Electronic Transactions on Numerical Analysis.
Volume 55, pp. 1–75, 2022.

ISSN 1068–9613.
DOI: 10.1553/etna_vol55s1

ETNA
Kent State University and
Johann Radon Institute (RICAM)

A SURVEY ON VARIATIONAL CHARACTERIZATIONS FOR NONLINEAR EIGENVALUE PROBLEMS*

JÖRG LAMPE† AND HEINRICH VOSS‡

Abstract. Variational principles are very powerful tools when studying self-adjoint linear operators on a Hilbert space $\mathcal{H}$. Bounds for eigenvalues, comparison theorems, interlacing results, and monotonicity of eigenvalues can be proved easily with these characterizations, to name just a few. In this paper we consider generalizations of these principles to families of linear, self-adjoint operators depending continuously on a scalar in a real interval.

Key words. nonlinear eigenvalue problem, variational characterization, iterative projection methods, AMLS, quantum dots, viscoelastic damping, total least-squares problems, fluid-solid interaction

AMS subject classifications. 35P30, 47A52, 47A75, 47J10, 65F15, 65F17

Contents.

1. Introduction. This paper considers the nonlinear eigenvalue problem of finding a parameter λ such that the linear system

$$T(\lambda)x = 0$$

has a nontrivial solution $x \neq 0$, where $T(\cdot)$ is a family of linear bounded operators on a real Hilbert space $\mathcal{H}$. It generalizes the linear eigenvalue problem $Ax = \lambda x$ and the generalized linear eigenvalue problem $Ax = \lambda Bx$, where A and B are linear operators on $\mathcal{H}$.

Nonlinear eigenvalue problems arise in a variety of applications in science and engineering, such as the dynamic analysis of structures, vibrations of fluid-solid structures, the electronic behavior of quantum dots, and viscoelastic oscillators, to name just a few. Due to its wide range of applications, the quadratic eigenvalue problem $T(\lambda)x = \lambda^2 Mx + \lambda Cx + Kx = 0$ is of particular interest [153, 225], but also polynomial [162, 163, 173], rational [217], and

*Received September 28, 2020. Accepted September 6, 2021. Published online on November 3, 2021. Recommended by Yousef Saad.

†Institute of Electrical Engineering, System Theory and Mathematics, University of Applied Sciences Cologne, Germany (joerg.lampe@rfh-koeln.de).

‡Institute of Mathematics, Hamburg University of Technology, D-21071 Hamburg, Germany (voss@tu-harburg.de).

ETNA
Kent State University and
Johann Radon Institute (RICAM)

more general eigenvalue problems appear. A standard approach for investigating or numerically solving polynomial eigenvalue problems is linearization, where the original problem is transformed into a generalized linear eigenvalue problem with the same spectrum. Details on linearization and structure preservation are discussed in [81, 106, 137, 156]. Likewise, a rational eigenvalue problem can be turned into a polynomial one by multiplying it with an appropriate scalar polynomial in λ. Notice, however, that important structural properties like symmetry and variational characterizations of eigenvalues may get lost. Moreover, the degree of the polynomial can become very large, and roots of the denominator produce spurious eigenvalues that may hamper the numerical solution. We do not consider polynomial or rational eigenvalue problems here.

In this paper we consider self-adjoint nonlinear eigenvalue problems that allow for a variational characterization of its eigenvalues generalizing the well-known minmax characterization of Poincaré [191] or Courant [50], Fischer [68], and Weyl [256] for linear eigenvalue problems. Recent surveys on general nonlinear eigenvalue are contained in [89] and [245].

Variational characterizations are very powerful tools when studying self-adjoint linear operators on a Hilbert space $\mathcal{H}$. Bounds for eigenvalues, comparison theorems, interlacing results, and monotonicity of eigenvalues can be proved easily with these characterizations, to name just a few.

The paper is organized as follows. Section 2 summaries generalizations of the variational characterization of eigenvalues for symmetric nonlinear eigenproblems. Section 3 presents various numerical methods for dense eigenvalue problems, and Section 4 discusses iterative projection methods for large sparse problems. Most of the methods in these two sections also apply to non-symmetric problems.

The following three sections are concerned with the localization of eigenvalues for dense problems. Hyperbolic matrix polynomials allow for a definite linearization and can therefore be solved by standard algorithms like the QR algorithm. In Section 5 we present a method for detecting whether a given matrix polynomial is hyperbolic or not. Section 6 generalizes Sylvester's law of inertia to symmetric nonlinear eigenvalue problems, which combined with the bisection method yields an easy way to locate eigenvalues on the real axis. In Section 7 we consider nonlinear low-rank modifications of symmetric eigenvalue problems.

The Automated Multi-Level Substructuring (AMLS) method was introduced by Bennighof [27] to reduce large symmetric eigenvalue problems to much smaller ones. It consists of a combination of elimination of variables and curtailment of the system in several steps. Section 8 contains an error bound for AMLS.

Regularization for large-scale problems by orthogonal projection for total least-squares problems based on symmetric eigenproblems and for the dual regularization of total least-squares problems are presented in Sections 9, 10, and 11.

The following Sections 12 and 13 take advantage of variational characterizations of eigenvalues to examine the electronic behavior of quantum dots and to study viscoelastic damping. Section 14 is dedicated to modeling vibrations of fluid-solid structures. Although this is not a symmetric eigenproblem, its eigenvalues allow for a variational characterization. The paper closes with conclusions.

2. Variational characterizations of eigenvalues for nonlinear eigenproblems. This section contains a summary on variational characterizations of eigenvalues of nonlinear eigenvalue problems generalizing the well-known minmax characterization of Poincaré [191] or Courant [50], Fischer [68], and Weyl [256] for linear eigenvalue problems. Variational characterizations are highly useful instruments for investigating self-adjoint linear operators on a Hilbert space $\mathcal{H}$. For instance, they can be employed to obtain bounds for eigenvalues and comparison theorems yielding results on the interlacing and the monotonicity of eigenvalues;

ETNA
Kent State University and
Johann Radon Institute (RICAM)

see Sections 5, 6, and 7 for detailed results on the location of eigenvalues. The other sections of this paper are all about particular applications, where variational characterization plays a crucial role.

We consider the nonlinear eigenvalue problem

$$T(\lambda)x = 0, \tag{2.1}$$

where $T(\lambda) : \mathcal{H} \to \mathcal{H}$, $\lambda \in J$, is a family of self-adjoint and bounded operators depending continuously on the parameter λ and $J \subset \mathbb{R}$ is an open real interval that may be unbounded. We stress the fact that in this section we are only concerned with real eigenvalues in J although $T(\cdot)$ may be defined on a larger subset of $\mathbb{C}$ and $T(\cdot)$ may have additional eigenvalues in $\mathbb{C} \setminus J$.

To generalize the variational characterization of eigenvalues, we first need a generalization of the Rayleigh quotient. To this end we assume that

(A_1) *for every fixed $x \in \mathcal{H}$, $x \neq 0$, the real equation*

$$f(\lambda; x) := \langle T(\lambda)x, x\rangle = 0$$

has at most one solution $\lambda =: p(x) \in J$,

with the inner product $\langle \cdot, \cdot \rangle$ as scalar product. This defines a *Rayleigh functional* p of (2.1) with respect to J, and we denote by $\mathcal{D}(p) \subset \mathcal{H}$ the domain of definition of p.

Generalizing the definiteness requirement for a linear pencil $T(\lambda) = \lambda B - A$, we further assume that $\lambda \mapsto \langle T(\lambda)x, x\rangle$ is increasing at the point $p(x)$, i.e.,

(A_2) *for every $x \in \mathcal{D}(p)$ and every $\lambda \in J$ with $\lambda \neq p(x)$, it holds that*

$$(\lambda - p(x))f(\lambda; x) > 0.$$

If p is defined on $\mathcal{D} = \mathcal{H} \setminus \{0\}$, then the problem $T(\lambda)x = 0$ is called *overdamped*. This notation is motivated by the finite-dimensional quadratic eigenvalue problem

$$T(\lambda)x = \lambda^2 Mx + \lambda Cx + Kx = 0,$$

where M, C, and K are Hermitian and positive definite matrices. If C is large enough such that $d(x) := (x^H Cx)^2 - 4(x^H Kx)(x^H Mx) > 0$ for every $x \neq 0$, then $T(\cdot)$ is overdamped. Generalizations of the minmax and maxmin characterizations of eigenvalues were proved by Duffin [56] for the quadratic case and by Rogers [196] for general overdamped problems. Infinite-dimensional overdamped eigenvalue problems were studied by Turner [226], Langer [139], and Weinberger [254], who proved generalizations of both the maxmin characterization of Poincaré and the minmax characterization of Courant, Fischer, and Weyl for quadratic (and by Turner [227] for polynomial) overdamped problems. The corresponding results on generalizations for general overdamped problems of infinite dimension were derived by Hadeler [91]. Similar results (weakening the compactness or smoothness requirements) are contained in the works of Rogers [197], Werner [255], Abramov [1], Hadeler [92], Markus [168], Maksudov and Gasanov [165], and Hasanov [100].

The key to the variational principle in the nonoverdamped case is an appropriate enumeration of the eigenvalues. In general, the natural enumeration, i.e., the first eigenvalue is the smallest one, followed by the second smallest one, etc. is not reasonable. Instead, the number of an eigenvalue λ of the nonlinear problem (2.1) is inherited from the location of the eigenvalue 0 in the spectrum of the operator $T(\lambda)$ based on the following consideration; cf. [248].

ETNA
Kent State University and
Johann Radon Institute (RICAM)

For $j \in \mathbb{N}$ and $\lambda \in J$, let

$$\mu_j(\lambda) := \sup_{V \in S_j} \min_{v \in V, v \neq 0} \frac{\langle T(\lambda)v, v\rangle}{\langle v, v\rangle},$$

where S_j is the set of all j-dimensional subspaces of $\mathcal{H}$. We assume that

(A_3) *if $\mu_n(\lambda) = 0$ for some $n \in \mathbb{N}$ and some $\lambda \in J$, then, for $j = 1, \ldots, n$, the supremum in $\mu_j(\lambda)$ is attained and $\mu_1(\lambda) \geq \mu_2(\lambda) \geq \cdots \geq \mu_n(\lambda)$ are the n largest eigenvalues of the linear operator $T(\lambda)$. Conversely, if $\mu = 0$ is an eigenvalue of the operator $T(\lambda)$, then $\mu_n(\lambda) = 0$ for some $n \in \mathbb{N}$.*

DEFINITION 2.1. *$\lambda \in J$ is an nth eigenvalue of $T(\cdot)$ if $\mu_n(\lambda) = 0$ for $n \in \mathbb{N}$.*

Condition (A_3) is satisfied for example if for every $\lambda \in J$ the supremum of the essential spectrum of $T(\lambda)$ is less than 0. The following stronger condition that for every $\lambda \in J$ there exists $\nu(\lambda) > 0$ such that $T(\lambda) + \nu(\lambda)I$ is a compact operator was used in Hadeler [91]. With this enumeration the following minmax characterization for eigenvalues was proved in [243, 248].

THEOREM 2.2 ([243, 248]). *Let J be an open interval in $\mathbb{R}$, and let $T(\lambda) : \mathcal{H} \to \mathcal{H}$, $\lambda \in J$, be a family of self-adjoint and bounded operators depending continuously on the parameter λ such that the conditions (A_1), (A_2), and (A_3) are satisfied. Then the following statements hold:*

(i) For every $\ell \in \mathbb{N}$, there is at most one ℓth eigenvalue of $T(\cdot)$, which can be characterized by

$$\lambda_\ell = \min_{\substack{V \in S_\ell, \\ V \cap \mathcal{D}(p) \neq \emptyset}} \sup_{v \in V \cap \mathcal{D}(p)} p(v). \tag{2.2}$$

(ii) If

$$\lambda_\ell := \inf_{\substack{V \in S_\ell, \\ V \cap \mathcal{D}(p) \neq \emptyset}} \sup_{v \in V \cap \mathcal{D}(p)} p(v) \in J$$

for some $\ell \in \mathbb{N}$, then λ_ℓ is the ℓth eigenvalue of $T(\cdot)$ in J, and (2.2) holds.

(iii) If there exist the kth and the ℓth eigenvalue λ_k and λ_ℓ in J ($k < \ell$), then J contains the jth eigenvalue λ_j ($k \leq j \leq \ell$) as well, and $\lambda_k \leq \lambda_j \leq \lambda_\ell$.

(iv) The minimum in (2.2) is attained for the invariant subspace of $T(\lambda_\ell)$ corresponding to its ℓ largest eigenvalues.

The proof is based on the following lemma which relates the supremum of p on a subspace V of $\mathcal{H}$ to the sign of the Rayleigh quotient of $T(\lambda)$ on V and which is useful in numerical methods for computing eigenvalues of the problem (2.1).

LEMMA 2.3. *Under the conditions (A_1), (A_2), and (A_3), let $\lambda \in J$, and assume that V is a finite-dimensional subspace of $\mathcal{H}$ such that $V \cap \mathcal{D}(p) \neq \emptyset$. Then*

$$\lambda \left\{ \begin{array}{c} < \\ = \\ > \end{array} \right\} \sup_{x \in V \cap \mathcal{D}(p)} p(x) \qquad \Leftrightarrow \qquad \min_{x \in V} \langle T(\lambda)x, x\rangle \left\{ \begin{array}{c} < \\ = \\ > \end{array} \right\} 0.$$

REMARK 2.4. We only consider the case when for every $\lambda \in J$ the supremum of the essential spectrum of $T(\lambda)$ is less than 0. In the same way we obtain, for the case when for every $\lambda \in J$ the infimum of $T(\lambda)$ exceeds 0, a maxinf characterization of the eigenvalues of $T(\cdot)$ in J if we replace (A_2) by

(A_2') $(\lambda - p(x))f(\lambda; x) < 0$ *for every $x \in \mathcal{D}(p)$ and $\lambda \in J$ such that $\lambda \neq p(x)$*

ETNA
Kent State University and
Johann Radon Institute (RICAM)

and (A_3) with

(A_3') *If*

$$\nu_m(\lambda) := \inf_{V\in S_m} \max_{x\in V, x\neq 0} \langle T(\lambda)x, x\rangle / \langle x, x\rangle = 0$$

for some $m \in \mathbb{N}$ and some $\lambda \in J$, then, for $j = 1, \ldots, m$, the infimum in $\nu_j(\lambda)$ is attained, and $\nu_1(\lambda) \leq \nu_2(\lambda) \leq \cdots \leq \nu_m(\lambda)$ are the m smallest eigenvalues of the linear operator $T(\lambda)$. Conversely, if $\nu = 0$ is an eigenvalue of the operator $T(\lambda)$, then $\nu_m(\lambda) = 0$ for some $m \in \mathbb{N}$.

If the eigenvalues of $T(\cdot)$ are now enumerated in decreasing order, i.e., $\lambda \in J$ is an mth eigenvalue of $T(\cdot)$ if $\nu_m(\lambda) = 0$, for $m \in \mathbb{N}$, then λ_m can be characterized by

$$\lambda_m = \max_{\substack{V\in S_m \\ V\cap\mathcal{D}(p)\neq\emptyset}} \inf_{v\in V\cap\mathcal{D}(p)} p(v).$$

In the following we consider only problem (2.1) under the conditions (A_1), (A_2), and (A_3) although the analogue results also hold under the conditions (A_1), (A_2'), and (A_3') with the modified enumeration given above. If the extreme eigenvalue λ_1 is contained in J, then the enumeration based on (A_3) is the natural ordering. For this case, Barston [19] proved the minmax characterization for some extreme real eigenvalues for the finite-dimensional quadratic eigenvalue problems. Abramov [2] and Hasanov [101] derived the minmax and maxmin characterizations for the extreme eigenvalues for pencils of waveguide type, which are certain quadratic eigenvalue problems depending on two parameters. For $T(\cdot)$ with $\lambda_1 \in J$, it can further be shown that the eigenspaces corresponding to eigenvalues in J are contained in $\mathcal{D}(p) \cup \{0\}$. Hence, the minmax characterization is of the following form:

THEOREM 2.5. *Let the conditions (A_1), (A_2), and (A_3) be satisfied, and assume that $\lambda_1 = \inf_{x\in\mathcal{D}(p)} p(x) \in J$ and $\lambda_n \in J$, for some $n \in \mathbb{N}$.*

If $j \in \{1, \ldots, n\}$ and $V \in S_j$ such that $\lambda_j = \sup_{x\in V\cap\mathcal{D}(p)} p(x)$, then $V \subset \mathcal{D}(p) \cup \{0\}$, and the characterization of λ_j can be replaced by

$$\lambda_j = \min_{\substack{V\in S_j \\ V\subset\mathcal{D}(p)\cup\{0\}}} \sup_{v\in V\cap\mathcal{D}(p)} p(v).$$

A generalization of the maxmin characterization of Courant, Fischer, and Weyl was proved in [234]:

THEOREM 2.6 ([234]). *Assume that the conditions (A_1), (A_2), and (A_3) are satisfied. If there exists an nth eigenvalue $\lambda_n \in J$ of $T(\lambda)x = 0$, then*

$$\lambda_n = \max_{\substack{V\in S_{n-1} \\ V^\perp\cap\mathcal{D}(p)\neq\emptyset}} \inf_{v\in V^\perp\cap\mathcal{D}(p)} p(v),$$

and the maximum is attained by $W := \operatorname{span}\{u_1, \ldots, u_{n-1}\}$, where u_j denotes an eigenvector corresponding to the jth-largest eigenvalue $\mu_j(\lambda_n)$ of $T(\lambda_n)$.

Essentially the same variational characterizations of Poincaré- and of Courant-Fischer-Weyl-type were derived by Mel'nik and Nazarov [174], where $T(\lambda)$ is a set of bounded self-adjoint operators depending continuously differentiably on λ, by Griniv and Mel'nik [83] for $T(\lambda) = A(\lambda) - I$, where $A(\lambda)$ is self-adjoint and compact, and by Binding, Eschwé, and H. Langer [40] for general bounded and self-adjoint $T(\lambda)$ depending continuously on λ. Eschwé and M. Langer [66] obtained these variational characterizations also for unbounded operators. In most of these papers the natural enumeration of the eigenvalues is used. However, in [66] it is shifted by the dimension of the maximal subspace on which the functions $\langle T(\cdot)x, x\rangle$ are negative on the whole interval J.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Hadeler [90, 91] proved Rayleigh's principle for differentiable overdamped problems. He showed that the eigenvectors are orthogonal with respect to the scalar product

$$[x,y] := \begin{cases} \left\langle \dfrac{T(p(x))-T(p(y))}{p(x)-p(y)}x, y \right\rangle & \text{if} \quad p(x) \neq p(y), \\ \langle T'(p(x))x, y\rangle & \text{if} \quad p(x) = p(y), \end{cases} \tag{2.3}$$

which is symmetric, definite, and homogeneous but in general not bilinear. For non-differentiable problems, the scalar product (2.3) can be modified for the case $p(x) = p(y)$ by setting $[x,y] := \langle x, y\rangle$. Then the generalized scalar product $[\cdot,\cdot]$ becomes discontinuous for $p(x) = p(y)$, but the continuity is not needed in the proof of the following Rayleigh's principle ([246]):

THEOREM 2.7 ([246]). *Under the conditions (A_1), (A_2), (A_3), assume that J contains $n \geq 1$ eigenvalues $\lambda_1 \leq \cdots \leq \lambda_n$ (where λ_i is an ith eigenvalue) with orthogonal (with respect to $[\cdot,\cdot]$) eigenvectors $x_1, \ldots, x_n$.*

If there exists $x \in \mathcal{D}(p)$ with $[x_i, x] = 0$, for $i = 1, \ldots, n$, then J contains an $(n+1)$st eigenvalue, and

$$\lambda_{n+1} = \inf\{p(x) \,:\, [x_i, x] = 0,\ i = 1, \ldots, n\}.$$

Here, we took advantage of the Rayleigh functional to obtain variational characterizations of eigenvalues. It can also be used to prove approximation properties of eigenvector approximations as was done by Schreiber and Schwetlick [204, 206].

3. Numerical methods for dense nonlinear eigenproblems. In the following sections we focus on numerical methods for small dense nonlinear symmetric eigenproblems, i.e.,

$$T(\lambda)x = 0, \tag{3.1}$$

assuming that $T(\lambda)$ is a family of Hermitian matrices, which is a necessary prerequisite for the variational characterization of Section 2. In Section 3.1, methods based on a corresponding scalar equation are discussed, whereas Section 3.2 considers numerical approaches based on Newton's method. It is noticeable that most of them apply also to non-symmetric eigenproblems. However, for iterative projection methods considered in the next section, the residual inverse iteration, Algorithm 3, and the safeguarded iteration, Algorithm 4, are of particular interest.

For polynomial or rational eigenproblems, a common approach is to use linearization and apply standard methods for solving linear eigenvalue problems [81, 106, 190]. However, in many applications, the polynomial eigenproblems possess some desirable structure that should be preserved and exploited in their numerical solution for reasons of efficiency, stability, and accuracy [67, 106, 161, 164].

Furthermore, in [163] an approach was introduced to construct linearizations of polynomial eigenvalue problems, which generalize the companion forms, and which gave rise to linearizations preserving symmetry [102], definiteness [85, 103, 183], and respecting palindromic and odd-even structures [162]. We do not review these types of problems here.

For general nonlinear eigenproblems, the classical approach is to formulate the eigenvalue problem as a system of nonlinear equations and to use variants of the Newton's method or the inverse iteration [13, 123, 137, 186, 188]. Thus, these methods are local and not guaranteed to converge, but as for linear eigenvalue problems, their basin of convergence can be enlarged by homotopy methods [52, 117, 155] or trust region strategies [258].

Generally, methods for solving dense eigenproblems require several factorizations of varying matrices to approximate one eigenvalue. Thus, they are only appropriate for relatively small eigenproblems (with dimensions up to 1000, depending on the computer in use). In case

ETNA
Kent State University and
Johann Radon Institute (RICAM)

of large and sparse problems, iterative projection methods from Section 4 are applicable. It should be noted that when solving large and sparse problems, also solvers for dense nonlinear eigenproblems discussed in this section are needed to solve the projected eigenproblems.

3.1. Methods based on a scalar equation. In this section, we consider methods that are based on smooth scalar functions $\phi(\lambda)$ such that the eigenvalue $\hat{\lambda}$ of interest of (3.1) is a root of ϕ. We apply Newton's method to solve $\phi(\lambda) = 0$ for some initial guess, but any other suitable method (e.g., a higher-order method) may be used instead to solve this equation.

The most natural choice is $\phi(\lambda) = \det T(\lambda)$. It was suggested by Kublanovskaya [123, 124] to use a QR-decomposition with column pivoting $T(\lambda)P(\lambda) = Q(\lambda)R(\lambda)$, where $P(\lambda)$ is a permutation matrix that 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 λ is an eigenvalue if and only if $r_{nn}(\lambda) = 0$. Applying Newton's method to this equation yields the iteration

$$\lambda_{k+1} = \lambda_k - \frac{1}{e_n^H Q(\lambda_k)^H T'(\lambda_k) P(\lambda_k) R(\lambda_k)^{-1} e_n} \tag{3.2}$$

for eigenvalue approximations of (3.1), where e_n denotes the nth unit vector. Approximations of left and right eigenvectors can be obtained from $y_k = Q(\lambda_k)e_n$ and $x_k = P(\lambda_k)R(\lambda_k)^{-1}e_n$.

An improved version of Kublanovskaya's method was suggested by Jain, Singhal, and Huseyin [112], who also proved quadratic convergence of their scheme. A similar approach was presented by Yang [259] via a representation of Newton's method using the LU factorization of $T(\lambda)$.

A careful analysis of the QR algorithm for banded matrices with a narrow bandwidth based on Newton's method and a new version based on a new data structure enabling a more efficient use of memory is described in [74]. The source code contained in [75] is publicly available.

The following method (originally applied to polynomial matrix-valued functions $T(\cdot)$) called nonlinear generalized Rayleigh quotient iteration (NGRQI) was introduced by Lancaster [136] and also applies to more general $T(\cdot)$. Let $T(\cdot)$ be holomorphic in a neighborhood Λ of an eigenvalue $\hat{\lambda}$ of $T(\cdot)$, and assume that a and b are not orthogonal to $\ker(T^H(\hat{\lambda}))$ and $\ker(T(\hat{\lambda}))$, respectively. For a given $\lambda \neq \hat{\lambda}$, let $v(\lambda)$ and $w(\lambda)$ be solutions of the linear systems

$$T(\lambda)v(\lambda) = a \quad \text{and} \quad T(\lambda)^H w(\lambda) = b, \tag{3.3}$$

and define

$$\phi(\lambda) := \frac{1}{b^H T(\lambda)^{-1} a}.$$

Then Newton's method applied to ϕ yields the NGRQI method with the generalized Rayleigh quotients

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

at $(\lambda_k, v(\lambda_k), w(\lambda_k))$. If $\hat{\lambda}$ is a simple eigenvalue, then $\hat{\lambda}$ is a simple root of ϕ, and the NGRQI method converges quadratically to $\hat{\lambda}$ [136].

Since the system matrices in (3.3) become ill-conditioned close to $\hat{\lambda}$, Schwetlick and Schreiber [206] considered an equivalent bordered version of (3.3),

$$\begin{bmatrix} T(\lambda) & a \\ b^H & 0 \end{bmatrix} \begin{bmatrix} s \\ \mu_s \end{bmatrix} = \begin{bmatrix} 0 \\ \alpha \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} T(\lambda)^H & b \\ a^H & 0 \end{bmatrix} \begin{bmatrix} t \\ \nu_t \end{bmatrix} = \begin{bmatrix} 0 \\ \alpha \end{bmatrix},$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

which has to be solved, where α is a scaling factor. The system matrices are nonsingular for λ close enough to a simple eigenvalue $\hat{\lambda}$ if a and b are not orthogonal to the right and left eigenvector, respectively. Unger [228] (and more generally Langer [140] for holomorphic operator functions in a Hilbert space) considered also the case of multiple eigenvalues $\hat{\lambda}$.

A further approach has been introduced by Andrew, Chu, and Lancaster [12]. They take advantage of **B**ordered matrices, **D**eleting one row or column, and **S**ubstituting one of the vectors b or c, which motivates the name BDS method. The general form of BDS methods, which—like the approach in [206]—avoids the solution of linear systems with nearly singular matrices, reads as follows: For fixed vectors b and c and $H(\lambda) := T(\lambda) + bc^H$, determine s, t, and ϕ such that

$$\begin{aligned} H(\lambda)s(\lambda) &= (1-\phi(\lambda))b, \qquad t(\lambda)^T H(\lambda) = (1-\phi(\lambda))c^H, \\ b^H t(\lambda) &= 1 = c^H s(\lambda), \quad \text{and} \qquad \phi(\lambda) = -t(\lambda)^H T(\lambda) s(\lambda) = 0. \end{aligned}$$

The function $\phi(\lambda)$ can be evaluated in the following way: By solving $H(\lambda)\tilde{s}(\lambda) = b$ and scaling the solution, one obtains

$$s(\lambda) = \tilde{s}(\lambda) / \left(c^H \tilde{s}(\lambda)\right) \quad \text{and} \quad \phi(\lambda) = 1 - \left(c^H \tilde{s}(\lambda)\right)^{-1},$$

and $t(\lambda)$ can be calculated similarly. Determining a root $\hat{\lambda}$ by Newton's method is obviously equivalent to NGRQI; the eigenvector approximations, however, are different from the natural choices $s(\lambda) = T(\lambda)^{-1}c$ and $t(\lambda)^H = b^H T(\lambda)^{-1}$.

Osborne [187] considers Newton's method for the complex function $\phi(\lambda)$ defined by

$$T(\lambda)u = \phi(\lambda)x, \quad s^H u = \kappa,$$

where κ is a given constant and x and s are given vectors. The corresponding basic iteration applies Newton updates for λ for determining roots of

$$\phi(\lambda) := \frac{\kappa}{s^H T(\lambda)^{-1} x}.$$

This approach generalizes the method (3.2), the inverse iteration, and a method proposed in [188]. It was proved that the rate of convergence is quadratic and that cubic convergence can be obtained if not only λ but also x and/or s are updated appropriately, thus unifying the results in [13, 123, 137, 186, 188].

3.2. Methods based on Newton's method. Applying Newton's method to the nonlinear system

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

where $v \in \mathbb{C}^n$, $v \neq 0$, is suitably chosen, yields the (nonlinear) inverse iteration method of Algorithm 1, which converges locally and quadratically for simple eigenpairs.

If $T(\cdot)$ is Hermitian and the conditions (A_1) and either (A_2) or $(A_2)'$ of Section 2 are satisfied and if the update of λ_{k+1} in step 3 is replaced with $\lambda_{k+1} \leftarrow p(x_{k+1})$, given that all $x_k \in \mathcal{D}(p)$, then one obtains the Rayleigh functional iteration method of Algorithm 2. This method converges locally and cubically for simple eigenpairs.

The cost for solving a linear system in each iteration step with a varying matrix can be avoided by using the residual inverse iteration method of Algorithm 3, which replaces a varying matrix $T(\lambda_k)$ by a fixed matrix $T(\lambda_0)$, at least for several iteration steps. This idea was introduced by Neumaier [180].

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Algorithm 1 Inverse iteration.

Require: initial pair (λ_0, x_0) and normalization vector v with $v^H x_0 = 1$
  for $k = 0, 1, 2, \ldots$ until convergence **do**
    solve $T(\lambda_k)x_{k+1} = T'(\lambda_k)x_k$ for x_{k+1}
    set $\lambda_{k+1} = \lambda_k - v^H x_k/(v^H x_{k+1})$
    normalize $x_{k+1} \leftarrow x_{k+1}/(v^H x_{k+1})$
  end for

Algorithm 2 Rayleigh functional iteration.

Require: initial pair (λ_0, x_0) and normalization vector v with $v^H x_0 = 1$
1: **for** $k = 0, 1, 2, \ldots$ until convergence **do**
2: solve $T(\lambda_k)x_{k+1} = T'(\lambda_k)x_k$ for x_{k+1}
3: set $\lambda_{k+1} = p(x_{k+1})$
4: normalize $x_{k+1} \leftarrow x_{k+1}/(v^H x_{k+1})$
5: **end for**

If $T(\cdot)$ is Hermitian and $\lambda_0 \in \mathbb{R}$, then the convergence can be improved by determining λ_{k+1} in step 2 via the Rayleigh functional, i.e., solving $x_k^H T(\lambda_{k+1})x_k = 0$ for λ_{k+1}, given the same conditions as for the Rayleigh functional iteration above.

If $T(\cdot)$ is twice differentiable and $\hat{\lambda}$ is algebraically simple, then the residual inverse iteration converges for all (λ_0, x_0) sufficiently close to $(\hat{\lambda}, \hat{x})$, and

$$\|x_{k+1} - \hat{x}\|/\|x_k - \hat{x}\| = \mathcal{O}(|\lambda_0 - \hat{\lambda}|) \qquad \text{and} \qquad |\lambda_{k+1} - \hat{\lambda}| = \mathcal{O}(\|x_k - \hat{x}\|^t),$$

where $t = 2$ in the Hermitian case if λ_k is updated via the Rayleigh functional and $t = 1$ otherwise [204].

If (3.1) allows for a variational characterization of its eigenvalues, then the safeguarded iteration in Algorithm 4, which aims at a particular eigenvalue, is a natural choice. The safeguarded iteration was introduced by Voss and Werner [249]. Under assumptions (A_1) and (A_2) of Section 2, it has the following convergence properties; cf. [249] or [183], which is accessible more easily.

THEOREM 3.1 ([249]). *Let $J \subset \mathbb{R}$ be an open interval, and let $T(\lambda) \in \mathbb{C}^{n\times n}$, $\lambda \in J$, be a family of Hermitian matrices allowing for the minmax characterization.*

(i) If $\lambda_1 := \inf_{x\in\mathcal{D}(p)} p(x) \in J$ and $x_0 \in \mathcal{D}(p)$, then the safeguarded iteration for $j = 1$ converges globally and monotonically decreasing to λ_1.

(ii) If $T(\lambda)$ is holomorphic in a neighborhood $U \subset \mathbb{C}$ of a jth eigenvalue of $T(\cdot)$ and λ_j is a simple eigenvalue, then the safeguarded iteration converges locally and quadratically to λ_j.

(iii) Under the conditions of (ii), the convergence is even cubic if $T'(\lambda)$ is positive definite for $\lambda \in U \cap J$ and x_k in step 3 of Algorithm 4 is chosen to be an eigenvector corresponding to the jth-largest eigenvalue of the generalized eigenproblem $T(\sigma_{k-1})x = \mu T'(\sigma_{k-1})x$.

REMARK 3.2. In every iteration step of the methods based on Newton's method, one has to solve a linear system. Szyld and Fei [220] discussed the local convergence of inexact versions of several of these methods (inverse iteration, Rayleigh quotient iteration, residual inverse iteration, single vector Jacobi-Davidson method) demonstrating that its order of local convergence can be preserved. When local symmetry of $T(\cdot)$ is present, the use of a nonlinear Rayleigh functional is shown to be fundamental in achieving a higher-order convergence

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Algorithm 3 Residual inverse iteration.

Require: initial pair (λ_0, x_0) and normalization vector w with $w^H x_0 = 1$
1: **for** $k = 0, 1, 2, \ldots$ until convergence **do**
2: solve $w^H T(\lambda_0)^{-1} T(\lambda_{k+1}) x_k = 0$ for λ_{k+1}
3: solve $T(\lambda_0) u_k = T(\lambda_{k+1}) x_k$ for u_k
4: set $v_{k+1} \leftarrow x_k - u_k$
5: normalize $x_{k+1} \leftarrow v_{k+1}/(w^H v_{k+1})$
6: **end for**

Algorithm 4 Safeguarded iteration.

Require: initial vector $x_0 \in \mathcal{D}(p)$
1: compute $\sigma_0 = p(x_0)$
2: **for** $k = 1, 2, \ldots$ until convergence **do**
3: determine an eigenvector x_k corresponding to the jth-largest eigenvalue of $T(\sigma_{k-1})$
4: evaluate the Rayleigh functional $\sigma_k := p(x_k)$, i.e., solve $x_k^H T(\sigma_k) x_k = 0$ for σ_k
5: **end for**

rate [220]. In [219, 221, 222] they showed that the convergence of Newton-like methods for degenerate eigenvalues in general is linear, but for semi-simple eigenvalues, convergence is at least quadratic.

REMARK 3.3. For eigenvalue problems satisfying the conditions of the minmax characterization in an interval J, eigenvectors corresponding to different eigenvalues are necessarily linearly independent. In the general case, however, it may even happen that different eigenvalues share the same eigenvector. Generalizing the notion of eigenspaces, Betcke and Kressner [32] introduced and studied invariant pairs that can be computed in a stable way. Taking advantage of this notion Beyn, Effenberger, and Kressner [39, 120] designed algorithms for computing several eigenpairs of invariant pairs.

The methods considered so far are constructed for computing one eigenvalue or eigenpair at a time. To determine more eigenpairs one can repeat the calculations with modified initial values, but some care has to be taken to prohibit the method to converge to already converged eigenpairs. A standard approach called deflation consists of mapping an already converged eigenvalue to ∞ while preserving the others [89].

4. Iterative projection methods for large nonlinear eigenproblems. For sparse linear eigenvalue problems

$$Ax = \lambda x,$$

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 then solving 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 original problem.

Two types of iterative projection methods are in use: methods that expand the subspaces independently of the eigenpair of the projected problem and take advantage of a normal form of A, like the Arnoldi, Lanczos, and rational Krylov methods, and methods that aim at a particular eigenpair and choose the expansion such that it has a high approximation potential for a wanted eigenvector, like the Jacobi-Davidson method. Today the Arnoldi

ETNA
Kent State University and
Johann Radon Institute (RICAM)

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 [16]. Implementations of the (implicitly restarted) Arnoldi method is available in the package ARPACK [145] and in MATLAB with the command `eigs`.

For general nonlinear eigenproblems

$$T(\lambda)x = 0, \quad T : \mathbb{R} \supset J \to \mathbb{R}^{n\times n}, \tag{4.1}$$

with J being an open interval which may be unbounded, a normal form like the Schur factorization does not exist. Therefore, generalizations of iterative projection methods to general nonlinear eigenproblems always have to be of the second type, i.e., aiming at an individual eigenpair. These methods are considered in the next sections.

4.1. Iterative projection methods. An iterative projection method for a nonlinear eigenvalue problem (4.1) has the form given in Algorithm 5.

