THE EXTENDED KRYLOV SUBSPACE METHOD AND ORTHOGONAL LAURENT POLYNOMIALS

CARL JAGELS* AND LOTHAR REICHEL†

Dedicated to Henk van der Vorst on the occasion of his 65th birthday.

Abstract. The need to evaluate expressions of the form $f(A)\boldsymbol{v}$, where A is a large sparse or structured symmetric matrix, $\boldsymbol{v}$ is a vector, and f is a nonlinear function, arises in many applications. The extended Krylov subspace method can be an attractive scheme for computing approximations of such expressions. This method projects the approximation problem onto an extended Krylov subspace $\mathbb{K}^{\ell,m}(A) = \text{span}\{A^{-\ell+1}\boldsymbol{v}, \ldots, A^{-1}\boldsymbol{v}, \boldsymbol{v}, A\boldsymbol{v}, \ldots, A^{m-1}\boldsymbol{v}\}$ of fairly small dimension, and then solves the small approximation problem so obtained. We review available results for the extended Krylov subspace method and relate them to properties of Laurent polynomials. The structure of the projected problem receives particular attention. We are concerned with the situations when $m = \ell$ and $m = 2\ell$.

1. Introduction. Let $A \in \mathbb{R}^{n\times n}$ be a large, possibly sparse or structured, symmetric matrix, and let $\boldsymbol{v} \in \mathbb{R}^n$. We are interested in computing approximations of expressions of the form

$$\boldsymbol{w} := f(A)\boldsymbol{v}, \tag{1.1}$$

where f is a nonlinear function defined on the spectrum $\{\lambda_j\}_{j=1}^n$ of A. The matrix $f(A)$ can be determined via the spectral factorization,

$$A = U\Lambda U^T, \quad U \in \mathbb{R}^{n\times n}, \quad U^TU = I_n, \quad \Lambda = \text{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n] \in \mathbb{R}^{n\times n}, \tag{1.2}$$

where I_n denotes the $n \times n$ identity matrix. Then

$$f(A) = Uf(\Lambda)U^T, \qquad f(\Lambda) = \text{diag}[f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)].$$

Functions of interest in applications include

$$f(t) := \exp(t), \qquad f(t) := \sqrt{t}, \qquad f(t) := \ln(t).$$

A recent thorough discussion on the evaluation of $f(A)$, as well as of (1.1), is provided by Higham [13]. Applications and numerical methods also are described in, e.g., [1, 2, 5, 7, 8, 9, 10, 14, 23]. An early discussion on the approximation of large-scale expressions of the form (1.1) is presented by van der Vorst [26]; see also [27, Chapter 11].

For small matrices A, one can evaluate expressions of the form (1.1) by first computing the spectral factorization (1.2), then evaluating $f(A)$ by using this factorization, and finally multiplying $f(A)$ by the vector $\boldsymbol{v}$. When f is rational and A is symmetric positive definite, it may be attractive to use the Cholesky factorization of A instead of the spectral factorization.

The computation of the spectral factorization of A is not attractive when this matrix is large and sparse. The present paper is concerned with this situation. Then one typically first reduces A to a small symmetric matrix T_m and evaluates $f(T_m)$,

*Department of Mathematics and Computer Science, Hanover College, Hanover, IN 47243, USA. E-mail: jagels@hanover.edu.

†Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA. E-mail: reichel@math.kent.edu.

e.g., by determining the spectral or Cholesky factorizations of T_m. For instance, m steps of the Lanczos process applied to A with initial vector $\boldsymbol{v}$ yields the decomposition

$$AV_m = V_m T_m + \boldsymbol{g}_m \boldsymbol{e}_m^T, \tag{1.3}$$

where $V_m = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m] \in \mathbb{R}^{n\times m}$, $V_m^T V_m = I_m$, $\boldsymbol{v}_1 = \boldsymbol{v}/\|\boldsymbol{v}\|$, $T_m := V_m^T A V_m \in \mathbb{R}^{m\times m}$ is symmetric and tridiagonal, $\boldsymbol{g}_m \in \mathbb{R}^n$, and $V_m^T \boldsymbol{g}_m = \boldsymbol{0}$. Here and below $\boldsymbol{e}_j = [0, \ldots, 0, 1, 0, \ldots, 0]^T$ denotes the jth axis vector and $\|\cdot\|$ the Euclidean vector norm. We tacitly assume that m is chosen small enough so that a decomposition of the form (1.3) exists. The columns of V_m form an orthonormal basis for the Krylov subspace

$$\mathbb{K}^m(A, \boldsymbol{v}) = \mathrm{span}\{\boldsymbol{v}, A\boldsymbol{v}, \ldots, A^{m-1}\boldsymbol{v}\}. \tag{1.4}$$

The expression (1.1) now can be approximated by

$$\boldsymbol{w}_m := V_m f(T_m)\boldsymbol{e}_1 \|\boldsymbol{v}\|; \tag{1.5}$$

see, e.g., [4, 10, 14, 21] for discussions on this approach. Indeed, if $\boldsymbol{g}_m = \boldsymbol{0}$, then $\boldsymbol{w}_m = \boldsymbol{w}$. Moreover, let $\mathbb{P}_{m-1}$ denote the set of all polynomials of degree at most $m-1$. Then $f \in \mathbb{P}_{m-1}$ implies that $\boldsymbol{w}_m = \boldsymbol{w}$; see, e.g., [10] or [22, Proposition 6.3].

The decomposition (1.3) and the fact that $\mathrm{range}(V_m) = \mathbb{K}^m(A, \boldsymbol{v})$ show that:

i) The columns $\boldsymbol{v}_j$ of V_m satisfy a three-term recurrence relation. This follows from the fact that T_m is tridiagonal. The vectors $\boldsymbol{v}_j$ therefore are quite inexpensive to compute; only one matrix vector-product evaluation with A and a few vector operations are required to compute $\boldsymbol{v}_{j+1}$ from $\boldsymbol{v}_j$ and $\boldsymbol{v}_{j-1}$.

ii) The columns $\boldsymbol{v}_j$ can be expressed as

$$\boldsymbol{v}_j = p_{j-1}(A)\boldsymbol{v}, \qquad j = 1, 2, \ldots, m, \tag{1.6}$$

for certain polynomials $p_{j-1} \in \mathbb{P}_{j-1}$. This property shows that the right-hand side of (1.5) is of the form $p(A)\boldsymbol{v}$, where $p \in \mathbb{P}_{m-1}$.

iii) The polynomials $p_0, p_1, \ldots, p_{m-1}$ are orthogonal with respect to the inner product

$$(q, r) := (q(A)\boldsymbol{v})^T (r(A)\boldsymbol{v}) = \boldsymbol{v}^T U q(\Lambda) r(\Lambda) U^T \boldsymbol{v} = \sum_{j=1}^{n} q(\lambda_j) r(\lambda_j) \omega_j^2, \tag{1.7}$$

with $U^T \boldsymbol{v} = [\omega_1, \omega_2, \ldots, \omega_n]^T$, which is defined for $q, r \in \mathbb{P}_d$, where d is the number of distinct eigenvalues of A. The property

$$(xq, r) = (q, xr) \tag{1.8}$$

secures that the orthogonal polynomials p_j satisfy a three-term recurrence relation. Hence, the three-term recurrence relation for the vectors $\boldsymbol{v}_j$ is a consequence of the fact that polynomials orthogonal with respect to an inner product defined by a nonnegative measure on the real axis satisfy such a recursion.

It follows from ii) that if f cannot be approximated accurately by a polynomial of degree $m-1$ on the spectrum of A, then, generally, the expression (1.5) will be a poor approximation of (1.1). For this reason Druskin and Knizhnerman [11] proposed the Extended Krylov Subspace (EKS) method, which allows for the approximation of f by a rational function with a fixed pole, say at the origin.

Let A be nonsingular and consider the extended Krylov subspace

$$\mathbb{K}^{\ell,m}(A,\boldsymbol{v}) = \text{span}\{A^{-\ell+1}\boldsymbol{v},\ldots,A^{-1}\boldsymbol{v},\boldsymbol{v},A\boldsymbol{v},\ldots,A^{m-1}\boldsymbol{v}\}. \tag{1.9}$$

Thus, $\mathbb{K}^{1,m}(A,\boldsymbol{v}) = \mathbb{K}^{m}(A,\boldsymbol{v})$. Druskin and Knizhnerman [11] showed that projecting the problem (1.1) onto the subspace (1.9), instead of onto (1.4), can be attractive for many functions f. An algorithm for computing such approximations also is presented in [11]. This algorithm first determines an orthonormal basis $\{\boldsymbol{q}_j\}_{j=1}^{\ell}$ for $\mathbb{K}^{\ell,1}(A,\boldsymbol{v})$. Since $\mathbb{K}^{\ell,1}(A,\boldsymbol{v}) = \mathbb{K}^{\ell}(A^{-1},\boldsymbol{v})$, this basis can be generated by the Lanczos process applied to A^{-1} with initial vector $\boldsymbol{v}$. In particular, a three-term recursion formula can be used; see i) above. Subsequently this basis is augmented to yield an orthonormal basis $\{\boldsymbol{q}_j\}_{j=1}^{\ell+m-1}$ of $\mathbb{K}^{\ell,m}(A,\boldsymbol{v})$. The augmentation also allows the use of a three-term recursion relation. A shortcoming of this algorithm for the EKS method is that the parameter ℓ has to be prespecified; the scheme does not allow for efficient computation of an orthonormal basis for $\mathbb{K}^{\ell+1,m}(A,\boldsymbol{v})$ from an available orthonormal basis for $\mathbb{K}^{\ell,m}(A,\boldsymbol{v})$.

