Eigenpair Extraction Using Krylov Subspaces

Jens-Peter M. Zemke
zemke@tu-harburg.de

Institut für Numerische Simulation
Technische Universität Hamburg-Harburg

21.08.2007

Outline

Krylov decompositions

We consider a given Krylov decomposition

$$AQ_k = Q_{k+1}\underline{C}_k = Q_k C_k + q_{k+1} c_{k+1,k} e_k^T. \quad (1)$$

We suppose that

$$A \in \mathbb{C}^{(n,n)} \quad \text{is a general square matrix,}$$

$$Q_{k+1} = \begin{pmatrix} Q_k & q_{k+1} \end{pmatrix} \in \mathbb{C}^{(n,k+1)} \quad \text{is a matrix of basis vectors,}$$

$$\underline{C}_k = \begin{pmatrix} C_k \\ c_{k+1,k} e_k^T \end{pmatrix} \in \mathbb{C}^{(k+1,k)} \quad \text{is an extended Hessenberg matrix.}$$

We do not consider perturbations.

We remark that important parts of the results carry over to general rectangular approximations $\underline{C}_k$ which not necessarily have to be Hessenberg.

QOR for linear systems

In context of the approximate solution of linear systems

$$Ax = r_0, \qquad A \in \mathbb{C}^{n\times n}, \quad r_0 \in \mathbb{C}^n, \tag{2}$$

we consider the QOR approximation obtained by

$$x_k = Q_k z_k, \quad z_k = C_k^{-1} e_1 \|r_0\|, \tag{3}$$

where we have used the starting vector

$$q_1 = \frac{r_0}{\|r_0\|}. \tag{4}$$

We remark that this approximation need not exist.

The residual is always parallel to the next basis vector.

QMR for linear systems

In context of the approximate solution of linear systems

$$Ax = r_0, \qquad A \in \mathbb{C}^{n \times n}, \quad r_0 \in \mathbb{C}^n, \tag{5}$$

we also consider the QMR approximation obtained by

$$\underline{x}_k = Q_k \underline{z}_k, \quad \underline{z}_k = \underline{C}_k^{\dagger} \underline{e}_1 \|r_0\|, \tag{6}$$

where we have again used the starting vector

$$q_1 = \frac{r_0}{\|r_0\|}. \tag{7}$$

We remark that this approximation always exists.

The residual is unrelated to the next basis vector.

Ritz-Galërkin

In context of the approximate computation of eigenpairs

$$Av = v\lambda, \qquad A \in \mathbb{C}^{n\times n}, \quad v \in \mathbb{C}^n, v \neq o \quad \lambda \in \mathbb{C}, \tag{8}$$

we consider the Ritz-Galërkin approach

$$(\theta, y = Q_k s), \qquad C_k s = s\theta, \quad e_k^T s = 1. \tag{9}$$

The residual of the Ritz pair $(\theta, y = Q_k s)$ is parallel to the next basis vector.

Relation to the QOR residual polynomials:

$$R_k(\theta) = \det(I_k - \theta C_k^{-1}) = 0 \quad \Rightarrow \quad \chi_k(\theta) = \det(C_k - \theta I_k) = 0. \tag{10}$$

Fact I: Ritz pairs are shift invariant.

Fact II: Ritz values are (unsymmetric) Rayleigh quotients of Ritz vectors.

Harmonic Ritz

Related to QMR are the quasi-kernel polynomials (Freund, Manteuffel & Otto),

$$\underline{R}_k(z) = \det(I_k - z\underline{C}_k^\dagger \underline{I}_k). \tag{11}$$

The roots of $\underline{R}_k$ approximate the eigenvalues of A.

This is a variation of Ritz-Galërkin (Morgan); eigenpairs are called harmonic Ritz pairs (Paige, Parlett & van der Vorst, Sleijpen & van der Vorst).

The harmonic Ritz pairs $(\underline{\theta}, \underline{y} = Q_k \underline{s})$ are eigenpairs of

$$(I_k, \underline{C}_k^\dagger \underline{I}_k) \equiv (\underline{C}_k^H \underline{C}_k, C_k^H) \equiv (R_k, Q_k^H), \qquad \underline{C}_k = \underline{Q}_k R_k, \quad Q_k = \underline{I}_k^T \underline{Q}_k. \tag{12}$$

Harmonic Ritz vectors are superior to Ritz vectors for the “interior” of the spectrum.

Possibility of infinite harmonic Ritz values: homogeneous form.

Shifted harmonic Ritz

Harmonic Ritz values and vectors are not shift-invariant.

We can consider a shift or "target" τ:

$$\underline{R}_k^{(\tau)}(z) = \det(I_k - (z+\tau)(\underline{C}_k - \tau \underline{I}_k)^{\dagger} \underline{I}_k) = 0. \tag{13}$$

Harmonic Ritz values are usually not the Rayleigh quotient of the corresponding harmonic Ritz vector.

We might replace the harmonic Ritz value by this Rayleigh quotient, the so-called ρ-value.

The question is which shift is "optimal" in which sense.

Choosing the shift as Ritz value gives infinite harmonic Ritz values.

Jia's refined approaches

In a couple of papers Jia advocated the use of refinement.

Given an approximate eigenpair $(\tilde{\theta}, \tilde{y} = Q_k \tilde{s})$, discard the vector $\tilde{s}$ and compute the minimizer (SVD)

$$\acute{s} = \underset{s \in \mathbb{C}^k, \|s\|=1}{\arg\min} \|(\tilde{\theta} \underline{I}_k - \underline{C}_k) s\|. \tag{14}$$

The obvious methods are refined Ritz, refined harmonic Ritz and refined ρ-values.