Algorithm 5 Iterative projection method for nonlinear eigenproblems.

Require: initial basis V with $V^HV = I$; set $m = 1$
1: **while** $m \leq$ the number of wanted eigenvalues **do**
2: compute an eigenpair (θ, y) of the projected problem $V^HT(\lambda)Vy = 0$.
3: determine the Ritz vector $x = Vy$, $\|x\| = 1$, and the residual $r = T(\theta)x$
4: **if** $\|r\| < \varepsilon$ **then**
5: accept an approximate eigenpair $\lambda_m = \theta$, $x_m = x$; increase $m \leftarrow m + 1$
6: reduce the search space V if necessary
7: choose an approximation (λ_m, x) of the next eigenpair, and compute $r = T(\lambda_m)x$
8: **end if**
9: expand the search space $V = [V, v_{\text{new}}]$
10: update the projected problem
11: **end while**

The main question is how to expand the search space $\mathcal{V} = \text{span}\{V\}$, i.e., how to determine the new search direction v_{new} in line 9 if the approximation of an eigenvalue by a solution of the projected problem is not sufficiently accurate.

Let θ be an eigenvalue of the projected problem

$$V^HT(\lambda)Vy = 0$$

and $x = Vy$ a corresponding Ritz vector. Then inverse iteration yields a suitable candidate

$$v := T(\theta)^{-1}T'(\theta)x$$

for the expansion of $\mathcal{V}$. To implement this expansion we have to solve a large linear system, where from step to step the system matrix varies. In a truly large problem, the exact solution v will not be accessible but only an inexact solution $\tilde{v} := v + e$ of $T(\theta)v = T'(\theta)x$, and the next iterate will be a solution of the projection of $T(\lambda)x = 0$ upon the expanded space $\tilde{\mathcal{V}} := \text{span}\{\mathcal{V}, \tilde{v}\}$.

If we assume that x is already a good approximation of an eigenvector of $T(\cdot)$, then v will be an even better approximation, and therefore the eigenvector we are looking for will be very close to the plane $E := \text{span}\{x, v\}$. We therefore neglect the influence of the orthogonal complement of x in $\mathcal{V}$ on the next iterate and discuss the nearness of the planes E and $\tilde{E} := \text{span}\{x, \tilde{v}\}$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

If the angle between these two planes is small, then the projection of $T(\lambda)$ upon $\tilde{\mathcal{V}}$ should be similar to the one upon span$\{\mathcal{V}, v\}$, and the approximation properties of the inverse iteration should be maintained. If this angle becomes large, then it is not surprising that the convergence properties of inverse iteration are not reflected by the projection method.

THEOREM 4.1 ([241]). *Let $\phi_0 = \arccos(x^T v)$ denote the angle between x and v, with $x, v \in \mathbb{R}^n$, $\|x\| = \|v\| = 1$, and the relative error of $\tilde{v}$ by $\varepsilon := \|e\|$.*

Then the maximal possible acute angle between the planes E and $\tilde{E}$ is

$$\beta(\varepsilon) = \begin{cases} \arccos\sqrt{1 - \varepsilon^2/\sin^2\phi_0} & \text{if } \varepsilon \le |\sin\phi_0|, \\ \frac{\pi}{2} & \text{if } \varepsilon \ge |\sin\phi_0|. \end{cases}$$

Obviously, for every $\alpha \in \mathbb{R}$, $\alpha \neq 0$, the plane E is also spanned by x and $x + \alpha v$. If $\tilde{E}(\alpha)$ is the plane which is spanned by x and a perturbed realization $x + \alpha v + e$ of $x + \alpha v$, then by the same arguments as in the proof of Theorem 4.1, the maximum angle between E and $\tilde{E}(\alpha)$ is

$$\gamma(\alpha, \varepsilon) = \begin{cases} \arccos\sqrt{1 - \varepsilon^2/\sin^2\phi(\alpha)} & \text{if } \varepsilon \le |\sin\phi(\alpha)|, \\ \frac{\pi}{2} & \text{if } \varepsilon \ge |\sin\phi(\alpha)|, \end{cases}$$

where $\phi(\alpha)$ denotes the angle between x and $x + \alpha v$. Since the mapping

$$\phi \mapsto \arccos\sqrt{1 - \varepsilon^2/\sin^2\phi}$$

decreases monotonically in the interval $\phi \in [\arcsin(\varepsilon), \pi/2]$, the expansion of the search space by an inexact realization of $x + \alpha v$ is most robust with respect to small perturbations if α is chosen such that x and $x + \alpha v$ are orthogonal, i.e.,

$$v = x - \frac{x^H x}{x^H T(\theta)^{-1} T'(\theta) x} T(\theta)^{-1} T'(\theta) x,$$

which yields a maximum acute angle between E and $\tilde{E}(\alpha)$,

$$\gamma(\alpha, \varepsilon) = \begin{cases} \arccos\sqrt{1 - \varepsilon^2} & \text{if } \varepsilon \le 1, \\ \frac{\pi}{2} & \text{if } \varepsilon \ge 1. \end{cases}$$

This expansion v of the current search space $\mathcal{V}$ can be obtained by the solution of the equation

$$\left(I - \frac{T'(\theta) x x^H}{x^H T'(\theta) x}\right) T(\theta) \left(I - x x^H\right) v = -r, \qquad v \perp x, \tag{4.2}$$

with $r = T(\theta)x$.

This is the so called correction equation of the Jacobi-Davidson method, which was derived in [38] generalizing the approach of Sleijpen and van der Vorst [211] for linear and polynomial eigenvalue problems. Hence, the Jacobi-Davidson method is the most robust realization of an expansion of a search space such that the direction of the inverse iteration is contained in the expanded space in the sense that it is least sensitive to inexact solves of linear systems $T(\theta)v = T'(\theta)x$.

Neglecting the orthogonalization with respect to the previous search space, the expansion in the direction of (4.2) is equivalent to expanding the search space by the direction of the

ETNA
Kent State University and
Johann Radon Institute (RICAM)

inverse iteration. A connatural expansion of search spaces utilized in the Nonlinear Arnoldi method [237] is based on the residual inverse iteration.

There are many variants of the Jacobi-Davidson and Nonlinear Arnoldi methods in the literature [34, 35, 37, 38, 58, 59, 69, 89, 105, 145, 172, 204, 210, 211, 238, 242]. A broad survey is given in the PhD thesis of Schreiber [204]. In the following, we consider only the Jacobi-Davidson and Nonlinear Arnoldi methods, together with an early version of the Rational Krylov method by Ruhe [202].

4.2. The Jacobi-Davidson method. The correction equation (4.2) in the Jacobi-Davidson method does not have to be solved exactly to maintain fast convergence, but usually a few steps of a Krylov subspace solver with an appropriate preconditioner suffice to obtain a good expansion direction of the search space. The natural generalization of the Jacobi-Davidson method for polynomial eigenvalue problems was suggested in [210, 212] and studied in [38, 238, 242] for general nonlinear eigenproblems.

In the correction equation (4.2), the operator $T(\theta)$ is restricted to map the subspace $x^\perp$ into itself. Hence, if $M \approx T(\theta)^{-1}$ is a preconditioner of $T(\theta)$, then a preconditioner for an iterative solver of (4.2) should be modified correspondingly to

$$\tilde{M} := \left(I - \frac{T'(\theta)xx^H}{x^H T'(\theta)x}\right) M \left(I - \frac{xx^H}{x^H x}\right).$$

It was already pointed out for linear eigenproblems in [211] that taking into account the projectors in the preconditioner, i.e., using $\tilde{M}$ instead of M in a preconditioned Krylov solver, increases the cost only slightly. Applying a preconditioned Krylov solver to (4.2) requires solving one linear system with M in every iteration step, and one additional solve with $T'(\theta)$ during the initialization.

A template for the Jacobi-Davidson method for the nonlinear eigenvalue problem $T(\lambda)x = 0$ is given in Algorithm 6. In the following we comment on some of its steps. A detailed discussion is contained in [38, 238, 242].

(i) In step 1 of Algorithm 6, prior information such as known approximate eigenvectors of the problem (4.1) or eigenvectors of the contiguous problems can be introduced into the algorithm. If no information on the eigenvectors is at hand and we are interested in eigenvalues close to the parameter $\sigma \in D$, then one can choose an initial vector at random, execute a few Lanczos or Arnoldi steps for the linear eigenproblem $T(\sigma)u = \theta u$ or $T(\sigma)u = \theta T'(\sigma)u$, and choose V as an orthonormal basis of the eigenvectors corresponding to eigenvalues small in modulus. Starting with a random vector without this preprocessing step usually will yield a value λ_m in step 4 that is far away from σ and will avert convergence.
Rational eigenvalue problems governing free vibrations of fluid-solid structures require a particular initial space, the choice of which is discussed in [236].

(ii) Preconditioning is key to a successful iterative solver. A comprehensive exposition of many useful preconditioning techniques can be found in [45, 203]. Sleijpen and van der Vorst [211] suggested to precondition by a few steps of BiCGStab (or GMRES in the non-symmetric case), which essentially costs one matrix-vector product in every iteration step and one additional matrix-vector product to initialize.

(iii) 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 3.

(iv) 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. In the linear case this is not a problem. Krylov subspace solvers construct an orthogonal basis of the ansatz space not aiming

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Algorithm 6 Nonlinear Jacobi-Davidson method.

Require: initial basis V with $V^HV = I$; set $m = 1$

determine a preconditioner $M \approx T(\sigma)^{-1}$, σ close to the first wanted eigenvalue
while $m \leq$ the number of wanted eigenvalues **do**
  compute an approximation of the mth wanted eigenvalue λ_m and the corresponding eigenvector y_m of the projected problem $V^HT(\lambda)Vy = 0$
  determine the Ritz vector $u = Vy_m$ and the residual $r = T(\lambda_m)u$
  if $\|r\|/\|u\| < \epsilon$ **then**
    accept an approximate eigenpair (λ_m, u); increase $m \leftarrow m + 1$;
    reduce the search space V if indicated
    determine a new preconditioner $M \approx T(\lambda_m)^{-1}$ if indicated
    choose an approximation (λ_m, u) of the next eigenpair
    compute the residual $r = T(\lambda_m)u$;
  end if
  Find an approximate solution of the correction equation

$$\left(I - \frac{T'(\lambda_m)uu^H}{u^HT'(\lambda_m)u}\right) T(\sigma) \left(I - \frac{uu^H}{u^Hu}\right) z = -r$$

  (by a preconditioned Krylov solver, e.g.)
  orthogonalize $z = z - VV^Hz$, $v = z/\|z\|$, and expand the subspace $V = [V, v]$
  update the projected problem
end while

at a particular eigenvalue, and one gets approximations of extreme eigenvalues without replication, at least if reorthogonalization 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. [69].

If $T(\lambda)$ is a family of real symmetric or Hermitian 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 the safeguarded iteration, and while approximating the mth eigenvalue, usually enough information about the next eigenvector is gathered to compute the $(m + 1)$st eigenvalue safely. This approach, discussed in [38], 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 σ_0 and approximate them one after the other by the method of successive linear problems. If already $m - 1$ eigenvalues of (4.1) closest to σ_0 have been determined and μ_ℓ is an approximation of the eigenvalue wanted next, then we iteratively perform the following three steps until convergence: we solve the linear eigenproblem $V^HT(\mu_\ell)Vy = \theta V^HT'(\mu_\ell)Vy$, choose the eigenvalue $\hat{\theta}$ such that $|\sigma_0 - (\mu_\ell - \hat{\theta})|$ is mth-smallest among the eigenvalues θ, and set $\mu_{\ell+1} = \mu_\ell - \hat{\theta}$.

A disadvantage of this method is the fact that consecutive eigenvalues λ_{m-1} and λ_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 [242]. Hence, this method should be used only if a small number of eigenvalues close to a parameter are of interest.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Quite often the nonlinear eigenvalue problem under consideration is a (small) perturbation of a linear eigenvalue problem. This happens for instance for rational eigenproblems governing the free vibrations of a structure using a viscoelastic constitutive relation to describe the behavior of the material [242]. It is well known that often the eigenmodes of the damped and undamped problems 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 mth-smallest eigenvalue ω_m^2, determine an approximate eigenvalue $\tilde{\omega}$ of the nonlinear projected problem from the complex equations $y^H V^H T(\omega) V y = 0$ or $e^H V^H T(\sigma)^{-1} T(\omega) V y = 0$ for some fixed vector $e \neq 0$, and correct it by one of the methods in Section 3.

(v) 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 eigenvector that 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 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 $X_m := \text{span}\{x_1, \ldots, x_m\}$. If an approximation of an eigenvector wanted next is obtained cheaply, then we add it to X_m. A local restart procedure which is particularly suitable if a very large number of eigenvalues or eigenvalues in the interior of the spectrum are desired is discussed in [166].

(vi) Some of the eigensolvers discussed in Section 3 can be used to get approximations of 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, then we continue with the current approximations.

(vii) v is orthogonalized with respect to the current search space $\mathcal{V}$ by the classical Gram-Schmidt method. It may be replaced by the modified Gram-Schmidt method for stability reasons. Notice, however, that the classical Gram-Schmidt procedure is able to use BLAS3 and thus can be faster than the modified Gram-Schmidt method due to the better use of cache.

4.3. The Nonlinear Arnoldi method. Expanding the current search space $\mathcal{V}$ by the direction $\hat{v} = x - T^{-1}(\sigma)T(\theta)x$, which is suggested by the residual inverse iteration, generates similar robustness problems as in the case of the inverse iteration. If $\hat{v}$ is close to the desired eigenvector, then an inexact evaluation of $\hat{v}$ spoils the favorable approximation properties of the residual inverse iteration.

Similarly as in the Jacobi-Davidson method, one could replace $\hat{v}$ by $z := x + \alpha\hat{v}$, where α is chosen such that $x^H z = 0$, and one could determine an approximation of z by solving a correction equation. However, since the new search direction is orthonormalized against the previous search space $\mathcal{V}$ and since x is contained in $\mathcal{V}$, we may choose the new direction $\tilde{v} = T(\sigma)^{-1}T(\theta)x$ as well. This direction satisfies the orthogonality condition $x^H\tilde{v} = 0$ at least in the limit as θ approaches a simple eigenvalue $\hat{\lambda}$ (cf. [240]), i.e.,

$$\lim_{\theta \to \hat{\lambda}} x^H T(\sigma)^{-1} T(\theta) x = 0.$$

For the linear problem $T(\lambda) = A - \lambda B$, the expansion $\tilde{v}$ is exactly the Cayley transform with

ETNA
Kent State University and
Johann Radon Institute (RICAM)

pole σ and zero θ. Since

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

and Krylov spaces are shift-invariant, the resulting projection method of expanding $\mathcal{V}$ by v is nothing else but the shift-and-invert Arnoldi method.

If the linear system $T(\sigma)v = T(\theta)x$ is too expensive to solve for v, then we may choose as new direction $v = MT(\theta)x$ with $M \approx T(\sigma)^{-1}$, and for the linear problem we obtain an inexact Cayley transform or a preconditioned Arnoldi method. The resulting iterative projection method, which was introduced in [170, 171] for quadratic eigenvalue problems and was studied in [235, 237] for general nonlinear eigenproblems, is referred to as the Nonlinear Arnoldi method in spite the fact that differently to the linear case, no Krylov space is determined in the course of the algorithm and no Arnoldi recursion holds.

Since the rate of convergence depends crucially on $|\sigma-\lambda|$, it is advisable to change the shift—or more generally the preconditioner M—in the course of the algorithm if convergence to the current eigenvalue is too slow.

A template for the preconditioned Nonlinear Arnoldi method with restarts and varying preconditioner is given in Algorithm 7.

Algorithm 7 Nonlinear Arnoldi method.

Require: initial shift σ and an initial basis V with $V^HV = I$; set $m = 1$
determine a preconditioner $M \approx T(\sigma)^{-1}$, σ close to the first wanted eigenvalue
while $m \leq$ the number of wanted eigenvalues **do**
  compute an appropriate eigenvalue θ and the corresponding eigenvector y of the projected problem $V^HT(\theta)Vy = 0$
  determine the Ritz vector $u = Vy$ and the residual $r = T(\theta)u$
  if $\|r\|/\|u\| < \epsilon$ **then**
    accept $\lambda_m = \theta$, $x_m = u$, increase $m \leftarrow m+1$
    determine a new preconditioner $M \approx T(\sigma)^{-1}$ if indicated
    restart if necessary
    choose an approximations θ and u of the next eigenvalue and eigenvector
    determine the residual $r = T(\theta)u$
  end if
  $\hat{v} = Mr$
  $v = \hat{v} - VV^H\hat{v}$,$\tilde{v} = v/\|v\|$, $V = [V, \tilde{v}]$
  reorthogonalize if necessary
  update the projected problem $V^HT(\theta)Vy = 0$
end while

Since the residual inverse iteration with fixed pole σ converges linearly and the contraction rate satisfies $\mathcal{O}(|\sigma-\lambda_m|)$, it is reasonable to update the preconditioner if the convergence (measured by the quotient of the last two residual norms before convergence) has become too slow. For several other recent variations and generalizations of the Arnoldi method for quadratic or general polynomial eigenvalue problems, see [17, 18, 70, 107, 154, 170, 171, 225].

REMARK 4.2. The LSTRS software for the efficient solution of large-scale Trust-Region subproblems was proposed by Rojas, Santos, and Sorensen; see [198, 213]. It is based on recasting the problem in terms of a parameter-dependent eigenvalue problem and adjusting a parameter iteratively. The essential effort in each iteration is the solution of an eigenvalue problem for the smallest eigenvalue of a bordered Hessian matrix (or the two smallest eigenval-

ETNA
Kent State University and
Johann Radon Institute (RICAM)

ues in the so-called hard case) and the associated eigenvector(s). Using the Nonlinear Arnoldi method to solve the eigenvalue problems makes it possible to recycle most of the information from previous iterations, which can substantially accelerate LSTRS [128].

4.4. Rational Krylov method. In some sense, Ruhe [201, 202] generalized the rational Krylov approach for linear eigenproblems to sparse nonlinear eigenvalue problems. His idea was to nest the linearization of problem (4.1) by Lagrangian interpolation and to solve the resulting linear eigenproblem by Arnoldi's method. Similarly to the rational Krylov process for linear eigenvalue problems, a sequence $\mathcal{V}_k$ of subspaces of $\mathbb{C}^n$ is constructed. At the same time Hessenberg matrices H_k, which approximate the projection of $T(\sigma)^{-1}T(\lambda_k)$ to $\mathcal{V}_k$, are updated. Here σ denotes a shift (which similarly as in the rational Krylov method for linear problems can be updated in the course of the algorithm) and λ_k an approximation of the wanted eigenvalue of (4.1). Then a Ritz vector x_k of H_k corresponding to an eigenvalue of small modulus approximates an eigenvector of the nonlinear problem from which a (hopefully) improved eigenvalue approximation of problem (4.1) is obtained.

The convergence results of this first version of the rational Krylov method for nonlinear problems were far from being satisfactory. To improve convergence, Ruhe in [202] proposed an inner iteration which enforces the residual $r_k = T(\sigma)^{-1}T(\lambda_k)x_k$ to be orthogonal to the search space $\mathcal{V}_k$. (This property is automatically satisfied for linear eigenproblems.) The inner iteration is presented heuristically not noticing that it actually is nothing else but a solver of the projected nonlinear eigenproblem $V_k^H T(\sigma)^{-1}T(\lambda)V_k s = 0$. Thus, the rational Krylov method for nonlinear eigenproblems can be interpreted as an iterative projection method [113]. The inner iteration can be replaced by any solver for dense nonlinear eigenproblems.

Algorithm 8 Rational Krylov method for nonlinear eigenproblems.

Require: initial vector $V_1 = [v_1]$ with $\|v_1\| = 1$, initial λ_1 and σ
1: **for** $k = 1, 2, \ldots$ until convergence **do**
2: solve the projected eigenproblem $V_k^H T(\sigma)^{-1}T(\lambda)V_k s = 0$ for (λ, s)
3: compute the Ritz vector $x = V_k s$ and the residual $r = T(\sigma)^{-1}T(\lambda)x$
4: orthogonalize $r = r - V_k V_k^H r$
5: expand the search space $V_{k+1} = [V_k\ ,\ r/\|r\|]$
6: **end for**

Although derived differently, the rational Krylov method expands the search space $\mathcal{V}_k$ in the same way as in the Arnoldi method for nonlinear eigenproblems introduced in [235, 237]. However, differently from the rational Krylov method, in the Arnoldi approach, the original problem $T(\lambda)x = 0$ is projected to $\mathcal{V}_k$. Thus, the Nonlinear Arnoldi method preserves symmetry properties of the problem (4.1), which can be exploited when solving the projected problems.

The inner iteration algorithm in the original rational Krylov method usually does not converge very fast, which makes the original rational Krylov method inferior to other iterative projection methods. However, there is one advantage of Ruhe's approach: The solvers for dense nonlinear eigenproblems need the explicit form of the projected problem whereas Ruhe's approach only needs a procedure that yields the vector $T(\sigma)^{-1}T(\lambda)x$ for a given vector x.

4.5. Numerical example. To demonstrate the numerical behavior of the iterative projection methods, we consider a delay differential equation [61, 93, 94, 182, 244]

$$u_t(x,t) = \Delta u(x,t) + a(x)u(x,t) + b(x)u(x,t-\tau), \qquad t > 0,\ x \in [0,\pi] \times [0,\pi],$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 4.1
Computation times for the 20 smallest eigenvalues of a discretized delay differential equation of dimension 39,601.

Preconditioner	Nonlinear Arnoldi		Jacobi-Davidson	
	# iter.	CPU	# iter.	CPU
LU	125	14.9	119	38.4
inc. LU, 10^{-3}	241	34.2	143	44.7
inc. LU, 10^{-2}	1001	245.0	177	58.2

where Δ is the Laplacian operator $\Delta := \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}$ and a and b are real functions of x. Semi-discretizing with finite differences with respect to x and the ansatz $u(x,t) = e^{\lambda t}v(x)$ yields the nonlinear eigenvalue problem

$$T(\lambda)v = \lambda v + Av + e^{-\lambda\tau}Bv = 0.$$

In [244] we tested both iterative projection methods, i.e., the Nonlinear Arnoldi and the Jacobi-Davidson procedure, for a problem of this type of dimension $n = 39,601$. Since $T(\lambda)$ is symmetric and the conditions of the minmax characterization are satisfied, the projected problems can be solved by the safeguarded iteration, and the eigenvalues can be determined safely one after the other.

We have computed the 20 smallest eigenvalues. For both methods an average of approximately 6 iterations are needed to find an eigenvalue. Notice, however, that for the Nonlinear Arnoldi method, only one solve with the preconditioner is needed to expand the search space, whereas the Jacobi-Davidson method requires the approximate solution of a correction equation.

Table 4.1 contains the CPU time for both methods, where we employed the LU factorization as well as incomplete LU factorizations for two cut-off levels, 10^{-3} and 10^{-2}, and did not reduce the search space during the iterations. It is observed that for an accurate preconditioner, the Nonlinear Arnoldi method is much faster than the Jacobi-Davidson method, whereas for a coarse preconditioner, the Jacobi-Davidson method is the clear winner. The same observation was made for many other examples; the Jacobi-Davidson method is more robust with respect to coarse preconditioners than the Nonlinear Arnoldi method. This can be explained by the motivation of the Jacobi-Davidson method in [241]: It aims at the expansion direction (containing the information of an inverse iteration step) in the most robust way, i.e., it is least sensitive to inexact solves. Despite requiring a smaller number of iterations, the Jacobi-Davidson method might be slower in terms of CPU time since one iteration is more expensive compared to one iteration of the Nonlinear Arnoldi method.

The CPU times in Table 4.1 correspond to the projection methods without restart. Figure 4.1 displays on the left the time consumption of the Nonlinear Arnoldi method with incomplete LU preconditioner with threshold 10^{-2} as well as the time that is required for solving the projected eigenvalue problems. It demonstrates the necessity of restarts since the superlinear time consumption is mainly caused by the eigensolvers. On the right, Figure 4.1 displays the behavior of the Nonlinear Arnoldi method if the method is restarted whenever the dimension of the search space exceeds 100 after the computation of an eigenvalue had been completed.

REMARK 4.3. Further numerical examples for the Jacobi-Davidson method and the Nonlinear Arnoldi method are shown in Example 12.3, where we considered a quantum dot problem of dimension $96,640$ with preconditioning and restarts, and in Example 14.4 of dimension $67,616$, applying AMLS to a fluid-solid vibration problem.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

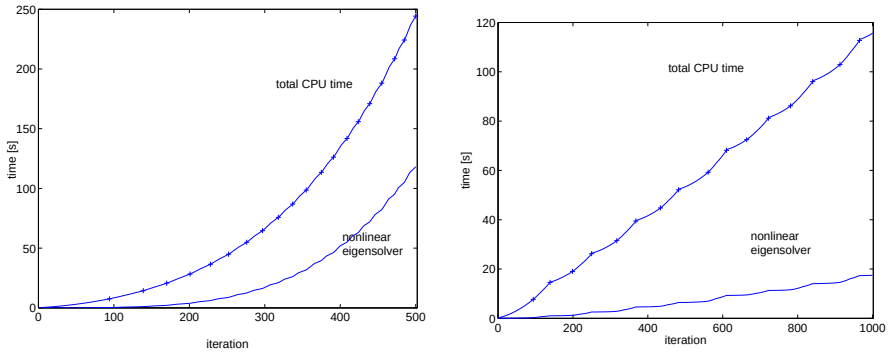


FIG. 4.1. *Time consumption of the Nonlinear Arnoldi method without (left) and with (right) restarts [244].*

5. Detecting hyperbolic and definite eigenproblems. Hyperbolic or, more generally, definite matrix polynomials are important classes of Hermitian matrix polynomials. They allow for a definite linearization and can therefore be solved by a standard algorithm for Hermitian matrices like the QR algorithm [85, 86, 103, 104]. They have only real eigenvalues, which can be characterized as minmax and maxmin values of Rayleigh functionals. In this section we present an easy way to test whether a given matrix polynomial is hyperbolic or definite or not [183]. We consider here only quadratic hyperbolic problems. Definite problems and polynomial eigenproblems of higher degree are discussed in [183].

5.1. Quadratic hyperbolic problems. For quadratic hyperbolic pencils Higham, Tisseur, and Van Dooren [104] proposed a method for testing hyperbolicity and constructing a definite linearization. Another method for detecting if a Hermitian quadratic matrix polynomial is hyperbolic, which is based on cyclic reduction, was introduced by Guo and Lancaster [87] and accelerated by Guo, Higham, and Tisseur [85]. Another method based on an improved arc algorithm for a Hermitian linearization of the quadratic pencil is studied in [86].

DEFINITION 5.1. *The quadratic matrix pencil*

$$Q(\lambda) := \lambda^2 A + \lambda B + C$$

with Hermitian matrices $A, B, C \in \mathbb{C}^{n\times n}$ is called hyperbolic if A is positive definite, and, for every $x \in \mathbb{C}^n$, $x \neq 0$, the quadratic polynomial

$$f(\lambda; x) := \lambda^2 x^H A x + \lambda x^H B x + x^H C x = 0$$

has two distinct real roots

$$p_\pm(x) = \frac{1}{2x^H A x}\left(-x^H B x \pm \sqrt{(x^H B x)^2 - 4(x^H A x)(x^H C x)}\right).$$

A hyperbolic quadratic matrix polynomial $Q(\cdot)$ has the following properties (cf. [168]): the ranges $\tilde{J}_\pm := p_\pm(\mathbb{C}^n \setminus \{0\})$ are disjoint real closed intervals with $\max \tilde{J}_- < \min \tilde{J}_+$ (this was proved by Duffin [56] for the overdamped case, and this is true for hyperbolic problems as well since the shifted pencil $Q(\lambda + \theta)$ is overdamped for sufficiently large θ), $Q(\lambda)$ is positive definite for $\lambda < \min \tilde{J}_-$ and $\lambda > \max \tilde{J}_+$, and it is negative definite for $\lambda \in (\max \tilde{J}_-, \min \tilde{J}_+)$.

Let J_- and J_+ be open intervals with $\tilde{J}_- \subset J_-$ and $\tilde{J}_+ \subset J_+$, respectively, with $J_- \cap J_+ = \emptyset$. Each of the intervals J_- and J_+ contains n eigenvalues

$$\lambda_n^- \leq \lambda_{n-1}^- \leq \cdots \leq \lambda_1^- < \lambda_1^+ \leq \cdots \leq \lambda_n^+$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

(notice that in J_-, the sign condition (A_2) from Section 2 is satisfied for $-Q(\lambda)$, and therefore the smallest eigenvalue is an nth eigenvalue), which can be characterized by (cf. Duffin [56])

$$\lambda_j^- = \max_{\dim V=j} \min_{x\in V, x\neq 0} p_-(x), \qquad \lambda_j^+ = \min_{\dim V=j} \max_{x\in V, x\neq 0} p_+(x).$$

The safeguarded iteration (cf. Algorithm 4) for λ_1^+ and for λ_1^- converges globally and monotonically decreasing and increasing, respectively, for every initial vector $x_0 \in \mathbb{C}^n \setminus \{0\}$.

This suggests the following Algorithm 9 for detecting whether a quadratic matrix polynomial is hyperbolic or not. In the upper sweep (lines 1–19), we determine sequences x_k and $\sigma_k := p_+(x_k)$ by the safeguarded iteration for p_+ aiming at λ_1^+, which is terminated if a discriminant $d(x_k) = (x_k^H B x_k)^2 - 4(x_k^H A x_k)(x_k^H C x_k)$ is negative (indicating that $Q(\lambda)$ is not hyperbolic) or a parameter μ is found such that $Q(\mu) < 0$, indicating that $Q(\lambda)$ is hyperbolic.

If the relative distance of σ_k and σ_{k-1} becomes very small and hyperbolicity is not disclosed in line 15, then we determine in the lower sweep (lines 20–35) sequences x_k and $\omega_k = p_-(x_k)$ by the safeguarded iteration for p_- aiming at λ_1^-. If there is a clear gap between J_- and J_+, i.e., if $Q(\lambda)$ is hyperbolic, then the matrix $Q(\mu)$, $\mu := 0.5(\min_j \sigma_j + \omega_k)$, will turn out to be negative definite after a few steps. However, it may happen that $\{\omega_k\}$ approaches $\min_j \sigma_j$ signalling that the gap is extremely small or even $\lambda_1^- = \lambda_1^+$.

5.2. Numerical considerations. Some remarks about Algorithm 9:

(i) Since the Rayleigh functional has similar approximation properties as the Rayleigh quotient in the linear case (i.e., an approximation x_k of an eigenvector with error $\mathcal{O}(\varepsilon)$ yields an approximation $\sigma_k = p_+(x_k)$ of the corresponding eigenvalue, the error of which satisfies $\mathcal{O}(\varepsilon^2)$), the eigenvector approximations x_k do not have to be computed very accurately.

(ii) Non-hyperbolicity is detected in lines 1, 7, and 28 if the discriminant $d(x_k)$ is negative and in lines 13 and 32 if the sequence σ_k and ω_k are not monotonically decreasing and increasing, respectively.

(iii) Hyperbolicity is detected if $Q(\lambda)$ is negative definite for some λ. If λ_1^+ is a simple eigenvalue, then the safeguarded iteration converges quadratically, and therefore (at least close to convergence), the increment $\rho_k := \sigma_{k-1} - \sigma_k$ will be greater than the error $\sigma_k - \lambda_1^+$. Moreover, ρ_k will converge to 0, and even if the gap $\lambda_1^+ - \lambda_1^-$ is small, a double step $\sigma_{k-1} + 2\rho_k = 2\sigma_k - \sigma_{k-1}$ is likely to hit the gap eventually. Therefore in line 15, the negative definiteness of $Q(\mu)$, $\mu := 2\sigma_k - \sigma_{k-1}$, is tested, which can be done by computing the Cholesky decomposition of $-Q(\mu)$. For not too small gaps between J_+ and J_-, this test often reveals that $Q(\lambda)$ is hyperbolic well before convergence of the safeguarded iteration.

(iv) Although for multiple eigenvalues the quadratic convergence of the safeguarded iteration is not proved, the double step strategy worked also fine for double eigenvalues λ_1^+ and λ_1^-; cf. [183, Example 3.3].

(v) The algorithm fails if both sequences $\{\sigma_k\}$ and $\{\omega_k\}$ converge and if their limits are very close to each other or even coincide. In the latter case, $Q(\lambda)$ is called weakly hyperbolic.

DEFINITION 5.2. *The pencil $Q(\lambda)$ is weakly hyperbolic if A, B, and C are Hermitian, A is positive definite, and*

$$\gamma := \min_{\|x\|=1} [(x^H B x)^2 - 4(x^H A x)(x^H C x)] \geq 0.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Algorithm 9 Detecting hyperbolicity.

Require: initial vector $x_0 \neq 0, \epsilon > 0$
if $d(x_0) = (x_0^H B x_0)^2 - 4(x_0^H A x_0)(x_0^H C x_0) < 0$ **then**
  STOP: $Q(\lambda)$ is not hyperbolic
end if
determine $\sigma_0 = p_+(x_0)$
for $k = 1, 2, \ldots$ until convergence **do**
  determine an eigenvector x_k of $Q(\sigma_{k-1})$ corresponding to its largest eigenvalue
  if $d(x_k) = (x_k^H B x_k)^2 - 4(x_k^H A x_k)(x_k^H C x_k) < 0$ **then**
    STOP: $Q(\lambda)$ is not hyperbolic
  end if
  determine $\sigma_k = p_+(x_k)$
  if $|(\sigma_k - \sigma_{k-1})/\sigma_k| \leq \epsilon$ **then**
    set $\sigma = \sigma_k, \omega_0 = p_-(x_k)$ and GOTO 20
  else if $\sigma_k > \sigma_{k-1}$ **then**
    STOP: $Q(\lambda)$ is not hyperbolic
  else if $Q(2\sigma_k - \sigma_{k-1})$ is negative definite **then**
    $\mu = 2\sigma_k - \sigma_{k-1}$
    STOP: $Q(\lambda)$ is hyperbolic
  end if
end for
for $k = 1, 2, \ldots$ until convergence **do**
  if $Q((\omega_{k-1} + \sigma)/2)$ is negative definite **then**
    $\mu = (\omega_{k-1} + \sigma)/2$
    STOP: $Q(\lambda)$ is hyperbolic
  else if $|(\sigma - \omega_{k-1})/\omega_{k-1}| \leq \epsilon$ **then**
    STOP: Hyperbolicity not detectable, $Q(\lambda)$ maybe weakly hyperbolic
  end if
  determine an eigenvector x_k of $Q(\omega_{k-1})$ corresponding to its largest eigenvalue
  if $d(x_k) = (x_k^H B x_k)^2 - 4(x_k^H A x_k)(x_k^H C x_k) < 0$ **then**
    STOP: $Q(\lambda)$ is not hyperbolic
  end if
  determine $\omega_k = p_-(x_k)$
  if $\omega_k < \omega_{k-1}$ **then**
    STOP: $Q(\lambda)$ is not hyperbolic
  end if
end for

REMARK 5.3. A weakly hyperbolic eigenvalue problem has $2n$ real eigenvalues, and if $\gamma = 0$ (i.e., $Q(\lambda)$ is not hyperbolic), then it holds that

$$\lambda_n^- \leq \lambda_{n-1}^- \leq \cdots \leq \lambda_1^- = \lambda_1^+ \leq \lambda_2^+ \leq \ldots \lambda_n^+.$$

Obviously, p_+ as defined in (6.3) is a Rayleigh functional of $Q(\lambda)$ with respect to the interval $\tilde{J}_+ := (\lambda_1^+, \infty)$ satisfying (A_1) and (A_2), and all eigenvalues in $\tilde{J}_+$ are minmax and maxmin values of p_+.

If $\sigma_{k-1} \in \tilde{J}_+$ and x_k is an eigenvector corresponding to the maximal eigenvalue of $Q(\sigma_{k-1})$, then $x_k^H Q(\sigma_{k-1}) x_k \geq 0$, and therefore the maximal solution σ_k of the quadratic equation $x_k^H Q(\lambda) x_k = 0$ satisfies $\sigma_k = p_+(x_k)$ or $\sigma_k = \lambda_1^+$. Hence, the safeguarded iterations either stop after a finite number of steps with $\sigma_k = \lambda_1^+$, or $\{\sigma_k\} \subset \tilde{J}_+$ is a

ETNA
Kent State University and
Johann Radon Institute (RICAM)

monotonically decreasing sequence converging to some $\hat\sigma \in \overline{\tilde J_+}$. In the latter case, we obtain in the same way as in the proof of Theorem 3.1 that $\hat\sigma = \lambda_1^+$. Likewise, the sequence $\{\omega_k\}$ constructed in the same way for the interval $(-\infty, \lambda_1^-)$ converges to λ_1^- and is monotonically increasing.