Recently, Simoncini [24] described an approach to generating orthonormal bases for the sequence of nested spaces

$$\mathbb{K}^{1,1}(A,\boldsymbol{v}) \subset \mathbb{K}^{2,2}(A,\boldsymbol{v}) \subset \ldots \subset \mathbb{K}^{m,m}(A,\boldsymbol{v}) \subset \ldots \subset \mathbb{R}^n. \tag{1.10}$$

The derivation uses numerical linear algebra techniques and reveals the existence of short recursion formulas for the orthonormal basis $\{\boldsymbol{q}_j\}_{j=1}^{2m-1}$ of $\mathbb{K}^{m,m}(A,\boldsymbol{v})$ when A is symmetric. These recursions are applied to determine bases for the nested spaces (1.10). Simoncini [24] also discusses the situation when A is a general square nonsingular matrix, but then there are no short recursion formulas, and describes an application to the solution of Lyapunov equations. Knizhnerman and Simoncini [17] apply the method in [24] to the approximation of expressions (1.1) and improve the error analysis in [11].

The present paper explores the connection between the EKS method and Laurent polynomials. The short recursion relations for the orthonormal basis $\{\boldsymbol{q}_j\}_{j=1}^{2m-1}$ of $\mathbb{K}^{m,m}(A,\boldsymbol{v})$ is a consequence of the short recursion relations for orthogonal Laurent polynomials. The latter recursions were first derived by Njåstad and Thron [18], and are reviewed by Jones and Njåstad [15]. We are particularly interested in the structure of the projected problem. Short recursion formulas for orthonormal bases for the nested Krylov subspaces

$$\mathbb{K}^{1,2}(A,\boldsymbol{v}) \subset \mathbb{K}^{2,4}(A,\boldsymbol{v}) \subset \ldots \subset \mathbb{K}^{m,2m}(A,\boldsymbol{v}) \subset \mathbb{R}^n \tag{1.11}$$

also are presented. These spaces are of interest when the evaluation of $A^{-1}\boldsymbol{w}$ for vectors $\boldsymbol{w} \in \mathbb{R}^n$ is significantly more cumbersome than the computation of $A\boldsymbol{w}$.

This paper is organized as follows. Section 2 discusses the situation when A is symmetric positive definite and determines the structure of the analog of the symmetric tridiagonal matrix T_m in the Lanczos decomposition (1.3) from the recursion formulas for Laurent polynomials. We also investigate the structure of the inverse of this matrix. Section 3 is concerned with symmetric indefinite matrices A. While we in Section 2 obtain pairs of three-term recursion formulas for the Laurent polynomials, the indefiniteness of A makes it necessary to use a five-term recursion formula in some instances. Recursion formulas for an orthonormal basis for extended Krylov subspaces of the form (1.11) are discussed in Section 4, and a few computed examples are presented in Section 5. Concluding remarks can be found in Section 6.

Error bounds for the computed rational approximants are derived in [4, 12, 17]. Many results on orthogonal rational functions can be found in [6]. The possibly first application of rational Krylov subspaces reported in the literature is to eigenvalue problems; see Ruhe [19, 20]. The extended Krylov subspace method of the present paper also can be applied in this context.

2. The positive definite case, $m = \ell$. We assume in this section that A is symmetric and positive definite. Let the Laurent polynomials $\phi_0, \phi_1, \phi_{-1}, \phi_2, \phi_{-2}, \dots$ of the form

$$\phi_j(x) := \begin{cases} x^j + \sum_{k=-j+1}^{j-1} c_{j,k} x^k, & j = 0, 1, \dots\ , \\ x^j + \sum_{k=j+1}^{-j} c_{j,k} x^k, & j = -1, -2, \dots\ , \end{cases} \tag{2.1}$$

be orthogonal with respect to the inner product (1.7). We refer to these Laurent polynomials as monic, because their leading coefficient is unity. The coefficients $c_{j,-j+1}$ of ϕ_j with $j \geq 1$, and $c_{j,-j}$ of ϕ_j with $j \leq 1$, are said to be trailing. Many properties of orthogonal Laurent polynomials are established in [15, 16, 18]. In particular, Njåstad and Thron [18] show that Laurent polynomials that are orthogonal with respect to a nonnegative measure on the real axis satisfy recursion relations with few terms. We will use these recursions in the present paper.

Introduce, analogously to (1.6), the vectors

$$\boldsymbol{v}_j := \frac{\phi_j(A)\boldsymbol{v}}{\|\phi_j(A)\boldsymbol{v}\|}, \qquad j = 0, 1, -1, 2, -2, \dots\ .$$

Due to the orthogonality of the ϕ_j with respect to the inner product (1.7), the vectors $\{\boldsymbol{v}_j\}_{j=-m+1}^{m}$ form an orthonormal basis for the extended Krylov subspace $\mathbb{K}^{m,m+1}(A, \boldsymbol{v})$. Analogously to the matrix V_m in the Lanczos decomposition (1.3), we define the matrices

$$\begin{aligned} V_{2m-1} &= [\boldsymbol{v}_0, \boldsymbol{v}_1, \boldsymbol{v}_{-1}, \boldsymbol{v}_2, \dots, \boldsymbol{v}_{m-1}, \boldsymbol{v}_{-m+1}] \in \mathbb{R}^{n\times(2m-1)}, \\ V_{2m} &= [V_{2m-1}, \boldsymbol{v}_m] \in \mathbb{R}^{n\times(2m)}. \end{aligned} \tag{2.2}$$

We are interested in the structure of the matrices

$$H_{2m-1} := V_{2m-1}^T A V_{2m-1} \tag{2.3}$$

and

$$G_{2m} := V_{2m}^T A^{-1} V_{2m}, \tag{2.4}$$

which are analogs of the symmetric tridiagonal matrix T_m in (1.3). The structure of H_{2m-1} and G_{2m} is a consequence of the recursion relations for the orthogonal Laurent polynomials ϕ_j. Simoncini [24] investigated the structure of H_{2m-1} by other means.

In order to expose the structure of H_{2m-1} and G_{2m}, we derive certain properties of orthogonal Laurent polynomials ϕ_j. The derivations allow us to introduce suitable notation and make the paper self-contained. For other proofs and related results, we refer to [15, 18, 24]. The following property of the trailing coefficients of the ϕ_j is required in our derivation of three-term recursion formulas for the vectors $\boldsymbol{v}_j$.

PROPOSITION 2.1. *Let the matrix A be definite. Then the coefficients $c_{j,-j+1}$ of ϕ_j, for $1 \le j \le m$, and the coefficients $c_{j,-j}$ of ϕ_j, for $-m+1 \le j \le -1$, are nonvanishing.*

Proof. We first show that $c_{j,-j+1} \neq 0$ for $j \ge 1$. Consider the Laurent polynomial $x^{-1}\phi_j(x)$, $j \ge 1$. By the definition of the inner product (1.7) and the definiteness of A, we have

$$(\phi_j, x^{-1}\phi_j) = \boldsymbol{w}_j^T A^{-1} \boldsymbol{w}_j \neq 0.$$

On the other hand,

$$(\phi_j, x^{-1}\phi_j) = (\phi_j, c_{j,-j+1}x^{-j} + \psi),$$

where ψ is a Laurent polynomial in $\mathrm{span}\{\phi_0, \phi_1, \phi_{-1}, \ldots, \phi_{-j+1}\}$. Hence,

$$(\phi_j, x^{-1}\phi_j) = c_{j,-j+1}(\phi_j, x^{-j}),$$

and therefore $c_{j,-j+1} \neq 0$.

The fact that the coefficients $c_{j,-j}$ are nonvanishing for $j \le -1$ follows similarly by considering $(\phi_j, x\phi_j) = \boldsymbol{w}_j^T A \boldsymbol{w}_j \neq 0$. □

Njåstad and Thron [18] refer to orthogonal Laurent polynomials with nonvanishing trailing coefficients as nonsingular, and show that their finite zeros are real and simple. Moreover, successive nonsingular Laurent polynomials have no common zeros; see also [15, 16] for related results.

Let $m > 0$ and suppose that $A^{-m}\boldsymbol{v} \notin \mathbb{K}^{m,m+1}(A, \boldsymbol{v})$. We would like to determine a vector $\boldsymbol{v}_{-m}$, such that

$$\{\boldsymbol{v}_0, \boldsymbol{v}_1, \boldsymbol{v}_{-1}, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{-m+1}, \boldsymbol{v}_m, \boldsymbol{v}_{-m}\}$$

is an orthonormal basis for $\mathbb{K}^{m+1,m+1}(A, \boldsymbol{v})$. The vector $\boldsymbol{v}_{-m}$ will be a multiple of $\phi_{-m}(A)\boldsymbol{v}$, where ϕ_{-m} is a Laurent polynomial of the form (2.1). In particular,

$$c_{m,-m+1}\phi_{-m}(x) - x^{-1}\phi_m(x) \in \mathrm{span}\{\phi_0, \phi_1, \phi_{-1}, \ldots, \phi_{-m+1}, \phi_m\} \tag{2.5}$$

and, therefore,

$$c_{m,-m+1}\phi_{-m}(x) - x^{-1}\phi_m(x) = -\sum_{k=-m+1}^{m} \gamma_{m,k}\phi_k(x),$$

where the Fourier coefficients are given by

$$\gamma_{m,k} = \frac{(x^{-1}\phi_m, \phi_k)}{(\phi_k, \phi_k)} = \frac{(\phi_m, x^{-1}\phi_k)}{(\phi_k, \phi_k)}. \tag{2.6}$$

Moreover, since

$$x^{-1}\phi_k(x) \in \mathrm{span}\{\phi_0, \phi_1, \phi_{-1}, \ldots, \phi_{m-1}, \phi_{-m+1}\}, \qquad k = -m+2, \ldots, m-1,$$

it follows that at most two of the Fourier coefficients are nonvanishing. Thus, we obtain

$$c_{m,-m+1}\phi_{-m}(x) = x^{-1}\phi_m(x) - \gamma_{m,m}\phi_m(x) - \gamma_{m,-m+1}\phi_{-m+1}(x), \tag{2.7}$$

which yields the three-term recursion relation

$$\delta_{-m}\boldsymbol{v}_{-m} = (A^{-1} - \beta_m I_n)\boldsymbol{v}_m - \beta_{-m+1}\boldsymbol{v}_{-m+1} \tag{2.8}$$

with $\beta_m = \gamma_{m,m}$ and $\delta_{-m} > 0$ a normalization factor to make $\boldsymbol{v}_{-m}$ a unit vector.

