ELA

http://math.technion.ac.il/iic/ela

STRUCTURE-PRESERVING SCHUR METHODS FOR COMPUTING SQUARE ROOTS OF REAL SKEW-HAMILTONIAN MATRICES[*]

ZHONGYUN LIU[†], YULIN ZHANG[‡], CARLA FERREIRA[‡], AND RUI RALHA[‡]

Abstract. The contribution in this paper is two-folded. First, a complete characterization is given of the square roots of a real nonsingular skew-Hamiltonian matrix W. Using the known fact that every real skew-Hamiltonian matrix has infinitely many real Hamiltonian square roots, such square roots are described. Second, a structure-exploiting method is proposed for computing square roots of W, skew-Hamiltonian and Hamiltonian square roots. Compared to the standard real Schur method, which ignores the structure, this method requires significantly less arithmetic.

1. Introduction. Given $A \in \mathbb{C}^{n \times n}$, a matrix X for which $X^2 = A$ is called a *square root* of A. The matrix square root is a useful theoretical and computational tool, one of the most commonly occurring matrix functions. See [11, 14, 16, 17, 20].

The theory behind the existence of matrix square roots is nontrivial and the feature which complicates this theory is that in general not all the square roots of a matrix A are functions of A. See [6, 20].

It is well known that certain matrix structures can be inherited by the square root. For example, a symmetric positive (semi)definite matrix has a unique symmetric positive (semi)definite square root [19]. The square roots of a centrosymmetric matrix are also centrosymmetric [23]. A nonsingular M-matrix has exactly one M-matrix as a square root. For an H-matrix with positive diagonal elements there exists one and only one square root which is also an H-matrix with positive diagonal elements

ELA

[21]. The principal square root of a centrosymmetric H-matrix with positive diagonal elements is a unique centrosymmetric H-matrix with positive diagonal entries [22]. Any real skew-Hamiltonian matrix has a real Hamiltonian square root [8].

For general matrices, an attractive method for computing matrix square roots is described by Björck and Hammarling [5]. This method uses the Schur decomposition but may require complex arithmetic. Higham [13] presented a modification of this method which enables real arithmetic to be used throughout when computing a real square root of a real matrix. This method has been extended to compute matrix pth roots [26] and general matrix functions [7].

In this paper, we will characterize the square roots of a real skew-Hamiltonian matrix W. We are mainly interested in square roots which are functions of the matrix and, as it is usually the case, we will not be concerned with singular matrices. We also propose a structure-exploiting method for computing square roots of W. This method uses the real skew-Hamiltonian Schur decomposition and requires significantly less arithmetic compared to the standard real Schur method. In [8], the problem of finding good numerical methods to compute Hamiltonian square roots for general skew-Hamiltonian matrices has been left as an open problem which, to our knowledge, has not been addressed until now. It is a basic tenet in numerical analysis that structure should be exploited allowing, in general, the development of faster and/or more accurate algorithms [4, 24].

We give some basic definitions and establish notation in Section 2. A description of the real Schur method and some results concerning the existence of real square roots are also presented in Section 2. In Section 3, we characterize the square roots of a nonsingular matrix W in a manner which makes clear the distinction between the square roots which are functions of W and those which are not. In Section 4, we present our algorithms for the computation of skew-Hamiltonian and Hamiltonian square roots. In Section 5, we present the results of some numerical experiments.

2. Definitions and previous results.

2.1. Square roots of a nonsingular matrix. It is a standard result that any matrix $A \in \mathbb{C}^{n \times n}$ can be expressed in the Jordan canonical form

$$(2.1) \qquad Z^{-1}AZ = J = \operatorname{diag}(J_1, J_2, \ldots, J_p),$$

where each $J_i = J_i(\lambda_k)$ is a Jordan block. We will use s to denote the number of distinct eigenvalues of A and so in (2.1) is $p \geq s$.

Given a scalar function f and a matrix $A \in \mathbb{C}^{n \times n}$ there are many different ways to define $f(A)$, a matrix of the same dimension of A, providing a useful generalization of a function of a scalar variable. The definition via Hermite interpolation defines

$$\boxed{\textbf{ELA}}$$

http://math.technion.ac.il/iic/ela

$f(A)$ to be a polynomial in the matrix A completely determined by the values of f on the spectrum of A. See [13, pp. 407–409].

Of particular interest is the scalar function $f(z) = z^{1/2}$ which is certainly defined on the spectrum of A if A is nonsingular. See [16, p. 3]. However, the square root function of A, $f(A)$, is not uniquely defined until one specifies which branch of the square root is to be taken in the neighborhood of each eigenvalue λ_i. Indeed, there are a total of 2^s matrices $f(A)$ when all combinations of branches for the square roots $f(\lambda_i)$, $i = 1, \ldots, s$, are taken. It is natural to ask whether these matrices are in fact square roots of A, that is, do we have $\left[f(A) \right]^2 = A$? Indeed, these matrices, which are polynomials in A by definition, are square roots of A [16, 20]. However, these square roots are not necessarily all the square roots of A.

The following result concerns the square roots of a Jordan block.

LEMMA 2.1. *For $\lambda_k \neq 0$ the Jordan block $J_i(\lambda_k)$ in (2.1) has precisely two upper triangular square roots*

$$(2.2) \quad L_k^{(j)} = L_k^{(j)}(\lambda_k) = \begin{bmatrix} f(\lambda_k) & f'(\lambda_k) & \cdot & \cdots & \frac{f^{(m_k-1)}(\lambda_k)}{(m_k-1)!} \\ & f(\lambda_k) & f'(\lambda_k) & \cdots & \frac{f^{(m_k-2)}(\lambda_k)}{(m_k-2)!} \\ & & \ddots & \ddots & \vdots \\ & & & f(\lambda_k) & f'(\lambda_k) \\ & & & & f(\lambda_k) \end{bmatrix}, \quad j = 1, 2,$$

where $f(\lambda) = \lambda^{1/2}$ and m_k is the order of J_i. The superscript j denotes the branch of the square root in the neighborhood of λ_k. Both square roots are functions of J_i and are the only square roots of J_i.

