The Behaviour of the Finite Precision Lanczos Algorithm

Jens–Peter M. Zemke

Talk at the Gregynog Workshop
on Computation and Analytic Problems
in Spectral Theory

11–16 July 1999

Overview

Properties of the exact recurrence

Idea: Orthogonal reduction of $A = A^T \in \mathbb{R}^{n \times n}$.

$$A \longrightarrow Q^T A Q = T \in \mathbb{R}^{n \times n},$$

$$T = \begin{pmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \ddots & \ddots & \\ & \ddots & \ddots & \beta_{n-1} \\ & & \beta_{n-1} & \alpha_n \end{pmatrix},$$

$$Q^T Q = I.$$

Dense matrices: Householder/Givens

+ Stable Algorithm
+ Eigenvalues accurate $\rightarrow O(\|A\|\varepsilon)$

– Operation count $O(n^3)$
– Storage amount $O(n^2)$

Sparse Matrices: Iterative implementation
$\Rightarrow$ Lanczos Algorithm

Compute an invariant subspace:

$$AQ = QT, \quad \begin{array}{ll} A \in \mathbb{R}^{n\times n} & \text{selfadjoint,} \\ T \in \mathbb{R}^{m\times m} & \text{tridiagonal,} \\ Q \in \mathbb{R}^{n\times m} & \text{orthonormal.} \end{array}$$

Lanczos Algorithm:

$$\begin{aligned} \beta_0 q_0 &\equiv 0 \\ q_1 &= ? \\ k &= 1 \end{aligned}$$

Iterate

$$\begin{aligned} \alpha_k &= q_k^T A q_k \\ r_k &= (A - \alpha_k I) q_k - \beta_{k-1} q_{k-1} \\ \beta_k &= \|r_k\|_2 \\ q_{k+1} &= r_k / \beta_k \\ k &= k+1 \end{aligned}$$

until $\beta_k = 0$.

Governing equation:

$$\boxed{\beta_k q_{k+1} = (A - \alpha_k I) q_{k-1} - \beta_{k-1} q_{k-1}}$$

With

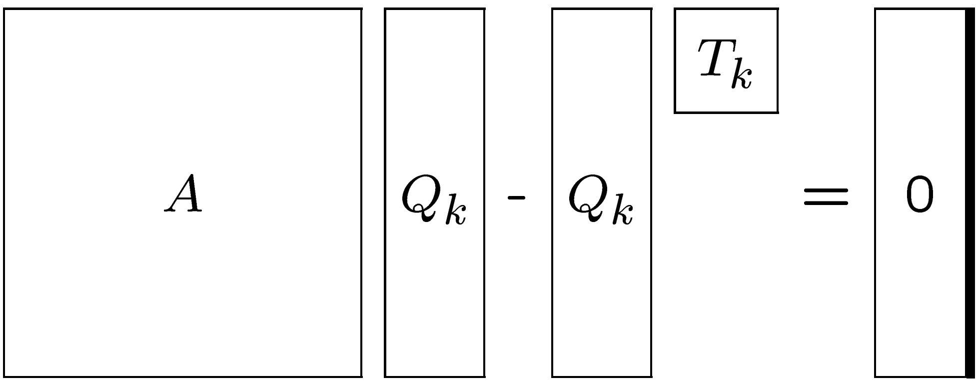
$$Q_k = \left[q_1, \ldots, q_n \right]$$

and

$$T_k = \begin{pmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \ddots & \ddots & \\ & \ddots & \ddots & \beta_{k-1} \\ & & \beta_{k-1} & \alpha_k \end{pmatrix}$$

one has in matrix form:

$$AQ_k - Q_k T_k = \beta_k q_{k+1} e_k^T = r_k e_k^T,$$



Eigendecomposition of T_k:

$$T_k S_k = S_k \Theta_k, \qquad \begin{array}{ll} S_k \in \mathbb{R}^{k\times k} & \text{orthogonal,} \\ \Theta_k \in \mathbb{R}^{k\times k} & \text{diagonal,} \end{array}$$

