

# COMPUTING INTERIOR EIGENVALUES OF MEDIUM SIZED GENERALIZED SYMMETRIC EIGENPROBLEMS

**Wolfgang Mackens and Heinrich Voss**

Section of Mathematics  
Technical University of Hamburg-Harburg, Germany  
e-mail: {mackens, voss}@tu-harburg.de

**Key words:** Generalized Eigenvalue Problem, Projection Method, Rational Krylov Space, Error Estimation, Shift-Invert Strategy, Iterative Refinement

---

**Abstract.** *We present an algorithmic framework to compute approximations to all eigenvalues of a generalized symmetric eigenvalue problem in a prespecified interval together with rigorous error bounds. The method needs the computation of some LU factorizations of the incorporated matrices. The ability to do this excludes the treatment of very large problems.*

---

## 1 Introduction

We consider the generalized eigenvalue problem

$$Kx = \lambda Mx, \tag{1}$$

where  $K, M \in \mathbb{R}^{(n,n)}$  are symmetric real matrices,  $M$  is positive definite and eigenvalues  $\lambda$  and corresponding eigenvectors  $x \in \mathbb{R}^n$  fulfilling (1) are sought.

Such problems arise in the analysis of the dynamic response of a linear structure using finite element methods, e.g., in which case  $K$  and  $M$  are the stiffness and mass matrices, respectively. Our interest centers on the approximation of all eigenvalues in a prespecified real interval  $[a, b]$ . This task is of interest whenever resonance of structures to outer excitations from a known frequency interval is studied such as in earthquake response analysis of buildings or resonances of ships and aircrafts to engine provided excitations.

For very large matrices this problem is addressed in [1] and [8], e.g. Very similar to the approach of Ericsson and Ruhe [3] the application of the method proposed in this paper is limited to not too large problems, since we require that the exact number  $N$  of eigenvalues in  $[a, b]$  can be determined in reasonable time. One should furthermore be able to solve linear equations with system matrices of the form  $K - \sigma M$ . Both requirements are fulfilled if one can compute an LU factorization of  $K - \sigma M$ . Then the number  $N$  can be found, e.g., with the aid of Sylvester's theorem of inertia from LU factorizations of  $K - aM$  and  $K - bM$ .

The guiding idea of the development of the proposed method originates from the desire to apply the error estimation procedure from [6] to estimate the errors of approximations of interior eigenvalues. The method from [6] is based on Krylov-Bogoliubov-Kahan and Kato-Temple bounds [14] and is meant to estimate errors (including those of multiple and clustered eigenvalues) at the two extremal ends of the spectrum. We make it work for interior eigenvalues through the application of a shift-invert-strategy from [3] (for a further discussion see as well [17]): With a suitably chosen value  $\sigma \in [a, b]$  different from an eigenvalue one approximates eigenpairs for the equation

$$L^T(K - \sigma M)^{-1}L\tilde{x} = \tilde{\lambda}\tilde{x} \tag{2}$$

with  $M = LL^T$  the Cholesky factorization of  $M$ . The eigenvectors  $\tilde{x}_i$  of the new system matrix

$$W(\sigma) := L^T(K - \sigma M)^{-1}L \tag{3}$$

are related to those  $x_i$  of the eigenproblem (1) through

$$\tilde{x}_i = L^T x_i \tag{4}$$

and the eigenvalues are transformed according to

$$\tilde{\lambda}_i = f(\lambda_i) := \frac{1}{\lambda_i - \sigma}. \tag{5}$$

Hence all eigenvalues in  $[a, \sigma)$  and in  $(\sigma, b]$  will become smallest or largest extremal eigenvalues of the new problem (2). In Ericsson's and Ruhe's approach [3] this property made them accessible to Lanczos approximation. In our setting they are thus prepared to undergo our error estimation procedure from [6].

Though the desire to apply the error estimate from [6] actually is the driving force of the development, we shall describe the resulting error estimation algorithm for the interior eigenvalues in the present paper in the late Section 4. This is due to our decision to report on the ingredients of the whole procedure after this introduction in the order of their actual appearance within the computational framework.

The error estimation procedure is described in [6] for eigenpair approximations gained by condensation procedures, but an inspection of that paper shows that the method works under the mere assumption that the approximate eigenpairs are Ritz approximations (cf. [14] or Section 2) of the extremal eigenvalues. As approximation space for the Ritz method we chose in this paper vectors

$$v_i := W(s_i)r, \quad i = 1, \dots, m, \quad (6)$$

with  $m \geq N$ ,  $W$  according to (3), a random vector  $r$  and suitably located shifts  $a \leq s_1 < s_2 < \dots < s_m \leq b$  which in the ideal case approximate the desired eigenvalues. We shall describe approaches to choose the shifts  $s_i, i = 1, \dots, m$ , together with their different pros and cons in Section 2.

Due to the well known relation

$$W(\sigma)W(\mu) = \frac{1}{\sigma - \mu} (W(\sigma) - W(\mu)) \quad (7)$$

the projected matrix  $V^T W(\sigma) V$  with  $V = [v_1, \dots, v_m]$  can be computed very efficiently [10]. We report on this computation in Section 3.

The main body of the algorithmic procedure is contained in the sections 2 to 4, where 4 describes the error estimation for the gained approximations.

Section 5 contains several numerical examples with practical FE-systems from the Harwell-Boeing test problem collection [12], illuminating some of the approaches.

In Section 6 we report on several variations and approaches to improve the basic algorithm, such as iterative refinements, coping with multiple eigenvalues and exploiting computed matrix factorizations.

The final Section 7 collects the main conclusions.

