The Symplectic Eigenvalue Problem, the Butterfly Form, the SR Algorithm, and the Lanczos Method

Peter Benner * Heike Faßbender †

August 1, 1997

Abstract

We discuss some aspects of the recently proposed symplectic butterfly form which is a condensed form for symplectic matrices. Any $2n \times 2n$ symplectic matrix can be reduced to this condensed form which contains $8n-4$ nonzero entries and is determined by $4n-1$ parameters. The symplectic eigenvalue problem can be solved using the SR algorithm based on this condensed form. The SR algorithm preserves this form and can be modified to work only with the $4n-1$ parameters instead of the $4n^2$ matrix elements. The reduction of symplectic matrices to symplectic butterfly form has a close analogy to the reduction of arbitrary matrices to Hessenberg form. A Lanczos-like algorithm for reducing a symplectic matrix to butterfly form is also presented.

Key words : butterfly form, symplectic Lanczos method, symplectic matrix, eigenvalues.
AMS(MOS) subject classifications : 65F15, 65F50

1 Introduction

The computation of eigenvalues and eigenvectors or deflating subspaces of symplectic pencils/matrices is an important task in applications like discrete linear-quadratic regulator problems, discrete Kalman filtering, computation of discrete stability radii, and the problem of solving discrete-time algebraic Riccati equations. See, e.g., [22, 24, 25, 29] for applications and further references. A matrix $M \in \mathbb{R}^{2n\times 2n}$ is called *symplectic* (or *J-orthogonal*) if

$$MJM^T = J \tag{1}$$

(or equivalently, $M^TJM = J$) and a *symplectic matrix pencil* $L - \lambda N, L, N \in \mathbb{R}^{2n\times 2n}$ is defined by the property

$$LJL^T = NJN^T, \tag{2}$$

*Universität Bremen, Fachbereich 3 – Mathematik und Informatik, 28357 Bremen, FRG. E-mail: peter@mathematik.uni-bremen.de This work was completed while this author was with the *Technische Universität Chemnitz–Zwickau* and was supported by *Deutsche Forschungsgemeinschaft*, research grant *Me 790/7-1 Singuläre Steuerungsprobleme*.

†Universität Bremen, Fachbereich 3 – Mathematik und Informatik, 28357 Bremen, FRG. E-mail: heike@mathematik.uni-bremen.de

where

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \tag{3}$$

and I_n is the $n \times n$ identity matrix. (Note that (2) is in general not equivalent to $L^T JL = N^T JN$.)

In most applications system-theoretic conditions are satisfied, which guarantee the existence of an n-dimensional invariant subspace (resp. deflating subspace) corresponding to the eigenvalues of the symplectic matrix M (resp. the symplectic pencil $L - \lambda N$) inside the open unit disk. This is the subspace one wishes to compute. The solution of the (generalized) symplectic eigenvalue problem with small and dense coefficient matrices has been the topic of numerous publications during the last 30 years. Even for these problems a numerically sound method, i.e., a strongly backward stable method in the sense of [7], is yet not known. The numerical computation of a deflating subspace is usually carried out by an iterative procedure like the QZ algorithm which transforms $L - \lambda N$ into a generalized Schur form, from which the deflating subspace can be read off. See, e.g., [29, 31]. The QZ algorithm is numerically backward stable but it ignores the symplectic structure. Thus the computed eigenvalues will in general not come in reciprocal pairs, although the exact eigenvalues have this property. Even worse, small perturbations may cause eigenvalues close to the unit disk to cross the unit circle such that the number of true and computed eigenvalues inside the open unit disk may differ. Hence it is crucial to make use of the symplectic structure.

Different structure-preserving methods which avoid the above mentioned problems have been proposed. Mehrmann [28] describes a symplectic QZ algorithm. This algorithm has all desirable properties, but its applicability is limited to the single input/output case due to the lacking reduction to symplectic J–Hessenberg form in the general case [1]. In [26], Lin uses the $S + S^{-1}$-transformation in order to solve the symplectic eigenvalue problem. But the method cannot be used to compute eigenvectors and/or invariant subspaces. Patel [34] shows that these ideas can also be used to derive a structure-preserving method for the generalized symplectic eigenvalue problem similar to Van Loan's square-reduced method for the Hamiltonian eigenvalue problem [37]. Based on the multishift idea presented in [1], he also describes a method working on a condensed symplectic pencil using implicit QZ steps to compute the stable deflating subspace of a symplectic pencil [33].

Using the analogy to the continuous-time case, i.e., Hamiltonian eigenvalue problems, Flaschka, Mehrmann, and Zywietz show in [16] how to construct structure-preserving methods for the symplectic eigenproblem based on the SR method [12, 27]. This method is a QR-like method based on the SR decomposition. In an initial step, the symplectic matrix is reduced to a more condensed form, the symplectic J-Hessenberg form. As in the general framework of GR algorithms [40], the SR iteration preserves the symplectic J-Hessenberg form at each step and is supposed to converge to a form from which eigenvalues and deflating subspaces can be read off. The authors note that "*...the resulting methods have significantly worse numerical properties than their corresponding analogues in the Hamiltonian case*" [16, abstract].

Recently, Banse and Bunse-Gerstner [4, 2, 3] presented a new condensed form for symplectic matrices which can be computed by an elimination process using elementary unitary and symplectic similarity transformations. The $2n \times 2n$ condensed matrix is symplectic, contains $8n - 4$ nonzero entries, and is determined by $4n - 1$ parameters. This condensed form, called *symplectic butterfly form*, can be depicted as a symplectic matrix of the following form:

$$\begin{bmatrix} \diagdown & \diagdown\!\!\!\diagdown\!\!\!\diagdown \\ \diagdown & \diagdown\!\!\!\diagdown\!\!\!\diagdown \end{bmatrix}.$$

The reduction of a symplectic matrix to butterfly form and also the existence of a numerically stable method to compute this reduction is strongly dependent on the first column of the transformation matrix that carries out the transformation. Once the reduction to butterfly form is achieved, the SR algorithm [12, 27] is a suitable tool for computing the eigenvalues/eigenvectors of a symplectic matrix. It preserves the butterfly form in its iterations and can be rewritten in a parameterized form that works with $4n-1$ parameters instead of the $(2n)^2$ matrix elements in each iteration. Hence, the symplectic structure, which will be destroyed in the numerical process due to roundoff errors, can easily be restored in each iteration for this condensed form.

In [2], a strict butterfly matrix is introduced in which the upper left diagonal matrix of the butterfly form is nonsingular. A strict butterfly matrix can be factored as

$$\begin{bmatrix} \diagdown & 0 \\ \diagdown & \diagdown \end{bmatrix} \begin{bmatrix} I & \diagdown\!\!\!\diagdown\!\!\!\diagdown \\ 0 & I \end{bmatrix}.$$

We will introduce an unreduced butterfly form in which the lower right tridiagonal matrix is unreduced. An unreduced butterfly matrix can be factored as

$$\begin{bmatrix} \diagdown & \diagdown \\ 0 & \diagdown \end{bmatrix} \begin{bmatrix} 0 & -I \\ I & \diagdown\!\!\!\diagdown\!\!\!\diagdown \end{bmatrix}.$$

Any unreduced butterfly matrix is similar to a strict butterfly matrix, but not vice versa. We will show that unreduced butterfly matrices have certain desirable properties which are helpful when examining the properties of the SR algorithm based on the butterfly form. A strict butterfly matrix does not necessarily have these properties.

In [2, 4] an elimination process for computing the butterfly form of a symplectic matrix is given which uses elementary unitary symplectic transformations as well as non-unitary symplectic transformations. Here, we also consider a structure-preserving symplectic Lanczos method which creates the symplectic butterfly form if no breakdown occurs. Such a symplectic Lanczos method will suffer from the well-known numerical difficulties inherent to any Lanczos method for nonsymmetric matrices. In [2], a symplectic look-ahead Lanczos algorithm is presented which overcomes breakdown by giving up the strict butterfly form. Unfortunately, so far there do not exist eigenvalue methods that can make use of that special reduced form. Standard eigenvalue methods as QR or SR have to be employed resulting in a full symplectic matrix after only a few iteration steps. We propose to employ an implicit restart technique instead of a look-ahead mechanism in order to deal with the numerical difficulties of the symplectic Lanczos method. This approach is based on the fundamental work of Sorensen [35].

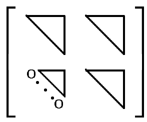
In Section 2, existence and uniqueness of the reduction of a symplectic matrix to butterfly form are reviewed. Unreduced butterfly matrices are introduced and their properties are presented. An SR algorithm based on the symplectic butterfly form is discussed in Section 3. The symplectic Lanczos method which reduces a symplectic matrix to butterfly form is derived in Section 4, where we also give the basic idea of an implicit restart for such a Lanczos process.

2 The Symplectic Butterfly Form

Here, we review the known results on existence and uniqueness of the reduction of a symplectic matrix to butterfly form and derive some new properties showing the analogy of the butterfly form to the Hessenberg form in generic chasing algorithms. As the reduction of a general matrix to upper Hessenberg form serves as a preparatory step for the QR algorithm, the reduction of a symplectic matrix to butterfly form can be used as a preparatory step for the SR algorithm. We will state results corresponding to those in the Hessenberg/QR-case for the symplectic butterfly form and the SR algorithm. Our main concern are symplectic matrices and the symplectic butterfly form, but we will briefly mention how the results presented here can be used for symplectic matrix pencils.

In order to state results concerning existence and uniqueness of the reduction of a symplectic matrix to symplectic butterfly form we need the following definitions. A matrix $A \in \mathbb{R}^{2n \times 2n}$ is called a *J-triangular* matrix if $A_{11}, A_{12}, A_{21}, A_{22} \in \mathbb{R}^{n \times n}$ are upper triangular matrices and A_{21} has a zero main diagonal, i.e.,

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} =$$



For a vector $v_1 \in \mathbb{R}^{2n}$ and $M \in \mathbb{R}^{2n \times 2n}$ define

$$K(M, v_1, \ell) := [v_1, M^{-1}v_1, M^{-2}v_1, \ldots, M^{-(\ell-1)}v_1, Mv_1, M^2v_1, \ldots, M^{\ell}v_1]. \tag{4}$$

(Note the similarity of this generalized Krylov matrix to the generalized ones in [9, 14, 15, 38].)