$$S_k = \left(s_{ij}^{(k)}\right)_{i,j\in\{1,\dots,k\}} = \left[s_1^{(k)}, \dots, s_k^{(k)}\right],$$

$$\Theta_k = \text{diag}\left(\theta_1^{(k)}, \dots, \theta_k^{(k)}\right).$$

Define Ritz pair:

$$\begin{array}{lll} y_j^{(k)} & = Q_k s_j^{(k)} & \text{(Ritz vector),} \\ \theta_j^{(k)} & & \text{(Ritz value).} \end{array}$$

Relation: Ritz pair $\longleftrightarrow$ Eigenpair?

$$\begin{array}{rl} & AQ_k - Q_k T_k = \beta_k q_{k+1} e_k^T \quad \Big| \cdot s_j^{(k)} \\ \Rightarrow & Ay_j^{(k)} - y_j^{(k)}\theta_j^{(k)} = \beta_k s_{kj}^{(k)} q_{k+1}. \end{array}$$

Eigendecomposition of A:

$$AV = V\Lambda, \quad \begin{array}{ll} V \in \mathbb{R}^{n \times n} & \text{orthogonal}, \\ \Lambda \in \mathbb{R}^{n \times n} & \text{diagonal}. \end{array}$$

$$\Lambda = \text{diag}\,(\lambda_1, \ldots, \lambda_n)$$

Residual bound:

$$\exists i : \left|\lambda_i - \theta_j^{(k)}\right| \leq \frac{\|Ay_j^{(k)} - y_j^{(k)}\theta_j^{(k)}\|_2}{\|y_j^{(k)}\|_2} = \beta_k \left|s_{kj}^{(k)}\right|.$$

Q_k is orthonormal, i.e.

$$\begin{aligned} \|Q_k s_j^{(k)}\|_2 &= \|[Q_k, Q_k^\perp][s_j^{(k)}; 0]\|_2 \\ &= \|[s_j^{(k)}; 0]\|_2 \\ &= \|s_j^{(k)}\|_2 \\ &= 1. \end{aligned}$$

Observation (numerical):

- The residual $\beta_k \left|s_{kj}^{(k)}\right|$ is not a sharp bound for simple eigenvalues.

Better bound (Temple/Kato):

No eigenvalue in $[\theta_j^{(k)}, \theta_j^{(k)} + \text{gap}]$. Then

$$0 \leq \lambda_i - \theta_j^{(k)} \leq \left(\beta_k s_{kj}^{(k)}\right)^2 \bigg/ \text{gap}.$$

Well separated eigenvalues (gap $= O(1)$)

$\Rightarrow$ convergence quadratic in the residual.

Cluster $\Lambda_i = \{\lambda_{i_1}, \ldots, \lambda_{i_l}\}$:

Convergence behaviour similar to case of simple eigenvalue until

$$\text{dist}\left(\Lambda_i, \theta_j^{(k)}\right)$$

reaches level

$$O(\text{diam}\,(\Lambda_i)\,\|A\|).$$

Eigenvalues sorted in ascending order:

$$\lambda_1 \leq \ldots \leq \lambda_n, \qquad \theta_1^{(k)} < \ldots < \theta_k^{(k)}.$$

A priori error bound (Kaniel-Saad):

$$0 \leq \frac{\theta_j^{(k)} - \lambda_j}{\lambda_n - \lambda_j} \leq \left(\frac{\sin \angle \left(q_1, V_j \right)}{\cos \angle \left(q_1, v_j \right)} \cdot \frac{\prod_{\nu=1}^{j-1} \left(\frac{\theta_\nu - \lambda_n}{\theta_\nu - \lambda_j} \right)}{T_{k-j}(1 + 2\gamma)} \right)^2.$$

$$V_j = \text{span} \left(v_1, \ldots, v_j \right)$$

Gap ratio:

$$\gamma = \frac{\lambda_j - \lambda_{j+1}}{\lambda_{j+1} - \lambda_n}.$$

Chebyshev polynomials T_k are used to dampen the unwanted part of the spectrum.