A similar argument shows that $c_{-m,m}\phi_{m+1}(x) - x\phi_{-m}(x)$ is a linear combination of $\phi_{-m}(x)$ and $\phi_m(x)$ and this gives the three-term recursion formula

$$\delta_{m+1}\boldsymbol{v}_{m+1} = (A - \alpha_{-m}I_n)\boldsymbol{v}_{-m} - \alpha_m\boldsymbol{v}_m. \tag{2.9}$$

The recursion relations (2.8) and (2.9) are the foundation for the following algorithm for computing an orthonormal basis for $\mathbb{K}^{m,m+1}(A,\boldsymbol{v})$. The algorithm is analogous to the standard Lanczos process for determining an orthonormal basis for the Krylov subspace (1.4).

Algorithm 2.1 (Orthogonalization process).

Input: *m, $\boldsymbol{v}$, functions for evaluating matrix-vector products and solving linear systems of equations with A;*
Output: *orthogonal basis $\{\boldsymbol{v}_k\}_{k=-m+1}^{m}$ of $\mathbb{K}^{m,m+1}(A,\boldsymbol{v})$;*
$\delta_0 := ||\boldsymbol{v}||$; $\boldsymbol{v}_0 := \boldsymbol{v}/\delta_0$;
$\boldsymbol{u} := A\boldsymbol{v}_0$; $\alpha_0 := \boldsymbol{v}_0^T\boldsymbol{u}$; $\boldsymbol{u} := \boldsymbol{u} - \alpha_0\boldsymbol{v}_0$;
$\delta_1 := ||\boldsymbol{u}||$; $\boldsymbol{v}_1 := \boldsymbol{u}/\delta_1$;
for $k = 1, 2, \ldots, m-1$ **do**
  $\boldsymbol{w} := A^{-1}\boldsymbol{v}_k$;
  $\beta_{-k+1} := \boldsymbol{v}_{-k+1}^T\boldsymbol{w}$; $\boldsymbol{w} := \boldsymbol{w} - \beta_{-k+1}\boldsymbol{v}_{-k+1}$;
  $\beta_k := \boldsymbol{v}_k^T\boldsymbol{w}$; $\boldsymbol{w} := \boldsymbol{w} - \beta_k\boldsymbol{v}_k$;
  $\delta_{-k} := ||\boldsymbol{w}||$; $\boldsymbol{v}_{-k} := \boldsymbol{w}/\delta_{-k}$;
  $\boldsymbol{u} := A\boldsymbol{v}_{-k}$;
  $\alpha_{-k} := \boldsymbol{v}_{-k}^T\boldsymbol{u}$; $\boldsymbol{u} := \boldsymbol{u} - \alpha_{-k}\boldsymbol{v}_{-k}$;
  $\alpha_k := \boldsymbol{v}_k^T\boldsymbol{u}$; $\boldsymbol{u} := \boldsymbol{u} - \alpha_k\boldsymbol{v}_k$;
  $\delta_{k+1} := ||\boldsymbol{u}||$; $\boldsymbol{v}_{k+1} := \boldsymbol{u}/\delta_{k+1}$;
end

The recursion coefficients generated by Algorithm 2.1 can be used to construct a matrix $\hat{H}_{2m-1} = [h_{j,k}] \in \mathbb{R}^{2m\times(2m-1)}$, such that

$$AV_{2m-1} = V_{2m}\hat{H}_{2m-1}, \tag{2.10}$$

where the matrices V_{2m-1} and V_{2m} are given by (2.2). The leading submatrix $H_{2m-1} \in \mathbb{R}^{(2m-1)\times(2m-1)}$ of $\hat{H}_{2m-1}$ is given by (2.3). We will now show that H_{2m-1} is pentadiagonal. The $(2k+1)^{\text{st}}$ column of AV_{2m-1} is $A\boldsymbol{v}_{-k}$, and by relation (2.9) with m replaced by k, or by the recursion formulas of Algorithm 2.1, we obtain

$$A\boldsymbol{v}_{-k} = \alpha_k\boldsymbol{v}_k + \alpha_{-k}\boldsymbol{v}_{-k} + \delta_{k+1}\boldsymbol{v}_{k+1}, \qquad k = 1, 2, \ldots, m-1. \tag{2.11}$$

Hence, the only nontrivial entries of the $(2k+1)^{\text{st}}$ column of H_{2m-1} are

$$h_{2k,2k+1} = \alpha_k, \qquad h_{2k+1,2k+1} = \alpha_{-k}, \qquad h_{2k+2,2k+1} = \delta_{k+1}.$$

Symmetry of H_{2m-1} yields two entries of the $(2k)^{\text{th}}$ column,

$$h_{2k+1,2k} = \alpha_k, \qquad h_{2k-1,2k} = \delta_{k-1}.$$

In order to determine the remaining nonvanishing entries of this column, we first rewrite relation (2.8) with m replaced by k,

$$\delta_{-k}\boldsymbol{v}_{-k} = A^{-1}\boldsymbol{v}_k - \beta_{-k+1}\boldsymbol{v}_{-k+1} - \beta_k\boldsymbol{v}_k. \tag{2.12}$$

Multiplying the above equation by A and making the appropriate substitutions for $A\boldsymbol{v}_{-k}$ and $A\boldsymbol{v}_{-k+1}$ yields

$$\begin{aligned}\beta_k A\boldsymbol{v}_k = & -\beta_{-k+1}\alpha_{k-1}\boldsymbol{v}_{k-1} - \beta_{-k+1}\alpha_{-k+1}\boldsymbol{v}_{-k+1}\\ & +(1-\beta_{-k+1}\delta_k - \alpha_k\delta_{-k})\boldsymbol{v}_k - \delta_{-k}\alpha_{-k}\boldsymbol{v}_{-k} - \delta_{-k}\delta_{k+1}\boldsymbol{v}_{k+1}.\end{aligned}$$

It follows from (2.12) that $\beta_k = \boldsymbol{v}_k^T A^{-1}\boldsymbol{v}_k$, and by the definiteness of A, we have $\beta_k \neq 0$. Hence,

$$\begin{aligned}A\boldsymbol{v}_k = & \, h_{2k-2,2k}\boldsymbol{v}_{k-1} + h_{2k-1,2k}\boldsymbol{v}_{-k+1} + h_{2k,2k}\boldsymbol{v}_k\\ & +h_{2k+1,2k}\boldsymbol{v}_{-k} + h_{2k+2,2k}\boldsymbol{v}_{k+1}\end{aligned} \tag{2.13}$$

for certain coefficients $h_{j,2k}$. Orthonormality of the vectors $\boldsymbol{v}_j$ and symmetry of A and H_{2m-1} now give

$$h_{2k-2,2k} = h_{2k,2k-2} = -\frac{\delta_{-k+1}\delta_k}{\beta_{k-1}}, \qquad h_{2k,2k} = \frac{1-\beta_{-k+1}\delta_k - \alpha_k\delta_{-k}}{\beta_k}.$$

Consequently, the odd-numbered columns of H_{2m-1} have at most three nontrivial elements and the even numbered columns contain at most five nonvanishing entries.

Example 2.1. The matrix H_{2m-1} is of the form

$$\begin{bmatrix}
\alpha_0 & \delta_1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0\\
\delta_1 & h_{2,2} & \alpha_1 & -\frac{\delta_{-1}\delta_2}{\beta_1} & 0 & 0 & 0 & 0 & \cdots & 0\\
0 & \alpha_1 & \alpha_{-1} & \delta_2 & 0 & 0 & 0 & 0 & \cdots & 0\\
0 & -\frac{\delta_{-1}\delta_2}{\beta_1} & \delta_2 & h_{4,4} & \alpha_2 & -\frac{\delta_{-2}\delta_3}{\beta_2} & 0 & 0 & \cdots & 0\\
0 & 0 & 0 & \alpha_2 & \alpha_{-2} & \delta_3 & 0 & 0 & \cdots & 0\\
0 & 0 & 0 & -\frac{\delta_{-2}\delta_3}{\beta_2} & \delta_3 & h_{6,6} & \alpha_3 & -\frac{\delta_{-3}\delta_4}{\beta_3} & \cdots & 0\\
0 & 0 & 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \cdots & 0\\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots\\
0 & 0 & 0 & 0 & \cdots & 0 & \star & \delta_{m-1} & * & \alpha_{m-1}\\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \alpha_{m-1} & \alpha_{-m+1}
\end{bmatrix},$$

where the entries marked by $\star$ and $*$ are $h_{2m-2,2m-4}$ and $h_{2m-2,2m-2}$, respectively. □

The matrix $\hat{H}_{2m-1}$ in (2.10) is given by

$$\hat{H}_{2m-1} = \begin{bmatrix} H_{2m-1} \\ \boldsymbol{h}_{2m-1}^T \end{bmatrix}$$

with

$$\boldsymbol{h}_{2m-1} = -\frac{\delta_{-m+1}\delta_m}{\beta_{m-1}}\boldsymbol{e}_{2m-2} + \delta_m \boldsymbol{e}_{2m-1} \in \mathbb{R}^{2m-1},$$

and we can write (2.10) in the form

$$AV_{2m-1} = V_{2m-1}H_{2m-1} + \boldsymbol{v}_m \boldsymbol{h}_{2m-1}^T. \tag{2.14}$$

This expression is analogous to the decomposition (1.3) obtained by the standard Lanczos process. Note that each leading principal submatrix of H_{2m-1} of even order is block-tridiagonal with block-size two and the matrix $\boldsymbol{v}_m \boldsymbol{h}_{2m-1}^T$ generically has two nonvanishing columns. Thus, our Lanczos-like process bears some similarity to the standard block Lanczos process with block-size two.

