Electronic Transactions on Numerical Analysis.
Volume 43, pp. 213-222, 2015.

ISSN 1068-9613.

ETNA
Kent State University
http://etna.math.kent.edu

A SUBSPACE ITERATION FOR SYMPLECTIC MATRICES*

ALEXANDER MALYSHEV†, MILOUD SADKANE‡, AND AHMED SALAM§

Dedicated to Lothar Reichel on the occasion of his 60th birthday

Abstract. We study the convergence behavior of an orthogonal subspace iteration for matrices whose spectrum is partitioned into three groups: the eigenvalues inside, outside, and on the unit circle. The main focus is on symplectic matrices. Numerical experiments are provided to illustrate the theory.

Key words. symplectic matrix, subspace iteration, invariant subspace

AMS subject classifications. 15A21, 65F15

1. Introduction. Every square matrix $W \in \mathbb{C}^{N\times N}$ can be block-factorized into the form

$$W = \begin{bmatrix} X_\infty & X_1 & X_0 \end{bmatrix} \begin{bmatrix} W_\infty & & \\ & W_1 & \\ & & W_0 \end{bmatrix} \begin{bmatrix} X_\infty & X_1 & X_0 \end{bmatrix}^{-1}, \tag{1.1}$$

where one or two matrices among W_∞, W_1, and W_0 may be empty. The spectrum of W consists of at most three groups: (i) eigenvalues $\lambda_1, \ldots, \lambda_{N_\infty}$ of W_∞ outside the unit circle, (ii) eigenvalues $\lambda_{N_\infty+1}, \ldots, \lambda_{N_\infty+N_1}$ of W_1 on the unit circle, and (iii) eigenvalues $\lambda_{N_\infty+N_1+1}, \ldots, \lambda_{N_\infty+N_1+N_0}$ of W_0 inside the unit circle. The corresponding invariant subspaces, which we denote by $\mathcal{X}_\infty = \text{range}(X_\infty)$, $\mathcal{X}_1 = \text{range}(X_1)$, and $\mathcal{X}_0 = \text{range}(X_0)$, are of particular importance in applications such as optimal control [6, Chapter 14], [8, Chapter 15], and the theory of parametric resonance [5, 14].

Many applications including those mentioned above deal with a symplectic structure. Recall that a matrix $W \in \mathbb{C}^{N\times N}$ is called J-symplectic if $W^*JW = J$, where $J \in \mathbb{C}^{N\times N}$ is an invertible skew-Hermitian matrix, i.e., $J^* = -J$. A standard choice for J is

$$J = \begin{bmatrix} 0_{N/2} & I_{N/2} \\ -I_{N/2} & 0_{N/2} \end{bmatrix}, \tag{1.2}$$

where N is even and $0_{N/2}$ and $I_{N/2}$ are respectively the zero and identity matrices of order $N/2$. When J is given by (1.2), the J-symplecticity of W yields the representation $W^{-1} = -JW^*J$. Moreover, if W is partitioned as

$$W = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}, \quad W_{ij} \in \mathbb{C}^{N/2\times N/2},$$

then

$$W^{-1} = \begin{bmatrix} W_{22}^* & -W_{12}^* \\ -W_{21}^* & W_{11}^* \end{bmatrix}, \tag{1.3}$$

and this formula can be used for an inexpensive computation of W^{-1}; see Algorithm 1.

*Received October 30, 2013. Accepted December 10, 2014. Published online on January, 23, 2015. Recommended by H. Sadok.

†University of Bergen, Department of Mathematics, Postbox 7800, 5020 Bergen, Norway (alexander.malyshev@math.uib.no).

‡Université de Brest, CNRS-UMR 6205, Laboratoire de Mathématiques de Bretagne Atlantique. 6, Av. Le Gorgeu. 29238 Brest Cedex 3., France (miloud.sadkane@univ-brest.fr).

§Laboratoire de Mathématiques Pures et Appliqués, Université du Littoral-Côte d'Opale. C.U. de la Mi-Voix, 50 rue F. Buisson, B.P. 699, 62228 Calais Cedex, France (Ahmed.Salam@lmpa.univ-littoral.fr).

The symplectic structure of W implies that $N_\infty = N_0 = N - N_\infty - N_1$, and the eigenvalues of W_∞ and W_0 in (1.1) are symmetric with respect to the unit circle. When all the blocks X_∞, X_1, and X_0 are nonempty, they satisfy the following J-orthogonalization properties:

$$X_\infty^* J X_\infty = 0, \quad X_0^* J X_0 = 0, \quad X_\infty^* J X_1 = 0, \quad X_0^* J X_1 = 0. \tag{1.4}$$

Furthermore, the matrices $X_0^* J X_\infty$ and $X_1^* J X_1$ are nonsingular, and the matrix

$$P_1 = X_1 (X_1^* J X_1)^{-1} X_1^* J \tag{1.5}$$

is the projector onto $\mathcal{X}_1$ along $\mathcal{X}_0 + \mathcal{X}_\infty$. More details are found in [14].