We now assume that $Q(\cdot)$ is hyperbolic and a parameter μ has been found such that $Q(\mu)$ is negative definite. Then the following transformation yields a definite linearization of $Q(\cdot)$. Shifting by μ yields a quadratic matrix polynomial

$$\tilde Q(\lambda) := Q(\lambda+\mu) = \lambda^2 A + \lambda(B+2\mu A) + (C + \mu^2 A + \mu B) =: \lambda^2 A + \lambda\tilde B + \tilde C,$$

where $\tilde C = \tilde Q(0) = Q(\mu)$ is negative definite and the well-known linearizations [85, 137]

$$L_1(\lambda) := \lambda \begin{bmatrix} A & 0 \\ 0 & -\tilde C \end{bmatrix} + \begin{bmatrix} \tilde B & \tilde C \\ \tilde C & 0 \end{bmatrix} \quad \text{and} \quad L_2(\lambda) := \lambda \begin{bmatrix} 0 & A \\ A & \tilde B \end{bmatrix} + \begin{bmatrix} -A & 0 \\ 0 & \tilde C \end{bmatrix} \tag{5.1}$$

of $\tilde Q(\lambda)$ are obviously definite. By employing the Cholesky factorization of $\operatorname{diag}\{A, -\tilde C\}$, it can be transformed to a standard eigenvalue problem and solved by the QR algorithm preserving the reality of its eigenvalues.

EXAMPLE 5.4. To compare our method to the cyclic reduction of Guo, Higham, and Tisseur [85], we use the following method for constructing quadratic matrix polynomials with prescribed eigenvalues and eigenvectors (cf. [138]): For (λ_j, v_j), $j = 1, \dots, 2n$, let

$$\begin{aligned} \Lambda_1 &:= \operatorname{diag}\{\lambda_1, \dots, \lambda_n\}, & \Lambda_2 &:= \operatorname{diag}\{\lambda_{n+1}, \dots, \lambda_{2n}\}, \\ V_1 &:= [v_1, \dots, v_n], & V_2 &:= [v_{n+1}, \dots, v_{2n}] \in \mathbb{R}^{n\times n}. \end{aligned}$$

Assume that V_1 and V_2 are nonsingular, $V_1V_1^T = V_2V_2^T$, and $\Gamma := V_1\Lambda_1V_1^T - V_2\Lambda_2V_2^T$ is nonsingular. Then the quadratic polynomial $Q(\lambda)$ with

$$\begin{aligned} A &= \Gamma^{-1}, & B &= -A(V_1\Lambda_1^2V_1^T - V_2\Lambda_2^2V_2^T)A, \\ C &= -A(V_1\Lambda_1^3V_1^T - V_2\Lambda_2^3V_2^T)A + B\Gamma B \end{aligned}$$

has eigenpairs (λ_j, v_j), $j = 1, \dots, 2n$.

We constructed a test set of 80 quadratic matrix functions $Q(\lambda) \in \mathbb{R}^{500\times 500}$ of this type, where λ_j, for $j = 1, \dots, 500$, are normally distributed with mean value -3 and standard variation 1, and, for $j = 501, \dots, 1000$, λ_j are uniformly distributed in $[-106, -6]$. If $\lambda_{\max} := \max_j \lambda_j > 0$, then the eigenvalues λ_j were shifted to the left by $1.1\lambda_{\max}$ (then all eigenvalues λ_j are negative, and the hyperbolic examples are even overdamped [85]; this is not needed in Algorithm 9 but is only used to compare it to the cyclic reduction in [85]). With random orthogonal matrices U_1, U_2 we chose $V_1 = U_1$ and $V_2 = V_1U_2$ so that $V_1V_1^T = V_2V_2^T$. For 51 of these examples $\max_{j=501,\dots,1000} \lambda_j < \min_{j=1,\dots,500} \lambda_j$, and the corresponding $Q(\lambda)$ are hyperbolic; actually they are overdamped; cf. [85]. For the remaining 29 problems, the matrix A turned out to be positive definite, but $Q(\lambda)$ was not hyperbolic.

Algorithm 9 detected the type of $Q(\lambda)$ in all examples correctly. The average CPU time on a Pentium D Computer with 3.2 GHz and 2 GB RAM was 0.65 seconds, minimal 0.47, and maximal 1.05 seconds. The safeguarded iteration required at least 2 steps, at most 3 steps, and the average number of steps was 2.06. The Nonlinear Arnoldi method constructed search spaces of minimal dimension 31, maximal dimension 67, and the average dimension was 44.8. Since we allowed for a maximal dimension of 100 of the search spaces, no restarts were necessary. Also the cyclic reduction algorithm of Guo, Higham, and Tisseur [85] detected the type of the pencils in all cases correctly. It required at most 23 iterations and at least no

ETNA
Kent State University and
Johann Radon Institute (RICAM)

iteration with an average of 8.95, and the average CPU time was 2.55 seconds, with minimal 0.27 and maximal 6.19 seconds. In 61 (resp. 28 non-hyperbolic) examples the safeguarded iteration was faster, whereas in 19 (resp. 1 non-hyperbolic) examples the cyclic reduction was the winner.

5.3. Definite polynomials. In a recent paper Higham, Mackey, and Tisseur [103] generalized the concept of hyperbolic quadratic polynomials, waiving the positive definiteness of the leading matrix A.

DEFINITION 5.5. *The quadratic matrix polynomial*

$$Q(\lambda) := \lambda^2 A + \lambda B + C$$

is definite if $A = A^H$, $B = B^H$, $C = C^H$ are Hermitian, there exists $\mu \in \mathbb{R} \cup \{\infty\}$ such that $Q(\mu)$ is positive definite, and for every fixed $x \neq 0$, the real equation

$$f(\lambda; x) := \lambda^2 x^H A x + \lambda x^H B x + x^H C x = 0$$

has two distinct roots in $\mathbb{R} \cup \{\infty\}$.

The following theorem was proved in [103]:

THEOREM 5.6 ([103]). *The Hermitian matrix polynomial $Q(\lambda)$ is definite if and only if any two (and hence all) of the following properties hold:*

(i) $d(x) := (x^H B x)^2 - 4(x^H A x)(x^H C x) > 0$ *for every* $x \in \mathbb{C}^n \setminus \{0\}$.
(ii) $Q(\eta) > 0$ *for some* $\eta \in \mathbb{R} \cup \{\infty\}$.
(iii) $Q(\xi) < 0$ *for some* $\xi \in \mathbb{R} \cup \{\infty\}$.

Hence, to detect that a pencil is definite, one has to find $\xi, \eta \in \mathbb{R} \cup \{\infty\}$ such that $Q(\xi) < 0 < Q(\eta)$. The article [183] presents an approach for this task, which is again based on the safeguarded iteration. The only additional problem is that one does not know in advance whether $\xi < \eta$ or $\eta < \xi$. The paper [183] discusses these two cases and demonstrates that by taking advantage of the safeguarded iteration, one can decide safely whether the problem is definite or not.

6. Sylvester's law of inertia. The inertia of a Hermitian matrix A is the triplet of nonnegative integers $\mathrm{In}(A) := (n_p, n_n, n_z)$, where n_p, n_n, and n_z are the number of positive, negative, and zero eigenvalues of A counting multiplicities. Sylvester's classical law of inertia [218] states that two Hermitian matrices $A, B \in \mathbb{C}^{n\times n}$ are congruent (i.e., $A = S^H B S$ for some nonsingular matrix $S \in \mathbb{C}^{n\times n}$) if and only if they have the same inertia $\mathrm{In}(A) = \mathrm{In}(B)$.

An obvious consequence of the law of inertia is the following corollary.

COROLLARY 6.1. *If A has an LDL^H factorization $A = LDL^H$, then n_p and n_n equals the number of positive and negative entries of D, respectively, and if only a block LDL^H factorization exists, where D is a block diagonal matrix with 1×1 and indefinite 2×2 blocks on its diagonal, then one has to increase the number of positive and negative 1×1 blocks of D by the number of 2×2 blocks to get n_p and n_n, respectively.*

Hence, the inertia of A can be computed easily, and this is particularly advantageous if the matrix is banded. If $B \in \mathbb{C}^{n\times n}$ is positive definite and $A - \sigma B = LDL^H$ is the block diagonal LDL^H factorization of $A - \sigma B$ for some $\sigma \in \mathbb{R}$, from which we get the inertia $\mathrm{In}(A - \sigma B) = (n_p, n_n, n_z)$ as described in the last corollary, then the generalized eigenvalue problem $Ax = \lambda Bx$ has n_n eigenvalues smaller than σ. Hence, the law of inertia yields a tool to locate eigenvalues of Hermitian matrices or definite matrix pencils. Combining it with bisection or the secant method, one can determine all eigenvalues in a given interval or initial approximations for fast eigensolvers, and it can be used to test whether a method has found all eigenvalues in an interval of interest or not.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

The law of inertia was first proved in 1858 by J. J. Sylvester [218], and several different proofs can be found in the literature [53, 80, 108, 181, 190], one of which is based on the minmax characterization of eigenvalues of Hermitian matrices. Here we discuss generalizations of the law of inertia to nonlinear eigenvalue problems allowing for a minmax characterization of its eigenvalues.

The following location result for real eigenvalues, generalizing Sylvester's law of inertia [218], was proved in [118].

THEOREM 6.2 ([118]). *Assume that $T : J \to \mathbb{R}^{n\times n}$ satisfies the conditions of the minmax characterization in Theorem 2.2.*

(i) Let $T(\cdot)$ be overdamped. For $\sigma \in J$, let (π, ν, δ) be the inertia of $T(\sigma)$. Then $T(\cdot)$ has π eigenvalues that are smaller than σ, ν eigenvalues that exceed σ, and if $\delta \neq 0$, then $\sigma = \lambda_{\pi+1} = \cdots = \lambda_{\pi+\delta}$ is an eigenvalue of geometric multiplicity δ.

(ii) Assume that $T(\mu)$ is negative definite for some $\mu \in J$, and for $\sigma > \mu$, let (π, ν, δ) be the inertia of $T(\sigma)$. Then $T(\cdot)$ has exactly π eigenvalues $\lambda_1 \leq \cdots \leq \lambda_\pi$ in J that are smaller than σ.

(iii) Let $\mu \in J$, and assume that for every r-dimensional subspace $V \subset \mathbb{R}^n$ with $V \cap \mathcal{D}(p) \neq \emptyset$, there exists $x \in V \cap \mathcal{D}(p)$ with $p(x) > \mu$. For $\sigma \in J$, $\sigma > \mu$, let (π, ν, δ) be the inertia of $T(\sigma)$. Then for $j = r, \ldots, \pi$, there exists a jth eigenvalue λ_j of $T(\cdot)$ in $[\mu, \sigma)$.

REMARK 6.3. Without using the minmax characterization of eigenvalues, Neumaier [181] proved part (iii) of Theorem 6.2 for matrices $T : J \to \mathbb{C}^{n\times n}$ that are Hermitian and (elementwise) differentiable in J with positive definite derivative $T'(\lambda)$, $\lambda \in J$. Obviously, such $T(\cdot)$ satisfies the conditions of the minmax characterization. Further location results generalizing Theorem 6.2 were obtained by Y. Nakatsukasa and V. Noferini [179] without using variational characterizations.

EXAMPLE 6.4. Consider the rational eigenvalue problem

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

where $K, M \in \mathbb{R}^{n\times n}$ are symmetric and positive definite, $C_j \in \mathbb{R}^{n\times k_j}$ has rank k_j, and $0 < \sigma_1 < \cdots < \sigma_p$, which models the free vibrations of certain fluid-solid structures; cf. [49].

In each interval $J_\ell := (\sigma_\ell, \sigma_{\ell+1})$, $\ell = 0, \ldots, p$, $\sigma_0 = 0$, $\sigma_{p+1} = \infty$, the function $f_\ell(\lambda, x) := x^T T(\lambda)x$ is strictly monotonically increasing, and therefore all eigenvalues in J_ℓ are minmax values of the Rayleigh functional p_ℓ corresponding to J_ℓ.

For the first interval J_0, item (ii) in Theorem 6.2 applies. Hence, if $\tau \in J_0$ and (n_p, n_n, n_z) is the inertia of $T(\tau)$, then there are exactly n_p eigenvalues in J_0 which are less than τ. Moreover, if $\tau_1 < \tau_2$ are contained in one interval J_j, then the number of eigenvalues in the interval (τ_1, τ_2) can be obtained from the inertia of $T(\tau_1)$ and $T(\tau_2)$ according to Theorem 6.2, item (iii).

6.1. Quadratic eigenvalue problems. We now apply Sylvester's law of inertia to quadratic matrix pencils

$$Q(\lambda) := \lambda^2 A + \lambda B + C, \tag{6.1}$$

with Hermitian matrices $A, B, C \in \mathbb{C}^{n\times n}$, where additional conditions guarantee that (some of) its real eigenvalues allow for a variational characterization and hence for a slicing of its spectrum using the inertia. Some of the presented results can be generalized to polynomials of higher degree.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

6.1.1. $C < 0$ and $A \geq 0$. Let C be negative definite and A positive semidefinite. Multiplying $Q(\lambda)x = 0$ by λ^{-1}, one gets the equivalent nonlinear eigenvalue problem

$$\tilde{Q}(\lambda)x := \lambda Ax + Bx + \lambda^{-1}Cx = 0. \tag{6.2}$$

Differentiating $f(\lambda; x) := x^H \tilde{Q}(\lambda)x$ with respect to λ yields

$$\frac{\partial}{\partial \lambda} f(\lambda; x) = x^H Ax - \lambda^{-2} x^H Cx > 0, \quad \text{for every } x \neq 0 \text{ and every } \lambda \neq 0.$$

Hence, the pencil $\tilde{Q}$ satisfies the conditions of the minmax characterization for both intervals $J_- := (-\infty, 0)$ and $J_+ := (0, \infty)$. For the corresponding Rayleigh functional $p_\pm$ with domain $\mathcal{D}_\pm$, it holds that $\lambda_1^+ = \inf_{x \in \mathcal{D}_+} p_+(x) \in J_+$ and $\lambda_n^- = \sup_{x \in \mathcal{D}_-} p_-(x) \in J_-$, and therefore the following location result follows from Theorem 6.2.

THEOREM 6.5 ([118]). *Let C be negative definite and A positive semidefinite.*

(i) For $\sigma > 0$, let $\mathrm{In}(\tilde{Q}(\sigma)) = (n_p, n_n, n_z)$ be the inertia of $\tilde{Q}(\sigma)$. Then the quadratic pencil (6.1) has n_p positive eigenvalues smaller than σ.

(ii) For $\sigma < 0$, let $\mathrm{In}(\tilde{Q}(\sigma)) = (n_p, n_n, n_z)$ be the inertia of $\tilde{Q}(\sigma)$. Then problem (6.2) has n_n negative eigenvalues larger than σ.

If A is positive definite, then $\tilde{Q}$ is overdamped with respect to J_+ and J_-, and therefore there exist exactly n positive and n negative eigenvalues. If $A \neq 0$ is positive semidefinite and $r = \mathrm{rank}(A)$, then ∞ is an infinite eigenvalue of multiplicity $n - r$, and there are only $n + r$ finite eigenvalues.

If B is positive definite, then the Rayleigh functional

$$p_+(x) = -2\frac{x^H Cx}{x^H Bx + \sqrt{(x^H Bx)^2 - 4(x^H Ax)(x^H Cx)}}$$

is defined on $\mathbb{C}^n \setminus \{0\}$. Hence, $(\tilde{Q}, J_+)$ is overdamped, and there exist n positive and r negative eigenvalues. Theorem 6.5 can be strengthened according to:

THEOREM 6.6 ([118]). *Assume that A is positive semidefinite, B is positive definite, and C is negative definite.*

(i) For $\sigma > 0$, let $\mathrm{In}(\tilde{Q}(\sigma)) = (n_p, n_n, n_z)$ be the inertia of $\tilde{Q}(\sigma)$. Then the quadratic pencil (6.1) has n_p positive eigenvalues less than σ, n_n finite eigenvalues exceeding σ, and if $n_z \neq 0$, then σ is an eigenvalue of $Q(\cdot)$ with multiplicity n_z.

(ii) For $\sigma < 0$, let $\mathrm{In}(\tilde{Q}(\sigma)) = (n_p, n_n, n_z)$ be the inertia of $\tilde{Q}(\sigma)$. Then (6.1) has n_n negative eigenvalues exceeding σ, $n_p - r$ eigenvalues smaller than σ, and if $n_z \neq 0$, then σ is an eigenvalue of $Q(\cdot)$ with multiplicity n_z.

REMARK 6.7. In [118] we discussed how to use Theorem 6.2 to slice the spectra of hyperbolic polynomial eigenvalue problems and in particular hyperbolic quadratic eigenvalue problems.

6.1.2. Nonoverdamped quadratic pencils. We consider the quadratic pencil (6.1), where A, B, and C are positive definite. Then, for $x \neq 0$, the two complex roots of the function $f(\lambda; x) := x^H Q(\lambda)x$ are

$$p_\pm(x) = \frac{1}{2x^H Ax}\left(-x^H Bx \pm \sqrt{(x^H Bx)^2 - 4(x^H Ax)(x^H Cx)}\right). \tag{6.3}$$

Let

$$\delta_- := \sup\{p_-(x) : p_-(x) \in \mathbb{R}\}, \qquad \delta_+ := \inf\{p_+(x) : p_+(x) \in \mathbb{R}\},$$
$$J_- := (-\infty, \delta_+), \qquad J_+ = (\delta_-, 0), \quad \text{and} \quad D_\pm := \{x \in \mathbb{C}^n : p_\pm(x) \in J_\pm\}.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

If $f(\lambda,x) > 0$, for $x \neq 0$, and $\lambda \in \mathbb{R}$, then $\delta_- = -\infty$ and $\delta_+ = \infty$, and the eigenvalue problem $Q(\lambda)x = 0$ has no real eigenvalues, but this does not need to be known in advance. Theorem 6.8 below applies to this case as well.

It is obvious that $-Q$ and Q both satisfy the conditions of the minmax characterization of their eigenvalues in J_- and J_+, respectively. Hence, all eigenvalues in J_- are minmax values of p_-

$$\lambda_j^- = \min_{\dim V=j,\ V\cap\mathcal{D}_-\neq\emptyset}\ \max_{x\in V\cap\mathcal{D}_-} p_-(x), \qquad j = 1,2,\dots$$

Taking advantage of the minmax characterization of the eigenvalues of $\tilde{Q}(\lambda) := -Q(-\lambda)$ in $\tilde{J} := J_+$ with the Rayleigh functional $\tilde{p} := -p_+$, we obtain the following maxmin characterization

$$\lambda_{2n+1-j}^+ = \max_{\dim V=j,\ V\cap\mathcal{D}_+\neq\emptyset}\ \min_{x\in V\cap\mathcal{D}_+} p_+(x), \qquad j = 1,2,\dots$$

of all eigenvalues of Q in J_+.

Hence, for $\sigma < \delta_+$ and for $\sigma > \delta_-$, we obtain slicing results for the spectrum of $Q(\cdot)$ from Theorem 6.2 (ii). If $\mathrm{In}(Q(\sigma)) = (n_p, n_n, n_z)$ and $\sigma < \delta_+$, then there exist n_n eigenvalues of $Q(\cdot)$ in $(-\infty,\sigma)$, and if $\sigma \in (\delta_-, 0)$, then there are n_n eigenvalues in $(\sigma, 0)$. However, δ_+ and δ_- are usually not known. The following theorem contains upper bounds of δ_- and lower bounds of δ_+, thus yielding subintervals of $(-\infty,\delta_+)$ and $(\delta_-, 0)$, where the above slicing applies.

THEOREM 6.8 ([118]). *Let $A, B, C \in \mathbb{C}^{n\times n}$ be positive definite, and let p_+ and p_- be defined in (6.3). Then it holds that*

(i)

$$\tilde{\delta}_+ := -\sqrt{\max_{x\neq 0}\frac{x^HCx}{x^HAx}} \le \delta_+ = \inf\{p_+(x) \ :\ p_+(x)\in\mathbb{R}\} \tag{6.4}$$

and

$$\delta_- = \sup\{p_-(x) \ :\ p_-(x)\in\mathbb{R}\} \le -\sqrt{\min_{x\neq 0}\frac{x^HCx}{x^HAx}} =: \tilde{\delta}_-. \tag{6.5}$$

(ii)

$$\hat{\delta}_+ := -2\max_{x\neq 0}\frac{x^HCx}{x^HBx} \le \delta_+ \quad \textit{and} \quad \delta_- \le -2\min_{x\neq 0}\frac{x^HCx}{x^HBx} =: \hat{\delta}_-. \tag{6.6}$$

EXAMPLE 6.9. The matrices A, B, and C were obtained by the following MATLAB expressions:

```
randn('state',0); A = eye(30); B = randn(30); C = randn(30);
                  B = B'*B; C = C'*C;
```

Then $Q(\lambda)x = 0$ has 36 real eigenvalues, 18 in the domain of p_- and 18 in the domain of p_+. So, Sylvester's law of inertia can be applied to all of them.

From Theorem 6.8 we obtain the following information: For both intervals the bounds from the estimates given in (6.4) and (6.5) are stronger than the ones from (6.6). 16 eigenvalues are contained in $(-\infty,\tilde{\delta}_+) = (-\infty, -\sqrt{\max(\lambda(C,A))}) = (-\infty, -10.5022)$ and 8 eigenvalues in $(\tilde{\delta}_-, 0) = (-\sqrt{\min(\lambda(C,A))}, 0) = (-0.3103, 0)$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

7. Low-rank modifications of symmetric eigenvalue problems. We consider a nonlinear low-rank modification of a symmetric eigenvalue problem

$$(A + \phi(\lambda)H)x = \lambda x, \tag{7.1}$$

where $A, H \in \mathbb{C}^{n\times n}$ are Hermitian matrices, H has low rank $k \ll n$, and ϕ is real-valued and continuous. We denote by $\alpha_1 \le \cdots \le \alpha_n$ the eigenvalues of A, and set $\alpha_j = -\infty$ for $j < 1$ and $\alpha_j = \infty$ for $j > n$.

7.1. Rank-1 modifications. We first consider the case $H = cc^H$, i.e., rank-1 modifications and generalize methods obtained by Huang, Bai, and Su in [109]. For constant modifications $B := A+\tau cc^H$ it is well known (cf. [80]) that for the eigenvalues $\beta_1 \le \cdots \le \beta_n$ of the matrix B, the following interlacing properties hold:

$$\begin{array}{llll} \text{If } \tau > 0, & \text{then} & \alpha_1 \le \beta_1 \le \alpha_2 \le \beta_2 \le \cdots \le \alpha_n \le \beta_n, & \text{and} \\ \text{if } \tau < 0, & \text{then} & \beta_1 \le \alpha_1 \le \beta_2 \le \alpha_2 \le \cdots \le \beta_n \le \alpha_n. & \end{array}$$

In case A has n distinct eigenvalues and $c \neq 0$ is not orthogonal to any of the corresponding eigenvectors, then, for $\tau > 0$, exactly one eigenvalue β_k is in each of the intervals (α_k, α_{k+1}), and, for $\tau < 0$, there is exactly one eigenvalue of B in (α_{k-1}, α_k), for $k = 1, \ldots, n$.

The following existence and inclusion results are contained in [250]. We consider only nonnegative modifications; the analogous results for nonpositive modifications are clear. For nonlinear rank-one modifications we immediately get the following theorem:

THEOREM 7.1 ([250]). *For $k \in \{1, \ldots, n-1\}$, let $\phi \in C[\alpha_k, \alpha_{k+1}]$ be nonnegative. Then the nonlinear eigenvalue problem*

$$(A + \phi(\lambda)cc^H)x = \lambda x \tag{7.2}$$

has an eigenvalue $\hat\lambda \in [\alpha_k, \alpha_{k+1}]$.

REMARK 7.2. Notice that differently from a constant rank-one modification of A, there does not necessarily exist an eigenvalue of (7.2) in $[\alpha_n, \infty)$ for $\phi \ge 0$. If c is an eigenvector of A corresponding to α_n, then $(A+\phi(\lambda)cc^T)c = (\alpha_n+\phi(\lambda)\|c\|^2)c$, and $\alpha_n+\phi(\lambda)\|c\|^2 \ge \alpha_n$ is the maximal eigenvalue of

$$(A + \phi(\lambda)cc^H)x = \mu x. \tag{7.3}$$

Hence, if $\phi(\lambda)\|c\|^2 > \lambda - \alpha_n$ for every $\lambda \ge \alpha_n$, then there does not exist an eigenvalue of (7.2) in $[\alpha_n, \infty)$.

The following theorem assures the existence of an eigenvalue in the extreme interval $[\alpha_n, \infty)$:

THEOREM 7.3. *Let $\phi \in C[\alpha_n, \infty)$ be nonnegative, and assume that there exists $\eta \ge \alpha_n$ such that for some $\delta > 0$ it holds that*

$$\frac{\phi(\lambda) - \phi(\eta)}{\lambda - \eta}\|c\|^2 \le 1 - \delta \qquad \textit{for every } \lambda > \eta. \tag{7.4}$$

Then the nonlinear eigenvalue problem (7.2) has an eigenvalue $\hat\lambda \in [\alpha_n, \infty)$, which is the largest eigenvalue of (7.3) with $\lambda = \hat\lambda$.

Notice that the condition (7.4) cannot be relaxed to $\delta = 0$. Choosing the function $\phi(\lambda) = (\lambda - \alpha_n + \exp(-\lambda))/\|c\|^2$ yields that the maximum eigenvalue $\mu_n = \lambda + \exp(-\lambda)$ of (7.3) has no fixed point in $[\alpha_n, \infty)$.

The uniqueness theorem also holds for the unbounded interval (α_n, ∞).

ETNA
Kent State University and
Johann Radon Institute (RICAM)

THEOREM 7.4 ([250]). *For $k \in \{1,\dots,n\}$, let $\phi \in C[\alpha_k, \alpha_{k+1}]$ be nonnegative, and assume that the condition*

$$\frac{\phi(\lambda)-\phi(\tilde\lambda)}{\lambda-\tilde\lambda}\|c\|^2 < 1$$

is satisfied for $\lambda, \tilde\lambda \in I := (\alpha_k, \alpha_{k+1})$, $\lambda \neq \tilde\lambda$. Then the problem (7.2) has at most one eigenvalue $\hat\lambda \in I$.

7.2. Low-rank modifications. We now consider low-rank modifications (7.1), where $A, H \in \mathbb{R}^{n\times n}$ are symmetric, H has low rank, and $\tau > 0$. Again we denote the eigenvalues of A by $\alpha_1 \le \cdots \le \alpha_n$, and we set $\alpha_j = \infty$ for $j > n$ and $\alpha_j = -\infty$ for $j < 1$. The inertia of H is denoted by (π, ν, ζ) and its eigenvalues by $\sigma_1 \le \cdots \le \sigma_\nu < 0 < \sigma_{n-\pi+1} \le \cdots \le \sigma_n$. The following bounds for the eigenvalues of $A + \tau H$ are due to Weyl [256] and can be found in the book of Parlett [190, Corollary 10.3.1].

THEOREM 7.5. *Let $\beta_1 \le \cdots \le \beta_n$ denote the eigenvalues of $B := A + \tau H$. Then for $\tau > 0$ it holds that*

$$\alpha_{i-\nu} \le \beta_i \le \alpha_{i+\pi}, \qquad \text{for } i = 1,\dots,n.$$

From this we immediately get the existence result for the nonlinear modification of A:

THEOREM 7.6. *For $k \in \{1+\nu,\dots,n-\pi\}$, let $\phi \in C[\alpha_{k-\nu}, \alpha_{k+\pi}]$ be nonnegative. Then the nonlinear eigenvalue problem*

$$(A + \phi(\lambda)H)x = \lambda x$$

has an eigenvalue $\hat\lambda \in [\alpha_{k-\nu}, \alpha_{k+\pi}]$, which is the kth-smallest eigenvalue of the linear eigenvalue problem

$$(A + \phi(\lambda)H)x = \mu x \tag{7.5}$$

with $\lambda = \hat\lambda$.

From the minmax characterization of eigenvalues one immediately obtains the following inequalities:

LEMMA 7.7 ([250]). *Let $\mu_k(\lambda)$ be the kth-smallest eigenvalue of (7.5), and denote by $\sigma_1 \le \sigma_2 \le \cdots \le \sigma_n$ the eigenvalues of H.*

(i) If $\phi(\lambda) \ge 0$, then it holds that

$$\alpha_1 + \phi(\lambda)\sigma_k \le \mu_k(\lambda) \le \alpha_n + \phi(\lambda)\sigma_k, \qquad k = 1,\dots,n.$$

(ii) For $\phi(\lambda) \ge \phi(\tilde\lambda)$, we have

$$\mu_k(\lambda) \le \mu_k(\tilde\lambda) + (\phi(\lambda)-\phi(\tilde\lambda))\sigma_n, \qquad \text{and}$$
$$\mu_k(\tilde\lambda) \le \mu_k(\lambda) - (\phi(\lambda)-\phi(\tilde\lambda))\sigma_1.$$

The following theorem guarantees the existence of eigenvalues of (7.1) in the extreme intervals.

THEOREM 7.8 ([250]).

(i) For $1 \le k \le \nu$, let $\phi \in C(-\infty, \alpha_{k+\pi}]$ be nonnegative, and assume that there exists $\eta \le \alpha_1$ such that for some $\delta > 0$ it holds that

$$\frac{\phi(\lambda)-\phi(\eta)}{\lambda-\eta}\sigma_k \le 1-\delta,$$

for every $\lambda < \eta$. Then the nonlinear eigenvalue problem (7.1) has an eigenvalue $\hat\lambda \in (-\infty, \alpha_{k+\pi}]$, which is the kth-smallest eigenvalue of (7.5) with $\lambda = \hat\lambda$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

(ii) For $n - \pi < k \le n$, let $\phi \in C[\alpha_{k-\nu}, \infty)$ be nonnegative, and assume that there exists $\eta \ge \alpha_n$ such that for some $\delta > 0$ it holds that

$$\frac{\phi(\lambda) - \phi(\eta)}{\lambda - \eta}\sigma_k \le 1 - \delta,$$

for every $\lambda > \eta$. Then the nonlinear eigenvalue problem (7.1) has an eigenvalue $\hat{\lambda} \in [\alpha_{k-\nu}, \infty)$, which is the kth-smallest eigenvalue of (7.5) with $\lambda = \hat{\lambda}$.

The uniqueness result obtains the following form:

THEOREM 7.9 ([250]). *For $k \in \{1, \dots, n\}$, let $\phi \in C[\alpha_{k-\nu}, \alpha_{k+\pi}]$ be nonnegative, and assume that*

$$\max\left(\frac{\phi(\lambda) - \phi(\tilde{\lambda})}{\lambda - \tilde{\lambda}}\sigma_n, \frac{\phi(\lambda) - \phi(\tilde{\lambda})}{\lambda - \tilde{\lambda}}\sigma_1\right) < 1, \qquad \lambda \ne \tilde{\lambda}, \tag{7.6}$$

holds in $[\alpha_{k-\nu}, \alpha_{k+\pi}]$.

Then problem (7.1) has at most one eigenvalue $\hat{\lambda} \in (\alpha_{k-\nu}, \alpha_{k+\pi})$, which is the kth-smallest eigenvalue of (7.5) with $\lambda = \hat{\lambda}$.

The global behavior of the spectrum of the nonlinear eigenvalue problem (7.1) is described in the next theorem.

THEOREM 7.10 ([250]). *Let $\phi : \mathbb{R} \to \mathbb{R}$ be a continuous and nonnegative function such that condition (7.6) holds for all $\lambda, \tilde{\lambda} \in \mathbb{R}$ with $\lambda \ne \tilde{\lambda}$, and assume that the conditions of Theorem 7.8 are satisfied.*

Then the nonlinear eigenvalue problem

$$(A + \phi(\lambda)H)x = \lambda x$$

has exactly n eigenvalues $\lambda_1 \le \cdots \le \lambda_n$, and $\lambda_k \in [\alpha_{k-\nu}, \alpha_{k+\pi}]$, $k = 1, \dots, n$.

The interval $[\alpha_k, \alpha_{k+1}]$ contains at most $\nu + \pi + 2$ eigenvalues λ_j, where it holds that $j \in \{k - \pi, k - \pi + 1, \dots, k + \nu, k + \nu + 1\}$.

7.3. Numerical considerations. We now consider numerical methods that are variants of safeguarded iterations for general nonlinear eigenvalue problems. Under the conditions of the minmax characterization (which are not necessarily satisfied here), it was shown in [183] that the safeguarded iteration converges quadratically, and the convergence is global for extreme eigenvalues but only local for the remaining ones. For the low-rank modifications of symmetric problems considered here, safeguarding can be included to guarantee global convergence for the interior eigenvalues.

Methods for rank-1 modifications are discussed in [250]. We only consider methods for computing a kth eigenvalue $\hat{\lambda}$ and a corresponding eigenvector $\hat{x}$ of

$$(A + \phi(\lambda)H)x = \lambda x,$$

i.e., an eigenvalue $\hat{\lambda}$ which is the kth-smallest eigenvalue of the linear problem

$$(A + \phi(\hat{\lambda})H)x = \mu x.$$

The proof of Theorem 7.9 demonstrates that for $\lambda \in I_k = (\alpha_{k-\nu}, \alpha_{k+\pi})$, it holds that

$$\lambda \le \hat{\lambda} \quad \Longleftrightarrow \quad \mu_k(\lambda) \ge \lambda,$$

which together with the following theorem yields a safeguarding.

THEOREM 7.11 ([250]). *Assume that condition (7.6) holds in an interval $I \subset \mathbb{R}$. Let*

$$f(\lambda, x) := x^H(A + \phi(\lambda)H - \lambda I)x.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

If $f(\tilde{\lambda},x)=0$ for some $\tilde{\lambda}\in I$ and $x\neq 0$, then it holds that

$$(\lambda-\tilde{\lambda})f(\lambda,x)<0 \qquad \textit{for every } \lambda\in I,\ \lambda\neq\tilde{\lambda}.$$

Hence an approximation of $\hat{\lambda}$ can be updated by the solution of $f(\lambda,x)=0$, and the following Algorithm 10 is obtained.

Algorithm 10 Determining an inner eigenpair of a low rank modified symmetric eigenproblem.

Require: initial bounds $\lambda_\ell := \alpha_{k-\nu}$, $\lambda_u := \alpha_{k+\pi}$ and initial guess $\lambda\in[\lambda_\ell,\lambda_u]$ of $\hat{\lambda}$
$\gamma = \max(0, \max_{\lambda\in[\lambda_\ell,\lambda_u]}\phi'(\lambda)\sigma_n, \max_{\lambda\in[\lambda_\ell,\lambda_u]}\phi'(\lambda)\sigma_1)$

1: determine an eigenpair (μ,x) corresponding to the kth-smallest eigenvalue of

$$(A+\phi(\lambda)H)x=\mu x$$

2: **while** $|\lambda-\mu|/(1-\gamma) >$ tol **do**
3: **if** $\mu>\lambda$ **then**
4: $\lambda_\ell=\lambda$
5: **else**
6: $\lambda_u=\lambda$
7: **end if**
8: **if** $f(\lambda_\ell,x)f(\lambda_u,x)>0$ **then**
9: $\lambda=0.5(\lambda_u+\lambda_\ell)$
10: **else**
11: solve $x^H(A+\phi(\lambda)H-\lambda I)x=0$ for λ
12: **end if**
13: determine an eigenpair (μ,x) corresponding to the kth-smallest eigenvalue of

$$(A+\phi(\lambda)H)x=\mu x$$

14: **end while**
15: $\hat{\lambda}:=\lambda$, $\hat{x}:=x$

Algorithm 10 computes an eigenpair $(\hat{\lambda},\hat{x})$ with $\hat{\lambda}$ as the kth-smallest eigenvalue of $B(\lambda):=A+\phi(\lambda)H$. Assuming the conditions given in Theorem 7.11 above and given that ϕ is continuously differentiable in a neighborhood of $\hat{\lambda}$, the local convergence of Algorithm 10 is quadratic.

