Challenges for Eigenvalue Computations in Breakthrough Applications

Agnieszka Międlar

University of Kansas
Department of Mathematics
amiedlar@ku.edu

47th Annual Spring Lecture Series
Department of Mathematical Sciences
Fulbright College of Arts & Sciences
University of Arkansas
May 5, 2022

Motivation

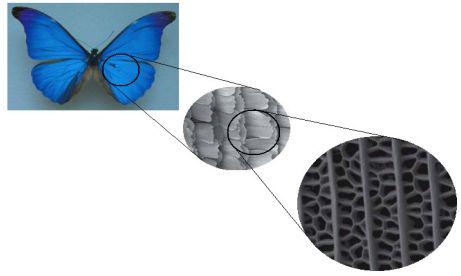
Nonlinear Eigenvalue Problem (NLEVP)

Given $T : \mathcal{D} \to \mathbb{C}^{n\times n}$, find $\lambda \in \mathcal{D} \subseteq \mathbb{C}$ and nonzero vectors $x, y \in \mathbb{C}^n$ such that

$$T(\lambda)x = 0, \quad y^H T(\lambda) = 0.$$

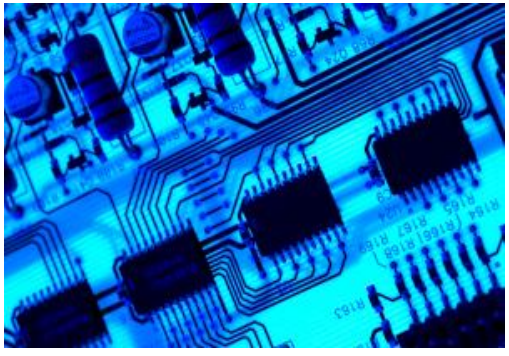
Standard EVP if $T(\lambda) = A - \lambda I, A \in \mathbb{C}^{n\times n}$.

Photonic Crystals



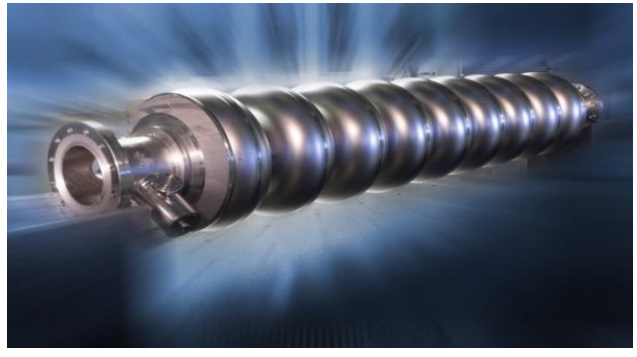
Copyright: Yablonovitch et al.

Nanoelectronics



Copyright: European Commission

Particle Accelerator Cavity



Copyright: Fermilab

We live in an eigen world. B. Broughton.

Linear and Nonlinear Eigenvalue Problems (NLEVPs)

Given a nonempty open set $\mathcal{D} \subseteq \mathbb{C}$ and a matrix-valued function $T : \mathcal{D} \to \mathbb{C}^{n\times n}$, the nonlinear eigenvalue problem consists of finding scalars $\lambda \in \mathcal{D}$ (*eigenvalues*) and nonzero vectors $x, y \in \mathbb{C}^n$ (right and left *eigenvectors*) satisfying

$$T(\lambda)x = 0, \quad y^H T(\lambda) = 0. \tag{1}$$

- In principle, T can be any function of λ, in practice, we consider polynomial, rational or exponential functions.
- In the following, we assume that $T \in H(\mathcal{D}, \mathbb{C}^{n\times n})$ (holomorphic in $\mathcal{D}$).

Remark:

The linear eigenvalue problem find $\lambda \in \mathbb{C}$ and a nonzero vector $x \in \mathbb{C}^n$ such that $Ax = \lambda x$ is a nonlinear eigenvalue problem with $T(\lambda) = A - \lambda I$.

A Classification of Eigenvalue Problems

problem		# **eigenvalues**
standard eigenvalue problem	$T(\lambda) = A - \lambda I$	n
generalized eigenvalue problem	$T(\lambda) = A - \lambda B$	n
quadratic eigenvalue problem	$T(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0$	$2n$
polynomial eigenvalue problem	$T(\lambda) = \sum_{i=0}^{d} \lambda^i A_i$	dn
nonlinear eigenvalue problem	$T(\lambda) = \sum_{i=0}^{d} f_i(\lambda) A_i$	∞

by M. Embree

Some Characteristics of Nonlinear Eigenvalue Problems

- The eigenvalues of (1) are solutions of $\det(T(z)) = 0$.
- T (even if regular, i.e., $\det(T(z)) \not\equiv 0$) may have infinitely many eigenvalues.
- The eigenvectors associated with distinct eigenvalues do not have to be linearly independent.
- The algebraic multiplicity of an isolated eigenvalue, although finite, is not bounded by the problem size n.

For a more detailed description, see:

quadratic EVPs [Tisseur, Meerbergen 2001], [Mehrmann, Voss 2004],

polynomial EVPs [Mackey, Mackey, Tisseur 2015],

nonlinear EVPs [Voss 2014] and [Güttel, Tisseur 2017].

Example (Güttel and Tisseur 2017)

Let $T(z) = \begin{bmatrix} e^{\imath z^2} & 1 \\ 1 & 1 \end{bmatrix}$. Eigenvalues of T are roots of $\det T(z) = e^{\imath z^2} - 1$, i.e.,

$$\lambda_k = \pm\sqrt{2\pi k}, \quad k = 0, \pm 1, \pm 2, \ldots.$$

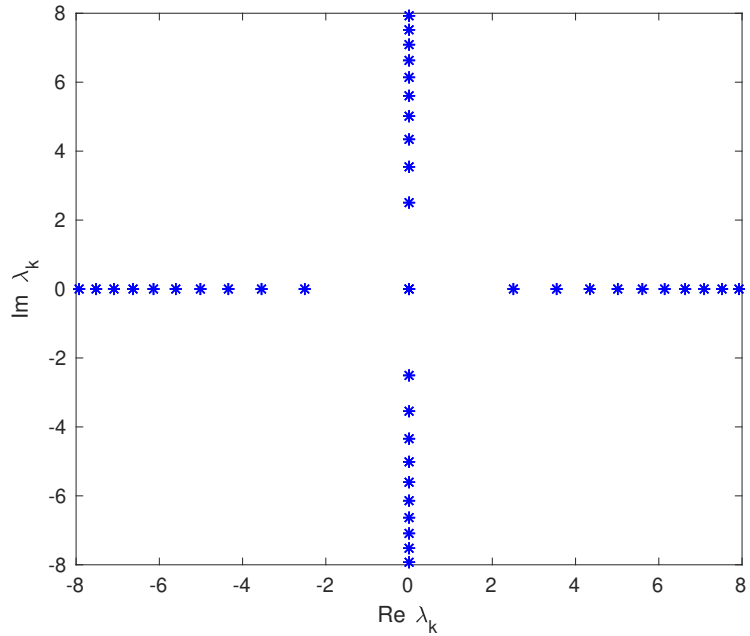
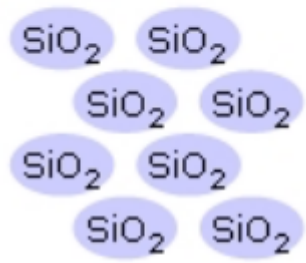
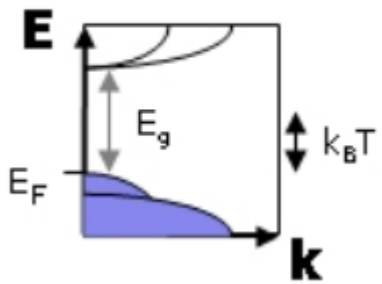


Figure: 42 (of infinitely many) eigenvalues of T. $\lambda_0 = 0$ has algebraic multiplicity 2, and all other eigenvalues are simple (algebraic multiplicity 1). Vector $[1, -1]^T$ is a right and left eigenvector for all eigenvalues λ.

Challenging Eigenvalue Computations





Electronic structure of an insulator

Copyright: Opensource Handbook of Nanoscience and Nanotechnology

Many-body Schrödinger equation

$$\widehat{H}\psi = E\psi,$$

ψ wave function,

$\widehat{H}$ Hamiltonian operator,

E total energy.

