# NUMERICAL METHODS FOR SPARSE NONLINEAR EIGENVALUE PROBLEMS

## Heinrich Voss

Hamburg University of Technology, Department of Mathematics,
D-21071 Hamburg, Federal Republic of Germany,
voss@tu-harburg.de

**Abstract**

This paper surveys numerical methods for general sparse nonlinear eigenvalue problems with special emphasis on iterative projection methods like Jacobi–Davidson, Arnoldi or rational Krylov methods. We briefly sketch a new approach to structure preserving projection methods, but we do not review the rich literature on polynomial eigenproblems which take advantage of a linearization of the problem.

## 1 Introduction

In this paper we consider the nonlinear eigenvalue problem

$$T(\lambda)x = 0 \tag{1}$$

where $T(\lambda) \in \mathbb{C}^{n \times n}$ is a family of matrices depending on a parameter $\lambda \in D$, and $D \subset \mathbb{C}$ is an open set. As in the linear case, $\lambda \in D$ is called an eigenvalue of problem (1) if equation (1) has a nontrivial solution $x \neq 0$. Then $x$ is called an eigenvector corresponding to $\lambda$.

A wide variety of applications requires the solution of a nonlinear eigenvalue problem. Quadratic problems

$$T(\lambda) := \lambda^2 M + \lambda C + K \tag{2}$$

arise in the dynamic analysis of structures such as damped vibrations of structures having a finite number of degrees of freedom [15], [42] (where the stiffness matrix $K$

and the mass matrix $M$ are symmetric and positive (semi-)definite, and the damping matrix is general), or vibrations of spinning structures yielding conservative gyroscopic systems [12], [31] (where $K = K^T$ and $M = M^T$ are positive (semi-)definite, and $C = -C^T$ is skew–symmetric), the study of corner singularities in anisotropic elastic materials [1], [2] (where $K = K^T$, $M = M^T$, $C = -C^T$, $M$ is positive definite and $K$ is negative definite), constrained least squares problems [14], and control of linear mechanical systems with a quadratic cost functional [38] (where $M$ and $K$ are Hamiltonian matrices and $C$ is skew–Hamiltonian). [55] surveys quadratic eigenvalue problems, its many applications, its mathematical properties, and some numerical solution techniques.

Polynomial eigenvalues

$$(3) \qquad T(\lambda)x = \sum_{j=0}^{k} \lambda^j A_j x = 0$$

of higher degree than two arise when discretizing a linear eigenproblem by dynamic elements [42], [57], [58] or by least squares elements [44], [45] (i.e. if one uses ansatz functions in a Rayleigh–Ritz approach which depend polynomially on the eigenparameter).

Rational eigenproblems

$$(4) \qquad T(\lambda)x = -Kx + \lambda Mx + \sum_{j=1}^{p} \frac{\lambda}{\sigma_j - \lambda} C_j x = 0$$

where $K = K^T$ and $M = M^T$ are positive definite and $C_j = C_j^T$ are matrices of small rank govern free vibration of plates with elastically attached masses [35], [54], [62] and vibrations of fluid solid structures [10], [41], [64]. A similar problem

$$(5) \qquad T(\lambda)x = -Kx + \lambda Mx + \lambda^2 \sum_{j=1}^{p} \frac{1}{\omega_j - \lambda} C_j x = 0$$

arises when a generalized linear eigenproblem is condensed exactly [40], [56]. Both problems only have real eigenvalues which can be characterized as minmax values of a Rayleigh functional [62].

A rational eigenproblem is obtained as well for the free vibrations of a structure if one uses a viscoelastic constitutive relation to describe the behaviour of a material [18], [19]. A finite element model obtains the form

$$(6) \qquad T(\omega) := \left( \omega^2 M + K - \sum_{j=1}^{k} \frac{1}{1 + b_j \omega} \Delta K_j \right)x = 0$$

where the stiffness and mass matrices $K$ and $M$ are positive definite, $k$ denotes the number of regions with different relaxation parameters $b_j$, and $\Delta K_j$ is an assemblage of element stiffness matrices over the region with the distinct relaxation constants.

Finally, a more general dependence on the eigenparameter appears in dynamic element methods when using non–polynomial ansatz functions [42] or in the stability analysis of vibrating systems under state delay feedback control [22], [23].

Most of the examples mentioned above are large and sparse, and typically only a small number of eigenvalues are of interest. Numerical methods have to exploit the sparseness fully to be efficient in storage and computing time.

For linear sparse eigenproblems $T(\lambda) = \lambda B - A$ very efficient methods are iterative projection methods (Lanczos method, Arnoldi method, Jacobi–Davidson method, e.g.), where approximations to the wanted eigenvalues and eigenvectors are obtained from projections of the eigenproblem to subspaces of small dimension which are expanded in the course of the algorithm. Essentially two types of methods are in use: methods which project the problem to a sequence of Krylov spaces like the Lanczos or the Arnoldi method [4], and methods which aim at a specific eigenpair expanding a search space by a direction which has a high approximation potential for the eigenvector under consideration like the Jacobi–Davidson method [4] or the Riccati method [8].

The Krylov subspace approaches take advantage of the linear structure of the underlying problem and construct an approximate incomplete Schur factorization (or incomplete spectral decomposition in the Hermitean case) from which they derive approximations to some of the extreme eigenvalues and corresponding eigenvectors, whereas the second type aims at the wanted eigenvalues one after the other using the Schur decomposition only to prevent the method from converging to eigenpairs which have been obtained already in a previous step.

For general nonlinear eigenproblems a normal form like the Schur factorization does not exist. Therefore, generalizations of Krylov subspace methods can be applied only to nonlinear problems if they are equivalent to a linear eigenproblem. It is well known that every polynomial eigenproblem can be linearized in several ways [15], [32], one of them being the straightforward manner which results in an eigenproblem for a block Frobenius matrix. However, applying a Krylov subspace method to a linearization always increases the dimension of the problem by the factor $k$ (the degree of the polynomial), and secondly symmetry properties which the original system may have in general are destroyed by a linearization.

In many applications the polynomial eigenproblem has some structure that should be reflected in its linearization, and should be exploited in its numerical solution for efficiency, stability and accuracy reasons. Bauchau [5] applied a two–sided Lanczos process (introduced in [7] for quadratic eigenproblems) to a symmetric/skew–symmetric linearization of a gyroscopic system thus preserving the property that the eigenvalues appear as purely imaginary pairs and avoiding complex arithmetic. More generally, Mehrmann and Watkins considered polynomial eigenproblems the spectrum of which have Hamiltonian structure, i.e. its eigenvalues appear in quadruples $\{\lambda, \overline{\lambda}, -\lambda, -\overline{\lambda}\}$ or in real or purely imaginary pairs $\{\lambda, -\lambda\}$ [1], [2], [38]. They studied a linearization that transforms the problem into a Hamiltonian/skew–Hamiltonian

pencil for which they developed a structure preserving skew–Hamiltonian, isotropic, implicitly restarted shift–and–invert Arnoldi algorithm called SHIRA [37].

Li and Ye [34] proposed a generalization of the Arnoldi method to the monic quadratic matrix polynomial $\lambda^2 I - \lambda A - B$ which does not take advantage of a linearization. Reducing the matrices $A$ and $B$ simultaneously to generalized Hessenberg matrices $H_k = Q_k^H A Q_k$ and $K_k = Q_k^H B Q_k$ by a sequence of orthogonal matrices $Q_k$ they derive at quadratic pencils $\theta^2 I - \theta H_k - K_k$ of much smaller dimension the Ritz pairs of which approximate eigenpairs of the original pencil. In [20] they generalized this approach to polynomial eigenproblems [20]. Bai [3] suggested for the monic quadratic pencil a projection to a generalized Krylov space which is spanned by mixed powers of the matrices $A$ and $B$

For general nonlinear eigenproblems Ruhe [47], [48],[50] generalized the rational Krylov approach for linear eigenproblems [49] by nesting the linearization of problem (1) by Lagrangean interpolation and the solution of the resulting linear eigenproblem by Arnoldi's method, where the Regula falsi iteration and the Arnoldi recursion are knit together. The name is a little misleading since no Krylov space is constructed but the method can be interpreted as a projection method where the search spaces are expanded by directions with high approximation potential for the eigenvector wanted next, namely by the vector obtained by some residual inverse iteration [28]. This method has the drawback, that potential symmetry properties of the underlying problem are destroyed which is not the case for the Arnoldi method in [59], [60], [61] which expands the search space by a different residual inverse iteration (again no Krylov space appears; the name is chosen because the method reduces to the shift–and–invert Arnoldi method if applied to a linear eigenproblem). Expanding the search space by an approximate inverse iteration one arrives at a Jacobi–Davidson method which was introduced by Sleijpen, Booten, Fokkema and van der Vorst [51] for polynomials and in [6] and [63] for general nonlinear eigenproblems.

In this paper we review the iterative projection methods for general (i.e. not necessarily polynomial) sparse nonlinear eigenproblems which generalize the Jacobi–Davidson approach for linear problems in the sense that the search space in every step is expanded by a vector with high approximation potential for the eigenvector wanted next. Although we have in mind sparse eigenproblems Section 2 summarizes methods for dense nonlinear eigenproblems which are needed in the iterative projection methods of Jacobi–Davidson, Arnoldi and rational Krylov type presented in Section 3. The paper closes with some numerical examples in Section 4 demonstrating the efficiency of the methods.

# 2   Methods for dense nonlinear eigenproblems

In this section we review methods for dense nonlinear eigenproblems. Typically, they require several factorizations of varying matrices to approximate one eigenvalue, and therefore, they are not appropriate for large and sparse problems. However, they

are needed within projection methods for sparse problems to solve the nonlinear projected problems of small dimension.

## 2.1 Solver of the characteristic equation

Clearly $\lambda$ is an eigenvalue of the nonlinear problem, if and only if it is a root of the characteristic equation

$$(7) \qquad \det T(\lambda) = 0.$$

Kublanovskaya [29], [30] proposed a method to solve (7) taking advantage of the QR decomposition. Let

$$(8) \qquad T(\lambda)P(\lambda) = Q(\lambda)R(\lambda)$$

be the QR factorization of $T(\lambda)$ where $P(\lambda)$ is a permutation matrix which is chosen such that the diagonal elements $r_{jj}(\lambda)$ of $R(\lambda)$ are decreasing in magnitude, i.e. $|r_{11}(\lambda)| \geq |r_{22}(\lambda)| \geq \cdots \geq |r_{nn}(\lambda)|$. Then $\lambda$ solves equation (7) if and only if

$$(9) \qquad f(\lambda) := r_{nn}(\lambda) = 0.$$

Applying Newton's method to this equation one obtains the following iteration method

$$(10) \qquad \lambda_{k+1} = \lambda_k - \frac{1}{e_n^H Q(\lambda_k)^H T'(\lambda_k)P(\lambda_k)R(\lambda_k)^{-1}e_n}.$$

for approximations to an eigenvalue of problem (7), where $e_n$ denotes the unit vector having a one in its last component. Approximations to left and right eigenvectors can be obtained from

$$y_k = Q(\lambda_k)e_n \quad \text{and} \quad x_k = P(\lambda_k)R(\lambda_k)^{-1}e_n.$$

Kublanovskaya's proof assumed that the elements $Q(\lambda)$ and $R(\lambda)$ in the factorization (7) are analytic which in general is not true, and it was assumed that the diagonal elements of $Q'(\lambda)^H Q(\lambda)$ are zero which can only be shown to be purely imaginary. Jain and Singhal [25] pointed out that in spite of these errors in the derivation Kublanovskaya's algorithm does work and provides quadratic convergence. Moreover, they modified the method in [25] and [26]. A similar approach was presented by Yang [69] who derived a representation of Newton's method for the characteristic equation (7) using the LU factorization of $T(\lambda)$.

5

## 2.2 Inverse iteration

For linear eigenproblems $Ax = \lambda x$ it is well known that inverse iteration is equivalent to Newton's method applied to the nonlinear system

$$\begin{pmatrix} Ax - \lambda x \\ v^H x - 1 \end{pmatrix} = 0$$

where $v \in \mathbb{C}^n$ is chosen suitably. Correspondingly for the nonlinear problem we obtain from

(11) $$F(x, \lambda) := \begin{pmatrix} T(\lambda)x \\ v^H x - 1 \end{pmatrix} = 0$$

by one step of Newton's method

(12) $$\begin{pmatrix} T(\lambda_k) & T'(\lambda_k)x_k \\ v^H & 0 \end{pmatrix} \begin{pmatrix} x_{k+1} - x_k \\ \lambda_{k+1} - \lambda_k \end{pmatrix} = - \begin{pmatrix} T(\lambda_k)x_k \\ v^H x_k - 1 \end{pmatrix}.$$

The first component yields

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

i.e. the direction of the new approximation to an eigenvector is

$$u_{k+1} := T(\lambda_k)^{-1}T'(\lambda_k)x_k.$$

Assuming that $x_k$ is already normalized by $v^H x_k = 1$ the second component of (12) reads $v^H x_{k+1} = v^H x_k$, and multiplying equation (13) by $v^H$ yields

$$\lambda_{k+1} = \lambda_k - \frac{v^H x_k}{v^H u_{k+1}}.$$

Hence, for nonlinear eigenproblems inverse iteration obtains the form given in Algorithm 1.

---
**Algorithm 1** Inverse iteration

---
1: Start with $\lambda_0$, $x_0$ such that $v^H x_0 = 1$
2: **for** $k = 0, 1, 2, \ldots$ until convergence **do**
3:    solve $T(\lambda_k)u_{k+1} = T'(\lambda_k)x_k$ for $u_{k+1}$
4:    $\lambda_{k+1} = \lambda_k - (v^H x_k)/(v^H u_{k+1})$
5:    normalize $x_{k+1} = u_{k+1}/v^H u_{k+1}$
6: **end for**

---

Inverse iteration has the following convergence properties.

**THEOREM 1**

Let $\lambda$ be an eigenvalue of $T(\cdot)$ with corresponding eigenvector $x$ such that $v^H x = 1$.

Assume that $T'(\lambda)$ is nonsingular and that 0 is an algebraically simple eigenvalue of $T'(\lambda)^{-1}T(\lambda)$. Then Algorithm 1 converges locally and quadratically to $(x, \lambda)$.