A quadratically convergent algorithm was proposed in [2] for computing the invariant subspaces $\mathcal{X}_\infty$, $\mathcal{X}_1$, and $\mathcal{X}_0$ of a symplectic matrix W. However, this algorithm necessitates the solution of a $3N \times 3N$ linear system at each iteration. In the present paper we propose a cheaper algorithm based on a variant of a subspace iteration suitable for symplectic matrices. It is widely known that (orthogonal) subspace iteration is reliable [9, 10, 11] even though its convergence rate is only linear. In our context, a classical variant of subspace iteration is formally implemented as follows: starting from $Q_0 = I$, where I denotes the identity matrix, compute iteratively Q_k by means of the QR factorizations $WQ_{k-1} = Q_k R_k$. If $Q_k = \begin{bmatrix} Q_k^{(\infty)} & Q_k^{(1)} & Q_k^{(0)} \end{bmatrix}$ is partitioned such that $Q_k^{(\infty)} \in \mathbb{C}^{N \times N_\infty}$, $Q_k^{(1)} \in \mathbb{C}^{N \times N_1}$, and $Q_k^{(0)} \in \mathbb{C}^{N \times N_0}$, then the convergence properties of the subspace iteration method (see, e.g., [10, Chapter 5]) guarantee that for large k, the subspaces $\text{range}(Q_k^{(\infty)})$, $\text{range}(Q_k^{(1)})$, and $\text{range}(Q_k^{(0)})$ approximate $\mathcal{X}_\infty$, $\mathcal{X}_1$, and $\mathcal{X}_0$, respectively. However, the detection of convergence can be nontrivial.

We propose another variant of subspace iteration:

ALGORITHM 1.

1. Set $Q_{0,1} = Q_{0,2} = I$.
2. For $k = 1, 2, \dots,$ $\begin{bmatrix} WQ_{1,k-1} \\ W^{-1} Q_{2,k-1} \end{bmatrix} = \begin{bmatrix} Q_{1,k} \\ Q_{2,k} \end{bmatrix} R_k.$

Here the matrices $Q_{1,k}, Q_{2,k} \in \mathbb{C}^{N \times N}$ satisfy the identity $Q_{1,k}^* Q_{1,k} + Q_{2,k}^* Q_{2,k} = I$, and the matrix $R_k \in \mathbb{C}^{N \times N}$ is upper triangular. Step 2 uses formula (1.3) for a symplectic matrix W. This variant has been used in the context of spectral dichotomy for a general matrix [1, 3], where the identity matrix was used in step 2 instead of W^{-1}. In such a case, Algorithm 1 is twice as slow; see the numerical experiments in Section 4.

We will show that for large k, the subspaces $\text{range}(Q_{1,k})$ and $\text{range}(Q_{2,k})$ approximate $\mathcal{X}_\infty + \mathcal{X}_1$ and $\mathcal{X}_0 + \mathcal{X}_1$, respectively, and that the intersection $\text{range}(Q_{1,k}) \cap \text{range}(Q_{2,k})$ approximates $\mathcal{X}_1$. An orthogonalization process is then applied to extract $\mathcal{X}_\infty$ and $\mathcal{X}_0$ from $\mathcal{X}_\infty + \mathcal{X}_1$ and $\mathcal{X}_0 + \mathcal{X}_1$.

The convergence behavior of Algorithm 1 is studied in Section 2. Roughly speaking, convergence is fast if the spectra of the blocks W_∞ and W_0 are well separated from the unit circle, W_1 is diagonalizable, and the condition number of $\begin{bmatrix} X_\infty & X_1 & X_0 \end{bmatrix}$ is not large.

The following notation and assumptions are used throughout the paper. The matrix W_1 is assumed to be diagonalizable and hence diagonal due to (1.1). The identity and null matrices of order p are denoted by I_p and 0_p or just I and 0 when the order is clear from the context. The 2-norm and Frobenius norm of a matrix A are denoted by $\|A\|_2$ and $\|A\|_F$. The transpose conjugate of A is denoted by A^*. A calligraphic letter $\mathcal{A}$ denotes the subspace $\text{range}(A)$ corresponding to a matrix A. The singular values of a matrix A are labeled in decreasing order, i.e., $\sigma_{\max}(A) = \sigma_1(A) \geq \sigma_2(A) \geq \ldots \geq \sigma_{\min}(A)$. If A is nonsingular, its condition number $\sigma_{\max}(A)/\sigma_{\min}(A)$ with respect to the 2-norm is denoted by $\text{cond}_2(A)$.

ETNA
Kent State University
http://etna.math.kent.edu

2. Convergence theory for Algorithm 1. The columns of the matrix $\begin{bmatrix} Q_{1,k} \\ Q_{2,k} \end{bmatrix}$ form an orthonormal basis of the linear space spanned by the columns of the matrix $\begin{bmatrix} W^k \\ W^{-k} \end{bmatrix}$. This fact follows from the following proposition.

PROPOSITION 2.1. *For the matrices in Algorithm 1, we have*

$$\begin{bmatrix} W^k \\ W^{-k} \end{bmatrix} = \begin{bmatrix} Q_{1,k} \\ Q_{2,k} \end{bmatrix} \bar{R}_k,$$

where $\bar{R}_k = R_k R_{k-1} \dots R_1$ *is upper triangular, and*

$$Q_{1,k}^* Q_{1,k} + Q_{2,k}^* Q_{2,k} = I.$$

Proof. Apply induction as follows:

$$\begin{bmatrix} W^k \\ W^{-k} \end{bmatrix} = \begin{bmatrix} W W^{k-1} \\ W^{-1} W^{-k+1} \end{bmatrix} = \begin{bmatrix} W Q_{1,k-1} \\ W^{-1} Q_{2,k-1} \end{bmatrix} \bar{R}_{k-1} = \begin{bmatrix} Q_{1,k} \\ Q_{2,k} \end{bmatrix} R_k \bar{R}_{k-1}. \qquad \square$$

Let us choose the block diagonal factorization

$$\begin{aligned} W &= X \begin{bmatrix} W_\infty & & \\ & W_1 & \\ & & W_0 \end{bmatrix} X^{-1} \\ &= \begin{bmatrix} X_\infty & X_1 & X_0 \end{bmatrix} \begin{bmatrix} W_\infty & & \\ & W_1 & \\ & & W_0 \end{bmatrix} \begin{bmatrix} X_\infty & X_1 & X_0 \end{bmatrix}^{-1} \end{aligned} \tag{2.1}$$

so that the columns of X_1 are eigenvectors of unit length corresponding to the eigenvalues of W on the unit circle and the orthonormality conditions

$$X_\infty^* X_\infty = I, \quad X_0^* X_0 = I \tag{2.2}$$

hold. Then the norm of X satisfies the bound $\|X\|_2 \le \sqrt{2 + N_1}$.