Further, let $P = P_n$ be the permutation matrix

$$P_n = [e_1, e_3, \ldots, e_{2n-1}, e_2, e_4, \ldots, e_{2n}] \in \mathbb{R}^{2n \times 2n}. \tag{5}$$

If the dimension of P_n is clear from the context, we leave off the superscript.

Theorem 2.1 *Let X be a $2n \times 2n$ nonsingular matrix. Let M and S be $2n \times 2n$ symplectic matrices and denote by v_1 the first column of S.*

a) There exists a $2n \times 2n$ symplectic matrix S and a J-triangular matrix R such that $X = SR$ if and only if all leading principal minors of even dimension of PX^TJXP^T are nonzero.

b) Let $X = SR$ and $X = \widetilde{S}\widetilde{R}$ be SR factorizations of X. Then there exists a symplectic matrix

$$D = \begin{bmatrix} C & F \\ 0 & C^{-1} \end{bmatrix}, \tag{6}$$

where $C = diag(c_1, \ldots, c_n)$, $F = diag(f_1, \ldots, f_n)$ such that $\widetilde{S} = SD^{-1}$ and $\widetilde{R} = DR$.

c) Let $K(M, v_1, n)$ be nonsingular. If $K(M, v_1, n) = SR$ is an SR decomposition then $S^{-1}MS$ is a butterfly matrix.

d) If $S^{-1}MS = B$ is a symplectic butterfly matrix then $K(M, v_1, n)$ has an SR decomposition $K(M, v_1, n) = SR$.

e) Let $S, \tilde{S} \in \mathbb{R}^{2n\times 2n}$ be symplectic matrices such that $S^{-1}MS = B$ and $\tilde{S}^{-1}M\tilde{S} = \tilde{B}$ are butterfly matrices. Then there exists a symplectic matrix D as in (6) such that $S = \tilde{S}D$ and $B = D\tilde{B}D^{-1}$.

Proof:

For the original statement and proof of a) see Theorem 11 in [13].

For the original statement and proof of b) see Proposition 3.3 in [10].

For the original statement and proof of c), d) and e) see Theorem 3.6 in [2]. $\surd$

The theorem introduces the *SR decomposition* of a matrix X. The SR decomposition has been studied, e.g., in [8, 10, 12]. Theorem 2.1 e) shows that the transformation to butterfly form is unique up to scaling with a matrix D as in (6). From the proof of c) it follows that the tridiagonal matrix in the lower right corner of the butterfly form is an unreduced tridiagonal matrix, that is, none of the upper and lower subdiagonal elements are zero. Similarly, one needs that these elements are nonzero to show in d) that R is nonsingular. Because of this we will call a symplectic matrix $B \in \mathbb{R}^{2n\times 2n}$ an *unreduced butterfly matrix* if

$$B = \begin{bmatrix} \mathcal{B}_1 & \mathcal{B}_2 \\ \mathcal{B}_3 & \mathcal{B}_4 \end{bmatrix} = \begin{bmatrix} \diagdown & \diagdown\!\diagdown\!\diagdown \\ \diagdown & \diagdown\!\diagdown\!\diagdown \end{bmatrix}, \tag{7}$$

where $\mathcal{B}_1, \mathcal{B}_3 \in \mathbb{R}^{n\times n}$ are diagonal matrices, $\mathcal{B}_2, \mathcal{B}_4 \in \mathbb{R}^{n\times n}$ are tridiagonal matrices, and $\mathcal{B}_4$ is unreduced, that is, the subdiagonal elements are nonzero.

Lemma 2.2 *If B as in (7) is an unreduced butterfly matrix, then $\mathcal{B}_3$ is nonsingular and B can be factored as*

$$\begin{bmatrix} \mathcal{B}_1 & \mathcal{B}_2 \\ \mathcal{B}_3 & \mathcal{B}_4 \end{bmatrix} = \begin{bmatrix} \mathcal{B}_3^{-1} & \mathcal{B}_1 \\ 0 & \mathcal{B}_3 \end{bmatrix} \begin{bmatrix} 0 & -I \\ I & \mathcal{B}_3^{-1}\mathcal{B}_4 \end{bmatrix} = \begin{bmatrix} \diagdown & \diagdown \\ 0 & \diagdown \end{bmatrix} \begin{bmatrix} 0 & -I \\ I & \diagdown\!\diagdown\!\diagdown \end{bmatrix}.$$

This factorization is unique. Note that $\mathcal{B}_3^{-1}\mathcal{B}_4$ is symmetric.

Proof :

The fact that B is symplectic implies $\mathcal{B}_1\mathcal{B}_4 - \mathcal{B}_3\mathcal{B}_2 = I$. Assume that $\mathcal{B}_3$ is singular, that is $(\mathcal{B}_3)_{jj} = 0$ for some j. Then the jth row of $\mathcal{B}_1\mathcal{B}_4 - \mathcal{B}_3\mathcal{B}_2 = I$ gives

$$(\mathcal{B}_1)_{jj}(\mathcal{B}_4)_{j,j-1} = 0, \quad (\mathcal{B}_1)_{jj}(\mathcal{B}_4)_{jj} = 1, \quad (\mathcal{B}_1)_{jj}(\mathcal{B}_4)_{j,j+1} = 0.$$

This can only happen for $(\mathcal{B}_1)_{jj} \neq 0$, $(\mathcal{B}_4)_{jj} \neq 0$, and $(\mathcal{B}_4)_{j,j-1} = (\mathcal{B}_4)_{j,j+1} = 0$, but $\mathcal{B}_4$ is unreduced. Hence $\mathcal{B}_3$ has to be nonsingular if $\mathcal{B}_4$ is unreduced. Thus, for an unreduced butterfly matrix we obtain

$$\begin{bmatrix} \mathcal{B}_3 & -\mathcal{B}_1 \\ 0 & \mathcal{B}_3^{-1} \end{bmatrix} \begin{bmatrix} \mathcal{B}_1 & \mathcal{B}_2 \\ \mathcal{B}_3 & \mathcal{B}_4 \end{bmatrix} = \begin{bmatrix} 0 & -I \\ I & \mathcal{B}_3^{-1}\mathcal{B}_4 \end{bmatrix}.$$

As both matrices on the left are symplectic, their product is symplectic and hence $\mathcal{B}_3^{-1}\mathcal{B}_4$ has to be a symmetric tridiagonal matrix. Thus

$$\begin{bmatrix} \mathcal{B}_1 & \mathcal{B}_2 \\ \mathcal{B}_3 & \mathcal{B}_4 \end{bmatrix} = \begin{bmatrix} \mathcal{B}_3^{-1} & \mathcal{B}_1 \\ 0 & \mathcal{B}_3 \end{bmatrix} \begin{bmatrix} 0 & -I \\ I & \mathcal{B}_3^{-1}\mathcal{B}_4 \end{bmatrix} = \begin{bmatrix} \ddots & \ddots \\ 0 & \ddots \end{bmatrix} \begin{bmatrix} 0 & -I \\ I & \ddots\ddots\ddots \end{bmatrix}.$$

The uniqueness of this factorization follows from the choice of signs of the identities. $\surd$

We will frequently make use of this decomposition and will denote it by

$$B_1 = \begin{bmatrix} \mathcal{B}_3^{-1} & \mathcal{B}_1 \\ 0 & \mathcal{B}_3 \end{bmatrix} = \left[\begin{array}{ccc|ccc} a_1^{-1} & & & b_1 & & \\ & \ddots & & & \ddots & \\ & & a_n^{-1} & & & b_n \\ \hline & & & a_1 & & \\ & & & & \ddots & \\ & & & & & a_n \end{array}\right], \tag{8}$$

$$B_2^{-1} = \begin{bmatrix} 0 & -I \\ I & \mathcal{B}_3^{-1}\mathcal{B}_4 \end{bmatrix} = \left[\begin{array}{ccc|cccc} & & & -1 & & & \\ & & & & \ddots & & \\ & & & & & \ddots & \\ & & & & & & -1 \\ \hline 1 & & & c_1 & d_2 & & \\ & \ddots & & d_2 & \ddots & \ddots & \\ & & \ddots & & \ddots & \ddots & d_n \\ & & 1 & & & d_n & c_n \end{array}\right], \tag{9}$$

$$
B = B_1 B_2^{-1} \;=\; \left[\begin{array}{cccc|cccc}
b_1 & & & & b_1c_1 - a_1^{-1} & b_1 d_2 & & \\
 & \ddots & & & b_2 d_2 & \ddots & \ddots & \\
 & & \ddots & & & \ddots & \ddots & b_{n-1}d_n \\
 & & & b_n & & & b_n d_n & b_n c_n - a_n^{-1} \\ \hline
a_1 & & & & a_1 c_1 & a_1 d_2 & & \\
 & \ddots & & & a_2 d_2 & \ddots & \ddots & \\
 & & \ddots & & & \ddots & \ddots & a_{n-1}d_n \\
 & & & a_n & & & a_n d_n & a_n c_n
\end{array}\right]. \qquad (10)
$$

From (8) – (10) we obtain

Corollary 2.3 *Any unreduced butterfly matrix* $B \in \mathbb{R}^{2n \times 2n}$ *can be represented by* $4n-1$ *parameters* $a_1, \ldots, a_n, d_2, \ldots, d_n \in \mathbb{R} \setminus \{0\}, b_1, \ldots, b_n, c_1, \ldots, c_n \in \mathbb{R}$.

Remark 2.4 *Any unreduced butterfly matrix is similar to an unreduced butterfly matrix with* $b_i = 1$ *and* $|a_i| = 1$ *for* $i = 1, \ldots, n$ *and* $sign(a_i) = sign(d_i)$ *for* $i = 2, \ldots n$ *(this follows from Theorem 2.1 e)).*

Remark 2.5 *We will have deflation if* $d_j = 0$ *for some* j. *Then the eigenproblem can be split into two smaller ones with unreduced symplectic butterfly matrices.*

The next result is well-known for Hessenberg matrices (e.g., [20, Theorem 7.4.4]) and will turn out to be essential when examining the properties of the SR algorithm based on the butterfly form.