**Proof**: Since Algorithm 1 was derived from Newton's method it suffices to prove that the Jacobian matrix $F'(x, \lambda)$ is nonsingular. Let

$$(14) \qquad F'(x, \lambda) \begin{pmatrix} z \\ \mu \end{pmatrix} = \begin{pmatrix} T(\lambda) & T'(\lambda) \\ v^H & 0 \end{pmatrix} \begin{pmatrix} z \\ \mu \end{pmatrix} = 0.$$

If $\mu = 0$ then it follows from the the first component $T(\lambda)z = 0$. Hence, $z = \alpha x$ for some $\alpha \in \mathbb{C}$, and $0 = v^H z = \alpha v^H x = \alpha$ yields $z = 0$.

If $\mu \neq 0$ then the first component reads $T(\lambda)z = -\mu T'(\lambda)x$. Multiplying by $T(\lambda)^{-1}$ yields $T(\lambda)^{-1}T'(\lambda)z = -\mu x \neq 0$, from which we obtain

$$(T(\lambda)^{-1}T'(\lambda))^2 z = -\mu T(\lambda)^{-1}T'(\lambda)x = 0,$$

contradicting the fact that 0 is assumed to be an algebraically simple eigenvalue of $T(\lambda)^{-1}T'(\lambda)$. $\qquad \square$

Clearly the normalization condition can be modified in each step arriving at Algorithm 2.

---

**Algorithm 2** Inverse iteration (modified)

---

1: Start with $\lambda_0$, $x_0$ such that $v_0^H x_0 = 1$
2: **for** $k = 0, 1, 2, \ldots$ until convergence **do**
3:     solve $T(\lambda_k)u_{k+1} = T'(\lambda_k)x_k$ for $u_{k+1}$
4:     $\lambda_{k+1} = \lambda_k - (v_k^H x_k)/(v_k^H u_{k+1})$
5:     normalize $x_{k+1} = u_{k+1}/v_{k+1}^H u_{k+1}$
6: **end for**

---

Ruhe [46] suggested to use $v_k = T(\lambda_k)^H y_k$ where $y_k$ is an approximation to a left eigenvector. Then the update for $\lambda$ becomes

$$\lambda_{k+1} = \lambda_k - \frac{y_k^H T(\lambda_k)x_k}{y_k^H T'(\lambda_k)x_k},$$

which is the Rayleigh functional for general nonlinear eigenproblems proposed by Lancaster [32], and which can be interpreted as one Newton step for solving the equation $f_k(\lambda) := y_k^H T(\lambda)x_k = 0$.

For linear Hermitean eigenproblems one receives even cubic convergence if $\lambda_k$ is updated by the Rayleigh quotient. The same holds true (cf. Rothe [44]) for symmetric nonlinear eigenproblems having a Rayleigh functional $p$ (cf. Subsection 2.5) if we replace statement 4. in Algorithm 2 by $\lambda_{k+1} = p(u_{k+1})$, where $p(u_{k+1})$ denotes the real root of $u_{k+1}^H T(\lambda)u_{k+1} = 0$ closest to $\lambda_k$.

## 2.3 Residual inverse iteration

For linear eigenproblems inverse iteration can be replaced by a simplified version $x_{k+1} = (A - \sigma I)^{-1} x_k$ with fixed $\sigma$ converging to an eigenvector corresponding to the eigenvalue of $A$ next to $\sigma$. The convergence is only linear but the method has the advantage that only one factorization of the matrix $A - \sigma I$ is necessary.

In contrast to the linear case replacing step 3: in Algorithm 1 by $x_{k+1} = T(\sigma)^{-1} T'(\lambda_k) x^k$ with a fixed shift $\sigma$ results in misconvergence. It is easily seen that this iteration converges to an eigenpair of the linear problem $T(\sigma) x = \gamma T'(\tilde{\lambda}) x$ ($\gamma \neq 0$ and $\tilde{\lambda}$ depending on the normalization condition) from which we can not recover an eigenpair of the nonlinear problem (1).

A remedy against this wrong convergence was proposed by Neumaier [39]. Assume that $T(\lambda)$ is twice continuously differentiable. Then for the increment in Algorithm 1 it holds

$$
\begin{aligned}
dx &= x_k - x_{k+1} = x_k + (\lambda_{k+1} - \lambda_k) u_{k+1} \\
&= x_k + (\lambda_{k+1} - \lambda_k) T(\lambda_k)^{-1} T'(\lambda_k) x_k \\
&= T(\lambda_k)^{-1} (T(\lambda_k) + (\lambda_{k+1} - \lambda_k) T'(\lambda_k)) x_k \\
&= T(\lambda_k)^{-1} T(\lambda_{k+1}) x_k + \mathcal{O}(|\lambda_{k+1} - \lambda_k|^2).
\end{aligned}
$$

Neglecting the second order term one gets

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

The advantage of this approach is that replacing $\lambda_k$ by a fixed shift $\sigma$ does not generate misconvergence.

To update the approximation to an eigenvalue $\hat{\lambda}$ Neumaier suggested $\lambda_{k+1} = p(x_k)$ if $T(\lambda)$ is a family of Hermitean matrices and $\hat{\lambda}$ is real, and the solution $\lambda_{k+1}$ of the equation

$$
v^H T(\sigma)^{-1} T(\lambda) x_k = 0
$$

which is closest to $\hat{\lambda}$ in the general case. For a fixed vector $v$ and $\sigma$ close to $\hat{\lambda}$ the vector $y = T(\sigma)^{-H} v$ can be considered as an approximate left eigenvector, and applying one Newton step to $y^H T(\lambda) x_k = 0$ results again in the general Rayleigh functional of Lancaster.

The residual inverse iteration method in Algorithm 3 has the following convergence properties which were proved by Neumaier [39].

**THEOREM 2**

Let $T(\lambda)$ be twice continuously differentiable, $\hat{\lambda}$ be a simple zero of $\det T(\lambda) = 0$, and let $\hat{x}$ be an eigenvector normalized by $v^H \hat{x} = 1$. Then the residual inverse iteration converges for all $\sigma$ sufficiently close to $\hat{\lambda}$, and it holds

$$
\frac{\|x_{k+1} - \hat{x}\|}{\|x_k - \hat{x}\|} = \mathcal{O}(|\sigma - \hat{\lambda}|) \quad \text{and} \quad |\lambda_{k+1} - \lambda_k| = \mathcal{O}(\|x_k - \hat{x}\|^q),
$$

**Algorithm 3** Residual inverse iteration

---

1: Let $v$ be a normalization vector and start with an approximations $\sigma$ and $x_1$ to
   an eigenvalue and corresponding eigenvector of (1) such that $v^H x_1 = 1$
2: **for** $k = 1, 2, \ldots$ until convergence **do**
3:   solve $v^H T(\sigma)^{-1} T(\lambda_{k+1}) x_\ell = 0$ for $\lambda_{k+1}$
     or set $\lambda_{k+1} = p(x_k)$ is $T(\lambda)$ is Hermitean
4:   compute the residual $r_k = T(\lambda_{k+1}) x_k$
5:   solve $T(\sigma) d_k = r_k$ for $d_k$
6:   set $z_{k+1} = x_k - d_k$
7:   normalize $x_{k+1} = z_{k+1}/v^H z_{k+1}$
8: **end for**

---

where $q = 2$ if $T(\lambda)$ is Hermitean, $\hat{\lambda}$ is real, and $\lambda_{k+1} = p(x_k)$ in step 3:, and $q = 1$ otherwise.

## 2.4   Successive linear approximations

A first order approximation of problem (1) is

$$(15) \qquad T(\lambda)x \approx (T(\tilde{\mu}) - \theta T'(\tilde{\mu}))x = 0, \ \theta = \tilde{\mu} - \lambda.$$

This suggests the method of successive linear problems in Algorithm 4 which was introduced by Ruhe [46], and which converges quadratically.

---

**Algorithm 4** Method of successive linear problems

---

1: Start with an approximation $\lambda_1$ to an eigenvalue of (1)
2: **for** $k = 1, 2, \ldots$ until convergence **do**
3:   solve the linear eigenproblem $T(\lambda_k)u = \theta T'(\lambda_k)u$
4:   choose an eigenvalue $\theta$ smallest in modulus
5:   $\lambda_{k+1} = \lambda_k - \theta$
6: **end for**

---

**THEOREM 3**
Let $T$ be twice continuously differentiable, and let $\hat{\lambda}$ be an eigenvalue of problem (1) such that $T'(\hat{\lambda})$ is nonsingular and 0 is an algebraically simple eigenvalue of $T'(\hat{\lambda})^{-1}T(\hat{\lambda})$. Then the method of successive linear problems converges quadratically to $\hat{\lambda}$.

**Proof**: Let $\hat{x}$ be an eigenvector corresponding to $\hat{\lambda}$, and let $v \in \mathbb{C}^n$ such that $v^H \hat{x} = 1$. Let $U(\hat{\lambda})$ be a neighbourhood of $\hat{\lambda}$, and $\Phi : \mathbb{C}^n \times \mathbb{C} \times U(\hat{\lambda}) \to \mathbb{C}^{n+1}$ be defined by

$$\Phi(x, \theta, \lambda) := \begin{pmatrix} T(\lambda)x - \theta T'(\lambda)x \\ v^H x - 1 \end{pmatrix}$$

9

Then $\Phi(\hat{x}, 0, \hat{\lambda}) = 0$, and the matrix

$$\frac{\partial}{\partial(x, \theta)} \Phi(\hat{x}, 0, \hat{\lambda}) = \begin{pmatrix} T(\hat{\lambda}) & -T'(\hat{\lambda})\hat{x} \\ v^H & 0 \end{pmatrix}$$

is nonsingular, which was already proved in Theorem 1.

By the implicit function theorem $\Phi(x, \theta, \lambda) = 0$ defines differentiable functions $x : U(\hat{\lambda}) \to \mathbb{C}^n$ and $\theta : U(\hat{\lambda}) \to \mathbb{C}$ on a neighbourhood of $\hat{\lambda}$ again denoted by $U(\hat{\lambda})$ such that

$$\Phi(x(\lambda), \theta(\lambda), \lambda) = 0 \quad \text{for every } \lambda \in U(\hat{\lambda}).$$

With this functions the method of successive linear problems can be rewritten as the fixed point iteration $\lambda_{n+1} = \phi(\lambda_n) := \lambda_n - \theta(\lambda_n)$, which converges quadratically if $\phi'(\hat{\lambda}) = 0$.

From the implicit function theorem it follows

$$\begin{aligned} \frac{d}{d\lambda} \begin{pmatrix} x \\ \theta \end{pmatrix} (\hat{\lambda}) &= -\frac{\partial}{\partial(x, \theta)} \Phi(\hat{x}, 0, \hat{\lambda})^{-1} \frac{\partial}{\partial\lambda} \Phi(\hat{x}, 0, \hat{\lambda}) \\ &= -\begin{pmatrix} T(\hat{\lambda}) & -T'(\hat{\lambda})\hat{x} \\ v^H & 0 \end{pmatrix}^{-1} \begin{pmatrix} T'(\hat{\lambda})\hat{x} \\ 0 \end{pmatrix} \end{aligned}$$

which yields $\theta'(\hat{\lambda}) = 1$, and therefore $\phi'(\hat{\lambda}) = 0$. $\qquad \square$

## 2.5 Safeguarded iteration

The numerical methods of the preceding subsections apply to general nonlinear eigenproblems, although for Hermitean problems and real eigenvalues inverse iteration and residual inverse iteration converge faster if the eigenvalue approximations are updated using the Rayleigh functional. In this subsection we consider the safeguarded iteration which applies only to Hermitean problems allowing a variational characterization of their eigenvalues [11], [43], [16], [17], [66], [65]. The method was introduced by Werner [68] for overdamped problems, and was studied by Werner and the author [67] for the nonoverdamped case.

Let $J \subset \mathbb{R}$ be an open interval which may be unbounded, and assume that $T(\lambda) \in \mathbb{C}^{n \times n}$ is a family of Hermitean matrices the elements of which are differentiable. We assume that for every $x \in \mathbb{C}^n \setminus \{0\}$ the real equation

$$(16) \qquad\qquad f(\lambda, x) := x^H T(\lambda) x = 0$$

has at most one solution $\lambda \in J$. Then equation (16) defines a functional $p$ on some subset $D \subset \mathbb{C}^n$ which obviously generalizes the Rayleigh quotient for linear pencils $T(\lambda) = \lambda B - A$, and which we call the Rayleigh functional of the nonlinear eigenvalue problem (1). We further assume that

$$(17) \qquad\qquad x^H T'(p(x)) x > 0 \quad \text{for every } x \in D$$

generalizing the definiteness requirement for linear pencils. By the implicit function theorem $D$ is an open set, and differentiating the identity $x^H T(p(x))x = 0$ one obtains, that the eigenvectors of (1) are stationary points of $p$.

Under the conditions above we proved in [66] a minmax principle for the nonlinear eigenproblem (1) if the eigenvalues are enumerated appropriately. $\lambda \in J$ is an eigenvalue of (1) if and only if $\mu = 0$ is an eigenvalue of the matrix $T(\lambda)$, and by Poincaré's maxmin principle there exists $m \in \mathbb{N}$ such that

$$0 = \max_{\dim V = m} \min_{x \in V, \ x \neq 0} \frac{x^H T(\lambda) x}{\|x\|^2}.$$

Then we assign this $m$ to $\lambda$ as its number and call $\lambda$ an $m$-th eigenvalue of problem (1).

With this enumeration it holds (cf. [66]) that for every $m \in \{1, \ldots, n\}$ problem (1) has at most one $m$-th eigenvalue in $J$, which can be characterized by

$$(18) \qquad \lambda_m = \min_{\dim V = m, D \cap V \neq \emptyset} \ \sup_{v \in D \cap V} p(v).$$

Conversely, if

$$(19) \qquad \lambda_m := \inf_{\dim V = m, D \cap V \neq \emptyset} \ \sup_{v \in D \cap V} p(v) \in J,$$

then $\lambda_m$ is an $m$-th eigenvalue of (1), and the characterization (18) holds. The minimum is attained by the invariant subspace of $T(\lambda_m)$ corresponding to its $m$ largest eigenvalues, and the supremum is attained by any eigenvector of $T(\lambda_m)$ corresponding to $\mu = 0$.

To prove this characterization we took advantage of the following relation

$$(20) \qquad \lambda \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} \lambda_m \quad \Leftrightarrow \quad \mu_m(\lambda) := \max_{\dim V = m} \ \min_{x \in V, \ x \neq 0} \frac{x^H T(\lambda) x}{\|x\|^2} \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} 0.$$

The enumeration of eigenvalues and the fact that the eigenvectors of (1) are the stationary vectors of the Rayleigh functional suggests the method in Algorithm 5 called safeguarded iteration for computing the $m$–th eigenvalue.