By Proposition 2.1, the linear space spanned by the columns of

$$\begin{bmatrix} W^k \\ W^{-k} \end{bmatrix} = \begin{bmatrix} X \begin{bmatrix} I & & \\ & W_1^k & \\ & & W_0^{2k} \end{bmatrix} \\ X \begin{bmatrix} W_\infty^{-2k} & & \\ & W_1^{-k} & \\ & & I \end{bmatrix} \end{bmatrix} \begin{bmatrix} W_\infty^k & & \\ & I & \\ & & W_0^{-k} \end{bmatrix} X^{-1} \tag{2.3}$$

coincides with the linear space spanned by the columns of the matrix $A_k + E_k$, where

$$A_k = \begin{bmatrix} X \begin{bmatrix} I & & \\ & W_1^k & \\ & & 0 \end{bmatrix} \\ X \begin{bmatrix} 0 & & \\ & W_1^{-k} & \\ & & I \end{bmatrix} \end{bmatrix}, \quad E_k = \begin{bmatrix} X \begin{bmatrix} 0 & & \\ & 0 & \\ & & W_0^{2k} \end{bmatrix} \\ X \begin{bmatrix} W_\infty^{-2k} & & \\ & 0 & \\ & & 0 \end{bmatrix} \end{bmatrix}. \tag{2.4}$$

ETNA
Kent State University
http://etna.math.kent.edu

Convergence of Algorithm 1 is based on the following result from the perturbation theory for the QR factorization.

THEOREM 2.2 (Sun [13]). *Consider QR factorizations* $A = \underset{\sim}{Q}\underset{\sim}{R}$ *and* $A + E = \widetilde{Q}\widetilde{R}$ *with the upper triangular factors having positive diagonals for matrices* A *and* $A + E$ *of full column rank. If* $\|A^\dagger\|_2\|E\|_2 < 1$*, where* $A^\dagger$ *is the Moore-Penrose pseudoinverse of* A*, then*

$$\|\widetilde{Q} - \underset{\sim}{Q}\|_F \leq \frac{\sqrt{2}\|A^\dagger\|_2\|E\|_F}{\sqrt{1 - \|A^\dagger\|_2\|E\|_2}}.$$

To apply this theorem, we need the following lemmas.

LEMMA 2.3. *For all* $k \geq 0$*, we have*

$$\|E_k\|_2 \leq \sqrt{\omega}\left(1 - \omega^{-1}\right)^k,$$

where $\omega = \max\left\{\|\sum_{k=0}^{\infty} W_0^k (W_0^k)^*\|_2, \|\sum_{k=0}^{\infty} W_\infty^{-k}(W_\infty^{-k})^*\|_2\right\} > 1.$

Proof. The norms of the matrix powers W_0^{2k} and W_∞^{-2k}, $k \geq 0$, decay as follows (see, e.g., [4]):

$$\|W_0^{2k}\|_2 \leq \sqrt{\|H_0\|_2}\left(1 - \frac{1}{\|H_0\|_2}\right)^k, \quad \|W_\infty^{-2k}\|_2 \leq \sqrt{\|H_\infty\|_2}\left(1 - \frac{1}{\|H_\infty\|_2}\right)^k,$$

where $H_0 = \sum_{k=0}^{\infty} W_0^k (W_0^k)^*$ and $H_\infty = \sum_{k=0}^{\infty} W_\infty^{-k}(W_\infty^{-k})^*$. Furthermore, from (2.2) and (2.4) we obtain $\|E_k\|_2^2 = \max\left\{\|W_0^{2k}\|_2^2, \|W_\infty^{-2k}\|_2^2\right\}$. □

LEMMA 2.4. *The Moore-Penrose pseudoinverse of* A_k *satisfies the bound*

$$\|A_k^\dagger\|_2 \leq \|X^{-1}\|_2.$$

Proof. From the properties of the singular values [12] and the matrix W_1, we have

$$\sigma_{\min}(A_k) \geq \sigma_{\min}\left(\begin{bmatrix} I & & \\ & W_1^k & \\ & & 0 \\ 0 & & \\ & W_1^{-k} & \\ & & I \end{bmatrix}\right)\sigma_{\min}(X) = \sigma_{\min}(X),$$

which yields the desired bound. □

Lemmas 2.3 and 2.4 show that A_k has full rank and $A_k + E_k$ has full rank for large k. To apply Theorem 2.2 for sufficiently large k, consider the QR decompositions

$$A_k = \underset{\sim}{Q}_k\underset{\sim}{R}_k, \quad A_k + E_k = \widetilde{Q}_k\widetilde{R}_k. \tag{2.5}$$

From (2.3) and (2.4) we know that $\widetilde{Q}_k = \begin{bmatrix} Q_{1,k} \\ Q_{2,k}\end{bmatrix}$. Let $\underset{\sim}{Q}_k = \begin{bmatrix} \underset{\sim}{Q}_{1,k} \\ \underset{\sim}{Q}_{2,k}\end{bmatrix}$. Then with the help of Lemmas 2.3, 2.4, and Theorem 2.2, we arrive at the following convergence estimate

$$\max\left\{\|Q_{1,k} - \underset{\sim}{Q}_{1,k}\|_2, \|Q_{2,k} - \underset{\sim}{Q}_{2,k}\|_2\right\} \leq \frac{\sqrt{2}\sqrt{N}\|X^{-1}\|_2\sqrt{\omega}(1 - \omega^{-1})^k}{\sqrt{1 - \|X^{-1}\|_2\sqrt{\omega}(1 - \omega^{-1})^k}}. \tag{2.6}$$

From this estimate and the fact that $\underset{\sim}{\mathcal{Q}}_{1,k} = \mathcal{X}_\infty + \mathcal{X}_1$ and $\underset{\sim}{\mathcal{Q}}_{2,k} = \mathcal{X}_1 + \mathcal{X}_0$, we obtain the following result.