## 2 Choice of the Ritz space

### 2.1 Rayleigh-Ritz methods in general

In the Rayleigh-Ritz approach to eigenpair approximation one "projects" an original (symmetric) eigenvalue problem

$$Ax = \lambda x, \quad A = A^T \in \mathbb{R}^{(n,n)}, x \in \mathbb{R}^n, \quad (8)$$

to a space of low dimension  $m \ll n$ , spanned by some linearly independent ansatz-vectors  $v_1, \dots, v_m$  through the following dimensional reduction process: With  $V := (v_1, \dots, v_m) \in \mathbb{R}^{(n,m)}$  assume  $x = Vy$  for some  $y \in \mathbb{R}^m$ . The demand that the projection of the residual  $AVy - \lambda Vy$  to  $\text{span}(V)$  be zero, leads to the obviously symmetric low dimensional eigenvalue problem

$$(V^T AV)y = \theta (V^T V)y \tag{9}$$

which is called the projection of (8) to  $\text{span}(V)$ . Notice that the problem is regular, since the reduced mass-matrix  $V^T V$  is symmetric and positive definite. We replaced  $\lambda$  from (8) by  $\theta$  to indicate that the eigenvalues  $\theta_1 \leq \theta_2 \leq \dots \leq \theta_m$  of the latter eigenproblem are, in general, different from the eigenvalues  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$  of (8).

Notice that for the special case  $m = 1$  the solution of the (then) one dimensional projected system (9) reduces to the computation of the Rayleigh-quotient

$$\theta = R(v) := \frac{v^T Av}{v^T v}. \tag{10}$$

Having found the eigenpairs  $(\theta_j, y_j), j = 1, \dots, m$  of (9), the  $\theta$ -values are called Ritz values for  $V$  and the corresponding eigenvector approximations  $\hat{x}_i := Vy_i, i = 1, \dots, m$  are called the Ritz vectors, together they are Ritz pairs.

From the minimax characterization of the eigenvalues (cf. [14], e.g.) one finds the following comparison between the eigenvalues  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$  and the Ritz values  $\theta_1 \leq \dots \leq \theta_m$ :

$$\lambda_i \leq \theta_i \leq \lambda_{n-m+i}, \quad i = 1, \dots, m. \tag{11}$$

Thus the  $i$ -th Ritz value is always an upper bound for the  $i$ -th eigenvalue, if both sets are assumed to be ordered by magnitude. Likewise the largest Ritz value is a lower bound for the largest eigenvalue. The second largest Ritz value is a lower bound for the second largest eigenvalue and so forth.

Notice that the resulting Ritz pairs of the Ritz procedure are independent of the specific basis  $v_1, \dots, v_m$  of  $\text{span}(V)$ . This justifies the name “projection of the eigenproblem to  $\text{span}(V)$ ” for the reduced problem (9). This has to be kept in mind, if we write

$$\theta_i = \theta_i(V) \text{ and } \hat{x}_i = \hat{x}_i(V), \quad i = 1, \dots, m \tag{12}$$

to indicate that the Ritz pairs are totally determined by the choice of  $V$ .

## 2.2 Some common Ritz approaches

There are lots of Ritz approaches in the literature. Each procedure is characterized by a specific method to build the space  $V$  which — very often — is complemented by a procedure to iteratively refine  $V$  with the aid of the previously computed Ritz pairs.

In all the approaches the aim of the procedure to determine  $V$  is to maximize the portions of the desired eigenvectors in the set of starting-vectors. Any attempt — whether computational or not — to adjust  $V$  to the directions of the desired eigenvectors is welcome.

In what is called eigenreanalysis [7] eigenvectors of similar eigenproblems are fed into the Ritz spaces. In nodal condensation approaches, nodal unit vectors are chosen such that they jointly represent a set of desired eigenvectors closely [4] and afterwards (as discussed in [11]) the quality of the resulting set of vectors is increased by application of a simultaneous inverse iteration step (when the smallest eigenvalues are of interest).

Repeated application of a joint multiplication with  $A$  or  $A^{-1}$  on a set of Ritz vector candidates is as well the basis of the “subspace iteration” [14] for the computation of the eigenpairs with eigenvalues of largest or smallest modulus, respectively. The number  $m$  of independent  $v$ -vectors is chosen therein at least one unit larger than the number  $N$  of desired eigenvectors. To approximate the lower part of the spectrum, e.g.,  $V$  is multiplied by  $A$ 's inverse. Afterwards the new set of  $v$ -vectors is normally formed from those  $m$  Ritz vectors  $\hat{x}_i(A^{-1}V)$  corresponding to the  $m$  smallest Ritz values  $\theta_1(A^{-1}V), \dots, \theta_m(A^{-1}V)$  together with  $m - N$  additional “guard-vectors”, which are randomly chosen for each new step. The Ritz vectors carry the information from the performed iteration steps while the guard vectors help not to overlook relevant eigendirections, which might not have been present in the starting  $V$ .

While the  $v$ -vectors of the previous steps are discarded in subspace iteration, they are kept in an increasing  $V$ -matrix in Krylov space iterations [16]. In case that the starting  $V$  consists of only one vector one is lead to ordinary Krylov space methods otherwise one is confronted with block-Krylov approaches. If the  $v$ -vectors are orthonormalized as they are computed, the methods are known as Lanczos methods. Advantageous of the Krylov methods are, e.g., an inexpensive way to compute the projected matrix  $V^TAV$  together with an efficient solution of the projected system and low memory requirements.

The previously treated methods only make use of the Ritz vectors of the previous step to build the  $V$ -matrix for the next step. It is, however, often very beneficial (though sometimes also more delicate through an increased sensitivity) to use the Ritz values as well. In the Rayleigh quotient iteration the (single) new  $v$ -vector is calculated through a shifted inverse iteration step

$$w := (A - \theta_1 I)^{-1} \hat{x}_1, \quad v_1 := \frac{w}{\|w\|_2}. \quad (13)$$

This shifted inverse iteration with Rayleigh quotient shift is known to converge globally to an eigenpair for almost every starting vector with a final cubic convergence rate [14] if the eigenvalue is simple. The delicacy lies in the fact that only starting vectors rather close to a certain eigenvector guaranty that the iteration will converge to the corresponding eigenpair. For other starting vectors it is – ad hoc – totally unclear to which eigenpair convergence will occur.

If the approximations (13) of previous steps are not discarded after their first use but are added to the previous ones to form the increasing set of vectors  $V$  one arrives at the rational Krylov space methods of Axel Ruhe [15]. Due to the Krylov structure of the space the projected matrices can again be formed very efficiently.

