, \ldots, m_n^{(n)})^T = ((s_{11}^{(n)})^2, \ldots, (s_{1n}^{(n)})^2)^T \tag{3.1}$$

is the solution of the linear system

$$V^{(n)} m^{(n)} = (1, (T_n)_{11}, (T_n^2)_{11}, \ldots, (T_n^{n-1})_{11})^T \tag{3.2}$$

$$\text{where } V^{(n)} = \begin{pmatrix} 1 & \ldots & 1 \\ \theta_1^{(n)} & & \theta_n^{(n)} \\ \vdots & & \vdots \\ (\theta_1^{(n)})^{n-1} & \ldots & (\theta_n^{(n)})^{n-1} \end{pmatrix},$$

is the Vandermonde matrix based on $\theta_1^{(n)}, \ldots, \theta_n^{(n)}$, X_{11} is the $(1,1)$ element of the matrix X. Indeed, $(T_n^k)_{11} = ((S_n \operatorname{diag}(\theta_j^{(n)}) S_n^T)^k)_{11} = (S_n (\operatorname{diag}(\theta_j^{(n)}))^k S_n^T)_{11} = \sum_{j=1}^{n} (\theta_j^{(n)})^k (s_{1j}^{(n)})^2$. Using

$$(T_n^k)_{11} = (T_n^{\lfloor k/2 \rfloor})_{1.} (T_n^{\lfloor (k+1)/2 \rfloor})_{.1},$$

where $X_{1.}$ and $X_{.1}$ denotes the first row and column of the matrix X, $\lfloor y \rfloor$ is the largest integer $\leq y$, it is not hard to see, that the value $(T_n^k)_{11}$ is thus determined by the values $\alpha_1, \alpha_2, \ldots, \alpha_{\lfloor \frac{k+1}{2} \rfloor}$, $\beta_2, \beta_3, \ldots, \beta_{\lfloor \frac{k+1}{2} \rfloor}$, and the right side in (3.2) is determined by $\alpha_1, \ldots, \alpha_{\lfloor n/2 \rfloor}$, $\beta_2, \ldots, \beta_{\lfloor n/2 \rfloor}$. We want explicit formulas for the weights (3.1). From (1.8)–(1.10)

$$\psi_{2,n}(\theta_j^{(n)}) \psi_{n-1}(\theta_j^{(n)}) = \beta^{(n)}, \quad \beta^{(n)} = \prod_{r=2}^{n} \beta_r^2, \quad j = 1, 2, \ldots, n, \tag{3.3}$$

$$m_j^{(n)} = (s_{1j}^{(n)})^2 = \frac{\psi_{2,n}(\theta_j^{(n)})}{\psi_n'(\theta_j^{(n)})} = \frac{\beta^{(n)}}{\psi_{n-1}(\theta_j^{(n)}) \psi_n'(\theta_j^{(n)})}, \tag{3.4}$$

and the orthonormality condition $\sum_{j=1}^{n} m_j^{(n)} = 1$ gives

$$\beta^{(n)} = \left[\sum_{j=1}^{n} (\psi_{n-1}(\theta_j^{(n)}) \psi_n'(\theta_j^{(n)}))^{-1} \right]^{-1}. \tag{3.5}$$

(3.4)–(3.5) was proved directly in [de Boor–78]. These relations determine the weights in (1.19) from the Ritz values at steps $n-1$ and n. The analogy of (3.4) for the squared bottom elements of the normalized eigenvectors can be found quite analogously

$$(s_{nj}^{(n)})^2 = \frac{\psi_{n-1}(\theta_j^{(n)})}{\psi_n'(\theta_j^{(n)})} = \frac{\beta^{(n)}}{\psi_{2,n}(\theta_j^{(n)})\psi_n'(\theta_j^{(n)})} \tag{3.6}$$

or

$$\delta_{nj}^2 = \frac{\prod_{r=2}^{n+1} \beta_r^2}{\psi_{2,n}(\theta_j^{(n)})\psi_n'(\theta_j^{(n)})}. \tag{3.7}$$

Inversely, explicit formulas for polynomials $\psi_1, \psi_2, \ldots, \psi_{n-1}$ based on $s_{1j}^{(n)}, \theta_j^{(n)}, j = 1, 2, \ldots, n$, were given by Cybenko [Cybenko–87].

Next result gives a relation among the innerproducts (1.19), cf. Lemma 5.9 of the paper [van der Sluis–87].

THEOREM 3.1. *Using the previous notation, the weights $m_r^{(k)}$, $r = 1, 2, \ldots, k$, in the k-th innerproduct, $1 \le k < n$, defined by the Lanczos run (exact or finite precision) for A, q^1, are related to the values $m_j^{(n)}, \theta_j^{(n)}, j = 1, 2, \ldots, n$, $\theta_r^{(k)}, r = 1, 2, \ldots, k$, by the formula*

$$m_r^{(k)} = \frac{1}{\psi_k'(\theta_r^{(k)})} \sum_{j=1}^{n} m_j^{(n)} \frac{\psi_k(\theta_j^{(n)})}{\theta_j^{(n)} - \theta_r^{(k)}} = \frac{1}{(\psi_k'(\theta_r^{(k)}))^2} \sum_{j=1}^{n} m_j^{(n)} \left(\frac{\psi_k(\theta_j^{(n)})}{\theta_j^{(n)} - \theta_r^{(k)}} \right)^2. \tag{3.8}$$

Especially, for $k = n-1$

$$\begin{aligned} m_r^{(n-1)} &= \frac{1}{\psi_{n-1}'(\theta_r^{(n-1)})} \sum_{j=1}^{n} \frac{\beta^{(n)}}{\psi_n'(\theta_j^{(n)})(\theta_j^{(n)} - \theta_r^{(n-1)})} = \\ &= \frac{\beta^{(n)}}{(\psi_{n-1}'(\theta_r^{(n-1)}))^2} \sum_{j=1}^{n} \frac{\psi_{n-1}(\theta_j^{(n)})}{\psi_n'(\theta_j^{(n)})(\theta_j^{(n)} - \theta_r^{(n-1)})^2}. \end{aligned} \tag{3.9}$$

Proof. (3.9) follows from (3.8), (3.4). We denote

$$\omega_{kr}(\lambda) = \prod_{\ell \ne r} (\lambda - \theta_\ell^{(k)}) = \frac{\psi_k(\lambda)}{(\lambda - \theta_r^{(k)})}. \tag{3.10}$$

Clearly $\omega_{kr}(\theta_r^{(k)}) = \psi_k'(\theta_r^{(k)})$. Let $p(\lambda)$ be any polynomial of degree at most $2k-1$.

Then we could write $p(\lambda)$ in the expansion form

$$p(\lambda) = \psi_k(\lambda)\tilde{p}(\lambda) + \tilde{\tilde{p}}(\lambda), \quad \tilde{p}(\lambda), \tilde{\tilde{p}}(\lambda) \text{ has degree at most } k-1,$$
$$\tilde{\tilde{p}}(\lambda) = \sum_{r=1}^{k} \nu_r \omega_{kr}(\lambda), \quad \nu_r = \frac{\tilde{\tilde{p}}(\theta_r^{(k)})}{\omega_{kr}(\theta_r^{(k)})} = \frac{p(\theta_r^{(k)})}{\omega_{kr}(\theta_r^{(k)})}.$$

Thus, using the fact that $\psi_k(\lambda)$ is orthogonal to $\tilde{p}(\lambda)$ with respect to the n-th innerproduct

$$\sum_{j=1}^{n} m_j^{(n)} p(\theta_j^{(n)}) = \sum_{j=1}^{n} m_j^{(n)} \tilde{\tilde{p}}(\theta_j^{(n)}) = \sum_{j=1}^{n} m_j^{(n)} \sum_{r=1}^{k} \frac{p(\theta_r^{(k)})}{\omega_{kr}(\theta_r^{(k)})} \omega_{kr}(\theta_j^{(n)}) =$$
$$= \sum_{r=1}^{k} \left[\frac{1}{\psi_k'(\theta_r^{(k)})} \sum_{j=1}^{n} m_j^{(n)} \frac{\psi_k(\theta_j^{(n)})}{(\theta_j^{(n)} - \theta_r^{(k)})} \right] p(\theta_r^{(k)}) = \sum_{r=1}^{k} m_r^{(n,k)} p(\theta_r^{(k)})$$

Moreover, $m_r^{(n,k)} > 0$, $r = 1, 2, \ldots, k$, $\sum_{r=1}^{k} m_r^{(n,k)} = 1$. To show this, we use the expansion $\omega_{k,r}(\lambda) = \omega_{k,r}(\theta_r^{(k)}) + (\lambda - \theta_r^{(k)})\kappa(\lambda)$, κ is polynomial of degree at most $k-2$. Then

$$\sum_{j=1}^{n} m_j^{(n)} \omega_{k,r}^2(\theta_j^{(n)}) = \omega_{k,r}(\theta_r^{(k)}) \sum_{j=1}^{n} m_j^{(n)} \omega_{k,r}(\theta_j^{(n)}) + \sum_{j=1}^{n} m_j^{(n)} \psi_k(\theta_j^{(n)}) \kappa(\theta_j^{(n)}),$$

but the last term vanishes because of orthogonality property. Then

$$m_r^{(n,k)} = \frac{1}{\omega_{k,r}(\theta_r^{(k)})} \sum_{j=1}^{n} m_j^{(n)} \omega_{k,r}(\theta_j^{(n)}) = \frac{1}{(\omega_{k,r}(\theta_r^{(k)}))^2} \sum_{j=1}^{n} m_j^{(n)} (\omega_{k,r}(\theta_j^{(n)}))^2 > 0.$$
$$\sum_{r=1}^{k} m_r^{(n,k)} = \sum_{r=1}^{k} m_r^{(n,k)} \psi_0(\theta_r^{(k)}) = \sum_{j=1}^{n} m_j^{(n)} \psi_0(\theta_j^{(k)}) = \sum_{j=1}^{n} m_j^{(n)} = 1,$$

where $\psi_0(1) \equiv 1$.

To finish the proof we need to show that $m_r^{(n,k)} \equiv m_r^{(k)}$. But the set of monic polynomials $1, \psi_1, \ldots, \psi_{k-1}$ is orthogonal with respect to both innerproducts defined by points $\{\theta_r^{(k)}\}_{r=1}^{k}$ and weights $\{m_r^{(k)}\}$ or $\{m_r^{(n,k)}\}$. This implies $m_r^{(k)} \equiv m_r^{(n,k)}$, $r = 1, 2, \ldots, k$, which can be directly proved as follows. The weights $m_r^{(*)}$, where $(*)$ is either (k) or (n,k), satisfy the relations

$$\sum_{r=1}^{k} m_r^{(*)} = 1$$
$$\sum_{r=1}^{k} m_r^{(*)} \psi_j(\theta_r^{(k)}) = 0, \quad j = 1, 2, \ldots k-1,$$

which can be written in the matrix form as

$$Xm^{(*)} = e^1, \tag{3.11}$$

where $m^{(*)} = (m_1^{(*)}, \ldots, m_k^{(*)})^T$, $e^1 = (1, 0, \ldots, 0)^T$,

$$X = \begin{pmatrix} 1 & \ldots & 1 \\ \psi_1(\theta_1^{(k)}) & \ldots & \psi_1(\theta_k^{(k)}) \\ \vdots & & \vdots \\ \psi_{k-1}(\theta_1^{(k)}) & \ldots & \psi_{k-1}(\theta_k^{(k)}) \end{pmatrix} \tag{3.12}$$

The matrix X is nonsingular (using row operations X can be transformed into the Vandermonde matrix). The difference vector $m^{(n,k)} - m^{(k)}$ thus solves homogeneous system with nonsingular matrix and must vanish. □

COROLLARY 3.1. *For any Jacobi matrix T_n which has T_k as its leading principal submatrix, $k = 1, 2, \ldots, n-1$,*

$$m_j^{(n)}(\psi_k(\theta_j^{(n)}))^2 \leq (\psi_k'(\theta_r^{(k)}))^2(\theta_j^{(n)} - \theta_r^{(k)})^2, \quad j = 1, 2, \ldots, n, r = 1, 2, \ldots, k. \tag{3.13}$$

Proof. From Theorem 3.1

$$\frac{1}{(\omega_{kr}(\theta_r^{(k)}))^2} \sum_{j=1}^{n} m_j^{(n)}(\omega_{kr}(\theta_j^{(n)}))^2 = m_r^{(k)} < 1$$

which gives

$$m_j^{(n)}(\omega_{kr}(\theta_j^{(n)}))^2 < (\omega_{kr}(\theta_r^{(k)}))^2$$

and multiplying both sides by $(\theta_j^{(n)} - \theta_r^{(k)})^2$ finishes the proof. □

One must be careful in the interpretation of Corollary 3.1. If there are two eigenvalues of T_k which forms a close pair, then for any eigenvalue $\theta_j^{(n)}$ of any Jacobi matrix which has T_k as its leading principal submatrix either $(\psi_k(\theta_j^{(n)}))^2$ is "as small as $(\psi_k'(\theta_r^{(k)}))^2(\theta_j^{(n)} - \theta_r^{(k)})^2$", or the weight $m_j^{(n)}$ is small. This is worthy of two comments.

First, if e.g. $\theta_r^{(k)}$ and $\theta_{r+1}^{(k)}$ form a close pair, then one intuitively expects $|\psi_k'(\theta_r^{(k)})| \ll 1$. But ψ_k is a monic polynomial of degree k and thus the value $|\psi_k'(\theta_r^{(k)})|$ can be very large. It may be relatively small, of course, in comparison e.g. to $|\psi_k'(\theta_{r-1}^{(k)})|$ in the case that $\theta_{r-1}^{(k)}$ is well separated from other eigenvalues.

Second, if $\theta_j^{(n)}$ was not approached by any $\theta_i^{(k)}, i = 1, 2, \ldots, k$, and there is a close pair $\theta_r^{(k)}, \theta_{r+1}^{(k)}$, then one could vaguely expect $|\psi_k'(\theta_j^{(n)})| \gg |\psi_k'(\theta_r^{(k)})|$ which implies $m_j^{(n)} \ll 1$

(supposing $|\theta_j^{(n)} - \theta_r^{(k)}| \leqq o(1)$). This conclusion is wrong, of course, because even in this case it can be $|\psi_k'(\theta_j^{(k)})| \ll |\psi_k'(\theta_r^{(k)}|$ due to inhomogeneity in the eigenvalue distribution. This shows very clearly, that the phenomena described here are determined by the distribution of *all eigenvalues* and one must be very careful in using arguments based on the local information only.

A simple but interesting relation can be derived directly from (1.8)-(1.10) and (3.3)

$$(3.14) \qquad (s_{1j}^{(n)})^2 (s_{nj}^{(n)})^2 = \frac{\beta^{(n)}}{\psi_n'(\theta_j^{(n)})}, \quad j = 1, 2, \ldots, n,$$

which gives

$$(3.15) \qquad \frac{m_j^{(n)}}{m_i^{(n)}} = \frac{\delta_{ni}^2}{\delta_{nj}^2} \left[\frac{\psi_n'(\theta_i^{(n)})}{\psi_n'(\theta_j^{(n)})} \right]^2 \quad 1 \leq i, j \leq n.$$

If now e.g. $\delta_{ni} \ll \delta_{nj}$ than either $|\psi_n'(\theta_j^{(n)})| \ll |\psi_n'(\theta_i^{(n)})|$, or $m_j^{(n)} \ll m_i^{(n)}$. We warn again that the values $\psi_n'(\theta_j^{(n)})$, $\psi_n'(\theta_i^{(n)})$ depend strongly on the global eigenvalue distribution and on the position of $\theta_j^{(n)}, \theta_i^{(n)}$ in the spectra. It cannot be simply stated without any additional assumption that e.g. $\theta_i^{(n)}$ well separated imply $\psi_n'(\theta_i^{(n)}) = o(1)$.

