Pergamon

Computers Math. Applic. Vol. 33, No. 10, pp. 23–40, 1997

0898-1221/97 $17.00 + 0.00

PII: S0898-1221(97)00074-6

The Shift-Inverted J-Lanczos Algorithm for the Numerical Solutions of Large Sparse Algebraic Riccati Equations

W. R. FERNG
Department of Appl. Math.
National Chiao Tung University, Taiwan, R.O.C.
ferng@math.nctu.edu.tw

WEN-WEI LIN
Institute of Appl. Math., National Tsing Hua University
Hsinchu, Taiwan, 30043, R.O.C.
wwlin@am.nthu.edu.tw

CHERN-SHUH WANG
National Center for High-Performance Computing
Hsinchu, Taiwan, R.O.C.
c00shu00@nchc.gov.tw

(Received December 1996; accepted January 1997)

Abstract—The goal of solving an algebraic Riccati equation is to find the stable invariant subspace corresponding to all the eigenvalues lying in the open left-half plane. The purpose of this paper is to propose a structure-preserving Lanczos-type algorithm incorporated with shift and invert techniques, named shift-inverted J-Lanczos algorithm, for computing the stable invariant subspace for large sparse Hamiltonian matrices. The algorithm is based on the J-tridiagonalization procedure of a Hamiltonian matrix using symplectic similarity transformations. We give a detailed analysis on the convergence behavior of the J-Lanczos algorithm and present error bound analysis and Paige-type theorem. Numerical results for the proposed algorithm applied to a practical example arising from the position and velocity control for a string of high-speed vehicles are reported.

Keywords—Riccati equation, Hamiltonian matrix, J-Lanczos algorithm, J-tridiagonalization, Sympletic matrix, SR factorization.

1. INTRODUCTION

The problem of solving the algebraic Riccati equation

$$-XNX + XA + A^{\top}X + K = 0, \tag{1.1}$$

where X, N, K, and A are real $n \times n$ matrices, $K = K^{\top} \geq 0$ (positive semidefinite) and $N = N^{\top} \geq 0$, frequently arises in optimal-control problems. It is assumed that (A, B) is stabilizable and (C, A) is detectable, where B and C are full rank factorizations of N and K, respectively [1]. Under these assumptions, equation (1.1) has a unique symmetric positive semidefinite solution. A well-known procedure is to compute the n-dimensional invariant subspace $\begin{bmatrix} Y \\ Z \end{bmatrix}$ corresponding

Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX

to the stable eigenvalues of the Hamiltonian matrix

$$M = \begin{bmatrix} A & N \\ K & -A^{\mathrm{T}} \end{bmatrix}.$$

The solution of equation (1.1) is then obtained by $X = -ZY^{-1}$ [1].

Many structure-preserving numerical algorithms have been proposed for computing the invariant subspace of a Hamiltonian matrix, thus for solving the algebraic Riccati equation (1.1), see, e.g., [2–7]. One type of these methods, [2,3] are based on the symplectic QR-type transformations in which the SR decomposition with symplectic similarity transformations is used to replace the usual QR decomposition. The other type of methods exploit the square of the Hamiltonian matrix (skew Hamiltonian) to compute the corresponding eigenvalues and use them to find the stable invariant subspaces. These algorithms are very efficient for problems of small or medium sizes, but they become inadequate for very large and sparse cases.

Since there is also a wide class in control theory, such as position and velocity control [8] or circulant system analysis [9], which leads to solve large sparse Hamiltonian eigenvalue problems, it is not practical to perform algorithms that require the modifications of the underlined Hamiltonian matrix for this type of applications. Hence, some Lanczos-type algorithms were proposed in [10–12] for computing large sparse Hamiltonian eigenvalue problems, in which the nonsymmetric look-ahead Lanczos algorithm is applied to reduce the Hamiltonian matrix to a block tridiagonal matrix without modifying the matrix itself.

In this paper, we present a structure-preserving Lanczos-type algorithm, named J-Lanczos algorithm, for solving large sparse Hamiltonian eigenvalue problems. In this algorithm, the Hamiltonian matrix M is partially reduced to a J-tridiagonal matrix using a sequence of symplectic similarity transformations. Just like the conventional Lanczos algorithm, information about M's extreme eigenvalues tends to emerge long before the J-tridiagonalization process is completed. The J-Ritz pairs (eigen-pairs of J-tridiagonal submatrices) computed by QR or symplectic QR-like algorithm [2] are used to approximate the extreme eigen-pairs of M.

The goal of solving the algebraic Riccati equation (1.1) is to find the stable invariant subspace corresponding to all the eigenvalues lying in the open left-half plane. Since the J-Lanczos algorithm converges to the extreme eigenvalues fast, there are two important aspects in practice. One is to develop a shift strategy for determining a sequence of shifts so that the J-Lanczos algorithm can be sequentially applied to the new shifted and inverted Hamiltonian matrices. The other is to determine how many shifts with how many J-Lanczos steps should be used. In practice, we begin with the zero shift and then we use the distribution density of the computed eigenvalues to predict the next shift and the number of the J-Lanczos steps. We name this approach the shift-inverted J-Lanczos algorithm.

Although the proposed J-Lanczos algorithm is mathematically equivalent to the the Lanczos-type algorithms in [10–12], derivation of J-Lanczos algorithm starts from a different point of view. Furthermore, an error bound analysis based on [13] which demonstrates the convergence behavior of the J-Lanczos algorithm is analyzed in depth in this paper for the J-Ritz values. We also present a variant Paige-type theorem [14] for the J-Lanczos algorithm which shows that the constructed J-Lanczos vectors will lose the symplecticity when some J-Ritz values begin to converge.

We organize this paper as follows. Some definitions that related to the so-called J-structure matrices are reviewed in the preliminary Section 2. In Section 3, we establish the existence theorem of the J-tridiagonalization of a Hamiltonian matrix and develop the J-Lanczos algorithm. The convergence analysis of the J-Ritz value and a variant Paige-type theorem for the J-Lanczos are presented in Section 4. Shift-invert strategies and numerical results for the proposed J-Lanczos method applied to a practical example arising from the position and velocity control for a string of high-speed vehicles [8] are discussed in Section 5. Concluding remarks are given in Section 6.

2. PRELIMINARIES

Herein, we denote the $n \times n$ identity matrix by I_n and define

$$J_n = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.$$

Note that $J_n^{-1} = J_n^{\top} = -J_n$. Let $\Pi_n \in \mathbb{R}^{2n\times 2n}$ be the permutation matrix

$$\Pi_n = [\, e_1, e_3, \dots, e_{2n-1}, e_2, e_4, \dots, e_{2n} \,], \tag{2.1}$$

where e_j is the j^{th} column of the identity matrix I_{2n}. If confusion is unlikely, the subscript n will be omitted.

A matrix $S \in \mathbb{R}^{2n\times 2m}$ $(n \geq m)$ is symplectic if $S^{\top} J_n S = J_m$. A matrix $M \in \mathbb{R}^{2n\times 2n}$ is Hamiltonian if and only if $(JM)^{\top} = JM$. The definitions for the J-structure matrices [2], SR factorization, and Krylov matrices that will be referred later in this paper are given as follows.

DEFINITION 2.1. *Let*

$$G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix}$$

be a $2n \times 2n$ matrix with $G_{ij} \in \mathbb{R}^{n\times n}$.

(i) *G is called a J-Hessenberg matrix if G_{11}, G_{21}, and G_{22} are upper triangular and G_{12} is upper Hessenberg. In addition, G is called an unreduced J-Hessenberg matrix if G_{12} is unreduced and G_{21} is nonsingular.*

(ii) *G is called a J-upper triangular matrix if G_{11}, G_{12}, and G_{22} are upper triangular and G_{21} is strictly upper triangular. In addition, G is J-strictly upper triangular if G_{11} and G_{22} are strictly upper triangular.*

(iii) *G is called a J-tridiagonal matrix if G_{11}, G_{21}, and G_{22} are diagonal and G_{12} is tridiagonal.*

DEFINITION 2.2. *Suppose $A \in \mathbb{R}^{2n\times 2m}$ $(n \geq m)$. The factorization $A = SR$, where $S \in \mathbb{R}^{2n\times 2m}$ is symplectic and $R \in \mathbb{R}^{2m\times 2m}$ is J_m-triangular is called an SR-factorization of A.*

DEFINITION 2.3. *Let $M \in \mathbb{R}^{2n\times 2n}$ be a Hamiltonian matrix. Given $x \in \mathbb{R}^{2n}$ and a positive integer j.*

(i) *The Krylov matrix of M with respect to x and j is defined by*

$$K_j = K[M, x, 2j] = \left[x, Mx, \dots, M^{j-1}x \mid M^j x, \dots, M^{2j-1}x\right].$$

(ii) *The Krylov subspace spanned by the columns of $K[M, x, 2j]$ is denoted by $K(M, x, 2j)$.*

3. J-TRIDIAGONALIZATION AND J-LANCZOS ALGORITHM

In this section, we establish the existence theorem of the J-tridiagonalization of a Hamiltonian matrix and develop the related J-Lanczos algorithm. This algorithm is equivalent to the algorithms proposed in [10–12], however, the derivation starts from a completely different point of view. First, the results of [3, Theorem 3.4, (i),(ii)] is generalized to a more general form.