Within the modified block Newton eigenanalysis from [9] (as well as to some extent in the condensation-projection method from [6] and a similar approach in [13]) both the

Ritz vectors and the Ritz values are used in the construction of the new  $V$ -set. Here  $m$  simultaneous shifted inverse iterations are carried out wherein each Ritz vector  $\hat{x}_i$  of a Ritz pair  $(\theta_i, \hat{x}_i)$  is subjected to its own inverse iteration

$$v_i^{\text{new}} := (A - \theta_i I)^{-1} v_i, \quad i = 1, \dots, m. \quad (14)$$

The convergence order of this procedure is locally at least quadratic in the case of simple eigenvalues (it can be interpreted as Newton's iteration to solve a certain nonlinear equation). As for its one dimensional ( $m = 1$ ) special case, the Rayleigh quotient iteration, it is not a priori clear to which pairs the global iteration will finally converge.

### 2.3 The Ritz projection of the present paper

In the present paper we start with a single random vector  $r \in \mathbb{R}^n$  to create a basis for a rational Krylov space in the form

$$v_i := W(s_i)r, \quad i = 1, \dots, m, \quad (15)$$

with  $W$  from (3), a set of  $m$  shifts

$$a = s_1 < s_2 < \dots < s_m = b \quad (16)$$

and  $m \geq N$ , where  $N$  is the exact number of eigenvalues in  $[a, b]$ , which is at the start determined from the inertia of  $K - aM$  and  $K - bM$ .

That the  $v$ -vectors (15) are indeed rational Krylov vectors is easily seen from

$$v_i := W(s_i)r := L^T(K - s_i M)^{-1} Lr = (L^{-1}KL^{-T} - s_i I)^{-1}r, \quad i = 1, \dots, m, \quad (17)$$

and the identity (7).

From (17) it follows that the  $v$ -vectors are of the form  $(A - s_i I)^{-1}r$  with

$$A = L^{-1}KL^{-T}, \quad (18)$$

and from (7) it follows by simple induction, that

$$\text{span}\{V\} = \text{span} \left\{ (A - s_1 I)^{-1}r, (A - s_2 I)^{-1}(A - s_1 I)^{-1}r, \dots, \prod_{i=1}^m (A - s_i I)^{-1}r \right\}.$$

Instead of using the same vector  $r$  within the definition of the  $v_i$ -vectors in (17) we could have chosen a different vector  $r_i$  for each single  $v_i$ , by which choice we could have tried to mimic the iteration step (14). There are two reasons for using just one single vector  $r$ : First, one has only little information about eigenvector structures corresponding to interior eigenvalues and hence there is normally no a priori knowledge on how to choose eigenvector approximations  $r_i$  from scratch (we shall indicate in Section 6, however, how such preinformation about eigenvectors can be incorporated into one vector  $r$  when it comes to an iterative refinement of gained Ritz pairs). Second, using only one  $r$ -vector is

beneficial in computing the projected matrices in Section 3 (If there are multiple eigenvalues one has to use more than one vector to resolve the eigenspaces, cf. Section 6, but we then use as few vectors as possible).

With the shifts  $s_i$ ,  $i = 1, \dots, m$ , we try to approximate the distribution of the eigenvalues in  $[a, b]$  as close as possible, hoping that a shifted inverse iteration step with  $(A - s_i I)^{-1}$  will amplify those eigenparts of  $r$  which correspond to the eigenvalues close to  $s_i$ . We always include the values  $a$  and  $b$  into the set of shifts, since we assume the inertia of  $(K - aM)$  and  $(K - bM)$  to be computed for the determination of  $N$ . Since we further assume that this is done by use of LU or  $LDL^T$  factorizations of these matrices the computation of  $(A - aI)^{-1}r$  and  $(A - bI)^{-1}r$  is nearly for free.

Moreover we always included the value  $\sigma$  needed in the projected eigenproblem of Ericsson's and Ruhe's shift-invert-matrix  $L^T(K - \sigma M)^{-1}L$  into the set of shifts since an LU factorization of this matrix has to be computed anyhow, and linear systems with this system matrix can be solved cheaply.

In this first draft of the algorithm we always chose  $\sigma$  to be the midpoint of the interval

$$\sigma := (a + b)/2. \tag{19}$$

Actually, in a more elaborate implementation we will use a more sophisticated choice of  $\sigma$ . However, its design is still under investigation.

For the choice of the remaining shifts basically two approaches are reasonable: The shifts are prespecified or they are constructed adaptively in the course of the algorithm.

The main advantage of placing the shifts at the very beginning is that the computation of the  $v_i$  can be carried out fully in parallel on several processors. This strategy makes sense if information about the distribution of the eigenvalues in  $[a, b]$  is at hand (like the asymptotic rules for the eigenvalues of discretized differential operators). In our numerical tests we placed the shifts equidistantly in  $[a, b]$ . Other placements of the shifts like fast Leja points (cf. [2]) are under investigation.

A severe disadvantage of prespecifying the shifts is that no use is made of the information about the spectrum that is gained from the inertia of  $A - s_i I$  obtained during the direct computation of  $(A - s_i I)^{-1}r$ . Moreover, as we observed in our numerical tests, equidistributed shifts tend to overlook part of the eigenvalues from eigenvalue clusters, and the projection procedure tends to become numerically unstable for larger dimensions  $m$  of the projected problem.

To exploit the information about the distribution of the eigenvalues received from the inertia of  $A - s_i I$  in an adaptive selection of the shifts we considered the following

**Sequential adaptive bisection strategy**

*We start with the trivial distribution  $s_1^{[0]} := a$  and  $s_2^{[0]} := b$  and the knowledge of the number of eigenvalues contained in this interval.*

*Within the  $(k + 1)$ st refinement step the partitioning*

$$a = s_1^{[k]} < s_2^{[k]} < \dots < s_{k+2}^{[k]} = b$$

together with the known numbers of eigenvalues in the various intervals is refined by introducing one additional shiftpoint  $s_{i+1}^{[k+1]}$  in the middle of one of the intervals  $[s_i^{[k]}, s_{i+1}^{[k]}]$ . This interval is chosen to be the interval containing the largest number of eigenvalues. If there are several of such intervals, choose one with largest length (the leftmost, e.g, if there are again more than one of the latter). Determine the inertia of  $(K - s_{i+1}^{[k+1]}M)$  when computing  $(A - s_{i+1}^{[k+1]}I)^{-1}r$  to determine the number of eigenvalues in the two newly created intervals.