Next theorem shows that the square roots of a nonsingular matrix $A \in \mathbb{C}^{n \times n}$ which are functions of A are "isolated" square roots, and, on the other hand, the square roots which are not functions of A form a finite number of parametrized families of matrices. See [13, p. 410]. Proofs can be found in [11].

THEOREM 2.2. *Let the nonsingular matrix $A \in \mathbb{C}^{n \times n}$ have the Jordan canonical form (2.1) and let $s \leq p$ be the number of distinct eigenvalues of A. Then A has precisely 2^s square roots which are functions of A, given by*

$$(2.3) \qquad X_j = Z \operatorname{diag}\left(L_1^{(j_1)}, L_2^{(j_2)}, \ldots, L_p^{(j_p)} \right) Z^{-1}, \quad j = 1, \ldots, 2^s,$$

corresponding to all possible choices of $j_1, \ldots, j_p$, $j_k = 1$ or $j_k = 2$, subject to the constraint that $j_i = j_k$ whenever $\lambda_i = \lambda_k$.

If $s < p$, A has square roots which are not functions of A; they form parametrized

ELA

http://math.technion.ac.il/iic/ela

families

$$(2.4) \quad X_j(U) = ZU \operatorname{diag} \left(L_1^{(j_1)}, L_2^{(j_2)}, \ldots, L_p^{(j_p)} \right) U^{-1} Z^{-1}, \qquad j = 2^s + 1, \ldots, 2^p,$$

where $j_k = 1$ or $j_k = 2$, U is an arbitrary nonsingular matrix which commutes with J, and for each j there exist i and k, depending on j, such that $\lambda_i = \lambda_k$ while $j_i \neq j_k$.

Note that formula in (2.3) follows from the fact that all square roots of A which are functions of A have the form

$$f(A) = f(ZJZ^{-1}) = Zf(J)Z^{-1} = Z \operatorname{diag} \left(f(J_i) \right) Z^{-1},$$

and from Lemma 2.1. The remaining square roots of A (if any), which cannot be functions of A, are given by (2.4).

2.2. Hamiltonian and skew-Hamiltonian matrices. Hamiltonian and skew-Hamiltonian matrices have properties that follow directly from the definition.

DEFINITION 2.3.   Let $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$, where I is the identity matrix of order n.

(1) A matrix $H \in \mathbb{R}^{2n \times 2n}$ is said to be *Hamiltonian* if $HJ = (HJ)^T$. Equivalently, H can be partitioned as

$$(2.5) \qquad H = \begin{bmatrix} A & G \\ F & -A^T \end{bmatrix}, \quad G = G^T, \quad F = F^T, \quad A, G, F \in \mathbb{R}^{n \times n}.$$

(2) A matrix $W \in \mathbb{R}^{2n \times 2n}$ is said to be *skew-Hamiltonian* if $WJ = -(WJ)^T$. Likewise, W can be partitioned as

$$(2.6) \quad W = \begin{bmatrix} A & G \\ F & A^T \end{bmatrix}, \quad G = -G^T, \quad F = -F^T, \quad A, G, F \in \mathbb{R}^{n \times n}.$$

These matrix structures induce particular spectral properties for H and W. Notably, the eigenvalues of H are symmetric with respect to the imaginary axis and the eigenvalues of W have even algebraic and geometric multiplicities.

DEFINITION 2.4.   A matrix $S \in \mathbb{R}^{2n \times 2n}$ is said to be *symplectic* if $SJS^T = J$.

If $U \in \mathbb{R}^{2n \times 2n}$ is symplectic and orthogonal it can be partitioned as

$$U = \begin{bmatrix} U_1 & U_2 \\ -U_2 & U_1 \end{bmatrix},$$

where $U_i \in \mathbb{R}^{n \times n}$, $i = 1, 2$.

ELA

http://math.technion.ac.il/iic/ela

Hamiltonian and skew-Hamiltonian structures are preserved if symplectic similarity transformations are used; if H is Hamiltonian (skew-Hamiltonian) and S is symplectic, then $S^{-1}HS$ is also Hamiltonian (skew-Hamiltonian). In the interest of numerical stability, the similarities should be orthogonal as well [3].

The first simplifying reduction of a skew-Hamiltonian matrix was introduced by Van Loan and Paige in [25, 27]. For any skew-Hamiltonian matrix W we can compute an orthogonal and symplectic matrix U such that W is brought to block-upper-triangular form

$$U^T W U = \begin{bmatrix} W_1 & W_2 \\ O & W_1^T \end{bmatrix},$$

where $W_2^T = -W_2$ and W_1 is upper Hessenberg. This is called the *symplectic Paige/Van Loan (PVL) form*. Subsequently, if the standard QR algorithm is applied to W_1 producing an orthogonal matrix Q and a matrix in real Schur form N_1 so that

$$W_1 = Q N_1 Q^T,$$

we attain the *real skew-Hamiltonian Schur decomposition* of W via $\mathcal{U} = U \begin{bmatrix} Q & O \\ O & Q \end{bmatrix}$.

LEMMA 2.5 (Real skew-Hamiltonian Schur form). *Let $W \in \mathbb{R}^{2n \times 2n}$ be skew-Hamiltonian. Then there exists an orthogonal matrix*

$$\mathcal{U} = \begin{bmatrix} \mathcal{U}_1 & \mathcal{U}_2 \\ -\mathcal{U}_2 & \mathcal{U}_1 \end{bmatrix}, \quad \mathcal{U}_1, \mathcal{U}_2 \in \mathbb{R}^{n \times n},$$

such that

$$\mathcal{U}^T W \mathcal{U} = \begin{bmatrix} N_1 & N_2 \\ 0 & N_1^T \end{bmatrix}, \quad N_2^T = -N_2,$$

and N_1 is in real Schur form.