(3.6)-(3.15) show that stabilization and clustering of Ritz values affect the weight of every other eigenvalue in the innerproduct (1.19). The strength of these connections are determined by the global eigenvalue distribution.

Previous results can be entirely described in the language of Jacobi matrices. In the rest of this section we present a few results concerning the Lanczos algorithm and its run in finite precision arithmetic. Subtracting

$$u_i^T \{ A z_j^{(k)} - \theta_j^{(k)} z_j^{(k)} = \beta_{k+1} s_{kj}^{(k)} q^{k+1} \}$$

and

$$(z_j^{(k)})^T \{ A u_i - \lambda_i u_i = 0 \}$$

gives useful result

$$(3.16) \qquad (\lambda_i - \theta_j^{(k)})(u_i, z_j^{(k)}) = \beta_{k+1} s_{kj}^{(k)} (u_i, q^{k+1}).$$

For the finite precision Lanczos run (3.16) is modified to

$$(3.17) \qquad (\lambda_i - \theta_j^{(k)})(u_i, z_j^{(k)}) = \beta_{k+1} s_{kj}^{(k)} (u_i, q^{k+1}) + u_i^T F_k s_j^{(k)},$$

where $\|u_i^T F_k s_j^{(k)}\| \le \varepsilon_4$, $\varepsilon_4 = k^{1/2}\|A\|\varepsilon_1$, see Introduction.

LEMMA 3.1. *For any step k of the exact or finite precision Lanczos process applied to A, q^1*

$$\sum_{j=1}^{k} s_{1j}^{(k)} z_j^{(k)} = q^1, \tag{3.18}$$

$$m_i^{1/2} = |(u_i, q^1)| = |\sum_{j=1}^{k} s_{1j}^{(k)} (u_i, z_j^{(k)})|. \tag{3.19}$$

Proof.

$$\sum_{j=1}^{k} s_{1j}^{(k)} z_j^{(k)} = Q_k \left(\sum_{j=1}^{k} s_{1j}^{(k)} s_j^{(k)} \right) = Q_k S_k s_0 = Q_k e_1 = q^1.$$

(3.19) immediately follows. □

(3.16)-(3.19) relates the weights in the original innerproduct to the weights in the k-th innerproduct and to the convergence of Ritz vectors. Next two lemmas assume exact arithmetic.

LEMMA 3.2. *For any step k of the Lanczos algorithm applied to A, q^1*

$$\sum_{j=1}^{k} \frac{\psi_k(\lambda_i)}{(\lambda_i - \theta_j^{(k)})\psi_k'(\theta_j^{(k)})} = 1. \tag{3.20}$$

Proof. Using (3.18), (3.16) and (3.14)

$$(u_i, q^1) = \sum_{j=1}^{k} s_{1j}^{(k)} (u_i, z_j^{(k)}) = \sum_{j=1}^{k} \frac{\beta_{k+1} s_{kj}^{(k)} s_{1j}^{(k)}}{\lambda_i - \theta_j^{(k)}} (u_i, q^{k+1}) =$$

$$= \sum_{j=1}^{k} \frac{\prod_{r=2}^{k+1} \beta_r}{(\lambda_i - \theta_j^{(k)})\psi_k'(\theta_j^{(k)})} (u_i, q^{k+1}).$$

From (1.18)

$$(u_i, q^{k+1}) = (u_i, \psi_k(A) q^1) / \prod_{r=2}^{k+1} \beta_r = (\psi_k(\lambda_i) u_i, q^1) / \prod_{r=2}^{k+1} \beta_r$$

which gives

$$(u_i, q^1) = (u_i, q^1) \sum_{j=1}^{k} \frac{\psi_k(\lambda_i)}{(\lambda_i - \theta_j^{(k)})\psi_k'(\theta_j^{(k)})}. \quad \square$$

Lemma 3.3 gives an analogy of (1.18) for Ritz vectors.

LEMMA 3.3. *The Ritz vectors determined at the step k of the Lanczos algorithm applied to A, q^1 are given by*

$$z_j^{(k)} = \operatorname*{sgn}_{kj} w_j^{(k)} / \|w_j^{(k)}\|, \quad w_j^{(k)} = \omega_{kj}(A)q^1, \quad \|w_j^{(k)}\| = \prod_{r=2}^{k} \beta_r / |s_{kj}^{(k)}| \tag{3.21}$$

where ω_{kj} is defined by (3.10) and $\operatorname*{sgn}_{kj} = s_{kj}^{(k)} / |s_{kj}^{(k)}|$, *or equivalently by*

$$z_j^{(k)} = \frac{s_{kj}^{(k)}}{\prod_{r=2}^{k} \beta_r} \sum_{i=1}^{N} (u_i, q^1)\omega_{kj}(\lambda_i)u_i = \frac{1}{s_{1j}^{(k)}\psi_k'(\theta_j^{(k)})} \sum_{i=1}^{n} (u_i, q^1)\omega_{kj}(\lambda_i)u_i \tag{3.22}$$

Proof. From [Parlett–80], Corollary (12-3-7)

$$\tilde{z}_j^{(k)} = \operatorname*{sgn}_{kj} z_j^{(k)} = w_j^{(k)} / \|w_j^{(k)}\|, \quad w_j^{(k)} = \omega_{kj}(A)q^1.$$

Then $(u_i, \tilde{z}_j^{(k)}) = \frac{\omega_{kj}(\lambda_i)}{\|w_j^{(k)}\|}(u_i, q^1)$. But from (3.16), (1.18)

$$(u_i, z_j^{(k)}) = \frac{\beta_{k+1} s_{kj}^{(k)}}{\lambda_i - \theta_j^{(k)}}(u_i, q^{k+1}) = \frac{s_{kj}^{(k)}}{\prod_{r=2}^{k} \beta_r}\omega_{kj}(\lambda_i)(u_i, q^1)$$

which comparing to the previous relation gives

$$\operatorname*{sgn}_{kj} \|w_j^{(k)}\| = \frac{\prod_{r=2}^{k} \beta_r}{s_{kj}^{(k)}} = \frac{1}{s_{1j}^{(k)}\psi_k'(\theta_j^{(k)})}. \quad \square$$

Next bound can be found e.g. in Paige's thesis [Paige–71]. The proof is very instructive, that is why it is included.

LEMMA 3.4. *For any Ritz vector* $z_j^{(k)}$, $j = 1, 2, \ldots, k$, *at any step* k *of the finite precision Lanczos process applied to* A, q^1

$$\|z_j^{(k)} - \xi_{ij}^{(k)} u_i\| \le \frac{\delta_{kj} + \varepsilon_4}{\min\limits_{r \ne i} |\lambda_r - \theta_j^{(k)}|} \tag{3.23}$$

where $\xi_{ij}^{(k)} = (z_j^{(k)}, u_i)$, $i = 1, 2, \ldots N$.

Proof. Let $z_j^{(k)} = \sum\limits_{i=1}^{N} \xi_{ij}^{(k)} u_i$ be the expansion of $z_j^{(k)}$ to the matrix A eigenvectors.

Then

$$\|z_j^{(k)} - \xi_{ij}^{(k)} u_i\| = \left[\sum_{r \ne i} (\xi_{rj}^{(k)})^2\right]^{1/2} \le \frac{\left[\sum\limits_{r \ne i} ((\lambda_r - \theta_j^{(k)}) \xi_{rj}^{(k)})^2\right]^{1/2}}{\min\limits_{r \ne i} |\lambda_r - \theta_j^{(k)}|} \le$$

$$\le \frac{\| \sum\limits_{r=1}^{N} \lambda_r \xi_{rj} u_r - \theta_j^{(k)} \sum\limits_{r=1}^{N} \xi_{rj} u_r \|}{\min\limits_{r \ne i} (\lambda_r - \theta_j^{(k)}|} = \frac{\|A z_j^{(k)} - \theta_j^{(k)} z_j^{(k)}\|}{\min\limits_{r \ne i} |\lambda_r - \theta_j^{(k)}|}$$

$$\le \frac{\delta_{kj} + \varepsilon_4}{\min\limits_{r \ne i} |\lambda_r - \theta_j^{(k)}|}. \quad \square$$

For the cluster of Ritz values, relation (3.24) in the following lemma complements (1.7).

LEMMA 3.5. *For any step* k *of the finite precision Lanczos process applied to* A, q^1

$$\delta_{kj} |(z_r^{(k)}, q^{k+1})| \le k |\theta_r^{(k)} - \theta_j^{(k)}| + \varepsilon_3, \quad 1 \le r, j \le k, r \ne j. \tag{3.24}$$

Proof. Let R_k be the strictly upper triangular matrix defined by

$$Q_k^T Q_k = R_k^T + \mathrm{diag}((q^i, q^i)) + R_k,$$

then

$$T_k R_k - R_k T_k = \beta_{k+1} Q_k^T q^{k+1} e^k + \delta R_k,$$

where $\|\delta R_k\| \le \|\delta R_k\|_F \le \varepsilon_3$, [Paige–76]. Multiplying by $(s_r^{(k)})^T$ from the left and $s_j^{(k)}$ from the right gives

$$(\theta_r^{(k)} - \theta_j^{(k)}) (s_r^{(k)})^T R_k s_j^{(k)} = \beta_{k+1} s_{ij}^{(k)} (z_r^{(k)}, q^{k+1}) + \varepsilon_{ij}^{(k)},$$

where $|\varepsilon_{ij}^{(k)}| \leq \varepsilon_3$ and considering $\|R_k\| \leq \|R_k\|_F \leq k$, the proof is finished. □

Lemma 3.5 says that if $\theta_j^{(k)}$ and $\theta_r^{(k)}$ form a close pair (whether this pair is separated from other Ritz values or not), then at least one value of δ_{kj}, $(z_r^{(k)}, q^{k+1})$ must be small (the property is of course symmetric). It will be used in Section 6. When $\|z_r^{(k)}\|$ is "small", which is the troubled case in the Paige's analysis, then δ_{kj} must be also "small", because from (1.3) and using (3.48) in [Paige–80]

$$\delta_{kj} \leq (\|Az_j^{(k)} - \theta_j^{(k)} z_j^{(k)}\| + \varepsilon_4)/\|q^{k+1}\| \leq 2\|A\| \ \|z_j^{(k)}\| + 2k^3\|A\|\varepsilon_2. \tag{3.25}$$

Finally, we recall the well known result describing the angles among Ritz vectors (see e.g. [Parlett–80], 13.4.7)

$$(\theta_r^{(k)} - \theta_j^{(k)})(z_r^{(k)}, z_j^{(k)}) = \beta_{k+1} s_{kj}^{(k)} (z_r^{(k)}, q^{k+1}) - \beta_{k+1} s_{kr}^{(k)} (z_j^{(k)}, q^{k+1}) + o(\varepsilon_3). \tag{3.26}$$

If $\theta_r^{(k)}$ is separated from $\theta_j^{(k)}$, the corresponding Ritz vectors do not vanish and both $s_{kj}^{(k)}$, $s_{kr}^{(k)}$ are "sufficiently small" or "sufficiently large", then the corresponding Ritz vectors are approximately orthogonal.

4. Numerical experiments. Experimental results discussed in Sections 5–8 were obtained for diagonal matrix $A = \mathrm{diag}(\lambda_i)$ and randomly generated initial vector r^0. Following [Strakoš–91, Greenbaum–92] we use

$$N = 24, \quad \lambda_1 = 0.1, \quad \lambda_N = 100,$$
$$\lambda_\nu = \lambda_1 + \frac{\nu - 1}{N - 1}(\lambda_N - \lambda_1) \cdot \rho^{N-\nu}, \quad \nu = 2, 3, \ldots, N-1,$$

where parameter $\rho \in \langle 0.5, 1 \rangle$. For some values of ρ (e.g. $\rho \in \langle 0.6, 0.8 \rangle$) the orthogonality is lost very quickly even for a very small dimension N. Our experiments were performed using Matlab, machine precision was $\varepsilon = 2.22 \times 10^{-16}$. Real runs of algorithm (1.1) were compared with the "exact" Lanczos run, i.e. process including double reorthogonalization of the each Lanczos vector against all previously computed ones [Greenbaum–92]. For each step i we monitored the values of $\beta_{i+1}, \prod_{j=2}^{i+1} \beta_j$, $\max_j |\psi_i'(\theta_j^{(i)})|$, $\|r^i\|/\|r^0\|$, $\max_j \|z_j^{(i)}\|$, which are on the figures denoted as bt, pbt, d, r and z, $\|r^i\|$ is the norm of the i-th residual for the corresponding conjugate gradient process solving

$$Ax = r^0, x^0 = 0.$$

Results of our computations will be presented in a graphic form as shown in Figure 4.1. It describes the step 30 of the Lanczos process for $\rho = 0.8$. The figure window is divided into two fields, separated by the virtual line $y = 10^{-9}$. There are two rows of cross-points in the lower field – the lower shows the distribution of original eigenvalues λ_r, the upper the distribution of Ritz values computed in step i. In the upper field there are four points above each Ritz value $\theta_j^{(i)}$:

$$
\begin{array}{ll}
* & \text{denotes } \|z_j^{(i)}\| / \max\limits_j \|z_j^{(i)}\| \\
o & \text{denotes } |s_{1j}^{(i)}| \\
+ & \text{denotes } |s_{ij}^{(i)}| \\
\cdot & \text{denotes } |\psi_i'(\theta_j^{(i)})| / \max\limits_j |\psi_i'(\theta_i^{(i)})|,
\end{array}
$$

values less then 10^{-9} are set to 10^{-9}. If there were two or three Ritz values approaching the same eigenvalue λ_r, then there would be a pair or a triple of the corresponding points above λ_r.

Figure 4.1 shows Ritz values in different stages of convergence. There are clusters of three Ritz values near λ_{24} and λ_{23} (at least one copy still does not have a small value of δ_{kj}), two copies approximating λ_{22}, first copy perturbed by the second copy approximating λ_{20} (c.f. [Parlett–80], p. 270). There are Ritz values with a moderate size of δ_{kj} (near $\lambda_{16}, \lambda_8 - \lambda_{13}$). Some of the eigenvalues are still not approximated by any Ritz value (e.g. λ_7). An example of the cluster development is shown in Figure 4.2, where crosspoints denote the position of Ritz values near the eigenvalue $\lambda_{19} = 13.24$. We see the "moving" of Ritz values caused by the forming of multiple copies near previously approximated eigenvalues or discovering the hidden eigenvalues (cf. [Parlett–81], graph of "bottom-pivot function" δ_j).