Kohn-Sham equation

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + v_{eff}[\rho](r)\right)\phi_i(r) = \epsilon_i\phi_i(r),$$

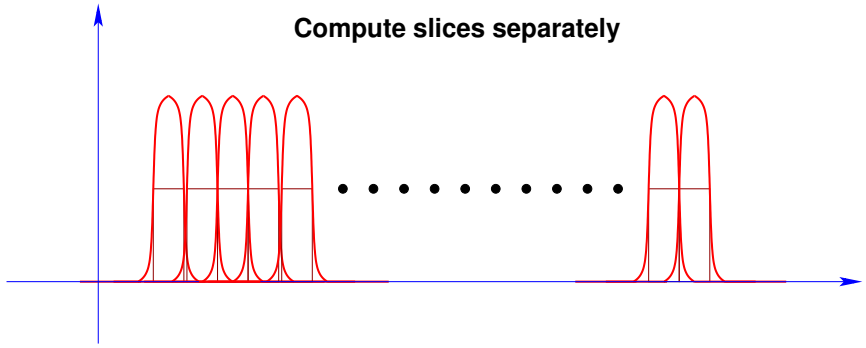
where

$$\rho(r) = \sum_{\text{occupied states}} |\phi_i(r)|^2.$$

Challenge: Hamiltonian of size $\sim$ 1 million, get 10% bands

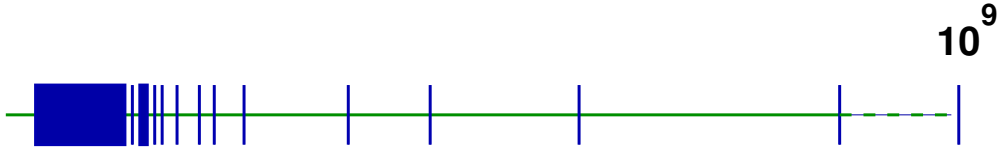
Divide and Conquer Approach and Interior Eigenvalues

Get the spectrum by 'slices'
and use polynomial filtering



However polynomial filters are ...

- **ineffective** if the spectrum looks like this



- **ruled out** when dealing with **generalized eigenvalue problems**.

Remedy:

- **rational filters** of the form $\widetilde{\rho}(z) = \sum_i \frac{\omega_i}{\sigma_i - z}$ are an **alternative**,
- exploit **Cauchy integral** representation of the spectral projector.

Brief Review of Cauchy Integrals and Contour Integration

Theorem (**Cauchy's Theorem (Cauchy-Goursat Theorem)**)

If $f(z)$ is analytic on a simply connected open region $\mathcal{D} \subset \mathbb{C}$ and if $\mathcal{C}$ is a closed contour in $\mathcal{D}$, then

$$\int_{\mathcal{C}} f(z)dz = 0.$$

Theorem (**Cauchy's Integral Formula**)

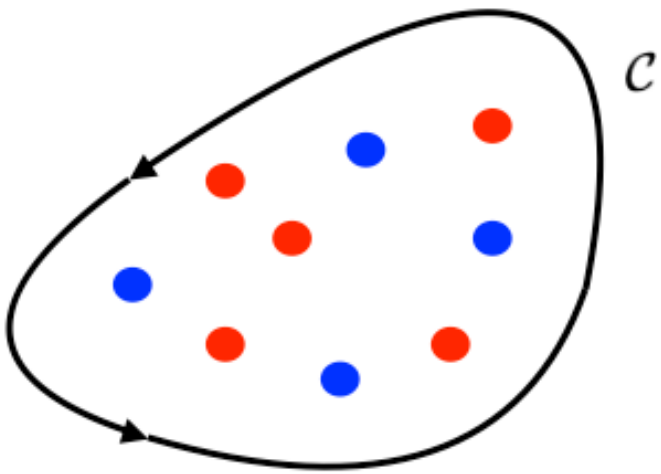
Suppose that $f(z)$ is analytic on a simply connected open region $\mathcal{D} \subset \mathbb{C}$ and $\mathcal{C}$ is a closed contour in $\mathcal{D}$, then

$$\frac{1}{2\pi\imath}\int_{\mathcal{C}} \frac{f(z)}{z-a}dz = \begin{cases} f(a), & \text{if } a \text{ lies inside } \mathcal{C} \\ 0, & \text{if } a \text{ lies outside } \mathcal{C}. \end{cases}.$$

Theorem (**Cauchy's Argument Principle**)

If $f(z)$ is a meromorphic function on and inside a closed contour $\mathcal{C}$, and $f(z)$ has no zeros or poles on $\mathcal{C}$, then

$$\frac{1}{2\pi\imath}\int_{\mathcal{C}}\frac{f'(z)}{f(z)}dz \;=\; (\#\ \textit{zeros of}\ f) - (\#\ \textit{poles of}\ f).$$



Definition ($f(A)$ via Cauchy Integral [Higham 2008])

For $A \in \mathbb{C}^{n \times n}$,

$$f(A) := \frac{1}{2\pi i} \int_{\mathcal{C}} f(z)(zI - A)^{-1} dz,$$

where f is analytic on and inside a closed contour $\mathcal{C}$ that encloses spectrum $\sigma(A)$ of matrix A.

This definition of a matrix function $f(A)$ establishes a **Cauchy Integral Formula for Matrices**.

What is the connection between Cauchy integrals and solving eigenvalue problems?

Counting Eigenvalues Inside $\mathcal{C}$

For $f(z) := \det(zI - A)$ the Cauchy's argument principle allows to determine the number of eigenvalues of A inside $\mathcal{C}$, i.e.,

$$|\sigma_{\mathcal{C}}(A)| := \frac{1}{2\pi\imath} \int_{\mathcal{C}} \frac{f'(z)}{f(z)} dz.$$

Inclusion regions for matrix eigenvalues [Brunaldi, Mellendorf 1994], [Beattie, Ipsen 2003], [Varga 2004], [Hochstenbach, Singer, Zachlin 2008], [Trefethen, Embree 2005].

Counting eigenvalues [Di Napoli, Polizzi, Saad 2014], [Bertrand, Philippe 2001], [Kamgnia, Philippe 2013].

Stochastic estimation methods for eigenvalue density [Futamura, Tadano, Sakurai 2010], [Maeda, Futamura, Sakurai 2011]

Approximating spectral densities [Wang 1994], [Silver, Roeder, Voter, Kress 1996], [Roeder, Silver, Drabold, Dong 1997], [Lin, Yang, Saad 2016], [Xi, Li, Saad 2018].

Suppose, that matrix $A \in \mathbb{C}^{n\times n}$ is diagonalizable, i.e.,

$$A = X\Lambda X^{-1} = \sum_{i=1}^{n} \lambda_i x_i y_i^H.$$

Now, for function f analytic on and inside a contour $\mathcal{C}$ enclosing $\sigma(A)$

$$\begin{aligned}
f(A) &= \frac{1}{2\pi\imath}\int_{\mathcal{C}} f(z)(zI - A)^{-1}dz \\
&= \frac{1}{2\pi\imath}\int_{\mathcal{C}} f(z)X(zI - \Lambda)^{-1}X^{-1}dz \\
&= X\begin{bmatrix} \frac{1}{2\pi\imath}\int_{\mathcal{C}} \frac{f(z)}{z-\lambda_1}dz & & \\ & \ddots & \\ & & \frac{1}{2\pi\imath}\int_{\mathcal{C}} \frac{f(z)}{z-\lambda_n}dz \end{bmatrix} X^{-1} \\
&= X\begin{bmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{bmatrix} X^{-1} = \sum_{i=1}^{n} f(\lambda_i)x_i y_i^*.
\end{aligned}$$

Cauchy Integral Representation of the Spectral Projector

Suppose that $f(z)$ is the indicator function, such that for $\mathcal{D} \subset \sigma(A)$

$$f(z) = \begin{cases} 1, & z \in \mathcal{D} \\ 0, & z \in \sigma(A) \backslash \mathcal{D} \end{cases} \quad \mathcal{D} \subset \sigma(A) \text{ and } \mathcal{C} = \partial\mathcal{D},$$

Then, for and $\mathcal{C} = \partial\mathcal{D}$

$$\begin{aligned} f(A) &= \frac{1}{2\pi\imath} \int_{\mathcal{C}} f(z)(zI - A)^{-1} dz \\ &= \sum_{i=1}^{n} f(\lambda_i) x_i y_i^* = H \\ &= \sum_{\lambda_i \in \mathcal{D}} x_i y_i^H =: P_{\mathcal{C}}. \end{aligned}$$

Therefore, the matrix $P_{\mathcal{C}}$ **is the spectral projector** onto the invariant subspace of matrix A associated with all eigenvalues of matrix A in $\mathcal{D}$, i.e., enclosed by the contour $\mathcal{C}$.

We have showed that the spectral projector P associated with the eigenvalues of matrix A enclosed by the contour $\mathcal{C}$ is given by

$$P_{\mathcal{C}} = \frac{1}{2\pi\imath} \int_{\mathcal{C}} (zI - A)^{-1} dz.$$

W.l.o.g. let us assume we are interested in eigenvalues located in $\mathcal{D} = [-1,\ 1]$. Then $\mathcal{C}$ can be taken (and is typically taken) to be a circle centered at the origin of the complex plane and radius unity.

Hence, using an n_c-point quadrature yields an **approximation of the spectral projector** $P_{\mathcal{C}}$ onto the invariant subspace associated with all eigenvalues of A enclosed by the contour $\mathcal{C}$, i.e.,

$$P_{\mathcal{C}} \approx \sum_{i=1}^{n_c} \omega_i \ (\sigma_i I - A)^{-1} =: \rho(A),$$

with integration **nodes** σ_i and **weights** ω_i, $i = 1, \ldots, n_c$.

How can we use this connection to develop a computationally feasible and effective eigenvalue solvers?

The Nonlinear FEAST Algorithm

with B. Gavin (LinkedIn) and E. Polizzi (UMass Amherst)

FEAST method [Polizzi 2009]

Task: Compute $\boldsymbol{n_e}$ eigenvalues of a linear eigenvalue problems $Ax = \lambda x$ inside contour $\mathcal{C}$.