We also can use the recursion relations (2.11) and (2.12) to derive a decomposition of the form

$$A^{-1}V_{2m} = V_{2m+1}\hat{G}_{2m} \tag{2.15}$$

for some matrix $\hat{G}_{2m} = [g_{j,k}] \in \mathbb{R}^{(2m+1)\times(2m)}$. We remark that the matrix $\hat{G}_{2m}$ has to have an even number of columns in order to accommodate the fact that $A^{-1}\boldsymbol{v}_{-k}$ is expressed as a linear combination of five orthogonal vectors. The decomposition (2.15) is analogous to (2.10).

The first $2m$ rows of $\hat{G}_{2m}$ make up the matrix G_{2m} given by (2.4). Arguing similarly as for H_{2m-1}, the nontrivial elements of the $(2k)^{\text{th}}$ column of G_{2m} are

$$g_{2k-1,2k} = \beta_{-k+1}, \qquad g_{2k,2k} = \beta_k, \qquad g_{2k+1,2k} = \delta_{-k},$$

and those of the $(2k-1)^{\text{st}}$ column are given by

$$\begin{aligned}
g_{2k-3,2k-1} &= -\frac{\delta_{k-1}\delta_{-k+1}}{\alpha_{-k+2}},\\
g_{2k-2,2k-1} &= \delta_{-k+1},\\
g_{2k-1,2k-1} &= \frac{1-\alpha_{k-1}\delta_{-k+1}-\beta_{-k+1}\delta_k}{\alpha_{-k+1}},\\
g_{2k,2k-1} &= \beta_{-k+1},\\
g_{2k+1,2k-1} &= -\frac{\delta_k\delta_{-k}}{\alpha_{-k+1}}.
\end{aligned}$$

Thus, the matrix G_{2m} is symmetric and pentadiagonal. Moreover, leading principal submatrices of even order are block-tridiagonal with block-size two.

The block-structure implies that the product of principal submatrices of H_{2m-1} and G_{2m} of (the same) even order is a rank-one modification of the identity. This property can be seen as follows. Assume for the moment that $2m-1 = n$ in (2.3) and $2m = n$ in (2.4). Then the matrix V_n in (2.3) and (2.4) is orthogonal, and we obtain that

$$H_nG_n = (V_n^TAV_n)(V_n^TA^{-1}V_n) = I_n.$$

Let $\tilde{H}_{2k}$ and $\tilde{G}_{2k}$ denote leading principal submatrices of order $2k$ of H_n and G_n, respectively. Due to the special form of the subdiagonal blocks, we have

$$\tilde{H}_{2k}\tilde{G}_{2k} = I_{2k} + \boldsymbol{e}_{2k}\boldsymbol{u}_{2k}^T, \tag{2.16}$$

where only the last two entries of $\boldsymbol{u}_{2k} \in \mathbb{R}^{2k}$ may be nonvanishing.

3. The indefinite case, $m = \ell$. In this section the nonsingular symmetric matrix A is not required to be definite. The derivation of the three-term recurrence formulas in Section 2 requires that the trailing coefficients of the Laurent polynomials ϕ_j be nonvanishing. This property followed from the definiteness of A. Now assume that, for some $k \geq 1$, the trailing coefficients of the Laurent polynomials $\phi_0, \phi_1, \phi_{-1}, \ldots, \phi_{-k+1}, \phi_k$ are nonvanishing, but that the trailing coefficient, $c_{-k,k}$, of ϕ_{-k} is zero. Thus,

$$\phi_{-k}(x) = x^{-k} + c_{-k,-k+1}x^{-k+1} + \ldots + c_{-k,k-1}x^{k-1}.$$

Njåstad and Thron [18] refer to orthogonal Laurent polynomials with vanishing trailing coefficient as singular, and show that two consecutive orthogonal Laurent polynomials cannot both be singular; see also [15, 16]. This result also follows from our discussion below.

Analogously to (2.7), we have

$$c_{k,-k+1}\phi_{-k}(x) = x^{-1}\phi_k(x) - \gamma_{k,k}\phi_k(x) - \gamma_{k,-k+1}\phi_{-k+1}(x), \tag{3.1}$$

where the coefficients $\gamma_{k,k}$ and $\gamma_{k,-k+1}$ are given by (2.6), and

$$\phi_k(x) = x^k + c_{k,k-1}x^{k-1} + \ldots + c_{k,-k+1}x^{-k+1}.$$

Comparing coefficients for the x^k-terms in the right-hand side and left-hand side of (3.1) shows that $\gamma_{k,k} = 0$. This is equivalent to $(x^{-1}\phi_k, \phi_k) = 0$; cf. (2.6).

Let $\psi \in \text{span}\{\phi_0, \phi_1, \phi_{-1}, \ldots, \phi_{k-1}, \phi_{-k+1}\}$. Then

$$0 = (x^{-1}\phi_k, \phi_k) = (\phi_k, x^{-k} + \psi) = (\phi_k, x^{-k})$$

and, therefore,

$$(\phi_k, x^{-1}\phi_{-k+1}) = (\phi_k, x^{-k}) = 0.$$

Since the left-hand side is proportional to $\gamma_{k,-k+1}$, cf. (2.6), it follows that $\gamma_{k,-k+1}$ vanishes. Thus, the recursion formula (3.1) simplifies to

$$c_{k,-k+1}\phi_{-k}(x) = x^{-1}\phi_k(x),$$

which, analogously to (2.8), yields

$$\delta_{-k}\boldsymbol{v}_{-k} = A^{-1}\boldsymbol{v}_k$$

or, equivalently,

$$A\boldsymbol{v}_{-k} = \frac{1}{\delta_{-k}}\boldsymbol{v}_k. \tag{3.2}$$

Thus, the only non-zero element of the $(2k+1)^{\text{st}}$ column of H_{2m-1} is $h_{2k,2k+1} = 1/\delta_{-k}$.

We turn to the recursion relation for ϕ_{k+1}. Since $c_{-k,k} = 0$, we must modify the technique used in Section 2. Instead of (2.5), we consider

$$\phi_{k+1}(x) - x\phi_k(x) \in \text{span}\{\phi_0, \phi_1, \phi_{-1}, \ldots, \phi_k, \phi_{-k}\}.$$

An argument similar to that of Section 2 shows that ϕ_{k+1} satisfies a five-term recursion formula

$$\begin{aligned}\phi_{k+1}(x) = x\phi_k(x) &- \gamma_{k+1,-k}\phi_{-k}(x) - \gamma_{k+1,k}\phi_k(x) \\ &- \gamma_{k+1,-k+1}\phi_{-k+1}(x) - \gamma_{k+1,k-1}\phi_{k-1}(x).\end{aligned} \tag{3.3}$$

This formula also is shown in [18]. It follows from (3.3) that the vector $\boldsymbol{v}_{k+1}$ satisfies a recursion relation of the form

$$\delta_{k+1}\boldsymbol{v}_{k+1} = A\boldsymbol{v}_k - a_{-k+1}\boldsymbol{v}_{-k+1} - a_{k-1}\boldsymbol{v}_{k-1} - a_{-k}\boldsymbol{v}_{-k} - a_k\boldsymbol{v}_k,$$

which we also express as

$$A\boldsymbol{v}_k = a_{k-1}\boldsymbol{v}_{k-1} + a_{-k+1}\boldsymbol{v}_{-k+1} + a_k\boldsymbol{v}_k + a_{-k}\boldsymbol{v}_{-k} + \delta_{k+1}\boldsymbol{v}_{k+1}. \tag{3.4}$$

The coefficients yield the entries of the $(2k)^{\text{th}}$ column of H_{2m-1} and are easy to determine from the expression above. Three of the coefficients have been evaluated previously, namely

$$a_{k-1} = h_{2k,2k-2}, \qquad a_{-k+1} = h_{2k,2k-1}, \qquad a_{-k} = h_{2k,2k+1},$$

and a_k is computed by means of an inner product,

$$a_k = h_{2k,2k} = \boldsymbol{v}_k^T A\boldsymbol{v}_k.$$

The recursion formulas of this section require that four n-vectors be retained in fast computer memory at any given time.

An examination of equation (3.3) reveals that the trailing coefficient is nonvanishing, and the next orthogonal vector, $\boldsymbol{v}_{-k-1}$, therefore can be computed by a three-term recursion formula analogous to (2.8).

Recall that $\beta_{k+1} = \boldsymbol{v}_{k+1}^T A^{-1}\boldsymbol{v}_{k+1}$. If $\beta_{k+1} \neq 0$, then the coefficients in the expansion

$$\begin{aligned} A\boldsymbol{v}_{k+1} = {} & h_{2k,2k+2}\boldsymbol{v}_k + h_{2k+1,2k+2}\boldsymbol{v}_{-k} + h_{2k+2,2k+2}\boldsymbol{v}_{k+1} \\ & + h_{2k+3,2k+2}\boldsymbol{v}_{-k-1} + h_{2k+4,2k+2}\boldsymbol{v}_{k+2} \end{aligned}$$

adhere to the same pattern as in the definite case, with the exceptions

$$h_{2k+1,2k+2} = 0, \qquad h_{2k+2,2k+2} = \frac{1 - \alpha_{k+1}\delta_{-k-1}}{\beta_{k+1}}.$$

These exceptions stem from (3.2). On the other hand, if β_{k+1} vanishes, then recursion formulas similar to those derived in the beginning of this section can be applied.