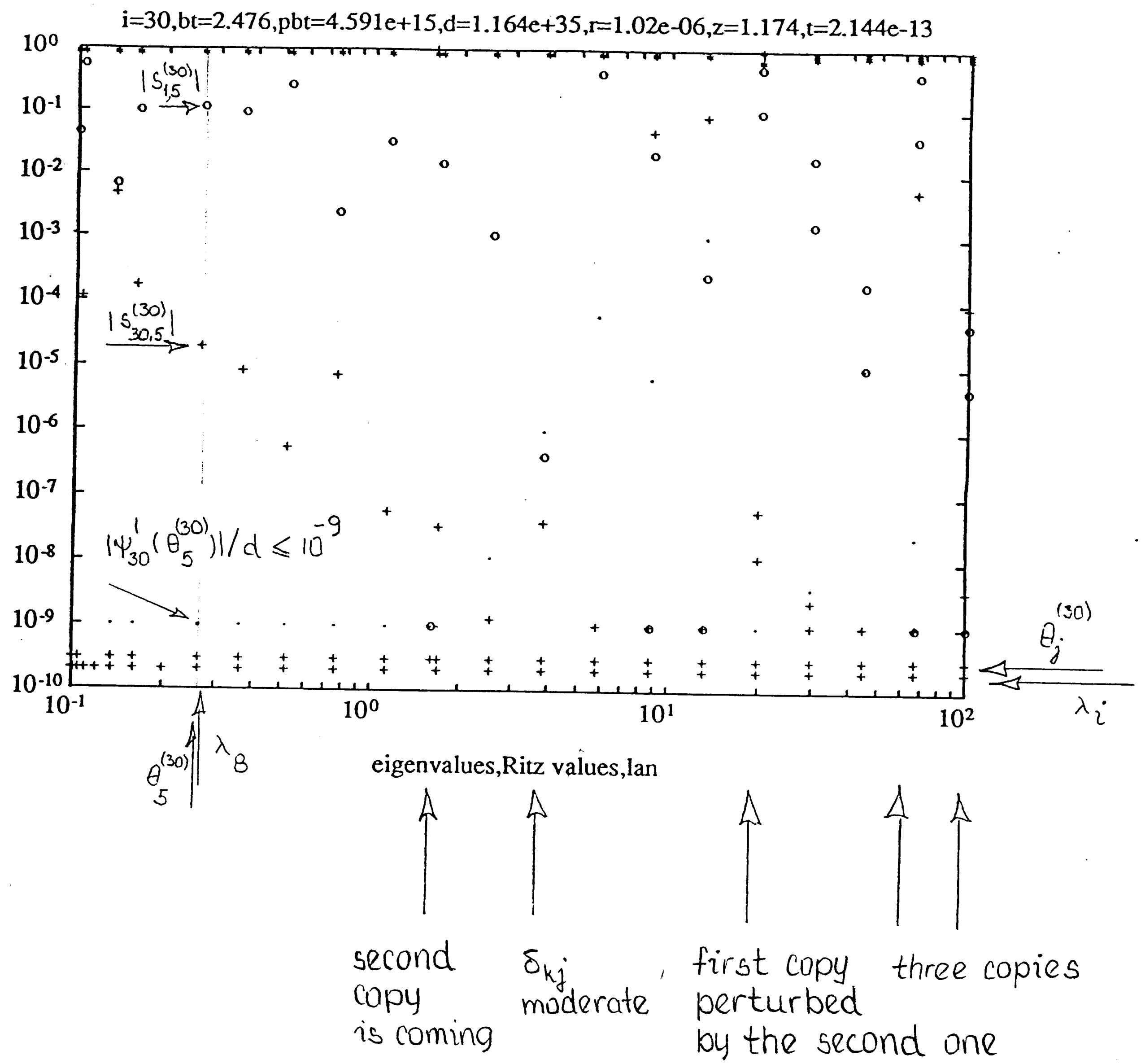
i=30,bt=2.476,pbt=4.591e+15,d=1.164e+35,r=1.02e-06,z=1.174,t=2.144e-13
$|s_{1,5}^{(30)}|$
$|s_{30,5}^{(30)}|$
$|\psi_{30}^{1}(\theta_5^{(30)})|/d \leq 10^{-9}$
$\theta_j^{(30)}$
λ_i
$\theta_5^{(30)}$
λ_8
eigenvalues,Ritz values,lan
second copy is coming
δ_{kj} moderate
first copy perturbed by the second one
three copies

Figure 4.1

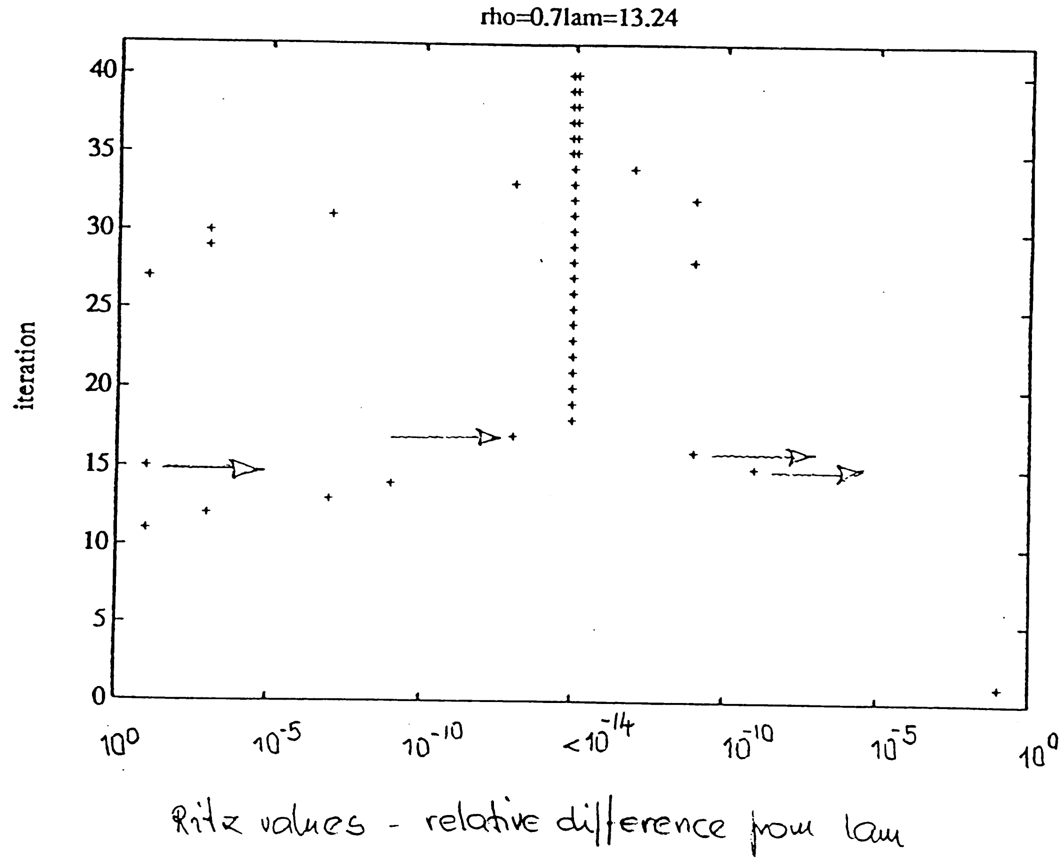

Figure 4.2

5. Approximation bounds based on Ritz values in several consecutive steps. This section gives bounds for δ_{kj} based on Ritz values in two or three consecutive steps.

First the upper and lower bounds for the bottom elements of the eigenvectors of Jacobi matrices is given (cf. [Hill–91], [Strakoš–90]).

LEMMA 5.1. *For any Jacobi matrix T_k*

$$(s_{k1}^{(k)})^2 \leq \frac{\theta_1^{(k)} - \theta_1^{(k-1)}}{\theta_1^{(k)} - \theta_2^{(k)}} \tag{5.1a}$$

$$(s_{kj}^{(k)})^2 \leq \frac{\theta_j^{(k)} - \theta_{j-1}^{(k-1)}}{\theta_j^{(k)} - \theta_{j-1}^{(k)}} \cdot \frac{\theta_j^{(k)} - \theta_j^{(k-1)}}{\theta_j^{(k)} - \theta_{j+1}^{(k)}} \qquad j = 2, \ldots, k-1 \tag{5.1b}$$

$$(s_{kk}^{(k)})^2 \leq \frac{\theta_k^{(k)} - \theta_{k-1}^{(k-1)}}{\theta_k^{(k)} - \theta_{k-1}^{(k)}} \tag{5.1c}$$

Proof. Immediately follows from rewriting (1.10) in terms of $\theta_j^{(k)}$ and $\theta_j^{(k-1)}$. $\square$

LEMMA 5.2. *For any Jacobi matrix T_k*

$$(5.2a)\qquad (s_{k1}^{(k)})^2 \geq \frac{\theta_1^{(k)} - \theta_1^{(k-1)}}{\sigma_k}$$

$$(5.2b)\qquad (s_{kj}^{(k)})^2 \geq \frac{4}{\sigma_k^2}(\theta_j^{(k)} - \theta_{j-1}^{(k-1)})(\theta_j^{(k-1)} - \theta_j^{(k)}), \qquad j = 2, \ldots, k-1$$

$$(5.2c)\qquad (s_{kk}^{(k)})^2 \geq \frac{\theta_k^{(k)} - \theta_{k-1}^{(k-1)}}{\sigma_k},$$

where $\sigma_k = \theta_k^{(k)} - \theta_1^{(k)}$.

Proof. For $j = 2, \ldots, k-1$

$$(s_{kj}^{(k)})^2 = \prod_{r=1}^{j-1} \frac{\theta_j^{(k)} - \theta_r^{(k-1)}}{\theta_j^{(k)} - \theta_r^{(k)}} \prod_{r=j+1}^{k} \frac{\theta_j^{(k)} - \theta_{r-1}^{(k-1)}}{\theta_j^{(k)} - \theta_r^{(k)}} =$$

$$= \frac{1}{\theta_j^{(k)} - \theta_1^{(k)}} \left[\prod_{r=1}^{j-2} \frac{\theta_j^{(k)} - \theta_r^{(k-1)}}{\theta_j^{(k)} - \theta_{r+1}^{(k)}}\right] (\theta_j^{(k)} - \theta_{j-1}^{(k-1)})(\theta_j^{(k)} - \theta_j^{(k-1)})$$

$$\left[\prod_{r=j+1}^{k-1} \frac{\theta_j^{(k)} - \theta_r^{(k-1)}}{\theta_j^{(k)} - \theta_r^{(k)}}\right] \frac{1}{\theta_j^{(k)} - \theta_k^{(k)}}$$

Using the interlacing property and $|(\theta_j^{(k)} - \theta_1^{(k)})(\theta_j^{(k)} - \theta_k^{(k)})| \leq \frac{1}{4}\sigma_k^2$ finishes the proof (the special cases are trivial). □

Lemma 5.2 is used for proving next result.

THEOREM 5.1. *For any Jacobi matrix $T_{k+1}, k > 1$*

$$\delta_{k1} \leq \sigma_{k+1}^{1/2}(\theta_1^{(k)} - \theta_1^{(k+1)})^{1/2} \tag{5.3a}$$

$$\delta_{kk} \leq \sigma_{k+1}^{1/2}(\theta_{k+1}^{(k+1)} - \theta_k^{(k)})^{1/2} \tag{5.3b}$$

$$\delta_{kj} \leq \frac{\sigma_{k+1}}{2}\left(\frac{\theta_j^{(k)} - \theta_j^{(k+1)}}{\theta_j^{(k+1)} - \theta_{j-1}^{(k)}}\right)^{1/2} \qquad j = 2, \ldots, k-1, \tag{5.3c}$$

$$\delta_{kj} \leq \frac{\sigma_{k+1}}{2}\left(\frac{\theta_{j+1}^{(k+1)} - \theta_j^{(k)}}{\theta_{j+1}^{(k)} - \theta_{j+1}^{(k+1)}}\right)^{1/2} \qquad j = 2, \ldots, k-1, \tag{5.3d}$$

and

(5.4a)

$$\delta_{k1} \leq \frac{\beta_{k+1}}{\beta_k}\sigma_{k-1}^{1/2}\frac{\theta_1^{(k-1)} - \theta_1^{(k)}}{(\theta_1^{(k-2)} - \theta_1^{(k-1)})^{1/2}}$$

(5.4b)

$$\delta_{kk} \leq \frac{\beta_{k+1}}{\beta_k}\sigma_{k-1}^{1/2}\frac{\theta_k^{(k)} - \theta_{k-1}^{(k-1)}}{(\theta_{k-1}^{(k-1)} - \theta_{k-2}^{(k-2)})^{1/2}}$$

(5.4c)

$$\delta_{kj} \leq \frac{\beta_{k+1}}{\beta_k}\frac{\sigma_{k-1}}{2}\frac{\theta_j^{(k-1)} - \theta_j^{(k)}}{[(\theta_j^{(k-1)} - \theta_{j-1}^{(k-2)})(\theta_j^{(k-2)} - \theta_j^{(k-1)})]^{1/2}} \qquad j = 2, \ldots, k-1$$

(5.4d)

$$\delta_{kj} \leq \frac{\beta_{k+1}}{\beta_k}\frac{\sigma_{k-1}}{2}\frac{\theta_j^{(k)} - \theta_{j-1}^{(k-1)}}{[(\theta_{j-1}^{(k-1)} - \theta_{j-2}^{(k-2)})(\theta_{j-1}^{(k-2)} - \theta_{j-1}^{(k-1)})]^{1/2}} \qquad j = 2, \ldots, k-1$$

where $\sigma_k = \theta_k^{(k)} - \theta_1^{(k)}$.

Proof. We prove (5.3c) and (5.4c), the other inequalities are given quite analogously. Using (1.11)

$$|\theta_\nu^{(\eta)} - \theta_\xi^{(\eta-1)}| \geq \delta_{\eta-1,\xi}|s_{\eta\nu}^{(\eta)}| = \frac{\beta_\eta}{\beta_{\eta+1}}|s_{\eta-1,\xi}^{(\eta-1)}|\delta_{\eta\nu} \tag{5.5}$$

Considering the first part of (5.5) with $\eta = k+1$, $\nu = \xi = j$

$$\delta_{kj} \leq \frac{\theta_j^{(k)} - \theta_j^{(k+1)}}{|s_{k+1,j}^{(k+1)}|}$$

and bounding $|s_{k+1,j}^{(k+1)}|$ by (5.2b) gives (5.3c).

Considering the second part of (5.5) with $\eta = k$, $\nu = \xi = j$

$$\delta_{kj} \leq \frac{\beta_{k+1}}{\beta_k} \frac{\theta_j^{(k-1)} - \theta_j^{(k)}}{|s_{k-1,j}^{(k-1)}|}$$

and using (5.2b) to bound $|s_{k-1,j}^{(k-1)}|$ gives (5.4c). $\square$

Note that (5.1) multiplied by β_{k+1}^2 relate the bound for δ_{kj} to the Ritz values computed at steps $k, k-1$, while (5.3) relate the bounds for δ_{kj} to the Ritz values at steps $k, k+1$ and (5.4) to the Ritz values at steps $k, k-1, k-2$. Thus, local properties of Ritz values in two or three subsequent steps can ensure convergence, cf. (1.4), (1.6), (2.1) and (3.23).

Using only neighbor Ritz values and two or three iterations substantially weakens the result. If, e.g. for $\theta_r^{(t-1)}$ $\delta_{t-1,r} \ll 1$ and $|\theta_r^{(t)} - \theta_r^{(t-1)}| = \sigma \leq \delta_{t-1,r}$, then using (2.1) there must be some $\theta_{j_l}^{(t+l)}$ within $2\delta_{t-1,r}$ of $\theta_r^{(t)}$ for any $l > 1$ but (5.1) gives

$$\delta_{tr} \leq \beta_{t+1} \frac{\sigma^{1/2}}{(\theta_{r+1}^{(t)} - \theta_r^{(t)})^{1/2}}$$

i.e. for $\theta_r^{(t)}$ well separated from $\theta_{r+1}^{(t)}$ the bound is $\sim \sigma^{1/2}$ only. Experiments confirmed that for the early stage of convergence (until δ_{kj} drops below $\sim \varepsilon^{1/2}$) (5.1) gives a sharp bound. Similarly, for outer Ritz values in the cluster perturbed by a new copy beginning to approach the cluster, the estimates (5.3) and (5.4) are realistic. Many relations similar to (5.1)–(5.4) can be developed, we choose the forms which looked simple enough.

It is natural to ask if (5.3), (5.4) proves the convergence of at least one Ritz value in any well separated cluster. The answer is negative. Let e.g. there are two close Ritz values at the steps k and $k-2$, followed by only one Ritz value at their neighborhood at the steps $k+1$ and $k-1$ and all these Ritz values are well separated from the others at the same steps, as is shown in the scheme

$k+1$		$+$	$\theta_{j+1}^{(k+1)}$
k	$\theta_j^{(k)}$	$++$	$\theta_{j+1}^{(k)}$
$k-1$		$+$	$\theta_j^{(k-1)}$
$k-2$	$\theta_{j-1}^{(k-2)}$	$++$	$\theta_j^{(k-2)}$
iteration step	Ritz values near $\theta_j^{(k)}, \theta_{j+1}^{(k)}$		

Without additional assumptions (5.1)–(5.4) give in this case no reasonable bound for δ_{kj} and $\delta_{k,j+1}$. A similar situation may occur e.g. after discovering a hidden eigenvalue in the spectra (cf. [Parlett–87] pp. 17–19, Figure 3) or forming multiple copies of the previously approximated eigenvalue. On the other hand one can expect a Ritz value the convergence of which is proved by (5.1)–(5.4) in the "history" of any cluster. This offers a justification for the statement that in the Lanczos process for A, q^1 every cluster of Ritz values approximates some eigenvalue λ_i of A (c.f. e.g. [Cullum–85]). This statement was in fact proved by Greenbaum's result [Greenbaum–89], it follows from the correspondence between the real and exact Lanczos runs. But this does not give a realistic bound for the precision of approximation.