PROPOSITION 2.5. *Algorithm 1 is linearly convergent, and*

$$\lim_{k\to\infty} \mathcal{Q}_{1,k} = \mathcal{X}_\infty + \mathcal{X}_1, \quad \lim_{k\to\infty} \mathcal{Q}_{2,k} = \mathcal{X}_1 + \mathcal{X}_0.$$

The above convergence analysis shows that convergence mainly depends on the condition number of the matrix X from the block-diagonalization (2.1) and on the decay of the powers of W_0 and W_∞^{-1}.

The next theorem characterizes the singular values of $\underset{\sim}{Q}_{1,k}$ and $\underset{\sim}{Q}_{2,k}$. For large k, this theorem and the estimate (2.6) characterize the singular values of $Q_{1,k}$ and $Q_{2,k}$. Such a characterization will be useful mainly for numerical purposes; see Section 4.

THEOREM 2.6. *The matrix $\underset{\sim}{Q}_{1,k}$ has N_∞ singular values equal to 1, N_0 singular values equal to 0, and N_1 singular values in the interval $[s, \sqrt{1-s^2}]$, where $s = \frac{1}{\sqrt{2}\,\mathrm{cond}_2(X)}$. The singular values $\sigma_i(\underset{\sim}{Q}_{2,k})$ equal $\sqrt{1-\sigma^2_{N-i+1}(\underset{\sim}{Q}_{1,k})}$.*

Proof. Since $\underset{\sim}{Q}^*_{1,k}\underset{\sim}{Q}_{1,k} + \underset{\sim}{Q}^*_{2,k}\underset{\sim}{Q}_{2,k} = I$, it is clear that $\sigma_i^2(\underset{\sim}{Q}_{1,k}) + \sigma^2_{N-i+1}(\underset{\sim}{Q}_{2,k}) = 1$. Also, since from (2.4) and (2.5) if follows that

$$\underset{\sim}{Q}_{1,k} = X \begin{bmatrix} I & & \\ & W_1^k & \\ & & 0 \end{bmatrix} \underset{\sim}{R}_k^{-1}, \quad \underset{\sim}{Q}_{2,k} = X \begin{bmatrix} 0 & & \\ & W_1^{-k} & \\ & & I \end{bmatrix} \underset{\sim}{R}_k^{-1},$$

it is clear that $\sigma_i(\underset{\sim}{Q}_{1,k}) = 0$ for $i > N_\infty + N_1$ and $\sigma_i(\underset{\sim}{Q}_{2,k}) = 0$ for $i > N_0 + N_1$, and therefore, $\sigma_i(\underset{\sim}{Q}_{1,k}) = 1$ for $1 \le i \le N_\infty$ and $\sigma_i(\underset{\sim}{Q}_{2,k}) = 1$ for $1 \le i \le N_0$.

From (2.4) and (2.5),

$$\underset{\sim}{R}_k^*\underset{\sim}{R}_k = \begin{bmatrix} I & & \\ & W_1^k & \\ & & 0 \end{bmatrix}^* X^*X \begin{bmatrix} I & & \\ & W_1^k & \\ & & 0 \end{bmatrix} + \begin{bmatrix} 0 & & \\ & W_1^{-k} & \\ & & I \end{bmatrix}^* X^*X \begin{bmatrix} 0 & & \\ & W_1^{-k} & \\ & & I \end{bmatrix},$$

which yields $\|R_k\|_2 \le \sqrt{2}\|X\|_2$.

Now, for $i \le N_\infty + N_1$,

$$1 = \sigma_i \begin{bmatrix} I & & \\ & W_1^k & \\ & & 0 \end{bmatrix} \le \sigma_i(\underset{\sim}{Q}_{1,k})\|X^{-1}\|_2\|R_k\|_2 \le \sigma_i(\underset{\sim}{Q}_{1,k})\sqrt{2}\,\mathrm{cond}_2(X).$$

Hence, $\sigma_i(\underset{\sim}{Q}_{1,k}) \ge \frac{1}{\sqrt{2}\,\mathrm{cond}_2(X)}$ for $i \le N_\infty + N_1$. Similarly, $\sigma_i(\underset{\sim}{Q}_{2,k}) \ge \frac{1}{\sqrt{2}\,\mathrm{cond}_2(X)}$ for $i \le N_0 + N_1$. $\square$

3. Algorithmic aspects. In this section we discuss some important issues that arise when implementing Algorithm 1, namely the approximation of the invariant subspaces $\mathcal{X}_\infty$, $\mathcal{X}_0$, and $\mathcal{X}_\infty$ from the matrices $Q_{1,k}$ and $Q_{2,k}$ and the stopping criterion.