Example 3.1. When $\beta_2 = 0$, the matrix $\hat{H}_7 \in \mathbb{R}^{8\times 7}$ is given by

$$\hat{H}_7 = \begin{bmatrix} \alpha_0 & \delta_1 & 0 & 0 & 0 & 0 & 0 \\ \delta_1 & h_{2,2} & \alpha_1 & -\frac{\delta_{-1}\delta_2}{\beta_1} & 0 & 0 & 0 \\ 0 & \alpha_1 & \alpha_{-1} & \delta_2 & 0 & 0 & 0 \\ 0 & -\frac{\delta_{-1}\delta_2}{\beta_1} & \delta_2 & h_{4,4} & 1/\delta_{-2} & \delta_3 & 0 \\ 0 & 0 & 0 & 1/\delta_{-2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \delta_3 & 0 & h_{6,6} & \alpha_3 \\ 0 & 0 & 0 & 0 & 0 & \alpha_3 & \alpha_{-3} \\ 0 & 0 & 0 & 0 & 0 & -\frac{\delta_{-3}\delta_4}{\beta_3} & \delta_4 \end{bmatrix}.$$

□

Since A is indefinite, the coefficient $\alpha_{-k+1} = \boldsymbol{v}_{-k+1}^T A \boldsymbol{v}_{-k+1}$ may vanish. In this situation, we use arguments similar to those for the case when $\gamma_{k,k} = 0$ to obtain that $\alpha_{k-1} = 0$ and $\delta_k \boldsymbol{v}_k = A\boldsymbol{v}_{-k+1}$. The vector $\boldsymbol{v}_{-k}$ then is computed from the five-term formula

$$(3.5) \quad \delta_{-k}\boldsymbol{v}_{-k} = A^{-1}\boldsymbol{v}_{-k+1} + b_k\boldsymbol{v}_k + b_{-k+1}\boldsymbol{v}_{-k+1} + b_{k-1}\boldsymbol{v}_{k-1} + b_{-k+2}\boldsymbol{v}_{-k+2}.$$

Analogously to the case discussed above, three of the coefficients have been determined previously, namely

$$b_k = 1/\delta_k, \qquad b_{k-1} = \delta_{-k+1}, \qquad b_{-k+2} = -\frac{\beta_{-k+2}\delta_{k-1}}{\beta_{k-1}}.$$

The remaining coefficient is computed by evaluating

$$b_{-k+1} = \boldsymbol{v}_{-k+1}^T A^{-1} \boldsymbol{v}_{-k+1}.$$

Note that, since $b_k > 0$ in (3.5), the trailing coefficient of ϕ_{-k} is non-zero and the vector $\boldsymbol{v}_{k+1}$ can be computed by using the three-term recursion formula (2.11), similarly as in the definite case. An expression for $A\boldsymbol{v}_k$ analogous to that found in (2.13) can be derived by multiplying equation (3.5) by A, making the appropriate substitutions for $A\boldsymbol{v}_j$, $j = -k+2, k-1, -k+1, -k$, and gathering terms associated with the same power. The entries in the $(2k)^{\text{th}}$ column of $\hat{H}_{2m-1}$ follow the same pattern as that in the definite case with the exceptions,

$$h_{2k,2k} = -\delta_k(b_{-k+1}\delta_{k-1} + \alpha_k\delta_{-k} + \delta_{-k+1}h_{2k,2k-2}), \qquad h_{2k+2,2k} = -\delta_k\delta_{-k}\delta_{k+1}.$$

Example 3.2. When $\alpha_{-2} = 0$, the matrix $\hat{H}_7 \in \mathbb{R}^{8\times 7}$ is given by

$$\hat{H}_7 = \begin{bmatrix} \alpha_0 & \delta_1 & 0 & 0 & 0 & 0 & 0 \\ \delta_1 & h_{2,2} & \alpha_1 & -\frac{\delta_{-1}\delta_2}{\beta_1} & 0 & 0 & 0 \\ 0 & \alpha_1 & \alpha_{-1} & \delta_2 & 0 & 0 & 0 \\ 0 & -\frac{\delta_{-1}\delta_2}{\beta_1} & \delta_2 & h_{4,4} & 0 & -\frac{\delta_{-2}\delta_3}{\beta_2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \delta_3 & 0 \\ 0 & 0 & 0 & -\frac{\delta_{-2}\delta_3}{\beta_2} & \delta_3 & h_{6,6} & \alpha_3 \\ 0 & 0 & 0 & 0 & 0 & \alpha_3 & \alpha_{-3} \\ 0 & 0 & 0 & 0 & 0 & -\delta_3\delta_{-3}\delta_4 & \delta_4 \end{bmatrix}.$$

□

4. The positive definite case, $m = 2\ell$**.** We derive short recursion formulas for orthogonal Laurent polynomials for the Krylov subspaces (1.11) and investigate the structure of the reduced problems. The matrix A is assumed to be positive definite. We consider the generation of orthogonal basis vectors $\boldsymbol{v}_j$ in an order commensurate with the nesting (1.11). To this end, introduce monic orthogonal Laurent polynomials $\phi_0, \phi_1, \phi_2, \phi_{-1}, \phi_3, \phi_4, \phi_{-2}, \phi_5, \dots$ of the form

$$\phi_j(x) := \begin{cases} x^j + \displaystyle\sum_{k=-\lfloor (j-1)/2 \rfloor}^{j-1} c_{j,k} x^k, & j = 0, 1, 2, \dots\ , \\ x^j + \displaystyle\sum_{k=j+1}^{-2j} c_{j,k} x^k, & j = -1, -2, \dots\ , \end{cases} \tag{4.1}$$

where $\lfloor \alpha \rfloor$ denotes the integer part of $\alpha \geq 0$. In particular, $\phi_0(x) = 1$. These polynomials are orthogonal with respect to the inner product (1.7), similarly as the Laurent polynomials (2.5) used in Sections 2 and 3, but they are of different form.

Define the vectors

$$\boldsymbol{v}_j := \frac{\phi_j(A)\boldsymbol{v}}{\|\phi_j(A)\boldsymbol{v}\|}, \qquad j = 0, 1, 2, -1, 3, 4, -2, 5, \dots\ . \tag{4.2}$$

Then

$$\{\boldsymbol{v}_0, \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_{-1}, \boldsymbol{v}_3, \dots, \boldsymbol{v}_{-m+1}, \boldsymbol{v}_{2m-1}\}$$

is an orthonormal basis for the extended Krylov subspace $\mathbb{K}^{m,2m}(A, \boldsymbol{v})$. We assume this basis to be available and describe how to compute an orthonormal basis for $\mathbb{K}^{m+1,2m+2}(A, \boldsymbol{v})$ by using recursion formulas with few terms. For ease of exposition, all Krylov subspaces considered are assumed to be of maximal dimension, i.e., $\dim(\mathbb{K}^{\ell,m}(A, \boldsymbol{v})) = \ell + m - 1$. Our derivation of the recursion relations is similar to that of Section 2 and some details therefore are omitted.

We show how to determine the vectors $\boldsymbol{v}_{2m}$, $\boldsymbol{v}_{-m}$, and $\boldsymbol{v}_{2m+1}$, defined by (4.2), in order. Since the orthogonal Laurent polynomials (4.1) are monic, we have

$$\phi_{2m}(x) - x\phi_{2m-1}(x) \in \operatorname{span}\{\phi_0, \phi_1, \phi_2, \phi_{-1}, \dots, \phi_{2m-2}, \phi_{-m+1}.\phi_{2m-1}\}. \tag{4.3}$$

This expression is orthogonal to all Laurent polynomials (4.1) except for ϕ_{2m-1}, ϕ_{-m+1}, and ϕ_{2m-2}. Let the $\gamma_{2m-1,j}$ denote the coefficient of ϕ_j in a Fourier expansion of the Laurent polynomial (4.3) in terms of the Laurent polynomials (4.1). Then the only nonvanishing coefficients in this expansion are $\gamma_{2m-1,2m-1}$, $\gamma_{2m-1,-m+1}$, and $\gamma_{2m-1,2m-2}$. This yields the four-term recursion relation

$$\begin{aligned} \delta_{2m}\boldsymbol{v}_{2m} = (A - \alpha_{2m-1,2m-1}I_n)\boldsymbol{v}_{2m-1} - \alpha_{2m-1,-m+1}\boldsymbol{v}_{-m+1} \\ -\alpha_{2m-1,2m-2}\boldsymbol{v}_{2m-2}, \qquad m > 2, \end{aligned} \tag{4.4}$$

with $\alpha_{j,k} := \boldsymbol{v}_j^T A \boldsymbol{v}_k$.

Next we consider the computation of $\boldsymbol{v}_{-m}$. Arguments similar to those used in the proof of Proposition 2.1 ensure that the coefficient $c_{2m,-m+1}$ of ϕ_{2m} is non-zero. It follows that

$$c_{2m,-m+1}\phi_{-m}(x) - x^{-1}\phi_{2m}(x) \in \operatorname{span}\{\phi_0, \phi_1, \phi_2, \phi_{-1}, \dots, \phi_{-m+1}, \phi_{2m-1}, \phi_{2m}\}.$$

Similarly as above, we find that all coefficients $\gamma_{2m,j}$ in the Fourier expansion of this expression in terms of the Laurent polynomials (4.1) vanish except for $\gamma_{2m,-m+1}$, $\gamma_{2m,2m-1}$, and $\gamma_{2m,2m}$. Here $\gamma_{2m,j}$ is the coefficient for ϕ_j. Thus,

$$\begin{aligned} c_{2m,-m+1}\phi_{-m}(x) = x^{-1}\phi_{2m}(x) - \gamma_{2m,2m}\phi_{2m}(x) \\ -\gamma_{2m,-m+1}\phi_{-m+1}(x) - \gamma_{2m,2m-1}\phi_{2m-1}(x), \end{aligned}$$

which yields the four-term recursion relation

$$(4.5)\quad \delta_{-m}\boldsymbol{v}_{-m} = (A^{-1} - \beta_{2m,2m}I_n)\boldsymbol{v}_{2m} - \beta_{2m,-m+1}\boldsymbol{v}_{-m+1} - \beta_{2m,2m-1}\boldsymbol{v}_{2m-1}$$

with $\beta_{j,k} := \boldsymbol{v}_j^T A^{-1}\boldsymbol{v}_k$.