---

**Algorithm 5** Safeguarded iteration

---

1: Start with an approximation $\sigma_1$ to the $m$-th eigenvalue of (1)
2: **for** $k = 1, 2, \ldots$ until convergence **do**
3:    determine an eigenvector $x_k$ corresponding to the $m$-largest eigenvalue of $T(\sigma_k)$
4:    solve $x_k^H T(\sigma_{k+1}) x_k = 0$ for $\sigma_{k+1}$
5: **end for**

---

The following theorem contains the approximation properties of the safeguarded iteration. It was already proved in [67] but because this technical report is not easily available we repeat its proof here.

**THEOREM 4**

(i) If $\lambda_1 := \inf_{x \in D} p(x) \in J$ and $x_1 \in D$ then the safeguarded iteration converges globally to $\lambda_1$.

(ii) If $\lambda_m \in J$ is a $m$-th eigenvalue of (1) which is simple then the safeguarded iteration converges locally and quadratically to $\lambda_m$.

(iii) Let $T(\lambda)$ be twice continuously differentiable, and assume that $T'(\lambda)$ is positive definite for $\lambda \in J$. If $x_k$ in step 3: of Algorithm 5 is chosen to be an eigenvector corresponding to the $m$ largest eigenvalue of the generalized eigenproblem $T(\sigma_k)x = \mu T'(\sigma_k)x$ then the convergence is even cubic.

**Proof**: (i): Assume that $x_{k-1} \in D$. Then $\sigma_k = p(x_{k-1}) \geq \lambda_1$, and (20) yields

$$(21) \qquad \mu_1(\sigma_k) = \max_{x \neq 0} \frac{x^H T(\sigma_k)x}{x^T x} = \frac{x_k^H T(\sigma_k)x_k}{x_k^H x_k} \geq 0.$$

Suppose that $x_k \notin D$. Then it follows from (21) that $x_k^H T(\lambda)x_k > 0$ for every $\lambda \in J$.

Let $\tilde{x} \in D$ be an eigenvector of $T$ corresponding to $\lambda_1$. Then we get from (17) $\tilde{x}^H T(\lambda)\tilde{x} < 0$ for every $\lambda \in J$, $\lambda < \lambda_1$. Hence for fixed $\lambda \in J$, $\lambda < \lambda_1$

$$q(t) := (\tilde{x} + t(x_k - \tilde{x}))^H T(\lambda)(\tilde{x} + t(x_k - \tilde{x})) = 0$$

has a solution $\tilde{t} \in (0,1)$, i.e. $w := \tilde{x} + \tilde{t}(x_k - \tilde{x}) \in D$ and $p(w) = \lambda < \lambda_1$ contradicting (18).

The monotonicity of $\{\sigma_k\}$ follows directly from the definition of $\sigma_{k+1}$, (21) and (17). Let $\hat{\sigma} := \lim_{k \to \infty} \sigma_k$ and let $\{x_{k_j}\}$ be a convergent subsequence of $\{x_k\}$, $x_{k_j} \to \hat{x} \neq 0$.

Then by the continuity of $T(\lambda)$

$$0 = x_{k_j}^H T(\sigma_{k_j+1})x_{k_j} \to \hat{x}^H T(\hat{\sigma})\hat{x},$$

i.e. $\hat{x} \in D$ and $p(\hat{x}) = \hat{\sigma}$, and we get from the continuous dependence of $\mu_1(\sigma)$ on $\sigma$

$$T(\hat{\sigma})\hat{x} = \lim_{j \to \infty} T(\sigma_{k_j})x_{k_j} = \lim_{j \to \infty} \mu_1(\sigma_{k_j})x_{k_j} = \mu_1(\hat{\sigma})\hat{x}.$$

Multiplying this equation by $\hat{x}^H$ yields $\mu_1(\hat{\sigma}) = 0$, and hence $\hat{\sigma} = \lambda_1$.

(ii): If $\lambda_m$ is a simple eigenvalue of $T$ then it is an easy consequence of the implicit function theorem that for $|\lambda - \lambda_m|$ small enough the function $\lambda \to x(\lambda)$ is defined and continuously differentiable, where $x(\lambda)$ denotes the suitably normalized eigenvector of $T(\lambda)$ corresponding to the $m$-largest eigenvalue. Because $D$ is an open

set, $h(\lambda) := p(x(\lambda))$ is defined in a neighbourhood of $\lambda_m$, and since the eigenvalues of $T$ are the stationary values of $p$, we get

$$h'(\lambda_m) = p'(x(\lambda_m))x'(\lambda_m)) = 0.$$

This proves the quadratic convergence of $\sigma_k = h(\sigma_{k-1})$ to $\lambda_m$.

(iii): Let $T'(\lambda)$ be positive definite and denote by $\mu(\lambda)$ the $m$-largest eigenvalue of the generalized eigenproblem $T(\lambda)x = \mu T'(\lambda)x$ and by $x(\lambda)$ a corresponding eigenvector which is suitably normalized such that $x(\cdot)$ is continuous. If $\lambda_m$ is an $m$-th eigenvalue of $T(\cdot)$ then $\mu(\lambda_m) = 0$, and differentiating $T(\lambda)x(\lambda) = \mu(\lambda)T'(\lambda)x(\lambda)$ yields

$$T'(\lambda_m)x(\lambda_m) + T(\lambda_m)x'(\lambda_m) = \mu'(\lambda_m)T'(\lambda_m)x(\lambda_m).$$

Multiplying by $x(\lambda_m)^H$ from the left we get $\mu'(\lambda_m) = 1$, and therefore

$$(22) \qquad\qquad\qquad T(\lambda_m)x'(\lambda_m) = 0.$$

If we define $h$ analogously to part (ii) by $h(\lambda) = p(x(\lambda))$ then as before $h'(\lambda_m) = 0$, and from

$$h''(\lambda_m) = -2\frac{x'(\lambda_m)^H T(p(x(\lambda_m)))x'(\lambda_m)}{x(\lambda_m)^H T'(p(x(\lambda_m)))x(\lambda_m)}$$

and (22) it follows $h''(\lambda_m) = 0$, i.e. the safeguarded iteration converges cubically. $\square$

## 3 Iterative projection methods

For sparse linear eigenvalue problems

$$(23) \qquad\qquad\qquad\qquad Ax = \lambda x$$

iterative projection methods like the Lanczos, Arnoldi, rational Krylov or Jacobi–Davidson method are very efficient. Here the dimension of the eigenproblem is reduced by projecting it to a subspace of much smaller dimension, and handling the reduced problem by a fast technique for dense problems. The subspaces are expanded in the course of the algorithm in an iterative way with the aim that some of the eigenvalues of the reduced matrix become good approximations of some of the wanted eigenvalues of the given large matrix.

Two types of iterative projection methods are in use: methods which expand the subspaces independently of the eigenpair of the projected problem and which use Krylov subspaces of $A$ or $(A - \sigma I)^{-1}$ for some shift $\sigma$ like the Arnoldi method or Lanczos method or rational Krylov method, and methods which aim at a particular eigenpair and choose the expansion $q$ such that it has a high approximation potential for a wanted eigenvector like the Jacobi–Davidson method or Riccati methods.

The Arnoldi method (and other Krylov subspace methods have similar properties) projects problem (23) onto the Krylov space

$$\mathcal{K}_k(A, v_1) = \text{span}\{v_1, Av_1, A^2 v_1, \dots, A^{k-1} v_1\}$$

where $v_1$ is an initial vector. It produces an orthogonal basis $V_k$ of $\mathcal{K}_k(A, v_1)$ such that the projected matrix is upper Hessenberg (which we now denote by $H_k$) such that

$$(24) \qquad\qquad AV_k = V_k H_k + f_k e_k^T$$

where $e_k \in \mathbb{R}^k$ is the $k$-th unit vector having a 1 in its last component, and $f_k$ is orthogonal to the columns of $V_k$, i.e. $V_k^H f_k = 0$. (24) demonstrates that $V_k^H A V_k = H_k$ is the orthogonal projection of $A$ to $\mathcal{K}_k(A, v_1)$.

If $(y, \theta)$ is an eigenpair of the projected problem, and $x = V_k y$ is the corresponding approximation to an eigenvector of the original problem (23) (which is called a Ritz vector corresponding to $\theta$), then it holds for the residuum

$$r := Ax - \theta x = AV_k y - \theta V_k y = V_k H_k y - \theta V_k y + f_k e_k^H y = (e_k^T y) f_k.$$

Hence, one obtains an error indicator $\|r\| = |e_k^T y| \cdot \|f_k\|$ for the eigenpair approximation $(x, \theta)$ without computing the Ritz vector $x$. If $A$ is Hermitean then this is even an error bound.

The Arnoldi method (together with its shifted and inverted and its restarted variants) is a standard solver for sparse linear eigenproblems. A detailed discussion is contained in [4]. Software implementing the (implicitly restarted) Arnoldi method is available in the package ARPACK [33] and the MATLAB command `eigs`.

The Arnoldi method converges to the extreme eigenvalues first. If one is interested in eigenvalues in the interior of the spectrum (eigenvalues close to a given shift $\sigma$, e.g.) one has to apply the Arnoldi method to a shifted and inverted matrix $(A - \sigma I)^{-1}$, i.e. one has to determine a factorization of $A - \sigma I$ which may be prohibitive for very large problems.

A remedy against this disadvantage is the Jacobi–Davidson method introduced by Sleijpen and van der Vorst [53]. Let $(x, \theta)$ be an approximation to an eigenpair obtained by a projection method with subspace $V$. We assume that

$$(25) \qquad\qquad \|x\| = 1, \ \theta = x^H Ax, \ r := Ax - \theta x \perp x.$$

Then the most desirable orthogonal correction $z$ solves the equation

$$(26) \qquad\qquad A(x + z) = \lambda(x + z), \quad z \perp x.$$

As $z \perp x$ the operator $A$ can be restricted to the subspace orthogonal to $x$ yielding $\tilde{A} := (I - xx^H)A(I - xx^H)$, and from $\theta = x^H Ax$ it follows

$$A = \tilde{A} + Axx^H + xx^H A - \theta xx^H.$$

Hence, we obtain from (26) and $\tilde{A}z = 0$

$$(27) \qquad (\tilde{A} - \lambda I)z = -r + (\lambda - \theta - x^H A z)x.$$

Since both, the left hand side and $r$ are orthogonal to $x$, the factor $\lambda - \theta - x^H A z$ must vanish, and therefore the correction $z$ has to satisfy $(\tilde{A} - \lambda I)z = -r$. Because $\lambda$ is unknown it is replaced by its approximation $\theta$, and we end up with the correction equation

$$(28) \qquad (I - xx^H)(A - \theta I)(I - xx^H)z = -r.$$

It can be shown that the expanded space $[V, z]$ for the Jacobi–Davidson method contains the direction $u = (A - \theta I)^{-1}x$ which is obtained by one step of inverse iteration [52]. One therefore can expect similar approximation properties, i.e. quadratic or even cubic convergence, if the problem is Hermitean.

Obviously, the Jacobi–Davidson method is aiming at a particular eigenpair (close to the shift $\theta$). If one is interested in more than one eigenvalue one has to use deflation based on the Schur decomposition of the matrix $A$ (cf. [13]).

As in the shift-and-invert Arnoldi method we have to solve a large linear system. However, numerical experiments demonstrate that problem (28) does not have to be solved exactly, but only approximately to maintain fast convergence. Normally only a small number of steps of a preconditioned Krylov subspace method are sufficient to obtain a good expansion $z$ for the subspace $V$ in the iterative projection method. Implementation details of the Jacobi–Davidson method for various types of linear eigenvalue problems can be found in [4]. Implementations in FORTRAN and MATLAB can be downloaded from the home page of Gerhard Sleijpen (http://www.math.ruu.nl/people/sleijpen).

Deriving the correction equation in the Jacobi–Davidson method one only uses the fact in (27) that $(\tilde{A} - \theta I)z$ and $r$ are orthogonal to $x$. One does not take advantage of the equation $\lambda - \theta - x^H A z = 0$. Brandts [8], [9] also took into account this equation, and developed the Riccati method which improves the Jacobi–Davidson method.

In the following we discuss generalizations of iterative projection methods to nonlinear eigenproblems. There are many papers on Arnoldi's method for polynomial and in particular quadratic eigenvalue problems taking advantage of their equivalence to linear eigenproblems of higher dimension, which will not be considered here. We will only briefly sketch in Subsection 3.1 a new and interesting approach by Mehrmann and Watkins who introduced a structure preserving Arnoldi methods for Hamiltonian eigenproblems. Further structure preserving projection methods for the special case of a conservative gyroscopic problem are contained [21], [36] and [55]. In Subsections 3.2 – 3.4 we concentrate on iterative projection methods for general nonlinear eigenproblems.

## 3.1 Polynomial eigenproblems

In this section we consider polynomial eigenproblems

$$(29) \qquad P(\lambda)x = \sum_{j=0}^{k} \lambda^j A_j x = 0$$

where $A_j \in \mathbb{C}^{n \times n}$ are given matrices. It is well known that problem (29) can be linearized, i.e. can be replaced by an equivalent linear eigenproblem. For instance the following linearization follows the same lines as the well known way replacing a $k$:th order linear system of differential equations by a first order system:

$$(30) \qquad Ay = \lambda By$$

where

$$A = \begin{pmatrix} -A_{k-1} & -A_{k-2} & \ldots & -A_1 & -A_0 \\ I & & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & I \end{pmatrix},$$

$$B = \begin{pmatrix} A_k & & & & \\ & I & & & \\ & & \ddots & & \\ & & & I & \\ & & & & I \end{pmatrix}, \text{ and } y = \begin{pmatrix} \lambda^{k-1}x \\ \lambda^{k-2}x \\ \vdots \\ \lambda x \\ x \end{pmatrix}.$$

Problem (30) can be solved by any eigensolver for sparse linear systems such as the Arnoldi or the Jacobi–Davidson method. However, this approach bears two disadvantages: first, the dimension grows from $n$ to $kn$, and second, symmetry properties which the original system (29) may have in general are destroyed by the linearization. Hence, this straightforward manner of linearization usually is not the best way to solve the polynomial eigenproblem (29).

For polynomial eigenproblems having certain symmetry properties appropriate linearizations and suitable projection methods preserving these symmetry properties are useful. Quite recently Mehrmann and Watkins suggested a method of this type for problem (29) the spectrum of which has Hamiltonian structure ([1], [2], [37], [38]).