Solution: Given $m_0 \geq n_e$ dimensional subspace represented by the column vectors of an initial $n \times m_0$ matrix Q **compute**

$$Q = \sum_{i=1}^{n_c} \omega_i (\sigma_i I - A)^{-1} Q \approx \frac{1}{2\pi\imath} \int_{\mathcal{C}} (zI - A)^{-1} \, Q \; dz = P_{\mathcal{C}} Q$$

and use it in the **Rayleigh-Ritz procedure**, i.e., solve a projected eigenvalue problem

$$Q^H A Q u = \lambda u.$$

Since, in general, the range of Q may initially not be a good approximation of the eigenspace of matrix A corresponding to the eigenvalues inside the contour $\mathcal{C}$, we need to **iteratively refine the subspace** Q.

FEAST Algorithm

Input: *Subspace Q (random) of dimension m_0 with $n \gg m_0 \geq m$, contour $\mathcal{C}$*

Output: *m eigenvalues of matrix A inside $\mathcal{C}$*

1. *Repeat until convergence*
2. *Compute $Q = P_{\mathcal{C}} Q$*
3. *Orthogonalize Q*
4. *Compute $A_Q = Q^H A Q$*
5. *Solve $A_Q U = U \Lambda$*
6. *Compute $Q = QU$*
7. *Check convergence of m wanted eigenvalues inside $\mathcal{C}$*

end

FEAST Characteristics

- Combines subspace iteration method with efficient **contour integration** technique used to **approximate a spectral projector** $P_{\mathcal{C}}$.
- **Accelerated subspace iteration** algorithm in conjunction with the Rayleigh-Ritz procedure, non-adaptive accelerator which is a rational matrix function that approximates the spectral projector, i.e.,

$$\rho(A) = \sum_{i=1}^{n_c} \omega_i (z_i I - A)^{-1}.$$

- Projection method onto **subspace of a fixed dimension**, unlike SS or Jacobi-Davidson method.
- Similar to **inverse subspace iteration**, but uses **multiple shifts** to accelerate convergence, i.e., executing inverse subspace iteration with matrix $(\sigma I - A)^{-1}$, is equivalent to using FEAST algorithm with a single quadrature node σ.

FEAST Characteristics

- Main building block is numerical integration resulting in solving independent linear systems, each for multiple right-hand sides, i.e., in practice the **subspace Q is computed** as

$$Q = \sum_{i=1}^{n_c} \omega_i Q_i,$$

  using the solutions Q_i of the n_c independent linear systems of the form

$$(\sigma_i - A)Q_i = V.$$

- Optimized sparse direct solvers, e.g., PARDISO, as well as iterative methods, can be used to solve these linear systems.
- Highly parallelizable.

Contour Integral Eigenvalue Solvers

- **Linear Eigenvalue Problems**

  SS method [Sakurai, Sugiura, 2003], [Sakurai, Tadano 2007].
  FEAST [Polizzi 2009], [Tang, Polizzi 2014], [Kestyn, Polizzi, Tang 2016].
  DD accelerated eigensolver [Kalantzis, Kestyn, Polizzi, Saad 2018].
  other [Xi, Saad 2016], [Kalantzis, Xi, Saad 2018], [Li, Xi, Erlandson, Saad 2019].

- **Nonlinear Eigenvalue Problems**

  Beyn's method [Beyn, 2010], [Beyn, Effenberger, Kressner 2011].
  SS-type methods [Asakura, Sakurai, Tadano, Ikegami, Kimura 2009], [Asakura, Sakurai, Tadano, Ikegami, Kimura 2010], [Yokota, Sakurai 2013].
  nonlinear FEAST [Gavin, M., Polizzi 2018].
  other [Embre, Gugercin, Brennan 2019], [El-Guide, M., Saad 2020], [Saad, El-Guide, M. 2020].

Newton's Method for Nonlinear Eigenvalue Problems

Note that solving a nonlinear eigenvalue problem (1) with normalization $c^H x = 1$ is equivalent to solving a nonlinear problem

$$F\left(\begin{bmatrix} x \\ \lambda \end{bmatrix}\right) = 0, \quad \text{with} \quad F\left(\begin{bmatrix} x \\ \lambda \end{bmatrix}\right) := \begin{bmatrix} T(\lambda)x \\ c^H x - 1 \end{bmatrix}.$$

Aplying Newton's method to F yields the following iteration

$$T(\lambda_k)x_{k+1} = -(\lambda_{k+1} - \lambda_k)T'(\lambda_k)x_k, \quad c^H x_{k+1} = 1,$$

which is equivalent to the nonlinear inverse iteration method.

Nonlinear Inverse Iteration [Unger 1950, Ruhe 1973]

$$\begin{aligned} \text{Solve} \quad T(\lambda_k)v_k &= T'(\lambda_k)x_k \quad \text{for} \quad v_k \\ \text{Set} \quad \lambda_{k+1} &= \lambda_k - \frac{c^H x_k}{c^H v_k} \\ \text{Compute} \quad x_{k+1} &= \frac{v_k}{\|v_k\|} \end{aligned}$$

Residual Inverse Iteration [Neumeier 1985]

Step 1: Given $w_k, x_k \in \mathbb{C}^n$ compute λ_{k+1} as a solution of

$$w_k^H T(\lambda) x_k = 0. \tag{2}$$

Step 2: Given a shift $\sigma \in \mathbb{C}$ replace $T'(\lambda_k)$ by a finite difference and $T(\lambda_k)^{-1}$ by $T(\sigma)^{-1}$, compute

$$x_{k+1} = x_k - T^{-1}(\sigma)T(\lambda_{k+1})x_k,$$

and normalize x_{k+1}.

Theorem (**Keldysh 1951**)

Suppose $T(z)$ has m semi-simple eigenvalues $\lambda_1, \ldots, \lambda_m$ (counting multiplicities) in the region $\mathcal{D} \subset \mathbb{C}$. Then

$$T(z)^{-1} = V(zI - \Lambda)^{-1} U^H + R(z), \tag{3}$$

- $V = [v_1, \ldots, v_m]$, $U = [u_1, \ldots, u_m]$,
- $\Lambda = diag(\lambda_1, \ldots, \lambda_m)$,
- $u_i^H T(\lambda_i) v_i = 1$,
- $R(z)$ *is analytic in* $\mathcal{D}$.

By Cauchy's integral formula

$$\frac{1}{2\pi\imath} \int_{\mathcal{C}} f(z) T(z)^{-1} \, dz = V f(\Lambda) U^H.$$

Contour Integration for Nonlinear Eigenvalue Problems

Nonlinear variants of the Sakurai-Sugiura (SS) method

- based on block Hankel matrices (SS-H method),
- for the polynomial eigenvalue problems [Asakura et al. 2009] and eigenvalue problems of analytic matrix functions $T(\lambda)$ [Asakura et al. 2009],
- cost efficient and highly scalable,
- the accuracy of obtained solutions is relatively low.

Beyn's integral method [Beyn 2010]

- uses the zeroth and first-order moments matrices to reduce an NLEVP with $m \ll n$ eigenvalues inside $\mathcal{C}$ to a linear eigenvalue problem of dimension m,
- highly accurate,
- based on the idea of Keldysh to probe a Jordan decomposition (conceptually very simple but known to be highly sensitive to perturbations).

Projection SS method [Yokota, Sakurai 2013]

- addresses the problem of low accuracy in the nonlinear SS-H method,
- uses the same contour integrals as the SS-H method, however, the approximate eigenpairs are extracted from the underlying subspaces via Rayleigh-Ritz procedure,
- gives better accuracy than the methods of [Asakura et al. 2009] and [Beyn 2010].

All presented methods use the moments of the Cauchy integral of $T^{-1}(\lambda)$, i.e.,

$$\frac{1}{2\pi i} \oint_{\mathcal{C}} z^k T^{-1}(z)\, dz, \quad k \geq 0.$$

- **Beyn's method** uses the $k = 0$ and $k = 1$ moments,
- **SS-type methods** use as many moments as necessary for convergence,
- **standard FEAST** uses only the $k = 0$ moment, which does not work in the case of the nonlinear $T(\lambda)$.

The Nonlinear FEAST (NLFEAST) Algorithm

Let us assume that the **nonlinear FEAST algorithm** should be a **generalization of nonlinear inverse iteration**

$$x_{k+1} = T(\lambda_k)^{-1} T'(\lambda_k) x_k, \tag{4}$$

that uses multiple shifts at once.

Hence, we would expect a nonlinear FEAST contour integral (for a single vector) to be of the form

$$q = \frac{1}{2\pi\imath} \oint_{\mathcal{C}} T(z)^{-1} T'(\lambda_k) x_k \; dz. \tag{5}$$

In practice, however, the update procedure in (4) does not converge to correct eigenpairs. Using a constant shift $z \neq \lambda_k$ in (4) results in the nonlinear inverse iterations converging to the eigenvalues of matrix $T(z)$.

The **residual inverse iteration** solves this problem by approximating the derivative $T'(\lambda_k)$ which, when inserted into (4), produces the following update rule

$$x_{k+1} = \frac{I - T(z)^{-1}T(\lambda_k)}{z - \lambda_k} x_k,$$

which combined with (5) yields

$$q = \frac{1}{2\pi\imath} \oint_{\mathcal{C}} \frac{I - T^{-1}(z)T(\lambda_k)}{z - \lambda_k} x_k \; dz. \tag{6}$$

In the case of linear eigenvalue problem, i.e., $T(\lambda) = A - \lambda I$, the approximation of the derivative is exact, and (6) is equal to the contour integral for the linear FEAST algorithm.