6. Clustered Ritz values and eigenvector approximations in the finite precision Lanczos computation. Let $C_i^{(k)} = \{\theta_\nu^{(k)}, \ldots, \theta_{\nu+\eta}^{(k)}\}$ be the cluster of Ritz values approximating λ_i at the step k of the finite precision Lanczos process for A, q^1,

$$d_i^{(k)} = \theta_{\nu+\eta}^{(k)} - \theta_\nu^{(k)} \text{ and } g_i^{(k)} = \min\{\theta_\nu^{(k)} - \theta_{\nu-1}^{(k)}, \theta_{\nu+\eta+1}^{(k)} - \theta_{\nu+\eta}^{(k)}\}$$

its diameter and separation, $d_i^{(k)} \ll g_i^{(k)}$, $d_i^{(k)} \ll 1$, $\eta \geq 1$ ($\theta_0^{(k)}$ and $\theta_{k+1}^{(k)}$ are formally set to $-\infty$ and $+\infty$ respectively). Let $J = \{\nu, \nu+1, \ldots, \nu+\eta\}$, $\lambda_i \in \langle \theta_\nu^{(k)}, \theta_{\nu+\eta}^{(k)} \rangle$. For any Ritz value $\theta_j^{(k)} \in C_i^{(k)}$ with $\delta_{kj} \ll 1$ the Ritz vector $z_j^{(k)}$ is a perturbed scalar multiple of u_i. If $\|z_j^{(k)}\| \gg 0$, then $z_j^{(k)}$ gives a good approximation of u_i (Lemma 3.4). But there is still no proof for the experimentally based conjecture that $\delta_{kj} \ll 1$ for each Ritz value in the cluster and the relation

$$\sum_{j \in J} \|z_j^{(k)}\|^2 \simeq \eta \tag{6.1}$$

proved by Paige ([Paige–71, 80]) gives no argument that $\|z_j^{(k)}\| \gg 0$ for $\theta_j^{(k)}$ with δ_{kj} small (cf. (3.25)). One can intuitively expect that the subspace spanned by the Ritz vectors corresponding to the cluster $C_i^{(k)}$ is just the slightly perturbed one-dimensional subspace generated by u_i, but considering (1.4) this is equivalent to the convergence proved by the small value δ_{kj} for each Ritz value in $C_i^{(k)}$. We will discuss some points related to this view on the problem.

From Lemma 3.5 and (1.7) immediately follows that if $\max_{j\in J} |(z_j^{(k)}, q^{k+1})| \neq 0$, then

$$\max_{j\in J} \delta_{kj} \leq \frac{kd_i^{(k)} + \varepsilon_3}{\max_{j\in J} |(z_j^{(k)}, q^{(k+1)})|} \tag{6.2}$$

i.e. if the orthogonality is lost in the direction of at least one Ritz vector corresponding to the Ritz value in the cluster, then each of the clustered Ritz values has small value δ_{kj}. We examine in detail the loss of orthogonality among the Ritz vectors $z_j^{(k)}$, Lanczos vector q^{k+1} and eigenvectors u_r.

For $\theta_j^{(k)} \neq \lambda_i$ the relation (3.17) together with (2.12) and (2.15) of [Paige–80] gives the bound

$$|(u_i, z_j^{(k)})| \leq \frac{\delta_{kj} + \varepsilon_3}{|\lambda_i - \theta_j^{(k)}|} \tag{6.3}$$

i.e. if δ_{kj} is small and $\theta_j^{(k)}$ approximate some λ_r well separated from λ_i (e.g. $|\lambda_r - \lambda_i| \gg \delta_{kj}$), then $z_j^{(k)}$ is approximately orthogonal to u_i.

Let $\theta_j^{(k)} \in C_i^{(k)}$ and $\delta_{kj} \gg 0$. Then q^{k+1} and u_i must be approximately orthogonal, because using (3.17)

$$|(u_i, q^{k+1})| \leq \frac{|\lambda_i - \theta_j^{(k)}| + \varepsilon_4}{\delta_{kj}} \leq \frac{d_i^{(k)} + \varepsilon_4}{\delta_{kj}}. \tag{6.4}$$

But then using (3.17) again for some $\theta_\ell^{(k)}$ well separated from λ_i

$$|(u_i, z_\ell^{(k)})| \leq \frac{\delta_{k\ell} + \varepsilon_4}{\max_{j\in J}(\delta_{kj})} \frac{d_i^{(k)} + \varepsilon_4}{|\lambda_i - \theta_\ell^{(k)}|} \leq \frac{\delta_{k\ell} + \varepsilon_4}{\max_{j\in J}(\delta_{kj})} \frac{d_i^{(k)} + \varepsilon_4}{g_i^{(k)}}, \tag{6.5}$$

i.e. if there is any $\theta_j^{(k)} \in C_i^{(k)}$ with $\delta_{kj} \gg 0$, then the eigenvector u_i is approximately orthogonal to any Ritz vector which has the Ritz value separated from λ_i. This is a little more than (3.26) would suggest. We hoped to use these results to prove that δ_{kj} is small for any Ritz value in $C_i^{(k)}$ but we failed to complete the proof. At the beginning of our work we hoped for the result justifying that for any Ritz value in $C_i^{(k)}$ δ_{kj} is less or equal to the quantity proportional to

$$\left(\frac{d_i^{(k)}}{g_i^{(k)}}\right)^{1/2}.$$

In our experiment using $\rho = 0.7$ and studying $C_{24}^{(17)} - C_{24}^{(21)}$ we receive e.g. approximate values shown in the next table:

step k	17	18	19	20	21
$(d_{24}^{(k)}/g_{24}^{(k)})^{1/2}$	10^{-1}	10^{-3}	10^{-5}	10^{-6}	10^{-7}
δ_{kk} (first copy)	$< 10^{-9}$	$< 10^{-9}$	10^{-8}	10^{-8}	10^{-7}
δ_{kk-1} (second copy)	5	10^{-1}	10^{-3}	10^{-5}	10^{-6}

It is clear that $\delta_{kk-1} \gg (d_{24}^{(k)}/g_{24}^{(k)})^{1/2}$ for $k = 17, 18, 19$. The maximal ratio between these two factors is of order $\|A\| = 100$. We feel that more experiments are needed.

7. Identification of spurious (ghost) eigenvalues. The approach using spurious eigenvalues was introduced and extensively studied by Cullum and Willoughby in many papers (e.g. [Cullum–85, 86,80]). They proposed a test for spuriosity based on theoretical arguments referring to the situation at the late stage of computation The occurrence of spurious Ritz values is examined in this section in a more general way.

Experimental results related to this subject and discussed below are shown in two series of figures. Figures 7.1–7.5 describe the results of the "exact" Lanczos (Figure 7.1, $i = 24$) and finite precision Lanczos (Figures 7.2–7.5, $i = 18 - 21$) for $\rho = 0.8$ and vector r^0 generated randomly. Figures 7.6–7.10 show similar Lanczos runs for $\rho = 0.8$ and the vector r^0 taken as the fourth power of the randomly generated vector (this gives a wide range of weights). Figure 7.1 and 7.6 show a final step of the "exact" Lanczos runs.

DEFINITION 7.1. Ritz value $\theta_j^{(k)}$ computed in the step k of the finite precision Lanczos process for A, q^1 is called *spurious*, if it is a result of the loss of orthogonality (caused by rounding errors) and has no analog in any step of the exact Lanczos process for A, q^1.

This definition is a bit vague but hopefully it describes sufficiently those Ritz values in the real Lanczos process which can never be found among the Ritz values (even slightly perturbed) of the exact process for the same data. We are not interested in those spurious Ritz values, which form multiple copies of some eigenvalue (we are not concerned about the multiplicity of eigenvalues). The main idea is to identify those spurious Ritz values, which do not correspond to any original eigenvalue.

Our discussion is based on the relations (3.7), (3.15). It is e.g. clear that if there appeared Ritz value $\theta_j^{(k)}$ with large δ_{kj} near some Ritz value $\theta_i^{(k)}$ with small δ_{ki} and

$|\psi'_k(\theta_j^{(k)})| \sim |\psi'_k(\theta_i^{(k)})|$, then the weight of $\theta_j^{(k)}$ would be small. On the other hand, if there occurred a Ritz value $\theta_j^{(k)}$ with large δ_{kj} and small weight $m_j^{(k)}$ near some Ritz value $\theta_i^{(k)}$ with substantially larger weight $m_i^{(k)}$, $m_i^{(k)} \gg m_j^{(k)}$, and $|\psi'_k(\theta_i^{(k)})|$, $|\psi'_k(\theta_j^{(k)})|$ are of the same order, then δ_{ki} must be small. In both cases $|\psi_{2,k}(\theta_j^{(k)})| \ll |\psi_{2,k}(\theta_i^{(k)})|$. So the Ritz value with the value $|\psi_{2,k}(\theta_j^{(k)})|$ relatively small in comparison to their neighbors has some interesting properties.

Definition 7.2. Ritz value $\theta_j^{(k)}$ computed in the step k of the finite precision Lanczos process for A, q^1 is called *cw-spurious*, if it is within a small multiple of the machine precision of the exact root of $\psi_{2,k}$.

We point out that it may be $|\psi_{2,k}(\theta_j^{(k)})| \gg 1$ for $\theta_j^{(k)}$ *cw*-spurious, because $\psi_{2,k}$ can be a high order polynomial.

In [Cullum–85] the equivalence between spurious and *cw*-spurious Ritz values is stated. We show that this relation is more complicated which doubt the existence of a simple (single number based) and reliable spuriosity test.

First, let a spurious Ritz value $\theta_j^{(k)}$ appears between two well separated Ritz values $\theta_{j-1}^{(k)}$, $\theta_{j+1}^{(k)}$ which just began to converge. One can expect $|\psi'_k(\theta_{j-1}^{(k)})| \sim |\psi'_k(\theta_j^{(k)})| \sim |\psi'_k(\theta_{j+1}^{(k)})|$ and thus (3.7) implies $|\psi_{2,k}(\theta_{j-1}^{(k)})| \sim |\psi_{2,k}(\theta_j^{(k)})| \sim |\psi_{2,k}(\theta_{j+1}^{(k)})|$ and $m_{j-1}^{(k)} \sim m_j^{(k)} \sim m_{j+1}^{(k)} \sim o(1) \gg 0$. Then either $\theta_j^{(k)}$ may not be marked *cw*-spurious, or $\theta_{j-1}^{(k)}$ and $\theta_{j+1}^{(k)}$ may be also marked *cw*-spurious. In any case it looks like a counterexample for the statement in [Cullum-85]. Moreover, $m_j^{(k)} \gg 0$. Ritz value $\theta_6^{(23)}$ in Figure 7.10 gives an example of such spurious Ritz value. We see that $|s_{1,6}^{(23)}| \sim |s_{1,6}^{(23)}| \sim 10^{-2}|s_{1,7}^{(23)}|$. Figure 7.11 shows how this spurious Ritz value change its position in the spectra to form a second copy approaching λ_{21}.

Second, if there were some eigenvalue in the original spectra with a small weight (like λ_{22} in figure 7.6), then the Ritz value "looking for" this eigenvalue might be marked *cw*-spurious even if they were result of the exact Lanczos process. This situation is discussed in [Cullum–85]. That may be the case of $\theta_{12}^{(15)}$ in Figure 7.7 ($i = 15$). Ritz value $\theta_7^{(16)}$ in Figure 7.8 ($i = 16$) is an example of another Ritz value changing its position in the spectra very quickly (in Figure 7.8 it finally approaches λ_{16} as $\theta_9^{(17)}$), which is caused by the difference in the weights m_{15} and m_{16}. None of these Ritz values are spurious (our statement is based on the experiment with double reorthogonalized Lanczos).

In [Cullum–85] an associated conjugate-gradient process is defined for the real Lanczos run. Then it is shown that the relative norm of the k-th residual r^k of this process converge asymptotically to zero and this argument is used to establish the spuriosity test. To our opinion the existence of spurious or *cw*-spurious Ritz values cannot be in general related to the small values of $\|r^k\|$. The spurious Ritz value $\theta_{13}^{(18)}$ shown in Figure 7.2 has a good chance to be a cw-spurious, but $\|r^k\| = 0.1744$.This value was computed using the formula

$$(7.1) \qquad \|r^k\|^2 = \frac{\prod_{l=2}^{k} \beta_l^2}{(\psi_{k-1}(0))^2},$$

cf. [Cullum–85], Lemma 4.5.2. For the spurious Ritz value $\theta_{10}^{(20)}$ in Figure 7.4 $\|r^k\| = 0.2141$. Figures 7.3 and 7.5 show how the position of these spurious Ritz values is changed in the next iterations.

This discussion shows that there is no hope for a simple (a single number based) and reliable test for spuriosity. It is necessary to combine several approaches, e.g. the test for cw-spuriosity proposed by Cullum with the test for stabilization based on 5.1. The general purpose procedure reliably marking any spurious Ritz value must be very complicated.