Other methods try to find a "good" shift in harmonic Ritz, use the minimization property (Rayleigh) of the ρ-value to discard the harmonic Ritz value,

$$\rho = \acute{\theta} = \underset{\theta \in \mathbb{C}}{\arg\min} \|(\theta \underline{I}_k - \underline{C}_k) \tilde{s}\|, \tag{15}$$

and then use refinement.

So what's next: Rayleigh of refined Rayleigh of refined Rayleigh of . . . ???

QMR eigenpairs

We proceed similar to the QMR approach when applied to linear systems,

$$\min_{z,y=Q_k s} \|zy - Ay\| \leqslant \|Q_{k+1}\| \cdot \min_{z,s} \|(z\underline{I}_k - \underline{C}_k)s\|. \tag{16}$$

We always suppose that the columns of Q_{k+1} have been scaled to unit length.

Definition (QMR eigenpair)

The pair $(\grave{\theta}, \grave{y} = Q_k \grave{s})$ is a QMR eigenpair, when

$$\|(\grave{\theta}\underline{I}_k - \underline{C}_k)\grave{s}\| = \min_{z\in\mathbb{C}, s\in\mathbb{C}^k, \|s\|=1} \mathrm{loc} \, \|(z\underline{I}_k - \underline{C}_k)s\|, \tag{17}$$

where “min loc” denotes a (not necessarily strict) local minimum.

Remark: The correct amplification factor due to a non-unitary basis is

$$\sigma_1(Q_{k+1})/\sigma_k(Q_k) \leqslant \kappa(Q_{k+1}). \tag{18}$$

QMR eigenpairs: characterization I

Denote the SVD of $z\underline{I}_k - \underline{C}_k$ by $U(z)\Sigma(z)V(z)^H = z\underline{I}_k - \underline{C}_k$.

Then the QMR eigenvalues can be characterized by

$$\grave{\theta} = \underset{z \in \mathbb{C}}{\arg\min\operatorname{loc}}\, \sigma_k(z), \tag{19}$$

and the QMR eigenvector $\grave{s} = v_k(\grave{\theta})$ can be chosen as any corresponding right singular vector. The QMR residual is given by $\sigma_k(\grave{\theta})$.

Results on real-analyticity of the SVD of complex matrices (Sun) can be used to obtain steepest descent (minus conjugate gradient):

$$z_{\text{new}} = z - \alpha\overline{u_k^H \underline{I}_k v_k} = z - \alpha v_k^H \underline{I}_k^H u_k \tag{20}$$

$$= z - \frac{\alpha}{\sigma_k} v_k^H \underline{I}_k^H (z\underline{I}_k - \underline{C}_k) v_k = z - \frac{\alpha}{\sigma_k} v_k^H (zI_k - C_k) v_k. \tag{21}$$

Setting $\alpha = \sigma_k$ yields alternating projections and is nearly optimal:

$$z_{\text{new}} = v_k^H C_k v_k. \tag{22}$$

QMR eigenpairs: characterization II (rough sketch)

The results on real-analyticity of the SVD of complex matrices (Sun) can also be used to adopt Newton's method. This poses problems with clustered and multiple smallest singular values and when far from the solution.

Instead we use the real-analyticity of

$$\sigma_k^2(s) \equiv \|(s^H C_k s \underline{I}_k - \underline{C}_k)s\|^2 = s^H (\underline{C}_k)^H \underline{C}_k s - |s^H C_k s|^2. \tag{23}$$

Stationary points on the unit sphere are necessarily singular vectors. When the associated singular value is simple, it is a stationary point of the singular value curve, thus $\grave{s} = s$ is a QMR eigenvector whenever we are on the kth curve. The associated QMR eigenvalue is $\grave{\theta} = \grave{s}^H C_k \grave{s}$.

We experimented with steepest descent and Newton's method for the minimization of real-analytic $\sigma_k(s)$ on the (complex) unit sphere (a Grassmannian) in the framework of optimization on Riemannian manifolds (Smith, Edelman + Arias & Smith, Manton).

A graphical representation

We associate with every real or complex approximate eigenpair $(\tilde{\theta}, \tilde{y} = Q_k \tilde{s})$ a point (z, w) in the plane $\mathbb{R}^2$ or $\mathbb{C} \times \mathbb{R}$:

$$z = \tilde{\theta}, \quad w = \frac{\|(\tilde{\theta}\underline{I}_k - \underline{C}_k)\tilde{s}\|}{\|\tilde{s}\|}. \tag{24}$$

The former gives the approximate eigenvalue, the latter gives the norm of the (quasi-)residual of the approximate eigenpair.

The norm of the residual of an eigenpair gives the backward error, i.e.,

$$w = \min \left\{ \|\Delta A\| \;:\; (A + \Delta A)\tilde{y} = \tilde{y}\,\tilde{\theta} \right\}. \tag{25}$$

Without additional knowledge a small backward error is the best we can achieve.

There exist “graphical” bounds for general and “Rayleigh” approximations.