The nonlinear FEAST is a generalization of the residual inverse iteration using multiple shifts.

From an initial set of approximate eigenvectors Q generate a refined subspace Q, i.e.,

$$Q = \frac{1}{2\pi\imath} \oint_{\mathcal{C}} \left(Q - T^{-1}(z) T(Q, \widetilde{\Lambda}) \right) (zI - \widetilde{\Lambda})^{-1} \, dz,$$

and solve a projected nonlinear eigenvalue problem $Q^H T(\lambda) Q u = 0$.

$T(Q, \widetilde{\Lambda})$ is the block residual, i.e., for the polynomial eigenvalue problem of degree d the block form of the residual function yields

$$T(Q, \widetilde{\Lambda}) = \sum_{i=0}^{d} A_i Q \widetilde{\Lambda}^i.$$

For a general nonlinear eigenvalue problem, the i^{th} column vector of the block residual $T(Q, \widetilde{\Lambda})$ is the residual vector $T(\widetilde{\lambda}_i) q_i$, where q_i is the i^{th} column vector of Q and $\widetilde{\lambda}_i$ is the corresponding Ritz value.

Nonlinear FEAST Characteristics

- Combines **residual inverse iteration methods** with efficient **contour integration** technique to construct a good projection space.
- **Accelerated projection method** in conjunction with the Rayleigh-Ritz procedure, non-adaptive accelerator which is a rational matrix function that generates a good projection space of a fixed dimension.
- Similar to **residual inverse iteration**, but uses **multiple shifts** to accelerate convergence, i.e., executing residual inverse iteration on (1) is equivalent to using nonlinear FEAST algorithm with a single quadrature node σ.
- Numerical integration results in **solving independent linear systems**, each for multiple right-hand sides, hence optimized sparse direct solvers as well as iterative methods, can be used.
- **Small nonlinear eigenvalue problems** can be solved efficiently.
- **Highly parallelizable**.

Convergence Analysis of the NLFEAST Algorithm

Given a **Rayleigh functional** $p(v)$ NLFEAST (with exact contour integration) is a fixed point iteration for

$$F_{FEAST}(v) = \beta(v)\frac{1}{2\pi\imath}\oint_{\mathcal{C}} \frac{1}{z - p(v)}(I - T(z)^{-1}T(p(v)))v \, dz.$$

Theorem (Fixed Point Equivalence for the NLFEAST)

Let $v \in \mathbb{C}^n$ and $c^H v \neq 0$ such that (2) has a unique solution $p(v)$ and suppose $w(v)$ satisfies $w(v)^H T(z)v \neq 0$. Then the following statements are equivalent: (i) $F_{FEAST}(v) = v$ and (ii) $T(p(v))v = 0$.

Theorem (Characterization of the Eigenvector Error)

Let (λ, x) be an eigenpair of problem (1) and w a function such that $w(x)^H T'(\lambda)x \neq 0$ and $w(x)T(\sigma)x \neq 0$. Then, there exists a neighborhood $\mathcal{N}(x)$ of x, such that for any pair $x_k, x_{k+1} \in \mathcal{N}(x)$ corresponding to one step of nonlinear FEAST iteration $x_{k+1} = F_{FEAST}(x_k)$, the eigenvector error satisfies

$$\|x_{k+1} - x\| = \mathcal{O}(\|x_k - x\|^2).$$

Numerical Examples

Example 1: Scattering resonances in 1D

As an example of open boundary quantum transmission problem, we consider the problem of scattering resonances in 1D with the following compactly supported finite square model potential

$$V(r) = \begin{cases} -V_0, & r \in [-L, L] \\ 0, & \text{otherwise} \end{cases},$$

with $V_0 > 0$ and width $2L = \pi\sqrt{2}$ [Cancès, Nectoux 2017].

We are interested to determine the **Siegert states** [Siegert 1939] $u \in H^1(-L, L)$ and the associated resonances $k \in \mathbb{C}$ such that for all $v \in H^1(-L, L)$

$$\int_{-L}^{L} u'v' + Vu\bar{v}'dx = \imath k(u(L)\bar{v}(L) + u(-L)\bar{v}(-L)) + k^2 \int_{-L}^{L} u\bar{v}dx.$$

In the case of $V : \mathbb{R} \to \mathbb{C}$ being bounded with compact support in $[-L, L]$, the set of discrete solutions $(u_i, k_i) \in H^1(-R, R) \times \mathbb{C}$ can be approximated using the finite element space $V_h \subset H^1(-L, L)$, i.e.,

$$T(k_h)u_h = \left(k_h^2 A_h + \imath k_h B_h - C_h\right) u_h = 0,$$

where

$$A_h = \frac{h}{6}\begin{bmatrix} 2 & 1 & 0 & 0 & \dots & 0 \\ 1 & 4 & 1 & 0 & \dots & 0 \\ 0 & 1 & 4 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & 1 & 4 & 1 \\ 0 & \dots & 0 & 0 & 1 & 2 \end{bmatrix}, B_h = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix} \in \mathbb{R}^{n+2 \times n+2},$$

and

$$C_h = \frac{1}{h}\begin{bmatrix} 1 & -1 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 \\ 0 & -1 & 2 & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & -1 & 2 & -1 \\ 0 & \dots & 0 & 0 & -1 & 1 \end{bmatrix} + V_0 A_h \in \mathbb{R}^{n+2 \times n+2}.$$

The associated linear eigenvalue problem has a form

$$\begin{bmatrix} -\imath B_h & C_h \\ I & O \end{bmatrix} \begin{bmatrix} k_h u_h \\ u_h \end{bmatrix} = k_h \begin{bmatrix} A_h & O \\ O & I \end{bmatrix} \begin{bmatrix} k_h u_h \\ u_h \end{bmatrix}. \tag{7}$$

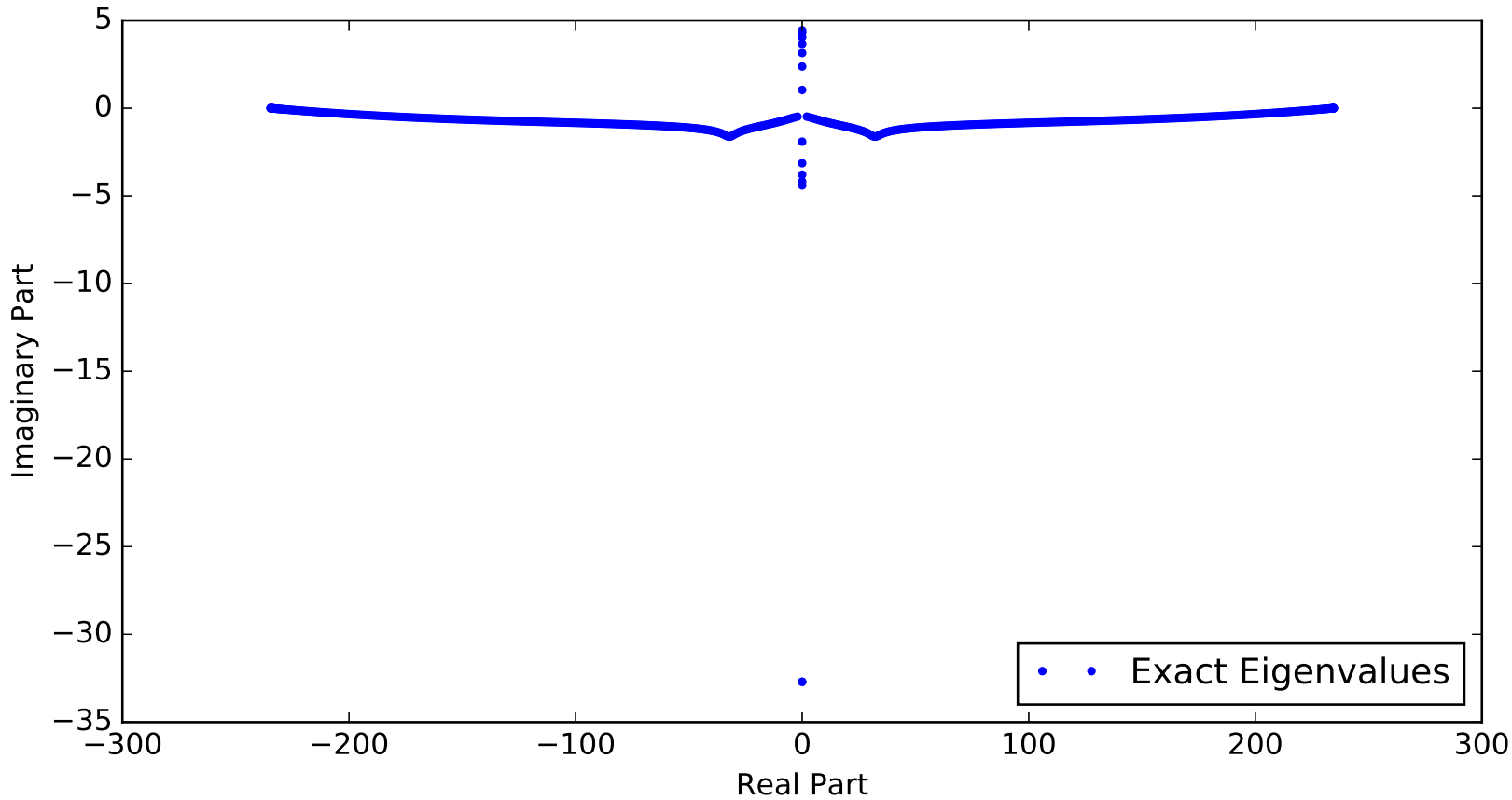


Figure: All $n = 302$ scattering resonances for potential $V_0 = 10$ obtained via linearization (7).