The obvious advantage is that the distribution of shifts is a better approximation of the distribution of eigenvalues in  $[a, b]$ . We observed that the projection method with this choice of shifts were able to resolve eigenvalue clusters. The disadvantage, however, is that the straightforward parallelization of the method mentioned above for prespecified shifts is excluded.

### 3 Computing the projected matrices

Having provided the  $m \geq N$  columns  $v_i := (A - s_i I)^{-1}r$ ,  $i = 1, \dots, m$ , of  $V \in \mathbb{R}^{n \times m}$  one has to compute the elements

$$v_i^T (A - \sigma I)^{-1} v_j \text{ and } v_i^T v_j \tag{20}$$

of the projected matrices

$$V^T (A - \sigma I)^{-1} V \text{ and } V^T V. \tag{21}$$

The latter elements are clearly just inner products of the known vectors  $v_i$ ,  $i = 1, \dots, m$ , which are easily calculable.

Not quite as obvious is the observation that the elements of the first matrix do not need  $m$  applications of  $(A - \sigma I)^{-1}$  to each of the  $v_i$  either. Instead the work to compute this matrix can be done only with the calculation of inner products among the  $v_i$ -vectors if these are complemented by the vector  $v_\sigma := (A - \sigma I)^{-1}r$  and if the Ericsson-Ruhe shift  $\sigma$  is not itself a member of the shift-set. (We shall deal with the minor modifications for the case that  $\sigma \in \{s_1, \dots, s_m\}$  in Section 6.2.) Notice that  $v_\sigma$  is in principle already available since one will certainly evaluate the inertia of  $(A - \sigma I)$ .

Since

$$\begin{aligned} (A - \sigma I)^{-1} v_j &= (A - \sigma I)^{-1} (A - s_j I)^{-1} r \\ &= (\sigma - s_j)^{-1} \cdot ((A - \sigma I)^{-1} r - (A - s_j I)^{-1} r) \\ &= (\sigma - s_j)^{-1} \cdot (v_\sigma - v_j), \end{aligned} \tag{22}$$

one has

$$v_i^T (A - \sigma I)^{-1} v_j = (\sigma - s_j)^{-1} (v_i^T v_\sigma - v_i^T v_j). \tag{23}$$

**Remarks:**

1. Notice that the lefthand side of (23) allows to interchange  $i$  and  $j$ . For a stable computation one would choose the index  $j$  on the right hand side such that  $|s_j - \sigma| \geq |s_i - \sigma|$ .
2. From (23) we see that the computation of the projected matrices can be done recursively during an adaptive determination of the shifts. If one more shift  $s_+$  is added to a set of existing shifts  $s_1 < s_2 < \dots < s_m$  together with a projection basis  $V$  and the corresponding projected matrices (21) the computation of the same data for the augmented shift set can be done by bordering of the projected matrices. This requires essentially the computation of  $v_+ := (A - s_+I)^{-1}r$  and the calculation of the  $(m + 2)$  inner products

$$v_+^T v_i, \quad i = 1, \dots, m, + \quad \text{and} \quad v_+^T v_\sigma$$

to allow the computation of the last row (or column) of

$$(V, v_+)^T (A - \sigma I)^{-1} (V, v_+) = \begin{pmatrix} V^T (A - \sigma I)^{-1} V & V^T (A - \sigma I)^{-1} v_+ \\ v_+^T (A - \sigma I)^{-1} V & v_+^T (A - \sigma I)^{-1} v_+ \end{pmatrix}.$$

3. In a practical implementation one will certainly solve the sequence of projected eigenvalue problems of increasing size, while constructing the shift set in an adaptive manner.

Since the computation of the error bounds can be realized with  $O(n)$  operations (cf. the sections 4 and 6.4), one will most probably even estimate the errors of the already gained approximations, in order to know when to stop the construction of  $v$ -elements.

In an advanced implementation it might even be useful to adjust the choice of new shift points to the intermediately computed eigenvalue approximations.

## 4 Error estimation procedure

By the shift-invert approach leading from the original eigenproblem (1) to the transformed problem (2) the spectral transformation  $f$  from (5) maps the interval of interest  $[s_1, s_m] = [a, b]$  to the union of the two semiinfinite intervals  $(-\infty, f(s_1)]$  and  $[f(s_m), \infty)$  as shown in Figure 1. Through this transformation the original eigenvalues  $\lambda_1 < \lambda_2 < \dots < \lambda_l$  in the interval  $[s_1, \sigma)$  and  $\lambda_{l+1} < \dots < \lambda_N$  in  $[\sigma, s_m]$  appear as the extreme eigenvalues  $f(\lambda_l) < \dots < f(\lambda_2) < f(\lambda_1)$  and  $f(\lambda_N) < \dots < f(\lambda_{l+1})$  of the transformed problem on the negative and positive side of the transformed spectrum, respectively. As extreme eigenvalues their Ritz approximations  $\rho_1 \leq \rho_2 \leq \dots \leq \rho_m$  from

$$(V^T (A - \sigma I)^{-1} V) y_i = \rho_i V^T V y_i \tag{24}$$

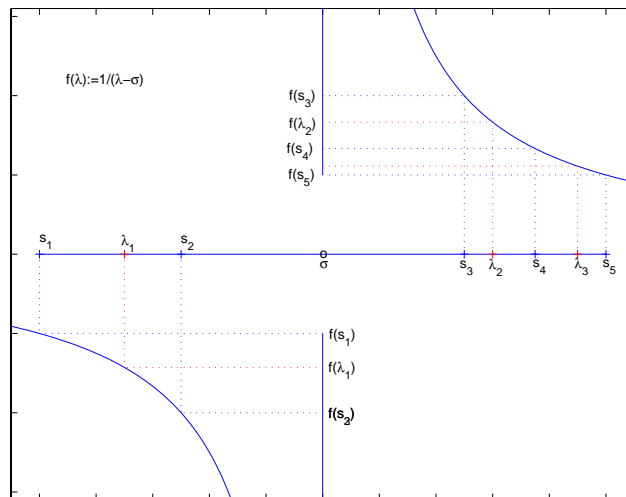


Figure 1: Shift-invert transformation

become accessible to the error estimation procedure from [6].