$\Rightarrow$ Convergence 'asymptotically' geometric for outer eigenvalues.

The finite precision recurrence

Relations in finite precision?

$\longrightarrow q_k, \alpha_k, \beta_k, \ldots$ denote computed quantities

Balance governing equation:

$$\beta_k q_{k+1} = (A - \alpha_k I) q_k - \beta_{k-1} q_{k-1} - \boxed{f_k}$$

Matrix form:

$$AQ_k - Q_k T_k = \beta_k q_{k+1} e_k^T + \boxed{F_k}$$

Disturbed residual:

$$AQ_k s_j^{(k)} - Q_k s_j^{(k)} \theta_j^{(k)} = \beta_k s_{kj}^{(k)} q_{k+1} + \boxed{F_k s_j^{(k)}}$$

Disturbed residual bound:

$$\exists i : \left|\lambda_i - \theta_j^{(k)}\right| \leq \frac{\beta_k \left|s_{kj}^{(k)}\right| + \boxed{O(\|A\|\varepsilon)}}{\|y_j^{(k)}\|_2}$$

$\Rightarrow$ unknown denominator.

Orthogonality of Lanczos vectors q_k,

$$Q_k^T Q_k - I_k = \text{?}$$

View matrix form

$$AQ_k - Q_k T_k = \beta_k q_{k+1} e_k^T \equiv E_k$$

as Sylvester equation for Q_k, i.e. as linear equation in $\mathbb{R}^{nk \times nk}$:

$$(I_k \otimes A - T_k \otimes I_n)\text{vec}(Q_k) = \text{vec}(E_k).$$

Eigenvalues:

$$\begin{aligned}\lambda_{ij}(I_k \otimes A - T_k \otimes I_n) &= \lambda_i(A) - \lambda_j(T_k) \\ &= \lambda_i - \theta_j^{(k)}.\end{aligned}$$

Condition:

$$\text{cond}_2(I_k \otimes A - T_k \otimes I_n) = \frac{\max_{i,j}\left(\left|\lambda_i - \theta_j^{(k)}\right|\right)}{\min_{i,j}\left(\left|\lambda_i - \theta_j^{(k)}\right|\right)}.$$

Errors are random

$\Rightarrow q_j$ loose orthogonality almost certainly.

Columns of Q_k become linear dependent

$\Rightarrow \|y_j^{(k)}\|_2 = \|Q_k s_j^{(k)}\|_2$ can be small.

Underlying structure in loss of orthogonality?

Explicit diagonalization:

$$AQ_k - Q_kT_k = \beta_k q_{k+1} e_k^T + F_k \quad \Big| \; v_i^T \cdot (*) \cdot s_j^{(k)},$$

$$(\lambda_i - \theta_j^{(k)}) v_i^T Q_k s_j^{(k)} = v_i^T q_{k+1} \beta_k s_{kj}^{(k)} + v_i^T F_k s_j^{(k)}.$$

Local error at step $k \to k+1$:

$$v_i^T q_{k+1} = \frac{(\lambda_i - \theta_j^{(k)}) v_i^T y_j^{(k)} - v_i^T F_k s_j^{(k)}}{\beta_k s_{kj}^{(k)}}.$$

Paige's result:

Loss of orthogonality $\Longleftrightarrow$ $F_k \neq 0$ and convergence

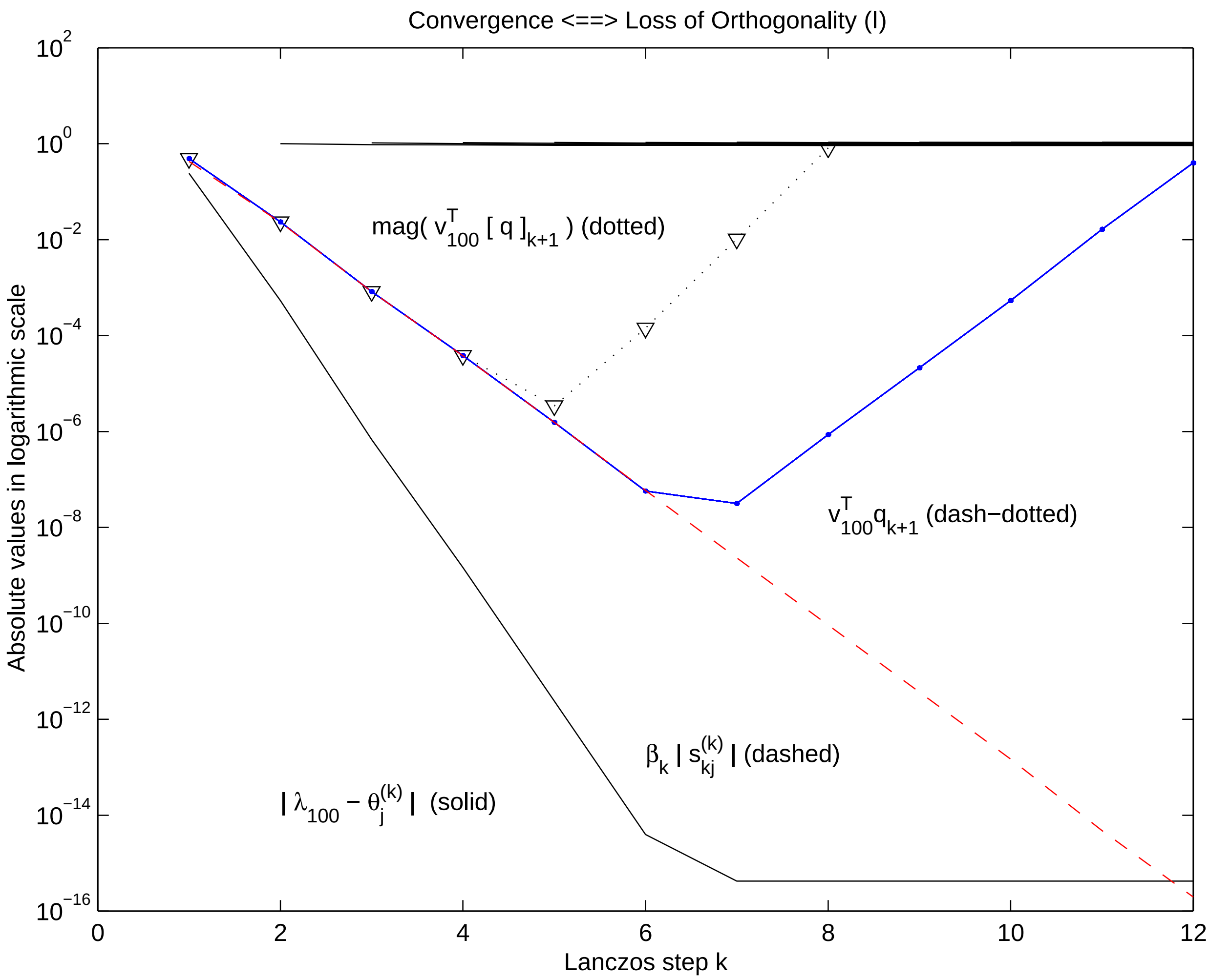
Absolute error level: $v_i^T F_k s_j^{(k)} = O(\|A\|\varepsilon)$.