Lastly, consider the computation of the vector $\boldsymbol{v}_{2m+1}$. Proposition 2.1 guarantees that the trailing coefficient of ϕ_{-m+1} is nonvanishing and therefore

$$c_{-m,2m}\phi_{2m+1}(x) - x\phi_{-m}(x) \in \text{span}\{\phi_0, \phi_1, \phi_2, \phi_{-1}, \dots, \phi_{2m}, \phi_{-m}\}.$$

All Fourier coefficients $\gamma_{-m,j}$ of this expression vanish with the exceptions of $\gamma_{-m,-m}$ and $\gamma_{-m,2m}$, where $\gamma_{-m,j}$ is the coefficient of ϕ_j. We conclude that

$$c_{-m,2m}\phi_{2m+1}(x) = x\phi_{-m}(x) - \gamma_{-m,-m}\phi_{-m}(x) - \gamma_{-m,2m}\phi_{2m}(x),$$

which, for $m > 2$, yields the three-term recursion relation

$$(4.6)\qquad \delta_{2m+1}\boldsymbol{v}_{2m+1} = (A - \alpha_{-m,-m}I)\boldsymbol{v}_{-m} - \alpha_{-m,2m}\boldsymbol{v}_{2m}.$$

The recursion relations (4.4), (4.5), and (4.6) are the foundation for the following algorithm for computing an orthonormal basis for $\mathbb{K}^{m,2m}(A,\boldsymbol{v})$.

ALGORITHM 4.1 (Orthogonalization process for $\mathbb{K}^{m,2m}(A,\boldsymbol{v})$.).
Input: *m, $\boldsymbol{v}$, functions for evaluating matrix-vector products and solving linear systems of equations with A;*
Output: *orthogonal basis $\{\boldsymbol{v}_k\}_{k=-m+1}^{2m}$ of $\mathbb{K}^{m,2m}(A,\boldsymbol{v})$;*
$\delta_0 := ||\boldsymbol{v}||$; $\boldsymbol{v}_0 := \boldsymbol{v}/\delta_0$;
$\boldsymbol{u} := A\boldsymbol{v}_0$; $\alpha_{0,0} := \boldsymbol{v}_0^T\boldsymbol{u}$; $\boldsymbol{u} := \boldsymbol{u} - \alpha_{0,0}\boldsymbol{v}_0$;
$\delta_1 := ||\boldsymbol{u}||$; $\boldsymbol{v}_1 := \boldsymbol{u}/\delta_1$;
$\boldsymbol{u} := A\boldsymbol{v}_1$; $\alpha_{1,0} := \boldsymbol{v}_0^T\boldsymbol{u}$; $\boldsymbol{u} := \boldsymbol{u} - \alpha_{1,0}\boldsymbol{v}_0$;
$\alpha_{1,1} := \boldsymbol{v}_1^T\boldsymbol{u}$; $\boldsymbol{u} := \boldsymbol{u} - \alpha_{1,1}\boldsymbol{v}_1$;
$\delta_2 := ||\boldsymbol{u}||$; $\boldsymbol{v}_2 := \boldsymbol{u}/\delta_2$;
for $k = 1, 2, \dots, m-1$ **do**
 $\boldsymbol{w} := A^{-1}\boldsymbol{v}_{2k}$;
 $\beta_{2k,2k-2} := \boldsymbol{v}_{2k-2}^T\boldsymbol{w}$; $\boldsymbol{w} := \boldsymbol{w} - \beta_{2k,2k-2}\boldsymbol{v}_{2k-2}$;
 $\beta_{2k,2k-1} := \boldsymbol{v}_{2k-1}^T\boldsymbol{w}$; $\boldsymbol{w} := \boldsymbol{w} - \beta_{2k,2k-1}\boldsymbol{v}_{2k-1}$;
 $\beta_{2k,2k} := \boldsymbol{v}_{2k}^T\boldsymbol{w}$; $\boldsymbol{w} := \boldsymbol{w} - \beta_{2k,2k}\boldsymbol{v}_{2k}$;
 $\delta_{-k} := ||\boldsymbol{w}||$; $\boldsymbol{v}_{-k} := \boldsymbol{w}/\delta_{-k}$;
 $\boldsymbol{u} := A\boldsymbol{v}_{-k}$;
 $\alpha_{-k,2k} := \boldsymbol{v}_{2k}^T\boldsymbol{u}$; $\boldsymbol{u} := \boldsymbol{u} - \alpha_{-k,2k}\boldsymbol{v}_{2k}$;
 $\alpha_{-k,-k} := \boldsymbol{v}_{-k}^T\boldsymbol{u}$; $\boldsymbol{u} := \boldsymbol{u} - \alpha_{-k,-k}\boldsymbol{v}_{-k}$;
 $\delta_{2k+1} := ||\boldsymbol{u}||$; $\boldsymbol{v}_{2k+1} := \boldsymbol{u}/\delta_{2k+1}$;
 $\boldsymbol{u} := A\boldsymbol{v}_{2k+1}$;
 $\alpha_{2k+1,2k} := \boldsymbol{v}_{2k}^T\boldsymbol{u}$; $\boldsymbol{u} := \boldsymbol{u} - \alpha_{2k+1,2k}\boldsymbol{v}_{2k}$;
 $\alpha_{2k+1,-k} := \boldsymbol{v}_{-k}^T\boldsymbol{u}$; $\boldsymbol{u} := \boldsymbol{u} - \alpha_{2k+1,-k}\boldsymbol{v}_{-k}$;
 $\alpha_{2k+1,2k+1} := \boldsymbol{v}_{2k+1}^T\boldsymbol{u}$; $\boldsymbol{u} := \boldsymbol{u} - \alpha_{2k+1,2k+1}\boldsymbol{v}_{2k+1}$;
 $\delta_{2k+2} := ||\boldsymbol{u}||$; $\boldsymbol{v}_{2k+2} := \boldsymbol{u}/\delta_{2k+2}$;
end

Given the orthonormal basis for the subspace $\mathbb{K}^{m,2m}(A,\boldsymbol{v})$, analogously to (2.2), we define the matrices

$$(4.7)\qquad \begin{aligned} V_{3m+1} &= [\boldsymbol{v}_0, \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_{-1}, \dots, \boldsymbol{v}_{2m}, \boldsymbol{v}_{-m+1}] \in \mathbb{R}^{n\times(3m+1)}, \\ V_{3m+2} &= [V_{3m+1}, \boldsymbol{v}_{2m+1}] \in \mathbb{R}^{n\times(3m+2)}. \end{aligned}$$

Similarly to the construction in Section 2, the recursion coefficients generated by Algorithm 4.1 can be used to determine a matrix $\hat{H}_{3m+1} = [h_{j,k}] \in \mathbb{R}^{(3m+2)\times(3m+1)}$, such that

$$AV_{3m+1} = V_{3m+2}\hat{H}_{3m+1}, \tag{4.8}$$

where the matrices V_{3m+1} and V_{3m+2} are given by (4.7). The leading submatrix $H_{3m+1} \in \mathbb{R}^{(3m+1)\times(3m+1)}$ of $\hat{H}_{3m+1}$ satisfies

$$H_{3m+1} = V_{3m+1}^T A V_{3m+1}. \tag{4.9}$$

We note that even though four-term recursions occur in Algorithm 4.1, the matrix H_{3m+1} is pentadiagonal. The $(3k-2)^{\text{th}}$ column of AV_{3m+1} is $A\boldsymbol{v}_{-k+1}$, and by the relation (4.6), with m replaced by $k-1$, or by the recursion formulas of Algorithm 4.1, we obtain, for $k = 2, 3, \ldots, m-1$,

$$A\boldsymbol{v}_{-k+1} = \alpha_{-k+1,2k-2}\boldsymbol{v}_{2k-2} + \alpha_{-k+1,-k+1}\boldsymbol{v}_{-k+1} + \delta_{2k-1}\boldsymbol{v}_{2k-1}. \tag{4.10}$$

Hence, the only nontrivial entries of the $(3k-2)^{\text{th}}$ column of H_{3m+1} are

$$h_{3k-3,3k-2} = \alpha_{-k+1,2k-2}, \qquad h_{3k-2,3k-2} = \alpha_{-k+1,-k+1}, \qquad h_{3k-1,3k-2} = \delta_{2k-1}.$$

The $(3k-1)^{\text{th}}$ column of AV_{3m+1} is $A\boldsymbol{v}_{2k-1}$, and by relation (4.4) with m replaced by k, we obtain, for $k = 2, 3, \ldots, m-1$,

$$\begin{aligned} A\boldsymbol{v}_{2k-1} = {} & \alpha_{2k-1,2k-2}\boldsymbol{v}_{2k-2} + \alpha_{2k-1,-k+1}\boldsymbol{v}_{-k+1} + \\ & \alpha_{2k-1,2k-1}\boldsymbol{v}_{2k-1} + \delta_{2k}\boldsymbol{v}_{2k}. \end{aligned}$$

It follows that the $(3k-1)^{\text{th}}$ column of H_{3m+1} only has the nontrivial entries

$$\begin{aligned} h_{3k-3,3k-1} &= \alpha_{2k-1,2k-2}, \qquad & h_{3k-2,3k-1} &= \alpha_{2k-1,-k+1}, \\ h_{3k-1,3k-1} &= \alpha_{2k-1,2k-1}, \qquad & h_{3k,3k-1} &= \delta_{2k+1}. \end{aligned}$$