A beautiful example

The first example is

$$\underline{C}_k = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{26}$$

Its Ritz values are given by

$$\theta_{1,3} = \mp\sqrt{2} \approx \mp 1.41421356, \quad \theta_2 = 0, \tag{27}$$

its harmonic Ritz values are given by

$$\underline{\theta}_{1,3} = \mp\sqrt{2} \approx \mp 1.41421356, \quad \underline{\theta}_2 = \infty, \tag{28}$$

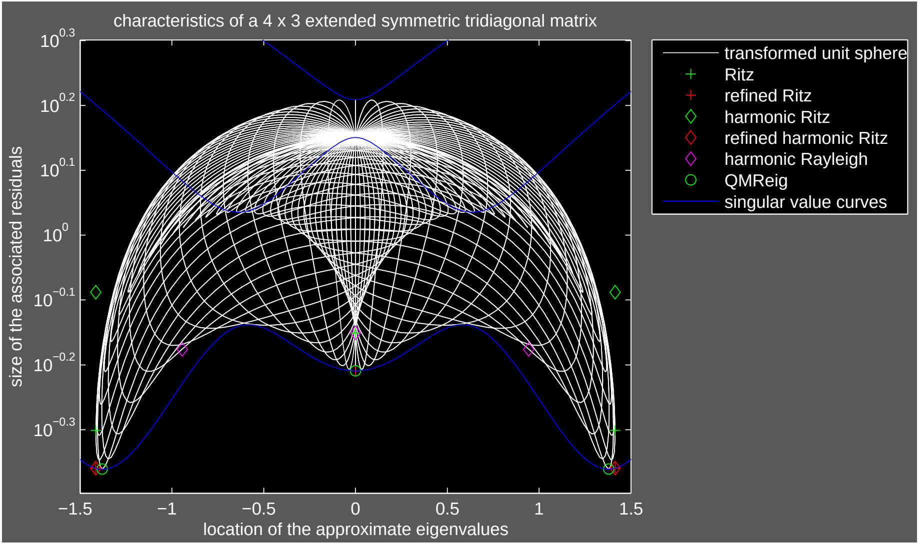
its ρ-values are given by

$$\rho_{1,3} = \mp\sqrt{2} \cdot \frac{2}{3} \approx \mp 0.9428090, \quad \rho_2 = 0, \tag{29}$$

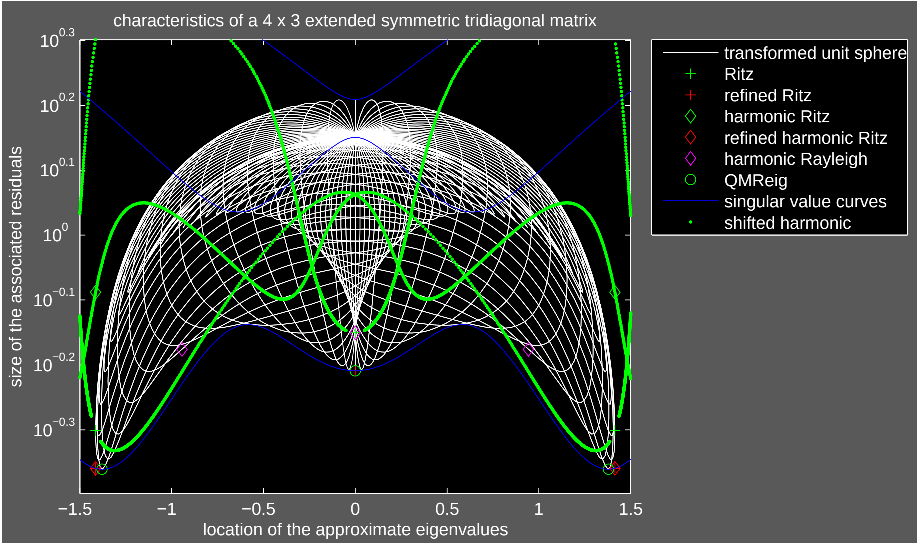
and its QMR eigenvalues are given by (where $y = 276081 + 21504\sqrt{2}i$)

$$\grave{\theta}_{1,3} = \mp\frac{\sqrt{2}}{16}\sqrt{113 + 2\Re\sqrt[3]{y}} \approx \mp 1.37898323557, \quad \grave{\theta}_2 = 0. \tag{30}$$

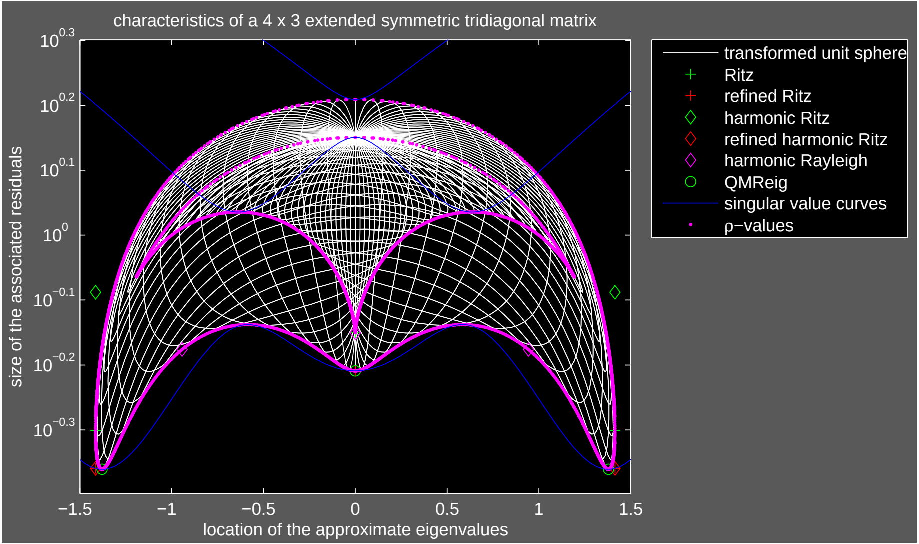
A beautiful example



A beautiful example



A beautiful example



A beautiful example

Short digression . . .