Replacing the eigenpair (μ,x) corresponding to the kth-smallest eigenvalue of the problem $(A+\phi(\lambda)H)x=\mu x$ in line 13 by the eigenpair (ν_λ,x) corresponding to the kth-largest eigenvalue of

$$(A+(\phi(\lambda)-\lambda\phi'(\lambda))H)x=\nu_\lambda(I-\phi'(\lambda)H)x,$$

one can even get cubic convergence. This follows from Theorem 3.1 (iii) and the relation $\nu_\lambda:=\lambda-\mu$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

EXAMPLE 7.12. Consider the nonlinear modification $(A + \phi(\lambda)H)x = \lambda Bx$ of

$$Ax := \frac{1}{h}\text{tridiag}\{-1, 2, -1\}x = \lambda\frac{h}{6}\text{tridiag}\{1, 4, 1\}x =: \lambda Bx,$$

where $A, B \in \mathbb{R}^{100\times 100}$ and H is a symmetric rank-three matrix with one negative and 2 positive eigenvalues. Let $\phi(\lambda) = 0.9(1 - \sin(\lambda))$ and

$$\gamma = \max(0, \max \phi'(\lambda)\sigma_n, \max \phi'(\lambda)\sigma_1) = 0.996$$

such that Theorem 7.6 applies. Although the initial value $\lambda = \alpha_{k-\nu}$ is quite far away from the eigenvalue under consideration, the method has no problem to converge. Table 7.1 contains the 5 smallest eigenvalues, the number of iterations, and the final residual. The initial value of λ is 0, for $k = 1$, and α_{k-1}, for $k = 2, 3, 4, 5$.

TABLE 7.1
Five smallest eigenvalues of Example 7.12 via Algorithm 10.

k	λ	iter	$(\alpha_{k-1}, \alpha_{k+2})$	res
1	3.33307908052349	3	(-∞ , 61.7167)	6.6e-11
2	23.6619469468822	3	(2.46745, 121.024)	8.4e-11
3	61.5752941605781	3	(22.2107, 200.192)	1.1e-10
4	121.024515451116	2	(61.7167, 299.299)	8.9e-11
5	200.192841286577	2	(121.024, 418.441)	8.6e-11

8. Automated Multi-Level Substructuring. The Automated Multi-Level Substructuring (AMLS) method has been developed by Bennighof [27, 28] to reduce the computational demands in a frequency response analysis involving large and complex models. An efficient implementation by Gao, Li, Yang, and Bai can be found in [73].

AMLS automatically divides a large finite element model into many substructures on a number of levels based on the sparsity structure of the system matrices. Assuming that the interior degrees of freedom depend quasistatically on the interface degrees of freedom and modeling the deviation from quasistatic dependence in terms of a small number of selected substructure eigenmodes, the size of the finite element model is reduced substantially while yielding satisfactory accuracy over a wide frequency range of interest.

Recent studies (e.g., [116, 122]) in the vibro-acoustic analysis of passenger car bodies, where very large FE models with more than one million degrees of freedom appear and several hundreds of eigenfrequencies and eigenmodes are needed, have shown that AMLS is considerably faster than Lanczos-type approaches.

We stress the fact that substructuring does not mean that it is obtained by a domain decomposition of a real structure, but it is understood in a purely algebraic sense, i.e., the dissection of the matrices can be derived by applying a graph partitioner to the matrix under consideration. However, because of its pictographic nomenclature, we will use terms like substructure or eigenmode from frequency response problems when introducing the AMLS method.

From a mathematical point of view AMLS is a projection method where the ansatz space is constructed exploiting Schur complements of submatrices and the truncation of spectral representations of subproblems. In this presentation, we will take advantage of the fact that the original eigenproblem is equivalent to a rational eigenvalue problem of the same dimension as the projected problem in AMLS, which can be interpreted as exact condensation of the original eigenproblem with respect to an appropriate basis. Its eigenvalues at the lower end of

ETNA
Kent State University and
Johann Radon Institute (RICAM)

the spectrum can be characterized as minmax values of a Rayleigh functional of this rational eigenproblem. Hence, comparing the Rayleigh quotient of the projected problem and the Rayleigh functional of the rational problem, we derive an a priori bound for the error of the AMLS method.

8.1. Substructuring of eigenproblems. We are concerned with the linear eigenvalue problem

$$Kx = \lambda Mx, \tag{8.1}$$

where $K \in \mathbb{R}^{n\times n}$ and $M \in \mathbb{R}^{n\times n}$ are symmetric and positive definite matrices.

We first consider one-level versions of the substructuring methods. Assume that the joint graph of the matrices K and M is partitioned into r substructures such that the rows and columns of K can be reordered in the following way:

$$K = \begin{bmatrix} K_{\ell\ell 1} & O & \dots & O & K_{\ell i 1} \\ O & K_{\ell\ell 2} & \dots & O & K_{\ell i 2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & \dots & K_{\ell\ell r} & K_{\ell i r} \\ K_{i\ell 1} & K_{i\ell 2} & \dots & K_{i\ell r} & K_{ii} \end{bmatrix},$$

and M after reordering has the same block form. Here $K_{\ell\ell j}$, $j = 1, \dots, r$, is the local stiffness matrix corresponding to the jth substructure, i denotes the set of interface vertices, and $K_{\ell i j}$ describes the interaction of the interface degrees of freedom and the jth substructure.

Distinguishing only between local and interface degrees of freedom, K and M have the following form:

$$K = \begin{bmatrix} K_{\ell\ell} & K_{\ell i} \\ K_{i\ell} & K_{ii} \end{bmatrix} \qquad \text{and} \qquad M = \begin{bmatrix} M_{\ell\ell} & M_{\ell i} \\ M_{i\ell} & M_{ii} \end{bmatrix}. \tag{8.2}$$

We transform the matrix K into block diagonal form using block Gaussian elimination, i.e., we apply the congruent transformation with

$$P = \begin{bmatrix} I & -K_{\ell\ell}^{-1} K_{\ell i} \\ 0 & I \end{bmatrix}$$

to the pencil (K, M), obtaining the equivalent pencil

$$(P^T K P, P^T M P) = \left(\begin{bmatrix} K_{\ell\ell} & 0 \\ 0 & \tilde{K}_{ii} \end{bmatrix}, \begin{bmatrix} M_{\ell\ell} & \tilde{M}_{\ell i} \\ \tilde{M}_{i\ell} & \tilde{M}_{ii} \end{bmatrix} \right). \tag{8.3}$$

We further transfer the pencil (8.3), taking advantage of a modal basis for the local degrees of freedom. To this end consider the substructure eigenvalue problems

$$K_{\ell\ell}\Phi = M_{\ell\ell}\Phi\Omega, \qquad \Phi^T M_{\ell\ell}\Phi = I, \tag{8.4}$$

where Ω is a diagonal matrix containing the eigenvalues of the ℓth substructure. Changing the basis for the local degrees of freedom to a modal one, i.e., applying the further congruent transformation $\operatorname{diag}\{\Phi, I\}$ to problem (8.3), one gets

$$\left(\begin{bmatrix} \Omega & 0 \\ 0 & \tilde{K}_{ii} \end{bmatrix}, \begin{bmatrix} I & \Phi^T \tilde{M}_{\ell i} \\ \tilde{M}_{i\ell}\Phi & \tilde{M}_{ii} \end{bmatrix} \right). \tag{8.5}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

In structural dynamics (8.5) is called the Craigh-Bampton form of the eigenvalue problem (8.1) corresponding to the partitioning (8.2); cf. [51]. In terms of linear algebra it results from a block Gaussian elimination to reduce K to block diagonal form and a diagonalization of the block $K_{\ell\ell}$ using a spectral basis.

Selecting some eigenmodes of the problem (8.4) (usually the ones according to eigenvalues which do not exceed a cut-off threshold) and dropping the rows and columns in (8.5) corresponding to the other modes, one arrives at the component mode synthesis method (CMS) introduced by Hurty [110] and Craigh and Bampton [51]. Hence, if the diagonal matrix Ω_1 contains at its diagonal the eigenvalues to be dropped and Φ_1 the corresponding eigenvectors and if Ω_2 and Φ_2 contain the eigenvalues and eigenvectors to keep, respectively, then the eigenproblem (8.5) can be rewritten as

$$\begin{bmatrix} \Omega_1 & 0 & 0 \\ 0 & \Omega_2 & 0 \\ 0 & 0 & \tilde{K}_{ii} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \lambda \begin{bmatrix} I & 0 & \tilde{M}_{\ell i 1} \\ 0 & I & \tilde{M}_{\ell i 2} \\ \tilde{M}_{i\ell 1} & \tilde{M}_{i\ell 2} & \tilde{M}_{ii} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{8.6}$$

with

$$\tilde{M}_{\ell ij} = \Phi_j^T (M_{\ell i} - M_{\ell\ell} K_{\ell\ell}^{-1} K_{\ell i}) = \tilde{M}_{i\ell j}^T, \qquad j = 1, 2,$$

and the CMS approximations of the eigenpairs of (8.1) are obtained from the reduced eigenvalue problem

$$\begin{bmatrix} \Omega_2 & 0 \\ 0 & \tilde{K}_{ii} \end{bmatrix} y = \lambda \begin{bmatrix} I & \tilde{M}_{\ell i 2} \\ \tilde{M}_{i\ell 2} & \tilde{M}_{ii} \end{bmatrix} y. \tag{8.7}$$

AMLS generalizes CMS in the following way. Again the graph of $|K|+|M|$ is partitioned into a small number of subgraphs, but more generally than in CMS, these subgraphs in turn are substructured on a number p of levels yielding a tree topology for the substructures. This induces the following partitioning of the index set $I = \{1, \ldots, n\}$ of the degrees of freedom: Let I_1 be the set of indices corresponding to interface degrees of freedom on the coarsest level, and for $j = 2, \ldots, p$, define I_j to be the set of indices of interface degrees of freedom on the jth level that are not contained in I_{j-1}. Finally, let I_{p+1} be the set of interior degrees of freedom on the finest level.

With these notations AMLS works as follows: Its first step is the CMS method with a cut-off frequency τ_1 applied to the finest substructuring, i.e., I_{p+1} is the set of local degrees of freedom and $\tilde{I}_{p+1} := \bigcup_{j=1}^{p} I_j$ is the set of interface degrees of freedom. After j steps, $1 \le j \le p-1$, one derives a reduced pencil

$$\left(\begin{bmatrix} \Omega_f & O & O \\ O & K_{\ell\ell}^{(j)} & K_{\ell i}^{(j)} \\ O & K_{i\ell}^{(j)} & K_{ii}^{(j)} \end{bmatrix}, \begin{bmatrix} M_{ff}^{(j)} & M_{f\ell}^{(j)} & M_{fi}^{(j)} \\ M_{\ell f}^{(j)} & M_{\ell\ell}^{(j)} & M_{\ell i}^{(j)} \\ M_{if}^{(j)} & M_{i\ell}^{(j)} & M_{ii}^{(j)} \end{bmatrix} \right),$$

where f denotes the degrees of freedom obtained in the spectral reduction in the previous steps, ℓ collects the indices in I_{p+1-j}, and i corresponds to the index set $\bigcup_{k=1}^{p-j} I_k$ of interface degrees of freedom on levels which are not yet treated. Applying the CMS method to the south-east 2×2 blocks of the matrices, i.e., annihilating the off-diagonal block $K_{\ell i}^{(j)}$ by block Gaussian elimination and reducing the set of ℓ-indices by spectral truncation with a cut-off frequency τ_{j+1}, one arrives at the next level.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

After p CMS steps we obtain the reduced problem

$$\left(\begin{bmatrix} \Omega_p & O \\ O & K_{\ell\ell}^{(p)} \end{bmatrix}, \begin{bmatrix} M_{ff}^{(p)} & M_{f\ell}^{(p)} \\ K_{\ell f}^{(p)} & M_{\ell\ell}^{(p)} \end{bmatrix} \right),$$

and a final spectral truncation of the lower-right blocks with cut-off frequency τ_{p+1} yields the reduction of problem (8.1) by AMLS.

We have chosen this unusual description of AMLS because it is very convenient for deriving the error bound in Section 8.2. Note that this description neglects the algorithmically important fact that all matrices $K_{\ell\ell}^{(j)}$ and $M_{\ell\ell}^{(j)}$ are block diagonal. Hence, the annihilation of the off-diagonal blocks $K_{\ell i}^{(j)}$ and the spectral reduction on each level is quite inexpensive. A matrix and variational analysis of AMLS is contained in [28]; implementation details can be found in [73, 116].

8.2. A priori bounds for AMLS. We first consider the component mode synthesis method (8.7). If λ is not a diagonal entry of Ω_1, then the first equation of (8.6) yields

$$x_1 = \lambda(\Omega_1 - \lambda I)^{-1} \tilde{M}_{\ell i 1} x_3,$$

and λ is an eigenvalue of (8.1) if and only if it is an eigenvalue of the rational eigenproblem

$$T(\lambda) y = 0, \tag{8.8}$$

where

$$T(\lambda) = -\begin{bmatrix} \Omega_2 & 0 \\ 0 & \tilde{K}_{ii} \end{bmatrix} + \lambda \begin{bmatrix} I & \tilde{M}_{\ell i 2} \\ \tilde{M}_{i\ell 2} & \tilde{M}_{ii} \end{bmatrix} + \lambda^2 \begin{bmatrix} 0 \\ \tilde{M}_{i\ell 1} \end{bmatrix} (\Omega_1 - \lambda I)^{-1} \begin{bmatrix} 0 & \tilde{M}_{\ell i 1} \end{bmatrix}.$$

We denote by

$$\underline{\omega} := \min \operatorname{diag} \Omega_1$$

the smallest eigenvalue of problem (8.4) neglected in the CMS method, which can be replaced by the cut-off threshold. Let $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ denote the eigenvalues of problem (8.1) ordered by magnitude, and let $m \in \mathbb{N}$ such that $\lambda_m < \underline{\omega} \le \lambda_{m+1}$. Then $\lambda_1, \ldots, \lambda_m \in J$ are the eigenvalues of the nonlinear eigenproblem (8.8) in J, and it can be shown that these eigenvalues satisfy the minmax principle.

The eigenvalues $\tilde{\lambda}_1 \le \tilde{\lambda}_2 \le \cdots \le \tilde{\lambda}_\nu$ of the reduced problem (8.7) are minmax values of the Rayleigh quotient $\rho(x)$ corresponding to problem (8.7). Comparing p and ρ on appropriate subspaces of $\mathbb{R}^\nu$, we arrive at the following a priori bound for the relative errors of the CMS approximations $\tilde{\lambda}_j$ to λ_j.

THEOREM 8.1 ([62]). *Let $K, M \in \mathbb{R}^{n\times n}$ be symmetric and positive definite, and let $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ be the eigenvalues of problem (8.1), which we assume to be ordered by magnitude.*

Denote by $\tilde{\lambda}_1 \le \tilde{\lambda}_2 \le \cdots \le \tilde{\lambda}_\nu$ the eigenvalues of the CMS approximation (8.7) of problem (8.1) corresponding to some partition of the graph $|K| + |M|$ and some cut-off threshold $\underline{\omega}$.

Assume that the interval $(0, \underline{\omega})$ contains m eigenvalues $\lambda_1, \ldots, \lambda_m$ of (8.1). Then it holds that

$$0 \le \frac{\tilde{\lambda}_j - \lambda_j}{\lambda_j} \le \frac{\lambda_j}{\underline{\omega} - \lambda_j} \le \frac{\tilde{\lambda}_j}{\underline{\omega} - \tilde{\lambda}_j}, \qquad j = 1, \ldots, m. \tag{8.9}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

REMARK 8.2. Numerical examples demonstrate that the error bound in (8.9) overestimates the true relative error of CMS by one or two orders of magnitude. The paper [62] contains an example that demonstrates that the bound cannot be improved without further assumptions.

REMARK 8.3. Bekas and Saad [25] identified the level-one version of AMLS as a linearization of a rational eigenproblem, which motivated them to suggest three modifications: a second-order approximation, expanding the projection space by Krylov subspaces, and a combination of these two modifications. For the multi-level substructuring method, however, these modifications do not seem to be useful.

REMARK 8.4. Yang et al. [257] considered the component mode synthesis method, and they obtained a simple heuristic for choosing spectral components from each substructure suggesting to drop all eigenpairs (ω, ϕ) of substructures in the reduction process such that

$$\rho_1(\omega) := \frac{\lambda_1}{\omega - \lambda_1} \leq \tau,$$

where λ_1 is the smallest eigenvalue of the problem under consideration and τ is a given tolerance. Theorem 8.1 guarantees that with this choice, the relative error of the CMS approximation $\tilde{\lambda}_1$ of λ_1 is less than τ.

Since AMLS can be understood as a sequence of p consecutive CMS steps and a terminating spectral truncation, it is clear how to obtain an a priori bound for the general AMLS method. Every reduction step in which a quasistatic/modal representation is obtained and the dimension is reduced by spectral truncation is identical to a CMS step utilizing the substructuring of the next level. Hence, one obtains the following error bound for AMLS.

THEOREM 8.5 ([62]). *Let K, M and λ_j, $j = 1, \ldots, n$, be given as in Theorem 8.1. Let the graph of $|K| + |M|$ be substructured with p levels, and denote by $\tilde{\lambda}_1^{(\nu)} \leq \tilde{\lambda}_2^{(\nu)} \leq \ldots$ the eigenvalues obtained by AMLS with a cut-off threshold $\underline{\omega}_\nu$ on the level ν.*

If $m \in \mathbb{N}$ such that $\lambda_m < \min_{\nu=0,\ldots,p} \underline{\omega}_\nu \leq \lambda_{m+1}$, then it holds that

$$\frac{\tilde{\lambda}_j - \lambda_j}{\lambda_j} \leq \prod_{\nu=0}^{p} \left(1 + \frac{\lambda_j^{(\nu)}}{\underline{\omega}_\nu - \lambda_j^{(\nu)}}\right) - 1, \qquad j = 1, \ldots, m.$$

Since the final problem is a projection of each of the intermediate eigenproblems in the AMLS reduction, it follows from the minmax characterization that $\lambda_j^{(\nu)} \leq \tilde{\lambda}_j$, for $\nu = 0, \ldots, p$. Therefore the a priori bound (8.9) can be replaced by the computable bound

$$\frac{\tilde{\lambda}_j - \lambda_j}{\lambda_j} \leq \prod_{\nu=0}^{p} \left(1 + \frac{\tilde{\lambda}_j}{\underline{\omega}_\nu - \tilde{\lambda}_j}\right) - 1, \qquad j = 1, \ldots, m.$$

8.3. AMLS reduction for nonlinear eigenproblems. We consider the nonlinear eigenvalue problem

$$T(\lambda)x = 0, \tag{8.10}$$

where $T(\lambda) \in \mathbb{C}^{n\times n}$ is a large and sparse matrix depending on a parameter $\lambda \in D \subset \mathbb{C}$. To generalize the AMLS method, we identify an essential linear part of $T(\cdot)$, i.e., we rewrite the problem (8.10) as

$$Kx - \lambda Mx - R(\lambda)x = 0,$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

where $K \in \mathbb{C}^{n\times n}$ and $M \in \mathbb{C}^{n\times n}$ are Hermitian and positive definite matrices and

$$R(\lambda) = K - \lambda M - T(\lambda)$$

is a perturbation of the linear eigenproblem $Kx = \lambda Mx$, which should be small in the eigenparameter set of interest.

Once the multi-level substructuring transformation of the linear pencil (K, M) has been accomplished with a given cut-off frequency, we obtain a matrix Φ_{AMLS} of substructure modes and a projected eigenproblem

$$\mathcal{K}y = \lambda \mathcal{M}y \tag{8.11}$$

of much smaller dimension, where $\mathcal{K} = \Phi_{\text{AMLS}}^H K \Phi_{\text{AMLS}}$ and $\mathcal{M} = \Phi_{\text{AMLS}}^H M \Phi_{\text{AMLS}}$.

This information can be used in two ways to solve the nonlinear eigenvalue problem approximately: First, we may project the nonlinear eigenproblem (8.10) to the subspace of $\mathbb{C}^n$ spanned by the substructure modes which were kept in the AMLS reduction, i.e.,

$$\Phi_{\text{AMLS}}^H T(\lambda)\Phi_{\text{AMLS}} y = \mathcal{K}y - \lambda \mathcal{M}y - \Phi_{\text{AMLS}}^H R(\lambda)\Phi_{\text{AMLS}} y = 0. \tag{8.12}$$

In particular this projection can be performed easily if the remainder $R(\lambda)$ has the form

$$R(\lambda) = \sum_{j=1}^{p} f_j(\lambda) C_j,$$

where $f_j(\lambda)$ are given complex functions and $C_j \in \mathbb{C}^{n\times n}$ are given matrices, which quite often have the same sparsity structure as the pencil (K, M) or some other simple structure. In this case the projection $\Phi_{\text{AMLS}}^H R(\lambda)\Phi_{\text{AMLS}}$ could be determined simultaneously with the matrices $\mathcal{K}$ and $\mathcal{M}$ in the course of the AMLS reduction.

Secondly, we may determine Ritz pairs $(\lambda_j, \Phi_{\text{AMLS}} y_j)$, $j = 1, \ldots, m$, of the linear problem $Kx = \lambda Mx$ corresponding to eigenvalues in the wanted region from the projected problem (8.11) and project the nonlinear problem to the subspace spanned by these Ritz vectors. Thus, we get

$$X^H T(\lambda) X z = \Lambda z - \lambda z - X^H R(\lambda) X z = 0, \tag{8.13}$$

where $\Lambda = \text{diag}\{\lambda_1, \ldots, \lambda_m\}$ and $X = (x_1, \ldots, x_m)$.

Problem (8.13) is equivalent to the projection of the problem (8.12) to the space spanned by the eigenvectors $y_1, \ldots, y_m$ of (8.11) corresponding to $\lambda_1, \ldots, \lambda_m$. Hence, we can expect that the first approach will yield better approximations. Examples, however, demonstrate that the loss of accuracy often is negligible.

In either case we arrive at a projected nonlinear eigenvalue problem of much smaller dimension which can be solved by an appropriate method, i.e., a dense solver if the projected problem is small or by linearization if the underlying problem is a polynomial eigenproblem or by an iterative projection method of Arnoldi- or Jacobi-Davidson type.

Applications of AMLS to gyroscopic eigenproblems including numerical experiments are contained in [63] and to applications to nonlinear eigenproblems (nonproportional damping, vibrations of fluid-solid structures) can be found in [64].

REMARK 8.6. We discussed the relation between AMLS and a subspace iteration, i.e., enhancing the eigenvector approximation from AMLS with a subspace iteration, in [251, 260] and preconditioning the subspace iteration with AMLS in [252].

ETNA
Kent State University and
Johann Radon Institute (RICAM)

9. Large-scale Tikhonov regularization via reduction by orthogonal projection. This section considers an approach for computing an approximate solution of a Tikhonov-regularized large-scale ill-posed least-squares problem

$$\min_{x\in\mathbb{R}^n} \|Ax-b\|, \tag{9.1}$$

with a severely ill-conditioned and possibly singular matrix $A \in \mathbb{R}^{m\times n}$, $m \geq n$, and an observation vector $b \in \mathbb{R}^m$.

Due to the ill-conditioning of A, the straightforward least-squares solution of (9.1) often does not yield a meaningful approximation of a solution, but it is necessary to stabilize the computation by regularization. One of the most popular regularizations methods is due to Tikhonov [224], which replaces (9.1) by

$$\min_{x\in\mathbb{R}^n} \left\{\|Ax-b\|^2 + \mu^{-1}\|Lx\|^2\right\} \tag{9.2}$$

with a scalar $\mu \in (0,\infty)$ and a general regularization matrix $L \in \mathbb{R}^{p\times n}$. The normal equations associated with (9.2) are given by

$$T(\lambda)x := (A^TA + \mu^{-1}L^TL)x = A^Tb. \tag{9.3}$$

If rank $\left[A^T, L^T\right] = n$, then for any $\mu > 0$, the problem (9.3) has the unique solution

$$x_\mu = (A^TA + \mu^{-1}L^TL)^{-1}A^Tb. \tag{9.4}$$

When the matrices A and L are small, the solutions x_μ of (9.4) can easily be determined for many values of $\mu > 0$ by first computing the generalized singular value decomposition (GSVD) [233] of the matrix pair $\{A, L\}$. For large-scale problems and a fixed $\mu > 0$, an approximation of x_μ can be determined by applying an iterative method, such as LSQR [189], to (9.3). However, generally, a suitable value of the parameter μ is not known a priori and has to be determined during the solution process. Many approaches to determining an appropriate value of μ, including the L-curve criterion [96, 97, 141], the discrepancy principle [65, 178], generalized cross validation [47, 78], and information criteria [10, 205], require the normal equations (9.3) to be solved repeatedly for many different values of the parameter μ. This can make applications of LSQR costly.

For the Tikhonov regularization problem in standard form, i.e., when $L = I$, approximate solutions of (9.4) can be computed by partial Lanczos bidiagonalization of A (see e.g., [42, 43, 44, 95]), i.e., by projecting the problem onto the Krylov subspace

$$\mathcal{K}_\ell(A^TA, A^Tb) = \text{span}\{A^Tb, (A^TA)A^Tb, \ldots, (A^TA)^{\ell-1}A^Tb\} \tag{9.5}$$

for some $\ell \geq 1$. Due to the shift invariance of Krylov subspaces, the subspace (9.5) can be used for several parameter values μ. A solution by partial Lanczos bidiagonalization can also be applied to the Tikhonov regularization problems (9.2) with $L \neq I$, provided that the regularization matrix can be transformed to standard form without too much effort. This transformation is carried out with the aid of the substitutions $y = Lx$ and $x = L_A^\dagger y$, where for $p \leq n$,

$$L_A^\dagger := (I - (A(I - L^\dagger L))^\dagger A)L^\dagger.$$

This matrix is referred to as the A-weighted pseudoinverse of L [60].

In [127] we proposed an iterative projection method that computes an approximate solution of (9.4) in a generalized Krylov subspace. The regularization parameter is determined

ETNA
Kent State University and
Johann Radon Institute (RICAM)

by the discrepancy principle. Given an estimate δ of the norm of the error in the vector b, i.e., $\delta \approx \|\Delta b\|$ with $b = b_{true} + \Delta b$, the regularization parameter $\mu = \mu(\delta)$ is determined such that the computed approximation $\tilde{x}_\mu$ of the solution x_μ satisfies

$$\|A\tilde{x}_\mu - b\| = \delta.$$

Introduce the function

$$\phi(\mu) := \|Ax_\mu - b\|^2,$$

where x_μ is given by (9.4), and let $\bar{\mu}$ satisfy

$$\phi(\bar{\mu}) = \delta^2.$$

The function $\phi(\mu)$ is convex and monotone. A numerical method for inexpensively computing upper and lower bounds for $\bar{\mu}$ when $L = I$ is described in [44]. Note that the evaluation of $\phi(\mu)$ is expensive when A is large. Assume that L is a regularization matrix such that the computation with $L_A^\dagger$ is costly.

Let $\mathcal{V}$ be a subspace of small dimension $k \ll n$, and let the columns of $V \in \mathbb{R}^{n\times k}$ be an orthonormal basis of $\mathcal{V}$. We propose to approximate $\phi(\mu)$ by the function

$$\phi(\mu; V) := \|Ax_\mu^k - b\|^2,$$

where x_μ^k is obtained by solving the Tikhonov problem (9.2) restricted to $\mathcal{V}$. Specifically, let

$$y_\mu^k := \text{argmin}_{y\in\mathbb{R}^k}\{\|AVy - b\|^2 + \mu_k^{-1}\|LVy\|^2\}, \qquad x_\mu^k := Vy_\mu^k.$$

The regularization parameter μ_k is determined as the zero of the function

$$f(\mu; V) := \|Ax_\mu^k - b\|^2 - \delta^2$$

and can be computed, e.g., by Newton's method, by rational inverse iteration (see [130, 133]), or by a cubically convergent zero finder [192]. Let μ_k be an available approximation of $\bar{\mu}$, and let

$$r_\mu^k = (A^TA + \mu_k^{-1}L^TL)x_\mu^k - A^Tb$$

be the residual of (9.3) corresponding to $x_\mu^k = Vy_\mu^k$. In the absence of round-off errors, r_μ^k is orthogonal to the search space $\mathcal{V}$. To enforce orthogonality in the presence of round-off errors, we reorthogonalize and obtain the expanded search space V_{new}, i.e.,

$$\tilde{r}_\mu^k := (I - VV^T)r_\mu^k, \qquad v_{new} := \tilde{r}_\mu^k/\|\tilde{r}_\mu^k\|, \qquad V_{new} := [V, v_{new}].$$

The resulting Generalized Krylov Subspace Tikhonov Regularization Method [127] is given in Algorithm 11.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Algorithm 11 Generalized Krylov Subspace Tikhonov Regularization Method.

Require: initial basis V_0, $V_0^T V_0 = I$
1: **for** $i = 0, 1, \ldots$ until convergence **do**
2: find the root μ_i of $f(\mu; V_i) = 0$
3: solve $(V_i^T T(\mu_i) V_i) y_\mu^i = V_i^T A^T b$
4: compute $r_\mu^i = T(\mu_i) V_i y_\mu^i - A^T b$
5: reorthogonalize (optional) $\tilde{r}_\mu^i = (I - V_i V_i^T) r_\mu^i$
6: normalize $v_{\text{new}} = \tilde{r}_\mu^i / \|\tilde{r}_\mu^i\|$
7: enlarge the search space $V_{i+1} = [V_i, v_{\text{new}}]$
8: **end for**
9: determine the approximate Tikhonov solution $x_\mu^i = V_i y_\mu^i$

EXAMPLE 9.1. The inverse heat equation `heat(5)` from Hansen's *Regularization Tools* [98] has been used to generate a test problem with ill-determined rank (i.e., with singular values that gradually decay to zero) and a numerically singular matrix $A \in \mathbb{R}^{400\times 200}$ and a vector b containing Gaussian noise with level $\sigma = 1 \cdot 10^{-2}$ together with a regularization matrix

$$L = L_1 = \begin{bmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix} \in \mathbb{R}^{(n-1)\times n}.$$

The initial search space is the Krylov subspace $\operatorname{span}\{V_0\} = \mathcal{K}_\ell(A^T A, A^T b)$ with $\ell = 7$.

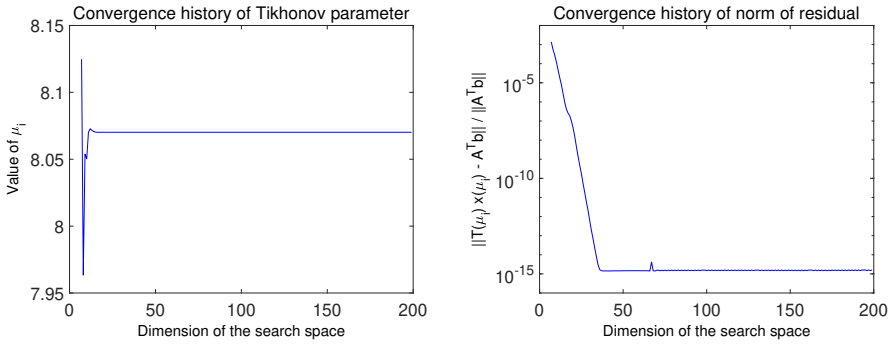


FIG. 9.1. *Convergence histories of Algorithm 11 for Example 9.1.*

On the left-hand side of Figure 9.1, the sequence $\{\mu_i\}$ is displayed, from which it can be observed that the value $\bar{\mu} \approx 8.07$ is approached very quickly as $\dim(\mathcal{V})$ increases. The corresponding relative residual norms, $\|r(x_\mu^i)\|/\|A^T b\|$, $i = 0, 1, 2, \ldots$, are displayed in Figure 9.1 on the right; cf. line 4 of Algorithm 11. Notice that the convergence of the regularization parameter is not monotonic; the sequence $\{\mu_i\}$ oscillates around $\bar{\mu}$ also when the dimension of $\mathcal{V}$ (i.e., $\dim(\mathcal{V}) = i + 7$) increases, but this is not visible in the figure.

10. Regularized total least-squares problems. Many problems in data estimation are governed by overdetermined linear systems $Ax \approx b$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $m \geq n$, where both the matrix A and the right-hand side b contain some noise. An appropriate approach to

ETNA
Kent State University and
Johann Radon Institute (RICAM)

this problem is the total least-squares (TLS) method [3, 79], which determines perturbations $\Delta A \in \mathbb{R}^{m\times n}$ of the coefficient matrix A and $\Delta b \in \mathbb{R}^m$ of the vector b such that

$$\|[\Delta A, \Delta b]\|_F^2 = \min! \qquad \text{subject to } (A+\Delta A)x = b + \Delta b, \tag{10.1}$$

where $\|\cdot\|_F$ denotes the Frobenius norm of a matrix. An overview of total least-squares methods and a comprehensive list of references is contained in [167, 230, 231, 232].

The TLS problem (10.1) can be analyzed in terms of the singular value decomposition (SVD) of A and $[A, b]$; cf. [80, 232]. For discretizations of ill-posed problems such as integral equations of the first kind (cf. [65, 84, 97]), least-squares or total least-squares methods often yield physically meaningless solutions, and regularization is necessary to stabilize the solution.

Motivated by Tikhonov regularization, a well-established approach is to add a quadratic constraint yielding the regularized total least squares (RTLS) problem

$$\|[\Delta A, \Delta b]\|_F^2 = \min! \qquad \text{subject to } (A+\Delta A)x = b + \Delta b,\ \|Lx\| \le \delta, \tag{10.2}$$

where δ is a regularization parameter and $L \in \mathbb{R}^{k\times n}$, $k \le n$, defines a (semi-)norm for the solution by which the size of the solution is bounded or a certain degree of smoothness for the solution can be imposed. Throughout this section it is assumed that the condition $\sigma_{min}([AK, b]) < \sigma_{min}(AK)$ holds, where K is an orthonormal basis of the kernel of L, which guarantees that a solution of the RTLS problem (10.2) is attained; cf. [22]. Notice that the condition is empty if the regularization matrix L is nonsingular. Stabilization by introducing a quadratic constraint was extensively studied in [23, 77, 88, 126, 129, 130, 131, 132, 133, 194, 207, 208]. Tikhonov regularization was considered in [22, 134].

It is assumed that the regularization parameter $\delta > 0$ is smaller than $\|Lx_{TLS}\|$, where x_{TLS} denotes the solution of the total least-squares problem (10.1); otherwise no regularization would be necessary. Then at the optimal solution of (10.2), the constraint $\|Lx\| \le \delta$ holds with equality. Under this condition Golub, Hansen, and O'Leary [77] derived the following first-order necessary conditions of the RTLS problem (10.2): The solution x_{RTLS} of (10.2) is a solution of the problem

$$(A^TA + \lambda_I I_n + \lambda_L L^TL)x = A^Tb, \tag{10.3}$$

where the parameters λ_I and λ_L are given by

$$\lambda_I = -\frac{\|Ax-b\|^2}{1+\|x\|^2}, \quad \lambda_L = \frac{1}{\delta^2}\Big(b^T(b-Ax) - \frac{\|Ax-b\|^2}{1+\|x\|^2}\Big). \tag{10.4}$$

This condition was used in the literature in two ways to solve problem (10.2): In [77, 88, 131, 194], the value of λ_I is chosen as a free parameter. Then for fixed λ_L, problem (10.3) is solved for (x, λ_I), and then λ_L is updated in a way that the whole process converges to the solution of (10.2). Conversely, in [129, 130, 207, 208], for a chosen parameter λ_I, problem (10.3) is solved for (x, λ_L), which yields a convergent sequence of updates for λ_I.

In the first case one has to determine in every iteration step the eigenvector of a symmetric matrix corresponding to its smallest eigenvalue, and in the latter approach, one has to find the rightmost eigenvalue and the corresponding eigenvector of a symmetric quadratic eigenproblem in every iteration step. Hence, in both cases one has to solve a sequence of eigenvalue problems, which converge as the methods approach the solution of (10.2).