Due to the minimax-result (11) for Ritz values the  $l$  smallest  $\rho$ -values represent consecutive upper bounds of the values  $f(\lambda_l) < \dots < f(\lambda_2) < f(\lambda_1)$ . Similarly the  $N - l$  largest  $\rho$ -values deliver lower bounds of the values  $f(\lambda_N) < \dots < f(\lambda_{l+1})$ . Bounds in the opposite directions are found from the Krylov-Bogoliubov error estimate. This is formulated in the following Corollary as a specialization of Kahan's Theorem (cf. [14], page 240). We state the latter a little bit later where we shall need it.

**Corollary** *Let the matrix  $T \in \mathbb{R}^{(n,n)}$  be symmetric. With an eigenvector approximation  $w \neq 0$  and an eigenvalue approximation  $\kappa$  let  $R := Tw - \kappa w$  be the residual of this eigenpair approximation. Then there exists an eigenvalue  $\lambda$  of  $T$  such that*

$$|\kappa - \lambda| \leq \|R\|_2 / \|w\|_2.$$

If we apply this result to the data  $T = (A - \sigma I)^{-1}$ ,  $w = Vy_i$  and  $\kappa = \rho_i$  from (24), we find centered intervals around the  $\rho_i$ -values of radius  $\delta_i = \|(A - \sigma I)^{-1}Vy_i - \rho_i Vy_i\|_2 / \|Vy_i\|_2 > 0$  each containing at least one of the eigenvalues  $f(\lambda_j)$  of  $(A - \sigma I)^{-1}$ . If the intervals are disjoint, clearly  $m$  different eigenvalues are located.

Since we know the numbers  $l$  and  $N - l$  of eigenvalues less than  $f(s_1)$  and larger than  $f(s_m)$ , respectively, we can infer that we have approximated all the desired eigenvalues images under  $f$  if equally large numbers of intervals  $[\rho_i - \delta_i, \rho_i + \delta_i]$  belong to the sets  $(-\infty, f(s_1)]$  and  $[f(s_m), \infty)$ , respectively. With a refined argumentation from [6] using the minimax properties one even gets the improved estimates that

$$f(\lambda_j) \in [\rho_{l-j+1} - \delta_{l-j+1}, \rho_{l-j+1}], \quad j = 1, \dots, l \quad (25)$$

and

$$f(\lambda_j) \in [\rho_{m-j+l+1}, \rho_{m-j+l+1} + \delta_{m-j+l+1}], \quad j = l + 1, \dots, N \quad (26)$$

as long as

$$\rho_l \leq f(s_1) \tag{27}$$

and

$$\rho_{m-N+l+1} \geq f(s_m), \tag{28}$$

are fulfilled, respectively.

The latter conditions (27) and (28) mean that as many  $\rho$ -approximations belong to the intervals  $(-\infty, f(s_1)]$  and  $(f(s_m), \infty)$  as there are eigenvalues in these sets, respectively. Notice that (27) and (28) are conditions on the approximation quality of the Ritz approximation, specifically on the quality of the  $V$  matrix. If these conditions are not satisfied, there are eigenvalues which may not be caught in the computed intervals. In this case one has to try to improve  $V$ 's quality. This can for example be tried by adding more shifts and corresponding  $v$ -vectors or by an iterative refinement of the  $r$ -vector (cf. Section 6.1). Both measures are not guaranteed to be successful. If there are multiple eigenvalues for example in the chosen interval  $[a, b]$  there is no chance to satisfy (27) and (28) unless additional  $r$  vectors are incorporated in a block Krylov fashion into the  $V$  matrix (cf. Section 6.3).

The above arguments on the secure inclusion of all eigenvalues fail, if part of the intervals from (25) or (26) intersect. As is demonstrated in Section 5 most of these estimates can be preserved with Kahan's Theorem if only the conditions (27) and (28) prevail.

**Kahan's Theorem** ([14], p. 240) *Let  $Q \in \mathbb{R}^{(n,p)}$  have orthonormal columns. Associated with  $Q$  and the symmetric matrix  $T$  are the projected matrix*

$$H := Q^T T Q$$

*and its residual matrix*

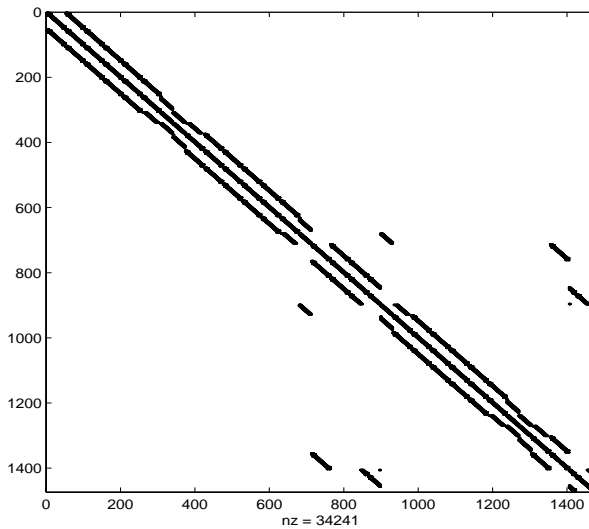
$$R := TQ - QH.$$

*Then there are  $p$  of  $T$ 's eigenvalues  $\{\lambda_{k_j} : j = 1, \dots, p\}$  which can be put in a one-to-one correspondence with the eigenvalues  $\theta_j, j = 1, \dots, p$  of  $H$  in such a way that*

$$|\theta_j - \lambda_{k_j}| \leq \|R\|_2, \quad j = 1, \dots, p.$$

This theorem has to be applied to all subsets of Ritz pairs  $(\rho_i, \hat{x}_i = Vy_i)$  whose corresponding intervals from (25) or (26) intersect. The result is – in principle – that after normalization of the Ritz vectors  $Vy_i$  the corresponding radii  $\delta_i$  are enlarged to the norm of the matrix  $R$  which contains in its columns all the residual vectors of the involved Ritz pairs.