There are twenty-two 22 complex scattering resonances lying inside the circle centered at 0.0 with radius $r = 15.5$. The nonlinear FEAST computes 10^{-10} accurate approximations of scattering resonances in 4 iterations using $n_c = 16$ integration nodes and the subspace of size $m_0 = 30$.

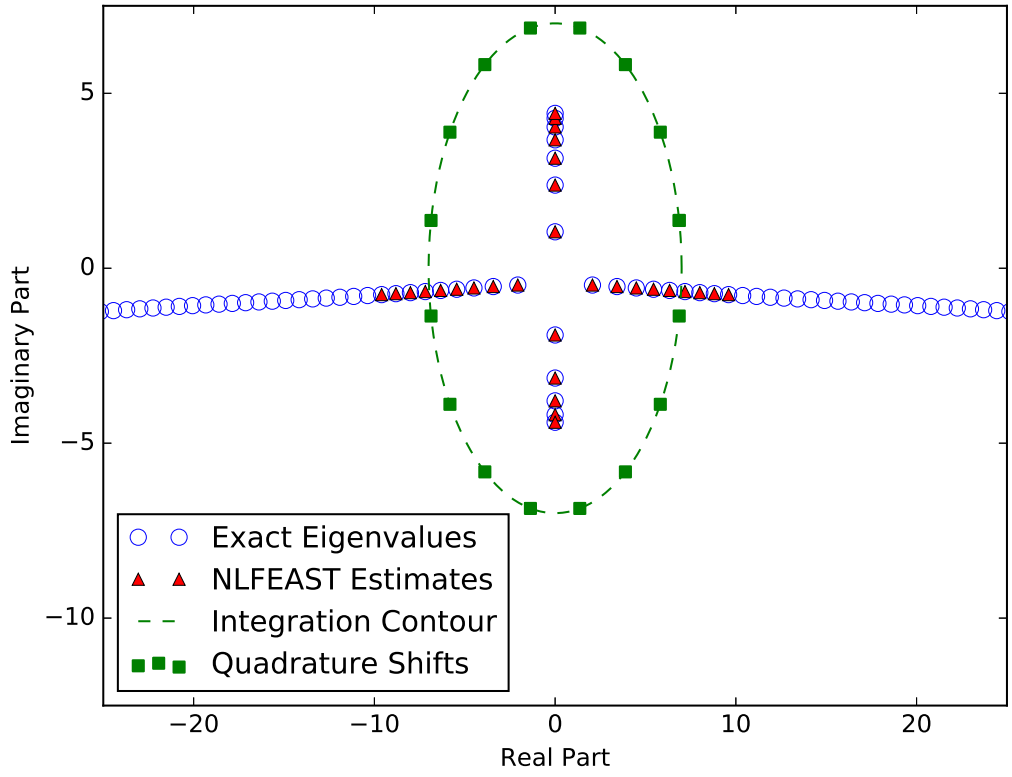


Figure: The nonlinear FEAST approximations of $m_0 = 30$ eigenvalues for the scattering resonance problem for potential $V_0 = 10$. NLFEAST captures all of the 22 eigenvalues inside of the integration contour, plus 8 eigenvalues that are closest to the integration contour while still being outside of it.

Example 2: Hadeler Problem

As an example of a general nonlinear eigenvalue problem, let us consider the **Hadeler problem** [Hadeler 1967], [Betcke, Higham, Mehrmann, Schröder, Tisseur 2013]:

$$T(\lambda) = (e^{\lambda} - 1)B_1 + \lambda^2 B_2 - B_0,$$

with the matrix elements of B_0, B_1, and B_2 being

$$B_0 = b_0 I, \quad B_1 = (b_{jk}^{(1)}), \quad B_2 = (b_{jk}^{(2)})$$

$$b_{jk}^{(1)} = (n + 1 - \max(j, k))jk, \quad b_{jk}^{(2)} = n\delta_{jk} + 1/(j + k),$$

where n is the dimension of the eigenvalue problem and b_0 is a parameter.

Let us consider $n = 200$ and $b_0 = 100$ following [Ruhe 1973].

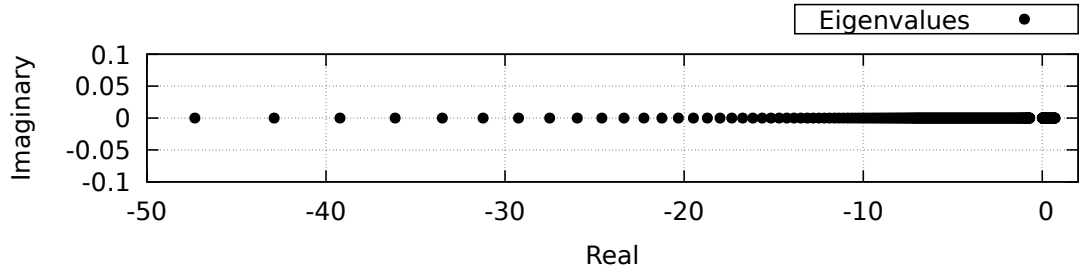
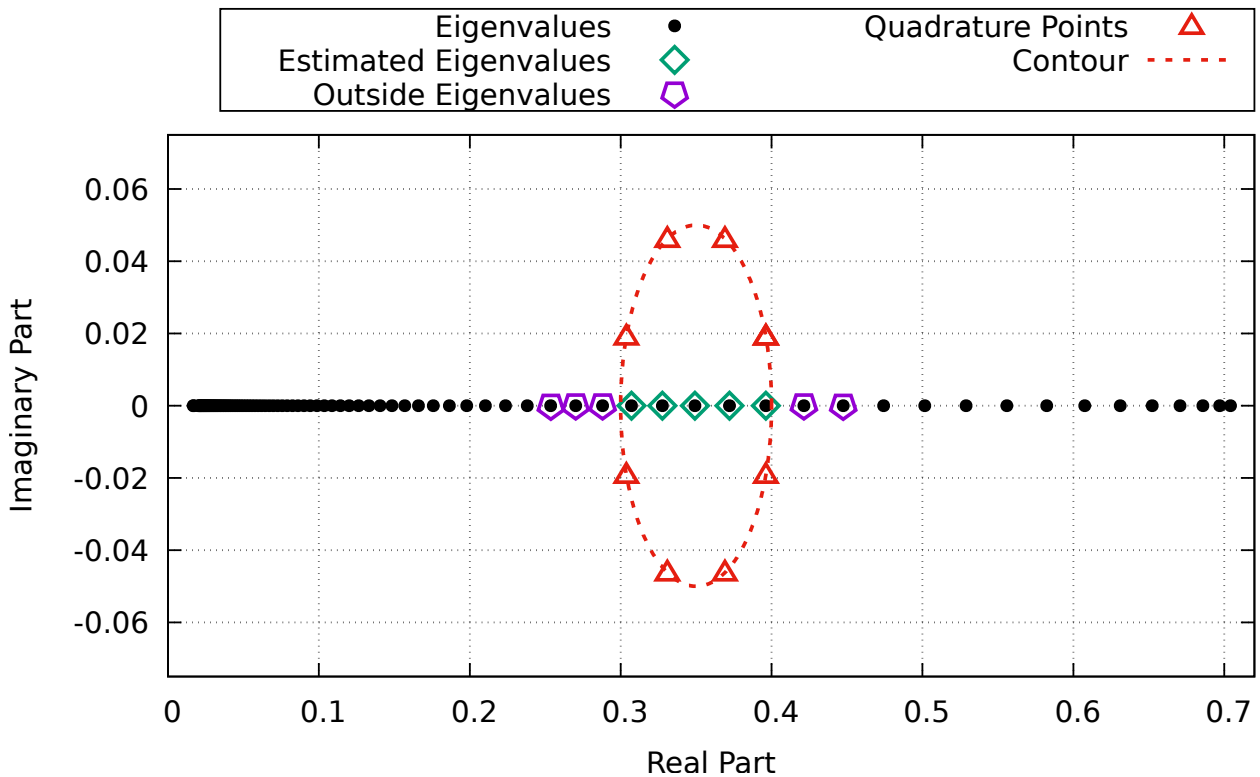
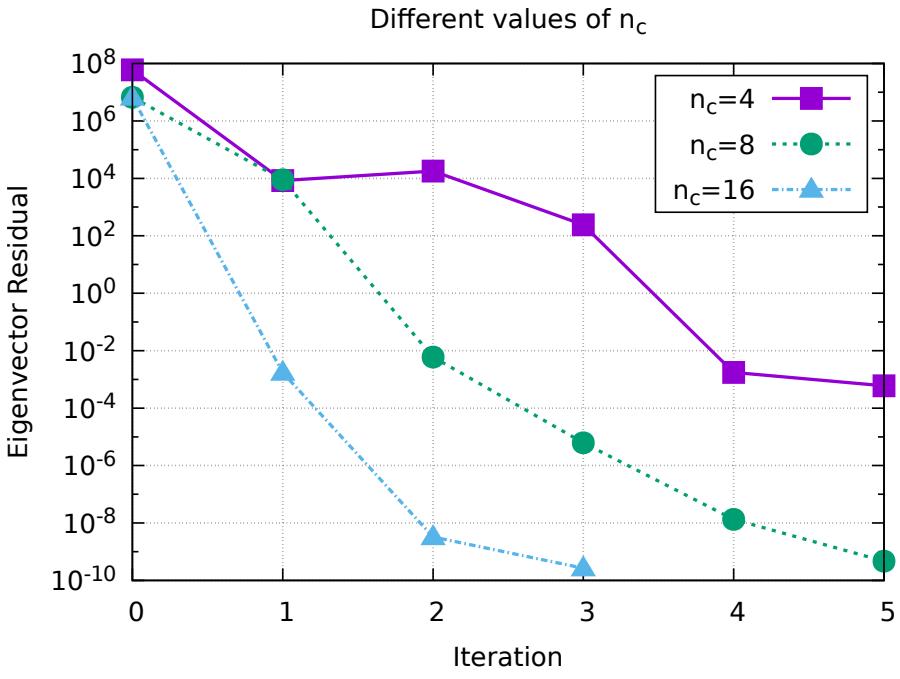