We now need to revise a variant of the Jordan canonical form (2.1) if A is real. For a nonreal eigenvalue λ of A, if $J_i(\lambda)$ appears in the Jordan canonical form with a certain multiplicity, then $J_i(\bar{\lambda})$ must also appear with the same multiplicity. See [18, p.150 ff.]. In general, every block pair of conjugate $n_k \times n_k$ Jordan blocks

(2.7) $$\begin{bmatrix} J_i(\lambda) & O \\ O & J_i(\bar{\lambda}) \end{bmatrix}$$

ELA

is similar to a real $2n_k \times 2n_k$ block matrix of the form

$$(2.8) \qquad C_{n_k}(a,b) = \begin{bmatrix} C(a,b) & I & & & \\ & C(a,b) & I & & \\ & & \ddots & \ddots & \\ & & & C(a,b) & I \\ & & & & C(a,b) \end{bmatrix}$$

where

$$(2.9) \qquad\qquad C(a,b) := \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

for $\lambda, \bar{\lambda} = a \pm ib$, $a, b \in \mathbb{R}$, $b \neq 0$. We call $C_k(a,b)$ a *real Jordan block*. These observations lead to the *real Jordan canonical form*.

THEOREM 2.6. [18, Theorem 3.4.5] *Each matrix $A \in \mathbb{R}^{n \times n}$ is similar (via a real similarity transformation) to a block diagonal real matrix of the form*

$$(2.10) \qquad J_R = \begin{bmatrix} C_{n_1}(a_1, b_1) & & & & & & \\ & \ddots & & & & & \\ & & C_{n_p}(a_p, b_p) & & & & \\ & & & J_{n_{p+1}}(\lambda_{p+1}) & & & \\ & & & & \ddots & & \\ & & & & & J_{n_{p+q}}(\lambda_{p+q}) \end{bmatrix},$$

where $\lambda_k = a_k + ib_k$, $a_k, b_k \in \mathbb{R}$, $k = 1, \ldots, p$, is a nonreal eigenvalue of A and λ_k, $k = p+1, \ldots, p+q$, is a real eigenvalue of A. Each real Jordan block $C_{n_k}(a_k, b_k)$ is of the form (2.8) and corresponds to a pair of conjugate Jordan blocks $J_{n_k}(\lambda_k)$ and $J_{n_k}(\bar{\lambda}_k)$ for a nonreal λ_k in the Jordan canonical form of A in (2.1). The real Jordan blocks $J_{n_k}(\lambda_k)$ are exactly the Jordan blocks in (2.1) with real λ_k. Notice that $2(n_1 + \cdots + n_p) + (n_{p+1} + \cdots + n_{p+q}) = n$. We call J_R a real Jordan matrix of order n, a direct sum of real Jordan blocks.

In [8] it is shown that every real skew-Hamiltonian matrix can also be reduced to a *real skew-Hamiltonian Jordan form* via a symplectic similarity. See also [9].

LEMMA 2.7.   [8, Theorem 1] *For every real skew-Hamiltonian matrix $W \in \mathbb{R}^{2n \times 2n}$, there exists a symplectic matrix $\Psi \in \mathbb{R}^{2n \times 2n}$ such that*

$$(2.11) \qquad\qquad \Psi^{-1} W \Psi = \begin{bmatrix} J_R & \\ & J_R^T \end{bmatrix},$$

where $J_R \in \mathbb{R}^{n \times n}$ is in real Jordan form (2.10) and is unique up to a permutation of real Jordan blocks.

$$\boxed{\textbf{ELA}}$$

To compute a real square root of W we propose a method which exploits the skew-Hamiltonian structure and uses the general real Schur method devised by Higham [13].

2.3. The real Schur method. When computing a real square root of a real matrix $A \in \mathbb{R}^{n \times n}$ it is desirable to work with real arithmetic since substantial computational savings may occur. Higham has extended the work of Björck and Hammarling's [5]. Given the reduction of A to the real Schur form

$$(2.12) \qquad Q^T A Q = R = \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1m} \\ & R_{22} & \dots & R_{2m} \\ & & \ddots & \vdots \\ & & & R_{mm} \end{bmatrix} \in \mathbb{R}^{n \times n},$$

where Q is real orthogonal and each block R_{ii} is either 1×1 or 2×2 with complex conjugate eigenvalues, the real Schur method then computes a square root Z of R and finally obtains a square root of A via the transformation $X = QZQ^T$. For details see [13, p. 412 ff.].

ALGORITHM 2.8. [Real Schur method]

1. compute a real Schur decomposition of A, $A = QRQ^T$;
2. compute a square root Z of R solving the equation $Z^2 = R$ via

$$Z_{ii}^2 = R_{ii}, \qquad 1 \le i \le m,$$

$$Z_{ii} Z_{ij} + Z_{ij} Z_{jj} = R_{ij} - \sum_{k=i+1}^{j-1} Z_{ik} Z_{kj}, \quad j = i+1, \dots, m.$$

[block fast recursion]
3. obtain a square root of A, $X = QZQ^T$.

The square root Z is real, and hence X is real, if and only if each of the blocks Z_{ii} is real. Z_{ii} can be computed efficiently since a 2×2 matrix with complex conjugate eigenvalues has four square roots, each of which is a function of the matrix. Two of the square roots are real and two are pure imaginary. See [13, pp. 414–417].

The cost of the real Schur method, measured in floating point operations (flops) may be broken down as follows. The real Schur factorization costs about $15n^3$ flops [12]. The computation of Z requires $n^3/6$ flops and the formation of $X = QZQ^T$ requires $3n^3/2$ flops [13, p. 418]. Only a fraction of the overall time is spent in computing the square root Z.

ELA

http://math.technion.ac.il/iic/ela

3. Square roots of a skew-Hamiltonian matrix. In this section, we present a detailed classification of the square roots of a skew-Hamiltonian matrix $W \in \mathbb{R}^{2n \times 2n}$ based on its real skew-Hamiltonian Jordan form (2.11). First, we discuss the square roots of J_R.