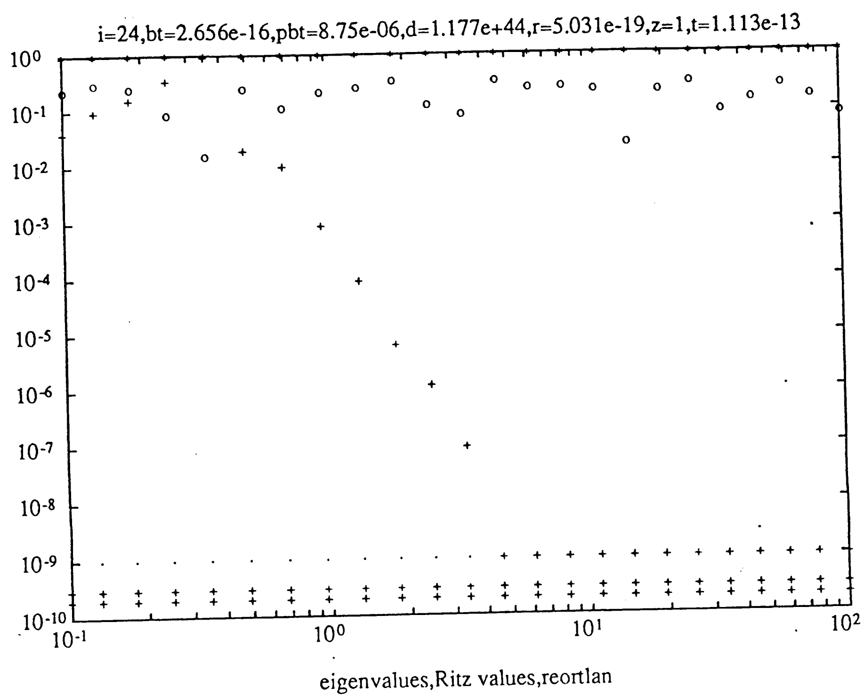

Figure 7.1

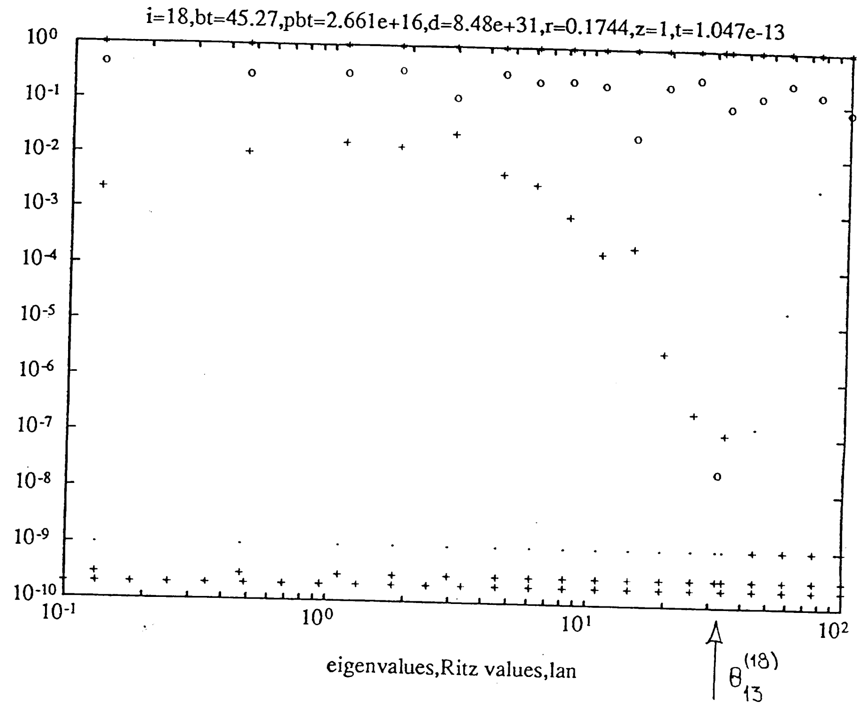

Figure 7.2

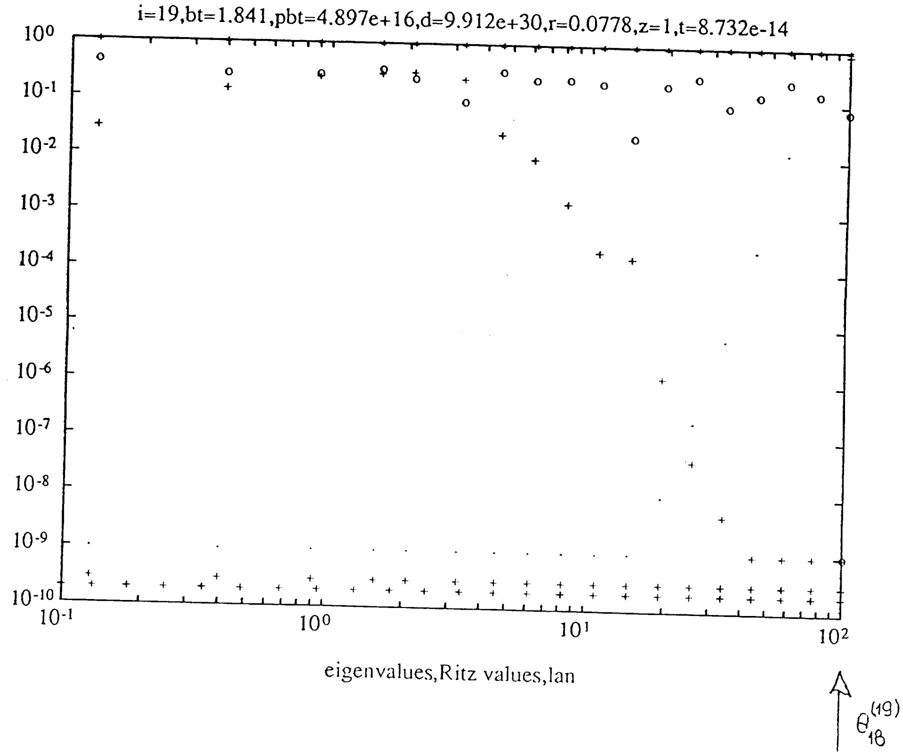

Figure 7.3

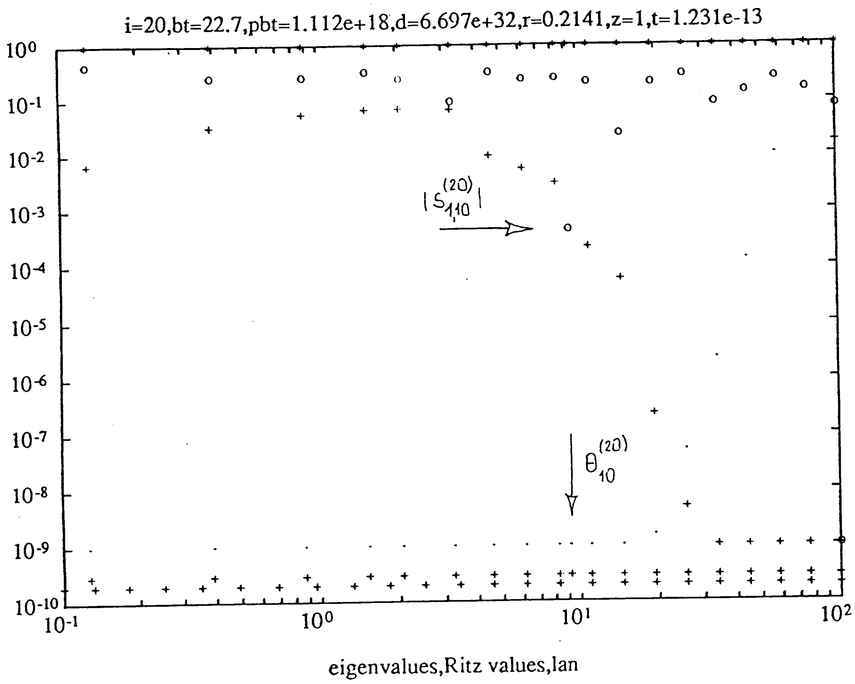

Figure 7.4

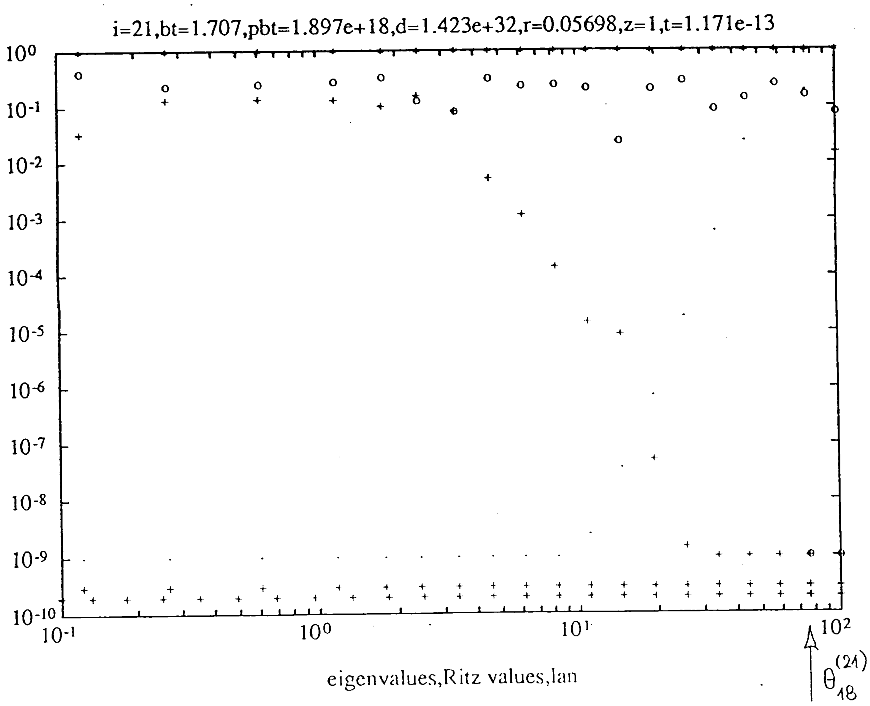

Figure 7.5

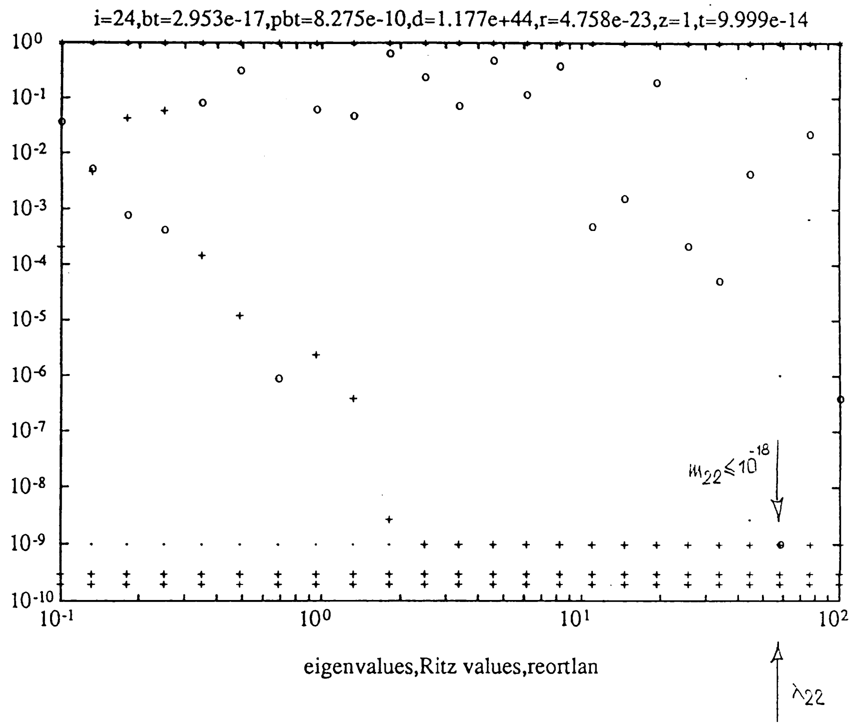

Figure 7.6

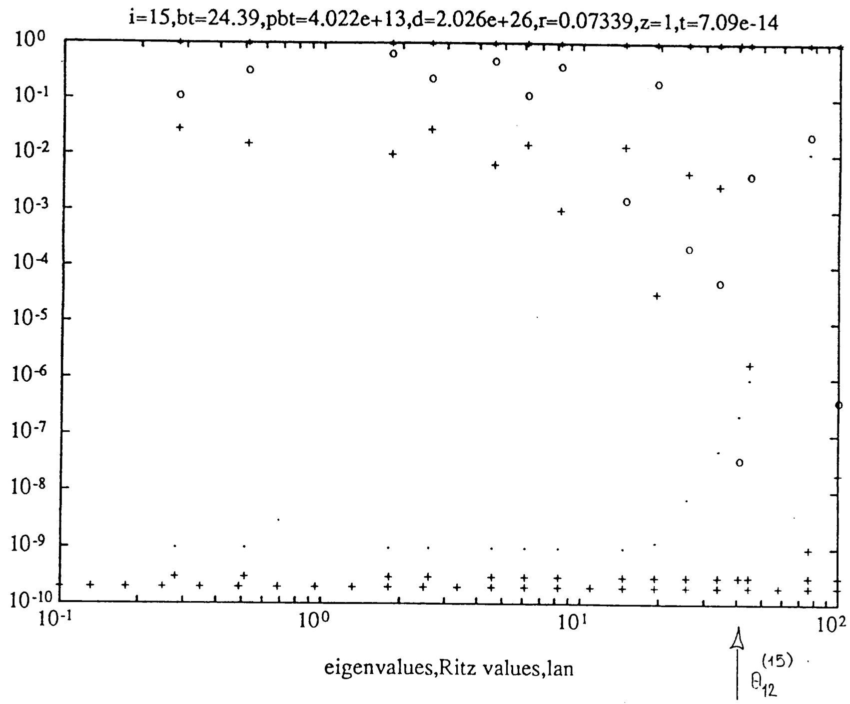

Figure 7.7

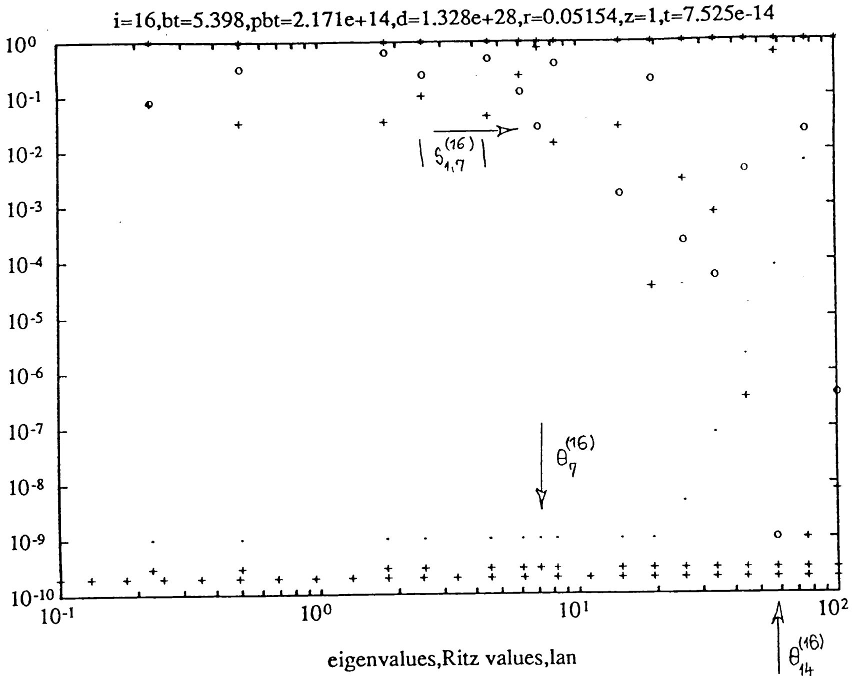

Figure 7.8

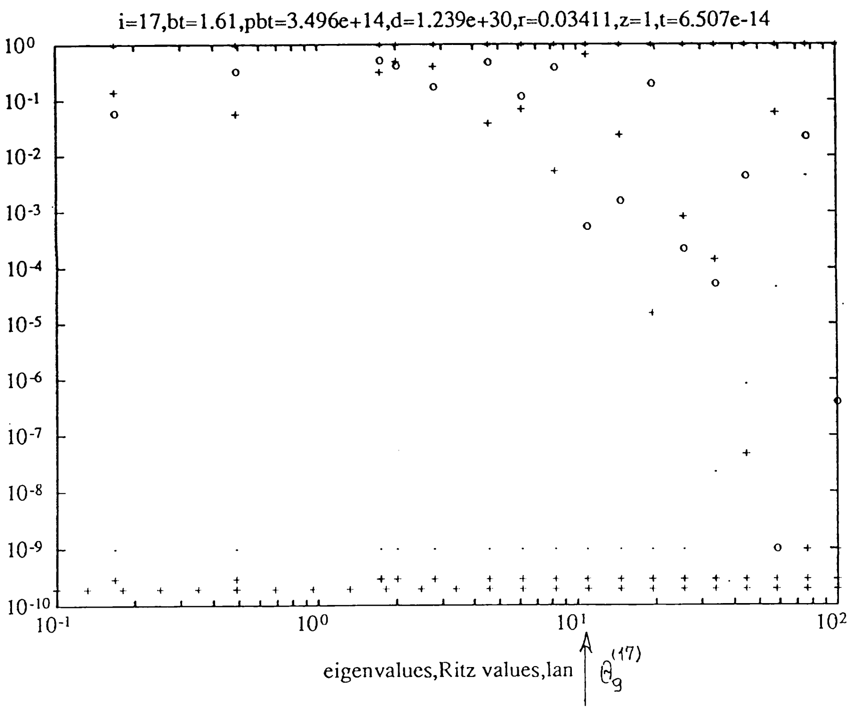

Figure 7.9

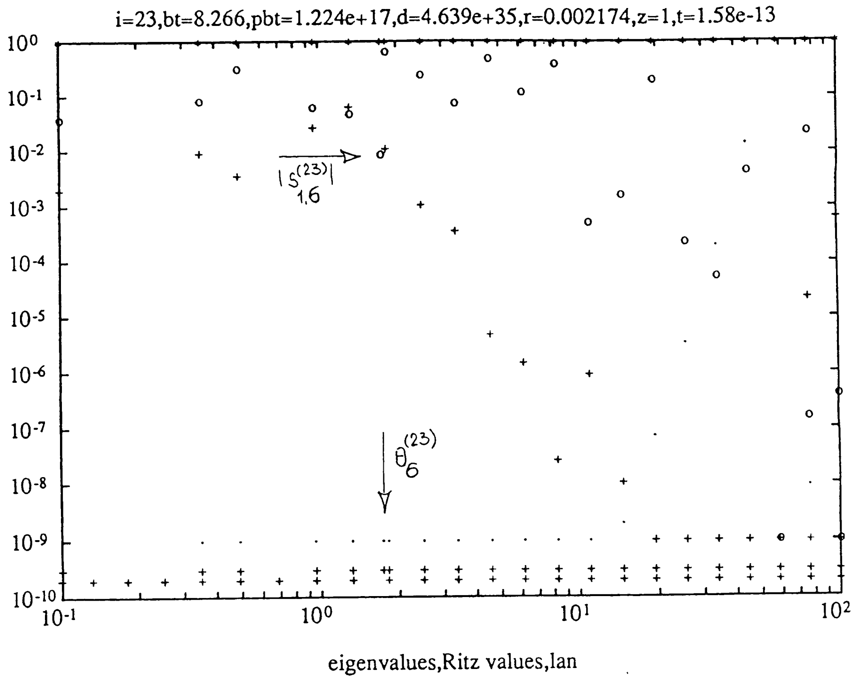

Figure 7.10

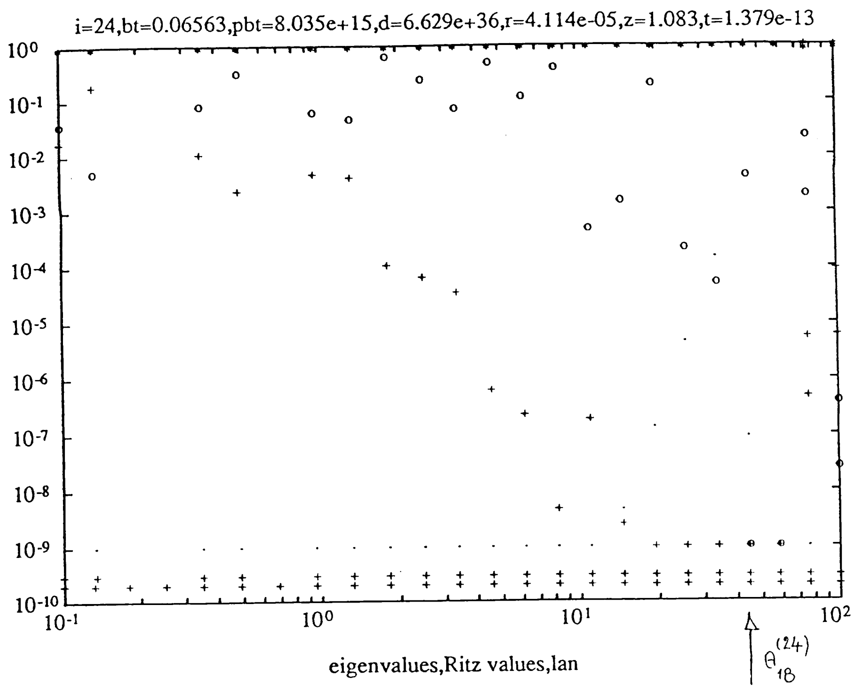

Figure 7.11

8. Stabilization of weights. It is very natural to ask a question about the stabilization of weights corresponding to the (well separated or clustered) Ritz value approaching some original eigenvalue. In general, it can be formulated in the next way.

Question 8.1. Let $(,)_n$, $n = 1, 2, \ldots$ be a sequence of innerproducts (1.19)–(1.20) defined by the exact or finite precision Lanczos process for A, q^1. In which way does the innerproduct $(,)_n$ approximate the original innerproduct $(,)$ defined by (1.13), (1.15)?

The next conjecture formulates the question more specifically.

Conjecture 8.1. Let $B_i^{(k)} = \{\theta_\nu^{(k)}, \ldots, \theta_{\nu+\eta}^{(k)}\}$ be a nonempty set of Ritz values approximating an eigenvalue λ_i at the step k of the exact or finite precision Lanczos process for A, q^1. Let these Ritz values be separated from the other Ritz values by $g_i^{(k)}$,

$$g_i^{(k)} = \min\{\theta_\nu^{(k)} - \theta_{\nu-1}^{(k)}, \quad \theta_{\nu+\eta+1}^{(k)} - \theta_{\eta+\nu}^{(k)}\},$$

and

$$\begin{aligned} d_i^{(k)} &= \qquad \theta_{\eta+\nu}^{(k)} - \theta_\nu^{(k)} \quad &&\text{for } \eta \neq 0, \\ d_i^{(k)} &= \min_{r=\nu,\nu+1} |\theta_\nu^{(k)} - \theta_r^{(k+1)}| \quad &&\text{for } \eta = 0. \end{aligned}$$

($\theta_0^{(k)}$ and $\theta_{k+1}^{(k)}$ are formally set to $-\infty$ and $+\infty$). Then

$$\sum_{r=\nu}^{\nu+\eta} m_r^{(k)} = \sum_{r=\nu}^{\nu+\eta} (s_{1r}^{(k)})^2 = m_i + h(d_i^{(k)}, g_i^{(k)}) \tag{8.1}$$

and $d_i^{(k)} \ll g_i^{(k)}$ implies $h(d_i^{(k)}, g_i^{(k)}) \ll m_i$.

In the language of Jacobi matrices a slightly different conjecture may be formulated.

Conjecture 8.2. Let T_k be any Jacobi matrix, $B_{\nu,\eta}^{(k)} = \{\theta_\nu^{(k)}, \ldots, \theta_{\nu+\eta}^{(k)}\}$ be a nonempty set of its eigenvalues, $J = \{\nu, \ldots, \nu + \eta\}$,

$$g_{\nu\eta}^{(k)} = \min\{\theta_\nu^{(k)} - \theta_{\nu-1}^{(k)}, \theta_{\nu+\eta+1}^{(k)} - \theta_{\nu+\eta}^{(k)}\},$$

where $\theta_0^{(k)}$ and $\theta_{k+1}^{(k)}$ are formally set to $-\infty$ and $+\infty$. Let T_t be any left principal submatrix of T_k for which there is a nonempty set $B_{\bar\nu\bar\eta}^{(t)} = \{\theta_{\bar\nu}^{(t)}, \ldots, \theta_{\bar\nu+\bar\eta}^{(t)}\}$ of its eigenvalues approaching some of the eigenvalues in $B_{\nu\eta}^{(k)}$, $\bar J = \{\bar\nu, \ldots, \bar\nu + \bar\eta\}$. Let

$$g_{\bar\nu\bar\eta}^{(t)} = \min\{\theta_{\bar\nu}^{(t)} - \theta_{\bar\nu-1}^{(t)}, \quad \theta_{\bar\nu+\bar\eta+1}^{(t)} - \theta_{\bar\nu+\bar\eta}^{(t)}\}.$$

$\theta_0^{(t)}$ and $\theta_{t+1}^{(t)}$ are formally set to $-\infty$ and $+\infty$. Let the distance $\sigma_{\nu\eta\bar\nu\bar\eta}^{(kt)}$ between $B_{\nu\eta}^{(k)}$ and $B_{\bar\nu\bar\eta}^{(t)}$ be defined as

$$\sigma_{\nu\eta\bar\nu\bar\eta}^{(kt)} = \max\{|\theta_j^{(k)} - \theta_r^{(t)}|, \quad j \in J, \quad r \in \bar J\}.$$

Then

$$\sum_{j=\nu}^{\nu+\eta}(s_{1j}^{(k)})^2 = \sum_{r=\bar\nu}^{\bar\nu+\bar\eta}(s_{1r}^{(t)})^2 + h(g_{\nu\eta}^{(k)}, g_{\bar\nu\bar\eta}^{(t)}, \sigma_{\nu\eta\bar\nu\bar\eta}^{(kt)}), \tag{8.2}$$

and

$$g_{\nu\eta}^{(k)}, g_{\bar\nu\bar\eta}^{(t)} \gg \sigma_{\nu\eta\bar\nu\bar\eta}^{(kt)} \quad \text{implies} \quad h(g_{\nu\eta}^{(k)}, g_{\bar\nu\bar\eta}^{(t)}, \sigma_{\nu\eta\bar\nu\bar\eta}^{(kt)}) \ll \sum_{r=\bar\nu}^{\bar\nu+\bar\eta}(s_{1r}^{(t)})^2.$$

In this section Conjecture 8.1 is proved for some special cases and preliminary results for the general case is discussed with some partial results.

We first examine the case corresponding to the exact Lanczos algorithm applied to A, q^1. Let $g_i = \min_{j\neq i} |\lambda_j - \lambda_i|$ be the separation of λ_i from the other eigenvalues. For Ritz values with small bottom elements of the corresponding eigenvectors the situation is simple.

LEMMA 8.1. *Exact Lanczos process for A, q^1 is considered. Let $\theta_j^{(k)}$ and $\theta_r^{(n)}$ approximate the same eigenvalue λ_i, $k < n < N$, let*

$$\delta = \max\{\delta_{kj}, \delta_{nr}\} \ll g_i.$$

Then

$$m_j^{(k)} = m_r^{(n)} + n^{1/2}\, \mathrm{o}\!\left(\frac{\delta}{g_i - \delta}\right). \tag{8.3}$$

Proof. From Lemma 3.4

$$|1 - \xi_{ij}^{(k)}| \le \|z_j^{(k)} - \xi_{ij}^{(k)} u_i\| \le \frac{\delta_{kj}}{\min_{\ell\neq i} |\lambda_\ell - \theta_j^{(k)}|} \le \frac{\delta}{g_i - \delta}$$

$$|1 - \xi_{ir}^{(n)}| \le \|z_r^{(n)} - \xi_{ir}^{(n)} u_i\| \le \frac{\delta}{g_i - \delta},$$