Since this enlargement of the  $\delta_i$  can again lead to further intersection of intervals, the finally resulting procedure is somewhat more subtle. It may be read up in [6].


 Figure 2: Structure plot of  $K$  and  $M$ 

## 5 Numerical Results

To demonstrate the properties of our approach we consider the generalized eigenvalue problem representing the consistent mass formulation for an FEM model of an ore car given by the matrices BCSSTK12 and BCSSTM12 of the Harwell–Boeing Collection.

$K$  and  $M$  are symmetric and positive definite matrices of dimension  $n = 1473$  with band width 651 and average band width 62. The structure plot of  $K$  is given in Figure 2.

To determine the eigenvalues in the interval  $[a, b] := [1.0e4, 4.0e4]$  we choose the shift  $\sigma = 2.5e4$ . From the LU factorizations of  $K - aM$  and  $K - bM$  we obtain that there are 4 eigenvalues of problem (1) less than  $1.0e4$ , and that the interval under consideration contains 8 eigenvalues  $\lambda_5 \leq \dots \leq \lambda_{12}$ , and the LU factorization of  $K - \sigma M$  yields  $\lambda_5, \lambda_6 \in [a, \sigma]$  and  $\lambda_7, \dots, \lambda_{12} \in [\sigma, b]$ .

The adaptive bisection strategy of Section 2.3 constructs 12 shifts given in Figure 3 on the left (marked by  $\times$ ) and finds 8 eigenvalues marked by  $+$  where the sixth  $+$ -sign corresponds to a cluster of 2 eigenvalues. Table 1 contains in its second column the approximate eigenvalues  $\rho_j$  obtained by the projection method, and from the minmax arguments it follows that  $\rho_j \leq \lambda_j$  for  $j = 5, 6$  and  $\rho_j \geq \lambda_j$  for  $j = 7, 8, \dots, 12$ . Column 3 shows the complementary bounds  $\mu_j$  obtained from Krylov-Bogoliubov’s error estimate (i.e. the Corollary from Section 4), and column 4 the relative distance of the bounds. In column 5 we add the actual relative errors of the approximate eigenvalues.

The intervals  $[\mu_{10}, \rho_{10}]$  and  $[\mu_{11}, \rho_{11}]$  overlap, and therefore they still may correspond to the same eigenvalue. Applying Kahan’s Theorem proves that there are actually 2 eigenvalues  $\lambda_j \in [\tilde{\mu}_j, \rho_j]$ ,  $j = 10, 11$ , satisfying the bounds of Table 2.

We compare these results to the approximate eigenvalues in the interval  $[a, b]$  which are obtained by the projection method using 12 equidistant shifts in  $[a, b]$ . As shown in Table

j	$\rho_j$	$\mu_j$	$\frac{ \rho_j - \mu_j }{\min(\rho_j, \mu_j)}$	$\frac{\lambda_j - \rho_j}{\lambda_j}$
5	1.066420908193e+04	1.068296264806e+04	1.76e-03	5.69e-06
6	1.546421047373e+04	1.547374609033e+04	6.17e-04	9.06e-07
7	2.528083490054e+04	2.528060232147e+04	9.20e-06	-7.60e-09
8	2.556234711165e+04	2.555582320599e+04	2.55e-04	-3.00e-06
9	2.697249608837e+04	2.697218404729e+04	1.16e-05	-1.79e-09
10	3.929927623030e+04	3.929292836430e+04		-8.85e-08
11	3.930143770207e+04	3.928862317145e+04		-3.58e-07
12	3.976176380949e+04	3.976134880228e+04	1.04e-05	-3.67e-10

Table 1: Bounds of eigenvalues in [1e4, 4e4]

j	$\tilde{\mu}_j$	$\rho_j$	$(\rho_j - \tilde{\mu}_j)/\tilde{\mu}_j$
10	3.928498184227e+04	3.929927623030e+04	3.64e-04
11	3.928713899219e+04	3.930143770207e+04	3.64e-04

Table 2: Resolution of bounds by Kahan's Theorem

3 the method finds only 7 eigenvalues in this interval, and it misses one of the clustered eigenvalues  $\lambda_{10}$  and  $\lambda_{11}$ .

Next we tested the method for the interval  $[a, b] = [1.0e6, 1.44e6]$  which contains the eigenvalues  $\lambda_{61}, \dots, \lambda_{72}$  of problem (1), 6 below the shift  $\sigma = 1.22e6$  and 5 above  $\sigma$ . The bisection method produces 15 shifts given in Figure 3 (marked as  $\times$ ). The projection method with these shifts detects all eigenvalues (marked as  $+$  in Figure 3 on the right) in  $[a, b]$ , and the intervals derived from the Krylov-Bogoliubov bounds do not overlap. Table 4 shows the results in the same way as for the first interval.

