Two Connections between the SR and HR Eigenvalue Algorithms*

Peter Benner[†] Heike Faßbender[†] David S. Watkins[‡]

April 24, 1997

Abstract

The SR and HR algorithms are members of the family of GR algorithms for calculating eigenvalues and invariant subspaces of matrices. This paper makes two connections between the SR and HR algorithms: (1) An iteration of the SR algorithm on a $2n \times 2n$ symplectic butterfly matrix using shifts μ_i, μ_i^{-1}, $i = 1, \ldots, k$, is equivalent to an iteration of the HR algorithm on an $n \times n$ tridiagonal sign-symmetric matrix using shifts $\mu_i + \mu_i^{-1}$, $i = 1, \ldots, k$. (2) An iteration of the SR algorithm on a $2n \times 2n$ J-tridiagonal Hamiltonian matrix using shifts μ_i, $-\mu_i$, $i = 1, \ldots, k$, is equivalent to an iteration of the HR algorithm on an $n \times n$ tridiagonal sign-symmetric matrix using shifts μ_i^2, $i = 1, \ldots, k$.

The SR algorithm [7, 6] and the HR algorithm [4, 5] are members of the family of GR algorithms [25] for calculating eigenvalues and invariant subspaces of matrices. The oldest member of the family is Rutishauser's LR algorithm [18, 19, 20] and the most widely used is the QR algorithm [9, 11, 23, 24, 26]. The SR and HR algorithms are useful because they preserve certain special structures. The SR algorithm preserves Hamiltonian and symplectic matrices, and the HR algorithm preserves sign-symmetric tridiagonal matrices. In this paper we prove two interesting connections between SR and HR algorithms. (1) An iteration of the SR algorithm on a $2n \times 2n$ symplectic butterfly matrix using shifts μ_i, μ_i^{-1}, $i = 1, \ldots, k$, is equivalent to an iteration of the HR algorithm on an $n \times n$ tridiagonal sign-symmetric matrix using shifts $\mu_i + \mu_i^{-1}$, $i = 1, \ldots, k$. (2) An iteration of the SR algorithm on a $2n \times 2n$ J-tridiagonal Hamiltonian matrix using shifts μ_i, $-\mu_i$, $i = 1, \ldots, k$, is equivalent to an iteration of the HR algorithm on an $n \times n$ tridiagonal sign-symmetric matrix using shifts μ_i^2, $i = 1, \ldots, k$.

---

*Submitted for publication in April, 1997.

[†]Universität Bremen, Fachbereich 3 – Mathematik and Informatik, 28357 Bremen, FRG. Email: peter@mathematik.uni-bremen.de, heike@mathematik.uni-bremen.de

[‡]Department of Pure and Applied Mathematics, Washington State University, Pullman, Washington 99164-3113. Email: watkins@wsu.edu. Mailing address: 6835 24th Avenue NE, Seattle, WA 98115-7037. Supported by the National Science Foundation under grant DMS-9403569.

1 SR Basics

The SR algorithm is applicable to real matrices of even dimensions $2n \times 2n$. Throughout the paper we express such matrices as block matrices

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

in which the blocks A_{ij} are always $n \times n$. Let

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix},$$

where I denotes the $n \times n$ identity matrix. A matrix $S \in \mathbb{R}^{2n \times 2n}$ is *symplectic* if $S^T J S = J$ (or equivalently, $SJS^T = J$). If S is symplectic, then S is nonsingular, and $S^{-1} = JS^TJ^T$. The symplectic matrices form a group. The eigenvalues of symplectic matrices occur in reciprocal pairs: If λ is an eigenvalue of S with right eigenvector x, then λ^{-1} is an eigenvalue of S with left eigenvector $(Jx)^T$. Symplectic eigenvalue problems arise in discrete-time control, filtering, and estimation problems (see, e.g., [12, 14, 16, 21] and the references given therein), and the computation of discrete stability radii [10].