because $|\lambda_i - \theta_j^{(k)}| \le \delta, |\lambda_i - \theta_r^{(n)}| \le \delta$, and consequently

$$z_j^{(k)} = \xi_{ij}^{(k)} u_i + \mathrm{o}\!\left(\frac{\delta}{g_i - \delta}\right) = \xi_{ir}^{(n)} u_i \cdot \frac{\xi_{ij}^{(k)}}{\xi_{nr}^{(n)}} + \mathrm{o}\!\left(\frac{\delta}{g_i - \delta}\right) = z_r^{(n)} + \mathrm{o}\!\left(\frac{\delta}{g_i - \delta}\right).$$

Using the definition of Ritz vectors and the fact that the matrix Q_n has orthonormal columns

$$Q_n \left[s_r^{(n)} - \begin{pmatrix} s_j^{(k)} \\ 0 \end{pmatrix} \right] = \mathrm{o}\!\left(\frac{\delta}{g_i - \delta} \right)$$

$$|s_{1r}^{(n)} - s_{1j}^{(k)}| \le \| s_{1r}^{(n)} - \begin{pmatrix} s_j^{(k)} \\ 0 \end{pmatrix} \| \le \|Q_n^T\| \,\mathrm{o}\!\left(\frac{\delta}{g_i - \delta} \right) \le n^{1/2}\, \mathrm{o}\!\left(\frac{\delta}{g_i - \delta} \right).$$

Because of

$$|m_r^{(n)} - m_j^{(k)}| = |(s_{1r}^{(n)} - s_{1j}^{(k)})(s_{1r}^{(n)} + s_{1j}^{(k)})|$$

the proof is finished. □

Similarly for well separated Ritz value $\theta_r^{(n)}$.

LEMMA 8.2. *Exact Lanczos process for A, q^1 is considered. Let $\theta_j^{(k)}$ and $\theta_r^{(n)}$ approximate the same eigenvalue λ_i, $k < n \le N$, let $\delta_{kj} \ll g_i$ and $\theta_r^{(n)}$ are well separated from the other Ritz values, i.e.*

$$\min_{\ell \ne r} |\theta_r^{(n)} - \theta_\ell^{(n)}| \ge \gamma g_i \gg \delta_{kj}^{1/2}.$$

Then

$$m_j^{(k)} = m_r^{(n)} + n^{1/2} \left[\frac{\beta_{n+1}}{(\gamma g_i)^{1/2}}\, \mathrm{o}\!\left(\frac{\delta_{kj}^{1/2}}{g_i - 2\delta_{kj}} \right) + \mathrm{o}\!\left(\frac{\delta_{kj}}{g_i - \delta_{kj}} \right) \right]. \tag{8.4}$$

Proof. Using (2.1), (5.1)

$$\delta_{nr} \le \beta_{n+1} \frac{(2\delta_{kj})^{1/2}}{(\gamma g_i)^{1/2}}$$

and repeating with slight modification the proof of Lemma 8.1 gives (8.4). □

For $n = N$ and $\theta_r^{(n)} = \lambda_i$ the result proved in Lemma 8.1 and 8.2 is rewritten in the next theorem.

THEOREM 8.1. *Exact Lanczos process for A, q^1 is considered. Let $\theta_j^{(k)}$, $1 \le k < N$, $1 \le j \le k$ is close to some λ_i so that*

$$\delta_{kj} \ll g_i = \min_{\ell \ne i} (\lambda_\ell - \lambda_i).$$

Then

$$m_j^{(k)} = m_i + N^{1/2}\, \mathrm{o}\!\left(\frac{\delta_{kj}}{g_i - \delta_{kj}} \right). \tag{8.5}$$

Let now the Ritz value $\theta_r^{(n)}$ is not well separated. Then from the interlacing property there can be at most one other Ritz value, say $\theta_{r+1}^{(n)}$, approaching λ_i and from the next lemma at least one of the values δ_{nr}, δ_{nr+1} must be large.

LEMMA 8.3. *Exact Lanczos process for A, q^1 is considered. Let $\theta_r^{(n)}$, $\theta_{r+1}^{(n)}$ approach the same eigenvalue λ_i, $1 < n < N$, let*

$$|\theta_r^{(n)} - \theta_{r+1}^{(n)}| \leq \vartheta g_i, \quad \vartheta \ll 1.$$

Then

$$\max\{\delta_{nr}, \delta_{nr+1}\} = \mathrm{o}\left(\frac{g_i(1-\vartheta)}{n^{1/2}}\right). \tag{8.6}$$

Proof. Let $\max\{\delta_{nr}, \delta_{nr+1}\} \leq \xi \frac{g_i(1-\vartheta)}{n^{1/2}}$, $\xi \ll 1$. Then analogously to the proof of Lemma 8.1

$$\begin{aligned}
|1 - \xi_{ir}^{(n)}| &\leq \|z_r^{(n)} - \xi_{ir}^{(n)} u_i\| \leq \frac{\xi}{n^{1/2}}, \\
|1 - \xi_{ir+1}^{(n)}| &\leq \|z_{r+1}^{(n)} - \xi_{ir}^{(n)} u_i\| \leq \frac{\xi}{n^{1/2}}, \\
z_{r+1}^{(n)} &= z_r^{(n)} + \mathrm{o}\left(\frac{\xi}{n^{1/2}}\right).
\end{aligned}$$

Then

$$\begin{aligned}
&s_{r+1}^{(n)} = s_r^{(n)} + \mathrm{o}(\xi) \quad \text{and} \\
&|(s_{r+1}^{(n)}, s_r^{(n)}) - 1| = \mathrm{o}(\xi)
\end{aligned}$$

which contradicts the orthonormality property of eigenvectors. □

One would like to prove that in this case one of the values δ_{nr}, δ_{nr+1} must be small, but the discussion of relations (5.1)–(5.4) in section 5 showed that this question remains open. We will analyze the situation in more detail. Let under the assumptions of Lemma 8.3 e.g. the original eigenvalue λ_{i+1} is not approached by any Ritz value at the step n and is well separated from the other eigenvalues so that

$$\left[\frac{\psi_n(\lambda_{i+1})}{\lambda_{i+1} - \theta_{r+1}^{(n)}}\right]^2 \gg (\psi_n'(\theta_{r+1}^{(n)}))^2.$$

Then for any subsequent step t, $n < t \le N$ at which λ_{i+1} is sufficiently closely approximated by some $\theta_\ell^{(t)}$ we receive from (3.13)

$$m_\ell^{(t)} \le \left[\frac{\psi_n'(\theta_{r+1}^{(n)})(\theta_\ell^{(t)} - \theta_{r+1}^{(n)})}{\psi_n(\theta_\ell^{(t)})}\right]^2 \sim \left[\frac{\psi_n'(\theta_{r+1}^{(n)})(\lambda_{i+1} - \theta_{r+1}^{(n)})}{\psi_n(\lambda_{i+1})}\right]^2 \ll 1$$

i.e. the weight of λ_{i+1} must be small. This gives an indication how the strange situation from Lemma 8.3 might eventually occur in practical computation. We have not found this case in our experiments. For $\theta_r^{(n)}$, $\theta_{r+1}^{(n)}$ sufficiently close

$$\left|\frac{\psi_n'(\theta_r^{(n)})}{\psi_n'(\theta_{r+1}^{(n)})}\right| \sim 1$$

and thus using (3.15) the weights $m_r^{(n)}$, $m_{r+1}^{(n)}$ are inversely proportional to the values δ_{nr}, δ_{nr+1}.