Assume that the matrices $A_j$ in (29) are real and alternating symmetric and skew-symmetric, i.e. $A_j^T = (-1)^j A_j$ or $A_j^T = (-1)^{j+1} A_j$, $j = 0, \ldots, k$, (a conservative gyroscopic system, e.g.) then it holds

$$P(\lambda)x = 0 \quad \iff \quad x^T P(-\lambda) = 0.$$

This means that the spectrum of $P(\lambda)x = 0$ has Hamiltonian symmetry, i.e. in general the eigenvalues appear in quadruples $\{\lambda, -\lambda, \bar{\lambda}, -\bar{\lambda}\}$ (notice that we assumed the matrices $A_j$ to be real), or in pairs if they are purely imaginary or real.

The following theorem the proof of which is obvious contains a linearization of (29) which is the basis of the structure preserving approach. This linearization without the alternating signs was already given by Lancaster [32].

**THEOREM 5**

$P(\lambda)x = 0$ has the same eigenvalues as the pencil $A - \lambda B$ where

$$
A = \left(\begin{array}{c|ccccc}
-A_0 & O & O & O & \ldots & O \\
\hline
O & -A_2 & -A_3 & -A_4 & \ldots & -A_k \\
O & A_3 & A_4 & & & O \\
O & -A_4 & & & & O \\
\vdots & \vdots & & & & \vdots \\
O & \pm A_k & O & O & \ldots & O
\end{array}\right),
$$

$$
B = \left(\begin{array}{ccccc|c}
A_1 & A_2 & A_3 & \ldots & A_{k-1} & A_k \\
\hline
-A_2 & -A_3 & -A_4 & \ldots & -A_k & O \\
A_3 & A_4 & & & O & O \\
-A_4 & & & & O & O \\
\vdots & & & & \vdots & \vdots \\
\pm A_k & O & O & \ldots & O & O
\end{array}\right).
$$

The pencil $A - \lambda B$ obviously is skew-symmetric/symmetric.

If the dimension of the linearized problem is even then multiplying by $J := \begin{pmatrix} O & I \\ -I & O \end{pmatrix}$ one gets an equivalent pencil

$$\lambda H_1 - H_2$$

where $(JH_1)^T = -JH_1$ (i.e. the matrix $H_1$ is skew-Hamiltonian) and $(JH_2)^T = JH_2$ (i.e. the matrix $H_2$ is Hamiltonian).

If $H_1$ can be represented in product form $H_1 = Z_1 Z_2$ such that $Z_2^T J = \pm J Z_1$ then this pencil is equivalent to the standard eigenvalue problem

$$Z_1^{-1} H_2 Z_2^{-1} - \lambda I =: W - \lambda I$$

in which the matrix $W$ can be easily shown to be Hamiltonian.

Applying Arnoldi's method to the skew-Hamiltonian matrix $-W$ then (in exact arithmetic) the Krylov space $\mathcal{K}_k(W, q_1)$ is isotropic, i.e.

$$x^T J y = 0 \quad \text{for all } x, y \in \mathcal{K}_k(W, q_1).$$

Therefore, to keep the Hamiltonian structure in the Arnoldi process one should not only orthogonalize against $q_1, \ldots, q_k$, but also against $Jq_1, \ldots, Jq_k$.

Moreover, for the same reason shifts have to used in pairs $\{\sigma, -\sigma\}$ (if eigenvalues are real or purely imaginary) or in quadruples $\{\sigma, -\sigma, \bar{\sigma}, -\bar{\sigma}\}$.

This results in the algorithm SHIRA, a structure-preserving skew-Hamiltonian, isotropic, implicitly restarted shift-and-invert Arnoldi algorithm proposed by Mehrmann and Watkins [37].

We do not discuss SHIRA in detail. We only consider for conservative gyroscopic pencils the basic steps of the reduction process, and we demonstrate that doubling the dimension in linearization does not yield larger linear systems to be solved in the shift–and–invert steps of the method.

Applying the structure preserving linearization to the conservative gyroscopic quadratic pencil

$$Q(\lambda) := \lambda^2 M + \lambda G + K$$

where $M^T = M$, $G^T = -G$ and $K^T = K$ yields the symmetric/skew-symmetric pencil

$$\begin{pmatrix} -K & O \\ O & -M \end{pmatrix} - \lambda \begin{pmatrix} G & M \\ -M & O \end{pmatrix},$$

and multiplying by $-J$ we obtain the skew-Hamiltonian/Hamiltonian pencil

$$\begin{pmatrix} O & M \\ -K & O \end{pmatrix} - \lambda \begin{pmatrix} M & O \\ G & M \end{pmatrix}$$

which with

$$\begin{pmatrix} M & O \\ G & M \end{pmatrix} = \begin{pmatrix} M & O \\ 0.5G & I \end{pmatrix} \begin{pmatrix} I & O \\ 0.5G & M \end{pmatrix} = Z_1 Z_2$$

can be reduced to the skew-Hamiltonian eigenproblem

$$Wy := \begin{pmatrix} M & O \\ 0.5G & I \end{pmatrix}^{-1} \begin{pmatrix} O & M \\ -K & O \end{pmatrix} \begin{pmatrix} I & O \\ 0.5G & M \end{pmatrix}^{-1} y = \lambda y.$$

In this case the eigenvalues are purely imaginary, and therefore appear in pairs $\{\lambda, -\lambda\}$. To preserve is structure shifts have to be used in pairs, too, and one has to apply the isotropic Arnoldi process to

$$
\begin{aligned}
R &= (W - \sigma I)^{-1}(W + \sigma I)^{-1} = Z_2(H_2 - \sigma H_1)^{-1}Z_1 Z_2(H_2 + \sigma H_1)^{-1}Z_1 \\
&= \begin{pmatrix} I & O \\ 0.5G & M \end{pmatrix} \begin{pmatrix} I & O \\ \sigma I & I \end{pmatrix} \begin{pmatrix} O & M^{-1} \\ -Q(\sigma)^{-1} & O \end{pmatrix} \times \\
&\quad \times \begin{pmatrix} M & O \\ G & M \end{pmatrix} \begin{pmatrix} O & M^{-1} \\ -Q(-\sigma)^{-1} & O \end{pmatrix} \begin{pmatrix} I & O \\ -\sigma I & I \end{pmatrix} \begin{pmatrix} M & O \\ 0.5G & I \end{pmatrix}
\end{aligned}
$$

which demonstrates that although the dimension of the eigenproblem is doubled one only has to solve linear systems the dimension of which is the original one.

## 3.2 An Arnoldi type method

In the following we consider iterative projection methods for the general nonlinear eigenproblem (1). We already pointed out that in this case the search spaces have to be expanded by directions having a high approximation potential for the eigenvector wanted next.

Inverse iteration converges quadratically for simple eigenvalues, and therefore is a suitable candidate. If $V$ is an orthonormal basis of the search space, if $(\lambda, y)$ is a solution of the projected problem

$$V^H T(\lambda) V y = 0,$$

and if $x = Vy$ is the corresponding Ritz vector, then the expansion $\tilde{v} = T(\lambda)^{-1} T'(\lambda) x$ would be a reasonable choice. However, in this case we would have to solve a high dimensional linear system in every iteration step where the coefficient matrix varies. One way out is the residual inverse iteration suggested by Neumaier [39], and given by $\tilde{x} = x - T(\sigma)^{-1} T(\lambda) x$ where $\sigma$ is a fixed shift (not to far away from the eigenvalue targeted at).

In iterative projection methods the new search direction is orthonormalized against the previous ansatz vectors. Since the Ritz vector $x$ is contained in the span of $V$ we may choose the new direction $v = T(\sigma)^{-1} T(\lambda) x$ as well. For the linear problem $T(\lambda) = A - \lambda B$ this is exactly the Cayley transform with pole $\sigma$ and zero $\lambda$, and since

$$(A - \sigma B)^{-1}(A - \lambda B) = I + (\lambda - \sigma)(A - \sigma B)^{-1} B$$

and Krylov spaces are shift-invariant the resulting projection method expanding $V$ by $v$ is nothing else but the shift-and-invert Arnoldi method.

If the linear system $T(\sigma)v = T(\lambda)x$ is too expensive to solve for $v$ we may choose as new direction $v = MT(\lambda)x$ with $M \approx T(\sigma)^{-1}$, and for the linear problem we obtain an inexact Cayley transform or a preconditioned Arnoldi method.

We therefore call the resulting iterative projection method which was studied in [36] for quadratic eigenvalue problems and in [60] and [59] for general nonlinear eigenproblems nonlinear Arnoldi method. We stress the fact that differently from the linear case no Krylov space is determined in the course of the algorithm.

Since we are interested in all eigenvalues in some region and since the speed of convergence is expected to depend crucially on $|\sigma - \lambda|$ it will be advisable to change the shift or more generally the preconditioner $M$ in the course of the algorithm if the convergence to the current eigenvalue becomes too slow. So actually we obtain a method which generalizes the rational Krylov method for linear problems in [49], and the name nonlinear rational Krylov method would be appropriate, too. However, since Ruhe [50] already introduced a rational Krylov method for nonlinear problems which differs from our method quite a bit we prefer the name nonlinear Arnoli method. We will summarize the rational Krylov method in Subsection 3.4 and comment on the differences of Ruhe's and our approach there.

**Algorithm 6** Nonlinear Arnoldi Method

1: start with an initial pole $\sigma$ and an initial basis $V$, $V^H V = I$;
2: determine preconditioner $M \approx T(\sigma)^{-1}$, $\sigma$ close to first wanted eigenvalue
3: $k = 1$
4: **while** $m \leq$ number of wanted eigenvalues **do**
5:   compute appropriate eigenvalue $\mu$ and corresponding eigenvector $y$ of the projected problem $T_V(\mu)y := V^H T(\mu)Vy = 0$.
6:   determine Ritz vector $u = Vy$ and residual $r_k = T(\mu)u$
7:   **if** $\|r_k\|/\|u\| < \epsilon$ **then**
8:     PRINT $\lambda_m = \mu$, $x_m = u$,
9:     **if** $m ==$ number of wanted eigenvalues **then**
10:       STOP
11:     **end if**
12:     $m = m + 1$
13:     **if** $(k > 1)$ & $(\|r_{k-1}\|/\|r_k\| > \text{tol})$ **then**
14:       choose new pole $\sigma$
15:       determine new preconditioner $M \approx T(\sigma)^{-1}$
16:     **end if**
17:     restart if necessary
18:     choose approximations $\mu$ and $u$ to next eigenvalue and eigenvector
19:     determine residual $r = T(\mu)u$
20:     $k = 0$
21:   **end if**
22:   $v = Mr$
23:   $v = v - VV^H v$ ,$\tilde{v} = v/\|v\|$, $V = [V, \tilde{v}]$
24:   reorthogonalize if necessary
25:   update projected problem $T_V(\mu) = V^H T(\mu)V$
26:   $k = k + 1$
27: **end while**

A template for the preconditioned Arnoldi method for nonlinear eigenvalue problems with restarts and varying preconditioner is contained in Algorithm 6. In the following we comment on some of its steps.

1: Here preinformation such as known approximate eigenvectors of problem (1) or eigenvectors of contiguous problems can be introduced into the algorithm.

If no information on eigenvectors is at hand, and we are interested in eigenvalues close to the parameter $\sigma \in D$, one can choose an initial vector at random, execute a few Arnoldi steps for the linear eigenproblem $T(\sigma)u = \theta u$ or $T(\sigma)u = \theta T'(\sigma)u$, and choose $V$ by orthogonalizing eigenvectors corresponding to small eigenvalues in modulus. Starting with a random vector without this preprocessing usually will yield a value $\mu$ in step 5: which is far away from $\sigma$ and will avert convergence.

For the rational eigenvalue problem (4) governing free vibrations of a plate with attached masses or a fluid–solid structure we discussed the choice of the initial space in [62]

2: In our numerical examples we used the $LU$ factorization of $T(\sigma)$ if this could be determined inexpensively and otherwise an incomplete $LU$ factorization, but every other preconditioner is fine.

3: $k$ counts the number of iterations for fixed $m$. This is only needed to measure the speed of convergence and to decide whether a new preconditioner is recommended in condition 13:

4: Every other stopping criterion can replace the requirement to determine $m$ eigenvalues.

5: Since the dimension of the projected problems are usually small they can be solved by any method for dense nonlinear eigenvalue problems discussed in Section 2.

A crucial point in iterative projection methods for general nonlinear eigenvalue problems when approximating more than one eigenvalue is to inhibit the method to converge to the same eigenvalue repeatedly. For linear eigenvalue problems this is no problem. Krylov subspace solvers construct an orthogonal basis of the ansatz space not aiming at a particular eigenvalue, and one gets approximations to extreme eigenvalues without replication (at least if re-orthogonalization is employed). If several eigenvalues are computed by the Jacobi–Davidson method then one determines an incomplete Schur factorization thus preventing the method from approaching an eigenvalue which was already obtained previously (cf. [13]). For nonlinear problems a similar normal form does not exist.

If $T(\lambda)$ is a family of real symmetric or Hermitean matrices and $D$ is a real interval such that the eigenvalues are maxmin values of a Rayleigh functional then the projected problems inherit this property. The eigenvalues can be determined one after the other by safeguarded iteration, and approximating the $m$-th eigenvalue usually enough information about the next eigenvector is gathered to compute the $(m + 1)$-th eigenvalue safely. This approach which was discussed in [60] (and in [6] for the nonlinear Jacobi–Davidson method) has the advantage that it is most unlikely that the method converges to an eigenvalue that has already been found previously.

Similarly, in the general case one can order the eigenvalues by their distance to a fixed parameter $\sigma_0$, and approximate them one after the other by the method of successive linear problems. If already $m-1$ eigenvalues of (1) closest to $\sigma_0$ have been determined, and $\mu_0$ is an approximation to the eigenvalue wanted next, we iteratively perform the following three steps until convergence: we solve the linear eigenproblem $V^H T(\mu_\ell) V y = \theta v^H T'(\mu_\ell) V y$, choose the eigenvalue $\hat{\theta}$ such that $|\sigma_0 - (\mu_\ell - \hat{\theta})|$ is $m$-smallest among the eigenvalues $\theta$, and set $\mu_{\ell+1} = \mu_\ell - \hat{\theta}$.

A disadvantage of this method is the fact that consecutive eigenvalues $\lambda_{m-1}$ and $\lambda_m$ usually will not be close to each other, and therefore, a preconditioner which was adequate for one eigenvalue can yield slow convergence of the iterative solver for the next eigenvalue. Hence, this method should be used only if a small number of eigenvalues close to a parameter is wanted.