A matrix

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

is said to be *J-triangular* if the submatrices R_{ij} are all upper triangular, and R_{21} is strictly upper triangular. If one performs a perfect shuffle of the rows and columns of a J-triangular matrix, one gets an upper triangular matrix. The product of J-triangular matrices is J-triangular. The nonsingular J-triangular matrices form a group.

For the purposes of this paper, a matrix $M \in \mathbb{R}^{2n \times 2n}$ will be called *trivial* if it is both symplectic and J-triangular. M is trivial if and only if it has the form

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

where C and F are diagonal matrices, C nonsingular.

Almost every matrix $A \in \mathbb{R}^{2n \times 2n}$ can be decomposed into a product $A = SR$, where S is symplectic and R is J-triangular [8]. If this *SR decomposition* exists, then other SR decompositions of A can be built from it by passing trivial factors back and forth between S and R. That is, if M is a trivial matrix, $\tilde{S} = SM$ and $\tilde{R} = M^{-1}R$, then $A = \tilde{S}\tilde{R}$ is another SR decomposition of A. If A is nonsingular, then this is the only way to create other SR decompositions. In other words, the SR decomposition is unique up to trivial factors.

The SR algorithm is an iterative algorithm that performs an SR decomposition at each iteration. If B is the current iterate, then a *spectral transformation function* q is chosen and the SR decomposition of $q(B)$ is formed, if possible:

$$q(B) = SR.$$

Then the symplectic factor S is used to perform a similarity transformation on B to yield the next iterate, which we will call $\hat{B}$:

$$\hat{B} = S^{-1}BS. \tag{1}$$

We shall assume throughout this paper that $q(B)$ is nonsingular. Nothing bad happens in the singular case [3, 25]; we are avoiding it here solely to simplify the discussion. Since $q(B)$ is nonsingular, S is determined up to a trivial factor, so $\hat{B}$ is determined up to similarity transformation by a trivial matrix.

2 HR Basics

Now consider matrices in $\mathbb{R}^{n\times n}$. A *signature matrix* is a diagonal matrix $D = \text{diag}\{d_1, \ldots, d_n\}$ such that each d_i is either 1 or -1. Given a signature matrix D, we say that a matrix $A \in \mathbb{R}^{n\times n}$ is *D-symmetric* if $(DA)^T = DA$. A tridiagonal matrix T is D-symmetric for some D if and only if $|t_{i+1,i}| = |t_{i,i+1}|$ for $i = 1, \ldots, n-1$. Every irreducible tridiagonal matrix is similar to a D-symmetric matrix (for some D) by a diagonal similarity with positive main diagonal entries. D-symmetric tridiagonal matrices can be generated by the unsymmetric Lanczos process [13], for example.

Almost every $A \in \mathbb{R}^{n\times n}$ has an *HR decomposition* $A = HU$, in which U is upper triangular, and H satisfies the hyperbolic property $H^TDH = \hat{D}$, where $\hat{D}$ is another signature matrix [5]. For nonsingular A the HR decomposition is unique up to a signature matrix. We can make it unique by insisting that the upper triangular factor U satisfy $u_{ii} > 0$, $i = 1, \ldots, n$. The *HR algorithm* [4, 5] is an iterative process based on the HR decomposition. Choose a spectral transformation function p for which $p(A)$ is nonsingular, and form the HR decomposition of $p(A)$, if possible:

$$p(A) = HU.$$

Then use H to perform a similarity transformation on A to get the next iterate:

$$\hat{A} = H^{-1}AH.$$

The HR algorithm has the following structure preservation property: If A is D-symmetric and $H^TDH = \hat{D}$, then $\hat{A}$ is $\hat{D}$-symmetric. If A is also tridiagonal, then so is $\hat{A}$.

3 The Symplectic Case

We return our focus to symplectic matrices in $\mathbb{R}^{2n\times 2n}$. Because the symplectic matrices form a group, the SR algorithm preserves symplectic structure. That is, if the initial matrix is symplectic, then all iterates will be symplectic.