From Lemma 3.3, relation (3.22)

$$s_{1\ell}^{(n)}(u_i, z_\ell^{(n)}) = \frac{\psi_n(\lambda_i)}{(\lambda_i - \theta_\ell^{(n)})\psi_n'(\theta_\ell^{(n)})}(u_i, q^1). \tag{8.7}$$

We give a bound for the terms $s_{1\ell}^{(n)}(u_i, z_\ell^{(n)})$, $\ell \ne r, r+1$. Let $\max\{\delta_{nr}, \delta_{nr+1}\} = \kappa g_i(1-\vartheta)/n^{1/2}$, $\kappa = \mathrm{o}(1)$. Then using (6.5) with $\varepsilon_4 = 0$

$$|(u_i, z_\ell^{(n)})| \le \delta_{n\ell}\frac{\vartheta n^{1/2}}{\kappa(1-\vartheta)g_i} \le \frac{n^{1/2}\beta_{n+1}\vartheta}{\kappa(1-\vartheta)g_i} \tag{8.8}$$

so

$$\left|\sum_{\substack{\ell=1\\ \ell\ne r,r+1}}^{n} s_{1\ell}^{(n)}(u_i, z_\ell^{(n)})\right| \le \left(\sum_{j=1}^{n}(s_{1\ell}^{(n)})^2\right)^{1/2}\left(\sum_{\substack{j=1\\ j\ne r,r+1}}^{n}(u_i, z_\ell^{(n)})^2\right)^{1/2} \le \frac{n\beta_{n+1}\vartheta}{\kappa(1-\vartheta)g_i}. \tag{8.9}$$

The result is summarized in the next theorem.

THEOREM 8.2. *Exact Lanczos process for A, q^1 is considered. Let $\theta_r^{(n)}, \theta_{r+1}^{(n)}$ approach the same eigenvalue $\lambda_i, 1 < n < N$, $1 \le r < n$, let*

$$|\theta_r^{(n)} - \theta_{r+1}^{(n)}| < \vartheta g_i, \quad g_i = \min_{\ell\ne i}|\lambda_\ell - \lambda_i|, \quad \vartheta \ll 1,$$
$$\max\{\delta_{nr}, \delta_{nr+1}\} = \kappa g_i(1-\vartheta)/n^{1/2}, \quad \kappa = \mathrm{o}(1).$$

Then

$$m_i^{1/2} = |s_{1r}^{(n)}(u_i, z_r^{(n)}) + s_{1r+1}^{(n)}(u_i, z_{r+1}^{(n)})| + \frac{n\beta_{n+1}}{\kappa} \circ\left(\frac{\vartheta}{(1-\vartheta)g_i}\right). \tag{8.10}$$

Proof. It is given by considering (8.9) with (3.19). □

Note that the terms $s_{1r}^{(n)}(u_i, z_r^{(n)})$ and $s_{1r+1}^{(n)}(u_i, z_{r+1}^{(n)})$ have the same sign, which follows from (8.7). Using (3.16) the relation (8.10) is rewritten as

$$m_i^{1/2} = \beta_{n+1}|(u_i, q^{n+1})| \left| \frac{s_{1r}^{(n)} s_{nr}^{(n)}}{\lambda_i - \theta_r^{(n)}} + \frac{s_{1r+1}^{(n)} s_{nr+1}^{(n)}}{\lambda_i - \theta_{r+1}^{(n)}} \right| + \frac{n\beta_{n+1}}{\kappa} \circ\left(\frac{\vartheta}{(1-\vartheta)g_i}\right). \tag{8.11}$$

That is all we are able to say about the case corresponding to the exact precision Lanczos run. We feel that the "point of discovery" idea introduced in [Parlett–87] might be used to cope with the case when λ_i is close to some other eigenvalues (and ideally – using the sequence of Jacobi matrices – it could be applied also to the finite precision case). This needs further extensive work.

Now we turn to the finite precision case. q^{k+1} cannot be expressed as polynomial in A applied to q^1, there are no analogies of the related results for a finite precision Lanczos run. Supposing well separation of the Ritz value at some step we easily derive the analogy of Lemma 8.2.

LEMMA 8.4. *Finite precision process for A, q^1 is considered. Let $\theta_j^{(k)}$, $\theta_r^{(n)}$ approximate the same eigenvalue $\lambda_i, 1 \leq k < n$, let*

$$\delta_{kj} \ll g_r^{(n)} = \min_{\ell \neq r} |\theta_\ell^{(n)} - \theta_r^{(n)}|.$$

Then

$$m_j^{(k)} = m_r^{(n)} + n^{1/2} \circ\left(\frac{\delta_{kj}}{g_r^{(n)} - \delta_{kj}}\right) \tag{8.12}$$

Proof. Assume hypothetical exact Lanczos process for T_n, e^1. (from Introduction this process produces exactly the same tridiagonal matrices like the original finite precision process for A, q^1 in the steps 1 thru n). Applying Theorem 8.1 the proof is finished. □

If we assume only a small values of bottom elements of the corresponding eigenvectors, things become more complicated. Let $\theta_j^{(k)}, \theta_r^{(n)}$ approximate the same eigenvalue λ_i, let

$\partial = \max\{\partial_{ij}^{(k)}, \partial_{ir}^{(n)}\}$, where $\partial_{ij}^{(k)}$, $\partial_{ir}^{(n)}$ are bounds for $|\lambda_i - \theta_j^{(k)}|$, $|\lambda_i - \theta_r^{(n)}|$ given by (1.6), let

$$\begin{aligned} n^{1/2}\|A\|\varepsilon_1 &\leq \delta = \max\{\delta_{kj}; \delta_{nr}\} \\ \partial &\ll g_i = \min_{\ell \neq i} |\lambda_\ell - \lambda_i|. \end{aligned}$$

Then from Lemma 3.4

$$\|z_j^{(k)} - \xi_{ij}^{(k)} u_i\| \leq \frac{2\delta}{g_i - \partial}, \|z_r^{(n)} - \xi_{ir}^{(n)} u_i\| \leq \frac{2\delta}{g_i - \partial}.$$

Let the columns of Q_n are linearly independent, let $0 < \sigma$ be the smallest singular value of Q_n, let $\xi = \min\{\xi_{ij}^{(k)}, \xi_{ir}^{(n)}\} \gg \left(\frac{\delta}{g_i - \partial}\right)^{1/2}$. Then

$$\begin{aligned} \left| \frac{m_r^{(n)}}{(\xi_{ir}^{(n)})^2} - \frac{m_j^{(k)}}{(\xi_{ir}^{(n)})^2} \right| &\leq \frac{2}{\xi} \left| \frac{s_{1r}^{(n)}}{\xi_{ir}^{(n)}} - \frac{s_{1j}^{(k)}}{\xi_{ij}^{(k)}} \right| \leq \frac{2}{\xi} \left\| \frac{s_r^{(n)}}{\xi_{ir}^{(n)}} - \frac{1}{\xi_{ij}^{(k)}} \begin{pmatrix} s_j^{(k)} \\ 0 \end{pmatrix} \right\| \leq \\ &\leq \frac{2}{\xi\sigma} \left\| Q_n \left(\frac{1}{\xi_{ir}^{(n)}} s_r^{(n)} - \frac{1}{\xi_{ij}^{(k)}} \begin{pmatrix} s_j^{(k)} \\ 0 \end{pmatrix} \right) \right\| = \frac{2}{\xi\sigma} \left\| \frac{z_r^{(n)}}{\xi_{ir}^{(n)}} - \frac{z_j^{(k)}}{\xi_{ij}^{(k)}} \right\| \leq \\ &\leq \frac{4}{\xi^2 \sigma} \frac{\delta}{g_i - \partial} \end{aligned}$$

and finally

$$\frac{m_r^{(n)}}{(\xi_{ir}^{(n)})^2} = \frac{m_j^{(k)}}{(\xi_{ij}^{(k)})^2} + \frac{4}{\xi^2\sigma} \circ\left(\frac{\delta}{g_i - \partial}\right), \tag{8.13}$$

which is some analogy of Lemma 8.1. This is of course a very weak result, because the assumption that the linear dependency among the columns of Q_n is well preserved is too strong. We show that if there are at some step n two Ritz values $\theta_r^{(n)}$, $\theta_{r+1}^{(n)}$ close to the same eigenvalue λ_i so that

$$n^{1/2}\|A\|\varepsilon_1 \leq \delta = \max\{\delta_{nr}, \delta_{nr+1}\} \ll g_i,$$

then the linear independency among the columns of Q_n is lost enough to make (8.13) worthless. Indeed, from the derivation of (8.13)

$$\sigma \leq \frac{2}{\xi} \left\| \frac{s_r^{(n)}}{\xi_{ir}^{(n)}} - \frac{s_{r+1}^{(n)}}{\xi_{ir+1}^{(n)}} \right\| \frac{\delta}{g_i - \partial} \leq \frac{4}{\xi^2} \frac{\delta}{g_i - \partial} \tag{8.14}$$

where the right inequality is valid because $s_r^{(n)}$ and $s_{r+1}^{(n)}$ are orthogonal. We note at this point one interesting experimental observation we have made. We met $\|z_r^{(n)}\| \ll \|z_{r+1}^{(n)}\|$

(or $\|z_r^{(n)}\| \gg \|z_{r+1}^{(n)}\|$) very rarely in our experiments and if it occurred, then the weights of the corresponding clustered Ritz values were almost equivalent. We have not clarified this observation theoretically.

In the finite precision case we have no termination property so no analogy of Theorem 8.1 can be proved in a way similar to the exact case.

The hard result proved by Greenbaum [Greenbaum–89, (8.21)] relates the original weights m_i, $i = 1, 2, \ldots, N$, to the weights of the corresponding clusters produced in the final step of the exact Lanczos process for $\bar{A}, \bar{q}_1$, where $\bar{A}$ is some larger matrix with the eigenvalues clustered around the original A's eigenvalues. It can be combined with Conjecture 8.2 to prove Conjecture 8.1 and the trouble with the finite termination is thus avoided. But there is still a need for a proof of Conjecture 8.2. We try to use (3.19) and develop some analogy of Theorem 8.2. This theorem is based on Lemma 8.3 which again cannot be extended to the finite precision Lanczos. Note that in the exact case the cluster of Ritz values might occur only occasionally and is "unstable", while in the finite precision case we cope with "stable" and "growing" clusters.

Let $B_i^{(k)} = \{\theta_\nu^{(k)}, \ldots, \theta_{\nu+\eta}^{(k)}\}$ be a nonempty set of Ritz values approximating λ_i defined in Conjecture 8.1, let $d_i^{(k)} \ll g_i^{(k)}$. Let J be the set of indices $\{\nu, \ldots, \nu+\eta\}$, $\eta \geq 1$, $\lambda_i \in \langle\theta_\nu^{(k)}, \theta_{\nu+\eta}^{(k)}\rangle$, $K = \{1, 2, \ldots, \nu-1, \nu+\eta+1, \ldots, k\}$. Using (3.19)

$$(u_i, q^1) = \sum_{j\in J} s_{1j}^{(k)}(u_i, z_j^{(k)}) + \sum_{j\in K} s_{1j}^{(k)}(u_i, z_j^{(k)}). \tag{8.15}$$

The first sum is determined by the cluster itself. The last sum account for the effect of the other Ritz values, which are supposed to be well separated from the cluster. For any $j \in K$ (using 3.17 and supposing $\delta_{kj} \geq \varepsilon_4$)

$$|(u_i, z_j^{(k)})| = \mathrm{o}(\delta_{kj}/g_i^{(k)}).$$

Let for a given small ζ, $K_+ = \{j \in K,\ \delta_{kj} \leq \zeta\frac{g_i^{(k)}}{k^{1/2}}, 0 < \zeta \ll 1\}$, $K_- = K \div K_+$. Then clearly

$$\left|\sum_{j\in K_+} s_{1j}^{(k)}(u_i, z_j^{(k)})\right| = \mathrm{o}(\zeta). \tag{8.16}$$

We are not able to say anything about the sum over K_- in the case $\delta_{k\ell} \ll 1$, $j \in J$. If for at least one $\ell \in J$ the value $\delta_{k\ell}$ is large, i.e.

$$\max_{\ell\in J}(\delta_{k\ell}) \geq \tau\frac{g_i^{(k)}}{k^{1/2}}, \quad \tau \gg 0,$$

then we have in fact the situation which is assured by Lemma 8.3 for the exact case and using (6.5)

$$\left| \sum_{j\in K_-} s_{1j}^{(k)}(u_i, z_j^{(k)}) \right| \leq k^{1/2} \frac{2\beta_{k+1}}{\max_{\ell \in J}(\delta_{k\ell})} \frac{d_i^{(k)} + \varepsilon_4}{g_i^{(k)}}$$

and supposing $d_i^{(k)} \geq \varepsilon_4$

$$|(u_i, g^1)| = |\sum_{j\in J} s_{1j}^{(k)}(u_i, z_j^{(k)})| + \text{o}(\zeta) + k\beta_{k+1}\,\text{o}\left(\frac{d_i^{(k)}}{\tau(g_i^{(k)})^2} \right), \tag{3.17}$$

i.e. if for at least one Ritz value in the sufficiently tight and well separated cluster δ_{kj} is not small, than the $|(u_i, q^1)|$ is a small perturbation of the sum

$$\left| \sum_{j\in J} s_{1j}^{(k)}(u_i, z_j^{(k)}) \right|.$$

This is of course a very partial result. The question remains open and needs further work.

REFERENCES

[1] F.V. Atkinson, *Discrete and Continuous Boundary Value Problems*, Academic Press, N.Y., 1964.

[2] C. de Boor and G.H. Golub, *The Numerically stable reconstruction of a Jacobi matrix from Spectral Data*, Linear algebra Appl. 21 (1978), pp. 245–260.

[3] C. de Boor and E.B. Saff, *Finite Sequences of Orthogonal Polynomials Connected by a Jacobi Matrix*, Linear Algebra Appl. 75 (1986), pp. 43–55.

[4] J. Cullum and R.A. Willoughby, *The Lanczos Phenomenon – an Interpretation Based upon Conjugate Gradient Optimization*, Linear Algebra Appl. 29 (1980), pp. 63–90.

[5] J. Cullum and R.A. Willoughby, *Lanczos Algorithms for Large Symmetric Eigenvalue Computation*, vol 1 Theory, Birghäuser, Boston, 1985.