Results for employing the sequence of QEPs are contained in Section 10.1, whereas results for the sequence of linear EVPs are contained in Section 10.2. Computational methods for solving RTLS problems based on eigenproblems are given in Section 10.3 and the corresponding numerical examples in Section 10.4.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

10.1. Regularized total least-squares based on a sequence of QEPs. It is well known (cf. [23, 232]) that the RTLS problem (10.2) is equivalent to

$$f(x) := \frac{\|Ax-b\|^2}{1+\|x\|^2} = \min! \qquad \text{subject to } \|Lx\|^2 \le \delta^2. \tag{10.5}$$

Let us first consider the parameter λ_I to be fixed for one iteration step, and let $\lambda := \lambda_L$ be a free parameter. The fixed parameter is updated and initialized as suggested in (10.4),

$$\lambda_I = \lambda_I(x^k) = -\frac{\|Ax^k-b\|^2}{1+\|x^k\|^2}.$$

The first-order optimality conditions then reads

$$B(x^k)x + \lambda L^T Lx = A^T b, \qquad \|Lx\|^2 = \delta^2, \tag{10.6}$$

with

$$B(x^k) = A^T A - f(x^k)I, \qquad f(x^k) = \frac{\|Ax^k-b\|^2}{1+\|x^k\|^2} = -\lambda_I(x^k),$$

which suggests the following Algorithm 12.

Algorithm 12 RTLSQEP.

Require: initial vector x^1.

1: **for** $k = 1, 2, \ldots$ until convergence **do**

2: with $B_k := B(x^k)$ solve

$$B_k x^{k+1} + \lambda L^T L x^{k+1} = A^T b, \quad \|Lx^{k+1}\|^2 = \delta^2 \tag{10.7}$$

for (x^{k+1}, λ) corresponding to the largest $\lambda \in \mathbb{R}$

3: **end for**

Sima, Van Huffel, and Golub [208] proposed to solve (10.6) via a quadratic eigenvalue problem similarly to the approach of Golub [76] for regularized least-squares problems. This motivates the name RTLSQEP of the algorithm. With an active constraint at the solution of (10.2), i.e., $\|Lx^*\|^2 = \delta^2$, the following global convergence result holds.

THEOREM 10.1 ([130]). *Any limit point x^* of the sequence $\{x^k\}$ constructed by Algorithm 12 is a global minimizer of the optimization problem (10.5) and thus of the RTLS problem (10.2).*

If L is square and nonsingular, then with $z = Lx^{k+1}$, problem (10.7) is equivalent to

$$W_k z + \lambda z := L^{-T} B_k L^{-1} z + \lambda z = L^{-T} A^T b =: h, \qquad z^T z = \delta^2.$$

Assuming that $W_k + \lambda I$ is positive definite and denoting $u := (W_k + \lambda I)^{-2} h$, one gets $h^T u = z^T z = \delta^2$, and the identity $h = \delta^{-2} h h^T u$ yields that $(W_k + \lambda I)^2 u = h$ is equivalent to the quadratic eigenvalue problem

$$T(\lambda)u := (W_k + \lambda I)^2 u - \delta^{-2} h h^T u = 0. \tag{10.8}$$

The choice of the rightmost eigenvalue can be motivated as the maximal Lagrange multiplier that minimizes an underlying quadratic function; cf. [71, 130].

ETNA
Kent State University and
Johann Radon Institute (RICAM)

In [129] it is shown that the rightmost eigenvalue $\hat{\lambda}$ of (10.8) is real and that $W_k + \hat{\lambda} I$ is positive semidefinite. We are considering the generic case of $W_k + \hat{\lambda} I$ being positive definite. In this case the solution of the original problem (10.7) is recovered from $z = (W_k + \hat{\lambda} I)u$ and $x^{k+1} = L^{-1} z$, where u is an eigenvector corresponding to $\hat{\lambda}$ which is scaled such that $h^T u = \delta^2$.

The case that $W_k + \hat{\lambda} I \geq 0$ is singular is discussed for constrained symmetric eigenproblems, i.e., with linear equality constraints for the eigenvector, in [72] and in [126] for solving the sequence (10.7).

REMARK 10.2. When $W_k + \hat{\lambda} I$ is singular it holds that $\|(W_k + \hat{\lambda} I)^{\dagger} h\| < \delta$ for any k happens if and only if the solution in this iteration step is nonunique. In this case the solutions can be constructed from the eigenspace corresponding to the rightmost eigenvalue $\hat{\lambda}$ of the QEP (10.8) and the minimum-norm solution of the consistent system $(W_k + \hat{\lambda} I)^{\dagger} h$. Thus, the RTLS problem (10.2) is nonunique if and only if at the limit value $f(x^*) = f^*$, with $W_{k \to \infty} := W(f^*)$, it holds that $\|(W(f^*) + \hat{\lambda} I)^{\dagger} h\| < \delta$.

If $\operatorname{rank}(L) = k < n$, then problem (10.7) has to be reduced to the range of L correspondingly, which does not effect the obtained results; see [129, 208].

With a symmetric matrix W and $h \in \mathbb{R}^n$, it holds that, for any fixed $x \neq 0$,

$$f(\lambda, x) := x^H T(\lambda) x = \lambda^2 \|x\|_2^2 + 2\lambda x^H W x + \|Wx\|_2^2 - |x^H h|^2 / \delta^2$$

is a parabola which attains its minimum at $\lambda = -\dfrac{x^H W x}{x^H x}$. Hence, we choose $J = (-\lambda_{\min}, \infty)$, where $\lambda_{\min}$ is the minimum eigenvalue of W. Then $f(\lambda, x) = 0$ has at most one solution $p(x) \in J$ for every $x \neq 0$, and the Rayleigh functional p of (10.8) corresponding to J is defined. Obviously, it holds that $x^H T'(p(x)) x > 0$ for every $x \in \mathcal{D}$, and the general conditions of the maxmin characterization are satisfied.

THEOREM 10.3 ([129]). *Let $\lambda_{\min}$ be the minimal eigenvalue of W and $x_{\min}$ be a corresponding eigenvector. Let $J = (-\lambda_{\min}, \infty)$, and denote by p the Rayleigh functional of $T(\cdot)$ and by $\mathcal{D}$ its domain of definition.*

(i) If $x_{\min}^T h = 0$ and $T(-\lambda_{\min})$ is positive semidefinite, then $\hat{\lambda} := -\lambda_{\min}$ is the maximal real eigenvalue of (10.8) and $x_{\min}$ is a corresponding eigenvector.

(ii) Otherwise, the maximal real eigenvalue is the unique eigenvalue $\hat{\lambda}$ of (10.8) in J, and it holds that

$$\hat{\lambda} = \max_{x \in \mathcal{D}} p(x).$$

(iii) $\hat{\lambda}$ is the rightmost eigenvalue of (10.8), i.e.,

$$\textit{real}(\lambda) \leq -\lambda_{\min} \leq \hat{\lambda} \qquad \textit{for every eigenvalue } \lambda \textit{ of } (10.8).$$

The following theorem characterizes the case that this rightmost eigenvalue of (10.8) is negative.

THEOREM 10.4 ([129]). *The maximal real eigenvalue $\hat{\lambda}$ of problem (10.8) is negative if and only if W is positive definite and*

$$\|W^{-1} h\| < \delta.$$

A negative rightmost eigenvalue of problem (10.8) may appear for any $\delta > 0$. Since it holds that $\|W^{-1} h\| = \|L(A^T A - f(x) I_n)^{-1} A^T b\|$, the condition of Theorem 10.4 can easily be fulfilled for a singular regularization matrix L and the vector $(A^T A - f(x) I_n)^{-1} A^T b$ in the nullspace of L. Small perturbations of this case show that a negative rightmost eigenvalue

ETNA
Kent State University and
Johann Radon Institute (RICAM)

may appear for a nonsingular matrix L as well. Theorem 10.4 demonstrates also that in the standard case $L = I$, the rightmost eigenvalue $\hat{\lambda}$ is always nonnegative if $\delta \leq \|x_{\text{TLS}}\|$; cf. [129].

Since the constraint is assumed to be active, the meaning of λ as Lagrange parameter implies $\hat{\lambda} > 0$ at any limit point x^*. Thus, a negative rightmost eigenvalue can only occur before convergence of Algorithm 12. The meaning of a negative Lagrange parameter can be understood by formulating equation (10.7) in Algorithm 12 as the following equivalent quadratic optimization problem (cf. [130]),

$$g(x; x^k) := \|Ax - b\|^2 - f(x^k)(1 + \|x\|^2) = \min! \quad \text{subject to} \quad \|Lx\|^2 = \delta^2, \tag{10.9}$$

yielding the same sequence of iterates $\{x^k\}$. By replacing the equality constraint in equation (10.9) by $\|Lx\|^2 \leq \delta^2$, a different globally convergent algorithm for solving RTLS problems is derived; cf. [24]. The computational effort per iteration is larger since the minimization problem with the inequality constraint cannot be solved by a QEP.

REMARK 10.5. A negative rightmost eigenvalue within Algorithm 12 occurs if and only if the minimum of the corresponding quadratic optimization problem (10.9) with inequality constraints is located in the interior of the feasible region $\|Lx\|^2 \leq \delta^2$.

EXAMPLE 10.6. Let

$$A = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad L = \begin{bmatrix} 0.95 & -1.74 \\ -0.94 & 1.73 \end{bmatrix}, \quad \delta = 0.99\|Lx_{TLS}\|.$$

For the unconstrained solution it holds that $x_{TLS} = x_{LS} = [1, 0.5]^T$, and the initial vector is chosen as $x^1 = [3, 2]^T \cdot (\delta / \|L[3, 2]^T\|) \approx [0.36, 0.24]^T$ with $\|Lx^1\| = \delta$ and $\|f(x^1)\| = 1.768$. The unique solution $x_{RTLS} \approx [0.9999, 0.5004]^T$ is close to x_{TLS}, with the corresponding function value $f(x_{RTLS}) \approx 1.763\text{e-}6$.

TABLE 10.1
Convergence history of $f(x^k)$.

iter	$f(x^k)$	rightmost λ^k	$f(\hat{x}^k)$
1	1.768e-0		
2	6.840e-1	-2.551515	6.632e-1 (interior)
3	7.793e-2	-0.841702	7.911e-2 (interior)
4	8.845e-4	-0.048918	8.937e-4 (interior)
5	1.874e-6	0.032408	1.874e-6 (bound.)
6	1.763e-6	0.033315	1.763e-6 (bound.)

Table 10.1 contains the convergence history of $f(x^k)$ and the rightmost eigenvalues λ^k determined by Algorithm 12 in the second and third column. In the last column the function values $f(\hat{x}^k)$ are displayed with the iterates $\hat{x}^k$ as minimizers of the quadratic optimization problem $g(x; x^k) = \min!$ s.t. $\|Lx\| \leq \delta$, showing that the occurrence of interior solutions corresponds to a negative sign of the rightmost λ^k of the QEPs.

10.2. Regularized total least-squares based on a sequence of EVPs. The second algorithm is based on keeping the parameter λ_L fixed for one iteration step and letting $\lambda := -\lambda_I$ be a free parameter. The following version of the first-order optimality conditions was proved by Renaut and Guo in [194].

THEOREM 10.7. *The solution x_{RTLS} of the RTLS problem (10.2) subject to the active constraint satisfies the augmented eigenvalue problem*

$$B(\lambda_L(x_{RTLS})) \begin{bmatrix} x_{RTLS} \\ -1 \end{bmatrix} = -\lambda_I(x_{RTLS}) \begin{bmatrix} x_{RTLS} \\ -1 \end{bmatrix}, \tag{10.10}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

with

$$B(\lambda_L) = M + \lambda_L N, \qquad M := [A,b]^T[A,b], \qquad N := \begin{bmatrix} L^TL & 0 \\ 0 & -\delta^2, \end{bmatrix}$$

and λ_L and λ_I as given in (10.4). Conversely, if $((\hat{x}^T,-1)^T,-\hat{\lambda})$ is an eigenpair of $B(\lambda_L(\hat{x}))$, where $\lambda_L(\hat{x})$ is recovered according to (10.4), then $\hat{x}$ satisfies (10.3) and $\hat{\lambda} = -f(\hat{x})$.

Algorithm 13 RTLSEVP.

Require: initial guess $\lambda_L^0 > 0$ and $B_0 = B(\lambda_L^0)$

1: **for** $k = 1, 2, \ldots$ until convergence **do**

2: solve

$$B_{k-1}y^k = \lambda y^k$$

for the eigenpair (y^k, λ) corresponding to the smallest λ

3: scale y^k such that $y^k = \begin{bmatrix} x^k \\ -1 \end{bmatrix}$

4: update $\lambda_L^k = \lambda_L(x^k)$ and $B_k = B(\lambda_L^k)$

5: **end for**

This condition suggested Algorithm 13 called RTLSEVP for obvious reasons. The choice of the smallest eigenvalue is motivated by aiming at

$$\lambda = -\lambda_I = f(x) = (\|Ax-b\|^2)/(1+\|x\|^2),$$

which is the function to be minimized; see (10.5) and (10.10). The straightforward idea in [88] to update λ_L in line 4 with (10.4), i.e.,

$$\lambda_L^{k+1} = \frac{1}{\delta^2}\left(b^T(b-Ax^{k+1}) - \frac{\|Ax^{k+1}-b\|^2}{1+\|x^{k+1}\|^2}\right)$$

does not lead in general to a convergent algorithm.

To enforce convergence, Renaut and Guo [194] proposed to determine a value θ such that the eigenvector $(x_\theta^T,-1)^T$ of $B(\theta)$ corresponding to the smallest eigenvalue of $B(\theta)$ satisfies the constraint $\|Lx_\theta\|^2 = \delta^2$, i.e., find a non-negative root $\hat{\theta}$ of the real function $g(\theta) := (\|Lx_\theta\|^2 - \delta^2)/(1+\|x_\theta\|^2)$. Then the corresponding eigenvector $(x_{\hat{\theta}}^T,-1)^T$ is a solution of (10.10). But the last component of an eigenvector corresponding to the smallest eigenvalue of $B(\theta)$ need not be different from zero, and then $g(\theta)$ is not necessarily defined. To fill this gap the following generalization has been stated in [131]:

DEFINITION 10.8. *Let $\mathcal{E}(\theta)$ denote the eigenspace of $B(\theta)$ corresponding to its smallest eigenvalue. Then*

$$g(\theta) := \min_{y\in\mathcal{E}(\theta)} \frac{y^TNy}{y^Ty} = \min_{(x^T,x_{n+1})^T\in\mathcal{E}(\theta)} \frac{\|Lx\|^2 - \delta^2 x_{n+1}^2}{\|x\|^2 + x_{n+1}^2}$$

is the minimal eigenvalue of the projection of the matrix N from (10.10) onto $\mathcal{E}(\theta)$.

This extends the definition of g to the case of eigenvectors with zero last components. The following theorem was shown in [131].

ETNA
Kent State University and
Johann Radon Institute (RICAM)

THEOREM 10.9. *The function $g : [0, \infty) \to \mathbb{R}$ has the following properties:*

(i) If $\sigma_{\min}([A, b]) < \sigma_{\min}(A)$, then $g(0) > 0$.
(ii) $\lim_{\theta\to\infty} g(\theta) = -\delta^2$.
(iii) If the smallest eigenvalue of $B(\theta_0)$ is simple, then g is continuous at θ_0.
(iv) g is monotonically not increasing on $[0, \infty)$.
(v) Let $g(\hat{\theta}) = 0$, and let $y \in \mathcal{E}(\hat{\theta})$ be such that $g(\hat{\theta}) = y^T N y/\|y\|^2$. Then the last component of y is different from 0.
(vi) g has at most one root.

Theorem 10.9 demonstrates that if $\hat{\theta}$ is a positive root of g, then $x := -y(1 : n)/y_{n+1}$ solves the RTLS problem (10.2), where y denotes an eigenvector of $B(\hat{\theta})$ corresponding to its smallest eigenvalue.

REMARK 10.10. If the smallest eigenvalue $\lambda_{\min}(\tilde{\theta})$ of $B(\tilde{\theta})$ is simple, then it follows from the differentiability of $\lambda_{\min}(\theta)$ and its corresponding eigenvector that

$$\frac{d\lambda_{\min}(B(\theta))}{d\theta}\Big|_{\theta=\tilde{\theta}} = g(\tilde{\theta}) \qquad \text{and} \qquad g(\tilde{\theta}) = 0 \Leftrightarrow \tilde{\theta} = \max_{\theta} \lambda_{\min}(B(\theta)).$$

Hence, searching for the root of $g(\theta)$ can be interpreted as searching for the maximum of the minimal eigenvalues of $B(\theta)$ with respect to θ.

REMARK 10.11. Notice that g is not necessarily continuous. A necessary condition for a jump discontinuity of g at θ_0 is a multiple smallest eigenvalue $\lambda_{\min}(B(\theta_0))$. The function g might not even have a root but jumps below zero at some θ_0. This indicates a nonunique solution of the RTLS problem (10.2); cf. [126, 131].

In this nonunique case, it follows by the interlacing theorem that $\lambda_{\min}$ is also the smallest eigenvalue of $A^T A + \theta_0 L^T L$ corresponding to an eigenvector v. Then $\bar{v} = (v^T, 0)^T \in \mathcal{E}(\theta_0)$ is an eigenvector of $B(\theta_0)$. For the Rayleigh quotient of N at $\bar{v}$ it holds that $\rho(\bar{v}) = (\bar{v}^T N \bar{v})/(\bar{v}^T \bar{v}) = \|Lv\|^2 > 0$; cf. Theorem 10.9. With $g(\theta_0) < 0$, there exist $w \in \mathcal{E}(\theta_0)$, with $\rho(w) = g(\theta_0) < 0$, and a nonzero last component. Thus, the space $\mathcal{E}(\theta_0)$ consists of two parts: the eigenspace corresponding to $\lambda_{\min}(B(\theta_{0-}))$ with only zero last components and the one-dimensional eigenspace corresponding to $\lambda_{\min}(B(\theta_{0+}))$ with nonzero last component. The RTLS solutions can be constructed from a linear combination of $\bar{v}$ and w such that $\rho(\alpha\bar{v} + \beta w) = 0$ with suitable parameters $\alpha, \beta \in \mathbb{R}$.

10.3. Computational methods for RTLS. Typically, the occurring eigenproblems in (10.8) and (10.10) are solved by inverse iteration, Rayleigh quotient iteration, the implicitly restarted Lanczos method presented in Section 3, or alternatively, by second-order Krylov subspace solvers that are briefly presented here. Thus, the only information that can be recycled from previous iterations in these methods is the eigenvector of the preceding step that can be used as initial vector. Much more information can be exploited when using general iterative projection methods such as the Nonlinear Arnoldi method from Section 4.3, which can be initialized with the entire search space of the previous eigenvalue problem.

10.3.1. RTLSQEP. In this section we discuss different approaches for solving the sequence of quadratic eigenvalue problems (10.8). For large-scale problems, iterative projection methods are quite efficient, where in each step the underlying problem (10.8) is projected to a search space $\mathcal{V} = \text{span}\{V\}$, which is expanded until the approximation obtained by solving the projected problem

$$V^T\Big((W_k + \lambda I)^2 - \delta^{-2} h_k h_k^T\Big) V u = 0$$

is sufficiently accurate. Expanding the subspace by some vector v obviously only requires appending a new vector $W_k v$ and a new component $h_k^T v$ to the current projected matrix $W_k V$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

and the vector $h_k^T V$, respectively. Hence, in these algorithms one does not need the explicit matrix W_k but only a procedure to evaluate $W_k v$ for a given vector v.

A straightforward approach for solving the QEP (10.8) at the kth iteration step of Algorithm 12 is linearization, i.e., solving

$$\begin{bmatrix} -2W_k & -W_k^2 + \delta^{-2} h_k h_k^T \\ I & 0 \end{bmatrix} \begin{bmatrix} \lambda u \\ u \end{bmatrix} = \lambda \begin{bmatrix} \lambda u \\ u \end{bmatrix},$$

for the maximal real eigenvalue and the corresponding u-part of the eigenvector, which is an eigenvector of (10.8). In [199, 200], methods especially designed for obtaining the rightmost eigenvalue of a matrix are presented, which are based on an approximation of the matrix exponential and which allows the rightmost eigenvalue to be captured more easily with an iterative projection method due to better separation. However, it is a drawback of linearization that symmetry properties of the quadratic problem are destroyed. Two structure-preserving linearizations are given in (5.1) in Section 5.2 leading to symmetric generalized linear eigenproblems. In the following, three structure-preserving methods suited for solving the sequence of QEPs are presented, which are not based on linearization.

A Krylov subspace-type method for monic QEP. Li and Ye [152] presented a Krylov subspace projection method for the monic quadratic eigenproblem $(\lambda^2 I - \lambda A_1 - A_0)u = 0$, with $A_1, A_0 \in \mathbb{R}^{n\times n}$. The method has particularly favorable properties if some linear combination of A_1 and A_0 is a matrix of small rank q. Then with $\ell + q + 1$ steps of an Arnoldi-type process, a matrix $Q \in \mathbb{R}^{n\times \ell+q+1}$ with orthonormal columns and two matrices $H_1 \in \mathbb{R}^{\ell+q+1\times \ell}$ and $H_0 \in \mathbb{R}^{\ell+q+1\times \ell}$ with lower bandwidth $q+1$ are determined such that

$$A_1 Q(:, 1:\ell) = Q(:, 1:\ell+q+1)H_1 \qquad \text{and} \qquad A_0 Q(:, 1:\ell) = Q(:, 1:\ell+q+1)H_0,$$

and approximations of eigenpairs of the quadratic eigenproblem are obtained from its orthogonal projection onto $\operatorname{span}\{Q(:, 1:\ell)\}$, which reads

$$(\lambda^2 I_\ell - \lambda H_1(1:\ell, :) - H_0(1:\ell, :))z = 0.$$

With regard to the QEP (10.8), it holds that $A_1 = 2W_k$ and $A_0 = W_k^2 - \delta^{-2} h_k h_k^T$. Usually no linear combination of A_1 and A_0 is of small rank, and thus the matrices H_0 and H_1 will become full. By applying $\ell + 2$ steps of the algorithm of Li and Ye with $A_1 = W_k$ and $A_0 = h_k h_k^T$, one obtains a matrix $Q \in \mathbb{R}^{n\times \ell+2}$ with orthonormal columns such that

$$A_i Q(:, 1:\ell) = Q(:, 1:\ell+2)H_i(1:\ell+2, 1:\ell), \qquad \text{for } i = 1, 2.$$

Hence, the orthogonal projection of problem (10.8) to $\mathcal{Q} := \operatorname{span}\{Q(:, 1:\ell)\}$ reads

$$\left(\lambda^2 I - 2\lambda H_1(1:\ell, 1:\ell) - \hat{H}_0\right)\tilde{u} = 0,$$

with $\hat{H}_0(1:\ell, 1:\ell) = \delta^{-2} H_0(1:\ell, 1:\ell) - H_1(1:\ell+2, 1:\ell)^T H_1(1:\ell+2, 1:\ell)$. As a consequence of $\operatorname{rank}\{A_0\} = q = 1$, it follows that H_1 and $\hat{H}_0$ are symmetric pentadiagonal matrices, and the cost for expanding the subspace $\mathcal{Q}$ by one vector is essentially one matrix-vector product (MatVec); cf. [152].

Second-Order Arnoldi Reduction for QEP. The Second-Order Arnoldi Reduction (SOAR) has been introduced by Bai and Su [18] for solving large scale QEP of the form $(\lambda^2 M + \lambda D + K)u = 0$, with $M, D, K \in \mathbb{R}^{n\times n}$, assuming a regular M. The main idea is based on the observation that the information of the Krylov space $\mathcal{K}_\ell$ of the linearization

$$\begin{bmatrix} A_1 & A_0 \\ I & O \end{bmatrix} \begin{bmatrix} \lambda u \\ u \end{bmatrix} = \lambda \begin{bmatrix} \lambda u \\ u \end{bmatrix}, \tag{10.11}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

with $A_1 = -M^{-1}D$, $A_0 = -M^{-1}K$, and the initial vector $[r_0^T, 0]^T$, with $r_0 \in \mathbb{R}^n$, is contained in the second-order Krylov space

$$\mathcal{G}_\ell(A_1, A_0; r_0) = \text{span}\{r_0, r_1, \ldots, r_{\ell-1}\},$$

where

$$\begin{aligned} r_1 &= A_1 r_0, \\ r_j &= A_1 r_{j-1} + A_0 r_{j-2}, \qquad \text{for } j \geq 2. \end{aligned}$$

The orthogonal projection of the QEP onto $\mathcal{G}_\ell(A_1, A_0; r_0)$ is the structure-preserving variant of projecting the linearized problem (10.11) onto $\mathcal{K}_\ell$.

Since the QEPs (10.8) are monic, there is no need to perform a LU decomposition of the matrix $M = I$, and the matrices $A_1 = -2W_k$ and $A_0 = -W_k^2 + \delta^{-2} h_k h_k^T$ are directly available. The current second-order Krylov space $\mathcal{G}_\ell(A_1, A_0; r_0)$ is expanded by $\tilde{q} := A_1 q_\ell + A_0 p_\ell$, where $p_\ell = Q_\ell s_\ell$ is some vector $p_\ell \in \text{span}\{Q_\ell\}$. Orthogonalization yields the direction of the new basis element

$$q_{\ell+1} = (I - Q_\ell Q_\ell^T)(A_1 q_\ell + A_0 p_\ell) = (I - Q_\ell Q_\ell^T)(-2W_k q_\ell - W_k^2 Q_\ell s_\ell + \delta^{-2} h_k h_k^T Q_\ell s_\ell),$$

where $W_k Q_\ell s_\ell$ can be updated from the previous step. Hence, expanding the search space $\mathcal{G}_\ell(A_1, A_0; r_0)$ requires 2 MatVecs. Thus, a single step of the SOAR method costs essentially twice as much as the one of the Krylov-type method of Li and Ye. A variant of SOAR that is approximating the second-order Krylov space $\mathcal{G}_\ell(A_1, A_0; r_0)$ only by $\mathcal{G}_\ell(A_1, \tilde{A}_0; r_0)$, with $\tilde{A}_0 = \delta^{-2} h_k h_k^T$, is suggested in [133], which reduces the cost to 1 MatVec per search space expansion. Since a sequence of converging QEPs has to be solved, it is favorable to use the solution vector of the preceding QEP as initial vector of the current Krylov subspace.

REMARK 10.12. It has been shown in [157] that the two-level orthogonal Arnoldi procedure (TOAR) for QEPs has a much better numerical stability behavior, i.e., it is backward stable when computing the corresponding orthonormal basis. In [121] and [229], this two-level orthogonal Arnoldi approach has been extended to general polynomial eigenvalue problems, which is beneficial in terms of numerical stability and memory savings compared to the higher-order methods using straightforward linearization in [70, 125].

Nonlinear Arnoldi method for QEPs. For the two Krylov-type subspace methods above, the only degree of freedom is the choice of the initial vector, whereas the Nonlinear Arnoldi method allows thick restarts, i.e., when solving $T_k(\lambda)u = 0$ in step k, Algorithm 7 can be initialized with the orthonormal basis V that was used in the preceding step when determining the solution $u^{k-1} = V\tilde{u}$ of $V^T T_{k-1}(\lambda) V\tilde{u} = 0$.

The projected eigenvalue problem

$$V^T T_k(\mu) V\tilde{u} = ((W_k + \mu I)V)^T((W_k + \mu I)V)\tilde{u} - \delta^{-2}(h_k^T V)^T(h_k^T V)\tilde{u} = 0$$

can be updated cheaply by appending one column and one entry to the current matrices and vector, respectively, at the essential cost of 1 MatVec with W_k in every iteration step. The determination of the residual $r = T_k(\mu)V\tilde{u}$ costs another MatVec with W_k.

The considerations above demonstrate that due to search space recycling, it is rather inexpensive to provide $V^T T_k(\lambda) V$ if $V^T T_{k-1}(\lambda) V$ is known. This suggests early updates, i.e., to leave the inner loop of the Nonlinear Arnoldi method for determining the rightmost eigenpair long before convergence. Fast convergence typically is obtained even without preconditioning, i.e., with $M = I$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 10.2
Example heat(1), average CPU time in seconds.

noise	n	QEPs LiYe	QEPs SOAR	QEPs NLArn	EVPs NLArn
1%	1000	0.47	0.63	0.36	0.19
	2000	1.19	1.02	0.99	0.60
	4000	4.68	3.78	3.88	2.65
10%	1000	0.46	0.45	0.32	0.19
	2000	1.18	0.99	0.98	0.61
	4000	4.67	3.73	3.92	2.54

10.3.2. RTLSEVP. Renaut and Guo [194] proposed to determine the minimum eigenvalue of $B(\theta_k)y = \big(M + \theta_k N\big)y = \lambda y$ in Algorithm 13 via the Rayleigh quotient iteration, initialized by the eigenvector found in the preceding iteration step. Although one uses information from the previous step, an obvious drawback of this method is the fact that each iteration step requires $\mathcal{O}(n^3)$ operations providing the LU factorization of $B(\theta_k)$. An efficient root-finding algorithm for solving $g(\theta) = 0$ is suggested and analyzed in [131, 133], also covering the case of a discontinuity at the root.

Nonlinear Arnoldi method for EVPs. Similarly to the approach in RTLSQEP, the entire information gathered in previous iteration steps can be employed solving (10.8) via the Nonlinear Arnoldi method with thick restarts applied to

$$T_k(\mu)u := (M + \theta_k N - \mu I)u = 0.$$

This time, in lines 1 and 8 of Algorithm 7, we aim at the minimum eigenvalue of $T_k(\mu)$. The projected problem

$$V^T T_k(\mu) V \tilde{u} = \big(([A,b]V)^T([A,b]V) + \theta_k V^T N V - \mu I\big)\,\tilde{u} = 0$$

can be updated efficiently if the search space is expanded by a new vector and if the iteration counter k is increased; i.e., a new θ_k is chosen. Thereby, the explicit form of the matrices M and N is not needed. In case of a sparse or structured regularization matrix L, the essential cost for determining the projected problem is 1 MatVec with $[A,b]$. The evaluation of the residual $r = T_k(\mu)V\tilde{u}$ in lines 5 and 11 costs another MatVec with $[A,b]^T$. Hence, one inner iteration step of the Nonlinear Arnoldi in RTLSEVP costs 2 MatVecs, which is half the cost of an inner iteration step of the Nonlinear Arnoldi applied in the RTLSQEP.

For the preconditioner in line 2, it is appropriate to chose $M \approx N^{-1}$, which usually can be implemented very cheaply and can be kept constant throughout the whole algorithm.

10.4. Numerical examples. To evaluate the performance of Algorithms 12 and 13 for large dimensions, we use 1D and 2D test examples from Hansen's *Regularization Tools* [98]. Two functions *heat(1)* and *tomo*, which are both discretizations of integral equations, are used to generate matrices $A \in \mathbb{R}^{m\times n}$, right-hand sides $b \in \mathbb{R}^m$, and solutions $x \in \mathbb{R}^n$ such that $Ax = b$. In all cases the matrices A and $[A,b]$ are ill-conditioned. In all examples, we let $m = 2n$, and a certain level of white noise has been added to the data. The numerical tests were run on a PentiumR4 computer with 3.4 GHz and 8GB RAM using MATLAB R2007b [133].

When using the RTLSQEP for problem *heat(1)* (see Table 10.2), roughly 100 MatVecs are carried out in about 3 outer iterations. This is the case for all tested eigensolvers, both noise levels, and different problem sizes. A matrix vector multiplication is the most expensive

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 10.3
Example tomo, average CPU time in seconds.

noise	n	QEPs LiYe	QEPs SOAR	QEPs NLArn	EVPs NLArn
1%	30x30	0.77	1.01	1.02	0.20
	40x40	2.62	2.55	2.07	0.54
	50x50	6.93	6.44	4.78	3.86
10%	30x30	0.77	1.02	1.00	0.21
	40x40	2.63	2.56	2.02	0.56
	50x50	6.89	6.38	4.80	3.83

operation within these algorithms, so the computation times are about equal. RTLSEVP requires approximately 50 MatVecs, which results in roughly half the computation time.

For the 2D problem *tomo* (see Table 10.3), RTLSQEP needs roughly 200–300 MatVecs due to a large number of outer iteration steps. The computation time is much shorter when using the RTLSEVP algorithm with about 60 MatVecs for the smaller problems and about 150 MatVecs for the 50×50 example. Note that in all examples the residuals and relative errors of the computed solutions are similar for the investigated methods.

11. Dual regularized total least-squares. Image reconstruction typically involves solving a (linear) inverse problem. In case the blurring matrix A and the blurred image b are contaminated by some noise, the total least-squares (TLS) method is an appropriate choice. Often this problem is ill-posed, thus regularization is necessary to stabilize the computed solution. Adding a quadratic constraint yields the regularized total least-squares (RTLS) problem (10.2) that has been discussed in Section 10. In case estimates for the norms of the errors in A and b are available, the dual RTLS (DRTLS) problem is obtained (cf. [135, 158, 159, 223]):

$$\|Lx\| = \min! \quad \text{subject to} \quad (A + \Delta A)x = b + \Delta b,\ \|\Delta b\| \le h_b,\ \|\Delta A\| \le h_A,$$

with given h_A, h_b, which requires solving a sequence of linear problems.

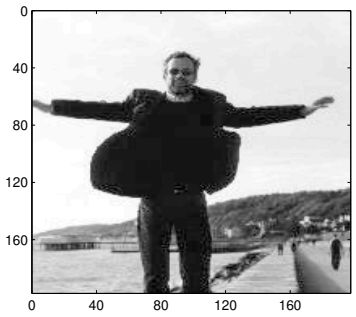
We consider a reconstruction of a greyscale image that is represented by an array of $n \times n$ pixels, with $n = 197$. The pixels are stored columnwise as a vector in $\mathbb{R}^N$ with $N = n^2$. Let the vector x_{true} represent the original image. A block Toeplitz blurring matrix $A_{true} \in \mathbb{R}^{N \times N}$ with Toeplitz blocks is determined by the function *blur* from [98] using the parameter values $band = 3$, which is the half-bandwidth of each $n \times n$ Toeplitz block, and $\sigma = 1.5$, which determines the width of the underlying Gaussian point spread function. Thus, the matrix $A_{true} \in \mathbb{R}^{38,809 \times 38,809}$ has $9.6 \cdot 10^5$ nonzero entries. Gaussian noise of level $\sigma = 10^{-4}$ is added to A_{true} and $b_{true} = A_{true} x_{true}$. Two regularization matrices L are compared, i.e., the first- and second-order discrete derivative operator for two space dimensions,

$$L_{1,2D} = \begin{bmatrix} L_1 \otimes I_n \\ I_n \otimes L_1 \end{bmatrix}, \quad L_{2,2D} = \begin{bmatrix} L_2 \otimes I_n \\ I_n \otimes L_2 \end{bmatrix},$$

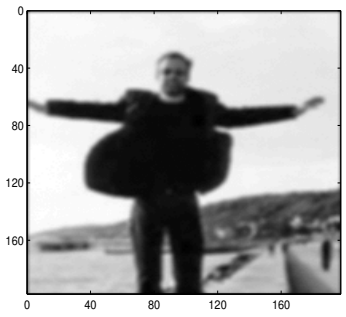
$$\text{with} \quad L_1 = \begin{bmatrix} -1 & 1 & & \\ & \ddots & & \\ & & -1 & 1 \end{bmatrix} \in \mathbb{R}^{(n-1) \times n}, \quad L_2 = \begin{bmatrix} -1 & 2 & -1 & & \\ & \ddots & \ddots & & \\ & & -1 & 2 & -1 \end{bmatrix} \in \mathbb{R}^{(n-2) \times n}.$$

The RTLSEVP Algorithm 13 from Section 10.2 and the generalized Krylov subspace dual RTLS method introduced in [135] have been applied for solving the RTLS and the corresponding dual RTLS problem, respectively.