3.1. Computation of $\mathcal{X}_\infty$, $\mathcal{X}_0$, and $\mathcal{X}_\infty$. The subspace $\mathcal{X}_1$ is the intersection of the subspaces $\mathcal{X}_\infty + \mathcal{X}_1$ and $\mathcal{X}_1 + \mathcal{X}_0$. From Proposition 2.5, the natural way to obtain such an intersection, once the iteration has been stopped, is to compute the intersection of the subspaces $\mathcal{Q}_{1,k}$ and $\mathcal{Q}_{2,k}$. This can be obtained from the SVDs of $Q_{1,k}$ and $Q_{2,k}$ [7, Chapter 12]. Let $Q_{1,k}$ and $Q_{2,k}$ be matrices whose columns form orthonormal bases of $\mathcal{Q}_{1,k}$ and $\mathcal{Q}_{2,k}$. Then the intersection can be obtained as the left or right singular vectors associated with the singular values of $Q^*_{2,k}Q_{1,k}$ that are equal to 1.

ETNA
Kent State University
http://etna.math.kent.edu

Let $\widetilde{X}_1$ be the matrix whose columns are formed by these singular vectors. Then we have $\widetilde{\mathcal{X}}_1 \approx \mathcal{X}_1$, and $\widetilde{P}_1 = \widetilde{X}_1(\widetilde{X}_1^* J \widetilde{X}_1)^{-1}\widetilde{X}_1^H J$ is an approximation of the projector P_1 given in (1.5). The J-orthogonalization properties (1.4) allow us to conclude that

$$\widetilde{\mathcal{X}}_\infty \equiv \text{range}\left((I - \widetilde{P}_1)Q_{1,k}\right) \approx \mathcal{X}_\infty \quad \text{and} \quad \widetilde{\mathcal{X}}_0 \equiv \text{range}\left((I - \widetilde{P}_1)Q_{2,k}\right) \approx \mathcal{X}_0.$$

If needed, the approximate projectors onto $\mathcal{X}_\infty$ and $\mathcal{X}_0$ can be constructed as follows: since by construction, the matrices $(I - \widetilde{P}_1)Q_{1,k}$ and $(I - \widetilde{P}_1)Q_{2,k}$ have only $N_\infty = N_0$ non-negligible singular values, the SVDs of $(I - \widetilde{P}_1)Q_{1,k}$ and $(I - \widetilde{P}_1)Q_{2,k}$ yield matrices $\widetilde{X}_\infty$ and $\widetilde{X}_0$ of $N_\infty = N_0$ orthonormal columns (left singular vectors associated with the larger singular values) which form an approximate basis of $\mathcal{X}_\infty$ and $\mathcal{X}_0$. The desired projectors are then given by $\widetilde{P}_\infty = \widetilde{X}_\infty(\widetilde{X}_0^* J \widetilde{X}_\infty)^{-1}\widetilde{X}_0^H J$ and $\widetilde{P}_0 = \widetilde{X}_0(\widetilde{X}_\infty^H J \widetilde{X}_0)^{-1}\widetilde{X}_\infty^H J$.

3.2. Stopping criterion. An important and difficult task is the choice of an effective criterion to stop the iteration of Algorithm 1. We have seen in Section 2 that the subspaces $\mathcal{Q}_{1,k}$ and $\mathcal{Q}_{2,k}$ converge to $\mathcal{X}_\infty + \mathcal{X}_1$ and $\mathcal{X}_1 + \mathcal{X}_0$, respectively. In particular, for large k, the matrices $Q_{1,k}$ and $Q_{2,k}$ become singular, the $N_0 = N_\infty$ smallest singular values of $Q_{1,k}$ and $Q_{2,k}$ converge to 0, the $N_\infty = N_0$ largest singular values converge to 1, and the remaining N_1 singular values oscillate in the open interval $(0, 1)$. This fact can be used as a stopping criterion.

For example, if the eigenvalues of W are all on the unit circle (i.e., $N_\infty = N_0 = 0$), then $\mathcal{Q}_{1,k}$ and $\mathcal{Q}_{2,k}$ converge to $\mathcal{X}_1$. Actually, in this case, (1.1) reduces to $W = X_1 W_1 X_1^{-1}$. It can easily be shown that then $\mathcal{Q}_{1,k} = \mathcal{Q}_{2,k} = \mathcal{X}_1$ for all $k \geq 1$. In other words, Algorithm 1 converges at the first iteration. In practice, if the singular values of $Q_{1,k}$ or $Q_{2,k}$ oscillate in $(0, 1)$, we may conclude that the eigenvalues of W are all on the unit circle. Then $\widetilde{X}_0 = \widetilde{X}_\infty = 0$, and $\widetilde{X}_1$ can be computed as explained above.

If W has no eigenvalue on the unit circle, i.e., $N_\infty = N_0 = N/2$, then $\mathcal{Q}_{1,k}$ converges to $\mathcal{X}_\infty$ and $\mathcal{Q}_{2,k}$ converges to $\mathcal{X}_0$. For large k, $Q_{1,k}$ and $Q_{2,k}$ have $N/2$ singular values close to 0 and $N/2$ singular values close to 1. A stopping criterion is obtained by monitoring the decrease to 0 of the $(\frac{N}{2} + 1)$-st largest singular value of $Q_{1,k}$. Another stopping criterion can be derived from the property $\lim_{k\to\infty} \sigma_i(Q_{1,k}Q_{2,k}^*) = 0$ for all i, and thus, the stopping criterion consists of checking if the norm $\|Q_{1,k}Q_{2,k}^*\|_2$ is less than some fixed threshold tol.