Loss of orthogonality occurs when

$$\lambda_i - \theta_j^{(k)} = O\left(v_i^T F_k s_j^{(k)}\right) = O(\|A\|\varepsilon)$$

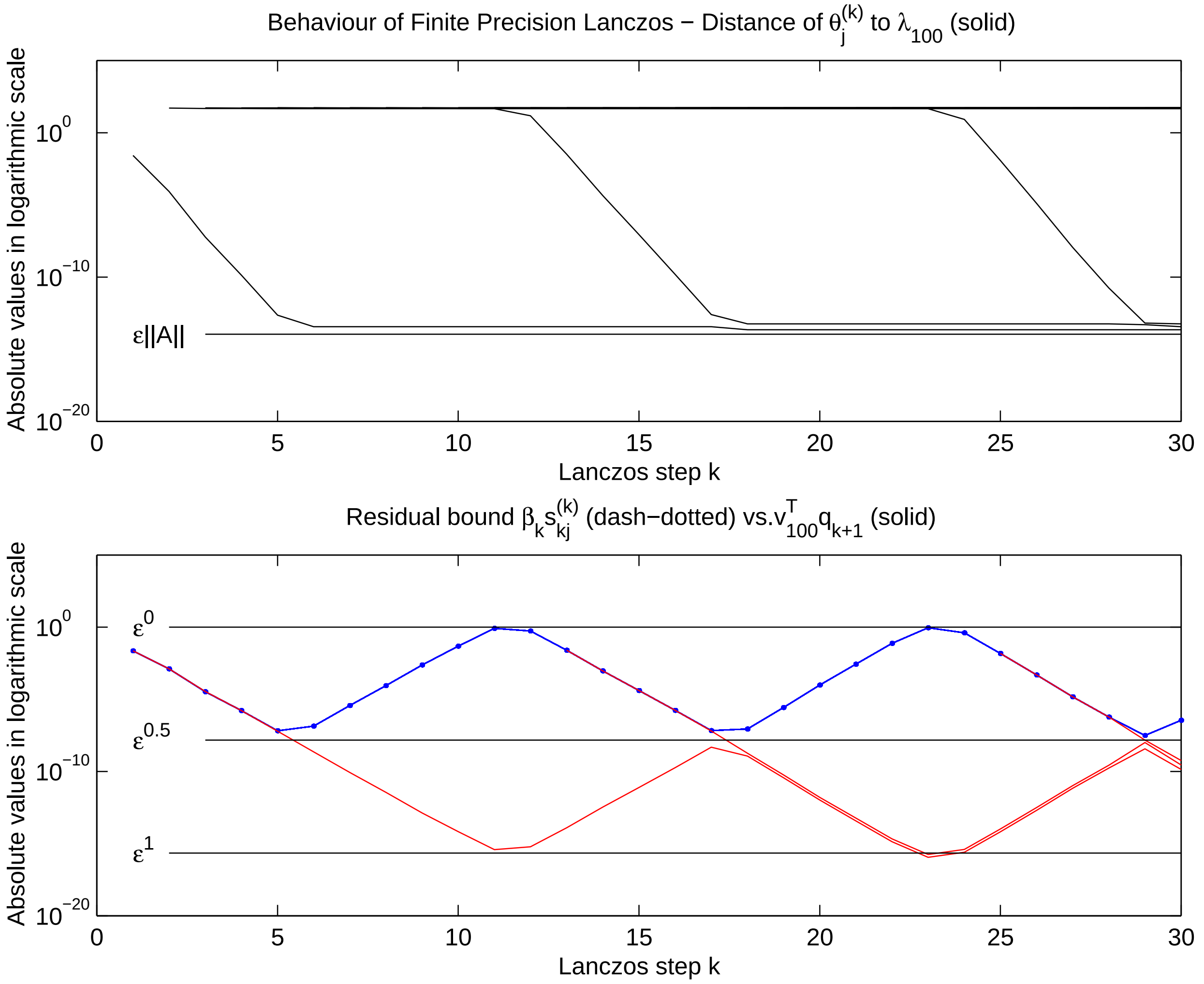
and is proportional to

$$\left(\beta_k s_{kj}^{(k)}\right)^{-1}.$$



A numerical example

Behaviour $\lambda_{\max} - \theta_j^{(k)}$ and $\beta_k s_{kj}^{(k)}$, $v_{\max}^T q_{k+1}$ for a random symmetric matrix $A \in [0,1]^{100\times 100}$.