Quite often the nonlinear eigenvalue problem under consideration is a (small) perturbation of a linear eigenvalue problem. In (6) we considered a rational eigenproblem governing the free vibrations of a structure using a viscoelastic constitutive relation to describe the behaviour of the material. It is well known that often the eigenmodes of the damped and the undamped problem do not differ very much although the eigenvalues do. Therefore, it is reasonable to determine an eigenvector $y$ of the undamped and projected problem $(\omega^2 V^H M V - V^H K V) y = 0$ corresponding to the $m$-smallest eigenvalue $\omega_m^2$, determine an approximate eigenvalue $\tilde{\omega}$ of the nonlinear projected problem from the complex equation $y^H V^H T(\omega) V y = 0$ or $e^H V^H T(\sigma)^{-1} T(\omega) V y = 0$, and correct it by one of the methods in Section 2.

13: Corresponding to Theorem 2 the residual inverse iteration with fixed pole $\sigma$ converges linearly, and the contraction rate satisfies $\mathcal{O}(|\sigma - \lambda_m|)$. We therefore update the preconditioner if the convergence measured by the quotient of the last two residual norms has become too slow.

In our numerical examples it happened that the condition in step 7: was fulfilled in the first step after having increased $m$. In this case the quotient of the last two residual norms does not say anything about the speed of convergence, and we do not update the preconditioner.

14: The new pole should not be chosen to close to an eigenvalue of $T(\cdot)$ because this would hamper the construction of the preconditioner. A general strategy cannot be given, but the proper way to choose a new pole depends on the problem under consideration and on the method in step 5: for solving the projected problem.

17: As the subspaces expand in the course of the algorithm the increasing storage or the computational cost for solving the projected eigenvalue problems may make it necessary to restart the algorithm and purge some of the basis vectors. Since a restart destroys information on the eigenvectors and particularly on the one the method is just aiming at we restart only if an eigenvector has just converged.

Since some of the solvers of the nonlinear projected eigenproblems in 5: take advantage of some enumeration of the eigenvalues it is natural to keep the eigenvectors that have been converged in the course of the algorithm. Otherwise this enumeration would be perturbed. We therefore continue with an orthonormal basis of $X_m := \mathrm{span}\{x_1, \ldots, x_m\}$. If an approximation to an eigenvector wanted next is obtained cheaply (cf. 18:) we add it to $X_m$.

18: Some of the eigensolvers discussed in Section 2 can be used to get approximations to the eigenvector and eigenvalue wanted next. In this case we continue with these approximations. If no information on the next eigenvalue and eigenvector can be gained cheaply we continue with the current approximations.

23: $v$ is orthogonalized with respect to the current search space $V$ by classical Gram–Schmidt. It may be replaced by modified Gram–Schmidt for stability reasons. Notice, however, that the classical Gram-Schmidt procedure is able to use BLAS3, and thus can be faster than classical Gram–Schmidt by a better use of cache.

24: If in statement 23: the norm of $v$ is reduced in the (classical or modified) Gram–Schmidt process by more than a modest factor $\kappa$, say $\kappa = 0.25$, then it is appropriate to repeat the Gram–Schmidt method once.

25: Often problem (1) has the form $T(\lambda) = \sum_{j=1}^{N} f_j(\lambda)C_j$ with differentiable complex functions $f_j$ and fixed matrices $C_j \in \mathbb{C}^{n \times n}$. Then the projected problem has the form

$$T_{V_k}(\lambda) = \sum_{j=1}^{N} f_j(\lambda)V_k^H C_j V_k =: \sum_{j=1}^{N} f_j(\lambda)C_{j,k},$$

and the matrices $C_{j,k}$ can be updated according to

$$C_{j,k} = \begin{pmatrix} C_{j,k-1} & V_{k-1}^H C_j \tilde{v} \\ \tilde{v}^H C_j V_{k-1} & \tilde{v}^H C_j \tilde{v} \end{pmatrix}.$$

## 3.3   A Jacobi–Davidson type method

Given a search space $V$, a solution $(\lambda, y)$ of the corresponding projected problem, and the Ritz vector $x = Vy$ in the last subsection we approximated the expansion of $V$ by the direction $v = T(\lambda)^{-1}T'(\lambda)x$ of inverse iteration by residual inverse iteration. A different way to avoid the solution of a large linear system in every iteration step is the Jacobi–Davidson approach.

A natural generalization of the Jacobi–Davidson method for linear eigenproblems which was already suggested in [51] for polynomial eigenvalue problems and which was studied in [63] and [6] for general nonlinear eigenproblems is the following one: Suppose that the columns of $V \subset \mathbb{C}^n$ form an orthonormal basis of the current search space, and let $(x, \theta)$ be a Ritz pair of (1) with respect to $V$, i.e.

$$\text{(31)} \qquad\qquad V^H T(\theta) V y = 0, \quad x = V y.$$

Then we consider the correction equation

$$\text{(32)} \qquad \left( I - \frac{px^H}{x^H p} \right) T(\theta) \left( I - \frac{xx^H}{x^H x} \right) z = -r, \quad z \perp x$$

where $p := T'(\theta)x$ and $r := T(\theta)x$. Restricting $T(\theta)$ in the correction equation to an operator from $x^\perp$ to $(T'(\theta)x)^\perp$ again guarantees that the direction of inverse iteration is contained in the subsequent search space. This can be seen as follows.

Equation (32) can be rewritten as

$$T(\theta)z - \alpha p = -r$$

where $\alpha$ has to be chosen such that $z \perp x$. Solving for $z$ we obtain

$$z = -x + \alpha T(\theta)^{-1} p = -x + \alpha T(\theta)^{-1} T'(\theta)x,$$

and $x = Vy$ yields $\tilde{z} := T(\theta)^{-1} T'(\theta)x \in \operatorname{span}[V, z]$.

Hence, as in the linear case the new search space $\operatorname{span}[V, z]$ contains the vector obtained by one step of inverse iteration with shift $\theta$ and initial vector $x$, and again we may expect quadratic or even cubic convergence of the resulting iterative projection method, if the correction equation (32) is solved exactly.

As in the linear case the correction equation is solved only approximately by a few steps of a Krylov solver with an appropriate preconditioner.

In the correction equation (32) the operator $T(\theta)$ is restricted to map the subspace $x^\perp$ to $(T'(\theta)x)^\perp$. Hence, if $K \approx T(\theta)$ is a preconditioner of $T(\theta)$ then a preconditioner for an iterative solver of (32) should be modified correspondingly to

$$\tilde{K} := (I - \frac{px^H}{x^H p})K(I - \frac{xx^H}{x^H x}).$$

With left-preconditioning equation (32) becomes

$$(33) \qquad \tilde{K}^{-1}\tilde{T}(\theta)z = -\tilde{K}^{-1}r, \quad z \perp x.$$

where

$$\tilde{T}(\theta) := (I - \frac{px^H}{x^H p})T(\theta)(I - \frac{xx^H}{x^H x}).$$

We apply a Krylov solver to equation (33) with initial guess $z = 0$. For the linear case this was already discussed in [53], and the transfer to equation (33) is straightforward.

Since the operator $\tilde{K}^{-1}\tilde{T}(\theta)$ maps the space $x^\perp$ into itself, and since the initial guess $z = 0$ is an element of $x^\perp$, all iterates are contained in this space, and therefore in each step we have to perform one matrix-vector product

$$(34) \qquad y = \tilde{K}^{-1}\tilde{T}(\theta)v$$

for some $v \in x^\perp$. To this end we first multiply $v$ by $\tilde{T}(\theta)$ which yields

$$\tilde{y} = (I - \frac{px^H}{x^H p})T(\theta)v = T(\theta)v - \frac{x^H T(\theta)v}{x^H p}p,$$

and then we solve

$$\tilde{K}y = \tilde{y}, \quad y \perp x.$$

This equation can be rewritten as

$$Ky - \alpha p = \tilde{y},$$

where $\alpha$ is determined from the condition $y \perp x$. Thus, we finally obtain

$$(35) \qquad y = K^{-1}\tilde{y} - \frac{x^H K^{-1}\tilde{y}}{x^H K^{-1}p}K^{-1}p$$

which demonstrates that taking into account the projectors in the preconditioner, i.e. using $\tilde{K}$ instead of $K$, raises the cost of the preconditioned Krylov solver only slightly. To initialize one has to solve the linear system $K\tilde{p} = p$ and to determine the scalar product $\alpha := x^H\tilde{p} = x^H K^{-1}p$. These computations have to be executed just once. Afterwards in each iteration step one has to solve only one linear system $Kw = \tilde{y}$ for $w$, one has to compute the scalar product $\beta := x^H w = x^H K^{-1}\tilde{u}$, and to perform one axpy $y = w - (\beta/\alpha)\tilde{y}$ to expand the Krylov space of $\tilde{K}^{-1}\tilde{T}(\theta)$.

A template for the Jacobi–Davidson method for the nonlinear eigenvalue problem (1) is given in Algorithm 7. The comments on the nonlinear Arnoldi in subsection 3.1 apply to this method, too. However, for the approximate solution of the correction equation we do not have a convergence result like Theorem 2 for the residual inverse iteration, and therefore the rule for updating the preconditioner in Algorithm 6 does not make sense here.

We solved the correction equation (32) by a few steps of preconditioned GMRES where we kept the preconditioner for a couple of eigenvalues. We terminated the solver of (32) in the $k$-th outer iteration for the $m$-th eigenvalue if the residual was reduced by at least $\tau_k = 2^{-k}$, and we allowed at most 10 steps of the solver. If the required accuracy $\tau_k$ was not met after at most 5 iteration steps we updated the preconditioner. However, we allowed at most one update for every eigenvalue $\lambda_m$.

---

**Algorithm 7** Nonlinear Jacobi–Davidson method

---

 1: Start with an initial basis $V$, $V^H V = I$; $m = 1$
 2: determine preconditioner $K \approx T(\sigma)^{-1}$, $\sigma$ close to first wanted eigenvalue
 3: **while** $m \le$ number of wanted eigenvalues **do**
 4:     compute approximation to $m$-th wanted eigenvalue $\lambda_m$ and corresponding eigenvector $x_m$ of projected problem $V^H T(\lambda) V x = 0$
 5:     determine Ritz vector $u = V x_m$ and residual $r = T(\lambda_m) u$
 6:     **if** $\|r\|/\|u\| < \epsilon$ **then**
 7:         PRINT approximate eigenpair $(\lambda_m, u)$; increase $m = m + 1$;
 8:         reduce search space $V$ if necessary
 9:         choose approximation $(\lambda_m, u)$ to next eigenpair
10:         compute residual $r = T(\lambda_m) u$;
11:     **end if**
12:     Find approximate solution of correction equation

$$(36) \qquad (I - \frac{T'(\lambda_m) u u^H}{u^H T'(\lambda_m) u}) T(\sigma) (I - \frac{u u^H}{u^H u}) t = -r$$

    (by preconditioned Krylov solver, e.g.)
13:     orthogonalize $t = t - V V^H t$, $v = t/\|t\|$, and expand subspace $V = [V, v]$
14:     determine new preconditioner $K \approx T(\lambda_m)^{-1}$ if necessary
15:     update projected problem
16: **end while**

---

Hwang, Lin and Mehrmann [24] considered a gyroscopic eigenproblem

$$(\lambda^2 M + \lambda(G + \varepsilon D) + K)x = 0$$

where $K = K^T$, $M = M^T$, $G = -G^T$ and $\varepsilon D$ represents the damping of the system. Since the damping is assumed to be small they suggested to determine eigenpairs $(\lambda_j, x_j)$ of the conservative gyroscopic problem $(\lambda^2 M + \lambda G + K)x = 0$ in the wanted region by the structure preserving method SHIRA from Subsection 3.1, and to improve these approximation by the Jacobi–Davidson method for the original quadratic problem.

## 3.4  Rational Krylov method

In [47], [48], [50] Ruhe generalized the rational Krylov approach for linear eigen-problems [49] to sparse nonlinear eigenvalue problems by nesting the linearization of problem (1) by Regula falsi and the solution of the resulting linear eigenproblem by Arnoldi's method, where the Regula falsi iteration and the Arnoldi recursion are knit together. Similarly as in the rational Krylov process a sequence $V_k$ of subspaces of $\mathbb{C}^n$ is constructed, and at the same time Hessenberg matrices $H_k$ are updated which approximate the projection of $T(\sigma)^{-1}T(\lambda_k)$ to $V_k$. Here $\sigma$ denotes a shift and $\lambda_k$ an approximation to the wanted eigenvalue of (1). Then a Ritz vector of $H_k$ correspond-ing to an eigenvalue of small modulus approximates an eigenvector of the nonlinear problem from which a (hopefully) improved eigenvalue approximation of problem (1) is obtained. Hence, in this approach the two numerical subtasks reducing the large dimension to a much smaller one and solving a nonlinear eigenproblem which are solved separately in the Arnoldi and the Jacobi–Davidson methods in Subsections 3.2 and 3.3 are attacked simultaneously. Following Ruhe's ideas this method was applied in [19] and [18] to the rational eigenvalue problem (6) governing damped vibrations of a structure.

Linearizing the nonlinear family $T(\lambda)$ by Lagrange interpolation between two points $\mu$ and $\sigma$ one gets

$$(37) \qquad T(\lambda) = \frac{\lambda - \mu}{\sigma - \mu}T(\sigma) + \frac{\lambda - \sigma}{\mu - \sigma}T(\mu) + \text{higher order terms.}$$

Keeping $\sigma$ fixed for several steps, iterating on $\mu$, neglecting the remainder in the Lagrange interpolation, and multiplying by $T(\sigma)^{-1}$ from the right one obtains

$$(38) \qquad T(\sigma)^{-1}T(\lambda_{j-1})w = \theta w \quad \text{with } \theta = \frac{\lambda_j - \lambda_{j-1}}{\lambda_j - \sigma}$$

predicting a singularity at

$$(39) \qquad \lambda_j = \lambda_{j-1} + \frac{\theta}{1 - \theta}(\lambda_{j-1} - \sigma).$$

If the dimension $n$ of problem (1) is small then this linear eigenproblem can be used to approximate an eigenvalue of the nonlinear problem, and choosing the smallest eigenvalue of (38) in modulus for every $j$ one can expect convergence to an eigenvalue close to the initial approximation $\lambda_1$.

For large and sparse matrices Ruhe suggested to combine the linearization (38) with a linear Arnoldi process. Assume that the method has performed $j$ steps, yielding approximations $\lambda_1, \ldots, \lambda_j$ to an eigenvalue, orthonormal vectors $v_1, \ldots, v_j$, and an upper Hessenberg matrix $H_{j,j-1} \in \mathbb{C}^{j \times (j-1)}$ such that the Arnoldi recursion

$$(40) \qquad T(\sigma)^{-1}T(\lambda_{j-1})V_{j-1} = V_j H_{j,j-1},$$

is fulfilled (at least approximately), where $V_j = [v_1, \ldots, v_j]$.