From Lemma 2.1, we get the following

COROLLARY 3.1. *For a real eigenvalue $\lambda_k \neq 0$ a canonical Jordan block $J_i(\lambda_k)$ has precisely two upper triangular square roots which are functions of J_i. These square roots are real if $\lambda_k > 0$ and they pure imaginary if $\lambda_k < 0$.*

We first examine the square roots of a 2×2 real Jordan block $C(a, b)$ in (2.9). We know that $C(a, b)$ has four square roots, each of which is a function of $C(a, b)$. Two of the square roots are real and two are pure imaginary.

LEMMA 3.2. *The real Jordan block $C_k(a, b)$ in (2.8) has precisely four block upper triangular square roots*

$$
F_k^{(j)} = \begin{bmatrix} F & F_1 & \cdot & \ldots & F_{k-1} \\ & F & F_1 & \ldots & F_{k-2} \\ & & \ddots & \ddots & \vdots \\ & & & F & F_1 \\ & & & & F \end{bmatrix}, \quad j = 1, \ldots, 4,
$$

where F is a square root of $C(a, b)$ and F_i, $i = 1, \ldots, k-1$, are the unique solutions of certain Sylvester equations. The superscript j denotes one of the four square roots of $C(a, b)$. These four square roots $F_k^{(j)}$ are functions of $C_k(a, b)$, two of them are real and two are pure imaginary.

Proof. Since $C_k(a, b)$ has 2 distinct eigenvalues and the Jordan form (2.7) has $p = 2$ blocks, from Theorem 2.2 we know that $C_k(a, b)$ has four square roots which are all functions of A.

Let X be a square root of $C_k(a, b)$ $(k > 1)$. It is not difficult to see that X inherits $C_k(a, b)$ block upper triangular structure,

$$
X = \begin{bmatrix} X_{11} & X_{12} & \ldots & X_{1,k} \\ & X_{22} & \ldots & X_{2,k} \\ & & \ddots & \vdots \\ & & & X_{kk} \end{bmatrix}
$$

where $X_{i,j}$ are all 2×2 matrices. Equating (i, j) blocks in the equation

$$
X^2 = C_k(a, b)
$$

ELA

http://math.technion.ac.il/iic/ela

we obtain

$$(3.1) \qquad X_{ii}^2 = C(a,b), \quad i = 1, \ldots, k,$$

$$(3.2) \qquad X_{ii}X_{i,i+1} + X_{i,i+1}X_{i+1,i+1} = I_2, \quad i = 1, \ldots, k-1$$

$$(3.3) \qquad X_{ii}X_{ij} + X_{ij}X_{jj} = -\sum_{l=i+1}^{j-1} X_{il}X_{lj}, \quad j = i+2, \ldots, k.$$

The whole of X is uniquely determined by its diagonal blocks. If F is one square root of $C(a,b)$, from (3.1) and to conform with the definition of $f(A)$ (the eigenvalues of X_{ii} must be the same), we have

$$(3.4) \qquad X_{ii} = F, \quad i = 1, \ldots, k.$$

Equations (3.2) and (3.3) are Sylvester equations and the condition for them to have a unique solution X_{ij} is that X_{ii} and $-X_{jj}$ have no eigenvalues in common and this is guaranteed.

From (3.2) we obtain the blocks X_{ij} along the first superdiagonal and (3.4) forces them to be all equal, say F_1,

$$X_{12} = X_{23} = \ldots = X_{k-1,k} = F_1.$$

This implies that the other superdiagonals obtained from (3.3) are also constant, say $F_{j-1}, j = 3, \ldots, k$,

$$X_{1j} = X_{2,j+1} = \ldots = X_{k-j+1,k} = F_{j-1}, \quad j = 3, \ldots, k.$$

Thus, since there are only exactly four distinct square roots of $C(a,b)$ which are functions of $C(a,b)$, $F = F^{(l)}$, $l = 1, \ldots, 4$, it follows that $C_k(a,b)$ will also have precisely four square roots which are functions of $C_k(a,b)$. If F is real then F_{j-1}, $j = 2, \ldots, k$, will also be real. If F is pure imaginary it can also be seen that F_{j-1}, $j = 2, \ldots, k$, will be pure imaginary too. $\square$

Next theorem combines Corollary 3.1 and Lemma 3.2 to characterize the square roots of a real Jordan matrix J_R.

THEOREM 3.3. *Assume that a nonsingular real Jordan matrix J_R in (2.10) has p real Jordan blocks corresponding to c distinct complex conjugate eigenvalue pairs and q canonical Jordan blocks corresponding to r distinct real eigenvalues.*

Then, J_R has precisely 2^{2c+r} square roots which are functions of J_R, given by

$$(3.5) \qquad X_j = \text{diag}\left(F_{n_1}^{(j_1)}, \ldots, F_{n_p}^{(j_p)}, L_{n_{p+1}}^{(i_1)}, \ldots, L_{n_{p+q}}^{(i_q)}\right), \quad j = 1, \ldots, 2^{2c+r},$$

ELA

http://math.technion.ac.il/iic/ela

corresponding to all possible choices of $j_1, \ldots, j_p$, $j_k = 1, 2, 3$ or 4, and $i_1, \ldots, i_q$, $i_k = 1$ or 2, subject to the constraint that $j_l = j_k$ and $i_l = i_k$ whenever $\lambda_l = \lambda_k$.