Figure: *Eigenvalues of the the Hadeler problem. All of the eigenvalues are real; there are n eigenvalues less than zero, and n eigenvalues greater than zero.*





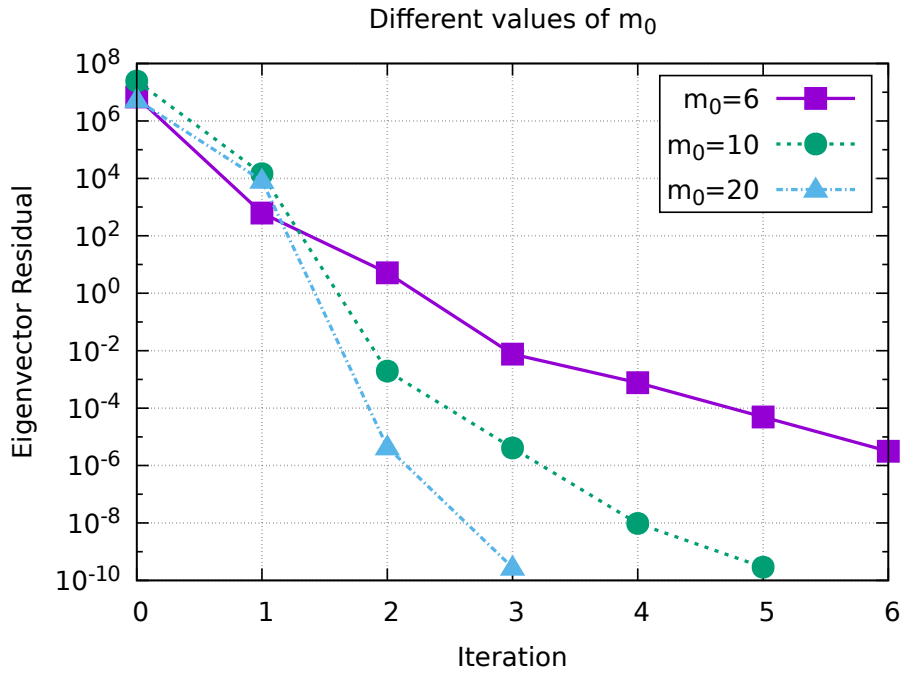


Figure: *Eigenvector residual versus iteration number for solving the Hadeler problem with NLFEAST. Five eigenvalues in the middle of the left half of the spectrum are calculated. The left plot shows results for several values of n_c, with $m_0 = 10$. The right plot shows results for several values of m_0, with $n_c = 8$.*

Rational Approximation Method for NLEVPs

with M. El-Guide(Mohammed VI Polytechnic University) and Y. Saad (University of Minnesota)

Rational Approximation of Matrix Value Function

Let us now consider a matrix-valued function $T(\lambda)$ of the form

$$T(\lambda) = f_0(\lambda)A_0 + f_1(\lambda)A_1 + f_2(\lambda)A_2 + \ldots + f_p(\lambda)A_p, \tag{8}$$

with holomorphic functions $f_0, \ldots, f_p : \Omega \to \mathbb{C}$ and constant coefficient matrices $A_0, \ldots, A_p \in \mathbb{C}^{n\times n}$. W.l.o.g. we assume $f_0(\lambda) = 1$ and $f_1(\lambda) = \lambda$, such that (8) has the form

$$T(\lambda) = -B_0 + \lambda A_0 + f_1(\lambda)A_1 + \ldots + f_p(\lambda)A_p. \tag{9}$$

Following the Cauchy integral representation of each function $f_j(\lambda)$ inside a region enclosed by a contour $\mathcal{C}$, we assume that **each $f_j(\lambda)$ in (9) is well approximated by a rational function**, i.e.,

$$f_j(\lambda) \approx \sum_{i=1}^{n_c} \frac{\omega_{ij}}{\lambda - \sigma_i}, \qquad \lambda \in \Omega,$$

where **σ_i's are quadrature nodes** located on the contour $\mathcal{C}$ and the **ω_{ij}'s** the corresponding quadrature **weights** of the n_C-point quadrature rule.

'Linearization' of the surrogate problem

Thus, with defining $B_i := \sum_{j=0}^{p} \omega_{ij} A_j$ and $v_i := \frac{x}{\sigma_i - \lambda}$, $i = 1, \ldots, n_c$ we obtain the following **surrogate eigenvalue problem**

$$\widetilde{T}(\lambda)x = (-B_0 + \lambda A_0)x - \sum_{i=1}^{n_c} B_i v_i, \tag{10}$$

and its **'linearization'**

$$\begin{bmatrix} (\lambda - \sigma_1)I & & & & I \\ & (\lambda - \sigma_2)I & & & I \\ & & \ddots & & \vdots \\ & & & (\lambda - \sigma_{n_c})I & I \\ -B_1 & -B_2 & \ldots & -B_{n_c} & \lambda A_0 - B_0 \end{bmatrix} \underbrace{\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n_c} \\ x \end{bmatrix}}_{w} = 0. \tag{11}$$

Eigenpairs of (10) provide good approximations to the eigenpairs of (1) with eigenvalues inside the contour $\mathcal{C}$.

Consequently, solutions of the surrogate eigenvalue problem $\widetilde{T}(\lambda)x = 0$ can be obtained by solving the generalized linear eigenvalue problem

$$\mathcal{A}w = \lambda \mathcal{M}w, \tag{12}$$

with

$$\mathcal{M} = \begin{bmatrix} I & & & \\ & I & & \\ & & \ddots & \\ & & & A_0 \end{bmatrix}, \quad \mathcal{A} = \begin{bmatrix} \sigma_1 I & & & -I \\ & \sigma_2 I & & -I \\ & & \ddots & \vdots \\ & & \sigma_{n_c} I & -I \\ B_1 & B_2 & \dots & B_{n_c} \quad B_0 \end{bmatrix}. \tag{13}$$

Note that matrices $\mathcal{M}$ and $\mathcal{A}$ are each of dimension $(n_c n + n) \times (n_c n + n)$, but they will never be stored explicitly.

Part I: Shift-and-invert on the full system

Using shifted inverse iteration, requires solving linear system with a shifted matrix $(\mathcal{A} - \sigma\mathcal{M})$ at each step, i.e., solving linear systems of the form

$$\begin{bmatrix} D & F \\ B^T & B_0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}. \tag{14}$$

Here, we can exploit the sparsity of D and F and easily form the following block LU factorization of matrix $\mathcal{A}$

$$L = \begin{bmatrix} I & 0 \\ BD^{-1} & I \end{bmatrix}, \qquad U = \begin{bmatrix} D & F \\ 0 & S \end{bmatrix},$$

where $S = C - BD^{-1}F$ is the **Schur complement** of the block B_0.
Solving (14) requires:

(1) solving the system $Sy = b - B^T D^{-1} a$,

(2) substituting y in the first part of (14) to obtain x from $Dx = a - Fy$.

Let $w^k = \begin{bmatrix} v^k \ ; \ x^k \end{bmatrix}$ be a vector obtained at the k-th step of the inverse iteration procedure. Each step of the shifted inverse power method (inverse iteration) requires solving the linear system

$$(\mathcal{A} - \sigma\mathcal{M})w^{k+1} = \mathcal{M}w^k \quad \text{or} \quad (\mathcal{A} - \sigma\mathcal{M}) \begin{bmatrix} v^{k+1} \\ x^{k+1} \end{bmatrix} = \begin{bmatrix} v^k \\ x^k \end{bmatrix}. \qquad (15)$$

System (15) is of the same form as that in (14) and it can be solved the same way, resulting in the following steps

$$\begin{aligned} x^{k+1} &= S(\sigma)^{-1}\left(A_0 x^k - B^T(D - \sigma I)^{-1} v^k\right), \\ v^{k+1} &= (D - \sigma I)^{-1}(v^k - F x^{k+1}). \end{aligned}$$

Part I: Projection method on the reduced system

We now introduce a projection method that works in $\mathbb{C}^n$, i.e., it **requires only vectors of the size of the original problem** (1).