Why are the ρ-values on the "borders" of the transformed unit sphere?

In the symmetric case it is easy to characterize these "borders" and to prove that the vectors defining them are indeed harmonic Ritz vectors for two certain shifts. These shifts as well as the harmonic Ritz values are expressed using stationary points along straight lines in the graphical interpretation . . .

Given the harmonic Ritz pair, it is even easier to find the direction along which the vector is a stationary point.

The QMR eigenvectors are harmonic Ritz vectors, the shifts are given by

$$\tau_{\pm} = \grave{\theta} \pm \sigma_k(\grave{\theta}). \tag{31}$$

The Ritz vectors are harmonic Ritz vectors with shifts $\tau_{\pm} = \pm\infty$.

A warning: loss of QMR eigenvalues; academic

The second example is

$$\underline{C}_k = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \qquad (32)$$

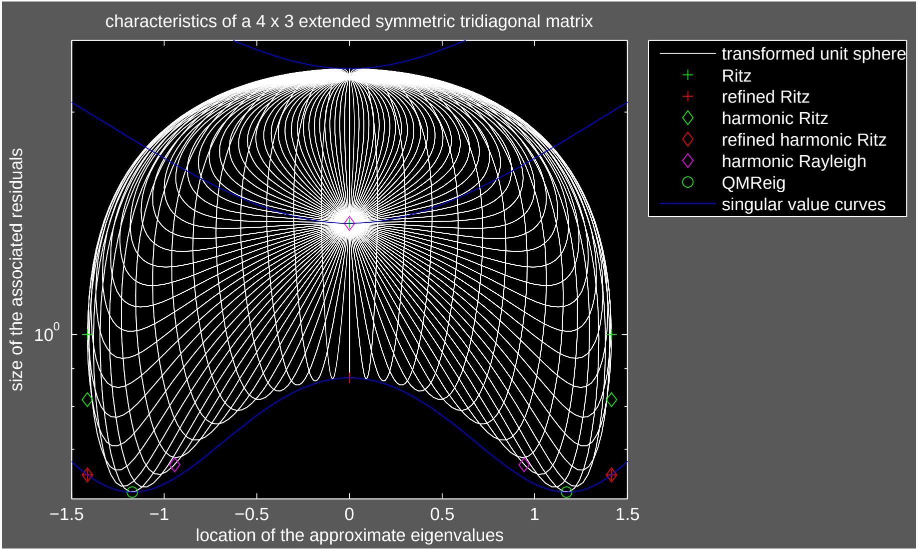
In this special case (one Ritz value being zero), replacing "2" by an arbitrary number does not alter the Ritz, harmonic Ritz and ρ-values.

Yet, this small change results in $\underline{C}_k$ having only two QMR eigenvalues, which are given by

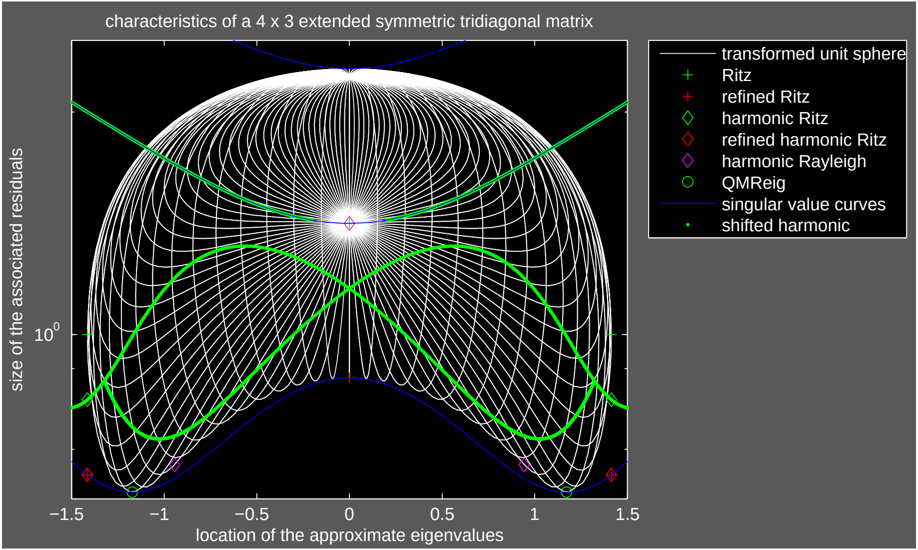
$$\grave{\theta}_{1,2} = \mp\sqrt{\frac{11}{8}} \approx \mp 1.17260393996. \qquad (33)$$

At zero a stationary point exists which is a maximum of the smallest singular value curve and a saddle point of the transformed sphere.