Updating the matrix $H_{j,j-1}$ according to the linear theory yields

$$(41) \qquad \tilde{H}_{j+1,j} = \begin{pmatrix} H_{j,j-1} & k_j \\ 0 & \|r_\perp\| \end{pmatrix}$$

where $k_j = V_j^H r_j$, $r_j = T(\lambda_j)v_j$, and $r_\perp = r_j - V_j V_j^H v_j$ which due to the nonlinearity of $T(\cdot)$ violates the next Arnoldi relation

$$T(\sigma)^{-1}T(\lambda_j)V_j = V_{j+1}\tilde{H}_{j+1,j}, \quad v_{j+1} = v_\perp/\|v_\perp\|.$$

To satisfy it at least approximately Ruhe took advantage of Lagrangean interpolation

$$A(\lambda_j) \approx \frac{\lambda_j - \sigma}{\lambda_{j-1} - \sigma}A(\lambda_{j-1}) - \frac{\lambda_j - \lambda_{j-1}}{\lambda_{j-1} - \sigma}I = \frac{1}{1-\theta}A(\lambda_{j-1}) - \frac{\theta}{1-\theta}I,$$

where $A(\lambda) := T(\sigma)^{-1}T(\lambda)$, and updated $H$ according to

$$(42) \qquad H_{j+1,j} = \begin{pmatrix} \frac{1}{1-\theta}H_{j,j-1} - \frac{\theta}{1-\theta}I_{j,j-1} & k_j \\ 0 & \|r_\perp\| \end{pmatrix}$$

arriving at a first version of the rational Krylov method in Algorithm 8.

---

**Algorithm 8** Rational Krylov method; preliminary version

---

1: Start with initial vector $v_1$ with $\|v_1\| = 1$, and initial $\lambda_1$ and $\sigma$
2: $r = T(\sigma)^{-1}T(\lambda_1)v_1$
3: **for** $j = 1, 2, \ldots$ until convergence **do**
4:    orthogonalize $h_j = V^H r$, $r_\perp = r - Vh_j$, $h_{j+1,j} = \|r_\perp\|$
5:    $\theta = \min$ eig $H_{j,j}$ with corresponding eigenvector $s$
6:    $\lambda_{j+1} = \lambda_j + \frac{\theta}{1-\theta}(\lambda_j - \sigma)$
7:    $H_{j+1,j} = \frac{1}{1-\theta}H_{j+1,j} - \frac{\theta}{1-\theta}I_{j+1,j}$
8:    $v_{j+1} = r_\perp/\|r_\perp\|$
9:    $r = T(\sigma)^{-1}T(\lambda_{j+1})v_{j+1}$
10: **end for**

---

Since the method turned out to be inefficient Ruhe [50] suggested to modify $\lambda$, $H$ and $s$ in an inner iteration until the residual $r = T(\sigma)^{-1}T(\lambda)V_j s$ is enforced to be orthogonal to $V_j$, and to expand the search space only after the inner iteration has converged.

If $H_{j,j}$ has already been updated according to step 7: then $H_{j,j}s = 0$, and with

$$k_j = V_j^H T(\sigma)^{-1}T(\lambda)V_j s = V_j^H r$$

we have approximately

$$T(\sigma)^{-1}T(\lambda)V_j \begin{bmatrix} I_{j-1} & \tilde{s} \\ 0 & s_j \end{bmatrix} = V_j[H_{j,j-1}, k_j] + re_j^T$$

where $\tilde{s}$ is the leading $j-1$ vector of $s$. Multiplying by the inverse of the matrix in brackets from the right and by $V_j^H$ from the left one gets the new Hessenberg matrix

$$\hat{H}_{j,j} = [H_{j,j-1}, \, k_j] \begin{bmatrix} I_{j-1} & -s_j^{-1}\tilde{s} \\ 0 & s_j^{-1} \end{bmatrix} = [H_{j,j-1}, \, -s_j^{-1}H_{j,j-1}\tilde{s} + s_j^{-1}k_j],$$

and $H_{j,j-1}\tilde{s} + s_j h_s = 0$ finally yields that the last column of $H_{j,j}$ has to be replaced by $h_j + s_j^{-1}k_j$. Thereafter $\lambda$ and $H$ have to be updated according to steps 5: – 7: of Algorithm 8, and these steps have to be repeated until (hopefully) the residual has become orthogonal to the search space $V_j$. The final version of the rational Krylov method including the inner iteration is contained in Algorithm 9

---

**Algorithm 9** Rational Krylov method; final version

---

1: start with initial vector $V = [v_1]$ with $\|v_1\| = 1$, initial $\lambda$ and $\sigma$ ; set $j = 1$
2: set $h_j = 0_j$; $s = e_j$; $x = v_j$;
3: compute $r = T(\sigma)^{-1}T(\lambda)x$ and $k_j = V_j^H r$
4: **while** $\|k_j\| >$ResTol **do**
5: $\quad$ orthogonalize $r = r - V_j^H k_j$
6: $\quad$ set $h_j = h_j + k_j s_j^{-1}$
7: $\quad$ $\theta = \min$ eig $H_{j,j}$ with corresponding eigenvector $s$
8: $\quad$ $x = V_j s$
9: $\quad$ update $\lambda = \lambda + \frac{\theta}{1-\theta}(\lambda - \sigma)$
10: $\quad$ update $H_{j,j} = \frac{1}{1-\theta}H_{j,j} - \frac{1}{1-\theta}I$
11: $\quad$ compute $r = T(\sigma)^{-1}T(\lambda)x$ and $k_j = V_j^H r$
12: **end while**
13: compute $h_{j+1,j} = \|r\|$
14: **if** $|h_{j+1,j}s_j| >$EigTol **then**
15: $\quad$ $v_{j+1} = r/h_{j+1,j}$; $j = j + 1$; GOTO 2:
16: **end if**
17: Accept eigenvalue $\lambda_i = \lambda$ and eigenvector $x_i = x$
18: If more eigenvalues wanted, choose next $\theta$ and $s$, and GOTO 8:

---

Ruhe motivated the inner iteration and the requirement to make sure that the residual is orthogonal to the search space only by analogy to the linear case where it is satisfied automatically not being aware that the inner iteration is nothing else but a solver of the projected problem

(43) $$V_j^H T(\sigma)^{-1}T(\lambda)V_j s = 0.$$

Hence, although motivated in a completely different way the rational Krylov method is an iterative projection method where the nonlinear eigenproblem $T(\sigma)^{-1}T(\lambda)x = 0$ is projected to a search space $V$, and $V$ is expanded by (the orthogonal complement

**Algorithm 10** Rational Krylov method, an iterative projection method

---
1: start with initial vector $V = [v_1]$ with $\|v_1\| = 1$, initial $\lambda$ and $\sigma$
2: **for** $j = 1, 2, \ldots$ until convergence **do**
3:   solve projected eigenproblem $V^H T(\sigma)^{-1} T(\lambda) V s = 0$ for $(\lambda, s)$
     by inner iteration
4:   compute Ritz vector $x = Vs$ and residual $r = T(\sigma)^{-1} T(\lambda) x$
5:   orthogonalize $r = r - V V^H r$
6:   expand searchspace $V = [V, r/\|r\|]$
7: **end for**

---

of) the residual $r = T(\sigma)^{-1} T(\lambda) V s$ of the Ritz pair (with respect to $V$), and one ends up with Algorithm 10.

The following observations are at hand: the inner iteration in step 3: can be replaced by any dense solver of Section 2, and numerical examples demonstrate [28] that the method can be accelerated considerably this way. But on the other hand, the solvers in Section 2 need the explicit form of the projected problem whereas the inner iteration of Ruhe only needs a procedure that yields the vector $T(\sigma)^{-1} T(\lambda) x$ for a given $x$. A disadvantage of the rational Krylov method is that symmetry properties which the original problem may have are destroyed if the projected problem (43) is considered instead of $V_j^H T(\lambda) V_j s = 0$ in the Arnoldi method or the Jacobi–Davidson algorithm.

# 4    Numerical examples

To test the methods we consider two types of problems, a finite element model of free vibrations of a fluid–solid structure, which is symmetric having a Rayleigh functional such that the projected problems can be solved by safeguarded iteration, and a rational eigenproblem governing damped vibrations of a structure which has non–real eigenvalues.

## 4.1    Vibrations of a fluid–solid structure

We consider a mathematical model which describes the problem governing free vibrations of a tube bundle immersed in a slightly compressible fluid under the following simplifying assumptions: The tubes are assumed to be rigid, assembled in parallel inside the fluid, and elastically mounted in such a way that they can vibrate transversally, but they can not move in the direction perpendicular to their sections. The fluid is assumed to be contained in a cavity which is infinitely long, and each tube is supported by an independent system of springs (which simulates the specific elasticity of each tube). Due to these assumptions, three-dimensional effects are neglected, and so the problem can be studied in any transversal section of the cavity.
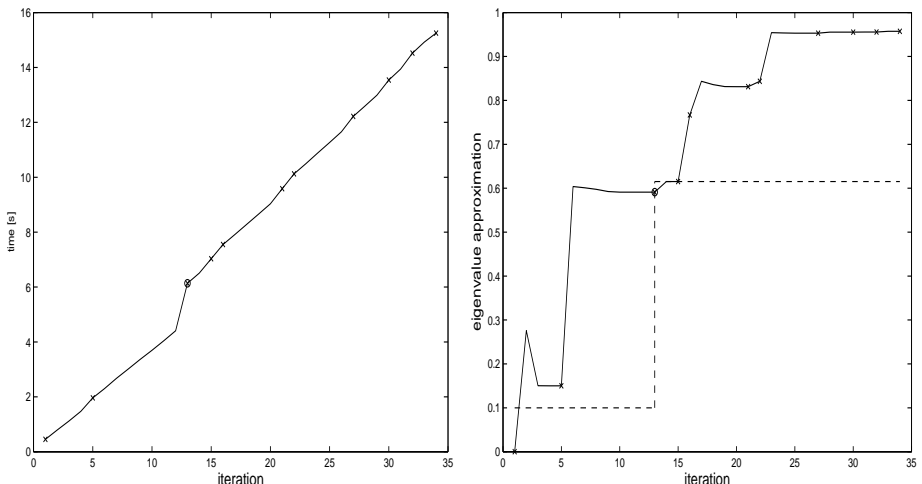
Fig. 1: Time consumption and convergence history for Arnoldi

Considering small vibrations of the fluid (and the tubes) around the state of rest, it can also be assumed that the fluid is irrotational.

Mathematically this problem can be described in the following way (cf. [41], [10]). Let $\Omega \subset \mathbb{R}^2$ (the section of the cavity) be an open bounded set with locally Lipschitz continuous boundary $\Gamma$. We assume that there exists a family $\Omega_j \neq \emptyset$, $j = 1, \ldots, p$, (the sections of the tubes) of simply connected open sets such that $\bar{\Omega}_j \subset \Omega$ for every $j$, $\bar{\Omega}_j \cap \bar{\Omega}_i = \emptyset$ for $j \neq i$, and each $\Omega_j$ has a locally Lipschitz continuous boundary $\Gamma_j$. With these notations we set $\Omega_0 := \Omega \setminus \bigcup_{j=1}^{p} \Omega_j$. Then the boundary of $\Omega_0$ consists of $p + 1$ connected components which are $\Gamma$ and $\Gamma_j$, $j = 1, \ldots, p$.

We denote by $H^1(\Omega_0) = \{u \in L^2(\Omega_0) : \nabla u \in L^2(\Omega_0)^2\}$ the standard Sobolev space equipped with the usual scalar product. Then the eigenfrequencies and the eigenmodes of the fluid-solid structure are governed by the following variational eigenvalue problem (cf. [41], [10])

*Find $\lambda \in \mathbb{R}$ and $u \in H^1(\Omega_0)$ such that for every $v \in H^1(\Omega_0)$*

$$(44) \qquad c^2 \int_{\Omega_0} \nabla u \cdot \nabla v \, dx = \lambda \int_{\Omega_0} uv \, dx + \sum_{j=1}^{p} \frac{\lambda \rho_0}{k_j - \lambda m_j} \int_{\Gamma_j} un \, ds \cdot \int_{\Gamma_j} vn \, ds.$$

Here $u$ is the potential of the velocity of the fluid, $c$ denotes the speed of sound in the fluid, $\rho_0$ is the specific density of the fluid, $k_j$ represents the stiffness constant of the spring system supporting tube $j$, $m_j$ is the mass per unit length of the tube $j$, and $n$ is the outward unit normal on the boundary of $\Omega_0$.

We consider the rational eigenvalue problem (44) where $\Omega$ is the rectangle $(0, 8) \times (0, 4)$, and the sections of the tubes are $(2, 2.2) \times (1, 1.2)$, $(5.8, 6.2) \times (1.8, 2.2)$ and $(5.8, 6) \times (2.8, 3)$. We assume that all constants in problem (44) are equal to 1.

Discretizing problem (44) by linear Lagrangean elements one gets a rational matrix eigenvalue problem

$$(45) \qquad\qquad T(\lambda)x := -Ax + \lambda Bx + \frac{\lambda}{1-\lambda}Cx = 0$$

where $C$ collects the contributions of all tubes. $A$, $B$, and $C$ are symmetric matrices, $A$ and $C$ are positive semidefinite, and $B$ is positive definite. In our example the dimension is $n = 22654$.

Problem (45) has 11 eigenvalues $\lambda_1 \leq \cdots \leq \lambda_{11}$ in the interval $J_1 = (0,1)$ (cf. [35]), and a large number of eigenvalues greater than 1, 10 of which are contained in the interval $(1,4)$.

We determined approximations to the eigenvalues in $[0,1)$ by the Arnoldi method (Algorithm 6), the Jacobi–Davidson method (Algorithm7), where in both cases the projected nonlinear eigenproblems where solved by safeguarded iteration, and by the rational Krylov method (Algorithm 10) where the projected rational eigenproblems were solved linearizing the equivalent quadratic eigenproblem $(1-\lambda)V^T T(\lambda)Vy = 0$. All three methods were able to find all 11 eigenvalues. The original rational Krylov method (Algorithm 9) as implemented by Jarlebring [27] turned out to depend very sensitively on the initial pole $\sigma$ and the initial approximation to an eigenvalue, and were able to find at most 8 eigenvalues in the interval $[0,1)$.