If $c + r < p + q$, then J_R has square roots which are not functions of J_R and they form $2^{2p+q} - 2^{2c+r}$ parameterized families given by

$$(3.6) \qquad X_j(\Omega) = \Omega \operatorname{diag}\left(F_{n_1}^{(j_1)}, \ldots, F_{n_p}^{(j_p)}, L_{n_{p+1}}^{(i_1)}, \ldots, L_{n_{p+q}}^{(i_q)} \right) \Omega^{-1},$$
$$j = 2^{2c+r} + 1, \ldots, 2^{2p+q},$$

where $j_k = 1, 2, 3$ or 4 and $i_k = 1$ or 2, Ω is an arbitrary nonsingular matrix which commutes with J_R and for each j there exist l and k depending on j, such that $\lambda_l = \lambda_k$ while $j_l \neq j_k$ or $i_l \neq i_k$.

Proof. The number of distinct eigenvalues is $s = 2c + r$ and, according to Theorem 2.2, J_R has precisely $2^s = 2^{2c+r}$ square roots which are functions of J_R. All square roots of J_R which are functions of J_R satisfy

$$f(J_R) = \begin{bmatrix} f(C_{n_1}) & & & & & \\ & \ddots & & & & \\ & & f(C_{n_p}) & & & \\ & & & f(J_{n_{p+1}}) & & \\ & & & & \ddots & \\ & & & & & f(J_{n_{p+q}}) \end{bmatrix}$$

and, according to Lemma 2.1 and Lemma 3.2 these are given by (3.5). The remaining square roots of J_R, if they exist, cannot be functions of J_R. Equation (3.6) derives from the second part of Theorem 2.2. □

Theorem 3.3, Lemma 3.2 and Corollary 3.1 give the next result.

COROLLARY 3.4. *Under the assumptions of* Theorem 3.3,

(1) *if J_R has a real negative eigenvalue, then J_R has no real square roots which are functions of J_R;*

(2) *if J_R has no real negative eigenvalues, then J_R has precisely 2^{c+r} real square roots which are functions of J_R, given by (3.5) with the choices of $j_1, \ldots, j_p$ corresponding to real square roots $F_{n_1}^{(j_1)}, \ldots, F_{n_p}^{(j_p)}$;*

(3) *if J_R has no real positive eigenvalues, then J_R has precisely 2^{c+r} pure imaginary square roots which are functions of J_R, given by (3.5) with the choices of $j_1, \ldots, j_p$ corresponding to pure imaginary square roots $F_{n_1}^{(j_1)}, \ldots, F_{n_p}^{j_p}$.*

With all these results we characterize the square roots of a real skew-Hamiltonian

W.

THEOREM 3.5. *Let $W \in \mathbb{R}^{2n \times 2n}$ be a nonsingular skew-Hamiltonian matrix with the real skew-Hamiltonian Jordan form in (2.11). Assume that J_R has p real Jordan blocks corresponding to c distinct complex conjugate eigenvalue pairs and q canonical Jordan blocks corresponding to r distinct real eigenvalues.*

Then W has precisely 2^{2c+r} square roots which are functions of W, given by

$$(3.7) \qquad Y_j = \Psi \operatorname{diag}\left(X_j, X_j^T\right) \Psi^{-1}, \quad j = 1, \ldots, 2^{2c+r},$$

where X_j is a square root of J_R given in (3.5).

W always has square roots which are not functions of W and they form $4^{2p+q} - 2^{2c+r}$ parameterized families given by

$$(3.8) \qquad Y_j(\Theta) = \Psi\Theta \operatorname{diag}\left(\widetilde{X}_j, \widehat{X}_j^T\right) \Theta^{-1}\Psi^{-1}, \quad j = 2^{2c+r} + 1, \ldots, 4^{2p+q},$$

where

$$\widetilde{X}_j = \operatorname{diag}\left(F_{n_1}^{(j_1)}, \ldots, F_{n_p}^{(j_p)}, L_{n_{p+1}}^{(i_1)}, \ldots, L_{n_{p+q}}^{(i_q)}\right),$$
$$\widehat{X}_j = \operatorname{diag}\left(F_{n_1}^{(j_{p+1})}, \ldots, F_{n_p}^{(j_{2p})}, L_{n_{p+1}}^{(i_{q+1})}, \ldots, L_{n_{p+q}}^{(i_{2q})}\right),$$

$j_k = 1, 2, 3$ or 4 and $i_k = 1$ or 2, Θ is an arbitrary nonsingular matrix which commutes with $\operatorname{diag}\left(J_R, J_R^T\right)$ and for each j there exist l and k depending on j, such that $\lambda_l = \lambda_k$ while $j_l \neq j_k$ or $i_l \neq i_k$.

Notice that W has $s = 2c + r$ distinct eigenvalues corresponding to $2(p + q)$ real Jordan blocks and $2(2p + q)$ canonical Jordan blocks. We always have $s \leq 2p + q$ and so $s < 2(2p + q)$. Thus, there are always square roots which are not functions of W.

Proof. This result is a direct consequence of Theorem 2.2 and Theorem 3.3. Notice that if X_j in (3.5) is a square root of J_R, then $\operatorname{diag}\left(X_j, X_j^T\right)$ is a square root of $\operatorname{diag}\left(J_R, J_R^T\right)$. Thus,

$$W = \Psi \begin{bmatrix} J_R & \\ & J_R^T \end{bmatrix} \Psi^{-1} = \Psi \begin{bmatrix} X_j & \\ & X_j^T \end{bmatrix} \begin{bmatrix} X_j & \\ & X_j^T \end{bmatrix} \Psi^{-1}$$
$$= \Psi \begin{bmatrix} X_j & \\ & X_j^T \end{bmatrix} \Psi^{-1} \Psi \begin{bmatrix} X_j & \\ & X_j^T \end{bmatrix} \Psi^{-1}.$$

Thus, $Y_j = \Psi \begin{bmatrix} X_j & \\ & X_j^T \end{bmatrix} \Psi^{-1}$ is a square root of W. Since X_j is a function of J_R, Y_j is a function of $\operatorname{diag}\left(J_R, J_R^T\right)$. This proves the first part of the theorem.