Lemma 2.6 *If* λ *is an eigenvalue of an unreduced symplectic butterfly matrix* $B \in \mathbb{R}^{2n \times 2n}$, *then its geometric multiplicity is one.*

Proof :

Since B is symplectic, B is nonsingular and its eigenvalues are nonzero. For any $\lambda \in \mathbb{C}$ we have $\operatorname{rank}(B - \lambda I) \geq 2n - 1$ because the first $2n-1$ columns of $B - \lambda I$ are linear independent. This can be seen by looking at the permuted expression $B_P - \lambda I = PBP^T - \lambda I =$

$$
\left[\begin{array}{cc|cc|cc|cc|cc}
b_1 - \lambda & b_1c_1 - a_1^{-1} & 0 & b_1 d_2 & & & & & & \\
a_1 & a_1 c_1 - \lambda & 0 & a_1 d_2 & & & & & & \\ \hline
0 & b_2 d_2 & b_2 - \lambda & b_2 c_2 - a_2^{-1} & 0 & b_2 d_3 & & & & \\
0 & a_2 d_2 & a_2 & a_2 c_2 - \lambda & 0 & a_2 d_3 & & & & \\ \hline
 & & 0 & b_3 d_3 & b_3 - \lambda & b_3 c_3 - a_3^{-1} & \ddots & & & \\
 & & 0 & a_3 d_3 & a_3 & a_3 c_3 - \lambda & & \ddots & & \\ \hline
 & & & & \ddots & & \ddots & & 0 & b_{n-1} d_n \\
 & & & & & \ddots & & \ddots & 0 & a_{n-1} d_n \\ \hline
 & & & & & & 0 & b_n d_n & b_n - \lambda & b_n c_n - a_n^{-1} \\
 & & & & & & 0 & a_n d_n & a_n & a_n c_n - \lambda
\end{array}\right].
$$

Obviously, the first two columns of the above matrix are linear independent as B is unreduced. We can not express the third column as a linear combination of the first two columns:

$$\begin{bmatrix} 0 \\ 0 \\ b_2 - \lambda \\ a_2 \end{bmatrix} = \beta_1 \begin{bmatrix} b_1 - \lambda \\ a_1 \\ 0 \\ 0 \end{bmatrix} + \beta_2 \begin{bmatrix} b_1 c_1 - a_1^{-1} \\ a_1 c_1 - \lambda \\ b_2 d_2 \\ a_2 d_2 \end{bmatrix}.$$

From the fourth row we obtain $\beta_2 = d_2^{-1}$. With this the third row yields

$$b_2 - \lambda = b_2.$$

As λ is an eigenvalue of B and is therefore nonzero, this equation can not hold. Hence the first three columns are linear independent. Similarly, we can see that the first $2n - 1$ columns are linear independent.

Hence, the eigenspaces are one-dimensional. $\surd$

Remark 2.7 *In [2] a slightly different point of view is taken in order to argue that a butterfly matrix can be represented by* $4n - 1$ *parameters. There, a* strict butterfly form *is introduced in which the upper left diagonal matrix* $\mathcal{B}_1$ *of the butterfly form is nonsingular. Then, using similar arguments as above, since*

$$\begin{bmatrix} \mathcal{B}_1^{-1} & 0 \\ -\mathcal{B}_3 & \mathcal{B}_1 \end{bmatrix} \begin{bmatrix} \mathcal{B}_1 & \mathcal{B}_2 \\ \mathcal{B}_3 & \mathcal{B}_4 \end{bmatrix} = \begin{bmatrix} I & \mathcal{B}_1^{-1}\mathcal{B}_2 \\ 0 & I \end{bmatrix}$$

and $\mathcal{B}_1^{-1}\mathcal{B}_2$ *is a symmetric tridiagonal matrix (same argument as used above), one obtains*

$$\begin{bmatrix} \mathcal{B}_1 & \mathcal{B}_2 \\ \mathcal{B}_3 & \mathcal{B}_4 \end{bmatrix} = \begin{bmatrix} \mathcal{B}_1 & 0 \\ \mathcal{B}_3 & \mathcal{B}_1^{-1} \end{bmatrix} \begin{bmatrix} I & \mathcal{B}_1^{-1}\mathcal{B}_2 \\ 0 & I \end{bmatrix} = \begin{bmatrix} \diagdown & 0 \\ \diagdown & \diagdown \end{bmatrix} \begin{bmatrix} I & \diagdown\!\!\diagdown\!\!\diagdown \\ 0 & I \end{bmatrix}.$$

Therefore a strict butterfly matrix can be represented by $4n - 1$ *parameters.*

Unfortunately, strict butterfly matrices do not have all the desirable properties of unreduced butterfly matrices. In particular, Lemma 2.6 does not hold for strict butterfly matrices as can be seen by the next example.

Example 2.8 *Let*

$$B = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right].$$

Then B *is a strict symplectic butterfly matrix that it is not unreduced. It is easy to see that the spectrum of* B *is given by* $\{1, 1\}$ *with geometric multiplicities two.*

From Remark 2.4 and (10) it also follows that any unreduced symplectic butterfly matrix is similar to a strict butterfly matrix. But Example 2.8 also shows that the converse does not hold.

Finally consider a symplectic matrix pencil $L - \lambda N$, that is $LJL^T = NJN^T$, where $L, N \in \mathbb{R}^{2n\times 2n}$. If N is nonsingular, then $M = N^{-1}L$ is a symplectic matrix. The results of this section can be applied to M. Assume that S transforms M to unreduced butterfly form : $S^{-1}MS = B = B_1B_2^{-1}$. Then the symplectic matrix pencil $L - \lambda N$ is equivalent to the matrix pencil $Q(L - \lambda N)S = B_2^{-1} - \lambda B_1^{-1}$

$$\begin{bmatrix} 0 & -I \\ I & \ddots \end{bmatrix} - \lambda \begin{bmatrix} \ddots & \ddots \\ 0 & \ddots \end{bmatrix},$$

where $Q = B_1^{-1}S^{-1}N^{-1}$. Such a pencil is called a *symplectic butterfly pencil*. If N is singular, then the pencil $L-\lambda N$ has at least one eigenvalue 0 and ∞. Assume that there are k eigenvalues 0 and k eigenvalues ∞. In a preprocessing step these eigenvalues can be deflated out using Algorithm 2.17 in [16]. In the resulting symplectic pencil $L'-\lambda N'$ of dimension $2(n-k)\times 2(n-k)$, N' is nonsingular. Hence we can build $M' = (N')^{-1}L'$ and transform it to butterfly form $B' = B_1'(B_2')^{-1}$. Thus, $L' - \lambda N'$ is similar to the symplectic butterfly pencil $(B_2')^{-1} - \lambda(B_1')^{-1}$. Adding k rows and columns of zeros to each block of B_1' and B_2', and appropriate entries on the diagonals, we can expand the symplectic butterfly pencil $(B_2')^{-1} - \lambda(B_1')^{-1}$ to a symplectic butterfly pencil $\widehat{B}_2 - \lambda\widehat{B}_1$ of dimension $2n \times 2n$ that is equivalent to $L - \lambda N$.

3 The SR Algorithm for Symplectic Butterfly Matrices

Based on the SR decomposition introduced in Theorem 2.1 a symplectic QR-like method for solving eigenvalue problems of arbitrary real matrices is developed in [10]. The QR decomposition and the orthogonal similarity transformation to upper Hessenberg form in the QR process are replaced by the SR decomposition and the symplectic similarity reduction to J-Hessenberg form. Unfortunately, a symplectic matrix in butterfly form is not a J-Hessenberg matrix so that we can not simply use the results of [10] for computing the eigenvalues of a symplectic butterfly matrix. But, as we will see in this section, an SR step preserves the butterfly form. If B is an unreduced symplectic butterfly matrix, $p(B)$ a polynomial such that $p(B) \in \mathbb{R}^{2n\times 2n}$, $p(B) = SR$, and if R is invertible, then $S^{-1}BS$ is a symplectic butterfly matrix again. This was already noted and proved in [2], but no results for singular $p(B)$ are given there. The next theorem shows that singular $p(B)$ are desirable (that is at least one shift is an eigenvalue of B), as they allow the problem to be deflated after one step.

First, we need to introduce some notation. Let $p(B)$ be a polynomial such that $p(B) \in \mathbb{R}^{2n\times 2n}$. Write $p(B)$ in factored form

$$p(B) := (B - \lambda_1 I_{2n})(B - \lambda_2 I_{2n}) \cdots (B - \lambda_k I_{2n}). \qquad (11)$$

From $p(B) \in \mathbb{R}^{2n\times 2n}$ it follows that if $\mu \in \mathbb{C}$, and $\mu \in \{\lambda_1, \ldots, \lambda_k\}$, then $\overline{\mu} \in \{\lambda_1, \ldots, \lambda_k\}$. $p(B)$ is singular if and only if at least one of the shifts λ_i is an eigenvalue of B. Let ν denote the number of shifts that are equal to eigenvalues of B. Here we count a repeated shift according to its multiplicity as a zero of p, except that the number of times we count it must not exceed its algebraic multiplicity (as an eigenvalue of B).

Lemma 3.1 *Let $B \in \mathbb{R}^{2n\times 2n}$ be an unreduced symplectic butterfly matrix. The rank of $p(B)$ in (11) is $2n-\nu$ with ν as defined above.*

Proof :

> Since B is an unreduced butterfly matrix, its eigenspaces are one-dimensional by Lemma 2.6. Hence, we can use the same arguments as in the proof of Lemma 4.4 in [39] in order to prove the statement of this lemma. $\surd$

In the following we will consider only the case that $\mathrm{rank}(p(B))$ is even. In a real implementation, one would choose a polynomial p such that each perfect shift is accompanied by its reciprocal, since the eigenvalues of a symplectic matrix always appear in reciprocal pairs. As noted before if $\mu \in \mathbb{C}$ is a perfect shift, then we will choose $\overline{\mu}$ as a shift as well. That is in that case, we will choose $\mu, \mu^{-1}, \overline{\mu}$ and $\overline{\mu}^{-1}$ as shifts. Further, if $\mu \in \mathbb{R}$ is a perfect shift, then we choose μ^{-1} as a shift as well. Because of this, $\mathrm{rank}(p(B))$ will always be even.

