

Weak asymptotics for Ritz values of rational Krylov spaces

Bernhard Beckermann

<http://math.univ-lille1.fr/~bbecker>

Laboratoire Paul Painlevé UMR 8524 (équipe ANO-EDP)
UFR Mathématiques, Université de Lille 1

Rolling Waves

Leuven, December 15-16, 2008

Work in progress and joint discussions with
Stefan Güttel, Raf Vandebril and Lothar Reichel

Dear Adhemar

After having passed so many good moments together, for instance in Luminy 1989,



in Leuven over all these years, in Antwerp in 1993...

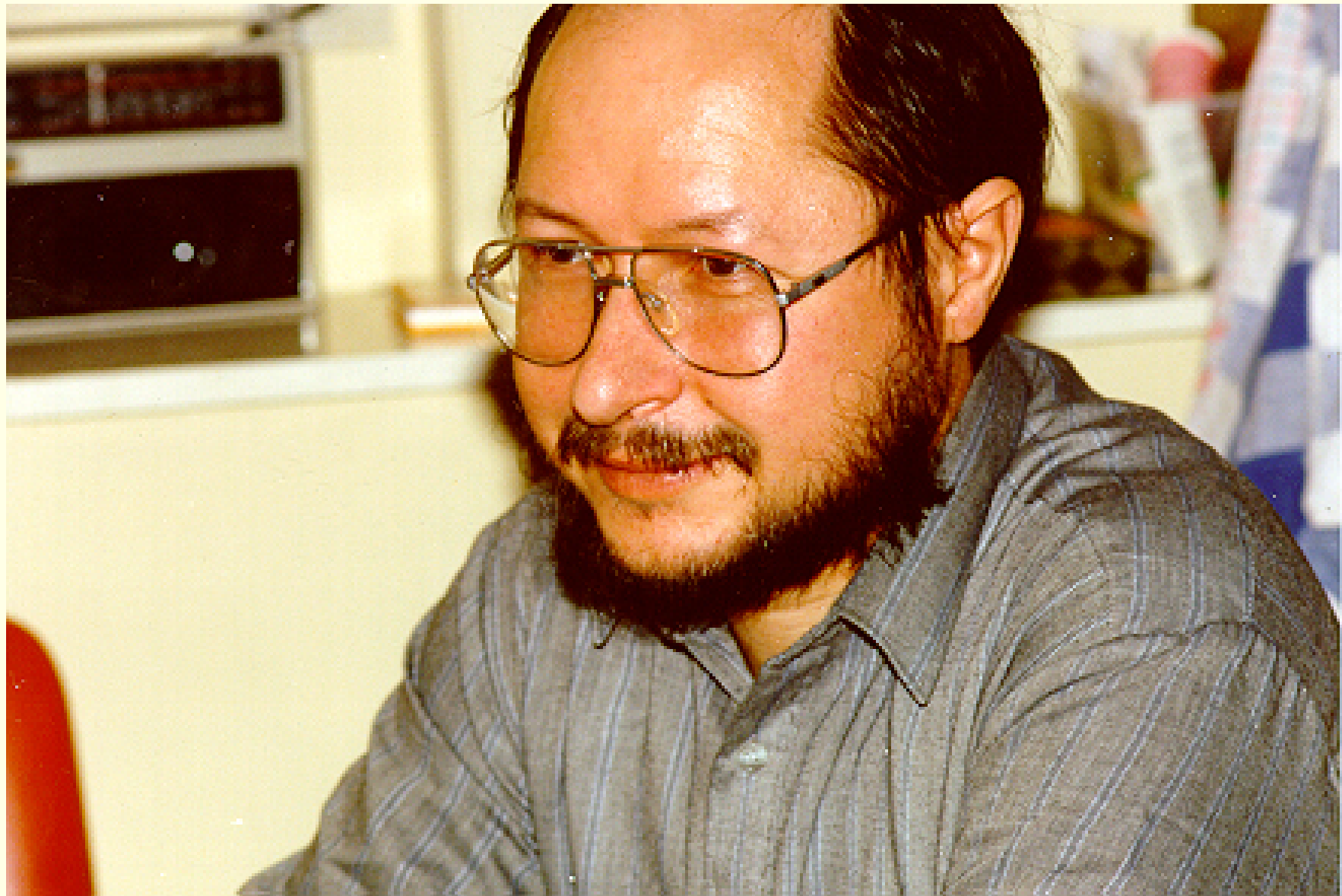


...at the ROLLS meetings in Leipzig 1996 and elsewhere, in Madrid (2002)