The second part follows from the second part of Theorem 2.2 and the fact that $\operatorname{diag}\left(\widetilde{X}_j, \widehat{X}_j^T\right)$ is a square root of $\operatorname{diag}\left(J_R, J_R^T\right)$. $\square$

$$\boxed{\textbf{ELA}}$$

It is easy to verify that if W is a skew-Hamiltonian matrix, then W^2 is also skew-Hamiltonian. This implies that any function of W, which is a polynomial by definition, is a skew-Hamiltonian matrix. Thus, all the square roots of W which are functions of W are skew-Hamiltonian matrices. The following result refers to the existence of real square roots of W.

COROLLARY 3.6. *Under the assumptions of* Theorem 3.5, *the following statements hold:*

(1) *if W has a real negative eigenvalue, then W has no real skew-Hamiltonian square roots;*

(2) *if W has no real negative eigenvalues, then W has precisely 2^{c+r} real skew-Hamiltonian square roots which are functions of W, given by* (3.7) *with the choices of $j_1, \ldots, j_p$ corresponding to real square roots $F_{n_1}^{(j_1)}, \ldots, F_{n_p}^{(j_p)}$.*

Observe that W may have real negative eigenvalues and yet still have a real square root; however, the square root will not be a function of W.

In [8, Theorem 2] it is shown that

LEMMA 3.7. *Every real skew-Hamiltonian matrix W has a real Hamiltonian square root.*

The proof is constructive and the key step is based in Lemma 2.7 - we can bring W into a real skew-Hamiltonian Jordan form (2.11) via a symplectic similarity. Further, it is shown that every skew-Hamiltonian matrix W has infinitely many real Hamiltonian square roots.

The following theorem gives the structure of those real Hamiltonian square roots.

THEOREM 3.8. *Let $W \in \mathbb{R}^{2n \times 2n}$ be a nonsingular skew-Hamiltonian matrix and assume the conditions in Theorem 3.5 hold.*

(1) *If W has no real negative eigenvalues, then W has real Hamiltonian square roots which are not functions of W and they form 2^{p+q} parameterized families given by*

$$(3.9) \qquad Y_j(\Theta) = \Psi\Theta \operatorname{diag}\left(X_j, -X_j{}^T\right)\Theta^{-1}\Psi^{-1}, \quad j = 1, \ldots, 2^{p+q},$$

where X_j denotes a real square root of J_R and Θ is an arbitrary nonsingular matrix which preserves the Hamiltonian structure and commutes with $\operatorname{diag}\left(J_R, J_R^T\right)$.

(2) *If W has some real negative eigenvalues, then W has real Hamiltonian square roots which are not functions of W and they form 2^{p+q} parameterized families*

given by

$$(3.10) \qquad Y_j(\Theta) = \Psi\Theta \begin{bmatrix} \widehat{X}_j & K_j \\ \widehat{K}_j & -\widehat{X}_j^T \end{bmatrix} \Theta^{-1}\Psi^{-1}, \quad j = 1, \dots, 2^{p+q},$$

where $\widehat{X}_j$ is a square root for the Jordan blocks of J_R which are not associated with real negative eigenvalues, K_j and $\widehat{K}_j$ are symmetric block diagonal matrices corresponding to the square roots of the real negative eigenvalues, and Θ is an arbitrary nonsingular matrix which preserves the Hamiltonian structure and commutes with $\mathrm{diag}\left(J_R, J_R^T\right)$.

Notice that since Hamiltonian structure is preserved under symplectic similarity transformations, Θ may be any symplectic matrix which commutes with $\mathrm{diag}\left(J_R, J_R^T\right)$. There are however other non symplectic matrices which may be used in equations (3.9) and (3.10) [10].

Proof. Equation (3.9) is a special case of Equation (3.8) in Theorem 3.5. If X_j is a real square root of J_R then $\mathrm{diag}\left(X_j, -X_j^T\right)$ is a Hamiltonian square root of $\mathrm{diag}\left(J_R, J_R^T\right)$ and Hamiltonian structure is preserved. There are 2^{p+q} real square roots of J_R which may be or not functions of J_R.

For the second part, assume that J_R in (2.10) has only one real negative eigenvalue, say $\lambda_k < 0$, $k > p$ corresponding to the real Jordan block J_{n_k}.

Let $\pm iM_{n_k}$ with $M_{n_k} \in \mathbb{R}^{n \times n}$ be the two pure imaginary square roots of J_{n_k}, which are upper triangular Toeplitz matrices. See Corollary 3.1 and (2.2). Observe that $(\pm iM_{n_k})^2 = -M_{n_k}^2 = J_{n_k}$. We will first construct a square root of $\mathrm{diag}\left(J_{n_k}, J_{n_k}^T\right)$ which is real and Hamiltonian. Let P_{n_k} be the reversal matrix of order n_k which satisfies $P_{n_k}^2 = I$ (the anti-diagonal entries are all 1's, the only nonzero entries). The matrices $P_{n_k}M_{n_k}$ and $M_{n_k}P_{n_k}$ are real symmetric and we have

$$\begin{bmatrix} & M_{n_k}P_{n_k} \\ -P_{n_k}M_{n_k} & \end{bmatrix}^2 = \begin{bmatrix} -M_{n_k}^2 & \\ & -P_{n_k}M_{n_k}^2 P_{n_k} \end{bmatrix} = \begin{bmatrix} J_{n_k} & \\ & J_{n_k}^T \end{bmatrix}.$$

Thus,

$$\begin{bmatrix} & M_{n_k}P_{n_k} \\ -P_{n_k}M_{n_k} & \end{bmatrix} \qquad \left(\text{and also } \begin{bmatrix} & -M_{n_k}P_{n_k} \\ P_{n_k}M_{n_k} & \end{bmatrix} \right)$$

is a real Hamiltonian square root of $\mathrm{diag}\left(J_{n_k}, J_{n_k}^T\right)$ which is not a function of $\mathrm{diag}\left(J_{n_k}, J_{n_k}^T\right)$.