Theorem 3.2 *Let $B \in \mathbb{R}^{2n\times 2n}$ be an unreduced symplectic butterfly matrix. Let $p(B)$ be a polynomial with $p(B) \in \mathbb{R}^{2n\times 2n}$ and $\mathrm{rank}(p(B)) = 2n-\nu =: 2k$. If $p(B) = SR$ exists, then $\widetilde{B} = S^{-1}BS$ is a symplectic matrix of the form*

$$\widetilde{B} = \left[\begin{array}{cc|cc} \diagdown & & \backslash\!\backslash\!\backslash & \\ & \Box & & \Box \\ \hline \diagdown & & \backslash\!\backslash\!\backslash & \\ & \Box & & \Box \end{array}\right] = \left[\begin{array}{cc|cc} \widetilde{B}_{11} & & \widetilde{B}_{13} & \\ & \widetilde{B}_{22} & & \widetilde{B}_{24} \\ \hline \widetilde{B}_{31} & & \widetilde{B}_{33} & \\ & \widetilde{B}_{42} & & \widetilde{B}_{44} \end{array}\right] \begin{array}{l} \}\,k \\ \}\,n-k \\ \}\,k \\ \}\,n-k \end{array},$$

$$\underbrace{\quad}_{k}\ \underbrace{\quad}_{n-k}\ \underbrace{\quad}_{k}\ \underbrace{\quad}_{n-k}$$

where $\begin{bmatrix} \widetilde{B}_{11} & \widetilde{B}_{13} \\ \widetilde{B}_{31} & \widetilde{B}_{33} \end{bmatrix}$ is a symplectic butterfly matrix and the eigenvalues of $\begin{bmatrix} \widetilde{B}_{22} & \widetilde{B}_{24} \\ \widetilde{B}_{42} & \widetilde{B}_{44} \end{bmatrix}$ are just the ν shifts that are eigenvalues of B.

In order to simplify the notation for the proof of this theorem and the subsequent derivations, we use in the following permuted versions of B, R, and S. Let

$$B_P = PBP^T, \quad R_P = PRP^T, \quad S_P = PSP^T, \quad J_P = PJP^T,$$

with P as in (5).

From $S^TJS = J$ we obtain

$$S_P^T J_P S_P = J_P = \left[\begin{array}{cc|cc|c|cc} 0 & 1 & & & & & \\ -1 & 0 & & & & & \\ \hline & & 0 & 1 & & & \\ & & -1 & 0 & & & \\ \hline & & & & \ddots & & \\ \hline & & & & & 0 & 1 \\ & & & & & -1 & 0 \end{array}\right]$$

while the permuted butterfly matrix B_P is of the form

$$B_P = \left[\begin{array}{cc|cc|cc|cc}
b_1 & b_1c_1 - a_1^{-1} & 0 & b_1d_2 & & & & \\
a_1 & a_1c_1 & 0 & a_1d_2 & & & & \\
\hline
0 & b_2d_2 & b_2 & b_2c_2 - a_2^{-1} & \ddots & & & \\
0 & a_2d_2 & a_2 & a_2c_2 & & \ddots & & \\
\hline
 & & \ddots & & \ddots & & 0 & b_{n-1}d_n \\
 & & & \ddots & & \ddots & 0 & a_{n-1}d_n \\
\hline
 & & & & 0 & b_nd_n & b_n & b_nc_n - a_n^{-1} \\
 & & & & 0 & a_nd_n & a_n & a_nc_n
\end{array}\right]. \tag{12}$$

Proof of Theorem 3.2:

B_P is an upper triangular matrix with two additional subdiagonals, where the second additional subdiagonal has a nonzero entry only in every other position (see (12)). Since R is a J-triangular matrix, R_P is an upper triangular matrix. In the following, we denote by Z^{2k} the first $2k$ columns of a $2n \times 2n$ matrix Z, while Z^{rest} denotes its last $2n - 2k$ columns. $Z^{2k,2k}$ denotes the leading $2k \times 2k$ principal submatrix of a $2n \times 2n$ matrix Z.

Now partition the permuted matrices B_P, S_P, J_P, and R_P as

$$\begin{array}{rclrcl}
B_P & = & [B_P^{2k} \mid B_P^{rest}], & S_P & = & [S_P^{2k} \mid S_P^{rest}], \\
J_P & = & [J_P^{2k} \mid J_P^{rest}], & R_P & = & \left[\begin{array}{c|c} R_P^{2k,2k} & X \\ \hline 0 & Y \end{array}\right],
\end{array}$$

where the matrix blocks are defined as before; $X \in \mathbb{R}^{2k \times 2(n-k)}$, $Y \in \mathbb{R}^{2(n-k) \times 2(n-k)}$.

First we will show that the first $2k$ columns and rows of $\widetilde{B}_P$ are in the desired form. We will need the following observations. The first $2k$ columns of $p(B_P)$ are linear independent, since B is unreduced. To see this, consider the following identity:

$$\begin{aligned}
p(B)K(B, e_1, n) = & \\
&= [p(B)e_1, p(B)B^{-1}e_1, \ldots, p(B)B^{-(n-1)}e_1, p(B)Be_1, \ldots, p(B)B^n e_1] \\
&= [p(B)e_1, B^{-1}p(B)e_1, \ldots, B^{-(n-1)}p(B)e_1, Bp(B)e_1, \ldots, B^n p(B)e_1] \\
&= K(B, p(B)e_1, n),
\end{aligned}$$

where we have used $p(B)B^r = B^r p(B)$ for $r = \pm 1, \ldots, \pm n$. From Theorem 2.1 d) we know that, since B is unreduced, $K(B, e_1, n)$ is a nonsingular upper J-triangular matrix. As $\text{rank}(p(B)) = 2k$, $K(B, p(B)e_1, n)$ has rank $2k$. If a matrix of the form $K(X, v, n) = [v, X^{-1}v, \ldots, X^{-(n-1)}v, Xv, \ldots, X^n v]$ has rank $2k$, then the columns

$$v, X^{-1}v, \ldots, X^{-(k-1)}v, Xv, \ldots, X^k v$$

are linear independent. Further we obtain

$$p(B) = K(B, p(B)e_1, n)(K(B, e_1, n))^{-1} =: [p_1, p_2, \ldots, p_{2n}].$$

Due to the special form of $K(B, e_1, n)$ (J-triangular!) and the fact that the columns 1 to k and $n+1$ to $n+k$ of $K(B, p(B)e_1, n)$ are linear independent, the columns

$$p_1, p_2, \ldots, p_k, p_{n+1}, p_{n+2}, \ldots, p_{n+k}$$

of $p(B)$ are linear independent. Hence the first $2k$ columns of $p(B_P) = Pp(B)P^T$ are linear independent.

The columns of S_P^{2k} are linear independent, since S_P is nonsingular. Hence the matrix $R_P^{2k,2k}$ is nonsingular, since

$$p(B_P)I^{2k} = S_P^{2k} R_P^{2k,2k}.$$

It follows that

$$S_P^{2k} = p(B_P)I^{2k}(R_P^{2k,2k})^{-1}. \qquad (13)$$

Moreover, since $\text{rank}(p(B)) = 2k$, we have that $\text{rank}(R_P) = 2k$. Since $\text{rank}(R_P^{2k,2k}) = 2k$, we obtain $\text{rank}(Y) = 0$ and therefore $Y = 0$. From this we see

$$R_P = \left[\begin{array}{c|c} R_P^{2k,2k} & X \\ \hline 0 & 0 \end{array}\right]. \qquad (14)$$

Further we need the following identities

$$\begin{aligned}
B_P p(B_P) &= p(B_P)B_P, & (15)\\
B_P^{-1} p(B_P) &= p(B_P)B_P^{-1}, & (16)\\
B_P^T J_P^{-1} &= J_P^{-1} B_P^{-1}, & (17)\\
B_P^{-1} &= J_P^{-1} B_P^T J_P, & (18)\\
S_P^T J_P^{-1} &= J_P^{-1} S_P^{-1}, & (19)\\
S_P^{-1} &= J_P^{-1} S_P^T J_P, & (20)\\
S_P J_P^{2k} &= S_P^{2k} J_P^{2k,2k}. & (21)
\end{aligned}$$

Equations (17) – (21) follow from the fact that B and S are symplectic while (15) – (16) result from the fact that Z and $p(Z)$ commute for any matrix Z and any polynomial p.

The first $2k$ columns of $\widetilde{B}_P$ are given by the expression

$$\begin{aligned}
\widetilde{B}_P^{2k} &= \widetilde{B}_P I^{2k} \\
&= S_P^{-1} B_P S_P I^{2k} \\
&= S_P^{-1} B_P S_P^{2k} \\
&= S_P^{-1} B_P p(B_P) I^{2k} (R_P^{2k,2k})^{-1} & \text{by} \quad (13)\\
&= S_P^{-1} p(B_P) B_P^{2k} (R_P^{2k,2k})^{-1} & \text{by} \quad (15)\\
&= R_P B_P^{2k} (R_P^{2k,2k})^{-1}
\end{aligned}$$

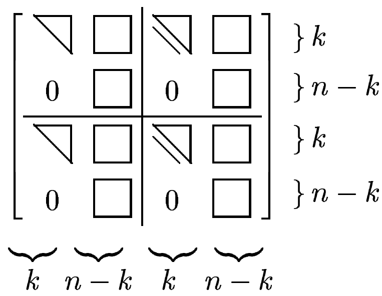
$$
= \left[\begin{array}{cc|cc|cc|cc}
x & x & x & x & \cdots & \cdots & x & x \\
x & x & x & x & \cdots & \cdots & x & x \\
\hline
0 & x & x & x & & & \vdots & \vdots \\
0 & x & x & x & & & \vdots & \vdots \\
\hline
 & & \ddots & & \ddots & & x & x \\
 & & & \ddots & & \ddots & x & x \\
 & & & & 0 & x & x & x \\
 & & & & 0 & x & x & x \\
 & & & & & & 0 & 0 \\
 & & & & & & 0 & 0 \\
\hline
0 & & & \cdots & & & & 0 \\
\vdots & & & & & & & \vdots \\
0 & & & \cdots & & & & 0
\end{array}\right].
$$