In the general case when there are eigenvalues inside, on, and outside the unit circle, for large k, $Q_{1,k}$ has N_∞ singular values close to 1, $N_0 = N_\infty$ singular values close to 0, and $N_1 = N - 2N_\infty$ singular values that oscillate in the open interval $(0, 1)$. The numerical experiments show that in general the convergence to 1 and 0 is fast when there is a sufficiently large gap between the eigenvalues on the unit circle and those outside or inside of it. A stopping criteria is obtained by monitoring the larger singular values of $Q_{1,k}$ (which must converge to 1). The iterations are stopped when the number of ones stabilizes to a value, say, $\widetilde{N}_\infty$, and $\sigma_{N-\widetilde{N}_\infty+1}(Q_{1,k}) < tol < \sigma_{N-\widetilde{N}_\infty}(Q_{1,k})$, where tol is some prescribed tolerance. Then we set $N_\infty = \widetilde{N}_\infty$. To keep the cost as small as possible, the computation of the singular values is carried out periodically, for example, every 10th iteration. Once the iteration is stopped, the matrices $\widetilde{X}_\infty$, $\widetilde{X}_1$, and $\widetilde{X}_0$ can be computed as explained in Section 3.1. Note that these matrices must satisfy approximately the properties (1.4) and (1.5), which can be used as an a posteriori verification of the accuracy of the computed matrices.

4. Numerical experiments. In this section we report some numerical experiments illustrating the convergence behavior of Algorithm 1. In all tests, the matrix J has the form (1.2). We also present comparisons with Algorithm 2 which consists of replacing, in step 2 of Algorithm 1, the matrix W^{-1} by the identity matrix.

ETNA
Kent State University
http://etna.math.kent.edu

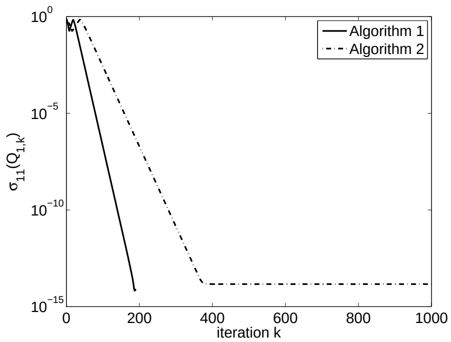

FIG. 4.1. *Convergence behavior of Algorithms 1 and 2 (Example 4.2).*

TABLE 4.1
Singular values of $Q_{1,k}$ (Example 4.2).

	iteration k				
	1	50	100	150	190
$\sigma_1(Q_{1,k})$	$9.983\ 10^{-1}$	1	1	1	1
$\sigma_{10}(Q_{1,k})$	$7.794\ 10^{-1}$	1	1	1	1
$\boldsymbol{\sigma_{11}(Q_{1,k})}$	$\mathbf{7.002\ 10^{-1}}$	$\mathbf{1.356\ 10^{-3}}$	$\mathbf{9.844\ 10^{-8}}$	$\mathbf{7.142\ 10^{-12}}$	$\mathbf{2.400\ 10^{-15}}$
$\sigma_{20}(Q_{1,k})$	$5.805\ 10^{-2}$	$4.770\ 10^{-21}$	$3.741\ 10^{-39}$	$2.814\ 10^{-60}$	$3.659\ 10^{-78}$

EXAMPLE 4.1. This example shows that Algorithm 1 converges at the first iteration when the eigenvalues of W are all on the unit circle. The matrix W is chosen block diagonal as $W = \texttt{blockdiag}\,(Q, Q)$, where Q is a 10×10 orthogonal matrix constructed with the MATLAB function `orthog`.

At the first iteration of Algorithm 1 ($k = 1$), the singular values of the matrix $Q_{1,k}$ are all equal to $7.0711\ 10^{-1}$ and remain the same during all iterations. From the discussion in Section 3.2, we conclude that the eigenvalues of W are all on the unit circle. The algorithm computes $\widetilde{X}_1$ being of order 20 satisfying $\|W\widetilde{X}_1 - \widetilde{X}_1\left(\widetilde{X}_1^* W \widetilde{X}_1\right)\| = 1.5504\ 10^{-15}$ and $\widetilde{P}_1$ satisfying $\|\widetilde{P}_1 - I_{20}\| = 1.3545\ 10^{-15}$.

EXAMPLE 4.2. This example shows that when W has no eigenvalues on the unit circle ($N = N_0 = N_\infty$), then the N_0 largest (smallest) singular values of $Q_{1,k}$ and $Q_{2,k}$ converge to 1 (0). The matrix W is given by $W = \texttt{blockdiag}\left(A, (A^{-1})^*\right)$, where A is a 10×10 upper triangular matrix whose strictly upper triangular part is chosen randomly in $(0, 1)$ and the diagonal elements are such that $A(k, k) = 1 + k/10$, $k = 1, \ldots, 10$. Therefore W has $N_0 = 10$ eigenvalues inside the unit circle and $N_\infty = 10$ eigenvalues outside of it. Following the discussion in Section 3.2, we show in Figure 4.1 and Table 4.1 (and especially in line 4 of this table) the convergence to 0 of the 11-th largest singular value of $Q_{1,k}$. With the stopping criterion discussed in Section 3.2 and $tol = 10^{-14}$, Algorithm 1 necessitates 190 iterations. The figure also exhibits the results of Algorithm 2. At iteration 431, $\sigma_{11}(Q_{1,k})$ computed by Algorithm 2 stagnates at $1.4681\ 10^{-14}$.