If $X_1 = \mathrm{diag}(F_{n_1}, \dots, F_{n_p})$ is a real square root of $\mathrm{diag}(C_{n_1}, \dots, C_{n_p})$, $X_2 = \mathrm{diag}(L_{n_{p+1}}, \dots, L_{n_{k-1}})$ is a real square root of $\mathrm{diag}(J_{n_{p+1}}, \dots, J_{n_{k-1}})$ and

ELA

http://math.technion.ac.il/iic/ela

858   Z. Liu, Y. L. Zhang, C. Ferreira, and R. Ralha

$X_3 = \text{diag}(L_{n_{k+1}}, \ldots, L_{n_{p+q}})$ is a real square root of $\text{diag}(J_{n_{k+1}}, \ldots, J_{n_{p+q}})$, then

$$
\begin{bmatrix}
X_1 & & & & & & \\
 & X_2 & & & & & \\
 & & O & & & & M_{n_k}P_{n_k} \\
 & & & X_3 & & & \\
 & & & & -X_1^T & & \\
 & & & & & -X_2^T & \\
 & -P_{n_k}M_{n_k} & & & & & O \\
 & & & & & & -X_3^T
\end{bmatrix}
=:
\begin{bmatrix}
\widehat{X}_j & K_j \\
\widehat{K}_j & -\widehat{X}_j^T
\end{bmatrix}
$$

is a real Hamiltonian square root of $\text{diag}\left(J_R, J_R^T\right)$. Notice that there are 2^{p+q} different square roots with this form. Thus,

$$
Y_j(\Theta) = \Psi\Theta
\begin{bmatrix}
\widehat{X}_j & K_j \\
\widehat{K}_j & -\widehat{X}_j^T
\end{bmatrix}
\Theta^{-1}\Psi^{-1}, \quad j = 1, \ldots, 2^{p+q},
$$

is a Hamiltonian square root of W.

If W has more than one real negative eigenvalue, the generalization is straightforward. ☐

4. Algorithms for computing square roots of a skew-Hamiltonian matrix. In this section, we will present a structure-exploiting Schur method to compute a real skew-Hamiltonian or a real Hamiltonian square root of a real skew-Hamiltonian matrix $W \in \mathbb{R}^{2n \times 2n}$ when W does not have real negative eigenvalues.

4.1. Skew-Hamiltonian square roots. First we obtain the PVL decomposition of $W \in \mathbb{R}^{n \times n}$ described in section 2.2,

$$
U^T W U =
\begin{bmatrix}
W_1 & W_2 \\
O & W_1^T
\end{bmatrix}, \quad W_2^T = -W_2,
$$

where U is symplectic-orthogonal and W_1 is upper Hessenberg. The matrix U is constructed as a product of elementary symplectic-orthogonal matrices. These are the $2n \times 2n$ Givens rotations matrices of the type

$$
\begin{bmatrix}
I_{j-1} & & & & \\
 & \cos\theta & & \sin\theta & \\
 & & I_{n-1} & & \\
 & -\sin\theta & & \cos\theta & \\
 & & & & I_{n-j}
\end{bmatrix}, \quad 1 \le j \le n,
$$

$$\boxed{\textbf{ELA}}$$

for some angle $\theta \in [-\pi/2, \pi/2]$, and the direct sum of two identical $n \times n$ Householder matrices

$$H_j \oplus H_j(\boldsymbol{v}, \beta) = \begin{bmatrix} I_n - \beta \boldsymbol{v}\boldsymbol{v}^T & \\ & I_n - \beta \boldsymbol{v}\boldsymbol{v}^T \end{bmatrix},$$

where $\boldsymbol{v}$ is a vector of length n with its first $j - 1$ elements equal to zero. A simple combination of these transformations can be used to zero out entries in W to accomplish the PVL form. See Algorithm 1 and Algorithm 5 in [2, pp. 4,10]. The product of the transformations used in the reductions is accumulated to form the matrix U.

Then the standard QR algorithm is applied to W_1 producing an orthogonal matrix Q and a quasi-upper triangular matrix N_1 in real Schur form (2.12) so that

$$W_1 = QN_1Q^T,$$

and we attain the real skew-Hamiltonian Schur decomposition of W,

$$\mathcal{T} = \mathcal{U}^T W \mathcal{U} = \begin{bmatrix} N_1 & N_2 \\ O & N_1^T \end{bmatrix}, \quad N_2 = -N_2^T,$$

where $\mathcal{U} = U \begin{bmatrix} Q & O \\ O & Q \end{bmatrix}$ and $N_2 = Q^T W_2 Q$. This procedure takes only approximately 20% of the computational cost the standard QR algorithm would require to compute the unstructured real Schur decomposition of W [2, p. 10].

Let

$$Z = \begin{bmatrix} X & Y \\ & X^T \end{bmatrix}, \qquad Y = -Y^T.$$

be a skew-Hamiltonian square root of $\mathcal{T}$. We can solve the equation $Z^2 = \mathcal{T}$ exploiting the structure. From

$$\begin{bmatrix} X & Y \\ & X^T \end{bmatrix} \cdot \begin{bmatrix} X & Y \\ & X^T \end{bmatrix} = \begin{bmatrix} X^2 & XY + YX^T \\ 0 & \left(X^T\right)^2 \end{bmatrix} = \begin{bmatrix} N_1 & N_2 \\ 0 & N_1^T \end{bmatrix},$$

we have

(4.1)
$$X^2 = N_1$$

and

(4.2)
$$XY + YX^T = N_2.$$

Equation (4.1) can be solved using Higham's real Schur method (see Algorithm 2.8) and it is not difficult to show that X inherits N_1's quasi-upper triangular structure.

$$\boxed{\textbf{ELA}}$$

Equation (4.2) is a Lyapunov equation which can be solved efficiently since X is already in quasi-upper triangular real Schur form and Y is skew-symmetric. The techniques are the same as for the Sylvester equation. See [1, 15].