Let $U = [u_1, u_2, \ldots, u_\nu]$ be an orthonormal basis of subspace which contains good approximations to eigenvectors of problem (1). Hence, an approximate eigenvector $\widetilde{x}$ can be expressed in this basis as $\widetilde{x} = Us$, with $s \in \mathbb{C}^\nu$. Then, a **Rayleigh-Ritz procedure applied to** (1) yields a projected problem

$$U^H \Big(-B_0 + \lambda A_0 + \sum_{i=1}^{n_c} \frac{B_i}{\lambda - \sigma_i} \Big) Us = 0. \tag{16}$$

This leads to a nonlinear eigenvalue problem in $\mathbb{C}^\nu$, namely:

$$\Big(-\widehat{B}_0 + \lambda \widehat{A}_0 + \sum_{i=1}^{n_c} \frac{\widehat{B}_i}{\lambda - \sigma_i} \Big) s = 0, \tag{17}$$

in which $\widehat{A}_0 = U^H A_0 U$, and $\widehat{B}_i = U^H B_i U$, for $i = 0, 1, \ldots, n_c$.

How to obtain a good subspace to perform the projection method?

Suppose we wish to perform a single step of the subspace iteration algorithm applied with shift-and-invert, we proceed as follows:

- at a given iteration step we have a certain basis $W = [w_1, w_2, \ldots, w_\nu]$ of the current subspace,
- we apply q **steps of shifted inverse iteration to each column** w_j, and denote the k-th iterate by $w_j^k = [v_j^k;\ x_j^k]$,
- after a column is processed by these q steps we **discard its top part and extract the U-part that will be used in the projection process**. This is done one column at a time, therefore **we only have to keep one vector of length** $(n_c + 1)n$,
- doing this for each column of U in succession constitutes one step of **reduced subspace iteration**.

Numerical Experiments

Example 1: Hadeler problem

We consider the *Hadeler problem* [Hadeler 1967], [Ruhe 1973], [Betcke, Higham, Mehrmann, Schröder, Tisseur 2013], [Higham, Porzio, Tisseur 2019]

$$T(\lambda) = (e^{\lambda} - 1)B_1 + \lambda^2 B_2 - B_0, \tag{18}$$

with the coefficient matrices

$$B_0 = b_0 I, \quad B_1 = (b_{jk}^{(1)}), \quad B_2 = (b_{jk}^{(2)}),$$

$$b_{jk}^{(1)} = (n + 1 - \max(j,k))jk, \quad b_{jk}^{(2)} = n\delta_{jk} + 1/(j+k),$$

of size $n = 200$ and $b_0 = 100$.

- The eigenvalues of (18) are real, n negative and n positive.
- The eigenvalues become better spaced as we move away from the origin and the smallest one is close to -48.
- Contour $\mathcal{C}$ is a circle centered at $c = -30$ with radius $r = 11.5$.

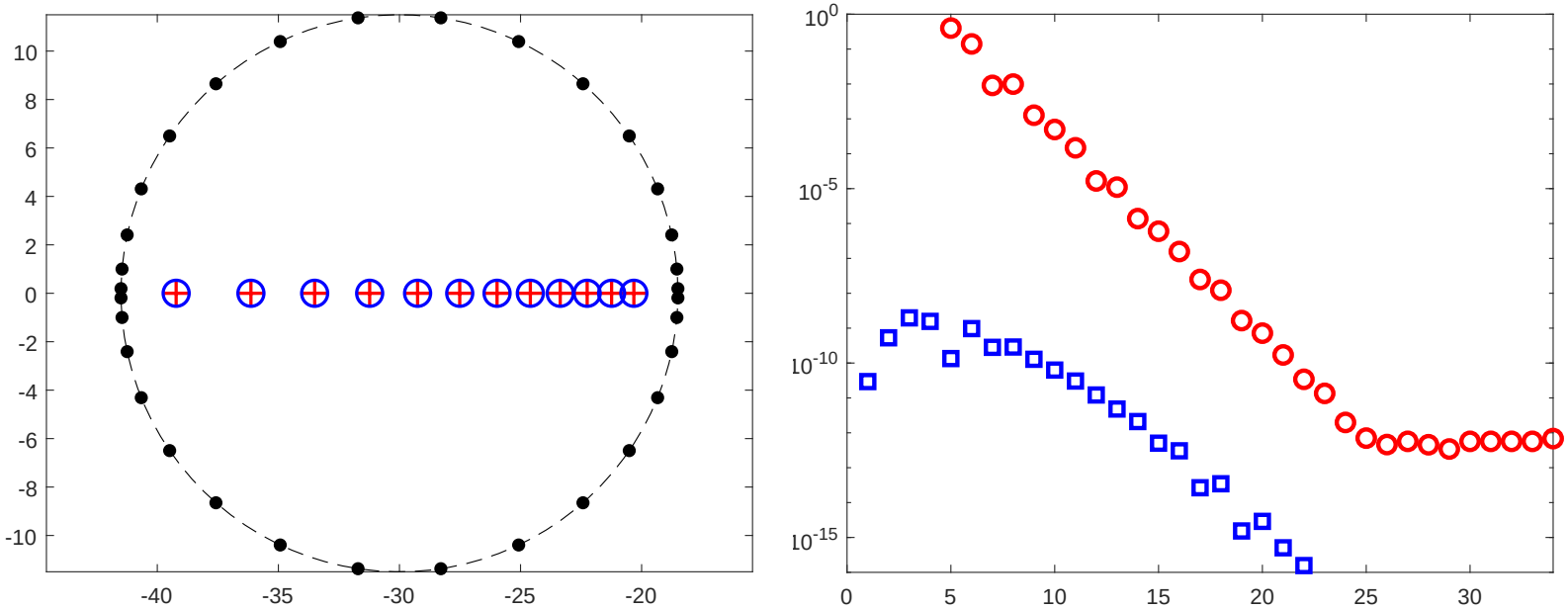


Figure: Left: Eigenvalues of the expanded problem (plus) and by Beyn's method (circle). Right: The errors e_{n_c} of the rational approximation of $e^\lambda - 1$ (square) and λ^2 (circle) versus the number of quadrature nodes n_c.

- We need $n_c = 32$ to get rational approximations of $f_1(\lambda) = e^\lambda - 1$ and $f_2(\lambda) = \lambda^2$ inside contour $\mathcal{C}$ up to the accuracy of 10^{-12}.
- Using $n_c = 32$ Gauss-Legendre quadrature nodes, 12 eigenvalues of (18) were computed **using shift-and-invert Arnoldi** applied to (12).
- For Beyn's method $n_c = 50$ was required to reach a backward error of the 12 eigenvalues smaller than 10^{-10}.

We repeat the same experiment using the **reduced subspace iteration**.

- Start with $\boldsymbol{\nu} = 40$ random vectors and apply $\boldsymbol{q} = 10$ steps of shifted inverse iteration to obtain a block of ν vectors each of size n.
- Orthogonalize the resulting ν vectors to **obtain an orthonormal basis U to perform the Rayleigh-Ritz** projection.
- We then **solve reduced (nonlinear) eigenvalue problem** (17) **of size** $(n_c + 1)\nu \ll (n_c + 1)n$ by solving an expanded problem.
- Before each restart of reduced subspace iteration, **select ν approximate eigenpairs whose eigenvalues are inside the contour $\mathcal{C}$**.
- We compare these results with the AAA algorithm [Nakatsukasa, Sete, Trefethen 2018] and the NLEIGS algorithm [Güttel, Van Beeumen, Meerbergen, Michiels 2014].

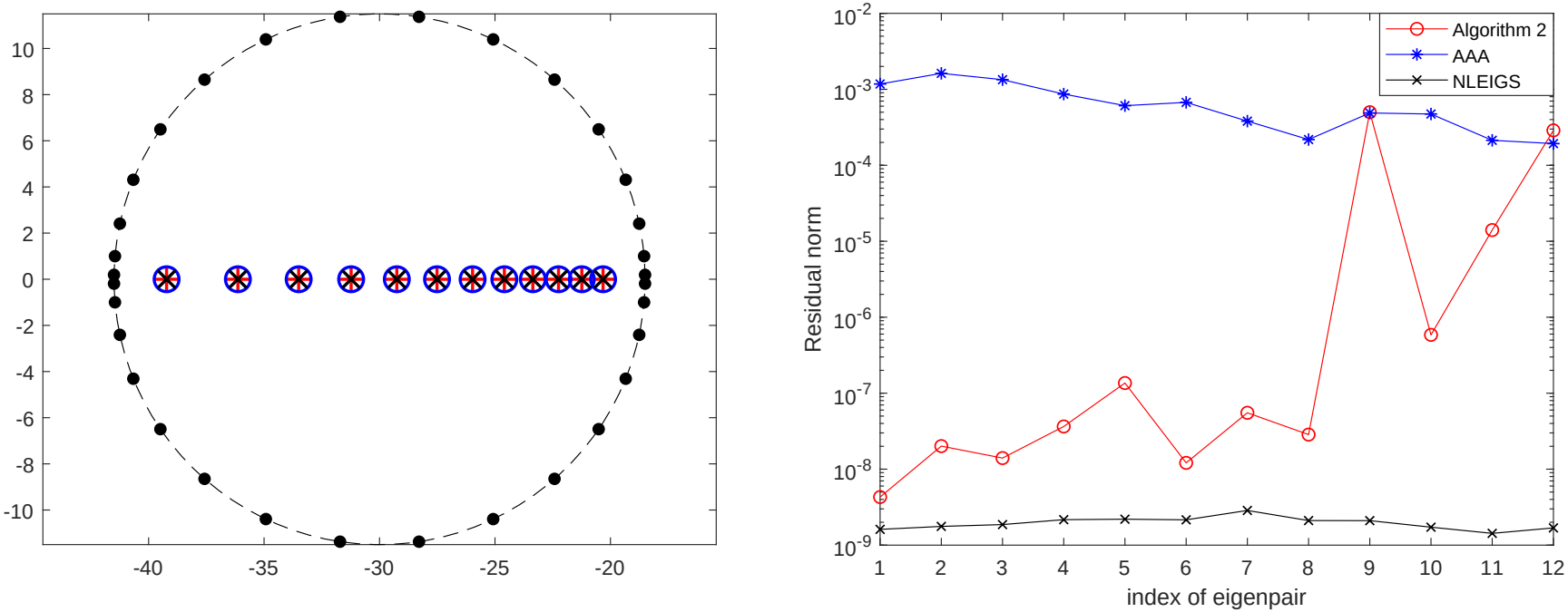


Figure: Left: Eigenvalues of (18) obtained by reduced subspace iteration (plus), the AAA algorithm (circle) and the NLEIGS algorithm (cross). Right: The residual norm $\|T(\lambda)x\|_\infty$ of the computed eigenpairs.