A symplectic matrix

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

is called a *butterfly* matrix if B_{11} and B_{21} are diagonal, and B_{12} and B_{22} are tridiagonal. Banse and Bunse-Gerstner [1, 2] showed that for every symplectic matrix M, there exist numerous symplectic matrices S such that $B = S^{-1}MS$ is a symplectic butterfly matrix. The SR algorithm preserves the butterfly form: If B is a symplectic butterfly matrix, then so is $\hat{B}$ in (1) [1, 2].

An *unreduced* symplectic butterfly matrix is one for which the tridiagonal submatrix B_{22} is irreducible [3]. Using the definition of a symplectic matrix, one easily verifies that if B is unreduced, then the diagonal submatrix B_{21} is nonsingular. This allows a decomposition of B into two simpler symplectic matrices:

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} B_{21}^{-1} & B_{11} \\ 0 & B_{21} \end{bmatrix} \begin{bmatrix} 0 & -I \\ I & T \end{bmatrix}, \tag{2}$$

where $T = B_{21}^{-1}B_{22}$ is an unreduced, symmetric, tridiagonal matrix. This decomposition of an unreduced symplectic butterfly matrix into a trivial matrix times a butterfly matrix of the special form $\begin{bmatrix} 0 & -I \\ I & T \end{bmatrix}$ is unique [3].

3.1 A Canonical Form for Symplectic Butterfly Matrices

We have noted that the symplectic butterfly form is preserved by the SR algorithm. The outcome of an SR iteration is not quite uniquely determined; it is determined up to a similarity transformation by a trivial (i.e. symplectic and J-triangular) matrix. It is therefore of interest to develop a canonical form for butterfly matrices under similarity transformations by trivial matrices. We restrict our attention to unreduced symplectic butterfly matrices, since every butterfly matrix can be decomposed into two or more smaller unreduced ones.

Theorem 1 *Let $\tilde{B}$ be an unreduced symplectic butterfly matrix. Then there exists a symplectic J-triangular matrix X such that $B = X^{-1}\tilde{B}X$ has the canonical form*

$$B = \begin{bmatrix} 0 & -D \\ D & T \end{bmatrix},$$

where D is a signature matrix, and T is a D-symmetric, unreduced tridiagonal matrix. D is uniquely determined, T is determined up to a similarity transformation by a signature matrix, and X is unique up to multiplication by a signature matrix of the form $\operatorname{diag}\{C, C\}$. The eigenvalues of T are $\lambda_i + \lambda_i^{-1}$, $i = 1, \ldots, n$, where λ_i, λ_i^{-1}, $i = 1, \ldots, n$ are the eigenvalues of B.

Proof. We are motivated by the decomposition (2), in which the nonsingular matrix B_{21} is used as a pivot to eliminate B_{11}. We now seek a similarity transformation that achieves a similar end. Let

$$X = \begin{bmatrix} Y^{-1} & -F \\ 0 & Y \end{bmatrix} \tag{3}$$

be a trivial matrix. We shall determined conditions on Y and F under which the desired canonical form is realized. Focusing on the first block column of the similarity transformation $B = X^{-1}\tilde{B}X$, we have

$$\begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} = \begin{bmatrix} Y & F \\ 0 & Y^{-1} \end{bmatrix} \begin{bmatrix} \tilde{B}_{11} \\ \tilde{B}_{21} \end{bmatrix} Y^{-1} = \begin{bmatrix} Y\tilde{B}_{11} + F\tilde{B}_{21} \\ Y^{-1}\tilde{B}_{21} \end{bmatrix} Y^{-1}.$$

We see that $B_{11} = 0$ if and only if $Y\tilde{B}_{11} + F\tilde{B}_{21} = 0$, which implies $F = -Y\tilde{B}_{11}\tilde{B}_{21}^{-1}$. Thus F is uniquely determined, once Y has been chosen. We have $B_{21} = Y^{-1}\tilde{B}_{21}Y^{-1}$, which shows that B_{21} and $\tilde{B}_{21}$ must have the same inertia. Thus the best we can do is take $B_{21} = D = \mathrm{sign}(\tilde{B}_{21})$, which is achieved by choosing $Y = |\,\tilde{B}_{21}\,|^{1/2}$.