Eigenvalues of largest moduli:

$$\lambda_{\max} = \lambda_{100} \approx 50.6, \qquad \lambda_{99} \approx 4.2.$$

Three phases of convergence to a given eigenvalue λ_i can be distinguished:

I Convergence (step 1-5):

$$\begin{aligned} \theta_j^{(k)} &\rightarrow \lambda_i, \\ \beta_k s_{kj}^{(k)} &\rightarrow O\left(\|A\|\sqrt{\varepsilon}\right), \\ v_i^T q_{k+1} &\rightarrow O\left(\sqrt{\varepsilon}\right). \end{aligned}$$

II Loss of orthogonality (step 5-11):

$$\begin{aligned} \theta_j^{(k)} &\approx \lambda_i, \\ \beta_k s_{kj}^{(k)} &\rightarrow O\left(\|A\|\varepsilon\right), \\ v_i^T q_{k+1} &\rightarrow O\left(1\right). \end{aligned}$$

III New Ritz value appears (step 11-17):

$$\begin{aligned} \theta_j^{(k)} &\approx \lambda_i, \\ \beta_k s_{kj}^{(k)} &\rightarrow O\left(\|A\|\sqrt{\varepsilon}\right), \\ v_i^T q_{k+1} &\rightarrow O\left(\sqrt{\varepsilon}\right). \end{aligned}$$

Interpretation based on:

Local loss of orthogonality:

$$v_i^T q_{k+1} = \frac{(\lambda_i - \theta_j^{(k)}) v_i^T y_j^{(k)} - v_i^T F_k s_j^{(k)}}{\beta_k s_{kj}^{(k)}}.$$

Stabilization of Ritz values:

$$\forall\ l > 0\ \exists\ i : \left| \theta_i^{(k+l)} - \theta_j^{(k)} \right| \leq \beta_k \left| s_{kj}^{(k)} \right|.$$

Thompson & McEnteggert (1968):