The projection methods with 15 equidistant shifts in  $[a, b]$  yields 10 eigenvalues and again

j	$\rho_j$	$(\lambda_j - \rho_j)/\lambda_j$
5	1.066426958541e+04	-1.15e-08
6	1.546422448367e+04	-4.66e-12
7	2.528083471552e+04	2.76e-10
8	2.556227133088e+04	3.69e-08
9	2.697249607884e+04	1.42e-09
10	3.929970356637e+04	1.10e-05
12	3.976175735764e+04	-1.62e-07

Table 3: Eigenvalue bounds for equidistant shifts

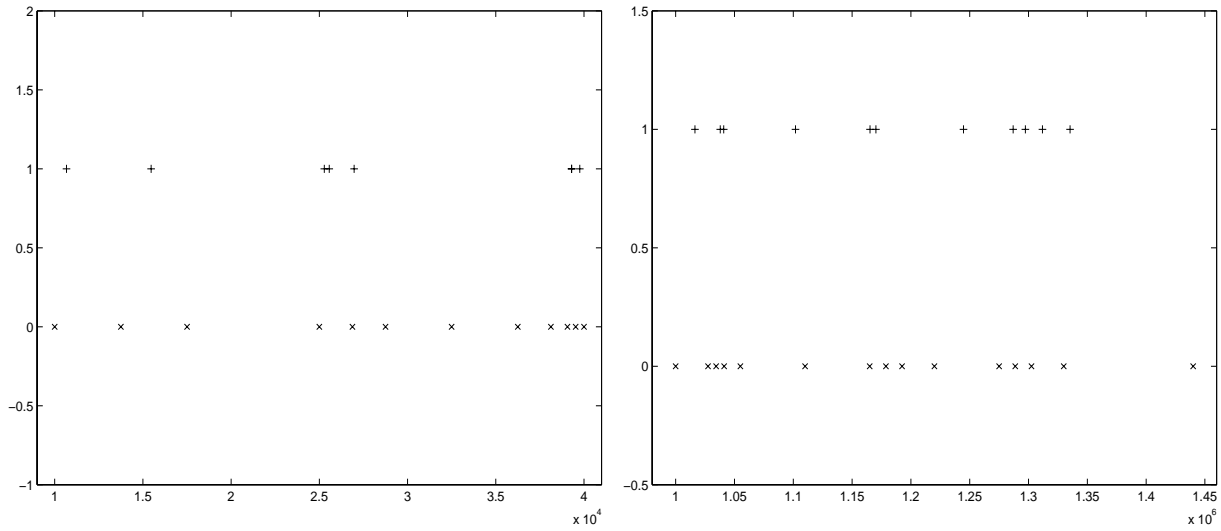


Figure 3: Distribution of eigenvalues (+) and shifts (x)

j	$\rho_j$	$\mu_j$	$\frac{ \rho_j - \mu_j }{\min(\rho_j, \mu_j)}$	$\frac{\lambda_j - \rho_j}{\lambda_j}$
61	1.016408372253e+06	1.016495643785e+06	8.59e-05	-4.24e-08
62	1.037914715458e+06	1.038057981099e+06	1.38e-04	-1.17e-07
63	1.040842551403e+06	1.040844460008e+06	1.83e-06	-2.09e-11
64	1.101893939733e+06	1.102674187896e+06	7.08e-04	-4.66e-06
65	1.165272879541e+06	1.165274042281e+06	9.98e-07	-2.08e-11
66	1.170236407647e+06	1.170243956172e+06	6.45e-06	-9.55e-10
67	1.244769630239e+06	1.244732428959e+06	2.99e-05	4.60e-08
68	1.286935395939e+06	1.286918571970e+06	1.31e-05	3.54e-09
69	1.297359894446e+06	1.297280179928e+06	6.14e-05	6.96e-08
70	1.311980121978e+06	1.309182467993e+06	2.14e-03	7.71e-05
71	1.335443602731e+06	1.332549286114e+06	2.17e-03	6.62e-05

 Table 4: Bounds of eigenvalues in  $[1e6, 1.44E6]$

it misses one eigenvalue close to  $\lambda_{62}$ .

## 6 Treating multiple eigenvalues and other refinements

### 6.1 Iterative Refinement of Ritz Pairs

As soon as in our numerical experiments the adaptive bisection process was able to resolve the eigenvalues (i.e. as many approximates were detected as there were eigenvalues) the quality of the approximates has found to be quite reasonable (cf. the Tables 1, 2 and 4, e.g.).

If these approximates still have to be improved we recommend the following

#### Iterative Refinement

With the determined Ritz vectors  $\hat{x}_i = Vy_i$  put

$$r_{new} := \sum_i \hat{x}_i, \tag{29}$$

use as shift set the values  $a$  and  $b$  together with the found eigenvalue approximations and redo the whole approximation process.

#### Remarks:

1. The iteration step is similar to a modified block Newton step (14) from [9] except that the individual eigenvector approximations bring their influence to bear through their sum (29). This makes it possible to use the efficient calculation of the projected matrices.
2. In a reanalysis situation [7] where eigenvectors  $x_i^{(s)}$ ,  $i = 1, \dots, q$  of one or several similar problems are known, one will certainly want to integrate this preinformation into the starting vector  $r$ , possibly analogously to (29) as a weighted sum

$$r := \sum_{i=1}^q w_i x_i^{(s)}.$$

### 6.2 Squeezing out $LDL^T$ information

While computing  $(A - s_i I)^{-1}r$  together with the inertia of  $(A - s_i I)$  an  $LDL^T$ -type factorization of  $(A - s_i I)$  or of  $(K - s_i M)$  to realize

$$L^T(K - s_i M)^{-1}L = (A - s_i I)^{-1}$$

will normally be determined. This allows to compute  $(A - s_i I)^{-1}r_+$  for further vectors  $r_+$  at comparably low expenses.

One would exploit this for example in the computation of the projection matrices (Section 3) if the Ericsson-Ruhe shift  $\sigma$  is taken from the  $s$ -shifts (as is the case in our numerical examples). In this case for  $\sigma = s_i$  the computation of

$$v_i^T(A - \sigma I)^{-1}v_i = v_i^T(A - \sigma I)^{-1}(A - \sigma I)^{-1}r$$

needs the computation of the additional vector  $(A - s_i I)^{-1}v_i$ . Using the factorization this is done inexpensively.

From this observation one obviously might want to further exploit the factorizations in the projection process. One idea to do so would be – as in in shifted Lanczos [3] or other rational Krylov approaches [15] – to incorporate several inverse iteration vectors

$$(A - s_i I)^{-1}r, (A - s_i I)^{-2}r, \dots, (A - s_i I)^{-k_i}r,$$

for each  $s_i$  with suitable multiplicities  $k_i$ . The efficient computation of the projected matrices from Section 3 is easily adapted to this setting.