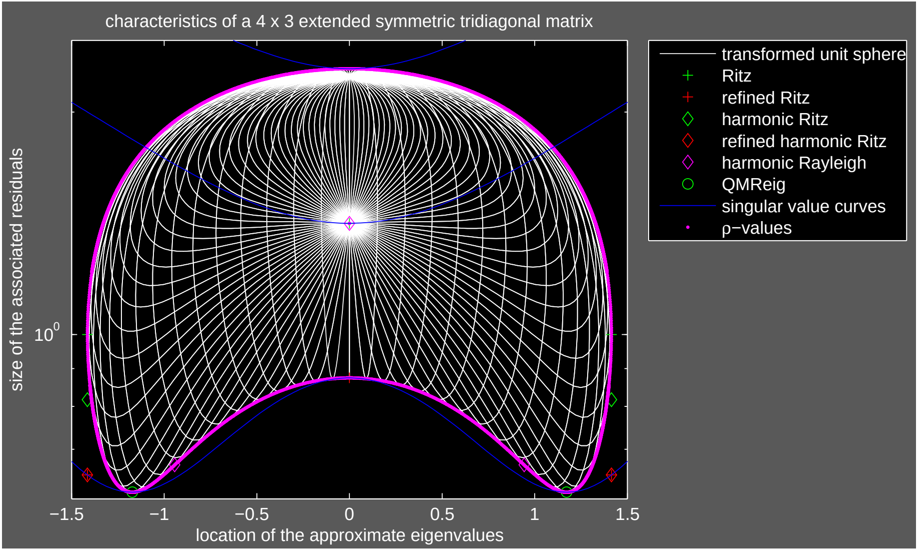
A warning: loss of QMR eigenvalues; academic



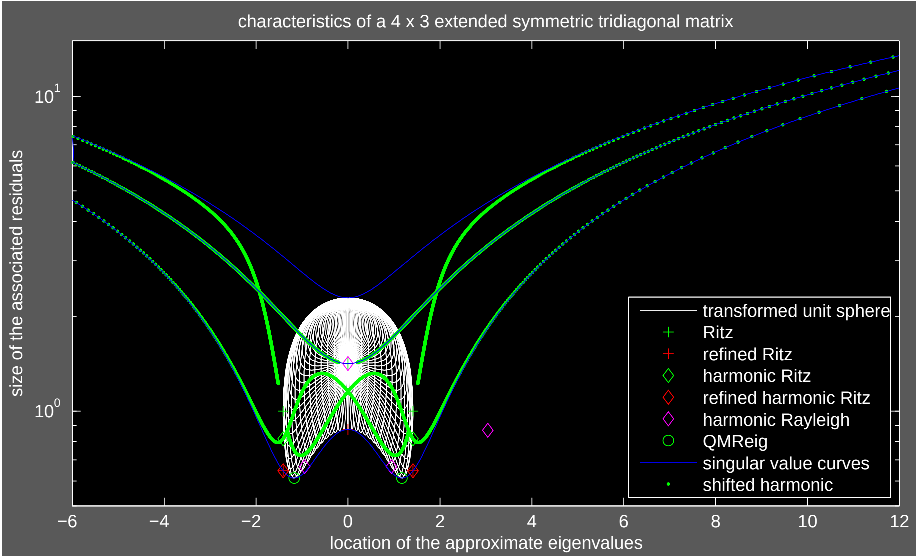
A warning: loss of QMR eigenvalues; academic



A warning: loss of QMR eigenvalues; academic



A warning: loss of QMR eigenvalues; academic



A warning: loss of QMR eigenvalue; realistic example

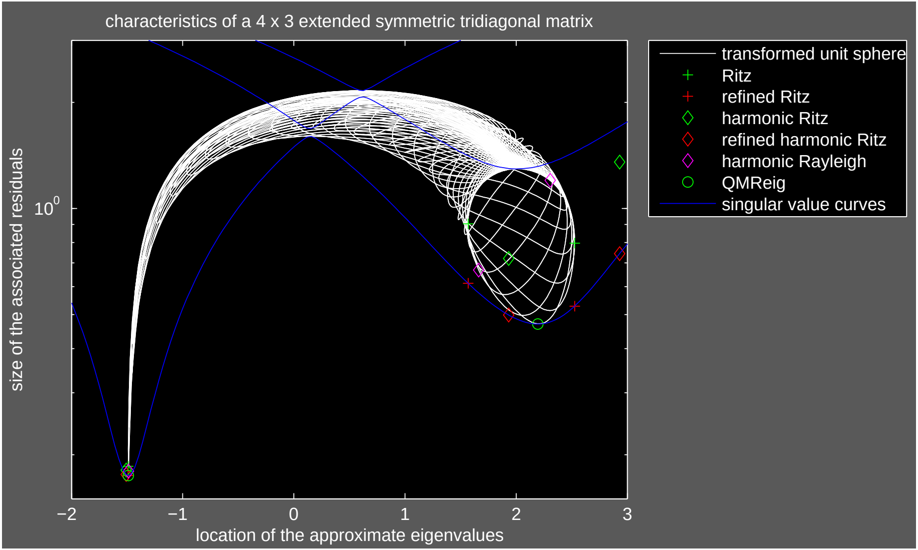
The third example is extended symmetric and generated using MATLAB's `randn` and `hess` functions and is approximately given by

$$\underline{C}_k \approx \begin{pmatrix} 0.46204801 & 1.75649255 & 0 \\ 1.75649255 & 0.23525002 & -0.70301190 \\ 0 & -0.70301190 & 1.90702012 \\ 0 & 0 & 1.21958322 \end{pmatrix}. \tag{34}$$