In summary, we should take X as in (3), where

$$Y = |\,\tilde{B}_{21}\,|^{1/2} \quad \text{and} \quad F = -Y\tilde{B}_{11}\tilde{B}_{21}^{-1}.$$

The resulting B has the desired form. The only aspect of the computation that is not completely straightforward is showing that $B_{12} = -D$. However, this becomes easy when one applies the following fact: If B is a symplectic matrix with $B_{11} = 0$, then $B_{12} = -B_{21}^{-1}$. The D-symmetry of T is also an easy consequence of the symplectic structure of B.

The uniqueness statements are easily verified.

If $\begin{bmatrix} x \\ y \end{bmatrix}$ is an eigenvector of B with eigenvalue λ, then $y \neq 0$, and $Ty = (\lambda + \lambda^{-1})y$. ♣

Remarks

1. The canonical form could be made unique by insisting that T's subdiagonal entries be positive: $t_{i+1,i} > 0$, $i = 1, \ldots, n-1$.
2. The decomposition (2) of the canonical form B is

$$B = \begin{bmatrix} 0 & -D \\ D & T \end{bmatrix} = \begin{bmatrix} D & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} 0 & -I \\ I & DT \end{bmatrix}.$$

3. Theorem 1 is a theoretical result. From the standpoint of numerical stability, it might not be advisable to transform a symplectic butterfly matrix into canonical form. In the process, the spectral information λ, λ^{-1} is condensed into T as $\lambda + \lambda^{-1}$. The original information can be recovered via the inverse transformation

$$\nu \to \left(\frac{\nu}{2}\right) \pm \sqrt{\left(\frac{\nu}{2}\right)^2 - 1}.$$

However, eigenvalues near ± 1 will be resolved poorly because this map is not Lipschitz continuous at $\nu = \pm 2$.

The behavior is similar to that of Van Loan's method for the Hamiltonian eigenvalue problem [22] (see also the remarks in Section 4.1). One may lose up to half of the significant digits as compared to the standard QR algorithm. For instance, try to compute the eigenvalues of the symplectic matrix

$$S = G^T \begin{bmatrix} 1+\delta & 0 \\ 0 & \frac{1}{1+\delta} \end{bmatrix} G,$$

where G is a randomly generated Givens rotation and δ is less than the square root of the machine precision, once by applying the QR algorithm to S and once to $S + S^{-1}$ followed by the inverse transformation given above.

4. Since

$$B^{-1} = JB^TJ^T = \begin{bmatrix} T^T & D \\ -D & 0 \end{bmatrix},$$

we have $B + B^{-1} = \text{diag}\{T^T, T\}$. Thus, forming T is equivalent to adding B^{-1} to B. The transformation $S \to S + S^{-1}$ was used in similar fashion in [15, 17] to compute the eigenvalues of a symplectic pencil.

5. In the proof we have shown that the eigenvectors of T can obtained from those of B. It is also possible to recover the eigenvectors of B from those of T: If $Ty = (\lambda + \lambda^{-1})y$ $(y \neq 0)$, then

$$\begin{bmatrix} -\lambda^{-1}Dy \\ y \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -\lambda Dy \\ y \end{bmatrix}$$

are eigenvectors of B associated with λ and λ^{-1}, respectively.

♣

3.2 Equivalence of the HR and Symplectic SR Algorithms

Normally if one speaks of a GR iteration with shifts $\mu_1, \ldots, \mu_k$, one means that the iteration is carried out using the polynomial spectral transformation $p(B) = (B - \mu_1 I) \cdots (B - \mu_k I)$. In the case of an SR iteration on a symplectic matrix B, it makes sense to do things differently. Since the eigenvalues occur in reciprocal pairs, and the shifts are chosen to approximate eigenvalues, it is reasonable to choose the shifts in reciprocal pairs. If one wishes to effect the shifts μ and μ^{-1}, one can employ the factor $(B - \mu I)(B - \mu^{-1}I)$. However, one can equally well multiply the factor by B^{-1} and use

$$(B - \mu I)(I - \mu^{-1}B^{-1}) = (B + B^{-1}) - (\mu + \mu^{-1})I$$

instead. This modification allows better exploitation of the symplectic structure without changing the outcome of the SR iteration.