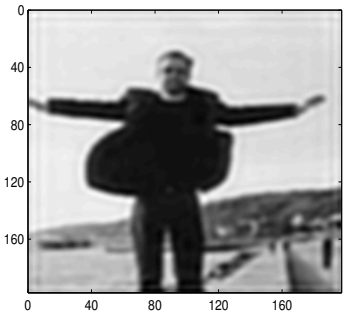
ETNA
Kent State University and
Johann Radon Institute (RICAM)



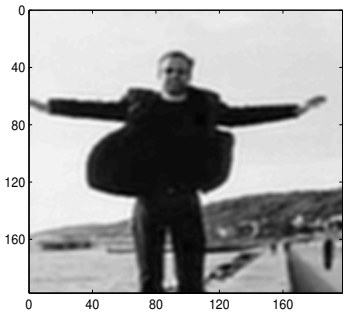
(a) Original picture



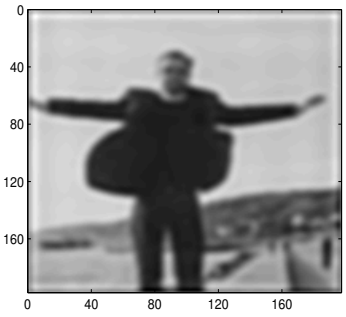
(b) Blurred and noisy picture



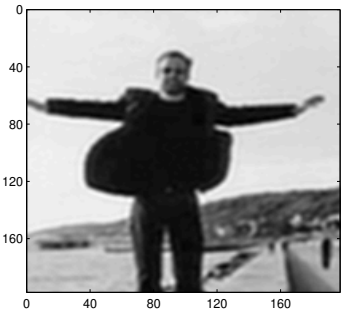
(c) Restored by RTLS with $L = L_{1,2D}$



(d) Restored by DRTLS with $L = L_{1,2D}$



(e) Restored by RTLS with $L = L_{2,2D}$



(f) Restored by DRTLS with $L = L_{2,2D}$

FIG. 11.1. *Original, blurred, and restored Lothar [135].*

Figure 11.1 displays the original (blur- and noise-free) image, the blurred and noisy image, and several reconstructions. The first row of Figure 11.1 depicts the original image as well as the blur- and noise-perturbed image. The relative error of the blurred and noisy image is $\|b - x_{true}\|/\|x_{true}\| = 20.46\%$. The images restored in the second row are obtained by using the discrete first-order derivative operator $L_{1,2D}$. The reconstructed images have relative errors of $\|x_{RTLS}^{L1,2D} - x_{true}\|/\|x_{true}\| = 7.45\%$ and $\|x_{DRTLS}^{L1,2D} - x_{true}\|/\|x_{true}\| = 6.20\%$, when using search space dimensions of 42 and 32, respectively. The last row of Figure 11.1

ETNA
Kent State University and
Johann Radon Institute (RICAM)

displays two restored images obtained with the discrete Laplace operator $L = L_{2,2D}$; the first one corresponds to RTLSEVP with a relative error of 9.55%, while the DRTLS restoration is $\|x_{DRTLS}^{L2,2D} - x_{true}\|/\|xtrue\| = 6.34\%$, both with search space dimensions of 41. The quality of the restorations obtained by the DRTLS method with $L_{1,2D}$ and $L_{2,2D}$ is about the same, whereas the corresponding restorations by RTLS are clearly inferior. We find the images obtained with $L_{1,2D}$ to be slightly sharper than the image determined with $L_{2,2D}$. Also the relative error is slightly smaller. In Section 10 total least-squares methods based on eigenproblems are discussed in more detail.

12. Electronic behavior of quantum dots. Semiconductor nanostructures have attracted tremendous attention in the past few years because of their unique physical properties and their potential for applications in micro- and optoelectronic devices. In such nanostructures, the free carriers are confined to a small region of space by potential barriers, and if the size of this region is smaller than the electron wavelength, then the electronic states become quantized at discrete energy levels. The ultimate limit of low-dimensional structures is the quantum dot, in which the carriers are confined in all three directions of space, and their electronic behavior is similar to that of an atom [20, 46, 169].

12.1. Kane formula of the electron effective mass. We consider the problem to determine a few relevant energy states of a quantum dot (InAs, e.g.) embedded in a matrix (GaAs, e.g.). According to the 8×8 $\mathbf{k} \cdot \mathbf{p}$ theory from [160], these are the eigenvalues λ of a linear eigenvalue problem

$$(H + V)\phi = \lambda\phi,$$

where ϕ contains as its components wave functions of the electron, heavy-, light-, and spin-orbit split-off hole bands (each of them appearing twice due to the spin), and the Hamiltonian operator H is an 8×8 matrix containing Hamiltonian operators of the subbands in its diagonal and the coupling of the subbands in its off-diagonal elements.

Projecting the 8×8 Hamiltonian onto the conduction band results in a single Schrödinger equation describing the electronic behavior of an electron in the quantum dot [20, 46, 151, 160],

$$-\nabla \cdot \left(\frac{\hbar^2}{2m(x,\lambda)} \nabla\Phi \right) + V(x)\Phi = \lambda\Phi, \qquad x \in \Omega_q \cup \Omega_m. \tag{12.1}$$

Here Ω_q and Ω_m is the region occupied by the quantum dot and the matrix (the surrounding material), respectively, $\hbar$ is the reduced Planck constant, and $V(x)$ is the confinement potential. The electron effective mass becomes energy dependent and is given by the Kane formula [115]:

$$m(x,\lambda) = \begin{cases} m_q(\lambda), & x \in \Omega_q, \\ m_m(\lambda), & x \in \Omega_m, \end{cases}$$

where

$$\frac{1}{m_j(\lambda)} = \frac{P_j^2}{\hbar^2} \left(\frac{2}{\lambda + g_j - V_j} + \frac{1}{\lambda + g_j - V_j + \delta_j} \right), \qquad j \in \{m, q\}, \tag{12.2}$$

V_j is the confinement potential, P_j the momentum, g_j the main energy gap, and δ_j the spin-orbit splitting in the jth region. Other types of effective masses (taking into account the effect of strain, e.g.) appear in the literature. They are all rational functions of λ where $1/m(x,\lambda)$ is monotonically decreasing with respect to λ.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Because the wave function Φ is essentially concentrated on the quantum dot, we assume homogeneous Dirichlet conditions on the outer boundary of the matrix. On the interface, we consider the so called BenDaniel-Duke conditions

$$\frac{1}{m_m}\frac{\partial\Phi}{\partial n}\bigg|_{\partial\Omega_m} = \frac{1}{m_q}\frac{\partial\Phi}{\partial n}\bigg|_{\partial\Omega_q}, \tag{12.3}$$

which guarantee continuity of the wave function Φ on the interface.

Multiplying (12.1) by $\Psi \in H_0^1(\Omega)$, $\Omega := \overline{\Omega}_q \cup \Omega_m$, and integrating by parts, one gets the weak form of the generating eigenvalue problem [240]:

Find $\lambda \in \mathbb{R}$ and $\Phi \in H_0^1(\Omega)$, $\Phi \neq 0$, such that for every $\Psi \in H_0^1(\Omega)$

$$\begin{aligned} a(\Phi,\Psi;\lambda) := {} & \frac{\hbar^2}{2}\int_{\Omega_q}\frac{1}{m_q(x,\lambda)}\nabla\Phi\cdot\nabla\Psi\,dx + \frac{\hbar^2}{2}\int_{\Omega_m}\frac{1}{m_m(x,\lambda)}\nabla\Phi\cdot\nabla\Psi\,dx \\ & + \int_{\Omega_q} V_q(x)\Phi\Psi\,dx + \int_{\Omega_m} V_m(x)\Phi\Psi\,dx = \lambda\int_{\Omega}\Phi\Psi\,dx =: \lambda b(\Phi,\Psi), \end{aligned} \tag{12.4}$$

which can be rewritten by the Lax-Milgram lemma as Schrödinger equation (12.1) with the effective Hamiltonian

$$\hat{H} = -\frac{\hbar^2}{2}\nabla\cdot\left(\frac{1}{m(\lambda,x)}\nabla\right) + V(x).$$

12.2. Full Approximation Method. To determine the relevant energy states and the corresponding wave functions, Li et al. [151] suggested the following method given in Algorithm 14, called the Full Approximation Method (FAM). Based on many examples, they reported that the method converges.

Algorithm 14 Full Approximation Method (FAM).

Require: initial energy level λ_k^0

1: determine the effective masses $m_q(\lambda_k^0)$ and $m_m(\lambda_k^0)$ for the quantum dot and the matrix
2: **for** $n = 1, 2, \ldots$ until convergence **do**
3: determine the kth-smallest eigenvalue λ_k^n and the corresponding eigenfunction Φ_k of the linear eigenvalue problem

$$-\nabla\cdot\left(\frac{\hbar^2}{2m_j(\lambda_k^{n-1})}\nabla\Phi\right) + V(x)\Phi = \lambda\Phi,\ x \in \Omega_q \cup \Omega_m, \tag{12.5}$$

with the Ben Daniel-Duke condition (12.3)
4: update the effective masses $m_j(\lambda_k^n)$, $j \in \{m, q\}$
5: **end for**

In [239] we proved linear convergence of the FAM and its enclosure properties such that error bounds are at hand in every iteration step.

THEOREM 12.1 ([239]). *Assume that $c_j := \lambda + g_j - V_j > 0$, for $j \in \{m, q\}$, and let λ_k and $\lambda_k(\lambda_k^{n-1}))$ be the kth-smallest eigenvalue of the nonlinear Schrödinger equation (12.1) and of the parameter-dependent equation (12.5), both with the BenDaniel-Duke condition (12.3) on the interface $\Omega_q \cap \Omega_m$ and homogeneous boundary condition on the outer boundary of Ω_m.*

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Let $\lambda_k^0 \geq 0$ be any initial value, and, for $n \in \mathbb{N}$, let $\lambda_k^n := \lambda_k(\lambda_k^{n-1})$ be the kth-smallest eigenvalue of (12.5). Then it holds that

$$\lambda_k^0 < \lambda_k^2 \leq \cdots \leq \lambda_k^{2n-2} \leq \lambda_k^{2n} \leq \lambda_k \leq \lambda_k^{2n+1} \leq \lambda_k^{2n-1} \leq \ldots \lambda_k^3 \leq \lambda_k^1,$$
$$\lim_{n\to\infty} \lambda_k^n = \lambda_k,$$

and the convergence is linear, i.e., there exists a constant C, $0 < C < 1$, and $N \in \mathbb{N}$ such that

$$|\lambda_k^n - \lambda_k| \leq C|\lambda_k^{n-1} - \lambda_k|, \qquad \text{for every } n \in \mathbb{N},\ n \geq N.$$

The convergence behavior of the FAM can be improved considerably if we take advantage of the Rayleigh functional. With a and b as in equation (12.4), the function $\lambda \mapsto a(\Phi, \Phi, \lambda)$ is monotonically decreasing and positive for every fixed $\Phi \in H_0^1(\Omega)$, $\Phi \neq 0$, and therefore the real equation

$$f(\lambda; \Phi) := \lambda b(\Phi, \Phi) - a(\Phi, \Phi, \lambda) = 0 \tag{12.6}$$

has a unique positive solution $\lambda =: p(\Phi)$. Hence, equation (12.6) defines a real functional p on $H_0^1(\Omega) \setminus \{0\}$, which is the Rayleigh functional of the nonlinear eigenvalue problem (12.4).

The Rayleigh functional is defined on the whole space $H_0^1(\Omega) \setminus \{0\}$, and the conditions of the minmax Theorem 2.2 are satisfied. Hence, the following theorem holds:

THEOREM 12.2.

(i) The Schrödinger equation (12.1) modeling the quantum dot with electron effective mass m_j given in (12.2) has a countable set of eigenvalues

$$0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \ldots,$$

which all have finite multiplicity, and the only cluster point of which is ∞.

(ii) The kth-smallest eigenvalue λ_k can be characterized as

$$\lambda_k = \min_{\dim V = k} \max_{u \in V, u \neq 0} p(u). \tag{12.7}$$

These properties suggest the modification of the Full Approximation Method for computing the kth energy level of problem (12.4) presented in Algorithm 15.

Algorithm 15 Modified Full Approximation Method (MFAM).

Require: initial energy level λ_k^0

1: determine the effective masses $m_q(\lambda_k^0)$ and $m_m(\lambda_k^0)$ for the quantum dot and the matrix
2: **for** $n = 1, 2, \ldots$ until convergence **do**
3: determine the kth-smallest eigenvalue $\tilde{\lambda}_k^n$ and the corresponding eigenfunction Φ_k of the linear eigenvalue problem

$$-\nabla \cdot \left(\frac{\hbar^2}{2m_j(\lambda_k^{n-1})} \nabla \Phi \right) + V(x)\Phi = \lambda \Phi,\ x \in \Omega_q \cup \Omega_m,$$

with the BenDaniel-Duke condition (12.3)
4: determine the Rayleigh functional $\lambda_k^n =: p(\Phi_k)$ at Φ_k
5: update the effective masses $m_j(\lambda_k^n)$, $j \in \{m, q\}$
6: **end for**

ETNA
Kent State University and
Johann Radon Institute (RICAM)

The Modified Full Approximation Method (MFAM) is nothing else but the fixed point iteration $\lambda_k^{n+1} = h(\lambda_k^n) := p(\Phi_k(\tilde{\lambda}_k^{n-1}))$, and since $\Phi(\lambda_k)$ is a stationary element of p, it follows that $h'(\lambda_k) = 0$. Hence [185], the Modified Full Approximation Method converges quadratically to λ_k, i.e., there exist some constant $C > 0$ such that

$$|\lambda_k^n - \lambda_k| \le C|\lambda_k^{n-1} - \lambda_k|^2 \qquad \text{for every } n \in \mathbb{N}.$$

12.3. Iterative projection methods for quantum dot problems. If the Schrödinger equation (12.1) is discretized by a Galerkin method (finite elements, e.g.), then one gets a rational matrix eigenvalue problem

$$S(\lambda)x := \lambda Mx - \frac{1}{m_q(\lambda)}A_q x - \frac{1}{m_m(\lambda)}A_m x - Bx = 0, \tag{12.8}$$

where

$$A_j = \Big(\int_{\Omega_j} \nabla\phi_k \cdot \nabla\phi_\ell \, dx\Big)_{k,\ell}, \qquad j \in \{q, m\},$$

$$M = \Big(\int_{\Omega} \phi_k\phi_\ell \, dx\Big)_{k,\ell} \qquad \text{and} \qquad B = \Big(V_q \int_{\Omega_q} \phi_k\phi_\ell \, dx + V_m \int_{\Omega_m} \phi_k\phi_\ell \, dx\Big)_{k,\ell},$$

and ϕ_k denotes a basis of the ansatz space. A_q, A_m, and B are symmetric and positive semidefinite, and M is positive definite, and for $\lambda \ge 0$, the matrix

$$\frac{\hbar^2}{2m_q(\lambda)}A_q + \frac{\hbar^2}{2m_m(\lambda)}A_m$$

is positive definite. Hence, the eigenvalues of the discretized problem (12.8) satisfy a minmax principle as well, and it follows from the minmax characterization (12.7) of the nonlinear Schrödinger equation that the kth-smallest eigenvalue of the discretized problem (12.8) is an upper bound for the corresponding eigenvalue of problem (12.1).

The dimension of the discretized Schrödinger equation usually will be quite large, and the FAM method will not be a reasonable choice since in every iteration step a linear eigenvalue problem has to be solved.

EXAMPLE 12.3. Consider a pyramidal quantum dot with width 12.4 nm and height 6.2 nm embedded in a cubic matrix of size 24.8 nm×24.8 nm×18.6 nm with the following parameters $P_q = 0.8503$, $g_q = 0.42$, $\delta_q = 0.48$, $V_q = 0$, $P_m = 0.8878$, $g_m = 1.52$, $\delta_m = 0.34$, and $V_m = 0.7$. This model was already treated by Hwang, Lin, Wang, and Wang in [111], who multiplied the rational eigenvalue problem by its common denominator to obtain a polynomial eigenvalue problem of degree 5. After discretizing with linear elements on an equidistant grid, they obtained a huge non-symmetric linear eigenvalue problem.

In [240] we considered the following finite element model: Using FEMLAB [48] we discretized (12.4) by cubic Lagrangian elements on a tetrahedral grid with 96,640 degrees of freedom such that 43,615 DoFs where located in the quantum dot, 43,897 DoFs in the matrix, and 9,128 DoFs on the interface.

We determined the 5 smallest eigenvalues of the rational eigenproblem (12.8) by the Nonlinear Arnoldi method and the Jacobi-Davidson method. We started the methods with a constant vector on $\overline{\Omega}_q \cup \Omega_m$ which is far away from an eigenvector, and we terminated the iteration for an eigenvalue if the residual norm was less than 10^{-8}. Table 12.1 contains

ETNA
Kent State University and
Johann Radon Institute (RICAM)

the approximations of the five smallest eigenvalues, the number of iterations to obtain these approximations, and the CPU times using MATLAB 7.0.4 on an AMD Opteron processor with 4 GByte RAM and 2.2 GHz. Notice that one gets upper bounds for the corresponding eigenvalues of problem (12.4).

TABLE 12.1
Five smallest eigenvalues of a finite element discretization of dimension 96,640.

dim	λ_1	λ_2	λ_3	λ_4	λ_5	CPU time
96,640	0.39779	0.57411	0.57411	0.68547	0.69714	
Arnoldi	44 it.	29 it.	29 it.	24 it.	21 it.	189 sec.
JD	9 it.	7 it.	9 it.	5 it.	6 it.	205 sec.

The following Figure 12.1 contains the convergence history of the Jacobi-Davidson and of the Nonlinear Arnoldi method. Notice that the second/third eigenvalue is a double eigenvalue, which does not influence the rate of convergence of both methods.

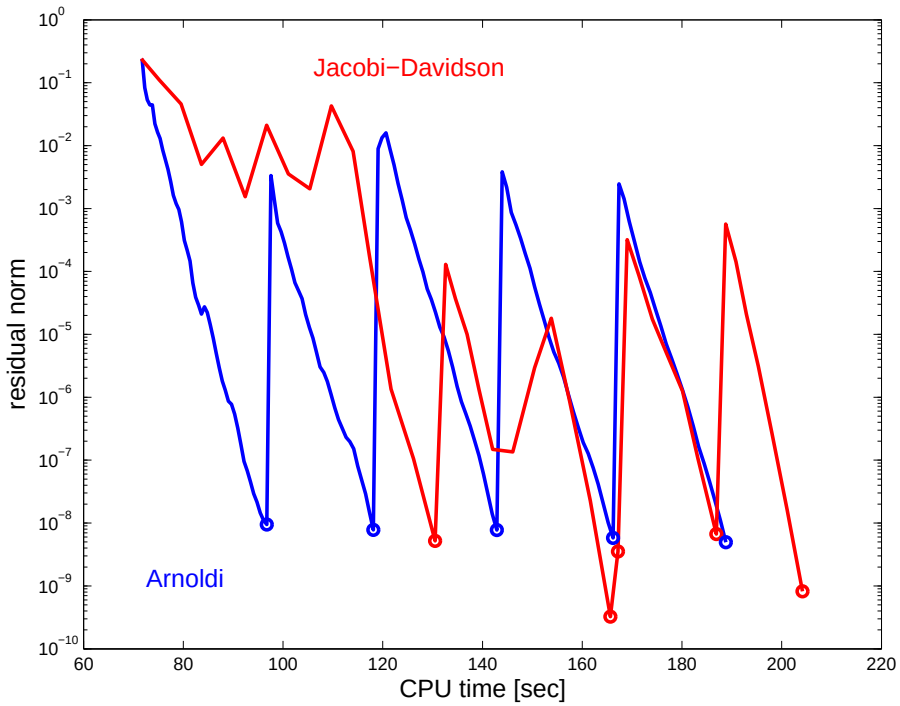


FIG. 12.1. *Convergence history of Example 12.3.*

We also solved the problem with an incomplete LU preconditioner with cut-off threshold τ. Table 12.2 contains the CPU time for determining the five smallest eigenvalues and for the preconditioner for several values of τ. The behavior is typical for many examples: For $\tau = 0.1$, the Jacobi-Davidson method is much faster than the Nonlinear Arnoldi method, but as τ descends, the Nonlinear Arnoldi method becomes the clear winner.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 12.2
Computational times for different incomplete LU preconditioners and solvers.

τ	JD	Arnoldi	precond.
0.1	261.4	1084.1	3.4
0.01	132.7	117.1	71.7
0.001	118.9	61.2	246.6

REMARK 12.4. We considered here the quantum dot problem (12.1), (12.2), and (12.3), which is overdamped. More general quantum dot problems are described in [33], where we considered stationary Schrödinger equations in the presence of spin-orbit splitting (a non-overdamped problem), in [34], which examines coupled quantum dots on wetting layers, and in [36] for the stationary Schrödinger equation governing electronic states of quantum dots and rings in magnetic fields.

13. Viscoelastic damping. The nature of energy dissipation mechanisms in a vibrating structure has always been very difficult to explain: damping models have been developed trying to fit experimental and mathematical results. The viscous approach proposed by Lord Rayleigh assuming that dissipative forces are proportional to the velocity of the systems degrees of freedom is the damping model used for the great majority of structural solid materials (metal, concrete, wood, glass, etc.) [4].

The weakness of the pure viscous model becomes evident when applying it to the so-called viscoelastic materials such as polymer derivatives and rubbers or rubber-like materials, which are characterized by a time-dependent constitutive model and by frequency-dependent Young's and shear moduli. Viscoelastic damping is introduced into the system assuming that the dissipative forces are proportional to the history of the velocity via kernel hereditary functions [4, 5, 114].

For small displacements, the most general form of a viscoelastic damped oscillator becomes [4, 5]

$$M\ddot{u}(t) + \int_0^t \mathcal{G}(t-\tau)\dot{u}(\tau)d\tau + Ku(t) = f(t), \tag{13.1}$$

together with initial conditions $u(0) = u_0$, $\dot{u}(0) = \dot{u}_0$, where $u \in \mathbb{R}^N$ is the displacement vector, $f \in \mathbb{R}^N$ is the forcing vector, $M \in \mathbb{R}^{N\times N}$ and $K \in \mathbb{R}^{N\times N}$ are the positive semi-definite mass and stiffness matrix, respectively, and $\mathcal{G} \in \mathbb{R}^{N\times N}$ is the symmetric kernel function of damping.

The modes of the system can be determined as non-trivial solutions of the free-motion problem. With functions of the form $u(t) = ue^{st}$, we get

$$T(s)u := \Big(s^2M + s\sum_{k=1}^{n} G_k(s)C_k + K\Big)u = 0, \tag{13.2}$$

where $C_k \in \mathbb{R}^{N\times N}$ are the symmetric coefficient matrices of frequency-dependent damping and $G_k(s)$ are the frequency-dependent nonviscous damping functions.

Many different damping models have been proposed to describe the dissipative behavior of viscoelastic materials, the Biot model [41], the inelastic displacement field model [146], exponential damping [4], and the generalized Maxwell model [193], to name just a few. In practice, mechanical engineering systems with two or more parts with significantly different levels of energy dissipation are encountered frequently in dynamical design, so these damping

ETNA
Kent State University and
Johann Radon Institute (RICAM)

systems often involve multiple damping models [55, 148, 150]. It was noticed in [119] that if the term $sG_k(s)C_k$ appearing in (13.2) is rational, then it can be transformed equivalently to $s\tilde{G}_k(s)\tilde{C}_k$, where

$$\tilde{G}_k(s) := \frac{\mu_k}{s+\mu_k} \tag{13.3}$$

is the exponential damping model.

The following results for the structure of the set of nonviscous eigenvalues of (13.2) with damping (13.3) are contained in [175, 176]. We consider the parameter-dependent problem

$$T(s;\gamma)u := \Big(s^2M + \gamma s\sum_{j=1}^{n}\frac{\mu_j}{s+\mu_j}C_j + K\Big)u = 0, \qquad \gamma > 0, \tag{13.4}$$

with $0 < \mu_1 < \mu_2 < \cdots < \mu_n$. Then it follows from Wagner and Adhikari [253] that for every $\gamma > 0$, with $r_j = \operatorname{rank}(C_j)$, the problem (13.4) has $r := \sum_{j=1}^{n} r_j$ real eigenvalues $\lambda(\gamma)$, and clearly each of them depends continuously on γ.

For sufficiently small $\gamma > 0$, the problem (13.4) has exactly r_j eigenvalues in the interval $I_j = (-\mu_j, -\mu_{j-1})$, which are close to $-\mu_j$, for $j = 1, \ldots, n$, with $\mu_0 = 0$, and the following monotonicity result holds:

LEMMA 13.1. *Let $0 < \gamma_1 < \gamma_2$, and assume that I_j contains an ith eigenvalue $\lambda_i^{(j)}(\gamma_2)$. Then it holds that*

$$\lambda_i^{(j)}(\gamma_1) \le \lambda_i^{(j)}(\gamma_2).$$

As γ increases all eigenvalues $\lambda_i^{(j)}$ in I_j grow, and this holds true in particular for the maximal eigenvalue $\lambda_1^{(j)}$. It may happen that $\lambda_1^{(j)}(\gamma)$ is bounded away from μ_{j-1} (this is for instance always the case for $j = 1$) or there exists a $\bar{\gamma}$ such that $\lim_{\gamma\to\bar{\gamma}-0}\lambda_1^{(j)}(\gamma) = \mu_{j-1}$. Then for $\gamma > \bar{\gamma}$ sufficiently close to $\bar{\gamma}$, there does no longer exist a first eigenvalue of $T(\cdot;\gamma)$ in I_j, but (due to the existence and continuity of eigenvalues of the linearization of $T(\cdot;\gamma)$ considered in [253]) a new eigenvalue of $T(\cdot;\gamma)$ appears in the interval I_{j-1}, and it follows from the maxmin characterization of the eigenvalues in I_{j-1} that this must be a $(r_{j-1}+1)$st eigenvalue unless the interval I_{j-1} is free of eigenvalues.

After the first eigenvalue $\lambda_1^{(j)}$ has passed μ_{j-1}, the same may happen for further eigenvalues. The following theorem summarizes preliminary results for the distribution of the real eigenvalues of $T(\cdot)$ found so far:

THEOREM 13.2 ([176]). *Consider the viscoelastic vibration problem*

$$T(\lambda)u := \Big(\lambda^2M + \lambda\sum_{j=1}^{n}\frac{\mu_j}{\lambda+\mu_j}C_j + K\Big)u = 0, \tag{13.5}$$

where the general conditions for K, M, C_j, and μ_j given in (13.1), (13.2), and (13.4) are satisfied. Assume that $r_j = \operatorname{rank}(C_j)$ and that the interval $I_j = (-\mu_j, -\mu_{j-1})$ contains s_j real eigenvalues of the problem (13.5).

Then the following statements are true:

- *(i) Each interval $I_j = (-\mu_j, -\mu_{j-1})$, $j = 1, \ldots, n$, contains at most N eigenvalues.*
- *(ii) If $r_j = N$, for every $j \in \{1, \ldots, n\}$, then each interval $(-\mu_j, -\mu_{j-1})$ contains exactly N eigenvalues.*
- *(iii) The interval $(-\mu_j, 0)$ contains at least $\sum_{k=1}^{j} r_k$ eigenvalues.*

ETNA
Kent State University and
Johann Radon Institute (RICAM)

(iv) In particular, the interval $I_1 = (-\mu_1, 0)$ *contains at least* r_1 *eigenvalues* $\lambda_k^{(1)}$, $k = 1, \dots, s_1$, $r_1 \le s_1 \le N$, *and all eigenvalues in* I_1 *can be characterized as*

$$\lambda_k^{(1)} = \max_{V \in S_k, V \subset D_1} \inf_{u \in V} p_1(u), \qquad k = 1, \dots, s_1.$$

(v) Assume that I_1 *contains* $s_1 > r_1$ *eigenvalues and* $r_2 > s_1 - r_1$. *Then the eigenvalues* $\lambda_k^{(2)} \in (-\mu_2, -\mu_1)$, $k = s_1 - r_1 + 1, \dots, s_2$, *allow for a variational characterization*

$$\lambda_k^{(2)} = \max_{V \in S_k, V \cap D_2 \neq \emptyset} \inf_{u \in V \cap D_2} p_2(u).$$

(vi) The interval $(-\mu_n, -\mu_j)$ *contains at most* $\sum_{k=j+1}^{n} r_k$ *eigenvalues; in particular, the interval* $(-\mu_n, \mu_{n-1})$ *contains at most* r_n *eigenvalues.*

(vii) If I_n *contains* s_n, $0 < s_n \le r_n$, *eigenvalues, then they allow for the characterization*

$$\lambda_k^{(n)} = \sup_{V \in S_k, V \cap D_n \neq \emptyset} \inf_{u \in V \cap D_n} p_n(u), \qquad k = r_n, r_n - 1, \dots, r_n - s_n + 1.$$

In the following we present improved results on whether all eigenvalues stay in an interval I_j or can leave it as the parameter γ increases.

THEOREM 13.3 ([176]). *Suppose that the value*

$$\theta_j := \left(\frac{\mu_j \min_{\|u\|=1} u^T (\sum_{k=j}^{n} C_k) u}{\mu_n - \mu_{j-1}} + \frac{\mu_{j-2} \max_{\|u\|=1} u^T (\sum_{k=1}^{j-2} C_k) u}{\mu_{j-2} - \mu_{j-1}} \right), \quad j = 2, \dots, n,$$

corresponding to I_j, $j = 2, \dots, n$, *is positive and that the set*

$$F_j = \{u : \ u^T C_j u \neq 0, \ u^T C_{j-1} u = 0\},$$

is nonempty. Then, at least the first eigenvalue in I_j *leaves this interval if*

$$\gamma > \frac{\max_{\|u\|=1} u^T \left(\mu_{j-1}^2 M + K\right) u}{\theta_j \mu_{j-1}}.$$

The following theorem provides a better understanding of the problem of staying in an interval I_j or leaving it.

THEOREM 13.4 ([176]). *With*

$$N_j(u) := \sum_{k=1}^{j-2} \frac{\mu_k u^T C_k u}{\mu_k - \mu_{j-1}} + \sum_{k=j}^{n} \frac{\mu_k u^T C_k u}{\mu_k - \mu_{j-1}},$$

the following statements hold:

(i) If $F_j^+ := \{u \in F_j : N_j(u) \le 0\} \neq \emptyset$, *then* I_j *contains at least one eigenvalue of* $T(\cdot; \gamma)$, *for every* $\gamma > 0$.

(ii) If $F_j^+ = \{u : \ u^T C_j u \neq 0\}$, *then no eigenvalue of* $T(\cdot; \gamma)$ *transfers to* I_{j-1} *as* γ *increases.*

(iii) If $F_j^- := \{u \in F_j : N_j(u) > 0\} \neq \emptyset$, *then at least one eigenvalue leaves the interval* I_j.

Theorem 13.5 demonstrates how the number of negative eigenvalues of the matrix $T(\cdot; \gamma)$ changes as an eigenvalue $\lambda(\gamma)$ crosses one of the poles $-\mu_j$.

THEOREM 13.5 ([176]). *The number of negative eigenvalues of the matrix* $T(\lambda; \gamma)$ *increases (at least) by one as* $\lambda(\gamma)$ *grows beyond* $-\mu_j$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

If for very small $\gamma > 0$, an interval I_j contains N eigenvalues $\lambda_1^{(j)} \geq \cdots \geq \lambda_N^{(j)}$, then as γ increases, one eigenvalue after another can cross μ_{j-1}, and then its number increases by one. But clearly this is impossible for $\lambda_N^{(j)}$, which is an Nth eigenvalue.

COROLLARY 13.6 ([176]). *Assume that $r_j = N$, for some $j \in \{1, \ldots, n\}$. Then I_j contains at least one eigenvalue $\lambda_N^{(j)}(\gamma)$ for every $\gamma > 0$, and the number of eigenvalues in the preceding intervals $\bigcup_{i=j+1}^{n}(-\mu_i, -\mu_{i-1})$ is $\sum_{i=j+1}^{n} r_i$, for every $\gamma > 0$, i.e., no eigenvalue can enter the interval $(-\mu_j, 0)$ from $(-\mu_n, -\mu_j)$.*

By comparing the eigenvalues of $T(\cdot;\gamma)$ and of

$$S(\lambda) := \lambda \sum_{i=1}^{n} \frac{\mu_i}{\mu_i + \lambda} C_i,$$

we can determine bounds for some of the eigenvalues which are not able to leave their initial interval I_j as γ increases.

THEOREM 13.7 ([176]). *Assume that $S(\cdot)$ has a kth eigenvalue $\kappa_k \in I_j$ and $T(\cdot)$ has a kth eigenvalue $\lambda_k^{(j)} \in I_j$. Then it holds that*

$$\kappa_k^{(j)} - \frac{1}{\gamma}\|\mu_{j+1}^2 M + K\|_2 \leq \lambda_k^{(j)} \leq \kappa_k^{(j)}.$$

We motivated problem (13.2) by viscoelastic damping appearing in very large engineering structures. However, usually the nonviscous damping does not appear in the entire structure but only in relatively small substructures. Therefore, the dimension of the numerical examples in the literature is usually quite small (3 in [6, 7, 8, 9, 149, 253], 4 in [82, 143, 144, 147, 209], and 5 in [142]). We consider an example of dimension 3, which gives us insight into the behavior of the real eigenvalues while the damping level is changing.

EXAMPLE 13.8. Let

$$M = I, \quad K = \begin{bmatrix} 3 & -2 & 0 \\ -2 & 3 & -2 \\ 0 & -2 & 3 \end{bmatrix}, \quad C_1 = e_1 e_1^T, \quad C_2 = e_2 e_2^T, \quad C_3 = I, \quad C_4 = e_3 e_3^T,$$

and $\mu = [1, 2, 3, 4]$.

Then $r_j := \operatorname{rank}(C_j) = 1$, for $j = 1, 2, 4$, and $r_3 = 3$, and for $\gamma > 0$, the rational eigenvalue problem

$$T(\lambda;\gamma) = \Big(\lambda^2 M + \lambda\gamma \sum_{j=1}^{4} \frac{\mu_j}{\lambda + \mu_j} C_j + K\Big)u = 0 \tag{13.6}$$

has $\sum_{j=1}^{4} r_j = 6$ real eigenvalues. Figure 13.1 displays these eigenvalues as γ grows from $\gamma = 0$ to $\gamma = 4$. For small γ there are r_j eigenvalues in $I_j := (-\mu_j, -\mu_{j-1})$ for $j = 1, 2, 3, 4$.

Theorem 13.3 enables us to predict whether an eigenvalue leaves its corresponding interval when γ becomes large. For example, for the second interval I_2, we have $F_2 = \{[0, 1, 0]^T\}$, and we easily calculate $\theta_2 = 4/3$. Actually, Figure 13.1 shows that the second eigenvalue has left I_2 already for some $\gamma < 0.7$. For the last interval I_4, Corollary 13.6 yields $F_4 = \emptyset$, and no eigenvalue can enter I_3 from I_4. For the interval I_3, we see that $F_3 = \{[1, 0, 0]^T, [0, 0, 1]^T\}$, and hence Lemma 13.3 yields that two eigenvalues leave the interval I_3 when γ tends to infinity.

The rational eigenvalue problem

$$S(\lambda)u := \lambda \sum_{i=1}^{4} \frac{\mu_i}{\mu_i + \lambda} C_i u = 0$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

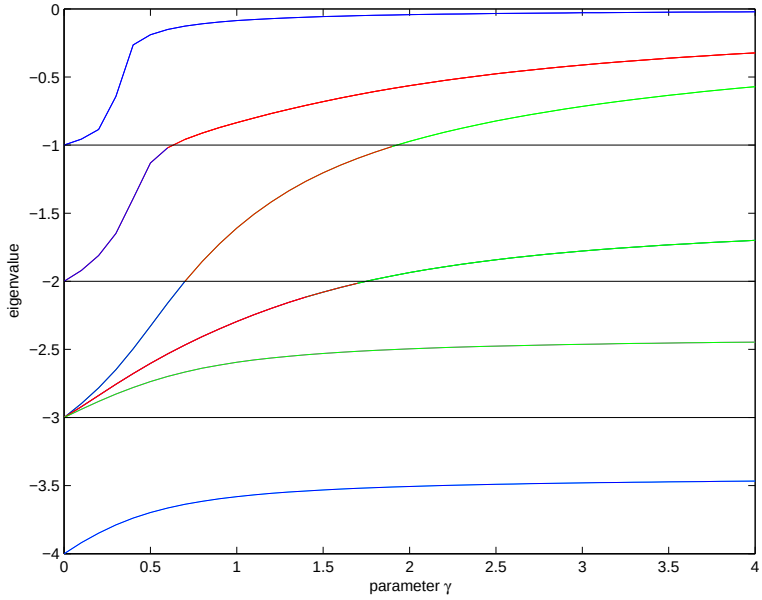


FIG. 13.1. *Real eigenvalue curves of Example 13.8. Blue, red, and green curves denote first, second, and third eigenvalues.*

has 3 negative eigenvalues $\kappa_k^{(j)}$ in $(-\mu_4, -\mu_2)$. Table 13.1 demonstrates that they are upper bounds for the corresponding eigenvalues $\lambda_k^{(j)}(\gamma)$ of (13.6) and that these eigenvalues get close to $\kappa_k^{(j)}$ for large γ.