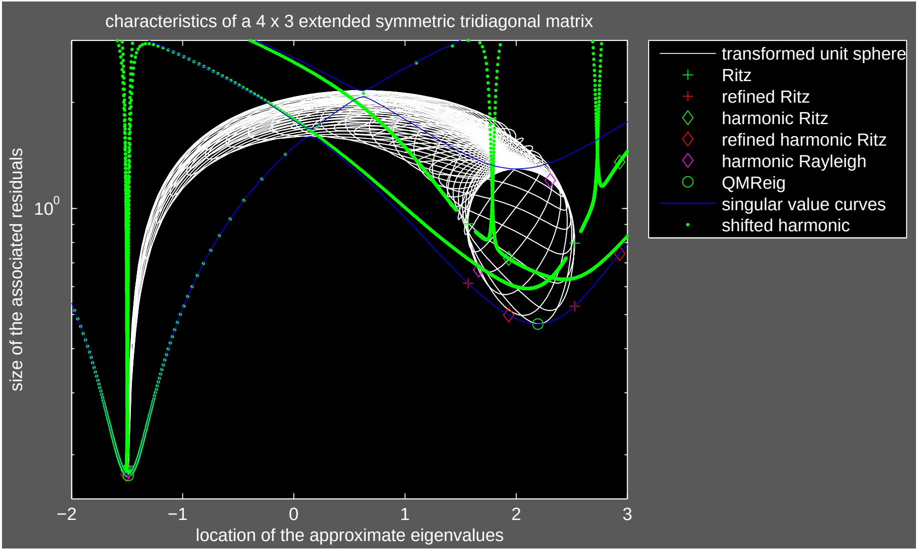
The computed Ritz, harmonic Ritz and ρ-values all differ. There are only two QMR eigenvalues. The smallest of all these and the norms of the eigenpair residuals (denoted by $n(\cdot,\cdot)$) are approximately given by

$$\begin{aligned}
\theta_1 &\approx -1.490413407713866, & n(\theta_1, s_1) &\approx 0.185432088955641\,7, \\
\underline{\theta}_1 &\approx -1.509143602001304, & n(\underline{\theta}_1, \underline{s}_1) &\approx 0.181039457164899\,5, \\
\rho_1 &\approx -1.487425797938723, & n(\rho_1, \underline{s}_1) &\approx 0.179732084050847\,2, \\
\grave{\theta}_1 &\approx -1.489367749116040, & n(\grave{\theta}_1, \grave{s}_1) &\approx 0.174658339265659\,0.
\end{aligned}$$

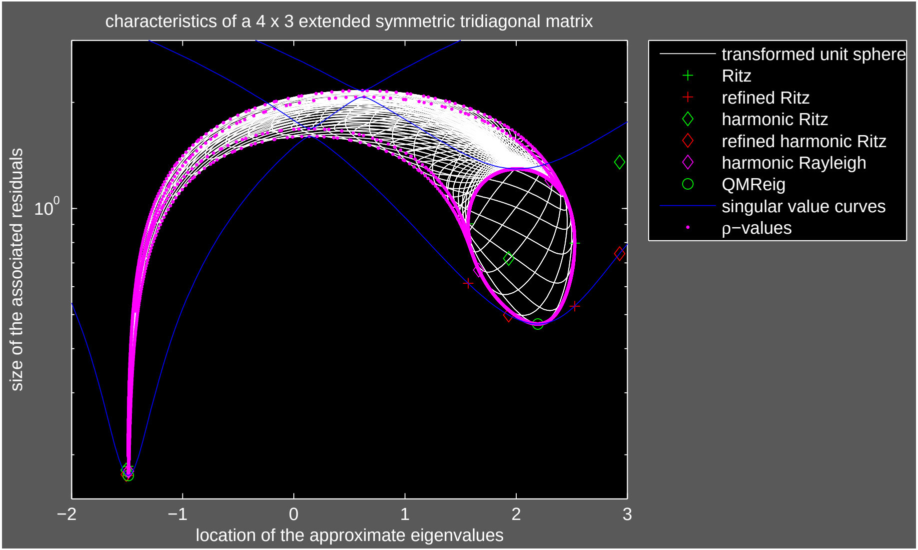
A warning: loss of QMR eigenvalue; realistic example



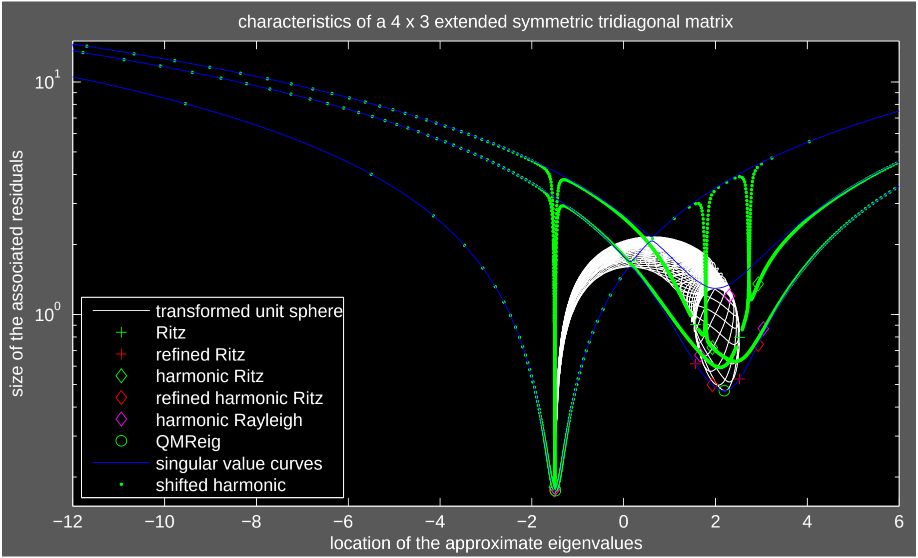
A warning: loss of QMR eigenvalue; realistic example



A warning: loss of QMR eigenvalue; realistic example



A warning: loss of QMR eigenvalue; realistic example



A Jordan block: infinitely many QMR eigenvalues

The next example is

$$\underline{C}_k = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{35}$$

We have Jordan blocks at $\theta = 0$, $\underline{\theta} = \infty$ and $\rho = 0$.