For the last equation we used (14), that $(R_P^{2k,2k})^{-1}$ is a $2k \times 2k$ upper triangular matrix, and that B_P is of the form given in (12). Hence

$$
\widetilde{B}_P = \left[\begin{array}{cc|cc|cc|cc|cc|cc|cc}
x & x & x & x & \cdots & \cdots & x & x & x & x & \cdots & \cdots & x & x \\
x & x & x & x & \cdots & \cdots & x & x & x & x & \cdots & \cdots & x & x \\
\hline
0 & x & x & x & & & \vdots & \vdots & x & x & \cdots & \cdots & x & x \\
0 & x & x & x & & & \vdots & \vdots & x & x & \cdots & \cdots & x & x \\
\hline
 & & \ddots & & \ddots & & x & x & x & x & \cdots & \cdots & x & x \\
 & & & \ddots & & \ddots & x & x & x & x & \cdots & \cdots & x & x \\
 & & & & 0 & x & x & x & x & x & \cdots & \cdots & x & x \\
 & & & & 0 & x & x & x & x & x & \cdots & \cdots & x & x \\
 & & & & & & 0 & 0 & x & x & \cdots & \cdots & x & x \\
 & & & & & & 0 & 0 & x & x & \cdots & \cdots & x & x \\
 & & & & & & & & x & x & \cdots & \cdots & x & x \\
 & & & & & & & & \vdots & \vdots & & & \vdots & \vdots \\
 & & & & & & & & x & x & \cdots & \cdots & x & x
\end{array}\right]
$$

and thus,

$\widetilde{B} \;=$

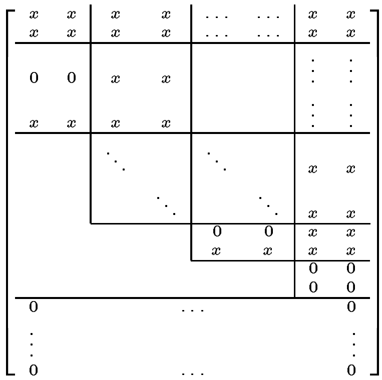


. (22)

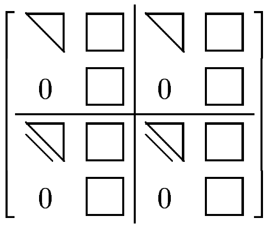
The first $2k$ columns of $(\widetilde{B}_P)^T$ are given by the expression

$$
\begin{aligned}
(\widetilde{B}_P)^T I^{2k} &= S_P^T B_P^T S_P^{-T} I^{2k} \\
&= S_P^T B_P^T J_P^{-1} S_P J_P I^{2k} && \text{by} \quad (20) \\
&= S_P^T B_P^T J_P^{-1} S_P J_P^{2k} \\
&= S_P^T B_P^T J_P^{-1} S_P^{2k} J_P^{2k,2k} && \text{by} \quad (13) \\
&= S_P^T B_P^T J_P^{-1} p(B_P) \left[\frac{(R_P^{2k,2k})^{-1}}{0}\right] J_P^{2k,2k} && \text{by} \quad (21)
\end{aligned}
$$

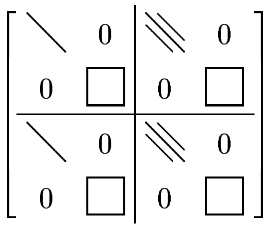
$$
\begin{aligned}
&= S_P^T J_P^{-1} B_P^{-1} p(B_P) \left[\frac{(R_P^{2k,2k})^{-1}}{0}\right] J_P^{2k,2k} && \text{by} \quad (17)\\
&= S_P^T J_P^{-1} p(B_P) B_P^{-1} \left[\frac{(R_P^{2k,2k})^{-1}}{0}\right] J_P^{2k,2k} && \text{by} \quad (16)\\
&= J_P^{-1} S_P^{-1} p(B_P) B_P^{-1} \left[\frac{(R_P^{2k,2k})^{-1}}{0}\right] J_P^{2k,2k} && \text{by} \quad (19)\\
&= J_P^{-1} R_P B_P^{-1} \left[\frac{(R_P^{2k,2k})^{-1}}{0}\right] J_P^{2k,2k} \\
&= (J_P^{-1} R_P J_P^{-1}) B_P^T (J_P \left[\frac{(R_P^{2k,2k})^{-1}}{0}\right] J_P^{2k,2k}) && \text{by} \quad (18)
\end{aligned}
$$

$=$  .

For the last equation we used (14), that $(R_P^{2k,2k})^{-1}$ a $2k \times 2k$ is an upper triangular matrix, and that B_P is of the form (12). Hence, we can conclude that

$\widetilde{B}^T =$  , (23)

where the blocks have the same size as before. Comparing (22) and (23) we obtain

$\widetilde{B} =$  .

This proves the first part of the theorem. The result about the eigenvalues now follows with the arguments as in the proof of Theorem 4.5 in [39]. There, a similar statement for a generic chasing algorithm is proved. $\surd$

Algorithm : SR algorithm for symplectic butterfly matrices
$B_1 := B =$ symplectic butterfly matrix
for $j = 1, 2, \ldots$ until satisfied
Choose polynomial p such that $p_j(B_j) \in \mathbb{R}^{2n\times 2n}$.
Compute $p_j(B_j) = S_j R_j = SR$ decomposition.
Set $B_{j+1} := S_j^{-1} B_j S_j$.

Table 1: SR algorithm for symplectic butterfly matrices.

Hence, assuming its existence, the SR decomposition and the SR step (that is, $B := S^{-1}BS$) possess many of the desirable properties of the QR step. An SR algorithm can thus be formulated similarly to the QR algorithm [8, 10]. In Table 1 we present a general SR algorithm for symplectic butterfly matrices.

There are different possibilities to choose the polynomial p_j in the algorithm given in Table 1, e.g.:

- single shift: $p(B) = B - \mu I$ for $\mu \in \mathbb{R}$;
- double shift: $p(B) = (B - \mu I)(B - \overline{\mu} I)$ for $\mu \in \mathbb{C}$, or
  $p(B) = (B - \mu I)(B - \frac{1}{\mu} I)$ for $\mu \in \mathbb{R}$;
- quadruple shift: $p(B) = (B - \mu I)(B - \overline{\mu} I)(B - \frac{1}{\mu} I)(B - \overline{\frac{1}{\mu}} I)$, for $\mu \in \mathbb{C}$.

In particular the double shift for $\mu \in \mathbb{R}$ and the quadruple shift for $\mu \in \mathbb{C}$ make use of the symmetries of the spectrum of symplectic matrices.

An algorithm for explicitly computing an SR decomposition for general matrices is presented in [10]. As with explicit QR steps, the expense of explicit SR steps comes from the fact that $p(B)$ has to be computed explicitly. A preferred alternative is the implicit SR step, an analogue to the Francis QR step [17, 20, 23]. The first implicit transformation S_1 is selected so that the first columns of the implicit and the explicit S are equivalent. That is, a symplectic matrix S_1 is determined such that

$$S_1^{-1} p(B) e_1 = \alpha_1 e_1, \qquad \alpha_1 \in \mathbb{R}.$$

Applying this first transformation to the butterfly matrix yields a symplectic matrix $S_1^{-1}BS_1$ with almost butterfly form having a small bulge. The remaining implicit transformations perform a bulge-chasing sweep down the subdiagonal to restore the butterfly form. That is, a symplectic matrix S_2 is determined such that $S_2^{-1}S_1^{-1}BS_1S_2$ is of butterfly form again.

Banse presents in [2] an algorithm to reduce an arbitrary symplectic matrix to butterfly form. Depending on the size of the bulge in $S_1^{-1}BS_1$, the algorithm can be greatly simplified to reduce $S_1^{-1}BS_1$ to butterfly form. The algorithm uses elementary symplectic Givens matrices [30]

$$G_k = \begin{bmatrix} C_k & -S_k \\ S_k & C_k \end{bmatrix},$$

where

$$C_k = I + (c_k - 1)e_k e_k^T, \qquad S_k = s_k e_k e_k^T, \qquad c_k^2 + s_k^2 = 1,$$

elementary symplectic Householder matrices [30]

$$H_k = \left[\begin{array}{cc|cc} I_{k-1} & & & \\ & P & & \\ \hline & & I_{k-1} & \\ & & & P \end{array}\right],$$

where

$$P = I_{n-k+1} - 2\frac{vv^T}{v^T v},$$

and elementary symplectic Gaussian elimination matrices [10]

$$L_k = \begin{bmatrix} W_k & V_k \\ 0 & W_k^{-1} \end{bmatrix},$$

where

$$W_k = I + (w_k - 1)(e_{k-1}e_{k-1}^T + e_k e_k^T), \qquad V_k = v_k(e_{k-1}e_k^T + e_k e_{k-1}^T).$$

As L_k is nonorthogonal, it might be ill-conditioned or might not even exist at all. This means that the SR decomposition of $p(B)$ does not exist or is close to the set of matrices for which an SR decomposition does not exist. As the set of these matrices is of measure zero [8], the polynomial p is discarded and an implicit SR shift with a random shift is performed as proposed in [10] in context of the Hamiltonian SR algorithm. For an actual implementation this might be realized by checking the condition number of L_k and performing an exceptional step if it exceeds a given tolerance.

The algorithm for reducing an arbitrary symplectic matrix to butterfly form as given in [2] can be summarized as given in Table 2 (in MATLAB-like notation). Note that pivoting is incorporated in order to increase numerical stability.