... or even recently in Lille (2008),





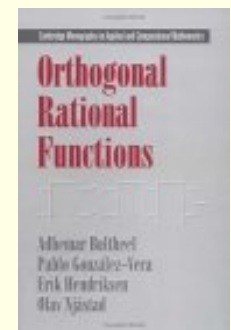
... I wish you a happy birthday ...

...and all the best for the future!!!

T_n are the zeros (i.e. the eigenvalues of J_n) of the orthonormal polynomial (OP) of degree n . These eigenvalues of T_n are called Arnoldi eigenvalue estimates (or Ritz values) with respect to the Krylov matrix $K_{N \times n}$ since some of these eigenvalues may be extraordinarily accurate approximations to some of the eigenvalues of A . It is well known that the approximation is better for the in absolute value largest eigenvalues of A . For the rational case, one might expect that some of the eigenvalues of T_n still are accurate approximations to some of the eigenvalues of A . In the more general rational case, however, the eigenvalues of T_n are not necessarily the zeros of the ORF φ_n , unless either $n = N$, $\mu_n = \infty$ or $q^n = \mathbf{0}$ (which for $n < N$ means that the Lanczos process breaks down at step n). Nevertheless, many observations strongly indicate that most of the zeros of φ_n do not differ much from the eigenvalues of T_n . Therefore, we believe that some of the zeros of φ_n are accurate approximations to some of the eigenvalues of A too. Consider for instance the case of one multiple pole $\mu_k = \mu \neq \infty$

Karl Deckers and Adhemar Bultheel,

Rational Krylov sequences and Orthogonal Rational Functions,
TW 499, KU Leuven (2007).