Thus, when we speak of an SR iteration with shifts μ_i, μ_i^{-1}, $i = 1, \dots, k$, applied to a symplectic matrix B, we shall mean that one forms the Laurent polynomial

$$q(B) = \prod_{i=1}^{k} (B - \mu_i I)(I - \mu_i^{-1} B^{-1}) = \prod_{i=1}^{k} \left((B + B^{-1}) - (\mu_i + \mu_i^{-1}) I \right),$$

next one performs the decomposition $q(B) = SR$, then one performs the similarity transformation $\hat{B} = S^{-1} B S$.

We are restricting ourselves to the nonsingular case, which means that none of the μ_i is allowed to be an eigenvalue of B. As we stated earlier, this is only for the sake of avoiding complications. Nothing bad happens in the singular case [3].

We allow complex shifts. However, if μ_i is not real, we insist that $\overline{\mu_i}$ should also appear in the list of shifts, so that $q(B)$ is real. In case that ℓ shifts are used in each SR or HR iteration step we say that the iteration is of *degree* ℓ.

Theorem 2 *Let* $B = \begin{bmatrix} 0 & -D \\ D & T \end{bmatrix}$ *be an unreduced symplectic butterfly matrix in canonical form. Then an* SR *iteration of degree* $2k$ *with shifts* μ_i, μ_i^{-1}, $i = 1, \dots, k$, *on* B *is equivalent to an* HR *iteration of degree* k *with shifts* $\mu_i + \mu_i^{-1}$, $i = 1, \dots, k$ *on the* D*-symmetric matrix* T.

Proof. The SR iteration has the form

$$q(B) = SR, \quad \hat{B} = S^{-1} B S,$$

where q is the Laurent polynomial

$$q(\lambda) = \prod_{i=1}^{k} \left[(\lambda + \lambda^{-1}) - (\mu_i + \mu_i^{-1}) \right].$$

Notice that $q(B) = p(B + B^{-1})$, where p is the ordinary polynomial

$$p(\nu) = \prod_{i=1}^{k} \left[\nu - (\mu_i + \mu_i^{-1}) \right].$$

Since $B + B^{-1} = \operatorname{diag}\{T^T, T\}$, we have

$$q(B) = \begin{bmatrix} p(T^T) & \\ & p(T) \end{bmatrix}. \tag{4}$$

An HR iteration on T with shifts $\mu_i + \mu_i^{-1}$, $i = 1, \dots, k$, has the form

$$p(T) = HU, \quad \hat{T} = H^{-1} T H.$$

U is upper triangular, H satisfies $H^T D H = \hat{D}$, where $\hat{D}$ is a signature matrix, and $\hat{T}$ is $\hat{D}$-symmetric.

Now let us relate this to the SR iteration on B. Since $Dp(T) = p(T^T)D$, we have

$$p(T^T) = DHUD = H^{-T}\hat{D}UD.$$

Thus

$$q(B) = \begin{bmatrix} p(T^T) & \\ & p(T) \end{bmatrix} = \begin{bmatrix} H^{-T} & \\ & H \end{bmatrix} \begin{bmatrix} \hat{D}UD & \\ & U \end{bmatrix}.$$

This is the SR decomposition of $q(B)$, for

$$S = \begin{bmatrix} H^{-T} & \\ & H \end{bmatrix}$$

is symplectic, and

$$R = \begin{bmatrix} \hat{D}UD & \\ & U \end{bmatrix}$$

is J-triangular. Using this SR decomposition to perform the SR iteration, we obtain

$$\hat{B} = S^{-1}BS = \begin{bmatrix} 0 & -\hat{D} \\ \hat{D} & \hat{T} \end{bmatrix}.$$

Thus the HR iteration on T is equivalent to the SR iteration on B. ♣

In principle we can compute the spectrum of a symplectic butterfly matrix by putting it into canonical form, calculating the eigenvalues of T, then inverting the transformation $\lambda \to \lambda + \lambda^{-1}$. Conversely, we can calculate the eigenvalues of a D-symmetric tridiagonal matrix T by embedding T and D in a symplectic butterfly matrix B, calculating the eigenvalues of B, and applying the transformation $\lambda \to \lambda + \lambda^{-1}$.