Algorithm : Reduction to butterfly form

input : $2n \times 2n$ symplectic matrix M
output : $2n \times 2n$ symplectic butterfly matrix M
 	
for $j = 1 : n-1$
    for $k = n : -1 : j+1$
        compute G_k such that $(G_k M)_{k+n,j} = 0$
        $M = G_k M G_k^{-1}$
    end
    if $j < n$
    then compute H_k such that $(H_k M)_{j+2:n,j} = 0$
            $M = H_k M H_k^{-1}$
    end
    compute L_{j+1} such that $(L_{j+1} M)_{j+1,j} = 0$
    $M = L_k M L_k^{-1}$
    if $\|M(j,j)\| > \|M(j+n,j)\|$
    then $p = j+n$
    else $p = j$
    end
    for $k = n : -1 : j+1$
        compute G_k such that $(M G_k)_{p,k} = 0$
        $M = G_k^{-1} M G_k$
    end
    if $j < n$
    then compute H_k such that $(M H_k)_{p,j+2+n:2n} = 0$
            $M = H_k^{-1} M H_k$
    end
end

Table 2: Reduction to butterfly form.

Let us assume for the moment that p is chosen to perform a quadruple shift. Then $p(B)e_1$ has eight nonzero entries

$$p(B)e_1 = [x, x, x, x, 0, \ldots, 0, x, x, x, x, 0, \ldots, 0]^T.$$

In order to compute S_1 such that $S_1^{-1} p(B)e_1 = \alpha e_1$, we have to eliminate the entries $n+1$ to $n+4$ by symplectic Givens transformations and the entries 2 to 4 by a symplectic Householder transformation. Hence $S_1^{-1}BS_1$ is of the form

$$
\left[\begin{array}{cccccccc|cccccccc}
x & + & + & + & & & & & x & x & + & + & + & & & \\
+ & x & + & + & & & & & x & x & x & + & + & & & \\
+ & + & x & + & & & & & + & x & x & x & + & & & \\
+ & + & + & x & & & & & + & + & x & x & x & & & \\
+ & + & + & + & x & & & & + & + & + & x & x & x & & \\
 & & & & & x & & & & & & & x & x & x & \\
 & & & & & & \ddots & & & & & & & & \ddots & \ddots \\
\hline
x & + & + & + & & & & & x & x & + & + & & & & \\
+ & x & + & + & & & & & x & x & x & + & & & & \\
+ & + & x & + & & & & & + & x & x & x & & & & \\
+ & + & + & x & & & & & + & + & x & x & x & & & \\
+ & + & + & + & x & & & & + & + & + & x & x & x & & \\
 & & & & & x & & & & & & & x & x & x & \\
 & & & & & & \ddots & & & & & & & & \ddots & \ddots
\end{array}\right],
$$

where a "+" denotes fill-in. Now the algorithm given in Table 2 can be used to reduce $S_1^{-1}BS_1$ to butterfly form again. Making use of all the zeros in $S_1^{-1}BS_1$ the given algorithm greatly simplifies. The resulting algorithm requires $O(n)$ floating point operations $(+, -, \times, /, \sqrt{\ })$ to restore the butterfly form.

Remark 3.3 *In order to implement such an implicit butterfly SR step, we do not need to form the intermediate symplectic matrices, but can apply the elimination matrices G_k, L_k, and H_k directly to the parameters $a_1, \ldots, a_n$, $b_1, \ldots, b_n$, $c_1, \ldots, c_n$, $d_2, \ldots, d_n$. In that case, we could also work directly with the symplectic butterfly pencil $B_1 - \lambda B_2$ with B_1, B_2 as in (8), (9).*

In [2] Banse also presents an algorithm to reduce a symplectic matrix pencil $L - \lambda N$, where L and N are symplectic matrices to a symplectic butterfly pencil $\widetilde{B}_1 - \lambda \widetilde{B}_2$. As in [2], strict butterfly matrices are used, this matrix pencil is of the form (see Remark 2.7)

$$
\begin{bmatrix} \diagdown & 0 \\ \diagdown & \diagdown \end{bmatrix} - \lambda \begin{bmatrix} I & \diagdown\!\!\diagdown\!\!\diagdown \\ 0 & I \end{bmatrix}.
$$

Hence we can not make direct use of this algorithm as our symplectic butterfly pencil is of the form

$$
\begin{bmatrix} \diagdown & \diagdown \\ 0 & \diagdown \end{bmatrix} - \lambda \begin{bmatrix} \diagdown\!\!\diagdown\!\!\diagdown & I \\ -I & 0 \end{bmatrix}
$$

but an algorithm based on our form can be derived in a similar way.

Working with the parameters is similar as in the parameterized SR algorithm given in [16] which is based on the parameterization of a symplectic J-Hessenberg matrix. This parameterization is determined by $4n - 1$ parameters. Besides using nonorthogonal elimination matrices, in order to obtain the parameterized version in [16], the explicit inversion of some of the matrix elements is necessary. Therefore, this parameterized SR algorithm "*...is highly numerically unstable*" [16]. We expect an implicit butterfly SR step to be more robust in the presence of roundoff errors as such explicit inversions can be avoided.

4 A Symplectic Lanczos Method for Symplectic Matrices

In this section, we describe a symplectic Lanczos method to compute the unreduced butterfly form (10) for a symplectic matrix M. A symplectic Lanczos method for computing a strict butterfly matrix is given in [2]. The usual nonsymmetric Lanczos algorithm generates two sequences of vectors. Due to the symplectic structure of M it is easily seen that one of the two sequences can be eliminated here and thus work and storage can essentially be halved. (This property is valid for a broader class of matrices, see [18].)

In order to simplify the notation we use in the following again the permuted versions of M and B as given by

$$M_P = PMP^T, \quad B_P = PBP^T, \quad S_P = PSP^T, \quad J_P = PJP^T,$$

with the permutation matrix P as in (5).

We want to compute a symplectic matrix S such that S transforms the symplectic matrix M to a symplectic butterfly matrix B. In the permuted version, $MS = SB$ yields

$$M_P S_P = S_P B_P. \tag{24}$$

Equivalently, as $B = B_1 B_2^{-1}$, we can consider

$$M_P S_P (B_2)_P = S_P (B_1)_P, \tag{25}$$

where

$$(B_1)_P = \left[\begin{array}{cc|c|cc} a_1^{-1} & b_1 & & & \\ 0 & a_1 & & & \\ \hline & & \ddots & & \\ \hline & & & a_n^{-1} & b_n \\ & & & 0 & a_n \end{array}\right] \tag{26}$$

$$(B_2)_P = \left[\begin{array}{cc|cc|cc|cc} c_1 & 1 & d_2 & 0 & & & & \\ -1 & 0 & 0 & 0 & & & & \\ \hline d_2 & 0 & c_2 & 1 & \ddots & & & \\ 0 & 0 & -1 & 0 & & \ddots & & \\ \hline & & \ddots & & \ddots & & d_n & 0 \\ & & & \ddots & & \ddots & 0 & 0 \\ \hline & & & & d_n & 0 & c_n & 1 \\ & & & & 0 & 0 & -1 & 0 \end{array}\right]. \tag{27}$$

The structure preserving Lanczos method generates a sequence of permuted symplectic matrices (that is, the columns of S^{2k} are J-orthogonal)

$$S_P^{2k} = [v_1, w_1, v_2, w_2, \ldots, v_k, w_k] \in \mathbb{R}^{2n \times 2k}$$

satisfying

$$M_P S_P^{2k} = S_P^{2k} B_P^{2k,2k} + d_{k+1}(b_{k+1}v_{k+1} + a_{k+1}w_{k+1})e_{2k}^T \tag{28}$$

or equivalently, as $B_P^{2k,2k} = (B_1^{2k,2k})_P (B_2^{2k,2k})_P^{-1}$ and $e_{2k}^T (B_2^{2k,2k})_P = -e_{2k-1}^T$, we have

$$M_P S_P^{2k} (B_2^{2k,2k})_P = S_P^{2k} (B_1^{2k,2k})_P - d_{k+1}(b_{k+1}v_{k+1} + a_{k+1}w_{k+1})e_{2k-1}^T. \tag{29}$$

Here, $B_P^{2k,2k} = P_k B^{2k,2k} P_k{}^T$ is a permuted $2k \times 2k$ symplectic butterfly matrix as in (24) and $(B_j^{2k,2k})_P = P_k (B_j^{2k,2k}) P_k{}^T$, $j = 1, 2$, is a permuted $2k \times 2k$ symplectic matrix of the form (26), resp. (27). The space spanned by the columns of $S^{2k} = P_n^T S_P^{2k} P_k$ is J–orthogonal, since $S_P^{2k^T} J_P^n S_P^{2k} = J_P^k$, where $P_j J^j P_j^T = J_P^j$ and J^j is a $2j \times 2j$ matrix of the form (3).

The vector $r_{k+1} := d_{k+1}(b_{k+1}v_{k+1} + a_{k+1}w_{k+1})$ is the *residual vector* and is J_P-orthogonal to the columns of S_P^{2k}, called *Lanczos vectors*. The matrix $B_P^{2k,2k} = J_P^{2k,2k}(S_P^{2k})^T J_P M_P S_P^{2k}$ is the J_P-orthogonal projection of M_P onto the range of S_P^{2k}. Equation (28) (resp. (29)) defines a length $2k$ Lanczos factorization of M_P. If the residual vector r_{k+1} is the zero vector, then equation (28) (resp. 29)) is called a *truncated Lanczos factorization* if $k < n$. Note that theoretically, r_{n+1} must vanish since $(S_P^{2n})^T J_P^n r_{n+1} = 0$ and the columns of S_P^{2n} form a J_P-orthogonal basis for $\mathbb{R}^{2n}$. In this case the symplectic Lanczos method computes a reduction to butterfly form.

Before developing the symplectic Lanczos method itself, we state the following theorem which explains that the symplectic Lanczos factorization is completely specified by the starting vector v_1.