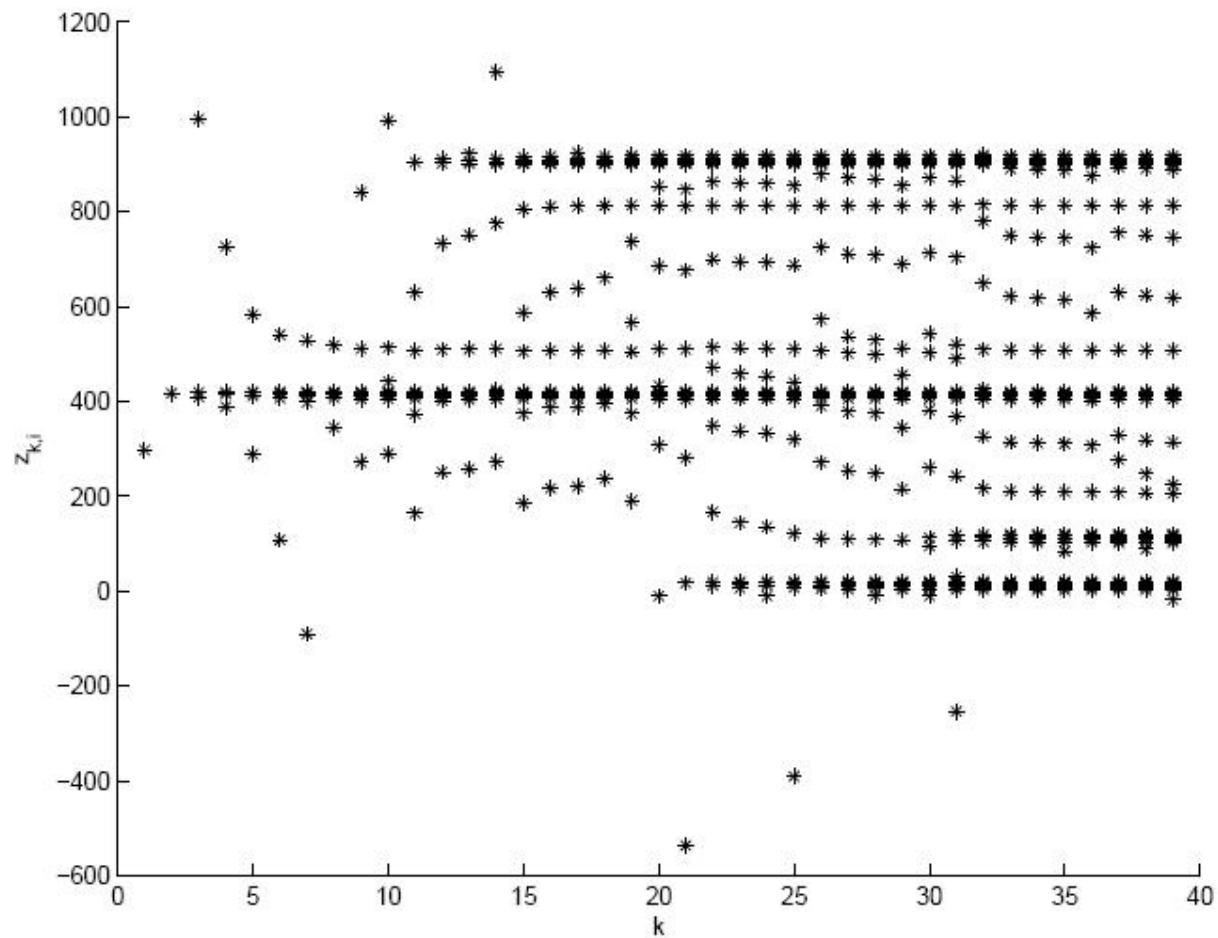


Figure 1: The zeros $z_{k,i}$ of $\varphi_k(z)$ for $k = 1, \dots, 39$ and $i = 1, \dots, k$ when A is given by (18) and the poles are given by (19).

... still the same paper!

Outline

- Ritz values are zeros of φ_n if...
 - Projections and Ritz values
 - Projections by rational Krylov sequences
 - ... the "next" pole should be ∞ ...
 - ... and the link with ORF...
- How do the "rational" Ritz values behave?
 - ... well, we stress the link with discrete OP...
 - ... a small bit of potential theory...
 - ... the preliminary results.

PART 1: Ritz values are zeros of φ_n if...

Projections and Ritz values

Given

- $A \in \mathbb{R}^{N \times N}$ "large sparse backslash matrix"
- $V_n \in \mathbb{R}^{N \times n}$ with orthonormal columns, $n \ll N$

we consider the projected counterpart $A_n = V_n^* A V_n \in \mathbb{C}^{n \times n}$.

Application 1: Do the n th **Ritz values** (spectrum $\Lambda(A_n)$ of A_n) approach part of $\Lambda(A)$?

Application 2: Can we approach the complicated $g(A)b$ by the simpler $V_n g(A_n) V_n^* b$?

Sometimes $V_n g(A_n) V_n^* b = p(A)b$ with p interpolating g at $\Lambda(A_n)$.

Example: $V_n = (v_1, v_2, \dots, v_n)$ orthonormal basis of Krylov subspace $\text{span}(b, Ab, A^2b, \dots, A^{n-1}b)$, computed by Arnoldi/Lanczos.

Rational Krylov spaces and rational Arnoldi

Given parameters $z_1, z_2, \dots \in \mathbb{C} \setminus \Lambda(A)$, say, $\neq 0$, let

$$\pi_k(z) = (1 - z/z_1)(1 - z/z_2)\dots(1 - z/z_k).$$

We consider the nested sequence of rational Krylov subspaces $\pi_k(A)^{-1}\text{span}(b, Ab, \dots, A^k b)$, with ONB v_1, v_2, \dots, v_{k+1} .

Put $v_1 = b/\|b\|$, and for $k = 1, \dots, n$ define v_{k+1} by orthogonalization of $(I - A/z_k)^{-1}Av_k$ against v_1, v_2, \dots, v_k , followed by normalization:

$$(I - A/z_k)^{-1}Av_k = h_{1,k}v_1 + \dots + h_{k,k}v_k + h_{k+1,k}v_{k+1}, \quad k = 1, 2, \dots, n.$$

Matrix language:

$$AV_n(I + H_n D_n) = V_n H_n + h_{n+1,n}(I - A/z_n)v_{n+1}(0, 0, \dots, 0, 1),$$

where $H_n = (h_{j,k}) \in \mathbb{R}^{n \times n}$ upper Hessenberg, $D_n = \text{diag}(1/z_1, 1/z_2, \dots, 1/z_n)$.