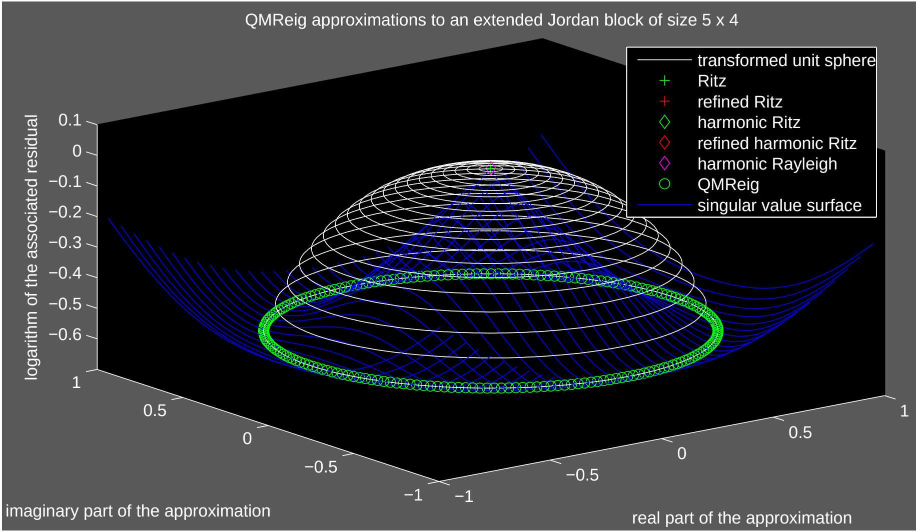
For $k \in \mathbb{N}$ this is an example of an infinite set of QMR eigenvalues,

$$\grave{\theta}_\phi = \cos\left(\frac{\pi}{k+1}\right) e^{i\phi}, \quad \phi \in [0, 2\pi). \tag{36}$$

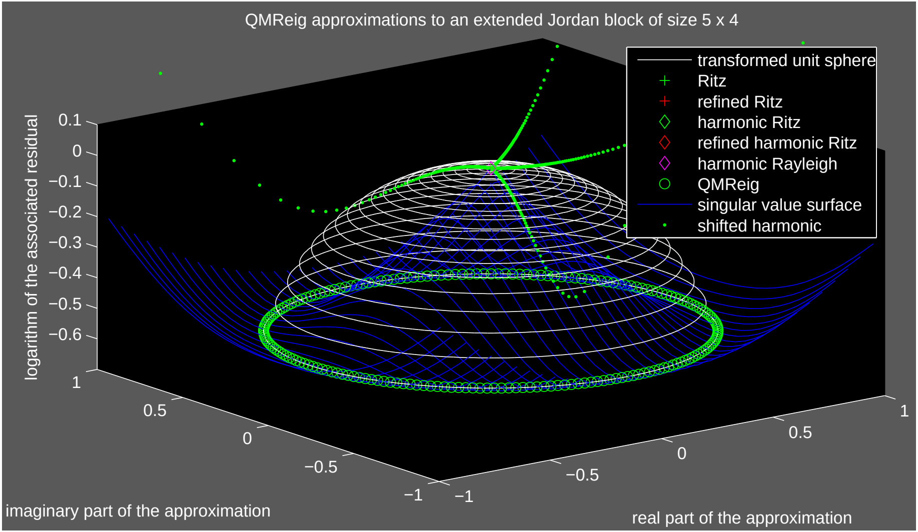
The residual of the corresponding QMR eigenpairs is given by

$$\|(\grave{\theta}_\phi \underline{I}_k - \underline{C}_k)\grave{s}_\phi\| = \sin\left(\frac{\pi}{k+1}\right) \quad \forall\ \phi. \tag{37}$$

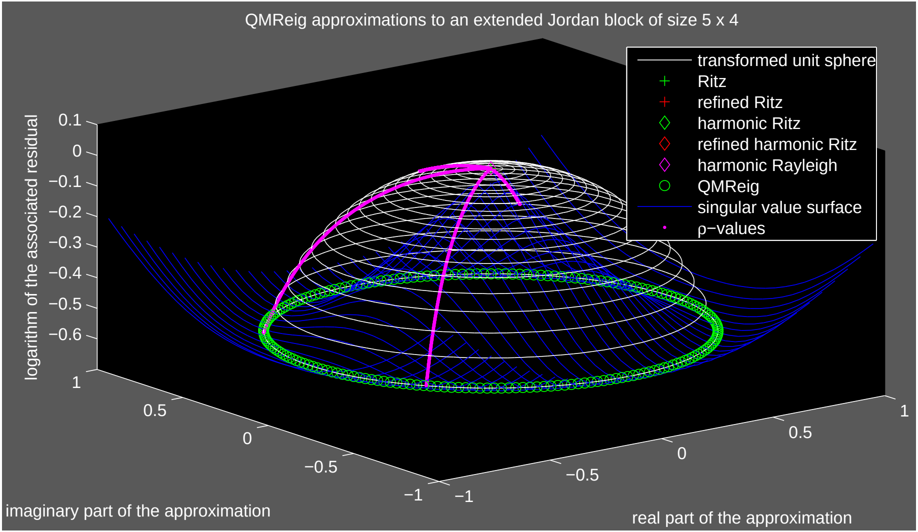
A Jordan block: infinitely many QMR eigenvalues