Theorem 4.1 *Let two length $2k$ Lanczos factorizations be given by*

$$\begin{aligned} M_P S_P^{2k} &= S_P^{2k} B_P^{2k,2k} + d_{k+1}(b_{k+1}v_{k+1} + a_{k+1}w_{k+1})e_{2k}^T \\ M_P \widehat{S_P}^{2k} &= \widehat{S_P}^{2k} \widehat{B_P}^{2k,2k} + \hat{d}_{k+1}(\hat{b}_{k+1}\hat{v}_{k+1} + \hat{a}_{k+1}\hat{w}_{k+1})e_{2k}^T, \end{aligned}$$

where S_P^{2k}, $\widehat{S_P}^{2k}$ have J_P-orthogonal columns, $B_P^{2k,2k}$, $\widehat{B_P}^{2k,2k}$ are permuted unreduced symplectic butterfly matrices with

$$\begin{aligned} (B_P^{2k,2k})_{jj} &= (\widehat{B_P}^{2k,2k})_{jj} &= 1, \\ |(B_P^{2k,2k})_{j+1,j}| &= |(\widehat{B_P}^{2k,2k})_{j+1,j}| &= 1, \end{aligned} \qquad \text{for} \quad j = 1, 3, 5, \ldots, 2k-1,$$

$$\begin{aligned} (B_P^{2k,2k})_{j+1,j-1} &> 0, \\ (\widehat{B_P}^{2k,2k})_{j+1,j-1} &> 0, \end{aligned} \qquad \text{for} \quad j = 3, 5, \ldots, 2k-1,$$

and

$$J_P^{2k,2k}(S_P^{2k})^T J_P (b_{k+1}v_{k+1} + a_{k+1}w_{k+1}) = J_P^{2k,2k}(\widehat{S_P}^{2k})^T J_P (\hat{b}_{k+1}\hat{v}_{k+1} + \hat{a}_{k+1}\hat{w}_{k+1}) = 0.$$

If the first columns of S_P^{2k} *and* $\widehat{S_P}^{2k}$ *are equal, then* $B_P^{2k,2k} = \widehat{B_P}^{2k,2k}$, $S_P^{2k} = \widehat{S_P}^{2k}$, *and*

$$d_{k+1}(b_{k+1}v_{k+1} + a_{k+1}w_{k+1}) \;=\; \hat{d}_{k+1}(\hat{b}_{k+1}\hat{v}_{k+1} + \hat{a}_{k+1}\hat{w}_{k+1}).$$

Proof :

This is a direct consequence of Theorem 2.1 e) and Remark 2.4. $\surd$

Next we will see how the factorization (28) (resp. (29)) may be computed. As this reduction is strongly dependent on the first column of the transformation matrix that carries out the reduction, we must expect breakdown or near-breakdown in the Lanczos process as they also occur in the reduction process to J-Hessenberg form, e.g., [10]. Assuming that no such breakdowns occur, a symplectic Lanczos method can be derived as follows.

Let $S_P = [v_1, w_1, v_2, w_2, \ldots, v_n, w_n]$. For a given vector v_1, a Lanczos method constructs the matrix S_P columnwise from the equations

$$M_P S_P (B_2)_P e_j = S_P (B_1)_P e_j, \qquad j = 1, 2, \ldots \,.$$

That is, for even numbered columns

$$\begin{aligned} & & M_P v_m &= b_m v_m + a_m w_m \\ &\Longleftrightarrow & a_m w_m &= M_P v_m - b_m v_m \\ & & &=: \widetilde{w}_m \end{aligned} \tag{30}$$

and for odd numbered columns

$$\begin{aligned} & & a_m^{-1} v_m &= M_P(d_m v_{m-1} + c_m v_m - w_m + d_{m+1} v_{m+1}) \\ &\Longleftrightarrow & d_{m+1} v_{m+1} &= -d_m v_{m-1} - c_m v_m + w_m + a_m^{-1} M_P^{-1} v_m \\ & & &=: \widetilde{v}_{m+1}. \end{aligned} \tag{31}$$

Note that $M_P^{-1} = -J_P M_P^T J_P$, since M is symplectic. Thus $M_P^{-1} v_m$ is just a matrix-vector-product with the transpose of M_P.

Now we have to choose the parameters a_m, b_m, c_m, d_{m+1} such that $S_P^T J_P S_P = J_P$ is satisfied, that is, we have to choose the parameters such that $v_{m+1}^T J_P w_{m+1} = 1$. One possibility is to choose

$$d_{m+1} = ||\widetilde{v}_{m+1}||_2, \qquad a_{m+1} = v_{m+1}^T J_P M_P v_{m+1}.$$

Premultiplying $\widetilde{v}_{m+1}$ by $w_m^T J_P$ and using $S_P^T J_P S_P = J_P$ yields

$$c_m = -a_m^{-1} w_m^T J_P M_P^{-1} v_m = a_m^{-1} v_m^T J_P M_P w_m.$$

Thus we obtain the algorithm given in Table 3.

There is still some freedom in the choice of the parameters that occur in this algorithm. Essentially, the parameters b_m can be chosen freely. Here we set $b_m = 1$. Likewise a different choice of the parameters a_m, d_m is possible.

Algorithm : Symplectic Lanczos method
Choose an initial vector $\widetilde{v}_1 \in \mathbb{R}^{2n}, \widetilde{v}_1 \neq 0$. Set $v_0 = 0 \in \mathbb{R}^{2n}$. Set $d_1 = \|\|\widetilde{v}_1\|\|_2$ and $v_1 = \frac{1}{d_1}\widetilde{v}_1$. for m = 1, 2, ... do (update of w_m) set $\widetilde{w}_m = M_P v_m - b_m v_m$ $a_m = v_m^T J_P M_P v_m$ $w_m = \frac{1}{a_m}\widetilde{w}_m$ (computation of c_m) $c_m = a_m^{-1} v_m^T J_P M_P w_m$ (update of v_{m+1}) $\widetilde{v}_{m+1} = -d_m v_{m-1} - c_m v_m + w_m + a_m^{-1} M_P^{-1} v_m$ $d_{m+1} = \|\|\widetilde{v}_{m+1}\|\|_2$ $v_{m+1} = \frac{1}{d_{m+1}}\widetilde{v}_{m+1}$

Table 3: Symplectic Lanczos Method

Choosing $b_m = 0$, a different interpretation of the algorithm in Table 3 can be given. The resulting butterfly matrix $B = S^{-1}MS$ is of the form

$$\begin{bmatrix} 0 & -A \\ A & T \end{bmatrix},$$

where A is a diagonal matrix and T is an unsymmetric tridiagonal matrix. As $S^{-1}MS = B$, we have $S^{-1}M^{-1}S = B^{-1}$ and

$$S^{-1}(M + M^{-1})S = B + B^{-1} = \begin{bmatrix} -T^T & 0 \\ 0 & T \end{bmatrix}.$$

Obviously there is no need to compute both T and $-T^T$. It is sufficient to compute the first n columns of S. This corresponds to computing the v_m in our algorithm. This case is not considered here any further. See also [26].

Note that only one matrix-vector product is required for each computed Lanczos vector w_m or v_m. Thus an efficient implementation of this algorithm requires $6n + (4nz + 32n)k$ flops[1], where nz is the number of nonzero elements in M_P and $2k$ is the number of Lanczos vectors

[1] (Following [20], we define each floating point arithmetic operation together with the associated integer indexing as a flop.)

computed (that is, the loop is executed k times). The algorithm as given in Table 3 computes an odd number of Lanczos vectors, for a practical implementation one has to omit the computation of the last vector v_{k+1} (or one has to compute an additional vector w_{k+1}).

In the symplectic Lanczos method as given above we have to divide by parameters that may be zero or close to zero. If such a case occurs for the normalization parameter d_{m+1}, the corresponding vector $\widetilde{v}_{m+1}$ is zero or close to the zero vector. In this case, a symplectic invariant subspace of M (or a good approximation to such a subspace) is detected. By redefining $\widetilde{v}_{m+1}$ to be any vector satisfying

$$\begin{aligned} v_j^T J_P \widetilde{v}_{m+1} &= 0, \\ w_j^T J_P \widetilde{v}_{m+1} &= 0, \end{aligned}$$

for $j = 1, \ldots, m$, the algorithm can be continued. The resulting butterfly matrix is no longer unreduced; the eigenproblem decouples into two smaller subproblems. In case $\widetilde{w}_m$ is zero (or close to zero), an invariant subspace of M_P with dimension $2m-1$ is found (or a good approximation to such a subspace). From (30) it is easy to see that in this case the parameter a_m will be zero (or close to zero).

Thus, if either v_{m+1} or w_{m+1} vanishes, the breakdown is benign. If $v_{m+1} \neq 0$ and $w_{m+1} \neq 0$ but $a_{m+1} = 0$, then the breakdown is serious. No reduction of the symplectic matrix to a symplectic butterfly matrix with v_1 as first column of the transformation matrix exists. On the other hand, an initial vector v_1 exists so that the symplectic Lanczos process does not encounter serious breakdown. However, determining this vector requires knowledge of the minimal polynomial of M. Thus, no algorithm for successfully choosing v_1 at the start of the computation yet exists.

Furthermore, in theory, the above recurrences for v_m and w_m are sufficient to guarantee the J-orthogonality of theses vectors. Yet, in practice, the J-orthogonality will be lost, re-J-orthogonalization is necessary, increasing the computational cost significantly.

The numerical difficulties of the symplectic Lanczos method described above are inherent to all Lanczos-like methods for nonsymmetric matrices. Different approaches to overcome these difficulties have been proposed. Taylor [36] and Parlett, Taylor, and Liu [32] were the first to propose a look–ahead Lanczos algorithm that skips over breakdowns and near–breakdowns. Freund, Gutknecht, and Nachtigal present in [19] a look–ahead Lanczos code that can handle look–ahead steps of any length. Banse adapted this method to the symplectic Lanczos method given in [2]. The price paid is that the resulting matrix is no longer of butterfly form, but has a small bulge in the butterfly form to mark each occurrence of a (near) breakdown. Unfortunately, so far there exists no eigenvalue method that can make use of that special reduced form.