These transformations are not necessarily advisable from the standpoint of numerical stability. The first will resolve eigenvalues near ± 1 poorly because, as we already mentioned, the inverse transformation is not Lipschitz continuous. The second transformation is perhaps less objectionable. However, any eigenvalues of T that are near zero will have poor relative accuracy, because cancellation will occur in the transformation $\lambda \to \lambda + \lambda^{-1}$.

4 The Hamiltonian Case

A matrix

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \in \mathbb{R}^{2n \times 2n}$$

is called *Hamiltonian* if it satisfies $(JA)^T = JA$. One easily checks that A is Hamiltonian if and only if $A_{22} = -A_{11}^T$, and A_{21} and A_{12} are symmetric. Eigenvalues of real, Hamiltonian matrices appear in plus-minus pairs: If λ is an eigenvalue of A with right eigenvector x, then $-\lambda$ is an eigenvalue of A with left eigenvector $(Jx)^T$. If A is Hamiltonian and S is symplectic, then $S^{-1}AS$

is Hamiltonian. Thus the Hamiltonian form is preserved under iterations of the SR algorithm. Hamiltonian eigenvalue problems arise in a variety of continuous-time control, filtering, and estimation problems, see, e.g., [12, 14, 16, 21] and the references given therein.

A Hamiltonian matrix is in J*-tridiagonal* form if A_{11}, A_{22}, and A_{21} are diagonal, and A_{12} is tridiagonal. There exist numerous symplectic matrices S such that $S^{-1}AS$ is J-tridiagonal [6]. The SR algorithm preserves the Hamiltonian J-tridiagonal form.

An *unreduced* J-tridiagonal matrix is one for which A_{21} is nonsingular, and A_{12} is unreduced, that is, its subdiagonal entries are all nonzero.

4.1 A Canonical Form for Hamiltonian J-tridiagonal Matrices

Just as we did in the symplectic case, we now introduce a canonical form for unreduced J-tridiagonal matrices under similarity transformations by trivial matrices.

Theorem 3 *Let $\tilde{A}$ be an unreduced Hamiltonian J-tridiagonal matrix. Then there exists a symplectic J-triangular matrix X such that $A = X^{-1}\tilde{A}X$ has the canonical form*

$$A = \begin{bmatrix} 0 & V \\ D & 0 \end{bmatrix},$$

where D is a signature matrix, and V is a symmetric, irreducible tridiagonal matrix. D is uniquely determined, V is determined up to a similarity transformation by a signature matrix, and X is unique up to multiplication by a signature matrix of the form $\text{diag}\{C, C\}$. *Let T denote the D-symmetric matrix DV. The eigenvalues of T are λ_i^2, $i = 1, \ldots, n$, where λ_i, $-\lambda_i$, $i = 1, \ldots, n$ are the eigenvalues of A.*

Proof. Just as in the proof of Theorem 1, we can show that the transforming matrix

$$X = \begin{bmatrix} Y^{-1} & -F \\ 0 & Y \end{bmatrix},$$

with

$$Y = |\,\tilde{A}_{21}\,|^{1/2} \quad \text{and} \quad F = -Y\tilde{A}_{11}\tilde{A}_{21}^{-1},$$

results in an A whose first block column is of the desired form. The fact that the other block column also has the desired form follows from the fact that A is Hamiltonian and other elementary considerations.

As in the symplectic case, the uniqueness statements are easily verified.