THEOREM 3.1. *Let $M \in \mathbb{R}^{2n\times 2n}$ be a Hamiltonian matrix and for a given $2n$-vector q_1, $K_m = K[M, q_1, 2m]$, $m \leq n$, be a Krylov matrix with $\operatorname{rank}(K_m) = 2m$. If $K_m \Pi_m = S_m R_m$ is an SR-factorization, then*

$$H_m = \left(J_m^{\top} S_m^{\top} J\right) M S_m \tag{3.1}$$

is an unreduced J_m-tridiagonal matrix such that

$$M S_m = S_m H_m + z_m e_{2m}^{\top} \tag{3.2}$$

and $(J_m^{\top} S_m^{\top} J) z_m = 0$, for a suitable $z_m \in \mathbb{R}^{2n}$.

PROOF. For any given q_1, one can find a vector $y_m \perp \text{Range}(K_m)$ and scalars $\alpha_0, \alpha_1, \ldots, \alpha_{2m-1} \in \mathbb{R}$ such that $M^{2m}q_1 = \sum_{i=0}^{2m-1} \alpha_i M^i q_1 + y_m$. Here y_m can be a zero vector. Let

$$C_m = \begin{bmatrix} 0 & & & & \alpha_0 \\ 1 & 0 & & & \alpha_1 \\ & \ddots & \ddots & & \vdots \\ & & \ddots & 0 & \vdots \\ & & & 1 & \alpha_{2m-1} \end{bmatrix}.$$

Then we have

$$MK_m = K_m C_m + y_m e_{2m}^\top. \tag{3.3}$$

Since $K_m \Pi_m = S_m R_m$ is an SR-factorization with R_m, a $2m \times 2m$ nonsingular J_m-triangular matrix equation (3.3) can be written as

$$MS_m = S_m R_m \Pi_m^\top C_m \Pi_m R_m^{-1} + \delta_m y_m e_{2m}^\top,$$

where $\delta_m = e_{2m}^\top R_m^{-1} e_{2m}$. Thus, by letting $\widetilde{y}_m = \delta_m (J_m^\top S_m^\top J) y_m$, we have

$$\left(J_m^\top S_m^\top J\right) MS_m = R_m \Pi_m^\top C_m \Pi_m R_m^{-1} + \widetilde{y}_m e_{2m}^\top \equiv H_m.$$

Since R_m and R_m^{-1} are J_m-triangular and $\Pi_m^\top C_m \Pi_m$ is J_m-Hessenberg, H_m is J_m-Hessenberg. But $(J_m^\top S_m^\top J) MS_m$ is Hamiltonian, therefore H_m is J_m-tridiagonal. Since R_m is nonsingular and C_m is companion it follows that H_m is unreduced.

To prove (3.2), we use (3.3)

$$\begin{aligned} MK_m &= K_m C_m + y_m e_{2m}^\top \\ &= K_m \left(\begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & \cdots & 0 & \alpha_0 \\ \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & \alpha_{2m-1} \end{bmatrix} \right) + y_m e_{2m}^\top \\ &\equiv K_m Z_m + M^{2m} q_1 e_{2m}^\top. \end{aligned}$$

Since $K_m \Pi_m = S_m R_m$, it follows that

$$\begin{aligned} MS_m = S_m &\left(R_m \Pi_m^\top Z_m \Pi_m R_m^{-1} + \left(J^\top S_m J_m\right)^\top M^{2m} q_1 e_{2m}^\top R_m^{-1} \right) \\ &+ \left(I - S_m \left(J^\top S_m J_m\right)^\top \right) M^{2m} q_1 e_{2m}^\top R_m^{-1}. \end{aligned}$$

Let

$$z_m = \left(I - S_m \left(J^\top S_m J_m\right)^\top \right) M^{2m} q_1 \gamma_m,$$

with $\gamma_m = e_{2m}^\top R_m^{-1} e_{2m}$. Then it is easy to see that $(J_m^\top S_m^\top J) z_m = 0$. ∎

THEOREM 3.2. *Let M be a Hamiltonian matrix and $S_m \in \mathbb{R}^{2n \times 2m}$, $m \le n$, be a symplectic matrix with $S_m e_1 = q_1$. If S_m satisfies*

$$MS_m = S_m H_m + z_m e_{2m}^\top,$$

where H_m is unreduced J_m-tridiagonal and $z_m \in \mathbb{R}^{2n}$, then $K[M, q_1, 2m]\Pi_m$ has an SR-factorization and rank $(K[M, q_1, 2m]) = 2m$.

PROOF. Since

$$Mq_1 = MS_m e_1 = \left(S_m H_m + z_m e_{2m}^\top\right) e_1 = S_m H_m e_1,$$

and $e_{2m}^{\mathsf{T}} H_m^{i-1} e_1 = 0$ for $i = 1, \ldots, 2m-1$, it is easy to show by induction hypothesis that

$$M^i q_1 = S_m H_m^i e_1.$$

Hence,

$$\begin{aligned} K\left[M, q_1, 2m\right] &= \left[q_1, M q_1, \ldots, M^{m-1} q_1 \mid M^m q_1, \ldots, M^{2m-1} q_1\right] \\ &= S_m \left[e_1, H_m e_1, \ldots, H_m^{m-1} e_1 \mid H_m^m e_1, \ldots, H_m^{2m-1} e_1\right] \\ &= S_m K\left[H_m, e_1, 2m\right]. \end{aligned}$$

Let $\widehat{H}_m = \Pi_m H_m \Pi_m^{\mathsf{T}}$. Then it is easy to verify that $\widehat{H}_m$ is upper Hessenberg and $K[H_m, e_1, 2m] = \Pi_m^{\mathsf{T}} K[\widehat{H}_m, e_1, 2m]$. Therefore,

$$K\left[M, q_1, 2m\right] \Pi_m = S_m \Pi_m^{\mathsf{T}} K\left[\widehat{H}_m, e_1, 2m\right] \Pi_m \equiv S_m R_m,$$

where $R_m \equiv \Pi_m^{\mathsf{T}} K[\widehat{H}_m, e_1, 2m] \Pi_m$. Since H_m is unreduced, using the same argument in the proof of [2, Theorem 3.4, (ii)], one can conclude that R_m is J-triangular and nonsingular. ∎

We comment that the previous two theorems hold for an arbitrary $2n \times 2n$ matrix M. The following existence and uniqueness theorems for J-tridiagonalization of a Hamiltonian matrix follow from the results in [2].

THEOREM 3.3. (EXISTENCE THEOREM). *If all leading principal minors of even dimension of $K[M, q_1, 2n]^{\mathsf{T}} J K[M, q_1, 2n]$ are nonzero, then there exists a symplectic matrix S with $Se_1 = q_1$ such that $H = S^{-1} M S$ is an unreduced J-tridiagonal matrix.*

THEOREM 3.4 (IMPLICIT SYMPLECTIC THEOREM). *Suppose M is a Hamiltonian matrix. Let S and $\widetilde{S}$ be two symplectic matrices with $Se_1 = \widetilde{S} e_1$. If $S^{-1} M S = H$ and $\widetilde{S}^{-1} M \widetilde{S} = \widetilde{H}$, where H and $\widetilde{H}$ are unreduced J-tridiagonal matrices, then there exists a matrix*

$$D = \begin{bmatrix} C & F \\ 0 & C^{-1} \end{bmatrix},$$

where C and F are $n \times n$ diagonal matrices such that $S = \widetilde{S} D$ and $H = D^{-1} \widetilde{H} D$.