TABLE 13.1
Bounds for real eigenvalues of Example 13.8.

j	k	$\lambda_k^{(j)}(4)$	$\lambda_k^{(j)}(10^4)$	$\kappa_k^{(j)}$
2	3	-1.699	-1.500065	-1.500000
3	3	-2.446	-2.400018	-2.400000
4	1	-3.467	-3.428586	-3.428571

14. Vibrations of fluid-solid structures. In this section vibrations of fluid-solid structures are considered that allow for a variational characterization of its eigenvalues although the corresponding eigenproblem is non-symmetric. This is an extension of the results from Section 2, where the considered eigenproblems are required to be self-adjoint. In Section 14.1 the corresponding variational characterization is derived, and in Section 14.2 we write down a finite-dimensional model. Section 14.3 discusses the Automated Multi-Level Substructuring for fluid-solid vibrations followed by a numerical example in Section 14.4.

14.1. Variational characterization of eigenvalues for a non-symmetric eigenproblem. Consider free vibrations of an elastic structure coupled with a fluid. The interaction between the structure and the fluid can significantly affect the response of the whole system and has to be taken into account properly.

Different formulations have been proposed to solve this problem: The pure displacement formulation [26], which leads to a simple symmetric eigenvalue problem, which, however, suffers from the presence of zero-frequency spurious circulation modes with no physical

ETNA
Kent State University and
Johann Radon Institute (RICAM)

meaning. These non-physical modes can be removed by modeling the fluid by the pressure field p and the structure by the displacement field u; cf. [54]. Thus, one arrives at a non-symmetric variational formulation of the problem.

Symmetric models of coupled fluid-structure vibration problems without spurious solutions have been achieved by describing the structure-acoustic system by a three-field formulation complementing the structural displacement and the fluid pressure with the fluid velocity potential [184] or the fluid displacement potential [177]. Finite element approximations based on this type of modeling are favoured today since one obtains symmetric matrix eigenvalue problems, and hence, variational characterizations of eigenvalues allow for using standard spectral approximation theory (see Babuska and Osborne [15]) to obtain convergence results for the eigenvalues and eigenvectors for Galerkin-type projection methods; cf. [30, 195].

In [216] we considered the elastoacoustic vibration problem describing the fluid by its pressure field and the structure by its displacement field. We proved that, although the resulting eigenvalue problem is non-symmetric, it shares many important properties with the symmetric model: taking advantage of a Rayleigh functional, its eigenvalues allow for the variational characterizations known from the linear theory. Namely, they can be characterized by Rayleigh's principle and are minmax and maxmin values of the Rayleigh functional.

We consider free vibrations of an elastic structure completely filled with a homogeneous, inviscid, and compressible fluid neglecting gravity effects. The fluid and the solid occupy Lipschitz domains $\Omega_f \subset \mathbb{R}^d$ and $\Omega_s \subset \mathbb{R}^d$, respectively, which we assume non-overlapping, $\Omega_f \cap \Omega_s = \emptyset$.

The boundary shall be divided by $\partial\Omega_s = \Gamma_D \cup \Gamma_I$ and $\partial\Omega_f = \Gamma_N \cup \Gamma_I$ into pairwise disjoint parts Γ_D, Γ_N, Γ_I, where Γ_D and Γ_N are Dirichlet- and Neumann-type boundaries and Γ_I is a common interface that is responsible for the coupling effect. The linear-elastic solid is modeled by its displacement function $u : \Omega_s \to \mathbb{R}^d$, $d = 2, 3$. The incompressible, inviscid, and homogeneous fluid is described by the relative pressure $p : \Omega_f \to \mathbb{R}$. This yields a formulation as a system of homogeneous time-independent partial differential equations

$$\begin{aligned} \operatorname{Div} \sigma(u) + \omega^2 \rho_s u &= 0 \qquad \text{in } \Omega_s, \\ \nabla^2 p + \frac{\omega^2}{c^2} p &= 0 \qquad \text{in } \Omega_f, \end{aligned}$$

with boundary conditions

$$u = 0 \quad \text{on } \Gamma_D \quad \text{and} \quad \nabla p \cdot n_f = 0 \quad \text{on } \Gamma_N$$

and interface conditions

$$\sigma(u)\, n - p\, n = 0 \quad \text{on } \Gamma_I \quad \text{and} \quad \omega^2 \rho_f u \cdot n + \nabla p \cdot n = 0 \quad \text{on } \Gamma_I,$$

where ω is the eigenfrequency of the vibrations, σ is the stress tensor of the solid, n_f is the unit normal vector on Γ_N, and n denotes the unit normal vector on Γ_I oriented towards the solid part. The interface boundary conditions are a consequence of an equilibrium of acceleration and force densities at the contact interface.

We assume that the fluid density $\rho_f > 0$ is constant in Ω_f and that the solid density $\rho_s : \Omega_s \to \mathbb{R}$ satisfies $0 < C_1 < \rho_s < C_2$, where C_1 and C_2 denote positive generic constants. To take into account homogeneous Dirichlet boundary conditions, we introduce the space $H^k_\Gamma(\Omega)$ for $\Gamma \subset \partial\Omega$ as the completion of $C^\infty_\Gamma(\Omega)$ in $H^k(\Omega)$, where $C^\infty_\Gamma(\Omega)$ denotes the space of infinitely differentiable functions u on Ω with $u = 0$ in a neighborhood of Γ.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

The variational forms can be obtained separately for the solid and the fluid and yield the following forms: Find $\lambda \in \mathbb{C}$ and nonzero $(u,p) \in H^1_{\Gamma_D}(\Omega_s)^d \times H^1(\Omega_f)$ such that

$$\int_{\Omega_s} \sigma(u) : \nabla v \, dx - \int_{\Gamma_I} p \, n \cdot v \, ds = \lambda \int_{\Omega_s} \rho_s u v \, dx, \tag{14.1a}$$

$$\int_{\Omega_f} \frac{1}{\rho_f} \nabla p \cdot \nabla q \, dx = \lambda \left(\int_{\Gamma_I} q \, n \cdot u \, ds + \int_{\Omega_f} \frac{1}{\rho_f c^2} p \, q \, dx \right), \tag{14.1b}$$

for all $(v,q) \in H^1_{\Gamma_D}(\Omega_s)^d \times H^1(\Omega_f)$, where $A : B = \sum_{ij} a_{ij} b_{ij}$ denotes the scalar matrix product.

We can immediately formulate the adjoint eigenvalue problem: Find $\lambda \in \mathbb{C}$ and nonzero $(u,p) \in H^1_{\Gamma_D}(\Omega_s)^d \times H^1(\Omega_f)$ such that

$$\int_{\Omega_s} \sigma(u) : \nabla v \, dx = \lambda \left(\int_{\Omega_s} \rho_s u v \, dx + \int_{\Gamma_I} p \, n \cdot v \, ds \right), \tag{14.2a}$$

$$\int_{\Gamma_I} q \, n \cdot u \, ds + \int_{\Omega_f} \frac{1}{\rho_f} \nabla p \cdot \nabla q \, dx = \lambda \int_{\Omega_f} \frac{1}{\rho_f c^2} p \, q \, dx, \tag{14.2b}$$

for all $(v,q) \in H^1_{\Gamma_D}(\Omega_s)^d \times H^1(\Omega_f)$.

For the linearized strain tensor ϵ in the solid, we assume that the strain-stress relationship satisfies

$$\sigma(v) : \nabla v \geq C_1 \epsilon(v) : \epsilon(v)$$

for some constant $C_1 > 0$, such that Korn's second inequality implies that a_s is a coercive bilinear form.

Problem (14.1) can be written in operator notation. The aim is to find $\lambda \in \mathbb{C}$ and nonzero $(u,p) \in H^1_{\Gamma_D}(\Omega_s)^d \times H^1(\Omega_f)$ so that

$$\begin{aligned} \mathcal{K}_s u + \mathcal{C} p &= \lambda \mathcal{M}_s u \\ \mathcal{K}_f p &= \lambda(-\mathcal{C}' u + \mathcal{M}_f p), \end{aligned}$$

where the operators are defined corresponding to the variational formulation in (14.1).

Some elementary properties of the fluid-solid interaction eigenvalue problem can be given as follows:

LEMMA 14.1.

(i) The eigenvalue problem and its adjoint problem have a zero eigenvalue with corresponding one-dimensional eigenspaces (u_0, p_0) and $(0, p_0)$, where $p_0 \equiv 1$ and u_0 is determined in the proof provided in [216].

(ii) The function (u,p) is an eigensolution of the right eigenvalue problem (14.1) *corresponding to an eigenvalue $\lambda \neq 0$ if and only if $(\lambda u, p)$ is an eigensolution of the adjoint eigenvalue problem* (14.2) *corresponding to the same eigenvalue.*

(iii) Eigenfunctions (u_1, p_1) and (u_2, p_2) of (14.1) *corresponding to distinct eigenvalues $\lambda_1 \neq \lambda_2$ are orthogonal with respect to the inner product*

$$\langle (u,p), (v,q) \rangle := a_s(u,v) + b_f(p,q).$$

(iv) Assume that (u_1, p_1) is an eigensolution of (14.1) *and $(\hat{u}_2, \hat{p}_2)$ an eigensolution of the adjoint problem* (14.2) *corresponding to the eigenvalues λ_1 and λ_2, respectively.*

ETNA
Kent State University and
Johann Radon Institute (RICAM)

If $\lambda_1 \neq \lambda_2$, then it holds that

$$a_s(\hat{u}_2, u_1) + c(\hat{u}_2, p_1) + a_f(\hat{p}_2, p_1) = b_s(\hat{u}_2, u_1) - c(u_1, \hat{p}_2) + b_f(p_1, \hat{p}_2) = 0.$$

If $\lambda_1 = \lambda_2$ and $(\hat{u}_2, \hat{p}_2) = (\lambda_1 u_1, p_1)$, then it holds that

$$\begin{aligned} a_s(\hat{u}_2, u_1) + c(\hat{u}_2, p_1) + a_f(\hat{p}_2, p_1) &\geq 0 \qquad \textit{and} \\ b_s(\hat{u}_2, u_1) - c(u_1, \hat{p}_2) + b_f(p_1, \hat{p}_2) &> 0. \end{aligned}$$

(v) The eigenvalue problem (14.1) has an infinite countable number of eigenvalues, which are all real and non-negative and which converge to infinity.

Lemma 14.1 states the relationship between the eigenfunctions of the problem (14.1) and the adjoint problem (14.2). The adjoint eigenfunction $(\lambda u, p)$ can be used as a test function in equation (14.1), so that we obtain

$$\lambda a_s(u, u) + \lambda c(u, p) + a_f(p, p) = \lambda^2 b_s(u, u) - \lambda c(u, p) + \lambda b_f(p, p)$$

for any eigensolution $(\lambda, (u, p))$, i.e., it is a zero of the function

$$g(\lambda, (u, p)) := \lambda^2 b_s(u, u) + \lambda \left(b_f(p, p) - a_s(u, u) - 2c(u, p) \right) - a_f(p, p).$$

If $b_s(u, u) > 0$, then this equation is quadratic in λ, and one can easily show that the maximal one is an eigenvalue λ of (14.1) corresponding to (u, p). This suggests to introduce an eigenvalue approximation for some general nonzero $(u, p) \in H^1_{\Gamma_D}(\Omega_s)^d \times H^1(\Omega_f)$, and we define the nonlinear Rayleigh functional as the maximal root of $g(\cdot, (u, p))$.

DEFINITION 14.2. *The functional $r : H^1_{\Gamma_D}(\Omega_s)^d \times H^1(\Omega_f) \backslash \{0\} \to \mathbb{R}$, where any nonzero $(u, p) \in H^1_{\Gamma_D}(\Omega_s)^d \times H^1(\Omega_f)$ is mapped to the maximal root of $g(\cdot, (u, p))$ is called the nonlinear Rayleigh functional, i.e.,*

$$r(u, p) = \begin{cases} \Delta + \sqrt{\Delta^2 + \frac{a_f(p,p)}{b_s(u,u)}} & \text{if } b_s(u, u) \neq 0, \\ \frac{a_f(p,p)}{b_f(p,p)} & \text{if } b_s(u, u) = 0, \end{cases}$$

where

$$\Delta = \frac{1}{2} \frac{-b_f(p, p) + a_s(u, u) + 2c(u, p)}{b_s(u, u)}.$$

Although fluid-solid eigenvalue problems are not self-adjoint, one obtains variational characterizations using the nonlinear Rayleigh functional. These results generalize variational principles known from the linear self-adjoint case.

THEOREM 14.3 ([216]). *Let $\lambda_1 \leq \lambda_2 \leq \cdots$ be the eigenvalues of (14.1) in ascending order and $(u_1, p_1), (u_2, p_2), \ldots$ the corresponding eigenfunctions. Then it holds that*

(i) (Rayleigh's principle)

$$\lambda_k = \min\{r(u, p) : \ a_s(u, u_j) + b_f(p, p_j) = 0, \ j = 1, \ldots, k-1\};$$

(ii) (minmax characterization)

$$\lambda_k = \min_{\substack{S_k \subset H^1_{\Gamma_D}(\Omega_s)^d \times H^1(\Omega_f) \\ \dim S_k = k}} \max_{0 \neq (u,p) \in S_k} r(u, p);$$

(iii) (maxmin characterization)

$$\lambda_k = \max_{\substack{S_{k-1} \subset H^1_{\Gamma_D}(\Omega_s)^d \times H^1(\Omega_f) \\ \dim S_{k-1} = k-1}} \min_{0 \neq (u,p) \in S^\perp_{k-1}} r(u, p),$$

where

$$S^\perp := \{(u, p) \in H^1_{\Gamma_D}(\Omega_s)^d \times H^1(\Omega_f) \ : \ a_s(u, v) + b_f(p, q) = 0 \textit{ for } (v, q) \in S\}.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

14.2. Discretizing fluid-solid problems. Discretizing (14.1) by finite elements respecting the domains of the fluid and of the structure yields the non-symmetric matrix eigenvalue problem [15, 99]

$$Kx := \begin{bmatrix} K_s & C \\ 0 & K_f \end{bmatrix} \begin{bmatrix} x_s \\ x_f \end{bmatrix} = \lambda \begin{bmatrix} M_s & 0 \\ -C^T & M_f \end{bmatrix} \begin{bmatrix} x_s \\ x_f \end{bmatrix} =: \lambda M x. \tag{14.3}$$

Here x_s is the structure displacement vector with s degrees of freedom, x_f is the fluid pressure vector with f degrees of freedom, and $\lambda = \omega^2$ denotes the eigenvalue. $K_s \in \mathbb{R}^{s\times s}$ and $K_f \in \mathbb{R}^{f\times f}$ are the symmetric stiffness matrices, with $K_s > 0$, $K_f \geq 0$, and $M_s \in \mathbb{R}^{s\times s}$ and $M_f \in \mathbb{R}^{f\times f}$ are the positive definite mass matrices of the structure and the fluid, respectively, and $C \in \mathbb{R}^{s\times f}$ describes the coupling of structure and fluid.

Several authors [11, 29, 30, 31, 57] emphasize that the non-symmetric structure of the problem (14.3) makes it inconvenient from the numerical point of view. However, since the discrete eigenproblem (14.3) inherits the variational characterization of its eigenvalues from (14.1), all methods in Chapters 3 and 4 apply. The paper [214] describes the behavior of the Nonlinear Arnoldi method and the Jacobi-Davidson method for a fluid-solid problem of dimension 143,082.

A common approach for solving problem (14.3) [14, 122] (for example in the automotive industry), which works well for weakly coupled systems is as follows: One first determines the eigenpairs of the symmetric and positive definite and semidefinite eigenvalue problems

$$K_s x_s = \omega_s^2 M_s x_s \qquad \text{and} \qquad K_f x_f = \omega_f^2 M_f x_f \tag{14.4}$$

by the Lanczos method or Automated Multi-Level Substructuring and then projects the problem (14.3) to $\operatorname{diag}\{X_s, X_f\}$, where the columns of X_s and X_f are the eigenmodes of the problem (14.4) not exceeding a given cut-off level. The projected problem

$$\begin{bmatrix} X_s^T K_s X_s & X_s^T C X_f \\ 0 & X_f^T K_f X_f \end{bmatrix} \begin{bmatrix} y_s \\ y_f \end{bmatrix} = \lambda \begin{bmatrix} X_s^T M_s X_s & 0 \\ -X_f^T C^T X_s & X_f^T M_f X_f \end{bmatrix} \begin{bmatrix} y_s \\ y_f \end{bmatrix}$$

has the same structure as the original problem but is of much smaller dimension. An example in [247], however, demonstrates that this approach is not appropriate for strongly coupled problems.

14.3. AMLS for fluid-solid interaction problems. Our AMLS variant for the fluid-solid interaction problem (14.3) is based on the symmetric eigenproblem

$$T(\lambda)x := \left(\begin{bmatrix} 0 & C & K_s & 0 \\ C^T & 0 & 0 & K_f \\ K_s & 0 & 0 & 0 \\ 0 & K_f & 0 & 0 \end{bmatrix} - \lambda \begin{bmatrix} M_s & 0 & 0 & 0 \\ 0 & M_f & 0 & 0 \\ 0 & 0 & K_s & 0 \\ 0 & 0 & 0 & K_f \end{bmatrix} \right) x = 0, \tag{14.5}$$

whose eigenpairs resemble those from (14.3). If $(\lambda^2, (x_s^T, x_f^T)^T)$ solves problem (14.3), then

$$(\pm\lambda, \begin{bmatrix} \lambda^2 x_s^T & \pm\lambda x_f^T & \pm\lambda x_s^T & x_f^T \end{bmatrix}^T)$$

are solutions of (14.5) unless $\lambda = 0$.

We introduce the following numbering of the eigenvalues of (14.5) in the interval $J := (-\underline{\omega}, \underline{\omega})$ of interest:

$$-\underline{\omega} < \cdots \leq \lambda_{-3} \leq \lambda_{-2} < 0 < \lambda_2 \leq \lambda_3 \leq \cdots < \underline{\omega}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

and $\lambda_{-1} = \lambda_1 = 0$. The matrix $T(0)$ has rank $n-2$ and a symmetric spectrum, i.e., the two zero eigenvalues have the enumeration $n/2$ and $n/2+1$, and in terms of the theory of nonlinear eigenvalue problems, the following variational characterization holds

$$\lambda_{+j} = \min_{\dim V = n/2+j} \sup_{x \in V \cap D} p(x) \qquad \text{and} \qquad \lambda_{-j} = \min_{\dim V = n/2-j+1} \sup_{x \in V \cap D} p(x).$$

The CMS a priori bounds in Theorem 14.4 can be derived by comparing the nonlinear Rayleigh functional of problem (14.3) with the Rayleigh quotient r of the truncated linear eigenproblem

$$\begin{bmatrix} \Lambda_2 & 0 \\ 0 & \hat{K}_{11} \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = \lambda \begin{bmatrix} I & \check{M}_{21} \\ \check{M}_{21}^T & \hat{M}_{11} \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} \tag{14.6}$$

on appropriate subspaces.

THEOREM 14.4 ([215]). *Denote by $0 = \lambda_{+1} < \lambda_{+2} \le \lambda_{+3} \le \cdots < \underline{\omega}$ one zero eigenvalue and the positive eigenvalues of the nonlinear eigenvalue problem $T(\lambda)x = 0$ and by $0 = \tilde{\lambda}_{+1} < \tilde{\lambda}_{+2} \le \tilde{\lambda}_{+3} \le \cdots < \underline{\omega}$ the corresponding eigenvalues of the truncated linear eigenproblem (14.6).*

Then for every $j \ge 2$ such that $\lambda_{+j}, \tilde{\lambda}_{+j} \in J := (-\underline{\omega}, \underline{\omega})$, it holds that

$$\lambda_{+j} - \frac{\lambda_{+j}^2}{\underline{\omega} + \lambda_{+j}} \le \tilde{\lambda}_{+j} \le \lambda_{+j} + \frac{\lambda_{+j}^2}{\underline{\omega} - \lambda_{+j}},$$

i.e.,

$$-\frac{\lambda_{+j}}{\underline{\omega} + \lambda_{+j}} \le \frac{\tilde{\lambda}_{+j} - \lambda_{+j}}{\lambda_{+j}} \le \frac{\lambda_{+j}}{\underline{\omega} - \lambda_{+j}}. \tag{14.7}$$

The upper bound in (14.7) for the relative error has the same structure as the error bound given in Theorem 8.1 for CMS applied to a definite eigenvalue problem. In the definite case, the lower bound is 0 due to the fact that CMS is a projection method and the eigenvalues under consideration are at the lower end of the spectrum. CMS for the indefinite eigenproblem (14.3) discards eigenmodes corresponding to eigenvalues which exceed a cut-off frequency in modulus. Since we are approximating an interior eigenvalue λ_{+j} of the eigenvalue problem (14.3) by an interior eigenvalue $\tilde{\lambda}_{+j}$ of the linear eigenvalue problem corresponding to (14.5), which may have different numbers with respect to the natural enumeration (smallest eigenvalue first one, second smallest second one, etc.), it may happen that $\tilde{\lambda}_{+j} < \lambda_{+j}$. The paper [215] contains an example that this can actually happen.

AMLS on p partitioning levels is mathematically equivalent to p CMS steps, so that in the CMS step at level $k = p, \ldots, 1$, the eigenmodes at level k are truncated and the eigenmodes on all other levels are retained. We denote by $\lambda_{+j}^{(k)}$ the eigenvalue approximation if the lowest k partitioning levels have been handled, i.e., $\lambda_{+j}^{(0)}$ denotes the exact eigenvalues and $\lambda_{+j}^{(p)}$ the approximation when the reduction process has terminated. Then we apply the CMS bound in Theorem 14.4 recursively and obtain the following error bound for AMLS.

THEOREM 14.5 ([215]). *Consider the AMLS algorithm for fluid-solid interaction eigenproblems on p levels. Denote by $\lambda_{+j}^{(k)}$ the eigenvalues after the k lowest partitioning levels have been handled ($k = 0, \ldots, p$), and assume that the cut-off frequency satisfies $\underline{\omega} > p\lambda_{+j}^{(0)} \ge 0$. Then the eigenvalues can be bounded by*

$$\frac{\underline{\omega}\lambda_{+j}^{(p)}}{\underline{\omega} + p\lambda_{+j}^{(p)}} \le \lambda_{+j}^{(0)} \le \frac{\underline{\omega}\lambda_{+j}^{(p)}}{\underline{\omega} - p\lambda_{+j}^{(p)}}.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

14.4. Numerical example. To evaluate the modified AMLS algorithm for fluid-solid interaction problems, we consider a two-dimensional model with 120,473 degrees of freedom. The solid is steel, and its discretization has 67,616 degrees of freedom. As fluid we consider water, whose discretization leads to 52,857 degrees of freedom. To investigate the coupling effects, the underlying geometry was chosen with a rather large interface between fluid and solid; cf. [215].

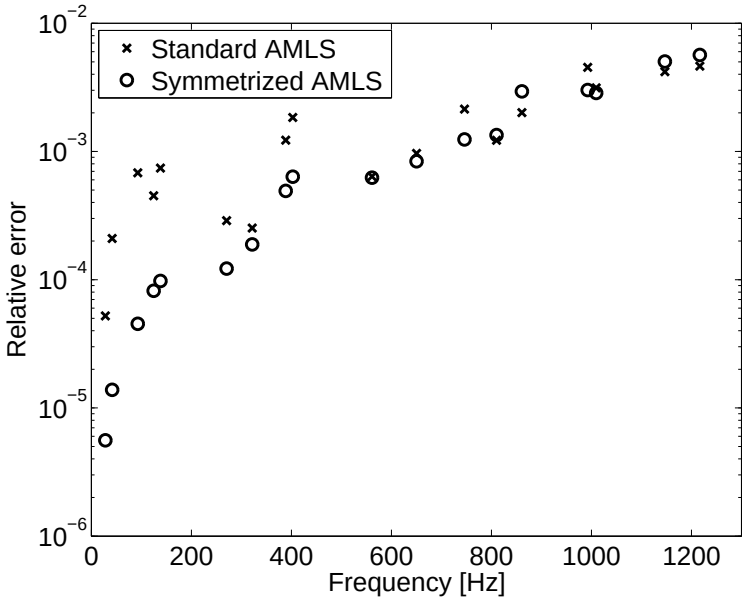


FIG. 14.1. *The relative error of adapted AMLS for fluid-solid interaction problems and the relative error of standard AMLS applied to fluid-solid interaction problems.*

We applied the AMLS variant described in Section 14.3 for the coupled fluid-solid problem and compared the eigenvalue approximations to those obtained from the standard AMLS. In both cases, the algorithm was performed on 10 sub-structuring levels and 751 structures using a cut-off frequency corresponding to 10,000 Hz on each partitioning level. The errors are displayed in Figure 14.1.

Eigenvalues with large accuracy improvements (e.g., $\lambda \approx 100$ Hz) turned out to belong to eigenforms with significant influence of the coupling. Eigenforms corresponding to larger eigenfrequencies were less influenced by the coupling, and in some cases, the eigenvalue approximations are slightly worse compared with the AMLS variant neglecting the coupling effects in the reduction process. In all cases, the eigenvalue approximations were of larger magnitude than the exact eigenvalues.

15. Conclusions. Variational characterization of eigenpairs is a powerful tool for linear eigenvalue problems to derive localization and monotonicity results, error bounds, comparison, and interlacing results for eigenvalues. In this paper we collect several generalizations for eigenvalue problems which are nonlinear with respect to the eigenparameter: we present a method for detecting hyperbolic matrix polynomials that allow for a definite linearization and can be solved by standard methods, we generalize Sylvester's law of inertia to locate eigenvalues and give bounds for low-rank modifications of eigenvalue problems. We analyze the electronic behavior of quantum dots, of viscoelastic operators, and the vibration of fluid-

ETNA
Kent State University and
Johann Radon Institute (RICAM)

solid structures. We present an error bound for automated multi-level substructuring for huge eigenvalue problems and present methods for the regularization of total least-squares problems. Numerical examples accompany the various methods.