The experiments were run under MATLAB 6.5 on an Intel Centrino M processor with 1.7 GHz and 1 GB RAM. Figures 1 to 3 show the time consumption and the convergence history of the three methods where in every case the initial pole was chosen to be $\sigma = 0.1$, and the iteration was terminated if the residual was less than $10^{-6}$. In all plots plus signs indicate found eigenvalues, and circles mark changes of the pole $\sigma$. The dashed lines indicate the pole in use.

Table 1 summarizes the properties of the iterative projection methods under consideration as applied to the symmetric nonlinear eigenproblem governing the free vibrations of the fluid–solid structure.

| Method | Iter. | LU fact. | CPU [s] | nlin.sol. [s] |
|---|---|---|---|---|
| Arnoldi | 34 | 2 | 14.93 | 0.13 |
| Jacobi–Davidson | 37 | 3 | 112.84 | 0.15 |
| rational Krylov | 40 | 2 | 70.80 | 0.22 |
| intervall $(1,4)$ | | | | |
| Arnoldi | 36 | 2 | 17.35 | 0.17 |
| Jacobi–Davidson | 37 | 5 | 125.87 | 0.23 |

Tabel 1: Fluid–solid structure

Enumerating the eigenvalues according to Subsection 2.5 the smallest eigenvalue in the interval $(1,4)$ turns out to be a 6:th eigenvalue. Starting the Arnoldi method
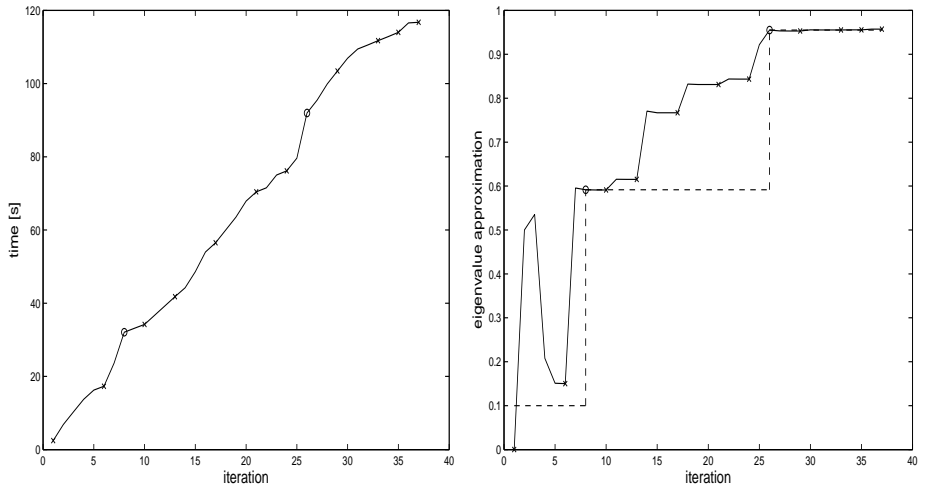
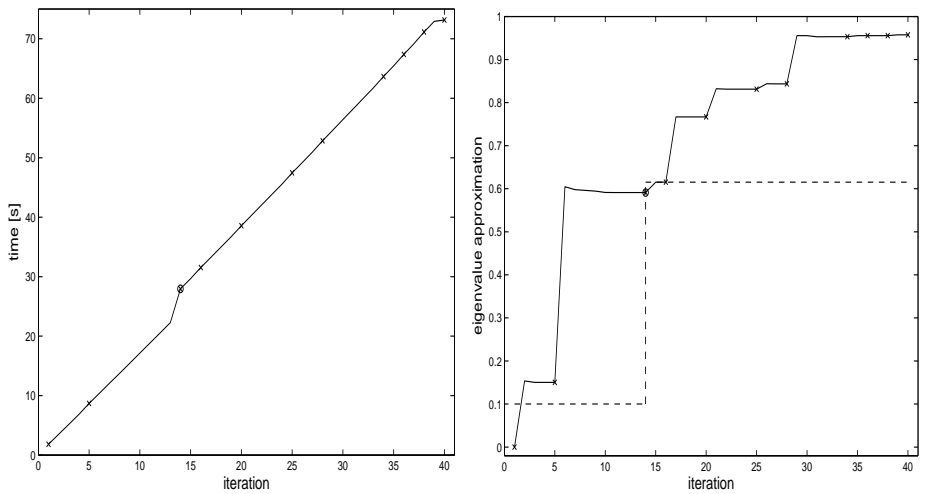Fig. 2: Time consumption and convergence history for Jacobi–Davidson



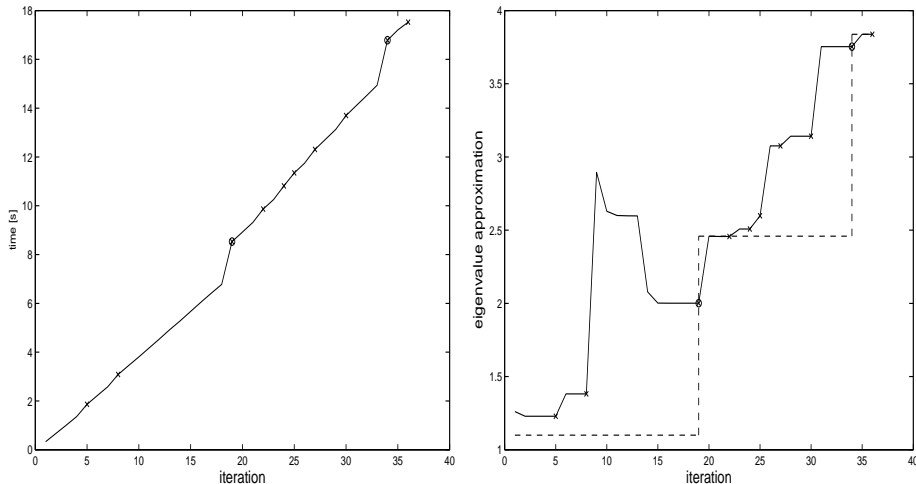Fig. 3: Time consumption and convergence history for rational Krylov

Fig. 4: Time consumption and convergence history for Arnoldi; Interval $(1, 4)$

and the Jacobi–Davidson method with the invariant subspace of the linear eigenproblem (cf. [62])

$$\left(A + \frac{\lambda}{\lambda - 1}C\right)x = \mu Bx, \quad \lambda = 1 + \varepsilon, \ \varepsilon > 0,$$

corresponding to the 6 largest eigenvalues, both methods were able to find all eigenvalues of problem (44) in the interval $(1, 4)$. The time consumption and the convergence histories are contained in Figures 4 and 5. The rational Krylov method destroys the symmetry of the problem, and the enumeration of Subsection 2.5 does not apply. Neither the implementation of Jarlebring nor the modification in Algorithm 10 was able to find more than one or two eigenvalues in the interval $(1, 4)$.

## 4.2  Damped vibrations of a structure

As a second example we consider the the free vibrations of a solid

$$\Omega := \{(x, y, z) \ : \ x^2/9 + y^2/4 + z^2 \leq 1, \ x \geq 0, \ z \geq 0\},$$

which is fixed at the boundary surfaces $\{(x, y, z) \in \Omega \ : \ x = 0, \ z = 0\}$. Assuming a density $\rho = 7800$, a Young's modulus $E = 2.1 * 10^{11}$ and a Poisson rate $\nu = 0.3$, and including nonproportional damping using the constitutive law of a standard linear viscoelastic solid with parameters $\Delta\nu = 0.27$, $\Delta E = 0.5 * 10^{11}$ and a relaxation constant $b = 10^{-4}$, the finite element method with linear Lagrangean elements generated a rational eigenproblem

$$(46) \qquad\qquad T(\omega) := \left(\omega^2 M + K - \frac{1}{1 + b\omega}\Delta K\right)x = 0$$
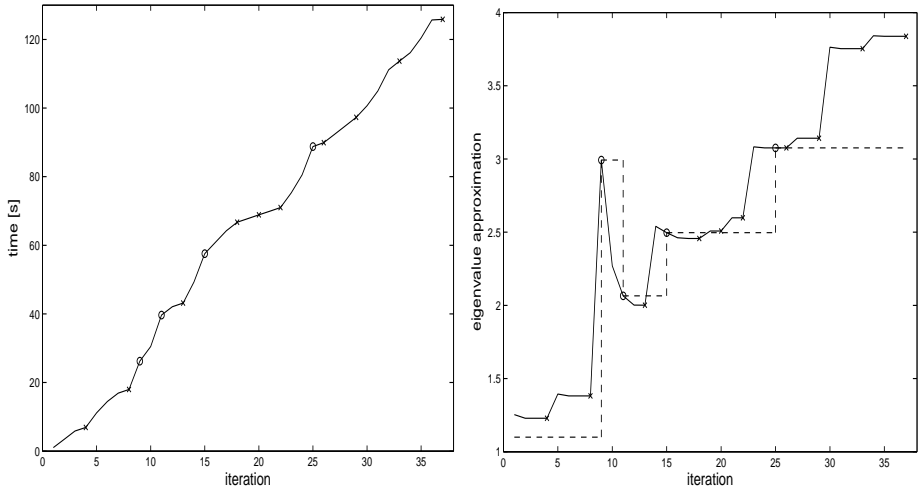
34

Fig. 5: Time consumpt. and conv. history for Jacobi–Davidson; Interval $(1, 4)$

of dimension 10704.

For symmetry reasons we determined only eigenvalues with negative imaginary part, and we computed 30 of them one after another with decreasing imaginary part. The nonlinear projected eigenproblems (after multiplying by $1 + b\omega$) were solved by linearization, and the iteration was terminated if the norm of the residual was less than $10^{-6}$.

The Arnoldi method without restarts needed 144 iteration steps, and a CPU time of 707.0 seconds to find all 30 eigenvalues with maximal negative imaginary part (i.e. the average number of iteration steps to determine an eigenvalue is less than 5). With a tolerance of $\text{tol} = 2 * 10^{-1}$ in step 13: of Algorithm 6 no update of the preconditioner was necessary.

The dominant share of the CPU time, namely 469.9 seconds was consumed by the solver of the projected nonlinear eigenproblems. Figure 6 displays the development of the time consumption of the entire iteration and the share of the nonlinear eigensolver. It demonstrates the necessity of restarts since the superlinear time consumption is mainly caused by the eigensolver.

We restarted the Arnoldi process if the dimension of the search space exceeded 50 with an orthogonal basis of the space spanned by the already determined eigenvectors. The method needed 3 restarts, and again all 30 eigenvalues with maximal negative imaginary part were found by the Arnoldi method requiring 139 iterations, and 4 updates of the preconditioner. The total CPU time was 199.6 seconds. Solving the projected eigenproblems and updating the preconditioners required 25.0 and 30.8 seconds, respectively. Figure 7 demonstrates the time consumption for this experiment.

Obviously, immediately after a restart the speed of convergence is slowed down. On the other hand this delay yields an update of the preconditioner accelerating the
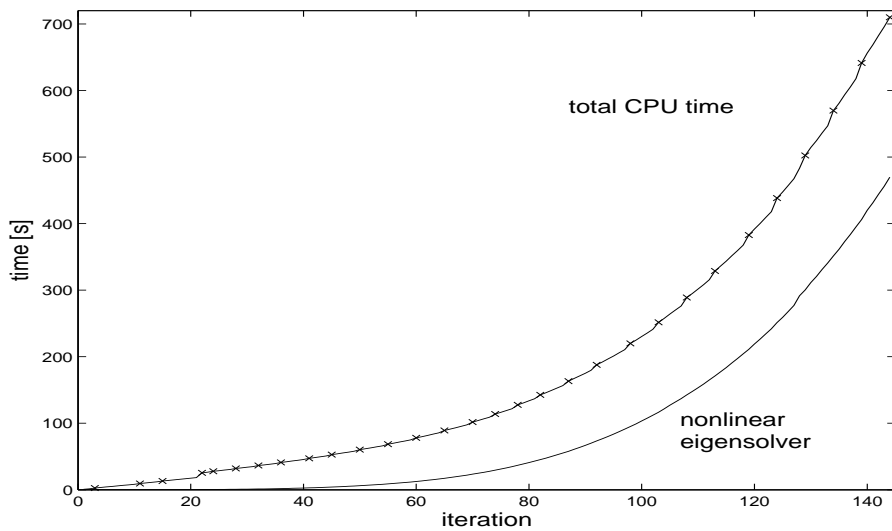
35
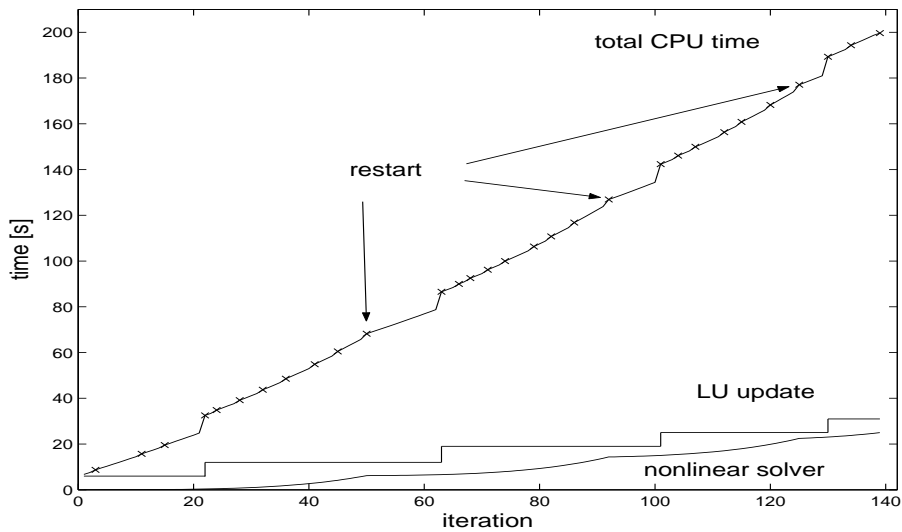
Fig. 6. Arnoldi method without restarts



Fig. 7. Arnoldi method with restarts

36

convergence, such that the total number of iteration steps is reduced from 144 to 139.