A different approach to deal with the numerical difficulties of the Lanczos process is to modify the starting vectors by an implicitly restarted Lanczos process (see the fundamental work in [11, 35]). The problems are addressed by fixing the number of steps in the Lanczos process at a prescribed value k which is dependent on the required number of approximate eigenvalues. J-orthogonality of the k Lanczos vectors is secured by re-J-orthogonalizing these vectors when necessary. The purpose of the implicit restart is to determine initial vectors such that the associated residual vectors are tiny. Given that a $2n \times 2k$ matrix S_P^{2k} is known such that

$$M_P S_P^{2k} = S_P^{2k} B_P^{2k,2k} + d_{k+1}(b_{k+1}v_{k+1} + a_{k+1}w_{k+1})e_{2k}^T \tag{32}$$

as in (28), an implicit Lanczos restart computes the Lanczos factorization

$$M_P \breve{S}_P^{2k} = \breve{S}_P^{2k} \breve{B}_P^{2k,2k} + \breve{d}_{k+1}(\breve{b}_{k+1}\breve{v}_{k+1} + \breve{a}_{k+1}\breve{w}_{k+1})e_{2k}^T \tag{33}$$

which corresponds to the starting vector

$$\breve{v}_1 = p(M_P)v_1$$

(where $p(M_P) \in \mathbb{R}^{2n\times 2n}$ is a polynomial) without having to explicitly restart the Lanczos process with the vector $\breve{v}_1$. Such an implicit restarting mechanism is derived in [5] analogous to the technique introduced in [6, 21, 35].

5 Concluding Remarks

Several aspects of the recently proposed, new condensed form for symplectic matrices, called the symplectic butterfly form, [4, 2, 3], are considered in detail. The $2n \times 2n$ symplectic butterfly form contains $8n - 4$ nonzero entries and is determined by $4n - 1$ parameters. The reduction to butterfly form can serve as a preparatory step for the SR algorithm, as the SR algorithm preserves the symplectic butterfly form in its iterations. Hence, its role is similar to that of the reduction of an arbitrary nonsymmetric matrix to upper Hessenberg form as a preparatory step for the QR algorithm. We have shown that an unreduced symplectic butterfly matrix in the context of the SR algorithm has properties similar to those of an unreduced upper Hessenberg matrix in the context of the QR algorithm. The SR algorithm not only preserves the symplectic butterfly form, but can be rewritten in terms of the $4n - 1$ parameters that determine the symplectic butterfly form. Therefore, the symplectic structure, which will be destroyed in the numerical computation due to roundoff errors, can be restored in each iteration step.

We have also briefly described an implicitly restarted symplectic Lanczos method which can be used to compute a few eigenvalues and eigenvectors of a symplectic matrix. The symplectic matrix is reduced to a symplectic butterfly matrix of lower dimension, whose eigenvalues can be used as approximations to the eigenvalues of the original matrix.

Acknowledgment Part of the work on this paper was carried out while the second author was visiting the University of California at Santa Barbara despite the fact that her office was located only 50 yards from the beach. She would like to thank Alan Laub for making this visit possible.

Both authors would like to thank David Watkins for many helpful suggestions which improved the paper significantly.

References

[1] G. S. Ammar and V. Mehrmann. On Hamiltonian and symplectic Hessenberg forms. *Linear Algebra Appl.*, 149:55–72, 1991.

[2] G. Banse. *Symplektische Eigenwertverfahren zur Lösung zeitdiskreter optimaler Steuerungsprobleme.* Dissertation, Universität Bremen, Fachbereich 3 – Mathematik und Informatik, Bremen, Germany, 1995.

[3] G. Banse. Condensed forms for symplectic matrices and symplectic pencils in optimal control. *ZAMM*, 76(Suppl. 3):375–376, 1996.

[4] G. Banse and A. Bunse-Gerstner. A condensed form for the solution of the symplectic eigenvalue problem. In U. Helmke, R. Menniken, and J. Sauer, editors, *Systems and Networks: Mathematical Theory and Applications*, pages 613–616. Akademie Verlag, 1994.

[5] P. Benner and H. Faßbender. An implicitly restarted symplectic Lanczos method for the symplectic eigenvalue problem. In preparation.

[6] P. Benner and H. Faßbender. An implicitly restarted symplectic Lanczos method for the Hamiltonian eigenvalue problem. *Linear Algebra Appl.*, to appear. *See also:* Tech. report SPC 95_28, Fak. f. Mathematik, Fak. f. Mathematik, TU Chemnitz–Zwickau, 09107 Chemnitz, FRG (1995).

[7] J.R. Bunch. The weak and strong stability of algorithms in numerical algebra. *Linear Algebra Appl.*, 88:49–66, 1987.

[8] A. Bunse-Gerstner. Matrix factorizations for symplectic QR-like methods. *Linear Algebra Appl.*, 83:49–77, 1986.

[9] A. Bunse-Gerstner and L. Elsner. Schur parameter pencils for the solution of the unitary eigenproblem. *Linear Algebra Appl.*, 154–156:741–778, 1991.

[10] A. Bunse-Gerstner and V. Mehrmann. A symplectic QR-like algorithm for the solution of the real algebraic Riccati equation. *IEEE Trans. Automat. Control*, AC-31:1104–1113, 1986.

[11] D. Calvetti, L. Reichel, and D. C. Sorensen. An implicitly restarted Lanczos method for large symmetric eigenvalue problems. *Electr. Trans. Num. Anal.*, 2:1–21, March 1994.

[12] J. Della-Dora. Numerical linear algorithms and group theory. *Linear Algebra Appl.*, 10:267–283, 1975.

[13] L. Elsner. On some algebraic problems in connection with general eigenvalue algorithms. *Linear Algebra Appl.*, 26:123–138, 1979.

[14] L. Elsner and K. Ikramov. On a normal form for normal matrices under finite sequences of unitary similarities. Submitted for publication, 1994.

[15] H. Faßbender. On numerical methods for discrete least-squares approximation by trigonometric polynomials. *Math. Comp.*, 66:719–741, 1997.

[16] U. Flaschka, V. Mehrmann, and D. Zywietz. An analysis of structure preserving methods for symplectic eigenvalue problems. *RAIRO Automatique Productique Informatique Industrielle*, 25:165–190, 1991.

[17] J. G. F. Francis. The QR transformation, Part I and Part II. *Comput. J.*, 4:265–271 and 332–345, 1961.

[18] R. Freund. Transpose-free quasi-minimal residual methods for non-Hermitian linear systems. In G. Golub et al., editor, *Recent advances in iterative methods. Papers from the IMA workshop on iterative methods for sparse and structured problems, held in Minneapolis, MN, February 24-March 1, 1992.*, volume 60 of *IMA Vol. Math. Appl.*, pages 69–94, New York, NY, 1994. Springer–Verlag.

[19] R. W. Freund, M. H. Gutknecht, and N. M. Nachtigal. An implementation of the look-ahead Lanczos algorithm for non-Hermitian matrices. *SIAM J. Sci. Comput.*, 14(1):137–158, January 1993.

[20] G.H. Golub and C.F. Van Loan. *Matrix Computations.* Johns Hopkins University Press, Baltimore, 2nd edition, 1989.

[21] E.J. Grimme, D.C. Sorensen, and P. Van Dooren. Model reduction of state space systems via an implicitly restarted Lanczos method. *Num. Alg.*, 12:1–31, 1996.

[22] D. Hinrichsen and N. K. Son. Stability radii of linear discrete-time systems and symplectic pencils. *Int. J. Robust Nonlinear Control*, 1:79–97, 1991.

[23] V. N. Kublanoskaja. On some algorithms for the solution of the complete eigenvalue problem. *U.S.S.R. Comput. Math. and Math. Phys.*, 3:637–657, 1961.

[24] P. Lancaster and L. Rodman. *The Algebraic Riccati Equation.* Oxford University Press, Oxford, 1995.

[25] A.J. Laub. Invariant subspace methods for the numerical solution of Riccati equations. In S. Bittanti, A.J. Laub, and J.C. Willems, editors, *The Riccati Equation*, pages 163–196. Springer-Verlag, Berlin, 1991.

[26] W.-W. Lin. A new method for computing the closed loop eigenvalues of a discrete-time algebraic Riccati equation. *Linear Algebra Appl.*, 6:157–180, 1987.

[27] V. Mehrmann. Der SR-Algorithmus zur Berechnung der Eigenwerte einer Matrix. Diplomarbeit, Universität Bielefeld, Bielefeld, FRG, 1979.

[28] V. Mehrmann. A symplectic orthogonal method for single input or single output discrete time optimal linear quadratic control problems. *SIAM J. Matrix Anal. Appl.*, 9:221–248, 1988.

[29] V. Mehrmann. *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution.* Number 163 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Heidelberg, July 1991.

[30] C.C. Paige and C.F. Van Loan. A Schur decomposition for Hamiltonian matrices. *Linear Algebra Appl.*, 14:11–32, 1981.

[31] T. Pappas, A.J. Laub, and N.R. Sandell. On the numerical solution of the discrete-time algebraic Riccati equation. *IEEE Trans. Automat. Control*, AC-25:631–641, 1980.

[32] B. N. Parlett, D. R. Taylor, and Z. A. Liu. A look-ahead Lanczos algorithm for unsymmetric matrices. *Math. Comp.*, 44(169):105–124, January 1985.

[33] R.V. Patel. Computation of the stable deflating subspace of a symplectic pencil using structure preserving orthogonal transformations. In *Proceedings of the 31st Annual Allerton Conference on Communication, Control and Computing*, University of Illinois, 1993.

[34] R.V. Patel. On computing the eigenvalues of a symplectic pencil. *Linear Algebra Appl.*, 188/189:591–611, 1993. See also: Proc. CDC-31, Tuscon, AZ, 1992, pp. 1921–1926.

[35] D. C. Sorensen. Implicit application of polynomial filters in a k-step Arnoldi method. *SIAM J. Matrix Anal. Appl.*, 13(1):357–385, January 1992.

[36] D. R. Taylor. *Analysis of the look ahead Lanczos algorithm.* PhD thesis, Center for Pure and Applied Mathematics, University of California, Berkeley, CA, 1982.

[37] C.F. Van Loan. A symplectic method for approximating all the eigenvalues of a Hamiltonian matrix. *Linear Algebra Appl.*, 16:233–251, 1984.

[38] D.S. Watkins. Some perspectives on the eigenvalue problem. *SIAM Rev.*, 35:430–471, 1993.

[39] D.S. Watkins and L. Elsner. Chasing algorithms for the eigenvalue problem. *SIAM J. Matrix Anal. Appl.*, 12:374–384, 1991.

[40] D.S. Watkins and L. Elsner. Convergence of algorithms of decomposition type for the eigenvalue problem. *Linear Algebra Appl.*, 143:19–47, 1995.