[6] J. Cullum and R.A. Willoughby, *A Practical Procedure for Computing Eigenvalues of Large Sparse Nonsymmetric Matrices*, In: Large Scale Eigv. Problems (J. Cullum and R.A. Willoughby, eds.), North Holland (1986), pp. 193–240.

[7] G. Cybenko, *An Explicit Formula for Lanczos Polynomials*, Linear Algebra Appl., 88/89 (1987), pp. 99–115.

[8] P. Deift and T. Nanda, *On the Determination of a Tridiagonal Matrix from its Spectrum and a Submatrix*, Linear Algebra Appl. 60 (1984), pp. 43–55.

[9] F.R. Gantmakher, M.G. Krein, *Ostsillatsionnyje matritsy i jadra i malyje kolebania mekhanicheskich sistem*, GITTL, Moskva (1950), in Russian.

[10] G.H. Golub and J.H. Welsch, *Calculation of Gauss quadrative rules*, Math. Comp. 23 (1969), pp. 221–230.

[11] G.H. Golub and Ch. V. Loan, *Matrix Computations*, The J. Hopkins University Press, Baltimore (1983,1989).

[12] W.B. Gragg and W.J. Harrod, *The Numerically Stable Reconstruction of Jacobi Matrices from Spectral Data*, Numer. Math 44 (1984), pp. 317–335.

[13] A. Greenbaum, *Behavior of Slightly Perturbed Lanczos and Conjugate Gradient Recurrences*, Linear Algebra Appl., 113 (1989), pp. 7–63.

[14] A. Greenbaum, Z. Strakoš, *Predicting behavior of finite precision Lanczos and conjugate gradient computations*, SIMAX, 13 (1992), pp. 121–137.

[15] O.H. HALD, *Inverse Eigenvalue Problems for Jacobi Matrices*, Linear Algebra Appl., 14 (1976), pp. 63–85.
[16] R.O. HILL, JR. AND B.N. PARLETT, *Refined Interlacing Properties*, SIMAX, 13 (1992), pp. 239–247.
[17] H. HOCHSTADT, *On the Construction of a Jacobi Matrix from Spectral Data*, Linear Algebra Appl. 8 (1974), pp. 435–446.
[18] H. HOCHSTADT, *The Construction of a Jacobi Matrix from Mixed Given Data*, Linear Algebra Appl. 28 (1979), pp. 113–115.
[19] C.C. PAIGE, *The Computation of Eigenvalues and Eigenvectors of Very Large Sparse Matrices (PhD thesis)*, U of London, London (1971).
[20] C.C. PAIGE, *Computational Variants of the Lanczos Method for the Eigenproblem*, J. Inst. Maths. Applics 10 (1972), pp. 373–381.
[21] C.C. PAIGE, *Error Analysis of the Lanczos Algorithm for Tridiagonalizing a Symmetric Matrix*, J. Inst. Maths. Applics 18 (1976), pp. 341–349.
[22] C.C. PAIGE, *Accuracy and effectiveness of the Lanczos algorithm for the symmetric eigenproblem*, Linear Algebra Appl. 34 (1980), pp. 235–258.
[23] B.N. PARLETT, *The Symmetric Eigenvalue Problem*, Prentice-Hall, New York, 1980.
[24] B.N. PARLETT AND J.K. REID, *Tracking the Progress of the Lanczos Algorithm for Large Symmetric Eigenproblems*, IMA Journal of Numerical Analysis 1 (1981), pp. 135–155.
[25] B.N. PARLETT, *Misconvergence in the Lanczos Algorithm*, Res. Rep. PAM-404, U. of California, Berkeley (1987).
[26] D.J. SCOTT, *How to Make the Lanczos Algorithm Converge Slowly*, Mathematics of Comp. 33 (1979), pp. 239–247.
[27] A. VAN DER SLUIS, H.A. VAN DER VORST, *The Convergence Behavior of Ritz Values in the Presence of Close Eigenvalues*, Linear Algebra Appl. 88/89 (1987), pp. 651–694.
[28] Z. STRAKOŠ, *A Note on the Rate of Convergence of the Conjugate Gradient Method and the Convergence of Ritz values (contributed paper)*, NATO ASI "Computer Algorithms for Solving Linear Alg. Systems: the State of the Art", Il Ciocco, Sept. 9–22 (1990).
[29] Z. STRAKOŠ, *On the Real Convergence Rate of the Conjugate Gradient Method*, Linear Algebra Appl. 154–156 (1991), pp. 535–549.
[30] Z. STRAKOŠ, *Lanczos Algorithm, Orthogonal Polynomials and Continued Fractions*, Tech. Rep. 500, ICIS CAS, Praha (1991).
[31] B. WENDROFF, *On Orthogonal Polynomials*, Proc. of the AMS 12 (1961), pp. 554–555.
[32] J.H. WILKINSON, *Algebraic Eigenvalue Problem*, Clarendon Press, Oxford (1965).

Recent IMA Preprints

#	Author/s	Title
866	**Gui–Qiang Chen and Tai–Ping Liu**, Zero relaxation and dissipation limits for hyperbolic conservation laws	
867	**Gui–Qiang Chen and Jian–Guo Liu**, Convergence of second-order schemes for isentropic gas dynamics	
868	**Aleksander M. Simon and Zbigniew J. Grzywna**, On the Larché–Cahn theory for stress-induced diffusion	
869	**Jerzy Łuczka, Adam Gadomski and Zbigniew J. Grzywna**, Growth driven by diffusion	
870	**Mitchell Luskin and Tsorng-Whay Pan**, Nonplanar shear flows for nonaligning nematic liquid crystals	
871	**Mahmoud Affouf**, Unique global solutions of initial-boundary value problems for thermodynamic phase transitions	
872	**Richard A. Brualdi and Keith L. Chavey**, Rectangular L-matrices	
873	**Xinfu Chen, Avner Friedman and Bei Hu**, The thermistor problem with zero–one conductivity II	
874	**Raoul LePage**, Controlling a diffusion toward a large goal and the Kelly principle	
875	**Raoul LePage**, Controlling for optimum growth with time dependent returns	
876	**Marc Hallin and Madan L. Puri**, Rank tests for time series analysis a survey	
877	**V.A. Solonnikov**, Solvability of an evolution problem of thermocapillary convection in an infinite time interval	
878	**Horia I. Ene and Bogdan Vernescu**, Viscosity dependent behaviour of viscoelastic porous media	
879	**Kaushik Bhattacharya**, Self-accommodation in martensite	
880	**D. Lewis, T. Ratiu, J.C. Simo and J.E. Marsden**, The heavy top: a geometric treatment	
881	**Leonid V. Kalachev**, Some applications of asymptotic methods in semiconductor device modeling	
882	**David C. Dobson**, Phase reconstruction via nonlinear least-squares	
883	**Patricio Aviles and Yoshikazu Giga**, Minimal currents, geodesics and relaxation of variational integrals on mappings of bounded variation	
884	**Patricio Aviles and Yoshikazu Giga**, Partial regularity of least gradient mappings	
885	**Charles R. Johnson and Michael Lundquist**, Operator matrices with chordal inverse patterns	
886	**B.J. Bayly**, Infinitely conducting dynamos and other horrible eigenproblems	
887	**Charles M. Elliott and Stefan Luckhaus**, 'A generalised diffusion equation for phase separation of a multi-component mixture with interfacial free energy'	
888	**Christian Schmeiser and Andreas Unterreiter**, The derivation of analytic device models by asymptotic methods	
889	**LeRoy B. Beasley and Norman J. Pullman**, Linear operators that strongly preserve the index of imprimitivity	
890	**Jerry Donato**, The Boltzmann equation with lie and cartan	
891	**Thomas R. Hoffend Jr., Peter Smereka and Roger J. Anderson**, A method for resolving the laser induced local heating of moving magneto-optical recording media	
892	**E.G. Kalnins, Willard Miller, Jr. and Sanchita Mukherjee**, Models of q-algebra representations: the group of plane motions	
893	**T.R. Hoffend Jr. and R.K. Kaul**, Relativistic theory of superpotentials for a nonhomogeneous, spatially isotropic medium	
894	**Reinhold von Schwerin**, Two metal deposition on a microdisk electrode	
895	**Vladimir I. Oliker and Nina N. Uraltseva**, Evolution of nonparametric surfaces with speed depending on curvature, III. Some remarks on mean curvature and anisotropic flows	
896	**Wayne Barrett, Charles R. Johnson, Raphael Loewy and Tamir Shalom**, Rank incrementation via diagonal perturbations	
898	**Mingxiang Chen, Xu-Yan Chen and Jack K. Hale**, Structural stability for time-periodic one-dimensional parabolic equations	
899	**Hong-Ming Yin**, Global solutions of Maxwell's equations in an electromagnetic field with the temperature-dependent electrical conductivity	
900	**Robert Grone, Russell Merris and William Watkins**, Laplacian unimodular equivalence of graphs	
901	**Miroslav Fiedler**, Structure-ranks of matrices	
902	**Miroslav Fiedler**, An estimate for the nonstochastic eigenvalues of doubly stochastic matrices	
903	**Miroslav Fiedler**, Remarks on eigenvalues of Hankel matrices	
904	**Charles R. Johnson, D.D. Olesky, Michael Tsatsomeros and P. van den Driessche**, Spectra with positive elementary symmetric functions	
905	**Pierre-Alain Gremaud**, Thermal contraction as a free boundary problem	
906	**K.L. Cooke, Janos Turi and Gregg Turner**, Stabilization of hybrid systems in the presence of feedback delays	
907	**Robert P. Gilbert and Yongzhi Xu**, A numerical transmutation approach for underwater sound propagation	
908	**LeRoy B. Beasley, Richard A. Brualdi and Bryan L. Shader**, Combinatorial orthogonality	
909	**Richard A. Brualdi and Bryan L. Shader**, Strong hall matrices	
910	**Håkan Wennerström and David M. Anderson**, Difference versus Gaussian curvature energies; monolayer versus bilayer curvature energies applications to vesicle stability	
911	**Shmuel Friedland**, Eigenvalues of almost skew symmetric matrices and tournament matrices	
912	**Avner Friedman, Bei Hu and J.L. Velazquez**, A Free Boundary Problem Modeling Loop Dislocations in Crystals	
913	**Ezio Venturino**, The Influence of Diseases on Lotka-Volterra Systems	
914	**Steve Kirkland and Bryan L. Shader**, On Multipartite Tournament Matrices with Constant Team Size	

915 **Richard A. Brualdi and Jennifer J.Q. Massey**, More on Structure-Ranks of Matrices

916 **Douglas B. Meade**, Qualitative Analysis of an Epidemic Model with Directed Dispersion

917 **Kazuo Murota**, Mixed Matrices Irreducibility and Decomposition

918 **Richard A. Brualdi and Jennifer J.Q. Massey**, Some Applications of Elementary Linear Algebra in Combinations

919 **Carl D. Meyer**, Sensitivity of Markov Chains

920 **Hong-Ming Yin**, Weak and Classical Solutions of Some Nonlinear Volterra Integrodifferential Equations

921 **B. Leinkuhler and A. Ruehli**, Exploiting Symmetry and Regularity in Waveform Relaxation Convergence Estimation

922 **Xinfu Chen and Charles M. Elliott**, Asymptotics for a Parabolic Double Obstacle Problem

923 **Yongzhi Xu and Yi Yan**, An Approximate Boundary Integral Method for Acoustic Scattering in Shallow Oceans

924 **Yongzhi Xu and Yi Yan**, Source Localization Processing in Perturbed Waveguides

925 **Kenneth L. Cooke and Janos Turi**, Stability, Instability in Delay Equations *Modeling* Human Respiration

926 **F. Bethuel, H. Brezis, B.D. Coleman and F. Hélein**, Bifurcation Analysis of Minimizing Harmonic Maps Describing the Equilibrium of Nematic Phases Between Cylinders

927 **Frank W. Elliott, Jr.**, Signed Random Measures: Stochastic Order and Kolmogorov Consistency Conditions

928 **D.A. Gregory, S.J. Kirkland and B.L. Shader**, Pick's Inequality and Tournaments

929 **J.W. Demmel, N.J. Higham and R.S. Schreiber**, Block *LU* Factorization

930 **Victor A. Galaktionov and Juan L. Vazquez**, Regional Blow-Up in a Semilinear Heat Equation with Convergence to a Hamilton-Jacobi Equation

931 **Bryan L. Shader**, Convertible, Nearly Decomposable and Nearly Reducible Matrices

932 **Dianne P. O'Leary**, Iterative Methods for Finding the Stationary Vector for Markov Chains

933 **Nicholas J. Higham**, Perturbation theory and backward error for $AX - XB = C$

934 **Z. Strakos and A. Greenbaum**, Open questions in the convergence analysis of the lanczos process for the real symmetric eigenvalue problem

935 **Zhaojun Bai**, Error analysis of the lanczos algorithm for the nonsymmetric eigenvalue problem

936 **Pierre-Alain Gremaud**, On an elliptic-parabolic problem related to phase transitions in shape memory alloys

937 **Bojan Mohar and Neil Robertson**, Disjoint essential circuits in toroidal maps

938 **Bojan Mohar**, Convex representations of maps on the torous and other flat surfaces

939 **Bojan Mohar and Svatopluk Poljak** Eigenvalues in combinatorial optimization

940 **Richard A. Brualdi, Keith L. Chavey and Bryan L. Shader**, Conditional sign-solvability

941 **Roger Fosdick and Ying Zhang**, The torsion problem for a nonconvex stored energy function

942 **René Ferland and Gaston Giroux**, An unbounded mean-field intensity model: Propagation of the convergence of the empirical laws and compactness of the fluctuations

943 **Wei-Ming Ni and Izumi Takagi**, Spike-layers in semilinear elliptic singular Perturbation Problems

944 **Henk A. Van der Vorst and Gerard G.L. Sleijpen**, The effect of incomplete decomposition preconditioning on the convergence of conjugate gradients

945 **S.P. Hastings and L.A. Peletier**, On the decay of turbulent bursts

946 **Apostolos Hadjidimos and Robert J. Plemmons**, Analysis of p-cyclic iterations for Markov chains

947 **ÅBjörck, H. Park and L. Eldén**, Accurate downdating of least squares solutions

948 **E.G. Kalnins, William Miller, Jr. and G.C. Williams**, Recent advances in the use of separation of variables methods in general relativity

949 **G.W. Stewart**, On the perturbation of LU, Cholesky and QR factorizations

950 **G.W. Stewart**, Gaussian elimination, perturbation theory and Markov chains

951 **G.W. Stewart**, On a new way of solving the linear equations that arise in the method of least squares

952 **G.W. Stewart**, On the early history of the singular value decomposition

953 **G.W. STewart**, On the perturbation of Markov chains with nearly transient states

954 **Umberto Mosco**, Composite media and asymptotic dirichlet forms

955 **Walter F. Mascarenhas**, The structure of the eigenvectors of sparse matrices

956 **Walter F. Mascarenhas**, A note on Jacobi being more accurate than QR

957 **Raymond H. Chan, James G. Nagy and Robert J. Plemmons**, FFT-based preconditioners for Toeplitz-Block least squares problems

958 **Zhaojun Bai**, The CSD, GSVD, their applications and computations

959 **D.A. Gregory, S.J. Kirkland and N.J. Pullman**, A bound on the exponent of a primitive matrix using Boolean rank

960 **Richard A. Brualdi, Shmuel Friedland and Alex Pothen**, Sparse bases, elementary vectors and nonzero minors of compound matrices

961 **J.W. Demmel**, Open problems in numerical linear algebra

962 **James W. Demmel and William Gragg**, On computing accurate singular values and eigenvalues of acyclic matrices

963 **James W. Demmel**, The inherent inaccuracy of implicit tridiagonal QR

964 **J.J.L. Velázquez**, Estimates on the $(N-1)$-dimensional Hausdorff measure of the blow-up set for a semilinear heat equation