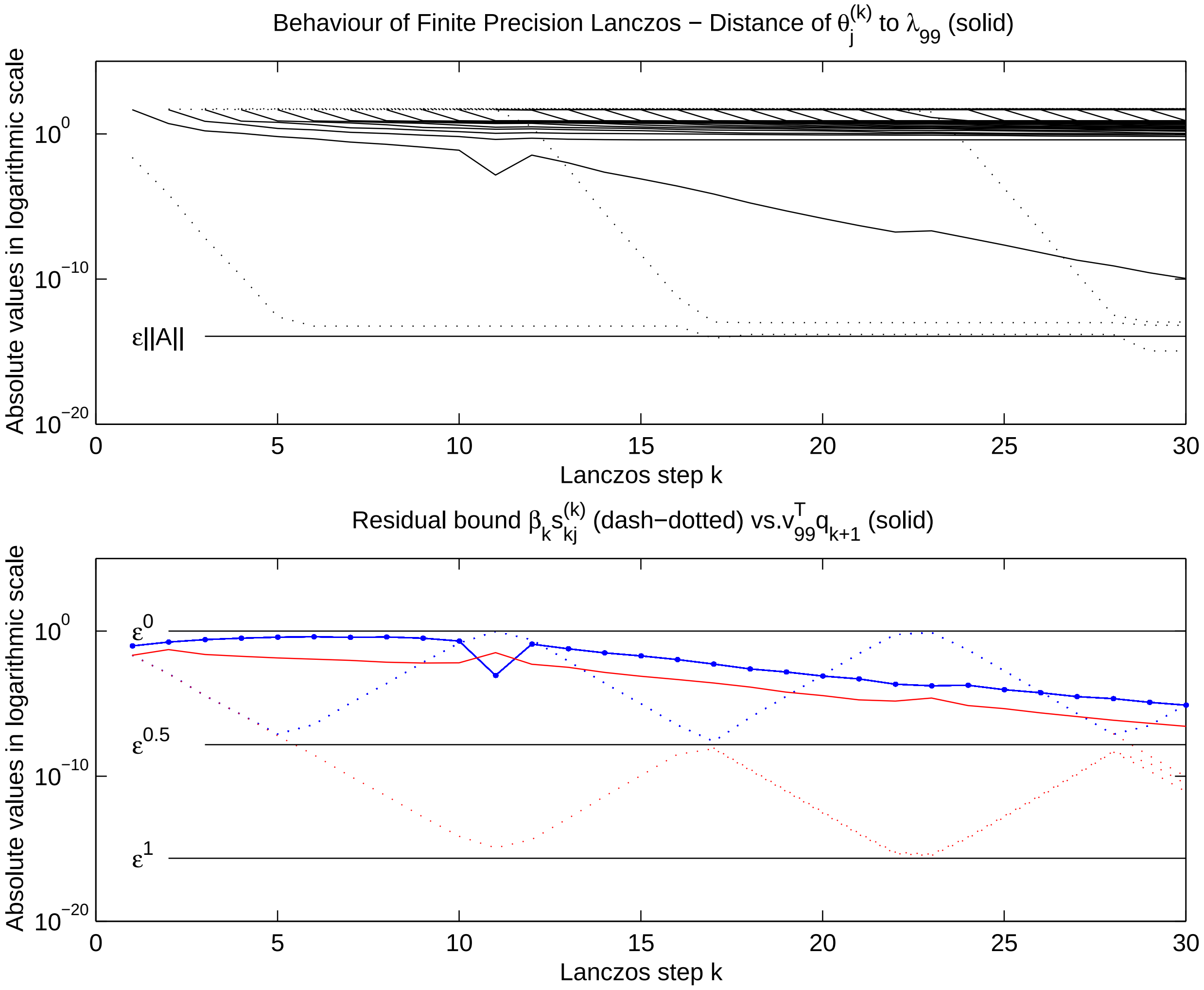
$$\left(s_{kj}^{(k)} \right)^2 = \prod_{i<j} \frac{\theta_j^{(k)} - \theta_i^{(k-1)}}{\theta_j^{(k)} - \theta_i^{(k)}} \cdot \prod_{i>j} \frac{\theta_j^{(k)} - \theta_{i-1}^{(k-1)}}{\theta_j^{(k)} - \theta_i^{(k)}}.$$

Perturbation theory:

$\Rightarrow$ Recurrence without stabilized Ritz value(s).