The nonvanishing entries of the $(3k)^{\text{th}}$ column are derived by multiplying expression (4.5) by the matrix A and replacing m by k. The derivation of an expression of $A\boldsymbol{v}_{2k}$ in terms of vectors $\boldsymbol{v}_j$ is analogous to the derivation of (2.13). We obtain

$$\begin{aligned} A\boldsymbol{v}_{2k} = {} & h_{3k-3,3k}\boldsymbol{v}_{2k-3} + h_{3k-2,3k}\boldsymbol{v}_{-k+1} + h_{3k-1,3k}\boldsymbol{v}_{2k-1} + \\ & h_{3k,3k}\boldsymbol{v}_{2k} + h_{3k+1,3k}\boldsymbol{v}_{-k} + h_{3k+2,3k}\boldsymbol{v}_{2k+1}, \end{aligned} \tag{4.11}$$

where we have used the fact that $\beta_{2k,2k} = \boldsymbol{v}_{2k}^T A^{-1}\boldsymbol{v}_{2k} > 0$, which follows from the positive definiteness of A. Orthonormality of the vectors $\boldsymbol{v}_j$ and symmetry of A and H_{3m+1} now give

$$h_{3k-3,3k} = h_{3k-2,3k} = 0, \qquad h_{3k-1,3k} = h_{3k,3k-1} = \delta_{2k+1},$$

as well as

$$h_{3k+1,3k} = -\frac{\delta_{-k}\alpha_{-k,-k}}{\beta_{2k,2k}}, \qquad h_{3k+2,3k} = -\frac{\delta_{-k}\delta_{2k+1}}{\beta_{2k,2k}},$$

$$h_{3k,3k} = \frac{1 - \beta_{2k,2k-1}\delta_{2k} - \alpha_{-k,2k}\delta_{-k}}{\beta_{2k,2k}}.$$

We also observe that, as a consequence of the symmetry of H_{3m+1},

$$h_{3k+1,3k} = h_{3k,3k+1} = \alpha_{-k,2k}, \qquad h_{3k+2,3k} = h_{3k,3k+2} = \alpha_{2k+1,2k}.$$

Example 4.1. Let $m = 3$. The matrix H_{10} is of the form

$$\begin{bmatrix} \alpha_{0,0} & \delta_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \delta_1 & \alpha_{1,1} & \delta_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \delta_2 & h_{3,3} & \alpha_{2,-1} & \alpha_{2,3} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha_{2,-1} & \alpha_{-1,-1} & \alpha_{-1,3} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha_{2,3} & \delta_3 & \alpha_{3,3} & \delta_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \delta_4 & h_{6,6} & \alpha_{4,-2} & \alpha_{4,5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha_{4,-2} & \alpha_{-2,-2} & \alpha_{-2,5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha_{4,5} & \delta_5 & \alpha_{5,5} & \delta_6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \delta_6 & h_{9,9} & \alpha_{6,-3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha_{6,-3} & \alpha_{-3,-3} \end{bmatrix}.$$

Moreover, the matrix $\hat{H}_{10}$ in (4.8) is given by

$$\hat{H}_{10} = \begin{bmatrix} H_{10} \\ \boldsymbol{h}_{10}^T \end{bmatrix}$$

with

$$\boldsymbol{h}_{10} = -\frac{\delta_{-3}\delta_7}{\beta_{6,6}}\boldsymbol{e}_9 + \delta_7 \boldsymbol{e}_{10} \in \mathbb{R}^{10}.$$

□

5. Numerical examples. The computations in this section are performed using MATLAB with about 15 significant decimal digits. In all examples, except when explicitly stated otherwise, $A \in \mathbb{R}^{1000\times 1000}$ and the vector $\boldsymbol{v} \in \mathbb{R}^{1000}$ has normally distributed random entries with mean zero and variance one. We will refer to the rational Lanczos method that uses the Krylov subspace $\mathbb{K}^{\ell,m}(A, \boldsymbol{v})$ as Lanczos(ℓ, m).

In all computed examples, we use Krylov subspaces of dimension 42. A reason for this is that 42 is divisible by both 2 and 3, and this slightly simplifies the implementation of the rational Krylov subspace methods considered. We determine the actual value $\boldsymbol{w}$, given by (1.1), as well as approximations

$$\hat{\boldsymbol{w}}_{42} = V_{42} f(H_{42}) \boldsymbol{e}_1 \|\boldsymbol{v}\|$$

obtained by the Lanczos$(21, 22)$ method of Sections 2-3 and by the Lanczos$(14, 29)$ method of Section 4. For comparison, we also compute the approximation $\boldsymbol{w}_{42}$, defined by (1.5) with $m = 42$, and evaluated by using the (standard) Lanczos decomposition (1.3) with $m = 42$. We refer to this method as Lanczos(42) in the tables, which display the errors $\|\boldsymbol{w} - \hat{\boldsymbol{w}}_{42}\|$ for Lanczos$(21, 22)$ and Lanczos$(14, 29)$, as well as the error $||\boldsymbol{w} - \boldsymbol{w}_{42}||$ for Lanczos(42), for several functions f.

All matrix functions are computed by means of the spectral decomposition of the matrix. For the function $f(x) = \exp(x)/x$, we evaluate (1.1) as $\exp(A)A^{-1}\boldsymbol{v}$, where $A^{-1}\boldsymbol{v}$ is computed by solving a linear system of equations. The rational Lanczos$(21, 22)$ method yields the approximation

$$\hat{\boldsymbol{w}}_{42} = V_{42} \exp(H_{42}) H_{42}^{-1} \boldsymbol{e}_1 \|\boldsymbol{v}\|,$$

with the symmetric and pentadiagonal matrix H_{42} defined by (2.3). The vector $H_{42}^{-1}\boldsymbol{e}_1$ is determined by evaluating the first column of the pentadiagonal matrix G_{42} given by (2.4). Computations with Lanczos$(14, 29)$ are carried out similarly. The standard Lanczos(42) method determines the Lanczos decomposition (1.3) with $m = 42$, which yields the approximation

$$\boldsymbol{w}_{42} = V_{42} \exp(T_{42}) T_{42}^{-1} \boldsymbol{e}_1 \|\boldsymbol{v}\|.$$

This expression is evaluated by first solving a linear system of equations for the vector $T_{42}^{-1}\boldsymbol{e}_1$.

The following examples show the approximations computed by using the rational Lanczos$(21, 22)$ and Lanczos$(14, 29)$ methods to be superior to approximations determined by the standard Lanczos(42) method. For most examples Lanczos$(14, 29)$ yields an as accurate approximation as Lanczos$(21, 22)$. This is interesting because for many matrices that arise in applications, matrix-vector products can be evaluated faster than solutions of linear systems of equations with the matrix.

$f(x)$	Lanczos(42)	Lanczos$(21, 22)$	Lanczos$(14, 29)$
$\exp(-x)$	$2.3 \cdot 10^{-6}$	$3.4 \cdot 10^{-15}$	$3.8 \cdot 10^{-15}$
$\sqrt{x}$	$1.3 \cdot 10^{0}$	$2.1 \cdot 10^{-2}$	$3.6 \cdot 10^{-2}$
$\exp(-\sqrt{x})$	$1.0 \cdot 10^{-3}$	$2.5 \cdot 10^{-13}$	$2.6 \cdot 10^{-13}$
$\ln(x)$	$1.8 \cdot 10^{-1}$	$3.4 \cdot 10^{-4}$	$7.1 \cdot 10^{-4}$
$\exp(-x)/x$	$2.4 \cdot 10^{-7}$	$3.5 \cdot 10^{-16}$	$3.9 \cdot 10^{-16}$

TABLE 5.1

Example 5.1: Errors in approximations of $f(A)\boldsymbol{v}$ determined by the standard and rational Lanczos methods for a symmetric positive definite tridiagonal matrix A.

Example 5.1. We compute approximations of $f(A)\boldsymbol{v}$ determined by the standard and rational Lanczos methods for the symmetric positive definite tridiagonal matrix $A = n^2[-1, 2, -1]$ of order $n = 1000$. The approximation errors are reported in Table 5.1. Note that the rational Lanczos methods yield significantly smaller approximation errors for many of the functions f than the standard Lanczos method. Moreover, both rational Lanczos methods Lanczos$(21, 22)$ and Lanczos$(14, 29)$ determine approximations of about the same quality. □

Example 5.2. Let $A = [a_{i,j}]$ be the symmetric positive definite Toeplitz matrix with entries $a_{i,j} = 1/(1 + |i - j|)$. Computed results are shown in Table 5.2. We remark that fast direct solution methods are available for linear systems of equations with this kind of matrix; see, e.g., [3, 25]. Approximations of (1.1) determined by the rational Lanczos methods Lanczos$(21, 22)$ and Lanczos$(14, 29)$ are seen to be of

$f(x)$	Lanczos(42)	Lanczos(21, 22)	Lanczos(14, 29)
$\exp(-x)$	$8.2 \cdot 10^{-15}$	$8.2 \cdot 10^{-15}$	$8.1 \cdot 10^{-15}$
$\sqrt{x}$	$1.9 \cdot 10^{-11}$	$1.0 \cdot 10^{-14}$	$1.0 \cdot 10^{-14}$
$\exp(-\sqrt{x})$	$1.9 \cdot 10^{-11}$	$6.9 \cdot 10^{-15}$	$7.0 \cdot 10^{-15}$
$\ln(x)$	$3.0 \cdot 10^{-10}$	$1.4 \cdot 10^{-14}$	$1.3 \cdot 10^{-14}$
$\exp(-x)/x$	$7.6 \cdot 10^{-9}$	$1.6 \cdot 10^{-14}$	$1.5 \cdot 10^{-14}$

Table 5.2

Example 5.2: Errors in approximations of $f(A)\boldsymbol{v}$ determined by the standard and rational Lanczos methods for a symmetric positive definite Toeplitz matrix A.