... but $v_1, v_2, \dots, v_n, A_n$ do not depend on z_n ! Formula simpler for $z_n = \infty$!

...the "next" pole should be ∞ ...

The artificial choice $z_n = \infty$ gives nicer formulas for A_n (LR & BB, 2008)

$$A_n = V_n^* A V_n = H_n (I + H_n D_n)^{-1},$$

compare with Fasino (2005), Bultheel & Deckers (2007).

Link with ORF

By definition, $v_k = \varphi_{k-1}(A)v_1$, with $\varphi_0(z) = 1$,

$$\frac{z}{1 - z/z_k} \varphi_{k-1}(z) = h_{1,k} \varphi_0(z) + \dots + h_{k,k} \varphi_{k-1}(z) + h_{k+1,k} \varphi_k(z),$$

where $\varphi_k = p_k/\pi_k$ with $\deg p_k \leq k$, and

$$v_j^* v_k = \delta_{j,k} = \langle \varphi_k, \varphi_j \rangle, \quad \langle p, q \rangle = \left(q(A)v_1 \right)^* p(A)v_1,$$

hence the φ_k are **Orthogonal Rational Functions** for poles z_1, \dots, z_{N-1} .

Denote by $\tilde{\varphi}_n$ the ORF for poles $z_1, z_2, \dots, z_{n-1}, z_n \leftarrow \infty$.

From the equation

$$\begin{aligned} z \left(\varphi_0(z), \dots, \varphi_{n-1}(z) \right) &= \left(\varphi_0(z), \dots, \varphi_{n-1}(z) \right) A_n \\ &\quad + \left(0, 0, \dots, 0, h_{n+1,n} \tilde{\varphi}_n(z) \right) (I + H_n D_n)^{-1} \end{aligned}$$

we see that $\Lambda(A_n) = \{\text{zeros of } \tilde{\varphi}_n\}$.

Simplification for symmetric A

If $A = A^*$ and real poles z_j then the machinery of ORF of Bultheel et al. gives

$$A_n = V_n^* A V_n = H_n (I + H_n D_n)^{-1} = J_n (I + \Delta_n J_n)^{-1}$$

with Jacobi matrix J_n and diagonal Δ_n .

Also for $A = A^*$ we have a simplified expression for our scalar product

$$\langle p, q \rangle = \left(q(A)v_1 \right)^* p(A)v_1 = \sum_{\lambda \in \Lambda(A)} w(\lambda)^2 p(\lambda) \overline{q(\lambda)}, \quad \langle 1, 1 \rangle = 1,$$

and the orthogonality for $\tilde{\phi}_n = \tilde{p}_n / \pi_n$ reads

$$\left\langle \frac{p}{\pi_{n-1}}, \tilde{\phi}_n \right\rangle = \sum_{\lambda \in \Lambda(A)} \frac{w(\lambda)^2}{|\pi_{n-1}(\lambda)|^2} p(\lambda) \tilde{p}_n(\lambda) = \begin{cases} 0 & \text{if } \deg p < n, \\ 1 & \text{if } p = \tilde{p}_n. \end{cases}$$

Asymptotic behavior of Ritz values?

... well, just like the behavior of zeros of discrete OP with varying weights!

... but there is no room for asymptotics since support $\Lambda(A)$ is finite!

Rakhmanov '96, Dragnev & Saff '97, Kuijlaars '00 and others:
look at sequences of scalar products/matrices!

But there are not enough Ritz values!

The orthogonality relation for $\tilde{\phi}_n = \tilde{p}_n/\pi_n$ tells us

- (a) Ritz values $\Lambda(A_n) \subset \text{conv}(\Lambda(A))$ eigenvalues,
- (b) two Ritz values are strictly separated by an eigenvalue,
- (c) there might be Ritz values far from $\Lambda(A)$
- (d) in regions where there are "few" eigenvalues
each one is matched by one (or sometimes 2) Ritz values.