With these theorems, we are able to derive a set of two-four-term recurrence formulae for J-tridiagonalization of Hamiltonian matrices. Suppose that, for a given Hamiltonian matrix M, there exists a symplectic matrix S such that $H = S^{-1} M S$ is unreduced J-tridiagonal. With column partitioning, we denote

$$S = \left[q_1, \ldots, q_n \mid q_{n+1}, \ldots, q_{2n}\right] \tag{3.4}$$

and

$$H = \left[\begin{array}{cccc|cccc} a_1 & & & & c_1 & b_1 & & \\ & \ddots & & & b_1 & \ddots & \ddots & \\ & & \ddots & & & \ddots & \ddots & b_{n-1} \\ & & & a_n & & & b_{n-1} & c_n \\ \hline k_1 & & & & -a_1 & & & \\ & \ddots & & & & \ddots & & \\ & & \ddots & & & & \ddots & \\ & & & k_n & & & & -a_n \end{array}\right], \tag{3.5}$$

with $k_i \neq 0$ for $i = 1,\dots,n$, and $b_i \neq 0$ for $i = 1,\dots,n-1$. Upon comparing columns in $MS = SH$, we obtain

$$Mq_i = a_i q_i + k_i q_{n+i}, \tag{3.6}$$

$$Mq_{n+i} = b_{i-1}q_{i-1} + c_i q_i + b_i q_{i+1} - a_i q_{n+i}, \tag{3.7}$$

for $i = 1,\dots,n$ with $b_0 \equiv 0$. By the implicit symplectic Theorem 3.4 and the symplecticity of S, if we require the following conditions hold:

$$\|q_i\|_2 = 1, \qquad q_i \perp q_{n+i}, \tag{3.8}$$

for $i = 1,\dots,n$, then the coefficients a_i, k_i, c_i, b_i and the J-Lanczos vectors q_{n+i}, q_{i+1} can be uniquely determined at the i^{th} step by the following identities. (Note that b_{i-1}, q_{i-1}, and q_i have been obtained in the previous steps.)

$$a_i = q_i^\top M q_i, \qquad \left(= q_{n+i}^\top J M q_i, \text{ later!}\right), \tag{3.9}$$

$$k_i = q_i^\top J M q_i, \tag{3.10}$$

$$q_{n+i} = \frac{(Mq_i - a_i q_i)}{k_i}, \tag{3.11}$$

$$c_i = -q_{n+i}^\top J M q_{n+i}, \tag{3.12}$$

$$r_i = Mq_{n+i} - b_{i-1}q_{i-1} - c_i q_i + a_i q_{n+i}, \tag{3.13}$$

$$b_i = \|r_i\|_2, \qquad \left(= q_{n+i+1}^\top J M q_{n+i}, \text{ later!}\right), \tag{3.14}$$

$$q_{i+1} = \frac{r_i}{b_i}. \tag{3.15}$$

By properly sequencing the formulae, we obtain the following J-tridiagonalization algorithm. Note that there is no loss of generality in choosing b_i to be positive due to Theorem 3.4. The q_i and q_{n+i} are called J-Lanczos vectors.

ALGORITHM 3.1. (J-TRIDIAGONALIZATION). Suppose $M \in \mathbb{R}^{2n\times 2n}$ is a Hamiltonian matrix. For a given nonzero vector q_1 with $\|q_1\|_2 = 1$, this algorithm computes the columns of the symplectic matrix S and entries of $H = S^{-1}MS$ such that H is a J-tridiagonal matrix.

Set $b_0 = 0$, $q_0 = 0$, $i = 1$.
$a_1 = q_1^\top M q_1$
$k_1 = q_1^\top J M q_1$
while $k_i \neq 0$
    $q_{n+i} = (Mq_i - a_i q_i)/k_i$
    $c_i = -q_{n+i}^\top J M q_{n+i}$
    $r_i = Mq_{n+i} - b_{i-1}q_{i-1} - c_i q_i + a_i q_{n+i}$
    $b_i = \|r_i\|_2 \ (= q_{n+i+1}^\top J M q_{n+i})$
    If $b_i = 0$, stop.
    $q_{i+1} = r_i/b_i$
    $i = i + 1$
    $a_i = q_i^\top M q_i \ (= q_{n+i}^\top M q_i)$
    $k_i = q_i^\top J M q_i$
end while

The iteration halts before complete J-tridiagonalization if the initial J-Lanczos vector q_1 is contained in a proper invariant subspace. This is a welcome event. However, the J-tridiagonalization procedure can also halts before the J-Lanczos vector, say, q_{n+j} can be constructed. Such termination does not guarantee an invariant subspace and is called a serious breakdown. The following theorem points out the conditions for these two situations and also proves that the matrix $S_j = [q_1,\dots,q_j \mid q_{n+1},\dots,q_{n+j}]$ constructed by Algorithm 3.1 (if it runs to the j^{th} step) is symplectic.

THEOREM 3.5. *Let $M \in \mathbb{R}^{2n\times 2n}$ be a Hamiltonian matrix and q_1 be a given unit vector. Let*

$$\Delta_j = \det\left(K_j^{\top} J K_j\right), \tag{3.16}$$

where $K_j \equiv K[M, q_1, 2j]$. Then the following statements hold.

(a) *$\Delta_j \neq 0$, $j = 1, \dots, m$, and* $\operatorname{rank}([K_m, M^{2m}q_1]) = 2m$ *for some $1 \le m \le n$ if and only if the J-tridiagonalization Algorithm 3.1 runs until $j = m$, i.e., $b_1 \dots b_{m-1}k_1 \dots k_m \neq 0$ and $b_m = 0$. Moreover, for $j = 1, \dots, m$, we have*

$$MS_j = S_j H_j + r_j e_{2j}^{\top}, \tag{3.17}$$

with

$$H_j = \left[\begin{array}{cccc|cccc} a_1 & & & & c_1 & b_1 & & \\ & \ddots & & & b_1 & \ddots & \ddots & \\ & & \ddots & & & \ddots & \ddots & b_{j-1} \\ & & & a_j & & & b_{j-1} & c_j \\ \hline k_1 & & & & -a_1 & & & \\ & \ddots & & & & \ddots & & \\ & & \ddots & & & & \ddots & \\ & & & k_j & & & & -a_j \end{array}\right] \tag{3.18}$$

and $S_j = [q_1, \dots, q_j \mid q_{n+1}, \dots, q_{n+j}]$ is symplectic, i.e., $S_j^{\top} J S_j = J_j$, and $\operatorname{Range}(S_j) = K(M, q_1, 2j)$.

(b) *$\Delta_j \neq 0$, $j = 1, \dots, m-1$, $\Delta_m = 0$, and* $\operatorname{rank}([K_{m-1}, M^{2m-2}q_1]) = 2m-1$ *for some $1 \le m \le n$ if and only if the Algorithm 3.1 runs until $j = m - 1/2$, i.e., $b_1 \dots b_{m-2}k_1 \dots k_{m-1} \neq 0$, $b_{m-1} \neq 0$, but $k_m = 0$.*

PROOF. Only if for part (a): since $k_1 = q_1^{\top} JMq_1 \neq 0$, from (3.9) and (3.11) we have $q_{n+1}^{\top} J q_{n+1} = -1$. By induction on j, suppose that the J-tridiagonalization iterations have produced $S_j = [q_1, \dots, q_j \mid q_{n+1}, \dots, q_{n+j}]$ for $j < m$, such that

$$\operatorname{Range}(S_j) = K(M, q_1, 2j) \tag{3.19}$$

and

$$S_j^{\top} J S_j = J_j. \tag{3.20}$$

It is easy to see from Algorithm 3.1 that (3.17) holds. Thus,

$$\left(J_j^{\top} S_j^{\top} J\right) MS_j = H_j + \left(J_j^{\top} S_j^{\top} J\right) r_j e_{2j}^{\top}. \tag{3.21}$$

Multiplying (3.6) by $q_{n+i}^{\top} J$ and $q_i^{\top} J$ from the left and using $q_{n+i}^{\top} J q_i = -1$, we have for $i = 1, \dots, j$,

$$a_i = -q_{n+i}^{\top} JMq_i \quad \text{and} \quad k_i = q_i^{\top} JMq_i,$$

as in (3.9) and (3.10), respectively. Also, multiplying (3.7) by $q_{n+i+1}^{\top} J$ from the left and using $q_{n+i+1}^{\top} J q_{i+1} = -1$, we have for $i = 1, \dots, j-1$,

$$b_i = -q_{n+i+1}^{\top} JMq_{n+i},$$

as in (3.14). Now from the J-tridiagonalization Algorithm 3.1 and $S_j^{\top} J S_j = J_j$, it follows that

$$\left(J_j^{\top} S_j^{\top} J\right) MS_j = \left[\begin{array}{c} -q_{n+1}^{\top} \\ \vdots \\ -q_{n+j}^{\top} \\ \hline q_1^{\top} \\ \vdots \\ q_j^{\top} \end{array}\right] JMS_j = H_j. \tag{3.22}$$