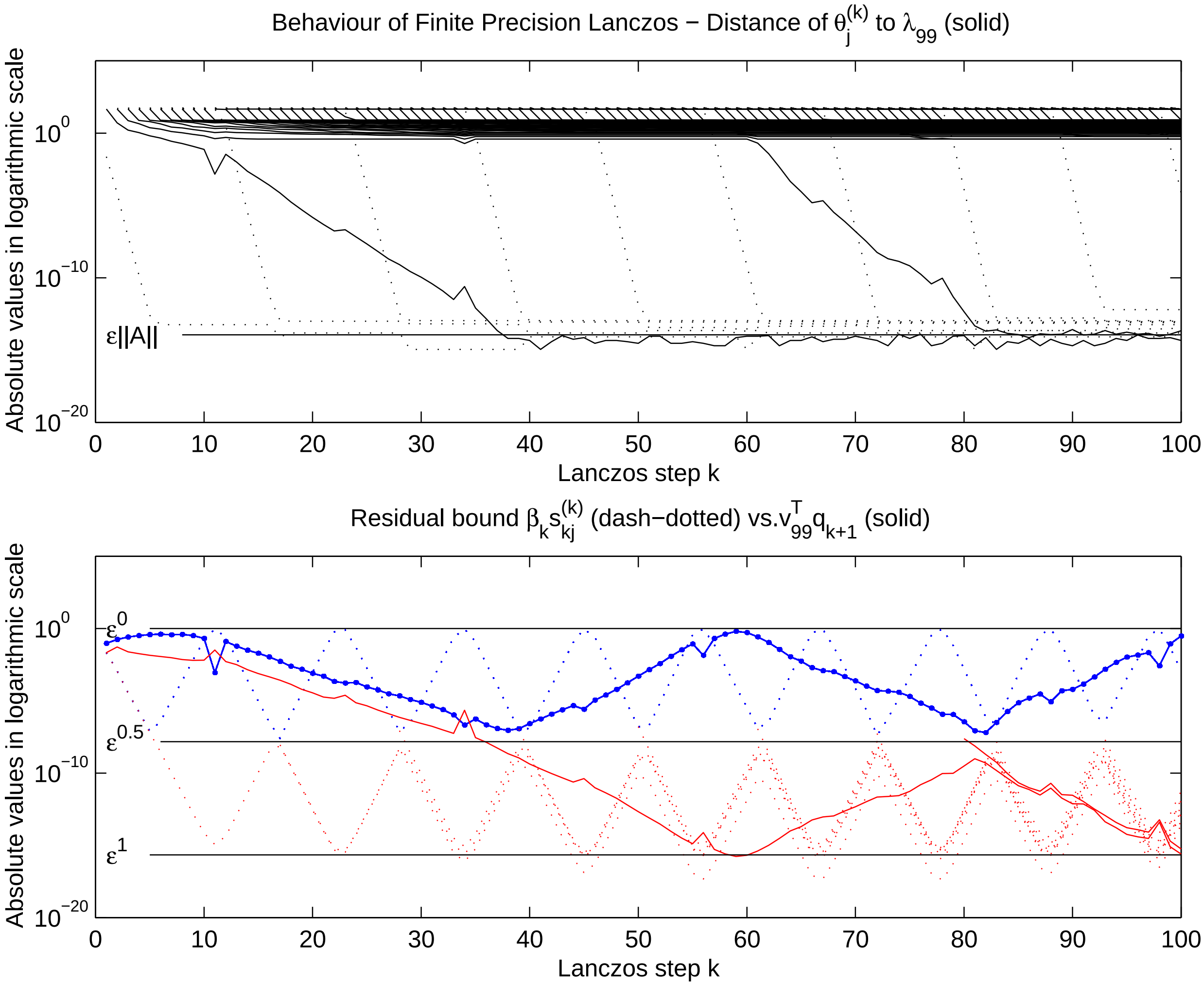
Secondary effects:

Convergence to eigenvalues close to stabilized Ritz values is perturbed.



Local behaviour becomes slightly perturbed.

Global behaviour governed by three phases.



Behaviour after perturbation occured

vs.

No perturbation occured

Perturbations dependent on absolute size.

Reorthogonalization techniques

Full Re**o**rthogonalization (LanFO):

Computes 'accurate' Q_k in $O(kn)$:

$$\|Q_k^T Q_k - I_k\| = O\left(\|A\|\varepsilon\right).$$

Semiorthogonalization techniques:

- **S**elective Re**o**rthogonalization (LanSO)
- **P**eriodic **R**eorthogonalization (LanPR)
- **P**artial **R**e**o**rthogonalization (LanPRO)

Indicator reaches $O(\|A\|\sqrt{\varepsilon})$

$\Rightarrow$ Invoke reorthogonalization.

Conclusion

- Lanczos' algorithm is **not** forward stable.
- Lanczos' algorithm **tends to 'forget'** the starting vector.
- Accuracy of Ritz pair dependent on number of Ritz pairs already accepted.
- Mixed **forward/backward analysis** gives useful insight.
- **Three phases model sufficient** to understand finite precision Lanczos.
- All known relations can be deduced without involved proofs.

Conclusion (extensions)

- Equivalent results for finite precision CG.

- The formula

$$v_i^T q_{k+1} = \frac{(\lambda_i - \theta_j^{(k)}) v_i^T y_j^{(k)} - v_i^T F_k s_j^{(k)}}{\beta_k s_{kj}^{(k)}}$$

  holds for **all** methods that

  - **compute the columns of the similarity transformation iteratively**.

- If these methods

  - **do not use reorthogonalization**

  we conclude that loss of convergence occurs **iff** the residual becomes small.