REFERENCES

[1] J. Š. ABRAMOV, *On the theory of nonlinear eigenvalue problems*, Dokl. Akad. Nauk SSSR, 212 (1973), pp. 11–14.
[2] ———, *Pencils of waveguide type and related extremal problems*, J. Math. Sci., 64 (1993), pp. 1278–1288.
[3] R. ADCOCK, *Note on the method of least squares*, The Analyst, 4 (1877), pp. 183–184.
[4] S. ADHIKARI, *Structural Dynamic Analysis with Generalized Damping Models: Analysis*, Wiley, Hoboken, 2014.
[5] ———, *Structural Dynamic Analysis with Generalized Damping Models: Identification*, Wiley, Hoboken, 2014.
[6] S. ADHIKARI AND B. PASCUAL, *Eigenvalues of linear viscoelastic systems*, J. Sound Vib., 325 (2009), pp. 1000–1011.
[7] ———, *Iterative methods for eigenvalues of viscoelastic systems*, J. Vib. Acoust., 132 (2011), Art. 021002, 7 pages.
[8] S. ADHIKARI AND N. WAGNER, *Analysis of asymmetric nonviscously damped linear dynamic systems*, J. Appl. Mech., 70 (2003), pp. 885–893.
[9] ———, *Direct time–domain integration method for exponentially damped linear systems*, Comput. & Structures, 82 (2004), pp. 2453–2461.
[10] H. AKAIKE, *A new look at the statistical model identification*, IEEE Trans. Automat. Control, 19 (1974), pp. 716–723.
[11] A. ALONSO, A. DELLO RUSSO, C. PADRA, AND R. RODRÍGUEZ, *A posteriori error estimates and a local refinement strategy for a finite element method to solve structural-acoustic vibration problems*, Adv. Comput. Math., 15 (2001), pp. 25–59.
[12] A.L. ANDREW, K. E. CHU, AND P. LANCASTER, *On the numerical solution of nonlinear eigenvalue problems*, Computing, 55 (1995), pp. 91–111.
[13] P. M. ANSELONE AND L. B. RALL, *The solution of characteristic value-vector problems by Newton's method*, Numer. Math., 11 (1968), pp. 38–45.
[14] ANSYS INC., *Theory Reference for ANSYS and ANSYS Workbench. Release 11.0 (2007)*.
[15] I. BABUŠKA AND J. OSBORN, *Eigenvalue problems*, in Handbook of Numerical Analysis, P. Ciarlet and J. Lions, eds., vol. II, North Holland, Amsterdam, 1991, pp. 641–923.
[16] Z. BAI, J. DEMMEL, J. DONGARRA, A. RUHE, AND H. VAN DER VORST, *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, SIAM, Philadelphia, 2000.
[17] Z.-J. BAI AND Y.-F. SU, *Second-order Krylov subspace and Arnoldi procedure*, J. Shanghai Univ., 8 (2004), pp. 378–390.
[18] Z. BAI AND Y. SU, *SOAR: a second-order Arnoldi method for the solution of the quadratic eigenvalue problem*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 640–659.
[19] E. M. BARSTON, *A minimax principle for nonoverdamped systems*, Internat. J. Engrg. Sci., 12 (1974), pp. 413–421.
[20] G. BASTARD, *Wave Mechanics Applied to Semiconductor Heterostructures*, EDP Sciences, Les Ulis, 1988.
[21] K.-J. BATHE, *Finite Element Procedures*, Prentice Hall, Englewood Cliffs, 1995.
[22] A. BECK AND A. BEN-TAL, *On the solution of the Tikhonov regularization of the total least squares problem*, SIAM J. Optim., 17 (2006), pp. 98–118.
[23] A. BECK, A. BEN-TAL, AND M. TEBOULLE, *Finding a global optimal solution for a quadratically constrained fractional quadratic problem with applications to the regularized total least squares*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 425–445.
[24] A. BECK AND M. TEBOULLE, *A convex optimization approach for minimizing the ratio of indefinite quadratic functions over an ellipsoid*, Math. Program., 118 (2009), pp. 13–35.
[25] C. BEKAS AND Y. SAAD, *Computation of smallest eigenvalues using spectral Schur complements*, SIAM J. Sci. Comput., 27 (2005), pp. 458–481.
[26] T. BELYTSCHKO, *Fluid-structure interaction*, Comput. & Structures, 12 (1980), pp. 459–469.
[27] J. BENNIGHOF AND M. KAPLAN, *Frequency sweep analysis using multi-level substructuring, global modes and in iteratio*, in Proceeding of the 39th Annual AIAA Structural Dynamics and Material Conference, AIAA Meeting Papers, AIAA, Reston, 1988.
[28] J. BENNIGHOF AND R. LEHOUCQ, *An automated multilevel substructuring method for eigenspace computation in linear elastodynamics*, SIAM J. Sci. Comput., 25 (2004), pp. 2084–2106.
[29] A. BERMÚDEZ, L. HERVELLA-NIETO, AND R. RODRIGUEZ, *Finite element computation of three dimensional elastoacoustic vibrations*, J. Sound Vibr., 219 (1999), pp. 279–306.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[30] A. BERMÚDEZ AND R. RODRÍGUEZ, *Analysis of a finite element method for pressure/potential formulation of elastoacoustic spectral problems*, Math. Comp., 71 (2002), pp. 537–552.
[31] A. BERMÚDEZ, R. RODRÍGUEZ, AND D. SANTAMARINA, *A finite element solution of an added mass formulation for coupled fluid-solid vibrations*, Numer. Math., 87 (2000), pp. 201–227.
[32] T. BETCKE AND D. KRESSNER, *Perturbation, computation and refinement of invariant pairs for matrix polynomials*, Linear Algebra Appl., 435 (2011), pp. 514–536.
[33] M. M. BETCKE AND H. VOSS, *Stationary Schrödinger equations governing electronic states of quantum dots in the presence of spin-orbit splitting*, Appl. Math., 52 (2007), pp. 267–284.
[34] ———, *Numerical simulation of electronic properties of coupled quantum dots on wetting layers*, Nanotechnology, 19 (2008), Art. 165204, 10 pages.
[35] ———, *Restarting projection methods for rational eigenproblems arising in fluid-solid vibrations*, Math. Model. Anal., 13 (2008), pp. 171–182.
[36] ———, *Analysis and efficient solution of stationary Schrödinger equation describing the electronic states of quantum dots and rings in magnetic field*, Commun. Comput. Phys., 11 (2012), pp. 1591–1617.
[37] ———, *Restarting iterative projection methods for Hermitian nonlinear eigenvalue problems with minmax property*, Numer. Math., 135 (2017), pp. 397–430.
[38] T. BETCKE AND H. VOSS, *A Jacobi–Davidson–type projection method for nonlinear eigenvalue problems*, Future Gen. Comput. Syst., 20 (2004), pp. 363–372.
[39] W.-J. BEYN, C. EFFENBERGER, AND D. KRESSNER, *Continuation of eigenvalues and invariant pairs for parameterized nonlinear eigenvalue problems*, Numer. Math., 119 (2011), pp. 489–516.
[40] P. BINDING, D. ESCHWÉ, AND H. LANGER, *Variational principles for real eigenvalues of self-adjoint operator pencils*, Integral Equations Operator Theory, 38 (2000), pp. 190–206.
[41] M. A. BIOT, *Linear thermodynamics and the mechanics of solids*, in Proceedings of the Third U.S. National Congress of Applied Mechanics, Brown University, Providence, R.I., June 11-14, 1958, American Society of Mechanical Engineers, New York, 1958, pp. 1–18.
[42] Å. BJÖRCK, *A bidiagonalization algorithm for solving large and sparse ill-posed systems of linear equations*, BIT, 28 (1988), pp. 659–670.
[43] D. CALVETTI, G.H. GOLUB, AND L. REICHEL, *Estimation of the L-curve via Lanczos bidiagonalization*, BIT, 39 (1999), pp. 603–619.
[44] D. CALVETTI AND L. REICHEL, *Tikhonov regularization of large linear problems*, BIT, 43 (2003), pp. 263–283.
[45] K. CHEN, *Matrix Preconditioning Techniques and Applications*. Cambridge University Press, Cambridge, 2005.
[46] S. CHUANG, *Physics of Optoelectronic Devices*, Wiley, New York, 1995.
[47] J. CHUNG, J.G. NAGY, AND D. P. O'LEARY, *A weighted-GCV method for Lanczos-hybrid regularization*, Electron. Trans. Numer. Anal., 28 (2008), pp. 149–167.
http://etna.ricam.oeaw.ac.at/vol.28.2007-2008/pp149-167.dir/pp149-167.pdf
[48] COMSOL, *FEMLAB Version 3.3*, Burlington, USA, 2006.
[49] C. CONCA, J. PLANCHARD, AND M. VANNINATHAN, *Fluid and Periodic Structures*, Masson, Paris, 1995.
[50] R. COURANT, *Über die Eigenwerte bei den Differentialgleichungen der mathematischen Physik*, Math. Z., 7 (1920), pp. 1–57.
[51] R. CRAIGH. AND M. BAMPTON, *Coupling of substructures for dynamic analysis*, AIAA J., 6 (1968), pp. 1313–1319.
[52] E. M. DAYA AND M. POTIER-FERRY, *A numerical method for nonlinear eigenvalue problems—application to vibrations of viscoelastic structures*, Comput. & Structures, 79 (2001), pp. 533–541.
[53] J. W. DEMMEL, *Applied Numerical Linear Algebra*, SIAM, Philadelphia, 1997.
[54] J.-F. DEÜ, W. LARBI, AND R. OHAYON, *Variational formulation of interior structural-acoustic vibration problem*, in Computational Aspects of Structural Acoustics and Vibrations, G. Sandberg, R. Ohayon, eds., Springer, Vienna, 2009, pp. 1-21.
[55] Z. DING, L. LI, J. KONG, AND L. QIN, *A modal projection-based reduction method for transient dynamic responses of viscoelastic systems with multiple damping models*, Comput. & Structures, 194 (2018), pp. 60–73.
[56] R. J. DUFFIN, *A minimax theory for overdamped networks*, J. Rational Mech. Anal., 4 (1955), pp. 221–233.
[57] R. G. DURÁN, L. HERVELLA-NIETO, E. LIBERMAN, R. RODRÍGUEZ, AND J. SOLOMIN, *Finite element analysis of the vibration problem of a plate coupled with a fluid*, Numer. Math., 86 (2000), pp. 591–616.
[58] C. EFFENBERGER, *Robust Solution Methods for Nonlinear Eigenvalue Problems*, PhD. Thesis, EPFL Lausanne, Lausanne, 2013.
[59] ———, *Robust successive computation of eigenpairs for nonlinear eigenvalue problems*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 1231–1256.
[60] L. ELDÉN, *A weighted pseudoinverse, generalized singular values, and constrained least squares problems*, BIT, 22 (1986), pp. 487–502.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[61] L. E. ÈL'SGOL'TS AND S. B. NORKIN, *Introduction to the Theory and Application of Differential Equations with Deviating Arguments*, Academic Press, New York, 1973.
[62] K. ELSSEL AND H. VOSS, *An a priori bounds for automated multilevel substructuring*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 386–397.
[63] ———, *Reducing huge gyroscopic eigenproblems by automated multi-level substructuring*, Arch. Appl. Mech., 76 (2006), pp. 171–179.
[64] ———, *Reducing sparse nonlinear eigenproblems by automated multi-level substructuring*, Adv. Engrg. Softw., 39 (2008), pp. 828–838.
[65] H. W. ENGL, M. HANKE, AND A. NEUBAUER, *Regularization of Inverse Problems*, Kluwer, Dortrecht, 1996.
[66] D. ESCHWÉ AND M. LANGER, *Variational principles for eigenvalues of self-adjoint operator functions*, Integral Equations Operator Theory, 49 (2004), pp. 287–321.
[67] H. FASSBENDER, *Structured eigenvalue problems—structure-preserving algorithms, structured error analysis*, in Handbook of Linear Algebra, 2nd ed., Chpt. 63, L. Hogben, ed., CRC Press, Boca Raton, 2014.
[68] E. FISCHER, *Über quadratische Formen mit reellen Koeffizienten*, Monatsh. Math. Phys., 16 (1905), pp. 234–249.
[69] D. R. FOKKEMA, G. L. G. SLEIJPEN, AND H. VAN DER VORST, *Jacobi-Davidson style QR and QZ algorithms for the reduction of matrix pencils*, SIAM J. Sci. Comput., 20 (1998), pp. 94–125.
[70] R. W. FREUND, *Padé-type model reduction of second-order and higher-order linear dynamical systems*, in Dimension Reduction of Large-Scale Systems, P. Benner, V. Mehrmann, and D. C. Sorensen, eds., vol. 45 of Lect. Notes Comput. Sci. Eng., Springer, Berlin, 2005, pp. 191–223.
[71] W. GANDER, *Least squares with a quadratic constraint*, Numer. Math., 36 (1980/81), pp. 291–307.
[72] W. GANDER, G. GOLUB, AND U. VON MATT, *A constrained eigenvalue problem*, Linear Algebra Appl., 114/115 (1989), pp. 815–839.
[73] W. GAO, X. S. LI, C. YANG, AND Z. BAI, *An implementation and evaluation of the AMLS method for sparse eigenvalue problems*, ACM Trans. Math. Software, 34 (2008), Art. 20, 28 pages.
[74] C. K. GARRETT, Z. BAI, AND R.-C. LI, *A nonlinear QR algorithm for banded nonlinear eigenvalue problems*, ACM Trans. Math. Software, 43 (2016), Art. 4, 19 pages.
[75] C. K. GARRETT AND R.-C. LI, *Unstructurally banded nonlinear eigenvalue software*. http://www.csm.ornl.gov/newsite/software.html
[76] G. H. GOLUB, *Some modified matrix eigenvalue problems*, SIAM Rev., 15 (1973), pp. 318–334.
[77] G. H. GOLUB, P. HANSEN, AND D. P. O'LEARY, *Tikhonov regularization and total least squares*, SIAM J. Matrix Anal. Appl., 21 (1999), pp. 185–194.
[78] G. H. GOLUB, M. HEATH, AND G. WAHBA, *Generalized cross-validation as a method for choosing a good ridge parameter*, Technometrics, 21 (1979), pp. 215–223.
[79] G. GOLUB AND C. VAN LOAN, *An analysis of the total least squares problem*, SIAM J. Numer. Anal., 17 (1980), pp. 883–893.
[80] ———, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, 1996.
[81] I. GOHBERG, P. LANCASTER, AND L. RODMAN, *Matrix Polynomials*, SIAM, Philadelphia, 2009.
[82] D. GOLLA AND P. HUGHES, *Dynamics of viscoelastic structures—a time domain finite element formulation*, Trans. ASME J. Appl. Mech., 52 (1985), pp. 897–906.
[83] R. GRINIV AND T. MEL'NIK, *On the singular Rayleigh functional*, Math. Notes, 60 (1996), pp. 97–100.
[84] C. W. GROETSCH, *Inverse Problems in the Mathematical Sciences*, Vieweg, Braunschweig, 1993.
[85] C.-H. GUO, N. J. HIGHAM, AND F. TISSEUR, *Detecting and solving hyperbolic quadratic eigenvalue problems*, SIAM J. Matrix Anal. Appl., 30 (2008/09), pp. 1593–1613.
[86] ———, *An improved arc algorithm for detecting definite Hermitian pairs*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 1131–1151.
[87] C.-H. GUO AND P. LANCASTER, *Algorithms for hyperbolic quadratic eigenvalue problems*, Math. Comp., 74 (2005), pp. 1777–1791.
[88] H. GUO AND R. A. RENAUT, *A regularized total least squares algorithm*, in Total Least Squares and Errors-in-Variables Modeling (Leuven, 2001), S. Van Huffel and P. Lemmerling, eds., Kluwer Acad. Publ., Dordrecht, 2002, pp. 57–66.
[89] S. GÜTTEL AND F. TISSEUR, *The nonlinear eigenvalue problem*, Acta Numer., 26 (2017), pp. 1–94.
[90] K. P. HADELER, *Mehrparametrige und nichtlineare Eigenwertaufgaben*, Arch. Rational Mech. Anal., 27 (1967), pp. 306–328.
[91] ———, *Variationsprinzipien bei nichtlinearen Eigenwertaufgaben*, Arch. Rational Mech. Anal., 30 (1968), pp. 297–307.
[92] ———, *Nonlinear eigenvalue problems*, in Numerische Behandlung von Differentialgleichungen, R. Ansorge, L. Collatz, G. Hämmerlin, and W. Törnig, eds., Internat. Schriftenreihe Numer. Math., Band 27, Birkhäuser, Basel, 1975, pp. 111–129.
[93] J. HALE, *Functional Differential Equations*, Springer, New York, 1971.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[94] ———, *Ordinary Differential Equations*, Dover, Mineola, 2009.
[95] M. HANKE, *On Lanczos based methods for the regularization of discrete ill-posed problems*, BIT, 41 (2001), pp. 1008–1018.
[96] P. C. HANSEN, *Analysis of discrete ill-posed problems by means of the L-curve*, SIAM Rev., 34 (1992), pp. 561–580
[97] ———, *Rank-Deficient and Discrete Ill-Posed Problems*, SIAM, Philadelphia, 1998.
[98] ———, *Regularization Tools version 4.0 for Matlab 7.3*, Numer. Algorithms, 46 (2007), pp. 189–194.
[99] C. HARRIS AND A. PIERSOL (EDS.), *Harris' Shock and Vibration Handbook*, McGraw Hill, New York, 2002.
[100] M. HASANOV, *An approximation method in the variational theory of the spectrum of operator pencils*, Acta Appl. Math., 71 (2002), pp. 117–126.
[101] ———, *The spectra of two-parameter quadratic operator pencils*, Math. Comput. Modelling, 54 (2011), pp. 742–755.
[102] N. J. HIGHAM, D. S. MACKEY, N. MACKEY, AND F. TISSEUR, *Symmetric linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl., 29 (2006/07), pp. 143–159.
[103] N. J. HIGHAM, D. S. MACKEY, AND F. TISSEUR, *Definite matrix polynomials and their linearization by definite pencils*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 478–502.
[104] N. HIGHAM, F. TISSEUR, AND P. VAN DOOREN, *Detecting a definite Hermitian pair and a hyperbolic or elliptic quadratic eigenvalue problem, and associated nearness problems*, Linear Algebra Appl., 351/352 (2002), pp. 455–474.
[105] M. HOCHSTENBACH AND Y. NOTAY, *The Jacobi–Davidson method*, GAMM-Mitt., 29 (2006), pp. 368–382.
[106] L. HOGBEN, *Handbook of Linear Algebra*, 2nd ed., CRC Press, Boca Raton, 2014.
[107] U. B. HOLZ, G. H. GOLUB, AND K. H. LAW, *A subspace approximation method for the quadratic eigenvalue problem*, SIAM J. Matrix Anal. Appl., 26 (2004/05), pp. 498–521.
[108] R. HORN AND C. JOHNSON, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.
[109] X. HUANG, Z. BAI, AND Y. SU, *Nonlinear rank-one modification of the symmetric eigenvalue problem*, J. Comput. Math., 28 (2010), pp. 218–234.
[110] W. C. HURTY, *Vibrations of structural systems by component-mode synthesis*, J. Engrg. Mech. Div., 86 (1960), pp. 51–69.
[111] T.-M. HWANG, W.-W. LIN, W.-C. WANG, AND AND W. WANG, *Numerical simulation of three dimensional quantum dot*, J. Comput. Phys., 196 (2004), pp. 208–232.
[112] N. K. JAIN, K. SINGHAL, AND K. HUSEYIN, *On the roots of functional lambda matrices*, Comput. Methods Appl. Mech. Engrg., 40 (1983), pp. 277–292.
[113] E. JARLEBRING AND H. VOSS, *Rational Krylov for nonlinear eigenproblems, an iterative projection method*, Appl. Math., 50 (2005), pp. 543–554.
[114] D. JONES, *Handbook of Viscoelastic Vibration Damping*. Wiley, Chichester 2001.
[115] E. KANE, *Band structure of indium antimonide*, J. Phys. Chem. Solids, 1 (1957), pp. 249–261.
[116] M. KAPLAN, *Implementation of Automated Multi-Level Substructuring for Frequency Response Analysis*, PhD. Thesis, University of Texas, Austin, 2001.
[117] R. KOLLÁR, *Homotopy methods for nonlinear eigenvalue pencilswith applications*, SIAM J. Math. Anal., 43 (2011), pp. 612–633.
[118] A. KOSTIĆ AND H. VOSS, *On Sylvester's law of inertia for nonlinear eigenvalue problems*, Electron. Trans. Numer. Anal., 40 (2013), pp. 82–93.
http://etna.ricam.oeaw.ac.at/vol.40.2013/pp82-93.dir/pp82-93.pdf
[119] ———, *Real eigenvalues of viscoelastic oscillators involving several damping models*, Sarajevo J. Math., 14 (2018), pp. 35–43.
[120] D. KRESSNER, *A block Newton method for nonlinear eigenvalue problems*, Numer. Math., 114 (2009), pp. 355–372.
[121] D. KRESSNER AND J. E. ROMAN, *Memory-efficient Arnoldi algorithms for linearizations of matrix polynomials in Chebyshev basis*, Numer. Linear Algebra Appl., 21 (2014), pp. 569–588.
[122] A. KROPP AND D. HEISERER, *Efficient broadband vibro-acoustic analysis of passenger car bodies using an FE-based component mode synthesis approach*, J. Comput. Acoustics, 11 (2003), pp. 139–157.
[123] V. N. KUBLANOVSKAYA, *On an approach to the solution of the generalized latent value problem for λ-matrices*, SIAM J. Numer. Anal., 7 (1970), pp. 532–537.
[124] ———, *Methods and algorithms for solving spectral problems for polynomial and rational matrices*, J. Math. Sci., 96 (1999), pp. 3085–3287.
[125] J. LAMPE, *Modellreduktion für sehr große, dünn besetzte Systeme zweiter Ordnung mit dem Arnoldi-verfahren*, Master thesis, Institute of Numerical Simulation, Hamburg University of Technology, Hamburg, 2005.
[126] ———, *Solving Regularized Total Least Squares Problems Based on Eigenproblems*, PhD. Thesis, Institute of Numerical Simulation, Hamburg University of Technology, Hamburg, 2010.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[127] J. LAMPE, L. REICHEL, AND H. VOSS, *Large-scale Tikhonov regularization via reduction by orthogonal projection*, Linear Algebra Appl., 436 (2012), pp. 2845–2865.
[128] J. LAMPE, M. ROJAS, D. C. SORENSEN, AND H. VOSS, *Accelerating the LSTRS algorithm*, SIAM J. Sci. Comput., 33 (2011), pp. 175–194.
[129] J. LAMPE AND H. VOSS, *On a quadratic eigenproblem occurring in regularized total least squares*, Comput. Statist. Data Anal., 52 (2008), pp. 1090–1102.
[130] ———, *Global convergence of RTLSQEP: a solver of regularized total least squares problems via quadratic eigenproblems*, Math. Model. Anal., 13 (2008), pp. 55–66.
[131] ———, *A fast algorithm for solving regularized total least squares problems*, Electron. Trans. Numer. Anal., 31 (2008), pp. 12–24.
http://etna.ricam.oeaw.ac.at/vol.31.2008/pp12-24.dir/pp12-24.pdf
[132] ———, *Efficient determination of the hyperparameter in regularized total least squares problems*, Appl. Numer. Math., 62 (2012), pp. 1229–1241.
[133] ———, *Solving regularized total least squares problems based on eigenproblems*, Taiwanese J. Math., 14 (2010), pp. 885–909.
[134] ———, *Large-scale Tikhonov regularization of total least squares*, J. Comput. Appl. Math., 238 (2013), pp. 95–108.
[135] ———, *Large-scale dual regularized total least squares*, Electron. Trans. Numer. Anal., 42 (2014), pp. 13–40.
http://etna.ricam.oeaw.ac.at/vol.42.2014/pp13-40.dir/pp13-40.pdf
[136] P. LANCASTER, *A generalised Rayleigh quotient iteration for lambda-matrices*, Arch. Rational Mech. Anal., 8 (1961), pp. 309–322.
[137] ———, *Lambda-Matrices and Vibrating Systems*, Dover Publications, Mineola, 2002.
[138] ———, *Inverse spectral problems for semisimple damped vibrating systems*, SIAM J. Matrix Anal. Appl., 29 (2006/07), pp. 279–301.
[139] H. LANGER, *Über stark gedämpfte Scharen im Hilbertraum*, J. Math. Mech., 17 (1967/1968), pp. 685–705.
[140] U. LANGER, *Untersuchungen zum Kummerschen Verfahren zur numerischen Behandlung nichtlinearer Eigenwertaufgaben*, Beiträge Numer. Math., 6 (1977), pp. 97–110.
[141] C. LAWSON AND R. HANSON, *Solving Least Square Problems*, Prentice-Hall, Englewood Cliffs, 1974.
[142] M. LÁZARO, *Nonviscous modes of nonproportionally damped viscoelastic systems*, J. Appl. Mech., 82 (2015), Art. 121011, 9 pages.
[143] M. LÁZARO AND J. L. PÉREZ-APARICIO, *Characterization of real eigenvalues in linear viscoelastic oscillators and the nonviscous set*, J. Appl. Mech., 81 (2013), Art. 021015, 14 pages.
[144] M. LÁZARO, J. L. PÉREZ-APARICIO, AND M. EPSTEIN, *Computation of eigenvalues in proportionally damped viscoelastic structures based on the fixed-point iteration*, Appl. Math. Comput., 219 (2012), pp. 3511–3529.
[145] R. B. LEHOUCQ, D. C. SORENSEN, AND C. YANG, *ARPACK Users' Guide*, SIAM, Philadelphia, 1998.
[146] G. LESIEUTRE, *Finite elements for dynamic modeling of uniaxial rods with frequency-dependent material properties*, Int. J. Sol. Struct., 29 (1992), pp. 1567–1579.
[147] L. LI AND Y. HU, *State-space methods for viscoelastic systems involving general damping models*, AIAA J., 54 (2016), pp. 3290–3295.
[148] L. LI, Y. HU, W. DENG, L. LÜ, AND Z. DING, *Dynamics of structural systems with various frequency-dependent damping models*, Front. Mech. Engng., 10 (2015) pp. 48–63.
[149] L. LI, Y. HU, AND X. WANG, *Improved approximate methods for calculating frequency response function matrix and response of MDOF systems with viscoelastic hereditary terms*, J. Sound Vibr., 332 (2013), pp. 3945–3956.
[150] L. LI, Y. HU, X. WANG, AND L. LÜ, *A hybrid expansion method for frequency response functions of non-proportionally damped systems*, Mech. Syst. Signal Proc., 42 (2014), pp. 31–41.
[151] Y. LI, O. VOSKOBOYNIKOV, C. P. LEE, AND S. M. SZE, *Energy and coordinate dependent effective mass and confined electron states in quantum dots*, Solid State Commun., 120 (2001), pp. 79–83.
[152] R.-C. LI AND Q. YE, *A Krylov subspace method for quadratic matrix polynomials with application to constrained least squares problems*, SIAM J. Matrix Anal. Appl., 25 (2003), pp. 405–428.
[153] X. LANG AND R.-C. LI, *The hyperbolic quadratic eigenvalue problem*, Forum Math., Sigma, 3 (2015), Art. E13, 93 pages.
[154] B.-S. LIAO, Z. BAI, L.-Q. LEE, AND K. KO, *Nonlinear Rayleigh-Ritz iterative method for solving large scale nonlinear eigenvalue problems*, Taiwanese J. Math., 14 (2010), pp. 869–883.
[155] S. LIAO, *Series solution of nonlinear eigenvalue problems by means of the homotopy analysis method*, Nonlinear Anal. Real World Appl., 10 (2009), pp. 2455–2470.
[156] J. LIESEN AND C. MEHL, *Matrix Polynomials*, in Handbook of Linear Algebra, 2nd ed., Chpt. 18, L. Hogben, ed., CRC Press, Boca Raton, 2014.
[157] D. LU, Y. SU, AND Z. BAI, *Stability analysis of the two-level orthogonal Arnoldi procedure*, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 195–214.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[158] S. LU, S. V. PEREVERZEV, AND U. TAUTENHAHN, *Dual regularized total least squares and multi-parameter regularization*, Comput. Methods Appl. Math., 8 (2008), pp. 253–262.
[159] ———, *Regularized total least squares: computational aspects and error bounds*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 918–941.
[160] J. M. LUTTINGER AND W. KOHN, *Motion of electrons and holes in perturbed periodic fields*, Phys. Rev., 97 (1955), pp. 869–883.
[161] D.S. MACKEY, *Structured Linearization of Matrix Polynomials*, PhD. Thesis, Manchester Institute for Mathematical Sciences, The University of Manchester, Manchester, 2006.
[162] D. S. MACKEY, N. MACKEY, C. MEHL, AND V. MEHRMANN, *Structured polynomial eigenvalue problems: good vibrations from good linearizations*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 1029–1051.
[163] ———, *Vector spaces of linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 971–1004.
[164] D. S. MACKEY, N. MACKEY, AND F. TISSEUR, *Structured tools for structured matrices*, Electron. J. Linear Algebra, 10 (2003), pp. 106–145.
[165] F. G. MAKSUDOV AND M. G. GASANOV, *On the variational theory of the spectrum of operator pencils*, Dokl. Akad. Nauk., 325 (1992), pp. 915–918.
[166] M. MARKIEWICZ AND H. VOSS, *A local restart procedure for iterative projection methods for nonlinear symmetric eigenproblems*, in Proc. of Algoritmy 2005, A. Handlovicova, Z. Kriva, K. Mikula, and D. Sevcovic, eds, Slovak University of Technology, Bratislava, 2005, pp. 212–221.
[167] I. MARKOVSKY AND S. VAN HUFFEL, *Overview of total least squares methods*, Signal Process., 87 (2007), pp. 2283–2302.
[168] A. MARKUS, *Introduction to the Spectral Theory of Polynomial Operator Pencils*, AMS, Providence, 1988.
[169] Y. MASUMOTO AND T. TAKAGAHARA (EDS.), *Semiconductor Quantum Dots*, Springer, Heidelberg, 2002.
[170] K. MEERBERGEN, *Locking and restarting quadratic eigenvalue solvers*, SIAM J. Sci. Comput., 22 (2000), pp. 1814–1839.
[171] ———, *Fast frequency response computation for Rayleigh damping*, Internat. J. Numer. Methods Engrg., 73 (2008), pp. 96–106.
[172] K. MEERBERGEN, C. SCHRÖDER, AND H. VOSS, *A Jacobi-Davidson method for two-real-parameter nonlinear eigenvalue problems arising from delay-differential equations*, Numer. Linear Algebra Appl., 20 (2013), pp. 852–868.
[173] V. MEHRMANN AND D. WATKINS, *Polynomial eigenvalue problems with Hamiltonian structure*, Electron. Trans. Numer. Anal., 13 (2002), pp. 106–118.
http://etna.ricam.oeaw.ac.at/vol.13.2002/pp106-118.dir/pp106-118.pdf
[174] T. MEL'NIK AND S. NAZAROV, *Asymptotics of the solution to the Neumann spectral problem in a domain of the "tooth-comb" type*, J. Math. Sci., 85 (1997), pp. 2326–2346.
[175] S. A. MOHAMMADI AND H. VOSS, *Variational characterization of real eigenvalues in linear viscoelastic oscillators*, Math. Mech. Solids, 23 (2018), pp. 1377–1388.
[176] ———, *On the distribution of real eigenvalues in linear viscoelastic oscillators*, Numer. Linear Algebra Appl., 26 (2019), Art. e2228, 13 pages.
[177] H. MORAND AND R. OHAYON, *Substructure variational analysis of the vibrations of coupled fluid-structure systems*, Internat. J. Numer. Methods Engrg., 14 (1979), pp. 741–755.
[178] V. A. MOROZOV, *On the solution of functional equations by the method of regularization*, Soviet Math. Dokl., 7 (1966), pp. 414–417.
[179] Y. NAKATSUKASA AND V. NOFERINI, *Inertia laws and localization of real eigenvalues for generalized indefinite eigenvalue problems*, Linear Algebra Appl., 578 (2019), pp. 272–296.
[180] A. NEUMAIER, *Residual inverse iteration for the nonlinear eigenvalue problem*, SIAM J. Numer. Anal., 22 (1985), pp. 914–923.
[181] ———, *Introduction to Numerical Analysis*, Cambridge University Press, Cambridge, 2001.
[182] S.-I. NICULESCU, *Delay Effects on Stability. A Robust Control Approach*, Springer, London, 2001.
[183] V. NIENDORF AND H. VOSS, *Detecting hyperbolic and definite matrix polynomials*, Linear Algebra Appl., 432 (2010), pp. 1017–1035.
[184] L. OLSON AND K.-J. BATHE, *Analysis of fluid-structure interaction. A direct symmetric coupled formulation based on the fluid velocity potential*, Comput. & Structures, 21 (1985), pp. 21–32.
[185] J. ORTEGA AND W. C. RHEINBOLD, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.
[186] M. R. OSBORNE, *A new method for the solution of eigenvalue problems*, Comput. J., 7 (1964), pp. 228–232.
[187] ———, *Inverse iteration, Newton's method, and nonlinear eigenvalue problems*, in Contribution of Dr. J. H. Wilkinson to Numerical Analysis, Symp. Proc. Series 19, Inst. Math. Appl., London, 1978, pp. 21–53.
[188] M. R. OSBORNE AND S. MICHAELSON, *The numerical solution of eigenvalue problems in which the eigenvalue parameter appears nonlinearly, with an application to differential equations*, Comput. J., 7 (1964), pp. 66–71.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[189] C. C. PAIGE AND M. A. SAUNDERS, *LSQR: an algorithm for sparse linear equations and sparse least squares*, ACM Trans. Math. Software, 8 (1982), pp. 43–71.
[190] B. N. PARLETT, *The Symmetric Eigenvalue Problem*, SIAM, Philadelphia, 1998.
[191] H. POINCARE, *Sur les Equations aux Derivees Partielles de la Physique Mathematique*, Amer. J. Math., 12 (1890), pp. 211–294.
[192] L. REICHEL AND A. SHYSHKOV, *A new zero-finder for Tikhonov regularization*, BIT, 48 (2008), pp. 627–643.
[193] F. RENAUT, J.-L. DION, G. CHEVALLIER, I. TAWFIC, AND R. LEMAIRE, *A new identification method of viscoelastic behavior: application to the generalized Maxwell model*, Mech. Syst. Signal Proc., 25 (2011), pp. 991–1010.
[194] R. A. RENAUT AND H. GUO, *Efficient algorithms for solution of regularized total least squares*, SIAM J. Matrix Anal. Appl., 26 (2004/05), pp. 457–476.
[195] R. RODRÍGUEZ AND J. E. SOLOMIN, *The order of convergence of eigenfrequencies in finite element approximations of fluid-structure interaction problems*, Math. Comp., 65 (1996), pp. 1463–1475.
[196] E. H. ROGERS, *A minimax theory for overdamped systems*, Arch. Rational Mech. Anal., 16 (1964), pp. 89–96.
[197] ———, *Variational properties of nonlinear spectra*, J. Math. Mech., 18 (1968), pp. 479–490.
[198] M. ROJAS, S. A. SANTOS, AND D. C. SORENSEN, *Algorithm 873: LSTRS: MATLAB software for large-scale trust-region subproblems and regularisation*, ACM Trans. Math. Software, 34 (2008), Art. 11, 28 pages.
[199] M. W. ROSTAMI, *New algorithms for computing the real structured pseudospectral abscissa and the real stability radius of large and sparse matrices*, SIAM J. Sci. Comput., 37 (2015), pp. S447–S471.
[200] M. W. ROSTAMI AND F. XUE, *Robust linear stability analysis and a new method for computing the action of the matrix exponential*, SIAM J. Sci. Comput., 40 (2018), pp. A3344–A3370.
[201] A. RUHE, *Rational Krylov: a practical algorithm for large sparse nonsymmetric matrix pencils*, SIAM J. Sci. Comput., 19 (1998), pp. 1535–1551.
[202] ———, *A rational Krylov algorithm for nonlinear matrix eigenvalue problems*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI), 268 (2000), pp. 176âĂŞ180.
[203] Y. SAAD, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.
[204] K. SCHREIBER, *Nonlinear Eigenvalue Problems: Newton-type Methods and Nonlinear Rayleigh Functionals*, PhD. Thesis, Institut für Mathematik, TU Berlin, Berlin, 2008.
[205] G. SCHWARZ, *Estimating the dimension of a model*, Ann. Statist., 6 (1978), pp. 461–464.
[206] H. SCHWETLICK AND K. SCHREIBER, *Nonlinear Rayleigh functionals*, Linear Algebra Appl., 436 (2012), pp. 3991–4016.
[207] D. SIMA, *Regularization Techniques in Model Fitting and Parameter Estimation*, PhD. Thesis, Faculteit Ingenieurswetenschappen, Katolieke Universiteit Leuven, Leuven, 2006.
[208] D. M. SIMA, S. VAN HUFFEL, AND G. H. GOLUB, *Regularized total least squares based on quadratic eigenvalue problem solvers*, BIT, 44 (2004), pp. 793–812.
[209] K. SINGH, *Eigenvalue and eigenvector computation for discrete and continuous structures composed of viscous materials*, Inter. J. Mech. Sci., 110 (2016), pp. 127–137.
[210] G. L. G. SLEIJPEN, A. G. L. BOOTEN, D. R. FOKKEMA, AND H. A. VAN DER VORST, *Jacobi-Davidson type methods for generalized eigenproblems and polynomial eigenproblems*, BIT, 36 (1996), pp. 595–633.
[211] G. SLEIJPEN AND H. VAN DER VORST, *A Jacobi-Davidson iteration method for linear eigenvalue problems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 401–425.
[212] G.L. SLEIJPEN, H. VAN DER VORST, AND M. VAN GIJZEN, *Quadratic eigenproblems are no problem*, SIAM News, 8 (1996), pp. 9–10.
[213] D. SORENSEN, *Minimization of a large-scale quadratic function subject to s spherical constraint*, SIAM J. Optim., 7 (1997), pp. 141–161.
[214] M. STAMMBERGER AND H. VOSS, *On an unsymmetric eigenvalue problem governing free vibrations of fluid-solid structures*, Electron. Trans. Numer. Anal., 36 (2009/10), pp. 113–125.
http://etna.ricam.oeaw.ac.at/vol.36.2009-2010/pp113-125.dir/pp113-125.pdf
[215] ———, *Automated multi-level sub-structuring for fluid-solid interaction problems*, Numer. Linear Algebra Appl., 18 (2011), pp. 411–427.
[216] ———, *Variational characterization of eigenvalues of a non-symmetric eigenvalue problem governing elastoacoustic vibrations*, Appl. Math., 59 (2014), pp. 1–13.
[217] Y. SU AND Z. BAI, *Solving rational eigenvalue problems via linearization*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 201–216.
[218] J. SYLVESTER, *A demonstration of the theorem that every homogeneous quadratic polynomial is reducible by orthogonal substitutions to the form of a sum of positive and negative squares*, Philos. Mag. (4), 4 (1852), pp. 138–142.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[219] D. B. SZYLD, E. VECHARYNSKI, AND F. XUE, *Preconditioned eigensolvers for large-scale nonlinear Hermitian eigenproblems with variational characterizations. II. Interior eigenvalues*, SIAM J. Sci. Comput., 37 (2015), pp. A2969–A2997.
[220] D. B. SZYLD AND F. XUE, *Local convergence analysis of several inexact Newton-type algorithms for general nonlinear eigenvalue problems*, Numer. Math., 123 (2013), pp. 333–362.
[221] ———, *Local convergence of Newton-like methods for degenerate eigenvalues of nonlinear eigenproblems. I. Classical algorithms*, Numer. Math., 129 (2015), pp. 353–381.
[222] ———, *Local convergence of Newton-like methods for degenerate eigenvalues of nonlinear eigenproblems. II. Accelerared algorithms*, Numer. Math., 129 (2015), pp. 383–403.
[223] U. TAUTENHAHN, *Regularization of linear ill-posed problems with noisy right hand side and noisy operator*, J. Inverse Ill-Posed Probl., 16 (2008), pp. 507–523.
[224] A. TIKHONOV, *Solution of incorrectly formulated problems and the regularization method*, Soviet Math. Dokl., 4 (1963), pp. 1035–1038.
[225] F. TISSEUR AND K. MEERBERGEN, *The quadratic eigenvalue problem*, SIAM Rev., 43 (2001), pp. 235–286.
[226] R. E. L. TURNER, *Some variational principles for a nonlinear eigenvalue problem*, J. Math. Anal. Appl., 17 (1967), pp. 151–160.
[227] ———, *A class of nonlinear eigenvalue problems*, J. Functional Analysis, 2 (1968), pp. 297–322.
[228] G. UNGER, *Convergence orders of iterative methods for nonlinear eigenvalue problems*, in Advanced Finite Element Methods and Applications, T. Apel and O. Steinbach, eds., vol. 66 of Lect. Notes Appl. Comput. Mech., Springer, Heidelberg, 2013, pp. 217–237.
[229] R. VAN BEEUMEN, K. MEERBERGEN, AND W. MICHIELS, *Compact rational Krylov methods for nonlinear eigenvalue problems*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 820–838.
[230] S. VAN HUFFEL (ED.), *Recent Advances in Total Least Squares Techniques and Errors-in-Variables Modeling*, SIAM, Philadelphia, 1997.
[231] S. VAN HUFFEL AND P. LEMMERLING (EDS.), *Total Least Squares and Errors-in-Variables Modeling*, Springer, Dordrecht, 2002.
[232] S. VAN HUFFEL AND J. VANDEVALLE, *The Total Least Squares Problem*, SIAM, Philadelphia, 1991.
[233] C. F. VAN LOAN, *Generalizing the singular value decomposition*, SIAM J. Numer. Anal., 13 (1976), pp. 76–83.
[234] H. VOSS, *A maxmin principle for nonlinear eigenvalue problems with application to a rational spectral problem in fluid-solid vibration*, Appl. Math., 48 (2003), pp. 607–622.
[235] ———, *An Arnoldi method for nonlinear symmetric eigenvalue problems*, in Online Proceedings of the SIAM Conference on Applied Linear Algebra, 2003.
https://archive.siam.org/meetings/la03/proceedings/
[236] ———, *Initializing iterative projection methods for rational symmetric eigenproblems*, Online Proceedings of the Dagstuhl Seminar Theoretical and Computational Aspects of Matrix Algorithms, 2003.
ftp://ftp.dagstuhl.de/pub/Proceedings/03/03421/03421.VossHeinrich.Other.pdf
[237] ———, *An Arnoldi method for nonlinear eigenvalue problems*, BIT, 44 (2004), pp. 387–401.
[238] ———, *A Jacobi–Davidson method for nonlinear eigenproblems*, in Computational ScienceâĂŤICCS 2004. Part II, M. Bubak, G. D. van Albada, P. M. A. Sloot, and J. J. Dongarra, eds, Lecture Notes in Computer Science 3037, Springer, Berlin, 2004, pp. 34–41.
[239] ———, *Numerical calculation of the electronic structure for three-dimensional quantum dots*, Comput. Phys. Comm., 174 (2006), pp. 441–446.
[240] ———, *Iterative projection methods for computing relevant energy states of a quantum dot*, J. Comput. Phys., 217 (2006), pp. 824–833.
[241] ———, *A new justification of the Jacobi-Davidson method for large eigenproblems*, Linear Algebra Appl., 424 (2007), pp. 448–455.
[242] ———, *A Jacobi-Davidson method for nonlinear and nonsymmetric eigenproblems*, Comput. & Structures, 85 (2007), pp. 1284–1292.
[243] ———, *A minmax principle for nonlinear eigenproblems depending continuously on the eigenparameter*, Numer. Linear Algebra Appl., 16 (2009), pp. 899–913.
[244] ———, *Iterative projection methods for large-scale nonlinear eigenvalue problems*, in Computational & Technology Reviews, Vol. 1, B.H.V. Topping, J. M. Adam, F. J. Pallarés, R. Bru, and M. L. Romero, eds., Saxe-Coburg Publications, Stirling, 2010, pp. 187–211.
[245] ———, *Nonlinear eigenvalue problems*, in Handbook of Linear Algebra, 2nd ed., Chpt. 60, L. Hogben, ed., CRC Press, Boca Raton, 2014
[246] ———, *Variational principles for eigenvalues of nonlinear eigenproblems*, in Numerical Mathematics and Advanced Applications—ENUMATH 2013, A. Abdulle, S. Deparis, D. Kressner, F. Nobile, and M. icasso, eds., vol. 103 of Lect. Notes Comput. Sci. Eng., Springer, Cham, 2015, pp. 305–313.
[247] H. VOSS AND M. STAMMBERGER, *Structural-acoustic vibration problems in the presence of strong coupling*, J. Pressure Vessel Techn., 135 (2013), Art. 011303, 8 pages.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[248] H. VOSS AND B. WERNER, *A minimax principle for nonlinear eigenvalue problems with applications to nonoverdamped systems*, Math. Methods Appl. Sci., 4 (1982), pp. 415–424.
[249] ———, *Solving sparse nonlinear eigenvalue problems*, Preprint 82/4, Inst. f. Angw. Math., Universität Hamburg, Hamburg, 1984.
[250] H. VOSS, K. YILDIZTEKIN, AND X. HUANG, *Nonlinear low rank modification of a symmetric eigenvalue problem*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 515–535.
[251] H. VOSS, J. YIN, AND P. CHEN, *Enhancing eigenvector approximations of huge gyroscopic eigenproblems for AMLS with subspace iteration*, Trans. Control Mech. Syst., 2 (2013), pp. 294–301.
[252] ———, *Preconditioning subspace iteration for large eigenvalue problems with automated multi-level substructuring*, Asian J. Math. Comput. Res., 10 (2016), pp. 136–150.
[253] N. WAGNER AND S. ADHIKARI, *Symmetric state–space methods for a class of nonviscously damped systems*, AIAA J., 41 (2003), pp. 951–956.
[254] H. F. WEINBERGER, *On a nonlinear eigenvalue problem*, J. Math. Anal. Appl., 21 (1968), pp. 506–509.
[255] B. WERNER, *Das Spektrum von Operatorenscharen mit verallgemeinerten Rayleighquotienten*, Arch. Rational Mech. Anal., 42 (1971), pp. 223–238.
[256] H. WEYL, *Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung)*, Math. Ann., 71 (1912), pp. 441–479.
[257] C. YANG, W. GAO, Z. BAI, X. LI, L. LEE, P. HUSBANDS, AND E. NG, *An algebraic sub-structuring method for large-scale eigenvalue calculations*, SIAM J. Sci. Comput., 27 (2005), pp. 873–892.
[258] C. YANG, J. C. MEZA, AND L.-W. WANG, *A trust region direct constrained minimization algorithm for the Kohn-Sham equation*, SIAM J. Sci. Comput., 29 (2007), pp. 1854–1875.
[259] H. YANG, *A method for eigenvalues of sparse λ–matrices*, Internat. J. Numer. Methods Engrg., 19 (1983), pp. 943–948.
[260] J. YIN, H. VOSS, AND P. CHEN, *Improving eigenpairs of automated multilevel substructuring with subspace iteration*, Comput. & Structures, 119 (2013), pp. 115–124.