- Reduced subspace iteration computes 12 eigenvalues of interest requiring $\ell = 6$ outer iterations.
- For the AAA and NLEIGS algorithms, the boundary circle is discretized by 100 equispaced points. 7 and 38 nodes are needed to approximate $T(\lambda)$ up to 10^{-12} inside the circle, respectively.

Example 2: Acoustic behavior of pump casing model

Figure: Geometry and BE mesh of the pump casing model partitioned into 3 479 652 triangles, leading to the nonlinear eigenvalue problem with 1 728 508 DoFs.

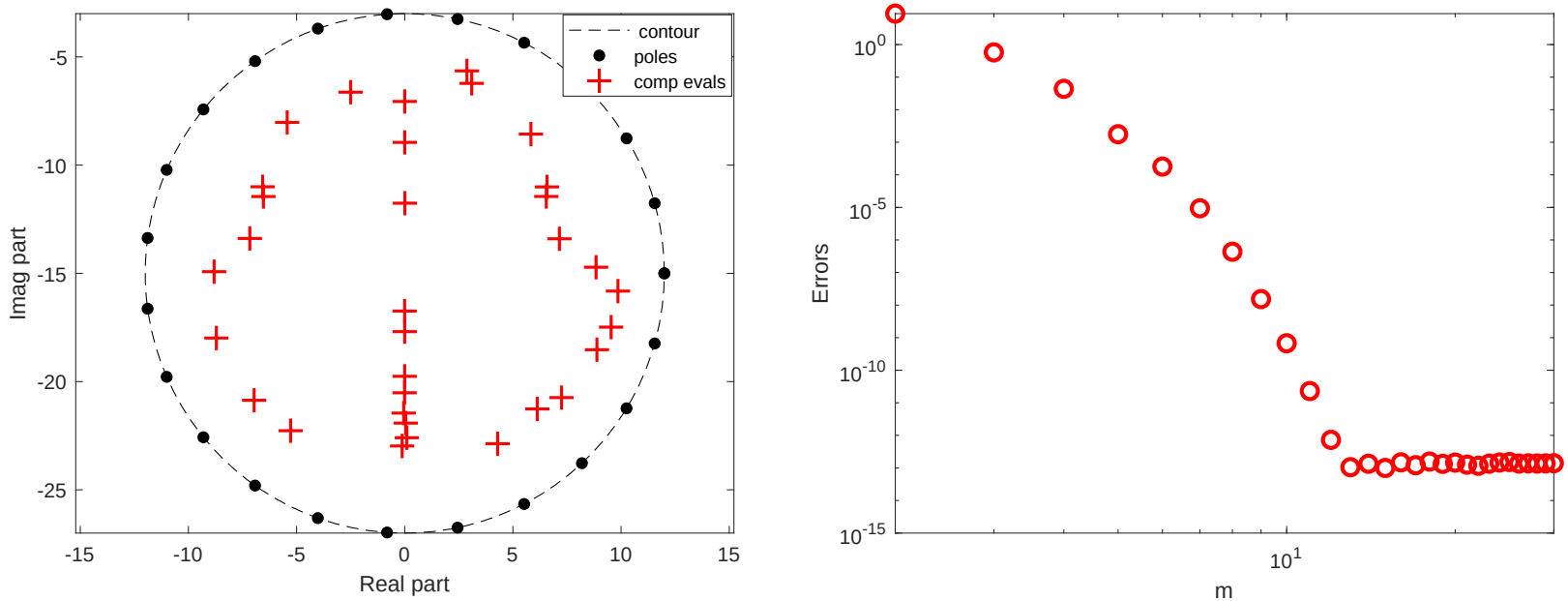
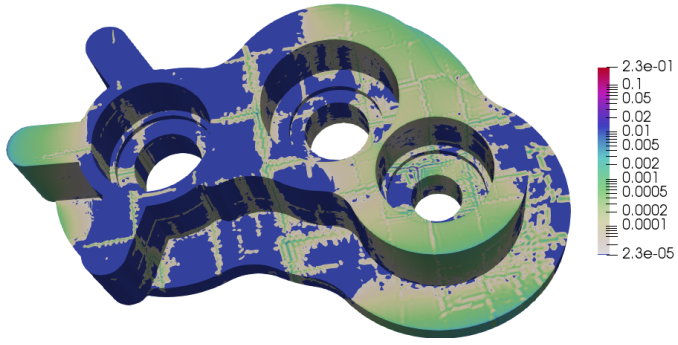
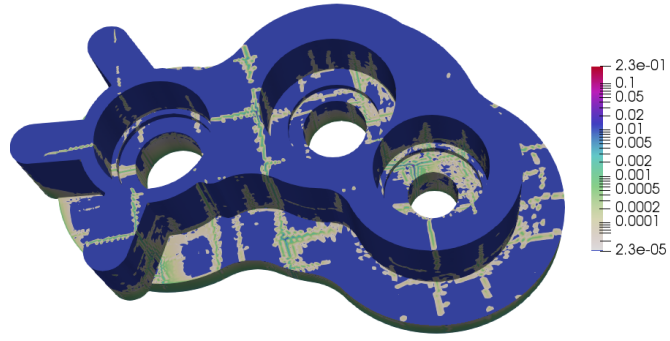


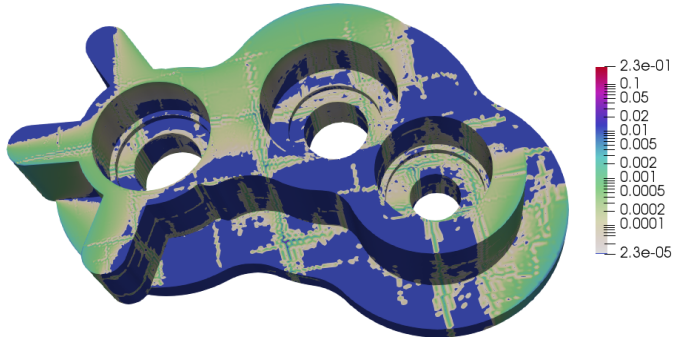
Figure: Left: 33 eigenvalues of the 3D Helmholtz problem on the pump model domain with Robin boundary conditions computed via (12). Right: Approximation errors versus the order n_c of the approximation inside a unit circle.



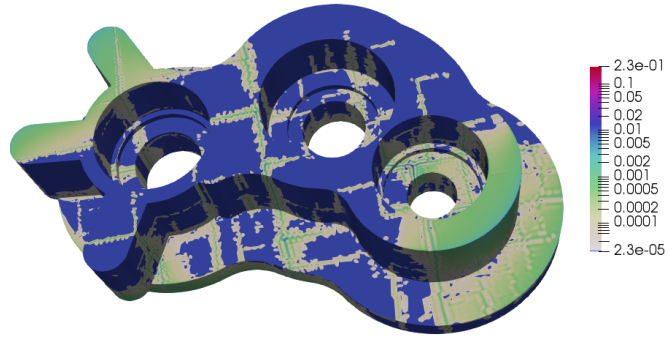
Eigenfrequency: 2.87 - 5.64i Hz



Eigenfrequency: 5.83 - 8.56i Hz



Eigenfrequency: 3.09 - 6.22i Hz



Eigenfrequency: -2.50 - 6.63i Hz

Figure: Eigenmodes of the pump model corresponding to four different eigenvalues calculated up to the relative residuals $\frac{\|T(\lambda)u\|_2}{\|u\|_2} \leq 10^{-9}$.

Few more details:

- $n_c = 24$ trapezoidal quadrature nodes,
- assembling **$\mathcal{H}$-matrix compressed versions of 24 matrices B_i and performing the matrix-vector multiplications with B_i** (efficiently parallelized on 32 cores with per-core memory limit of ≈ 31 GB) took 7.3 hours,
- reduced subspace iteration was used to solve the same problem with 10 704 **triangles, initial subspace size $\nu = 40$, $\ell = 20$ outer iterations and $q = 10$** steps of inverse iteration method,
- **parallelism** has been exploited for **computing matrices B_i and $\widehat{B}_i$**, and to **obtain an approximate subspace** $span\{U\}$.

Thank you very much for your attention.