Consequently, from (3.21) and (3.22), we have

$$\left(J_j^\top S_j^\top J\right) r_j = 0. \tag{3.23}$$

Since $b_1 \dots b_{m-1} k_1 \dots k_m \neq 0$, from (3.23) and (3.15) we have

$$q_{j+1}^\top J S_j = 0. \tag{3.24}$$

By induction hypothesis (3.19) and from (3.24), it follows that

$$q_{j+1} \perp \text{ Range}(JS_j) = \text{ Range}\left(JK\left[M, q_1, 2j\right]\right). \tag{3.25}$$

Since $q_{j+1} \in$ Range$([q_1, Mq_1, \dots, M^{2j}q_1])$, from $q_{j+1}^\top J q_{j+1} = 0$ and (3.25), we have

$$q_{j+1}^\top J M^{2j} q_1 = 0. \tag{3.26}$$

From (3.13), (3.15), and (3.11), we derive

$$q_{j+1} = \frac{1}{b_j}\left(Mq_{n+j} - b_{j-1}q_{j-1} - c_j q_j + a_j q_{n+j}\right) \in K\left(M, q_1, 2j+1\right) \tag{3.27}$$

and

$$q_{n+j+1} = \frac{1}{k_{j+1}}\left(Mq_{j+1} - a_{j+1}q_{j+1}\right) \in K\left(M, q_1, 2(j+1)\right). \tag{3.28}$$

From (3.25), (3.26), (3.28), and $M^\top J = -JM$, it follows that

$$q_{n+j+1}^\top J M^i q_1 = \frac{1}{k_{j+1}}\left(Mq_{j+1} - a_{j+1}q_{j+1}\right)^\top J M^i q_1 = 0, \tag{3.29}$$

for $i = 0, 1, \dots, 2j-1$. Thus,

$$q_{n+j+1}^\top J S_j = 0. \tag{3.30}$$

From (3.24),(3.30), and $q_{n+j+1}^\top J q_{j+1} = -1$, we can show that

$$S_{j+1} = \left[q_1, \dots, q_{j+1} \mid q_{n+1}, \dots, q_{n+j+1}\right]$$

is symplectic, i.e.,

$$S_{j+1}^\top J S_{j+1} = J_{j+1},$$

and from (3.27) and (3.28), we have Range$(S_{j+1}) = K(M, q_1, j+1)$ with full column rank.

Follow from Theorem 3.2 that $K_m\Pi_m$ has an SR factorization $K_m\Pi_m = S_m R_m$. Hence, it implies that all leading principal minors of even dimension of Δ_m are nonzero [15, Theorem 11]. Moreover, since $b_m = 0$, from (3.13) and (3.14), it follows that rank$([K_m, M^{2m}q_1]) = 2m$.

If for part (a): from assumptions and [15, Theorem 11], it follows that $K_m\Pi_m$ has a SR factorization $K_m\Pi_m = S_m R_m$. From Theorem 3.1 there is an unreduced J_m-tridiagonal matrix H_m such that (3.2) holds. If we require the columns of $S_m \equiv [q_1, \dots, q_m \mid q_{n+1}, \dots, q_{n+m}]$ satisfy $\|q_i\|_2 = 1$ and $q_i \perp q_{n+i}$ for $i = 1, \dots, m$, then the entries o H_m in (3.18) and the vectors q_i, q_{n+i} for $i = 1, \dots, m$ are uniquely determined by (3.9)–(3.15). Thus, we have $b_1 \dots b_{m-1} k_1 \dots k_m \neq 0$. Moreover, since rank$([K_m, M^{2m}q_1]) = 2m$ from (3.13), it follows that $b_m = 0$.

(b) From the proof of (a), we have that $\Delta_j \neq 0$ for $j = 1, \dots, m-1$ and rank$([K_{m-1}, M^{2m-2}q_1])$ $= 2m-1$ if and only if $b_1 \dots b_{m-2} k_1 \dots k_{m-1} \neq 0$, $b_{m-1} \neq 0$. Consequently, from [15, Theorem 11], Theorems 3.1 and 3.2, it follows that $\Delta_m \neq 0$ if and only if $k_m \neq 0$. ∎

This theorem shows that under some mild condition for the initial vector q_1, the J-tridiagonalization Algorithm 3.1 computes a symplectic matrix S_j which partially reduces the Hamiltonian matrix M to a J-tridiagonal matrix H_j. The eigenvalues of H_j are called the J-Ritz values and are used to approximate the eigenvalues of M. The following result provides a computable criteria to check the acceptance of an approximate J-Ritz pair.

THEOREM 3.6. *Suppose that j steps of J-tridiagonalization Algorithm 3.1 have been performed and H_j has no pure imaginary eigenvalue. Let*

$$U_j^{-1} H_j U_j = \left[\begin{array}{ccc|ccc} \theta_1 & & & & & \\ & \ddots & & & & \\ & & \theta_j & & & \\ \hline & & & -\theta_1 & & \\ & & & & \ddots & \\ & & & & & -\theta_j \end{array}\right] = \Lambda_j \tag{3.31}$$

be the J-diagonalization of the J-tridiagonal matrix H_j, where $U_j = [u_1, \dots, u_j \mid u_{j+1}, \dots, u_{2j}]$ is symplectic with $\|u_i\|_2 = \|u_{j+i}\|_2 = 1$, $i = 1, \dots, j$. If $Y_j = [y_1, \dots, y_j \mid y_{n+1}, \dots, y_{n+j}] = S_j U_j$, then the following identities hold:

$$\|My_i - \theta_i y_i\|_2 = |\beta_{j,i}| \tag{3.32}$$

with $\beta_{j,i} = b_j u_{2j,i}$, and

$$\|My_{n+i} + \theta_i y_{n+i}\|_2 = |\beta_{j,j+i}| \tag{3.33}$$

with $\beta_{j,j+i} = b_j u_{2j,j+i}$ for $i = 1, \dots, j$, where $U_j = (u_{k,l})$. Note that y_i and y_{n+i} are called the J-Ritz vectors corresponding to the J-Ritz values θ_i and $-\theta_i$, respectively.

PROOF. Since $MS_j = S_j H_j + r_j e_{2j}^{\top}$, it follows that

$$MS_j U_j = S_j U_j U_j^{-1} H_j U_j + r_j e_{2j}^{\top} U_j,$$

that is,

$$MY_j = Y_j \Lambda_j + r_j \left(e_{2j}^{\top} U_j\right).$$

Thus,

$$My_i = \theta_i y_i + r_j \left(e_{2j}^{\top} U_j e_i\right) \tag{3.34}$$

and

$$My_{n+i} = -\theta_i y_{n+i} + r_j \left(e_{2j}^{\top} U_j e_{j+i}\right), \tag{3.35}$$

for $i = 1, \dots, j$. The results follow by taking 2-norm and recalling that $\|r_j\|_2 = |b_j|$. ∎

From (3.34) and (3.35), we have two residual vectors

$$u_{2j,i} r_j = My_i - \theta_i y_i \tag{3.36}$$

and

$$u_{2j,j+i} \left(J r_j\right)^{\top} = \left(y_{n+i}^H J\right) M - \theta_i \left(y_{n+i}^H J\right). \tag{3.37}$$

Applying the results in [16] to (3.36) and (3.37), it follows that $(\theta_i, y_i, y_{n+i}^H J)$ is an eigen-triplet of $M - E$. The norm of the perturbation E satisfies

$$\|E\|_2 \le |b_j| \max_{i,j} \left\{ \frac{|u_{2j,i}|}{\|y_i\|_2}, \frac{|u_{2j,j+i}|}{\|y_{n+i}\|_2} \right\}. \tag{3.38}$$

Furthermore, from (3.38) and the results in [16], we can estimate the distance from θ_i to an eigenvalue, say, $\lambda^{(i)}$ of M by

$$\begin{aligned} |\lambda^{(i)} - \theta_i| &\le \frac{\|y_{n+i}\|_2 \|y_i\|_2}{|y_{n+i}^H J y_i|} \|E\|_2 + O\left(\|E\|_2^2\right) \\ &\le \frac{|b_j|}{|u_{j+i}^H J_j u_i|} \max\left\{|u_{2j,i}|\, \|y_{j+i}\|_2, |u_{2j,j+i}|\, \|y_i\|_2\right\} + O\left(\|E\|_2^2\right) \\ &\le \frac{|b_j|\, \|S_j\|_2}{|u_{j+i}^H J_j u_i|} \max\left\{|u_{2j,i}|, |u_{2j,j+i}|\right\} + O\left(\|E\|_2^2\right). \end{aligned} \tag{3.39}$$

We conclude this section by summarizing the J-Lanczos method in the following algorithm.