First numerical experiments have been promising. The variation of the approach is under further investigation.

### 6.3 Multiple Eigenvalues

If a multiple eigenvalue of (1) belongs to the interval  $[a, b]$  the projection to a Krylov space has no chance to detect this and it will not meet the requirements (27) and (28). The projection method will only be able to detect a one dimensional representation of the higher dimensional eigenspace which is determined by the projection of  $r$  to this eigenspace.

In order to resolve higher dimensional eigenspaces one has to produce  $v$ -Vectors based on several ansatz-vectors  $r_1, \dots, r_k$  where  $k$  has to be as large as the dimensions of the concerned eigenspaces.

The efficient computation of the projected matrices in Section 3 can be adapted to this situation where the additional computation of  $(A - \sigma I)^{-1}r$  has to be extended to the computation of the similar vectors  $(A - \sigma I)^{-1}r_i$  which is again inexpensively performed using the factorization of  $(A - \sigma I)$ .

Within the iterative refinement process from 6.1 the redefinition of  $r$  in (29) has to be expanded to a redefinition of the  $k$  vectors  $r_1, \dots, r_k$ . This has to be in such a way that the projections of these vectors to all the eigenspaces span these spaces. Strategies to secure this property are being tested. Methods to identify multiplicities during the projection through the use of subspace-iteration type random guard-vectors are under investigation.

### 6.4 $O(n)$ -Computation of Error-Estimates

As a final remark we notice that the computation of  $(A - \sigma I)^{-1}\hat{x}_i$  with the Ritz vector  $\hat{x}_i$  for the Krylov-Bogoliubov-Kahan error estimates can be implemented as an  $O(n)$ -operation

procedure. The Ritz vector  $\hat{x}_i$  is a combination  $\hat{x}_i = Vy_i$  of the  $v$ -vectors  $v_j = (A - s_j I)^{-1}$  and hence

$$(A - \sigma I)^{-1} \hat{x}_i = \sum_j y_i^{(j)} (A - \sigma I)^{-1} v_j$$

such that – due to (22) –

$$(A - \sigma I)^{-1} \hat{x}_i = \sum_j y_i^{(j)} (\rho_i - s_j)^{-1} (v_\sigma - v_j).$$

## 7 Conclusions

We presented a transfer of Ritz approximation methods with the Krylov-Bogoliubov-Kahan based error estimation from [6] for the extremal eigenvalues of symmetric generalized eigenvalue problems to the approximate computation of all eigenvectors in a prescribed interval in the interior of the spectrum.

After a shift-invert strategy which transforms the relevant eigenvalues into extremal ones the problem is projected onto a rational Krylov space.

The method proved to be promising for the case of single eigenvalues. Indications on how to apply the method in the presence of multiple eigenvalues are given. Further improvements like an iterative refinement approach are added.

## References

- [1] J. Baglama, D. Calvetti, and L. Reichel: Iterative methods for the computation of a few eigenvalues of a large symmetric matrix. *BIT* **36** (1996) 400-421
- [2] J. Baglama, D. Calvetti, and L. Reichel: Fast Leja points. *Electronic Transactions on Numerical Analysis* **7** (1998) 124-140
- [3] T. Ericsson and A. Ruhe: The spectral transformation Lanczos method for the numerical solution of large sparse generalized symmetric eigenvalue problems. *Mathematics of Computation* **35** (1980) 1251 - 1268
- [4] M. A. Gockel (ed.): *MSC/NASTRAN. Handbook of Dynamic Analysis*. The MacNeal-Schwendler Corp., Los Angeles 1983
- [5] G. H. Golub, and C. F. Van Loan, *Matrix Computations*. The John Hopkins University Press, Baltimore, Maryland, 3rd edition, 1996
- [6] T. Hitziger, W. Mackens, and H. Voss: A condensation-projection method for generalized eigenvalue problems. pp 239 - 282 in H. Power, and C. A. Brebbia (eds): *High Performance Computing in Engineering 1*, Computational Mechanics Publications, Southampton, 1995

- [7] B. Hofferek: Reanalysis approaches in eigenvalue approximations, Report, Section of Mathematics, Technical University of Hamburg-Harburg, in preparation
- [8] R. B. Lehoucq and D. C. Sorensen: Deflation techniques for an implicitly restarted Arnoldi iteration. *SIAM J. Matrix Anal. Appl.* **17** (1996) 789 - 821
- [9] R. Lösche, H. Schwetlick, and G. Timmermann: A modified block Newton iteration for approximating an invariant subspace of a symmetric matrix. Report IOKOMO-o3-96, Dept. of Mathematics, Technical University of Dresden 1996, to appear in *Linear Algebra Appl.*
- [10] W. Mackens, and H. Voss: A projection method for computing the minimum eigenvalue of a symmetric positive definite Toeplitz matrix. *Linear Algebra and its Applications* **275-276** (1998) 401 - 415
- [11] W. Mackens, and H. Voss: *Non nodal condensation of eigenvalue problems.*, *ZAMM* **79** (1999) 4, 243 - 255
- [12] Matrix Market, <http://math.nist.gov/MatrixMarket>
- [13] M. Papadopoulos, and E. Garcia: Improving in model reduction schemes using the system equivalent reduction expansion process. *AIAA J.* **34** (1996), 2217 - 2219
- [14] B. N. Parlett: *The Symmetric Eigenvalue Problem*. Classics in Applied Mathematics, SIAM, Philadelphia 1998
- [15] A. Ruhe: Rational Krylov: A practical algorithm for large sparse nonsymmetric matrix pencils. *SIAM J. Sci. Comput.* **19** (1998) 1535 - 1551
- [16] Y. Saad: *Numerical Methods for Large Eigenvalue Problems*. Manchester University Press, Manchester 1992
- [17] D. S. Scott: The advantages of inverted operators in Rayleigh-Ritz approximations. *SIAM J. Sci. Stat. Comput.* **3** (1982)