The Jacobi–Davidson method and the rational Krylov method according to Algorithm 10 are much slower than the Arnoldi method, but they show similar behaviour: without restart a substantial share of the total CPU time is consumed by the solver of the projected nonlinear eigenproblems, and both methods can be accelerated by restarts. Details about these methods are contained in Table 2. The rational Krylov method with inner iteration converges for this problem, although it is very slow. To determine the 30 wanted eigenvalues 366 iterations are necessary requiring 2248.5 seconds.

| Method | Iter. | LU fact. | CPU [s] | nlin.sol. [s] |
|---|---|---|---|---|
| Arnoldi | 144 | 2 | 707.0 | 469.9 |
| Arnoldi, restarted | 139 | 5 | 199.6 | 25.0 |
| Jacobi–Davidson | 111 | 9 | 1050.5 | 161.2 |
| Jacobi–Davidson, restarted | 109 | 12 | 914.4 | 18.9 |
| rational Krylov | 147 | 3 | 1107.1 | 465.3 |
| rational Krylov, restarted | 147 | 4 | 647.8 | 28.5 |

Tabel 2: Damped vibrations of a structure

# References

[1] T. Apel, V. Mehrmann, and D. Watkins. Structured eigenvalue methods for the computation of corner singularities in 3D anisotropic elastic structures. *Comput. Meth. Appl. Mech. Engrg.*, 191:4459 – 4473, 2002.

[2] T. Apel, V. Mehrmann, and D. Watkins. Numerical solution of large scale structured polynomial and rational eigenvalue problems. Technical report, Fakultät für Mathematik, TU Chemnitz, 2003.

[3] Z. Bai. Second order Krylov subspace and its application for solving the quadratic eigenvalue problem, 2003. Lecture presented at the Dagstuhl Seminar on Theoretical and Computational Aspects of Matrix Algorithms.

[4] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H.A. van der Vorst, editors. *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide.* SIAM, Philadelphia, 2000.

[5] O. A. Bauchau. A solution of the eigenproblem for undamped gyroscopic systems with the Lanczos algorithm. *Internat. J. Numer. Math. Engrg.*, 23:1705–1713, 1986.

[6] T. Betcke and H. Voss. A Jacobi–Davidson–type projection method for nonlinear eigenvalue problems. Technical Report 47, Section of Mathematics, Hamburg University of Technology, 2002. To appear in Future Generation Computer Systems.

[7] M. Borri and P. Mantegazza. Efficient solution of quadratic eigenproblems arising in dynamic analysis of structures. *Comput. Meth. Appl. Mech. Engrg.*, 12:19–31, 1977.

[8] J. Brandts. The Riccati method for eigenvalues and invariant subspaces of matrices with inexpensive action. *Lin. Alg. Appl.*, 358:333–363, 2003.

[9] J. Brandts. Solving eigenproblems: from Arnoldi via Jacobi-Davidson to the Riccati method. *Lecture Notes in Computer Science*, 2542:167–173, 2003.

[10] C. Conca, J. Planchard, and M. Vanninathan. Existence and location of eigenvalues for fluid-solid structures. *Comput.Meth.Appl.Mech.Engrg.*, 77:253 – 291, 1989.

[11] R.J. Duffin. A minmax theory for overdamped networks. *J.Rat.Mech.Anal.*, 4:221 – 233, 1955.

[12] R.J. Duffin. The Rayleigh–Ritz method for dissipative and gyroscopic systems. *Quart.Appl.Math.*, 18:215 – 221, 1960.

[13] D.R. Fokkema, G.L.G. Sleijpen, and H.A. Van der Vorst. Jacobi-Davidson style QR and QZ algorithms for the partial reduction of matrix pencils. *SIAM J.Sci.Comput.*, 20:94 – 125, 1998.

[14] W. Gander, G.H. Golub, and U. von Matt. A constrained eigenvalue problem. *Lin.Alg.Appl.*, 114–115:815 – 839, 1989.

[15] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. Academic Press, New York, 1982.

[16] K. P. Hadeler. Variationsprinzipien bei nichtlinearen Eigenwertaufgaben. *Arch.Rat.Mech.Anal.*, 30:297 – 307, 1968.

[17] K. P. Hadeler. Nonlinear eigenvalue problems. In R. Ansorge, L. Collatz, G. Hämmerlin, and W. Törnig, editors, *Numerische Behandlung von Differentialgleichungen*, ISNM 27, pages 111–129. Birkhäuser, Stuttgart, 1975.

[18] P. Hager. *Eigenfrequency Analysis. FE-Adaptivity and a Nonlinear Eigenvalue Problem*. PhD thesis, Chalmers University of Technology, Göteborg, 2001.

[19] P. Hager and N.E. Wiberg. The rational Krylov algorithm for nonlinear eigenvalue problems. In B.H.V. Topping, editor, *Computational Mechanics for the Twenty-First Century*, pages 379 – 402. Saxe–Coburg Publications, Edinburgh, 2000.

[20] L. Hoffnung, R.-C. Li, and Q. Ye. Krylov type subspace methods for matrix polynomials. Technical Report 2002-08, Department of Mathematics, University of Kentucky, 2002.

[21] U.B. Holz. *Subspace Approximation Methods for Perturbed Quadratic Eigenvalue Problems*. PhD thesis, Department of Mathematics, Stanford University, 2002.

[22] H.Y. Hu, E.H. Dowell, and L.N. Virgin. Stability estimation of high dimensional vibrating systems under state delay feedback control. *J. Sound and Vibration*, 214:497 – 511, 1998.

[23] H.Y. Hu and Z.H. Wang. *Dynamics of Controlled Mechanical Systems with Delayed Feedback*. Springer, Berlin, 2002.

[24] T.-M. Hwang, W.-W. Lin, and V. Mehrmann. Numerical solution of quadratic eigenvalue nproblems with structure preserving methods. *SIAM J. Sci. Comput.*, 24:1283 – 1302, 2003.

[25] N. K. Jain and K. Singhal. On Kublanovskaya's approach to the solution of the generalized latent value problem for functional $\lambda$–matrices. *SIAM J.Numer.Anal.*, 20:1062–1070, 1983.

[26] N. K. Jain, K. Singhal, and K. Huseyin. On roots of functional lambda matrices. *Comput.Meth.Appl.Mech.Engrg.*, 40:277–292, 1983.

[27] E. Jarlebring. Krylov methods for nonlinear eigenvalue problems. Master's thesis, Royal Institute of Technology, Dept.Numer.Anal.Comput.Sci., Stockholm, 2003.

[28] E. Jarlebring and H. Voss. Rational Krylov for nonlinear eigenproblems, an iterative projection method. Technical Report 69, Section of Mathematics, Hamburg University of Technology, 2003. Submitted to Appl.Math.

[29] V.N. Kublanovskaya. On an application of Newton's method to the determination of eigenvalues of $\lambda$-matrices. *Dokl. Akad. Nauk. SSR*, 188:1240 – 1241, 1969.

[30] V.N. Kublanovskaya. On an approach to the solution of the generalized latent value problem for $\lambda$-matrices. *SIAM. J. Numer. Anal.*, 7:532 – 537, 1970.

[31] P. Lancaster. Strongly stable gyroscopic systems. *Electr.J.Lin.Alg.*, 5:53 – 66, 1999.

[32] P. Lancaster. *Lambda–matrices and Vibrating Systems.* Dover Publications, Mineola, New York, 2002.

[33] R.B. Lehoucq, D.C. Sorensen, and C. Yang. *ARPACK Users' Guide. Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods.* SIAM, Philadelphia, 1998.

[34] R.-C. Li and Q. Ye. A Krylov subspace method for quadratic matrix polynomials with application to constrained least squares problems. Technical report, Department of Mathematics, University of Kentucky, 2002.

[35] L. Mazurenko and H. Voss. On the number of eigenvalues of a rational eigenproblem. Technical Report 61, Section of Mathematics, Hamburg University of Technology, 2003. Submitted to SIAM J.Numer.Anal.

[36] K. Meerbergen. Locking and restarting quadratic eigenvalue solvers. *SIAM J. Sci. Comput.*, 22:1814 – 1839, 2001.

[37] V. Mehrmann and D. Watkins. Structure-preserving methods for computing eigenpairs of large sparse skew-Hamiltonian/Hamiltonian pencils. *SIAM J. Sci. Comput.*, 22:1905 – 1925, 2001.

[38] V. Mehrmann and D. Watkins. Polynomial eigenvalue problems with Hamiltonian structure. *Electronic Transactions on Numerical Analysis*, 13:106 – 118, 2002.

[39] A. Neumaier. Residual inverse iteration for the nonlinear eigenvalue problem. *SIAM J. Numer. Anal.*, 22:914 – 923, 1985.

[40] N. Petersmann. *Substrukturtechnik und Kondensation bei der Schwingungsanalyse*, volume 76 of *Fortschrittberichte VDI, Reihe 11: Schwingungstechnik.* VDI Verlag, Düsseldorf, 1986.

[41] J. Planchard. Eigenfrequencies of a tube bundle placed in a confined fluid. *Comput.Meth.Appl.Mech.Engrg.*, 30:75 – 93, 1982.

[42] J.S. Przemieniecki. *Theory of Matrix Structural Analysis.* McGraw–Hill, New York, 1968.

[43] E. H. Rogers. A minimax theory for overdamped systems. *Arch.Rat.Mech.Anal.*, 16:89–96, 1964.

[44] K. Rothe. *Lösungsverfahren für nichtlineare Matrixeigenwertaufgaben mit Anwendungen auf die Ausgleichselementmethode.* Verlag an der Lottbek, Ammersbek, 1989.

[45] K. Rothe. Least squares element method for boundary eigenvalue problems. *Internat.J.Numer.Meth.Engrg.*, 33:2129–2143, 1992.

[46] A. Ruhe. Algorithms for the nonlinear eigenvalue problem. *SIAM J. Numer. Anal.*, 10:674 – 689, 1973.

[47] A. Ruhe. Computing nonlinear eigenvalues with spectral transformation Arnoldi. *ZAMM*, 76:S2: 17 – 20, 1996.

[48] A. Ruhe. Rational Krylov for nonlinear eigenvalues, 1996. Talk presented at the Householder XIII Conference at Pontresina, Switzerland.

[49] A. Ruhe. Rational Krylov, a practical algorithm for large sparse nonsymmetric matrix pencils. *SIAM J. Sci. Comp.*, 19(5):1535–1551, 1998.

[50] A. Ruhe. A rational Krylov algorithm for nonlinear matrix eigenvalue problems. *Zapiski Nauchnyh Seminarov POMI*, 268:176 – 180, 2000.

[51] G.L. Sleijpen, G.L. Booten, D.R. Fokkema, and H.A. van der Vorst. Jacobi-Davidson type methods for generalized eigenproblems and polynomial eigenproblems. *BIT*, 36:595 – 633, 1996.

[52] G.L. Sleijpen and H.A. van der Vorst. The Jacobi–Davidson method for eigenproblems and its relation with accelerated inexact newton schemes. In S.D. Margenov and P.S. Vassilevski, editors, *Iterative Methods in Linear Algebra*, volume 3 of *IMACS Series in Computational and Applied Mathematics*, pages 377–389, New Brunswick, 1996. IMACS.

[53] G.L. Sleijpen and H.A. van der Vorst. A Jacobi-Davidson iteration method for linear eigenvalue problems. *SIAM J.Matr.Anal.Appl.*, 17:401 – 425, 1996.

[54] S. I. Solov'ëv. Eigenvibrations of a plate with elastically attached loads. Preprint SFB393/03-06, Sonderforschungsbereich 393 an der Technischen Universität Chemnitz, Technische Universität, D-09107 Chemnitz, Germany, 2003. Available at http://www.tu-chemnitz.de/sfb393/Files/PS/sfb03-06.ps.gz.

[55] F. Tisseur and K. Meerbergen. The quadratic eigenvalue problem. *SIAM Review*, 43:235 – 286, 2001.

[56] H. Voss. An error bound for eigenvalue analysis by nodal condensation. In J. Albrecht, L. Collatz, and W. Velte, editors, *Numerical Treatment of Eigenvalue Problems, Vol. 3*, volume 69 of *International Series on Numerical Mathematics*, pages 205–214, Basel, 1984. Birkhäuser.

[57] H. Voss. A new justification of finite dynamic element methods. In J. Albrecht, L. Collatz, W. Velte, and W. Wunderlich, editors, *Numerical Treatment of Eigenvalue Problems, Vol. 4*, ISNM 83, pages 232 – 242, Basel, 1987. Birkhäuser.

[58] H. Voss. Free vibration analysis by finite dynamic element methods. In I. Marek, editor, *Proceedings of the Second International Symposium on Numerical Analysis*, Teubner Texte zur Mathematik 107, pages 295 – 298, Leipzig, 1988. Teubner.

[59] H. Voss. An Arnoldi method for nonlinear eigenvalue problems. Technical Report 56, Section of Mathematics, Hamburg University of Technology, 2002. Submitted to BIT Numerical Mathematics.

[60] H. Voss. An Arnoldi method for nonlinear symmetric eigenvalue problems. In *Online Proceedings of the SIAM Conference on Applied Linear Algebra, Williamsburg*, http://www.siam.org/meetings/laa03/, 2003.

[61] H. Voss. An Arnoldi type projection method for nonlinear symmetric eigenvalue problems. *Proc.Appl.Math.Mech*, 3:555 – 556, 2003.

[62] H. Voss. Initializing iterative projection methods for rational symmetric eigenproblems. In *Online Proceedings of the Dagstuhl Seminar Theoretical and Computational Aspects of Matrix Algorithms, Schloss Dagstuhl 2003*, ftp://ftp.dagstuhl.de/pub/Proceedings/03/03421/03421.VoszHeinrich.Other.pdf, 2003.

[63] H. Voss. A Jacobi–Davidson method for nonlinear eigenproblems. Technical Report 66, Section of Mathematics, Hamburg University of Technology, 2003. Submitted to Proceedings of ICCS'2004, Krakow,Poland.

[64] H. Voss. A rational spectral problem in fluid–solid vibration. *Electronic Transactions on Numerical Analysis*, 16:94 – 106, 2003.

[65] H. Voss. Variational characterization of eigenvalues of nonlinear eigenproblems. In M. Kocandrlova and V. Kelar, editors, *Proceedings of the International Conference on Mathematical and Computer Modelling in Science and Engineering*, pages 379 – 383, Prague, 2003. Czech Technical University.

[66] H. Voss and B. Werner. A minimax principle for nonlinear eigenvalue problems with applications to nonoverdamped systems. *Math.Meth.Appl.Sci.*, 4:415–424, 1982.

[67] H. Voss and B. Werner. Solving sparse nonlinear eigenvalue problems. Technical Report 82/4, Inst. f. Angew. Mathematik, Universität Hamburg, 1982.

[68] B. Werner. *Das Spektrum von Operatorenscharen mit verallgemeinerten Rayleighquotienten*. PhD thesis, Fachbereich Mathematik, Universität Hamburg, 1970.

[69] W.H. Yang. A method for eigenvalues of sparse λ-matrices. *Internat. J. Numer. Meth. Engrg.*, 19:943 – 948, 1983.