ALGORITHM 3.2. (J-LANCZOS). Given a Hamiltonian matrix $M \in \mathbb{R}^{2n\times 2n}$. For a given unit vector q_1, this algorithm computes the columns of the symplectic matrix S_j and entries of the $2j \times 2j$ J-tridiagonal matrix H_j such that $MS_j = S_jH_j + r_je_{2j}^\top$ using Algorithm 3.1. Then the algorithm computes the J-Ritz values and J-Ritz vectors to approximate the extreme eigen-pairs of M. Stop criterion is based on (3.32) and (3.33).

Given $q_1 \neq 0$ with $\|q_1\|_2 = 1$ and tolerance $\epsilon > 0$.
Set $b_{-1} = 0$, $q_{-1} = 0$, $j = 1$.
while $k_j \neq 0$ and $b_j \neq 0$
  Compute $a_j, k_j, q_{n+j}, c_j, b_j$, and q_{j+1} by Algorithm 3.1.
  Compute $U_j^{-1}H_jU_j = \Lambda_j$ as in (3.31) by using symplectic QR like algorithm [2] or QR algorithm.
  for $i = 1, \ldots, j$,
    if $|\beta_{j,i}| \leq \epsilon$ and $|\beta_{j,j+i}| \leq \epsilon$
    accept (θ_i, y_i), $(-\theta_i, y_{n+i})$ and their conjugate pairs as the desired eigen-pairs.
  end for
  if the desired eigen-pairs are satisfied, then stop
  else $j = j + 1$
end while

In the next section, we present an error bound for the J-Ritz values obtained from the J -Lanczos algorithm and prove a variant Paige-type theorem showing that convergence of the J-Ritz pairs implies loss of symplecticity.

4. ERROR BOUND ANALYSIS AND PAIGE-TYPE THEOREM

Let H_n be the J-tridiagonal matrix obtained from applying n iterations of the J-Lanczos algorithm to a Hamiltonian matrix M and H_m be a J-principal submatrix of H_n. Hereinafter, $\wp^k$ denotes the set of polynomials of degree less than or equal to k. The following lemma can be obtained immediately.

LEMMA 4.1. *Let e_i denotes the i^{th} column of identity matrix of suitable dimension. Then, for $i = 1, \ldots, 4m-1$, the following identities hold.*

(i) $e_1^\top H_n^i e_1 = e_1^\top H_m^i e_1$.
(ii) $e_{n+1}^\top H_n^i e_{n+1} = e_{m+1}^\top H_m^i e_{m+1}$.
(iii) $e_{n+1}^\top H_n^i e_1 = e_{m+1}^\top H_m^i e_1$.

For simplicity, we assume that both H_n and H_m here are J-diagonalizable, that is, $H_n = X^H \Lambda Y$ and $H_m = P^H \Theta Q$, where

$$\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n \mid -\lambda_1, \ldots, -\lambda_n) \equiv \begin{bmatrix} \Lambda_1 & 0 \\ 0 & -\Lambda_1 \end{bmatrix} \tag{4.1}$$

and

$$\Theta = \operatorname{diag}(\theta_1, \ldots, \theta_m \mid -\theta_1, \ldots, -\theta_m) \equiv \begin{bmatrix} \Theta_1 & 0 \\ 0 & -\Theta_1 \end{bmatrix}. \tag{4.2}$$

Let $E_n = [e_1, e_{n+1}] \in \mathbb{R}^{2n\times 2}$ and $E_m = [e_1, e_{m+1}] \in \mathbb{R}^{2m\times 2}$. With the decompositions of H_n and H_m above and apply Lemma 4.1, one can verify that

$$E_n^\top f(H_n)E_n = E_m^\top f(H_m)E_m,$$

for all $f \in \wp^{4m-1}$. This implies

$$\begin{bmatrix} x_1^H \\ x_{n+1}^H \end{bmatrix} \begin{bmatrix} f(\Lambda_1) & 0 \\ 0 & f(-\Lambda_1) \end{bmatrix} [y_1 \quad y_{n+1}] = \begin{bmatrix} p_1^H \\ p_{m+1}^H \end{bmatrix} \begin{bmatrix} f(\Theta_1) & 0 \\ 0 & f(-\Theta_1) \end{bmatrix} [q_1 \quad q_{m+1}]. \tag{4.3}$$

Here x_i, y_i, p_i, and q_i are the i^{th} column of X, Y, P, and Q, respectively. Denote $x_1 \equiv \begin{bmatrix} x_1^{(1)} \\ x_1^{(2)} \end{bmatrix}$, where $x_1^{(1)}, x_1^{(2)} \in \mathbb{R}^n$. Using the similar notations for y_1, p_1, q_1, x_{n+1}, y_{n+1}, p_{n+1}, and q_{n+1}, equation (4.3) becomes

$$\begin{aligned} &\left(x_1^{(1)} + x_1^{(2)}\right)^H (f(\Lambda_1) + f(-\Lambda_1)) \left(y_1^{(1)} + y_1^{(2)}\right) \\ &\qquad = \left(p_1^{(1)} + p_1^{(2)}\right)^H (f(\Theta_1) + f(-\Theta_1)) \left(q_1^{(1)} + q_1^{(2)}\right), \end{aligned} \tag{4.4}$$

for all $f \in \wp^{4m-1}$. By the property of $f(\Lambda_1) + f(-\Lambda_1)$, there is an even polynomial g with degree $\le 4m-2$ such that $g(\Lambda_1) = f(\Lambda_1) + f(-\Lambda_1)$. Hence, (4.4) can be rewritten as

$$\sum_{i=1}^{n} g(\lambda_i) (\bar{x}_{i,1} + \bar{x}_{n+i,1}) (y_{i,1} + y_{n+i,1}) = \sum_{i=1}^{m} g(\theta_i) (\bar{p}_{i,1} + \bar{p}_{m+i,1}) (q_{i,1} + q_{m+i,1}). \tag{4.5}$$

Now, let $\sigma_1 = \{\lambda_2, \dots, \lambda_n\}$ and $\hat{\sigma}_1 = \{\theta_2, \dots, \theta_m\}$. Suppose $\sigma_1 \cup \hat{\sigma}_1 = S_1 \cup S_2$ with $S_1 \cap S_2 = \phi$. Define

$$\delta_1(S_2) = \max\left\{ |x^2 - \theta_1^2| \prod_{\mu \in S_2} \frac{|x^2 - \mu^2|}{|\lambda_1^2 - \mu^2|} : x \in \sigma_1 \cup \hat{\sigma}_1 \right\} \tag{4.6}$$

and

$$\varepsilon_1^{(k)}(S_1) = \inf_{p \in \wp^k,\, p(\lambda_1^2) = 1} \max_{x \in S_1} \left| p\left(x^2\right) \right|. \tag{4.7}$$

With above definitions and notations, we establish an error bound for the J-Ritz values.

THEOREM 4.2. *Assume that* $|\lambda_1 - \theta_1| = \min_{1 \le j \le m} |\lambda_1 - \theta_j|$. *If* $s = |S_2| \le m - 2$ *holds, then*

$$\begin{aligned} |\lambda_1 - \theta_1| \le{}& \frac{\varepsilon_1^{(2m-s-2)}(S_1)\delta_1(S_2)}{|\lambda_1 + \theta_1| \, |x_{1,1} + x_{n+1,1}| \, |y_{1,1} + y_{n+1,1}|} \\ &\times \left(\sum_{i=2}^{n} |x_{i,1} + x_{n+i,1}| \, |y_{i,1} + y_{n+i,1}| + \sum_{i=2}^{m} |p_{i,1} + p_{m+i,1}| \, |q_{i,1} + q_{m+i,1}| \right). \end{aligned} \tag{4.8}$$

PROOF. Let

$$g(x) = \left(x^2 - \theta_1^2\right) p\left(x^2\right) \prod_{\mu \in S_2} \left(x^2 - \mu^2\right),$$

where $p \in \wp^{2m-s-2}$ with $p(\lambda_1^2) = 1$. Substituting $g(x)$ into (4.5), we obtain

$$\begin{aligned} &\left(\lambda_1^2 - \theta_1^2\right) p\left(\lambda_1^2\right) \prod_{\mu \in S_2} \left(\lambda_1^2 - \mu^2\right) (\bar{x}_{1,1} + \bar{x}_{n+1,1}) (y_{1,1} + y_{n+1,1}) \\ &\quad = - \sum_{\lambda_i \in S_1} \left(\lambda_i^2 - \theta_1^2\right) p\left(\lambda_i^2\right) \prod_{\mu \in S_2} \left(\lambda_i^2 - \mu^2\right) (\bar{x}_{i,1} + \bar{x}_{n+i,1}) (y_{i,1} + y_{n+i,1}) \\ &\qquad + \sum_{\theta_i \in S_1} \left(\theta_i^2 - \theta_1^2\right) p\left(\theta_i^2\right) \prod_{\mu \in S_2} \left(\theta_i^2 - \mu^2\right) (\bar{p}_{i,1} + \bar{p}_{m+i,1}) (q_{i,1} + q_{m+i,1}). \end{aligned}$$

From (4.6), we have

$$\begin{aligned} |\lambda_1 - \theta_1| \le{}& \frac{1}{|\lambda_1 + \theta_1| \, |x_{1,1} + x_{n+1,1}| \, |y_{1,1} + y_{n+1,1}|} \max_{x \in S_1} p\left(x^2\right) \delta_1(S_2) \\ &\times \left(\sum_{i=2}^{n} |x_{i,1} + x_{n+i,1}| \, |y_{i,1} + y_{n+i,1}| + \sum_{i=2}^{m} |p_{i,1} + p_{m+i,1}| \, |q_{i,1} + q_{m+i,1}| \right). \end{aligned}$$

Since $p \in \wp^{2m-s-2}$ with $p(\lambda_1^2) = 1$ is arbitrary, from definition (4.7) we get the error bound (4.8). ∎

We comment that (4.8) gives a new bound for $|\lambda_1 - \theta_1|$ when compared with the bound in [13]. Further analysis of the magnitude of the right-hand side in (4.8) is referred to [13] for details.