If $\begin{bmatrix} x \\ y \end{bmatrix}$ is an eigenvector of A with eigenvalue λ, then $y \neq 0$, and $Ty = DVy = \lambda^2 y$. ♣

Remarks

1. The canonical form could be made unique by insisting that either T's or V's subdiagonal entries be positive.

2. From the standpoint of numerical stability, it might not be advisable to transform a Hamiltonian J-tridiagonal matrix into canonical form. In the process, the spectral information $\pm\lambda$ is condensed into T as λ^2. Any small eigenvalues of $\tilde{A}$ are transformed to tiny eigenvalues of T, which are then extremely vulnerable to roundoff errors in any subsequent computations on T.

3. We note that $A^2 = \mathrm{diag}\{T^T, T\}$. Thus, forming T is tantamount to squaring A. Squaring a Hamiltonian matrix to compute its eigenvalues is also the basis of Van Loan's square-reduced method [22]. An error estimate for retrieving the eigenvalues of a Hamiltonian matrix from their squares computed by Van Loan's method is given in [22] and indicates that one may lose up to half the significant digits as compared to a numerically backward stable method as the QR algorithm. The same limitations in accuracy apply if we transform a Hamiltonian J-tridiagonal matrix into canonical form and compute its eigenvalues via those of T.

4. As in the symplectic case, the eigenvectors of A can be recovered from those of T. If $Ty = \lambda^2 y$, $(y \neq 0)$, then $\begin{bmatrix} \pm\lambda D y \\ y \end{bmatrix}$ are eigenvectors of A associated with eigenvalues $\pm\lambda$.

♣

4.2 Equivalence of the HR and Hamiltonian SR Algorithms

Consider an SR iteration on a Hamiltonian matrix A. Since the eigenvalues occur in plus-minus pairs, it is reasonable to choose the shifts in plus-minus pairs. If we wish to effect an SR iteration of degree $2k$ with shifts $\pm\mu_i$, $i = 1, \ldots, k$, we use the polynomial

$$q(A) = \prod_{i=1}^{k} (A - \mu_i I)(A + \mu_i I) = \prod_{i=1}^{k} A^2 - \mu_i^2 I.$$

Again we restrict ourselves to the nonsingular case for simplicity. We also insist that complex shifts be present in conjugate pairs, so that $q(A)$ is real.

Theorem 4 *Let $A = \begin{bmatrix} 0 & V \\ D & 0 \end{bmatrix}$ be an unreduced Hamiltonian J-tridiagonal matrix in canonical form. Then an SR iteration of degree $2k$ with shifts $\pm\mu_i$, $i = 1, \ldots, k$, on A is equivalent to an HR iteration of degree k with shifts μ_i^2, $i = 1, \ldots, k$ on the D-symmetric matrix $T = DV$.*

Proof. The SR iteration has the form

$$q(A) = SR, \quad \hat{A} = S^{-1}AS,$$

where q is the polynomial

$$q(\lambda) = \prod_{i=1}^{k} (\lambda - \mu_i)(\lambda + \mu_i).$$

Notice that $q(A) = p(A^2)$, where

$$p(\nu) = \prod_{i=1}^{k} \left(\nu - \mu_i^2\right).$$

Since $A^2 = \mathrm{diag}\{T^T, T\}$, we have $q(A) = \mathrm{diag}\{p(T^T), p(T)\}$, in analogy with (4). Proceeding as in the proof of Theorem 2, we recall that an HR iteration on T with shifts μ_i^2, $i = 1, \ldots, k$, has the form

$$p(T) = HU, \quad \hat{T} = H^{-1}TH.$$

U is upper triangular, H satisfies $H^T DH = \hat{D}$, where $\hat{D}$ is a signature matrix, and $\hat{T}$ is $\hat{D}$-symmetric. Just as in the proof of Theorem 2, we have $p(T^T) = DHUD = H^{-T}\hat{D}UD$, so

$$q(A) = \begin{bmatrix} p(T^T) & \\ & p(T) \end{bmatrix} = \begin{bmatrix} H^{-T} & \\ & H \end{bmatrix} \begin{bmatrix} \hat{D}UD & \\ & U \end{bmatrix},$$

which is an SR decomposition of $q(A)$. Using this SR decomposition to perform the SR iteration, we obtain

$$\hat{A} = S^{-1}AS = \begin{bmatrix} 0 & \hat{V} \\ \hat{D} & 0 \end{bmatrix},$$

where $\hat{T} = \hat{D}\hat{V}$. Thus the HR iteration on T is equivalent to the SR iteration on A. ♣