But what means "**few**"? Quantification of Arno Kuijlaars (2000) for poles $z_1 = z_2 = \dots = \infty$:

Look at sequences of matrices $(A^{[N]})_{N \geq 0}$ with an "asymptotic eigenvalue distribution" described by some Borel measure σ . Solve some extremal problem in potential theory with σ as constraint, then "few" means that the constraint is active!

Let $A^{[N]} = (A^{[N]})^* \in \mathbb{C}^{N \times N}$, and $b^{[N]} \in \mathbb{C}^N$ together with finite poles $Z^{[N]} = \{z_1^{[N]}, z_2^{[N]}, \dots\}$ s.t.

(H1) spectra in compactum: $\exists S$ compact s.t. $\forall N : \Lambda(A^{[N]}) \subset S$.

(H2) asymptotic spectrum: $\chi_N(\Lambda(A^{[N]})) := \frac{1}{N} \sum_{\lambda \in \Lambda(A^{[N]})} \delta_\lambda \rightarrow \sigma$.

(H3) asymptotic behavior of poles: $\chi_N(Z^{[N]}) \rightarrow \nu$.

(H4) regularity: logarithmic potentials U^ν, U^σ continuous.

(H5) sufficiently large eigencomponents: $\forall \Lambda(A^{[N]}) \ni \lambda^{[N]} \rightarrow \bar{\lambda}$ with e.v. $v^{[N]}$ we have $\left| \frac{(v^{[N]})^* b^{[N]}}{\|v^{[N]}\| \|b^{[N]}\|} \right|^{1/N} \rightarrow 1$.

(H6) sufficiently separated eigenvalues: $\forall \Lambda(A^{[N]}) \ni \lambda^{[N]} \rightarrow \bar{\lambda}$ we have $\frac{1}{N} \sum_{\lambda \in \Lambda(A^{[N]}) \setminus \{\lambda^{[N]}\}} \log |\lambda^{[N]} - \lambda| \rightarrow \int \log |\bar{\lambda} - \lambda| d\sigma(\lambda)$.

(H7) poles sufficiently separated from spectrum: $\forall \Lambda(A^{[N]}) \ni \lambda^{[N]} \rightarrow \bar{\lambda}$ we have $\frac{1}{N} \sum_{\lambda \in Z^{[N]}} \log |\lambda^{[N]} - \lambda| \rightarrow \int \log |\bar{\lambda} - \lambda| d\nu(\lambda)$.

where as usual logarithmic potential $U^\mu(z) := \int \log\left(\frac{1}{|x-z|}\right) d\mu(x)$ and logarithmic energy $I(\mu) := \int \int \log\left(\frac{1}{|x-y|}\right) d\mu(x) d\mu(y)$.

The main Theorem

Under the assumptions (H1)–(H7) we have for the $n(N)$ th rational Ritz values for $n(N)/N \rightarrow t \in (0, \sigma(S))$

- (i) $\chi_N(\Lambda(A_{n(N)}^{[N]})) \rightarrow \mu$
- (ii) $\forall \Lambda(A^{[N]}) \ni \lambda^{[N]} \rightarrow \bar{\lambda} : \limsup_{N \rightarrow \infty} \text{dist}(\lambda^{[N]}, A_{n(N)}^{[N]})^{1/N} \leq \exp(U^{\mu-\nu}(\lambda) - F)$

with μ being the unique minimizer of $I(\mu - \nu)$ in

$$\mathcal{M}_t^\sigma = \{\rho \text{ Borel measure of mass } t : \rho \geq 0, \sigma - \rho \geq 0\}.$$

$\mu \in \mathcal{M}_t^\sigma$ and $F \in \mathbb{R}$ are uniquely characterized by the equilibrium conditions

$$U^{\mu-\nu} \leq F \text{ in } \mathbb{C}, \text{ with equality in } \text{supp}(\sigma - \mu).$$

Case "few poles": if $\sigma - \nu \geq 0$ then $\mu = \nu$ but no geometric convergence.

Case "many poles": if I open with $\nu|_I - \sigma|_I > 0$ then $I \cap \text{supp}(\sigma - \mu) = \emptyset$.

Further conclusions

... happy birthday, Adhemar!