For the roundoff error analysis in the following, we prove a variant Paige-type theorem [14] which shows that the convergence of a J-Ritz pair implies loss of symplecticity and that duplicated J-Ritz pairs can occur.

Suppose that by the end of the j^{th} step, the J-Lanczos algorithm has produced S_j, the matrix of J-Lanczos vectors H_j, the J-tridiagonal matrix embodying the two-four-term recurrence formulae (3.9)–(3.15), and the residual vector r_j. For convenience of discussion, we suppose that H_j has no pure imaginary eigenvalue. If the effects of roundoff errors are taken into account, then two fundamental relations can be formulated by

$$MS_j - S_jH_j = r_je_{2j}^\top + F_j \tag{4.9}$$

and

$$J_j - S_j^\top JS_j = C_j^\top - C_j, \tag{4.10}$$

where F_j and C_j are the corresponding roundoff error matrices.

Suppose that the coefficients a_i, k_j, and c_i determined by (3.9), (3.10), and (3.12), respectively, are locally arithmetic exact. In addition, we assume that the following conditions are maintained in the J-Lanczos algorithm.

(A1) Local orthogonality: $q_i^\top q_{n+i} = 0$, $i = 1, \dots, j$.
(A2) J-unity and unity: $q_{n+i}^\top Jq_i = -1$ and $q_i^\top q_i = 1$, $i = 1, \dots, j$.
(A3) Local symplecticity: $q_{n+i}^\top Jq_{i-1} = 0$ and $q_{n+i}^\top Jq_{i+1} = 0$, $i = 1, \dots, j-1$.
(A4) The J-diagonalization of H_j is exact, namely there is a $2j \times 2j$ symplectic matrix U_J such that $U_j^{-1}H_jU_j = \Theta_j = \operatorname{diag}(\theta_1, \dots, \theta_j \mid -\theta_1, \dots, -\theta_j)$.

Let the matrix C_j in (4.10) be partitioned by

$$C_j = \begin{bmatrix} C_{11}^{(j)} & C_{12}^{(j)} \\ C_{21}^{(j)} & C_{22}^{(j)} \end{bmatrix}, \tag{4.11}$$

where $C_{i,k}^{(j)} \in \mathbb{R}^{n\times n}$, $i, k = 1, 2$, are upper triangular. From the skew symmetry of $J_j - Sj^\top JS_j$ and Assumptions (A2) and (A3), it is easy to check that all $C_{i,k}^{(j)}$, $i, k = 1, 2$, are strictly upper triangular. In addition, both $C_{12}^{(j)}$ and $C_{21}^{(j)}$ have zero subdiagonals. Multiplying (3.7) by $q_i^\top J$ from the left, for $i = 1, \dots, j$, and using (A2), one has

$$q_i^\top JMq_{n+i} = b_{i-1}q_i^\top Jq_{i-1} + b_iq_i^\top Jq_{i+1} - a_i. \tag{4.12}$$

Applying induction hypothesis that $q_i^\top Jq_{i-1} = 0$ to (4.12) and from (3.9), it follows that $q_i^\top Jq_{i+1} = 0$. We can conclude that $C_{11}^{(j)}, C_{12}^{(j)}, C_{21}^{(j)}$ are strictly upper triangular with zero first subdiagonals and $C_{22}^{(j)}$ is strictly upper triangular.

The Paige-type theorem is presented and proved in the following theorem.

THEOREM 4.3. *Suppose that H_j, S_j, and r_j constructed by the J-Lanczos algorithm satisfy (4.9) and (4.10). Suppose Assumptions (A1)–(A4) hold. Let $K_j \in \mathbb{R}^{2j\times 2j}$ be a J-strictly upper triangular matrix such that*

$$S_j^\top JF_j - F_j^\top J^\top S_j = K_j - K_j^\top, \tag{4.13}$$

and let

$$\Gamma_j = U_j^H K_j U_j. \tag{4.14}$$

Then, for $i = 1, \ldots, j$, the J-Ritz vector $y_i = S_j u_i$, and $y_{j+i} = S_j u_{j+i}$ satisfy

$$\frac{q_{j+1}^{\top} J y_i = \gamma_{\hat{\imath},j+i}}{b_j(u_{2j,j+i}) \equiv \gamma_{\hat{\imath},j+i}/\beta_{j,j+i}} \tag{4.15}$$

and

$$\frac{q_{j+1}^{\top} J y_{j+i} = \gamma_{j+\hat{\imath},i}}{b_j(u_{2j,i}) \equiv \gamma_{j+\hat{\imath},i}/\beta_{j,i}}, \tag{4.16}$$

for some $\hat{\imath}$ with $1 \le \hat{\imath} \le j$, where $\gamma_{l,k}$ and $u_{l,k}$ denote the $(l,k)^{\text{th}}$ entry of Γ_j and U_j, respectively. Moreover, for $i \ne k$, $i, k \in I_j \equiv \{i \mid I_m(\theta_i) \ge 0, i = 1, \ldots, j\}$, the following identity hold:

$$(\theta_k - \theta_i)\, \bar{y}_{j+i}^{H} J y_k = -\gamma_{\hat{k},j+k}\left(\frac{u_{2j,j+i}}{u_{2j,j+k}}\right) + \gamma_{j+\hat{\imath},i}\left(\frac{u_{2j,k}}{u_{2j,i}}\right) + \gamma_{\hat{k},j+i} - \gamma_{j+\hat{\imath},k}, \tag{4.17}$$

for some $\hat{\imath}$, $\hat{k}$ with $1 \le \hat{\imath}, \hat{k} \le j$,

PROOF. Premultiply (4.9) by $S_j^{\top} J$ to get

$$S_j^{\top} J M S_j = S_j^{\top} J S_j H + S_j^{\top} J r_j e_{2j}^{\top} + S_j^{\top} J F_j. \tag{4.18}$$

To eliminate $S_j^{\top} J M S_j$ from (4.18), one can take transpose and then apply (4.10) and (4.13) to derive

$$\begin{aligned} & e_{2j}\left(r_j^{\top} J^{\top} S_j\right) - \left(S_j^{\top} J r_j\right) e_{2j}^{\top} \\ &= \left(-J_j + S_j^{\top} J S_j\right) H_j + H_j^{\top}\left(J_j^{\top} - S_j^{\top} J_j^{\top} S_j\right) + S_j^{\top} J F_j - F_j^{\top} J^{\top} S_j \\ &= \left(C_j H_j + H_j^{\top} C_j\right) - \left(C_j^{\top} H_j + H_j^{\top} C_j^{\top}\right) + K_j - K_j^{\top}. \end{aligned} \tag{4.19}$$

Since H_j is J-tridiagonal and C_j has the special form as discussed above, it is easily seen that each submatrix of $C_j H_j$ and $H_j^{\top} C_j$ is strictly upper triangular according to the 4-block partition shown in (4.11). Similarly, each submatrix of $C_j^{\top} H_j$ and $H_j^{\top} C_j^{\top}$ is strictly lower triangular. Furthermore, since K_j is J-strictly upper triangular, we have

$$\left(S_j^{\top} J^{\top} r_j\right) e_{2j}^{\top} = C_j H_j + H_j^{\top} C_j + K_j. \tag{4.20}$$

From Assumption (A4), we have

$$H_j u_i = \theta_i u_i, \quad H_j u_{j+i} = -\theta_i u_{j+i}, \qquad i = 1, \ldots, j. \tag{4.21}$$

For convenience, we denote $\bar{u}_i$, $\bar{\theta}_i$, $\bar{u}_{j+i}$, and $-\bar{\theta}_i$ by $u_{\hat{\imath}}$, $\theta_{\hat{\imath}}$, $u_{j+\hat{\imath}}$, and $-\theta_{\hat{\imath}}$, respectively, for some $\hat{\imath}$ with $1 \le \hat{\imath} \le j$ and let

$$y_i = S_j u_i, \quad y_{j+i} = S_j u_{j+i}, \qquad i = 1, \ldots, j. \tag{4.22}$$

Premultiplying $u_{j+\hat{\imath}}^{H}$ and postmultiplying u_i to equation (4.20), and from (3.15), (4.14), (4.21), and (4.22), one can derive

$$q_{j+1}^{\top} J y_{j+i} \beta_{j,i} = \theta_i u_{j+\hat{\imath}}^{H} C_j u_i - \theta_i u_{j+\hat{\imath}}^{H} C_j u_i + u_{j+\hat{\imath}}^{H} K_j u_i \equiv \gamma_{j+\hat{\imath},i}. \tag{4.23}$$

This proves (4.15). To prove (4.16), it is sufficient to consider by premultiplying $\bar{u}_{j+i}^{H}$ and postmultiplying u_k to (4.20).