In principle we can compute the spectrum of a Hamiltonian, J-tridiagonal matrix by putting it into canonical form, calculating the eigenvalues of $T = DV$, then taking square roots. We have already noted the dangers of this approach. Conversely, we can calculate the eigenvalues of a D-symmetric tridiagonal matrix T by embedding $V = DT$ and D in a Hamiltonian J-tridiagonal matrix A in canonical form, calculating the eigenvalues of A, and squaring them.

5 Conclusions

We have derived connections between the HR iteration for sign-symmetric matrices and the SR algorithms for symplectic butterfly and Hamiltonian J-tridiagonal matrices. Transforming symplectic butterfly and Hamiltonian J-tridiagonal matrices into the canonical forms introduced in Sections 3 and 4, it can be shown that the SR iterations for the so obtained matrices with a special choice of shifts are equivalent to an HR iteration on a sign-symmetric matrix of half the size. Using this approach it is possible to obtain the eigenvalues and eigenvectors of symplectic butterfly and Hamiltonian J-tridiagonal matrices by applying the HR algorithm to the associated sign-symmetric matrix.

The results are mainly of theoretical interest, as the resulting methods suffer from a possible loss of half the significant digits during the transformation to canonical form.

References

[1] G. Banse. *Symplektische Eigenwertverfahren zur Lösung zeitdiskreter optimaler Steuerungsprobleme*. PhD thesis, University of Bremen, 1995.

[2] 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.

[3] P. Benner and H. Faßbender. The symplectic eigenvalue problem, the butterfly form, the SR algorithm, and the Lanczos method. Submitted for publication, 1996.

[4] M. A. Brebner and J. Grad. Eigenvalues of $Ax = \lambda Bx$ for real symmetric matrices A and B computed by reduction to pseudosymmetric form and the HR process. *Linear Algebra Appl.*, 43:99–118, 1982.

[5] A. Bunse-Gerstner. An analysis of the HR algorithm for computing the eigenvalues of a matrix. *Linear Algebra Appl.*, 35:155–173, 1981.

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

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

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

[9] J. G. F. Francis. The QR transformation, parts I and II. *Computer J.*, 4:265–272, 332–345, 1961.

[10] 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.

[11] V. N. Kublanovskaya. On some algorithms for the solution of the complete eigenvalue problem. *USSR Comput. Math. and Math. Phys.*, 3:637–657, 1961.

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

[13] C. Lanczos. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *J. Res. Nat. Bur. Stand.*, 45:255–281, 1950.

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

[15] 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.

[16] 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.

[17] 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.

[18] H. Rutishauser. Der Quotienten-Differenzen-Algorithmus. *Z. angew. Math. Physik*, 5:233–251, 1954.

[19] H. Rutishauser. *Der Quotienten-Differenzen-Algorithmus.* Number 7 in Mitt. Inst. angew. Math. ETH. Birkhäuser, Basel, 1957.

[20] H. Rutishauser. Solution of eigenvalue problems with the LR-transformation. *Nat. Bur. Standards Appl. Math. Series*, 49:47–81, 1958.

[21] V. Sima. *Algorithms for Linear-Quadratic Optimization*, volume 200 of *Pure and Applied Mathematics.* Marcel Dekker, Inc., New York, NY, 1996.

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

[23] D. S. Watkins. Understanding the QR algorithm. *SIAM Rev.*, 24:427–440, 1982.

[24] D. S. Watkins. *Fundamentals of Matrix Computations.* John Wiley and Sons, New York, 1991.

[25] D. S. Watkins and L. Elsner. Convergence of algorithms of decomposition type for the eigenvalue problem. *Linear Algebra Appl.*, 143:19–47, 1991.

[26] J. H. Wilkinson. *The Algebraic Eigenvalue Problem,.* Clarendon Press, Oxford University, 1965.