At iteration $k = 190$, Algorithm 1 computes matrices $\widetilde{X}_0$ and $\widetilde{X}_\infty$ each of size 20×10 whose columns are orthonormal and satisfy $\|W\widetilde{X}_0 - \widetilde{X}_0\left(\widetilde{X}_0^* W \widetilde{X}_0\right)\| = 8.3316\ 10^{-16}$ and $\|W\widetilde{X}_\infty - \widetilde{X}_\infty\left(\widetilde{X}_\infty^* W \widetilde{X}_\infty\right)\| = 1.7496\ 10^{-15}$. The computed projectors $\widetilde{P}_0$, $\widetilde{P}_\infty$ sat-

ETNA
Kent State University
http://etna.math.kent.edu

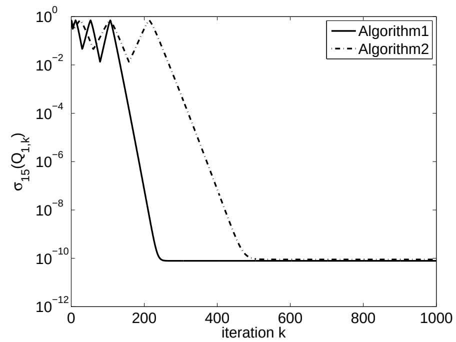

FIG. 4.2. *Convergence behavior of Algorithms 1 and 2 (Example 4.3).*

TABLE 4.2
Singular values of $Q_{1,k}$ (Example 4.3).

	iteration k		
	1	50	100
$\sigma_1(Q_{1,k})$	$9.991\ 10^{-1}$	1	1
$\sigma_6(Q_{1,k})$	$7.268\ 10^{-1}$	$8.410\ 10^{-1}$	$9.496\ 10^{-1}$
$\sigma_7(Q_{1,k})$	$7.071\ 10^{-1}$	$7.071\ 10^{-1}$	$7.071\ 10^{-1}$
$\sigma_{14}(Q_{1,k})$	$7.071\ 10^{-1}$	$7.071\ 10^{-1}$	$7.071\ 10^{-1}$
$\boldsymbol{\sigma_{15}(Q_{1,k})}$	$\mathbf{6.867\ 10^{-1}}$	$\mathbf{5.410\ 10^{-1}}$	$\mathbf{3.133\ 10^{-1}}$
$\sigma_{20}(Q_{1,k})$	$4.101\ 10^{-2}$	$1.702\ 10^{-13}$	$1.485\ 10^{-19}$

	iteration k		
	150	200	250
$\sigma_1(Q_{1,k})$	1	1	1
$\sigma_6(Q_{1,k})$	1	1	1
$\sigma_7(Q_{1,k})$	$7.071\ 10^{-1}$	$7.071\ 10^{-1}$	$7.071\ 10^{-1}$
$\sigma_{14}(Q_{1,k})$	$7.071\ 10^{-1}$	$7.071\ 10^{-1}$	$7.071\ 10^{-1}$
$\boldsymbol{\sigma_{15}(Q_{1,k})}$	$\mathbf{6.216\ 10^{-4}}$	$\mathbf{7.020\ 10^{-8}}$	$\mathbf{8.380\ 10^{-11}}$
$\sigma_{20}(Q_{1,k})$	$5.518\ 10^{-25}$	$1.409\ 10^{-23}$	$1.638\ 10^{-19}$

isfy $\mathrm{trace}(\widetilde{P}_0) = \mathrm{trace}(\widetilde{P}_\infty) = 10$, $\|\widetilde{P}_0^2 - \widetilde{P}_0\| = 9.5835\ 10^{-16}$, $\|\widetilde{P}_\infty^2 - \widetilde{P}_\infty\| = 4.1679\ 10^{-16}$, and $\|\widetilde{P}_0 + \widetilde{P}_\infty - I_{20}\| = 9.5835\ 10^{-16}$.

EXAMPLE 4.3. This example reveals the convergence behavior when W has eigenvalues inside, on, and outside the unit circle. The matrix W is given by $W = \mathtt{blockdiag}\left(A, (A^{-1})^*\right)$ with $A = \mathtt{blockdiag}\left(A_0, A_1\right)$ where A_0 is a 6×6-Jordan block corresponding to the eigenvalue 0.9 and A_1 is a 4×4 orthogonal matrix constructed with the MATLAB function `orthog`. Therefore, W has $N_0 = 6$ eigenvalues inside the unit circle, $N_\infty = 6$ eigenvalues outside the unit circle, and $N_1 = 8$ eigenvalues on the unit circle. The singular values $\sigma_{15}, \sigma_{16}, \ldots, \sigma_{20}$ should converge to 0. The stopping criterion discussed in Section 3.2 is used with $tol = 10^{-10}$. Algorithm 1 required 250 iterations while Algorithm 2 required 500 iterations. The convergence behavior of both algorithms is shown in Figure 4.2. Table 4.2 illustrates the convergence of the singular values of $Q_{1,k}$ computed by Algorithm 1.

ETNA
Kent State University
http://etna.math.kent.edu

TABLE 4.3
Singular values of $Q_{1,k}$ (Example 4.4).