Next, by premultiplying $\bar{u}_{j+i}^{H}$ and postmultiplying u_k to (4.18), for $i \ne k$, $i, k \in I_j \equiv \{i \mid \operatorname{Im}(\theta_i) \ge 0, i = 1, \ldots, j\}$, one obtains

$$\bar{u}_{j+i}^{H} S_j^{\top} J M S_j u_k = \theta_k \bar{u}_{j+i}^{H} S_j^{\top} J S_j u_k + \bar{u}_{j+i}^{H} S_j^{\top} J q_{j+1} b_j e_{2j}^{\top} u_k + \bar{u}_{j+i}^{H} S_j^{\top} J F_j u_k. \tag{4.24}$$

From Theorem 3.6 and (4.22), we have

$$\bar{y}_{j+i}^H JMy_k = \theta_k \bar{y}_{j+i}^H Jy_k + \bar{y}_{j+i}^H Jq_{j+1}\beta_{j,k} + \bar{y}_{j+i}^H JF_j u_k. \tag{4.25}$$

Similarly,

$$\bar{y}_k^H JMy_{j+i} = -\theta_i \bar{y}_k^H Jy_{j+i} + \bar{y}_k^H Jq_{j+1}\beta_{j,j+i} + \bar{y}_k^H JF_j u_{j+i}. \tag{4.26}$$

But

$$\bar{y}_k^H JMy_{j+i} = y_k^\top M^\top J^\top y_{j+i} = \left(y_k^\top M^\top J^\top y_{j+i}\right)^\top = \bar{y}_{j+i}^H JMy_k.$$

Subtracting (4.25) from (4.26), one has

$$(\theta_k - \theta_i)\bar{y}_{j+i}^H Jy_k = -q_{j+1}^\top Jy_k\beta_{j,j+i} + q_{j+1}^\top Jy_{j+i}\beta_{j,k} + \bar{y}_k^H JF_j u_{j+i} - \bar{y}_{j+i}^H JF_j u_k.$$

From (4.13)–(4.15), we have

$$\begin{aligned}(\theta_k - \theta_i)\bar{y}_{j+i}^H Jy_k &= -\frac{\gamma_{\hat{k},j+k}}{\beta_{j,j+k}}\beta_{j,j+i} + \frac{\gamma_{j+\hat{i},i}}{\beta_{j,i}}\beta_{j,k} + \bar{u}_k^H \left(S_j^\top JF_j - F_j^\top J^\top S_j\right) u_{j+i} \\ &= -\gamma_{\hat{k},j+k}\left(\frac{u_{2j,j+i}}{u_{2j,j+k}}\right) + \gamma_{j+\hat{i},i}\left(\frac{u_{2j,k}}{u_{2j,i}}\right) + \gamma_{\hat{k},j+i} - \gamma_{j+\hat{i},k},\end{aligned}$$

which proves (4.17). ∎

REMARKS. To conclude this section, we summarize the following comments.

(a) Equation (4.15) and (4.16) shows that in the J-Lanczos algorithm, if for some i, y_i approximates a desired eigenvector, i.e., $|\beta_{j,j+i}|$ is sufficiently small, then the quantity $q_{j+1}^\top Jy_i = \gamma_{j+\hat{i},i}/\beta_{j,j+i} \approx O(\epsilon)/O(\epsilon) \approx O(1)$. That is, $q_{j+1}^\top JS_j u_i \approx O(1)$. This means that the symplecticity between q_{j+1} and JS_j is lost. Hence, a resymplectization process after the j^{th} step should be performed.

(b) In view of equation (4.17), since the right-hand side is fairly small if $\theta_k \approx \theta_i$, the quantity $\bar{y}_{j+i}^H Jy_k$, and consequently, $y_{j+i}^H Jy_k$ may not be small. Thus, $y_{j+i} \approx y_{j+k}$ can happen. This means that in the J-Lanczos algorithm, a duplicated production of convergent J-Ritz pair is possible.

5. A PRACTICAL EXAMPLE AND NUMERICAL RESULTS

In this section, we discuss the numerical aspects on applying the proposed J-Lanczos algorithm to solving high-order Riccati equation arising from position and velocity control for a string of high-speed vehicles [8]. The matrices in the associated Riccati equation $-XNX + XA + A^\top X + K = 0$ in this practical example are given by

$$\begin{aligned} N &= \operatorname{diag}(1,0,1,0,\dots,0,1), \\ K &= \operatorname{diag}(0,10,0,10,\dots,10,0), \end{aligned}$$

and

$$A = \begin{bmatrix} A_{11} & A_{12} & & & & \\ & A_{22} & A_{23} & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & & 0 \\ & & & & A_{m-1,m-1} & -1 \\ & & & & 0 \quad 0 & -1 \end{bmatrix},$$

with

$$A_{ii} = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad A_{i,i+1} = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix}.$$

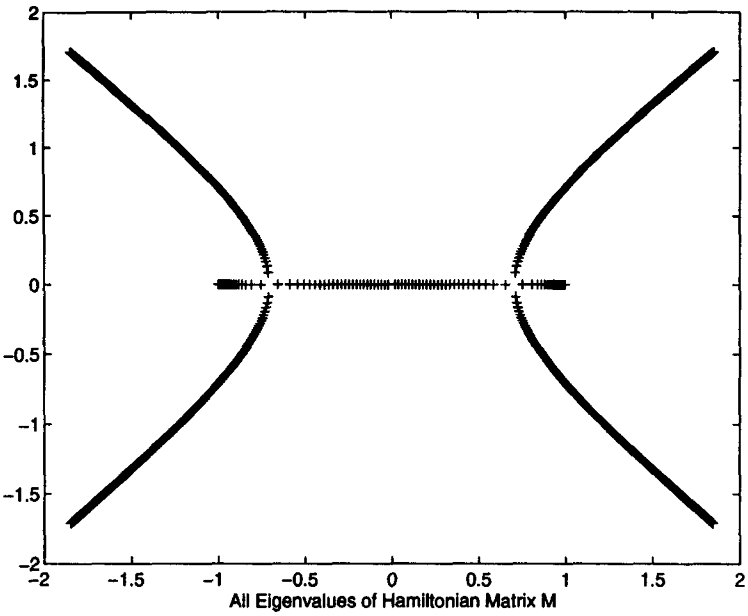

Figure 1. Distribution of all eigenvalues of the associated Hamiltonian matrix on the complex plane in our test problem.

For a string of $m = 501$ vehicles, it is necessary to solve the Riccati equation of order $n = 2m - 1 = 1001$, and consequently, the associated Hamiltonian matrix is of order 2002 with eigenvalue distribution as shown in Figure 1. The J-Lanczos Algorithm 3.2 incorporated with shift and invert techniques is implemented in MATLAB to solve the problem on a Sun SPARC-10 workstation with 32 MB of main memory.

Since all Lanczos-type algorithms converge fast for approximating some extreme eigenvalues, but not all eigenvalues. For solving Riccati equations, one has to compute all eigen-pairs of the associated Hamiltonian matrix. Hence, some shift and invert technique has to be considered and incorporated into the J-Lanczos algorithm.