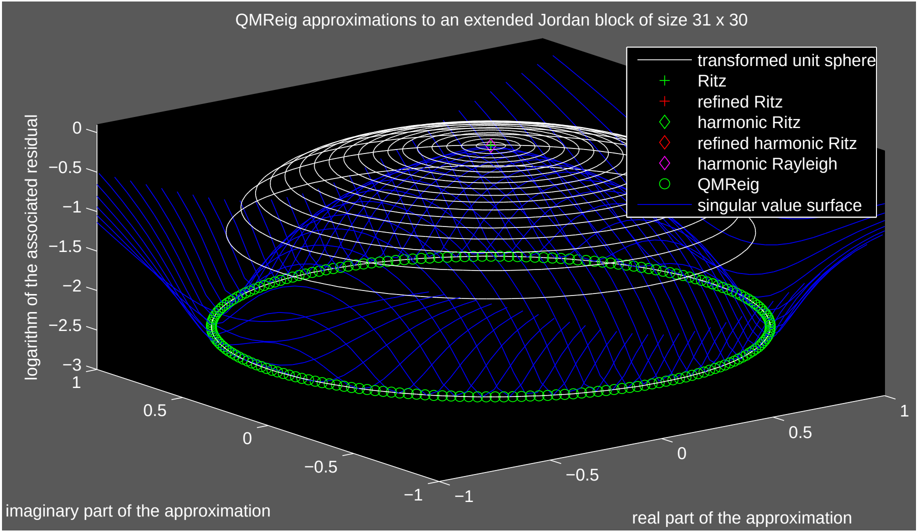
A Jordan block: infinitely many QMR eigenvalues



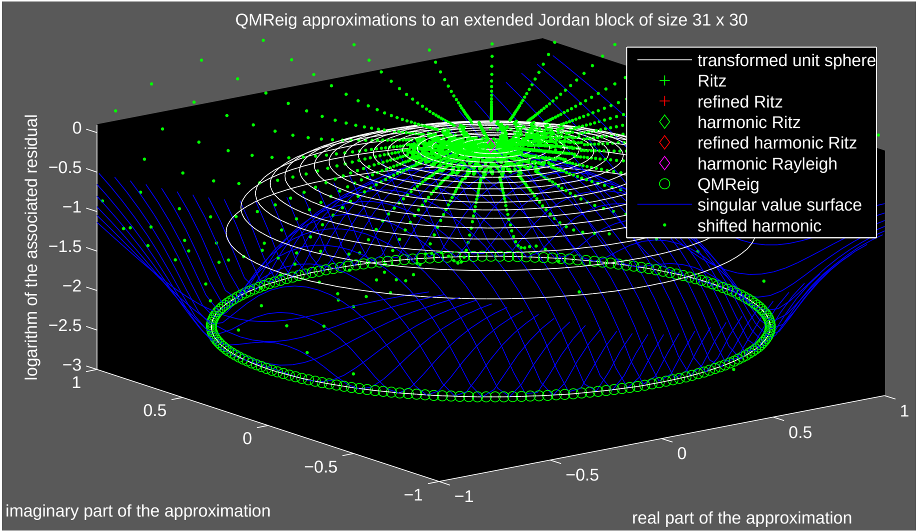
A Jordan block: infinitely many QMR eigenvalues



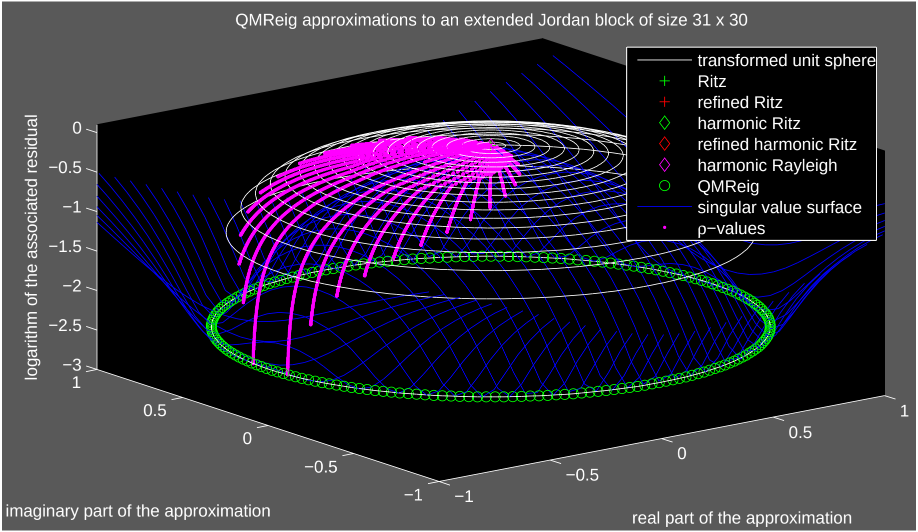
A Jordan block: infinitely many QMR eigenvalues



A Jordan block: infinitely many QMR eigenvalues



A Jordan block: infinitely many QMR eigenvalues



Always remember: It's only locally optimal . . .