$f(x)$	Lanczos(42)	Lanczos(21, 22)	Lanczos(14, 29)
$\sqrt{x}$	$2.0 \cdot 10^{-2}$	$3.7 \cdot 10^{-5}$	$5.0 \cdot 10^{-5}$
$\exp(-\sqrt{x})$	$1.3 \cdot 10^{-2}$	$3.6 \cdot 10^{-7}$	$2.1 \cdot 10^{-6}$
$\ln(x)$	$5.7 \cdot 10^{-2}$	$1.4 \cdot 10^{-5}$	$2.7 \cdot 10^{-5}$

Table 5.3

Example 5.3: Errors in approximations of $f(A)\boldsymbol{v}$ determined by the standard and rational Lanczos methods for a symmetric positive definite matrix $A = I + X^T X$, with X randomly generated.

higher accuracy than approximations obtained with the standard Lanczos method. Both rational Lanczos methods yield approximants of about the same accuracy. □

Example 5.3. Let $A = I + X^T X$, where $X \in \mathbb{R}^{1000\times1000}$ has randomly generated normally distributed entries with zero mean and variance one. Table 5.3 displays computed results and shows approximations of expressions (1.1) computed with the rational Lanczos methods to be more accurate than approximations determined by the standard Lanczos method. □

$f(x)$	Lanczos(42)	Lanczos(21, 22)	Lanczos(14, 29)
$\exp(x)$	$2.4 \cdot 10^{-2}$	$1.3 \cdot 10^{-7}$	$3.6 \cdot 10^{-6}$
$\exp(x)/x$	$2.5 \cdot 10^{-2}$	$3.0 \cdot 10^{-8}$	$5.1 \cdot 10^{-7}$

Table 5.4

Example 5.4: Errors in approximations of $f(A)\boldsymbol{v}$ determined by the standard and rational Lanczos methods for the symmetric negative definite matrix $A = -(I + X^T X)$, with X randomly generated.

Example 5.4. The matrix used in this example is of the form Let $A = -(I+X^T X)$, where $X \in \mathbb{R}^{1000\times1000}$ is generated similarly as in Example 5.3. Table 5.4 shows the errors in approximations of (1.1) determined by the rational and standard Lanczos methods. □

Example 5.5. The matrix used in this example is symmetric indefinite and of the form

$$A = \begin{bmatrix} B & C \\ C^T & -B \end{bmatrix},$$

where B is a tridiagonal symmetric Toeplitz matrix of order 500 with a typical row $[-1, 2, -1]$. All entries of $C \in \mathbb{R}^{500\times500}$ are zero with the exception of the entry 1 in the lower left corner of the matrix. Table 5.5 shows the error in approximations of (1.1) determined by the rational and standard Lanczos methods. The standard Lanczos method is seen to be unable to determine an accurate approximation of $f(t) = \exp(t)/t$. □

$f(x)$	Lanczos(42)	Lanczos(21, 22)	Lanczos(14, 29)
$\exp(x)$	$2.4 \cdot 10^{-13}$	$4.0 \cdot 10^{-10}$	$2.8 \cdot 10^{-13}$
$\exp(x)/x$	$2.1 \cdot 10^{3}$	$2.8 \cdot 10^{-10}$	$3.8 \cdot 10^{-10}$

TABLE 5.5

Example 5.5: Errors in approximations of $f(A)\boldsymbol{v}$ determined by the standard and rational Lanczos methods for a symmetric indefinite matrix.

$f(x)$	Lanczos(42)	Lanczos(21, 22)	Lanczos(14, 29)
$1/\sqrt{x}$	$1.4 \cdot 10^{-2}$	$5.6 \cdot 10^{-13}$	$2.7 \cdot 10^{-12}$

TABLE 5.6

Example 5.6: Errors in approximations of $f(A)\boldsymbol{v}$ determined by the standard and rational Lanczos methods for a symmetric positive definite matrix.

Example 5.6. The matrix used in this example is obtained from the discretization of the self-adjoint differential operator $L(u) = \frac{1}{10}u_{xx} - 100u_{yy}$ in the unit square. Each derivative is approximated by the standard three-point stencil with 40 equally spaced interior nodes in each space dimension. Homogeneous boundary conditions are used. This yields a 1600×1600 symmetric positive definite matrix A. The initial vector $\boldsymbol{v}$ for the polynomial and rational Lanczos processes is chosen to be the unit vector with all entries 1/40. Table 5.6 shows the errors in approximations of (1.1) determined by the standard and rational Lanczos methods. □

6. Conclusion and extension. The computed examples of Section 5 show that for many approximation problems (1.1) rational Lanczos methods can give significantly higher accuracy with the same number of steps than the standard Lanczos method. This is in agreement with the analyses presented in [4, 11, 12, 17]. Rational Lanczos methods require the solution of linear systems of equations, and it depends on the size, sparsity or structure of A if the solution of these systems is feasible. Structures, besides sparsity, that makes it possible to solve large linear systems of equations fairly rapidly include bandedness and semiseparability; see Vandebril et al. [25] for an authoritative treatment of the latter. In particular, Toeplitz matrices are semiseparable.

Many matrices of interest in applications allow faster computations of matrix-vector products than solution of linear systems of equations. It therefore can be of interest to use a rational Lanczos method that requires fewer linear systems to be solved than matrix-vector product evaluations. Section 4 illustrates that rational Lanczos methods for Krylov subspaces of the form $\mathbb{K}^{\ell,2\ell}(A, \boldsymbol{v})$ can be implemented with short recursion formulas, and the computed examples of Section 5 show that these rational Lanczos methods are competitive with regard to accuracy. We are presently investigating properties of rational Lanczos methods for Krylov subspaces $\mathbb{K}^{\ell,m}(A, \boldsymbol{v})$ with different ratios m/ℓ.

Acknowledgment. The authors would like to thank Olav Njåstad for reference [18].

REFERENCES

[1] M. Afanasjew, M. Eiermann, O. G. Ernst, and S. Güttel, A generalization of the steepest descent method for matrix functions, Electron. Trans. Numer. Anal., 28 (2008), pp. 206–222.

[2] E. J. Allen, J. Baglama, and S. K. Boyd, Numerical approximation of the product of the square root of a matrix with a vector, Linear Algebra Appl., 310 (2000), pp. 167–181.
[3] G. S. Ammar and W. B. Gragg, Superfast solution of real positive definite Toeplitz systems, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 61–76.
[4] B. Beckermann and L. Reichel, Error estimation and evaluation of matrix functions via the Faber transform, submitted for publication, 2008.
[5] M. Benzi and N. Razouk, Decay bounds and $O(n)$ algorithms for approximating functions of sparse matrices, Electron. Trans. Numer. Anal., 28 (2007), pp. 16–39.
[6] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad, Orthogonal Rational Functions, Cambridge University Press, Cambridge, 1999.
[7] D. Calvetti and L. Reichel, Exponential integration methods for large stiff systems of differential equations, in Iterative Methods in Scientific Computing IV, eds. D. R. Kincaid and A. C. Elster, IMACS Series in Computational and Applied Mathematics, vol. 5, IMACS, New Brunswick, 1999, pp. 237–243.
[8] D. Calvetti and L. Reichel, Lanczos-based exponential filtering for discrete ill-posed problems, Numer. Algorithms, 29 (2002), pp. 45–65.
[9] D. Calvetti, L. Reichel, and Q. Zhang, Iterative exponential filtering for large discrete ill-posed problems, Numer. Math., 83 (1999), pp. 535–556.
[10] V. L Druskin and L. A. Knizhnerman, Two polynomial methods for the computation of functions of symmetric matrices, USSR Comput. Math. Math. Phys., 29 (1989) pp. 112–121.
[11] V. Druskin and L. Knizhnerman, Extended Krylov subspace approximations of the matrix square root and related functions, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 755–771.
[12] V. Druskin, L. Knizhnerman, and M. Zaslavsky, Solution of large scale evolutionary problems using rational Krylov subspaces with optimized shifts, submitted for publication, 2008.
[13] N. J. Higham, Functions of Matrices: Theory and Computation, SIAM, Philadelphia, 2008.
[14] M. Hochbruck and C. Lubich, On Krylov subspace approximations to the matrix exponential operator, SIAM J. Numer. Anal., 34 (1997), pp. 1911–1925.
[15] W. B. Jones and O. Njåstad, Orthogonal Laurent polynomials and strong moment theory: a survey, J. Comput. Appl. Math., 105 (1999), pp. 51–91.
[16] W. B. Jones, W. J. Thron, and O. Njåstad, Orthogonal Laurent polynomials and the strong Hamburger moment problem, J. Math. Anal. Appl., 98 (1984), pp. 528–554.
[17] L. Knizhnerman and V. Simoncini, A new investigation of the extended Krylov subspace method for matrix function evaluations, submitted for publication, 2008.
[18] O. Njåstad and W. J. Thron, The theory of sequences of orthogonal L-polynomials, in Padé Approximants and Continued Fractions, eds. H. Waadeland and H, Wallin, Det Kongelige Norske Videnskabers Selskab, Skrifter 1, 1983, pp. 54–91.
[19] A. Ruhe, Rational Krylov sequence methods for eigenvalue computations, Linear Algebra Appl., 58 (1984), pp. 391–405.
[20] A. Ruhe, Rational Krylov: a practical algorithm for large sparse nonsymmetric matrix pencils, SIAM J. Sci. Comput., 19 (1998), pp. 1535–1551.
[21] Y. Saad, Analysis of some Krylov subspace approximations to the matrix exponential operator, SIAM J. Numer. Anal., 29 (1992), pp. 209–228.
[22] Y. Saad, Iterative Methods for Sparse Linear Systems, 2nd ed., SIAM, Philadelphia, 2003.
[23] T. Schmelzer and L. N. Trefethen, Evaluating matrix functions for exponential integrators via Carathéodory-Fejér approximation and contour integrals, Electron. Trans. Numer. Anal., 29 (2007), pp. 1–18.
[24] V. Simoncini, A new iterative method for solving large-scale Lyapunov matrix equations, SIAM J. Sci. Comput., 29 (2007), pp. 1268–1288.
[25] R. Vandebril, M. Van Barel, and N. Mastronardi, Matrix Computations and Semiseparable Matrices, Vol. 1: Linear Systems, Johns Hopkins University Press, Baltimore, 2007.
[26] H. A. van der Vorst, An iterative solution method for solving $f(A)x = b$, using Krylov subspace information obtained for the symmetric positive definite matrix A, J. Comput. Appl. Math., 18 (1987), pp. 249–263.
[27] H. A. van der Vorst, Iterative Krylov Methods for Large Linear Systems, Cambridge University Press, Cambridge, 2003.