An important consideration is how to preserve the Hamiltonian structure of the shift-inverted transformed matrices. Since a transformation matrix can usually be represented by a rational function of matrix M, say,

$$f(M) = \sum_{j=-\infty}^{\infty} c_j M^j, \qquad c_j \in \mathbb{C}, \text{ for all } j. \tag{5.1}$$

To preserve the Hamiltonian structure, we require that $(Jf(M))^H = Jf(M)$. Since M is Hamiltonian $(JM^j)^H = (-1)^{j+1} JM^j$. Hence,

$$J \sum_j \bar{c}_j (-1)^{j+1} M^j = J \sum_j c_j M^j.$$

Write $c_j = \alpha_j + i\beta_j$, where $i = \sqrt{-1}$ and $\alpha_j, \beta_j \in \mathbb{R}$ for all j. By comparing the coefficients,

$$c_j = \begin{cases} \alpha_j, & \text{if } j \text{ is odd,} \\ i\beta_j, & \text{if } j \text{ is even.} \end{cases}$$

Since a Hamiltonian matrix is first reduced to a J-tridiagonal matrix by the J-Lanczos algorithm using real symplectic similarity transformations, it requires that all even term coefficients of the considered rational matrix function $f(M)$ to be zero. For practical implementation, the following three types of analytic matrix function are considered.

1. Choose
$$f(M) = M^{-1},$$
whenever the desired eigenvalues are of the smallest modulus.
2. Choose
$$\begin{aligned} f(M) &= \left(M \pm \delta^2 M^{-1}\right)^{-1} \\ &= \left(M^2 \pm \delta^2 I\right)^{-1} M, \end{aligned}$$
where $\delta \in \mathbb{R}$ and $\delta > 0$. The choice of $\pm$ sign depends on the desired eigenvalues whether are close to the numbers $\pm\delta$ or pure imaginary numbers $\pm i\delta$.
3. Choose
$$\begin{aligned} f(M) &= \left(M^3 + bM + cM^{-1}\right)^{-1} \\ &= \left(M^4 + bM^2 + cI\right)^{-1} M, \end{aligned}$$
where $b = 2(\beta^2 - \alpha^2)$ and $c = (\alpha^2 + \beta^2)^2$ with $\alpha, \beta \in \mathbb{R}$, $\alpha, \beta > 0$, whenever the desired eigenvalues are close to complex numbers $\pm(\alpha \pm i\beta)$.

Table 1. Summary of numerical results with real shift δ.

J-Lanczos Iteration Numbers	Total Number of Shifts	Total Number of Eigenvalues	Total Time	Time Per Eigenvalue
j	δ	λ	sec.	sec.
20	5	58	1706.0	29.4
30	5	118	2499.2	21.6
40	4	112	3429.5	30.6
50	3	122	3623.0	29.7

Table 2. Summary of numerical results with complex shifts $\pm(\alpha \pm \beta i)$.

J-Lanczos Iteration Numbers	Total Number of Shifts	Total Number of Eigenvalues	Total Time	Time Per Eigenvalue
j	$\pm(\alpha \pm \beta i)$	λ	sec.	sec.
30	60	1276	77003.0	60.3
50	40	1560	65992.8	42.3
80	10	902	19760.3	21.9

In this test problem, it is easy to check that (A, B) is stabilizable and (C, A) is detectable. Thus, the corresponding Hamiltonian matrix M has no pure imaginary eigenvalues. Therefore, the shifts considered are either real or complex only. In Table 1, we summarize the results of the J-Lanczos algorithm with real shift δ and the Hamiltonian transformation matrix $(M+\delta M^{-1})^{-1}$. In Table 2, we summarize the results for complex shift $\pm(\alpha \pm i\beta)$ to $(M^3 + bM + cM^{-1})^{-1}$. In the implementation, complex shifts along a straight line on the complex plane with an argument angle $\frac{\pi}{4}$ were actually performed. As one can see from the results that the J-Lanczos algorithm is most efficient when 30 iterations are taken with real shifts and 80 iterations with complex shifts. Besides, we observed that the relation between the number of convergent eigenvalues ν and the number of J-Lanczos iterations j is approximately $\nu = 1.4 \times j - 25$.

Some observations and comments are in order.

(a) The dominant computations of this shift-inverted J-Lanczos method is the LU-factorizations of the shift matrices and the associated triangular solvers. In this particular test suite, the Hamiltonian matrix and all shift matrices are banded. Therefore, fast band factorization routine and storage format are easy to implement.

(b) It requires more J-Lanczos iterations in the complex shift cases than the real shift ones because the complex eigenvalues are more clustered than the real eigenvalues.

(c) In the implementations, we use $\varepsilon = 10^{-10}$ (see Algorithm 3.2) as the stopping criteria and obtained the computed solution $\hat{X}$ to the algebraic Riccati equation with residual $\| -\hat{X}N\hat{X} + \hat{X}A + A^{\mathsf{T}}\hat{X} + K\|_2 = 1.6 \times 10^{-5}$.

6. CONCLUSIONS

In this paper, we derived the J-Lanczos algorithm from the J-tridiagonalization procedure of a Hamiltonian matrix using symplectic similarity transformations. We also gave a detailed analysis on the convergence behavior of the J-Lanczos algorithm and presented error bound analysis and Paige-type theorem.

For very large and sparse Hamiltonian matrices, the general QR method [1] or structure-preserving numerical methods proposed in [2,3,5–7] for computing the stable invariant subspaces become inadequate when storage and computational effort are big concern. For example, one would not be able to solve the position and velocity control problem discussed in this paper on a regular workstation with any symplectic QR-type algorithms because of the storage constraint. Alternatively, the proposed structure-preserving J-Lanczos method can efficiently solve this problem with high accuracy.

Finally, we would like to comment that, unlike the serial oriented symplectic QR-type algorithms, parallel implementation of J-Lanczos algorithm with different shift-invert steps is straightforward.

REFERENCES

1. A. Laub, A Schur method for solving algebraic Riccati equations, *IEEE Trans. Auto. Contr.* **24**, 913–921, (1979).
2. A. Bunse-Gerstner and V. Mehrmann, A symplectic QR like algorithm for the solution of real algebraic Riccati equation, *IEEE Trans. on Auto. Contr.* **31**, 1104–1113, (1986).
3. A. Bunse-Gerstner, R. Byers and V. Mehrmann, A chart of numerical methods for structured eigenvalue problems, *SIAM J. Matr. Anal. Appl.* **13**, 419–453, (1992).
4. C.F. Van Loan, A symplectic method for approximating all the eigenvalues of a Hamiltonian matrix, *Lin. Alg. Appl.* **61**, 233–251, (1984).
5. G. Ammar and V. Mehrmann, On Hamiltonian and symplectic Hessenberg forms, *Lin. Alg. Appl.* **149**, 55–72, (1991).
6. H.-G. Xu, *Solving Algebraic Riccati Equation via Skew-Hamiltonian Matrices*, Ph.D. Thesis, Fudan Univ., Shanghai, China, (1991).
7. R.V. Patel, Z. Lin and P. Misra, Computation of stable invariant subspaces of Hamiltonian matrices, *SIAM J. Matr. Anal. Appl.* **15**, 284–298, (1994).
8. M. Athans, W.S. Levine and A. Levis, A system for the optimal and suboptimal position and velocity control for a string of high-speed vehicles, In *Proc. Fifth Int. Alalogue Computation Meetings*, Lausanne, Switzerland, (September 1967).
9. J. Wall, *Control and Estimation for Large-Scale Systems Having Spatial Symmetry*, Ph.D. Dissertation, Massachusetts Inst. of Tech. Electron. Syst. Lab. Rep. ESL-TH-842, (August 1978).
10. R.W. Freund, Lanczos-type algorithms for structured non-Hermitian eigenvalue problems, In *Proceeding of the Cornelius Lanczos International Centenary Conference*, (Edited by J.D. Brown, M.T. Chu, D.C. Ellison and R.J. Plemmons), (1994).
11. R.W. Freund and V. Mehrmann, A symplectic look-ahead Lanczos algorithm for the Hamiltonian eigenvalue problem, AT&T Numerical Analysis Manuscript, Bell Laboratories, Murray Hill, 1994.
12. R.W. Freund and H. Zha, Simplifications of the nonsymmetric Lanczos process and a new algorithm for Hermitian indefinite linear system, AT&T Numerical Analysis Manuscript, Bell Laboratories, Murray Hill, (1994).
13. Q. Ye, A convergence analysis for nonsymmetric Lanczos algorithms, *Math. Comput.* **56**, 677–691, (1991).

14. C.C. Paige, *The Computation of Eigenvalues and Eigenvectors of Very Large Sparse Matrices*, Ph.D. Thesis, Univ. of London, (1971).
15. L. Elsner, On some algebraic problems in connection with general eigenvalue problems, *Lin. Alg. Appl.* **26**, 123–138, (1979).
16. W. Kahan, B.N. Parlett and E. Jiang, Residual bounds on approximate eigensystems of nonnormal matrices, *SIAM J. Numer. Anal.* **19**, 470–484, (1982).