	iteration k				
	1	100	200	400	600
$\sigma_1(Q_{1,k})$	$9.9999\ 10^{-1}$	1	1	1	1
$\sigma_6(Q_{1,k})$	$9.1297\ 10^{-1}$	1	1	1	1
$\sigma_7(Q_{1,k})$	$8.9088\ 10^{-1}$	1	1	1	1
$\sigma_8(Q_{1,k})$	$8.1080\ 10^{-1}$	1	1	1	1
$\sigma_9(Q_{1,k})$	$8.0423\ 10^{-1}$	$9.9995\ 10^{-1}$	$9.9999\ 10^{-1}$	1	1
$\sigma_{10}(Q_{1,k})$	$7.4643\ 10^{-1}$	$9.9995\ 10^{-1}$	$9.9999\ 10^{-1}$	1	1
$\sigma_{11}(Q_{1,k})$	$6.6546\ 10^{-1}$	$1.0148\ 10^{-2}$	$5.0373\ 10^{-3}$	$2.5094\ 10^{-3}$	$1.6708\ 10^{-3}$
$\sigma_{12}(Q_{1,k})$	$5.9432\ 10^{-1}$	$9.8486\ 10^{-3}$	$4.9623\ 10^{-3}$	$2.4906\ 10^{-3}$	$1.6625\ 10^{-3}$
$\sigma_{13}(Q_{1,k})$	$5.8532\ 10^{-1}$	$7.6121\ 10^{-7}$	$9.4452\ 10^{-8}$	$1.1763\ 10^{-8}$	$3.4809\ 10^{-9}$
$\sigma_{14}(Q_{1,k})$	$4.5423\ 10^{-1}$	$7.3872\ 10^{-7}$	$9.3046\ 10^{-8}$	$1.1675\ 10^{-8}$	$3.4635\ 10^{-9}$
$\sigma_{15}(Q_{1,k})$	$4.0802\ 10^{-1}$	$4.5589\ 10^{-14}$	$4.5589\ 10^{-14}$	$4.5589\ 10^{-14}$	$4.5589\ 10^{-14}$
$\sigma_{20}(Q_{1,k})$	$3.5487\ 10^{-3}$	$6.7017\ 10^{-18}$	$6.7017\ 10^{-18}$	$6.7017\ 10^{-18}$	$6.7017\ 10^{-18}$

At iteration $k = 250$, Algorithm 1 computes matrices $\widetilde{X}_0$, $\widetilde{X}_1$, and $\widetilde{X}_\infty$ of sizes 20×6, 20×8, and 20×6, respectively, whose columns are orthonormal and which satisfy $\|W\widetilde{X}_0 - \widetilde{X}_0\left(\widetilde{X}_0^* W \widetilde{X}_0\right)\| = 8.9496\ 10^{-16}$, $\|W\widetilde{X}_1 - \widetilde{X}_1\left(\widetilde{X}_1^* W \widetilde{X}_1\right)\| = 1.4938\ 10^{-15}$, and $\|W\widetilde{X}_\infty - \widetilde{X}_\infty\left(\widetilde{X}_\infty^* W \widetilde{X}_\infty\right)\| = 1.4164\ 10^{-15}$.

The computed projectors $\widetilde{P}_0$, $\widetilde{P}_1$, $\widetilde{P}_\infty$ satisfy $\text{trace}(\widetilde{P}_0) = \text{trace}(\widetilde{P}_\infty) = 6$, $\text{trace}(\widetilde{P}_1) = 8$, $\|\widetilde{P}_0^2 - \widetilde{P}_0\| = 3.2880\ 10^{-16}$, $\|\widetilde{P}_1^2 - \widetilde{P}_1\| = 3.5060\ 10^{-16}$, $\|\widetilde{P}_\infty^2 - \widetilde{P}_\infty\| = 4.6857\ 10^{-16}$, and $\|\widetilde{P}_0 + \widetilde{P}_1 + \widetilde{P}_\infty - I_{20}\| = 5.1514\ 10^{-16}$.

EXAMPLE 4.4. This example shows that the diagonalizability condition imposed on W_1 is necessary for the extraction of the desired invariant subspaces. Without this condition, the behavior of Algorithm 1 is unpredictable. We consider the preceding example where A_0 is a 4×4-Jordan block for the eigenvalue 1. Table 4.3 displays the singular values of $Q_{1,k}$.

It seems that, during the iterations, the 8 eigenvalues on the unit circle moved off the unit circle preserving the symmetry of the spectrum. See the singular values $\sigma_i(Q_{1,k})$, $i = 7, 8, \ldots, 14$, in Table 4.3, where we observe that 4 eigenvalues moved out of the unit circle and 4 moved inside of it.

5. Conclusions.

We have studied the behavior of a variant of a subspace iteration suited for symplectic matrices. The study reveals the parameters responsible for the convergence rate. The algorithm is attractive due to its simplicity and the low cost of a single iteration.

REFERENCES

[1] A. YA. BULGAKOV AND S. K. GODUNOV, *Circular dichotomy of a matrix spectrum*, Siberian Math. J., 29 (1988), pp. 59–70.

[2] M. DOSSO AND M. SADKANE, *A spectral trichotomy method for symplectic matrices*, Numer. Algorithms, 52 (2009), pp. 187–212.

[3] S. K. GODUNOV, *Problem of the dichotomy of the spectrum of a matrix*, Siberian Math. J., 27 (1986), pp. 649–660.

[4] ———, *Modern Aspects of Linear Algebra*, AMS, Providence, 1998.

[5] S. K. GODUNOV AND M. SADKANE, *Spectral analysis of symplectic matrices with application to the theory of parametric resonance*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 1083–1096.

[6] I. GOHBERG, P. LANCASTER, AND L. RODMAN, *Indefinite Linear Algebra and Applications*, Birkhäuser, Basel, 2005

[7] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, 1996.
[8] P. LANCASTER AND L. RODMAN, *Algebraic Riccati Equations*, Clarendon Press, Oxford, 1995.
[9] B. N. PARLETT, *The Symmetric Eigenvalue Problem*, SIAM, Philadelphia, 1998.
[10] Y. SAAD, *Numerical Methods for Large Eigenvalue Problems*, Manchester University Press, Manchester, 1992.
[11] G. W. STEWART, *Simultaneous iteration for computing invariant subspaces of non-Hermitian matrices*, Numer. Math., 25 (1976), pp. 123–136.
[12] G. W. STEWART AND J.-G. SUN, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.
[13] J.-G. SUN, *On perturbation bounds for the QR factorization*, Linear Algebra Appl., 215 (1995), pp. 95–111.
[14] V. A. YAKUBOVICH AND V. M. STARZHINSKII, *Linear Differential Equations with Periodic Coefficients. Vol. 1 & 2*, Wiley, New York, 1975.