If the partitions of $X = (X_{ij})$, $Y = (Y_{ij})$ and $N_2 = (N_{ij})$ are conformal with N_1 block structure,

$$
X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1m} \\ & X_{22} & \cdots & X_{2m} \\ & & \ddots & \vdots \\ & & & X_{mm} \end{bmatrix}, \qquad
Y = \begin{bmatrix} Y_{11} & -Y_{21}^T & \cdots & -Y_{m1}^T \\ Y_{21} & Y_{22} & \cdots & -Y_{m2}^T \\ \vdots & \vdots & \ddots & \vdots \\ Y_{m1} & Y_{m2} & \cdots & Y_{mm} \end{bmatrix},
$$

$$
N_2 = \begin{bmatrix} N_{11} & -N_{21}^T & \cdots & N_{m1}^T \\ N_{21} & N_{22} & \cdots & -N_{m2}^T \\ \vdots & \vdots & \ddots & \vdots \\ N_{m1} & N_{m2} & \cdots & N_{mm} \end{bmatrix}, \qquad
\begin{aligned} Y_{ii} = -Y_{ii}^T, & \quad N_{ii} = -N_{ii}^T, \\ & i = i, \ldots, m, \end{aligned}
$$

then, from (4.2), we have

$$
\sum_{k=i}^{m} X_{ik} Y_{kj} + \sum_{k=j}^{m} Y_{ik} X_{jk}^T = N_{ij}
$$

and

$$
X_{ii} Y_{ij} + Y_{ij} X_{jj}^T = N_{ij} - \sum_{k=i+1}^{m} X_{ik} Y_{kj} - \sum_{k=j+1}^{m} Y_{ik} X_{jk}^T.
$$

These equations may be solved successively for $Y_{mm}, Y_{m,m-1}, \ldots, Y_{m1}$, $Y_{m-1,m-1}, Y_{m-1,m-2}, \ldots, Y_{m-1,1}, \ldots, Y_{22}, Y_{21}$ and Y_{11}. We have to solve

$$
X_{ii} Y_{ij} + Y_{ij} X_{jj}^T = N_{ij} - \sum_{k=i+1}^{m} X_{ik} Y_{kj} - \sum_{k=j+1}^{i} Y_{ik} X_{jk}^T + \sum_{k=i+1}^{m} Y_{ki}^T X_{jk}^T,
$$

(4.3) $i = m, m-1, \ldots, 1$
             $j = i, i-1, \ldots, 1.$

Since X_{ii} are of order 1 or 2, each system (4.3) is a linear system of order 1,2 or 4 and is usually solved by Gaussian elimination with complete pivoting. The solution is unique because X_{ii} and $-X_{jj}^T$ have no eigenvalues in common. See Section 2.3.

ALGORITHM 4.1. [Skew-Hamiltonian real Schur method]

1. compute a real skew-Hamiltonian Schur decomposition of W,

$$
\mathcal{T} = \mathcal{U}^T W \mathcal{U} = \begin{bmatrix} N_1 & N_2 \\ 0 & N_1^T \end{bmatrix};
$$

ELA

http://math.technion.ac.il/iic/ela

2. use Algorithm 2.8 to compute a square root X of N_1, $X^2 = N_1$;

3. solve the Sylvester equation $XY + YX^T = N_2$ using (4.3) and form

$$Z = \begin{bmatrix} X & Y \\ & X^T \end{bmatrix};$$

4. obtain the skew-Hamiltonian square root of W, $\mathcal{X} = \mathcal{U}Z\mathcal{U}^T$.

The cost of the real skew-Hamiltonian Schur method for $W \in \mathbb{R}^{2n \times 2n}$ is measured in flops as follows. The real skew-Hamiltonian Schur factorization of W costs about $3(2n)^3$ flops [12, 2]. The computation of X requires $n^3/6$ flops, the computation of the skew-symmetric solution Y requires about n^3 flops [12, p. 368] and the formation of $\mathcal{X} = \mathcal{U}Z\mathcal{U}^T$ requires $3(2n)^3/2$ flops. The total cost is approximately $5(2n)^3$ flops. Comparing with the overall cost of Algorithm 2.8, the unstructured real Schur method, which is about $17 \times (2n)^3$ flops, Algorithm 4.1 requires considerably fewer floating point operations.

4.2. Hamiltonian square roots. Analogously, let Z be a Hamiltonian square root of $\mathcal{T}$,

$$Z = \begin{bmatrix} X & Y \\ & -X^T \end{bmatrix}, \qquad Y = Y^T$$

(which is not a function of $\mathcal{T}$). To solve the equation $Z^2 = \mathcal{T}$, observe that, from

$$\begin{bmatrix} X & Y \\ & -X^T \end{bmatrix} \cdot \begin{bmatrix} X & Y \\ & -X^T \end{bmatrix} = \begin{bmatrix} X^2 & XY - YX^T \\ 0 & \left(X^T\right)^2 \end{bmatrix} = \begin{bmatrix} N_1 & N_2 \\ 0 & N_1^T \end{bmatrix}$$

it follows

(4.4) $$X^2 = N_1$$

and

(4.5) $$XY - YX^T = N_2.$$

Equation (4.4) can be solved using Higham's real Schur method and Equation (4.5) is a singular Sylvester equation with infinitely many symmetric solutions. See [8, Proposition 7]. Again, the structure can be exploited and we have to solve

$$X_{ii}Y_{ij} - Y_{ij}X_{jj}^T = N_{ij} - \sum_{k=i+1}^{m} X_{ik}Y_{kj} + \sum_{k=j+1}^{i} Y_{ik}X_{jk}^T - \sum_{k=i+1}^{m} Y_{ki}^T X_{jk}^T,$$

(4.6) $$i = m, m-1, \ldots, 1$$
$$j = i, i-1, \ldots, 1.$$

The solution of the linear system (4.6) may not be unique but it always exists.

ELA

http://math.technion.ac.il/iic/ela

ALGORITHM 4.2. [Hamiltonian real Schur method]

1. compute a real skew-Hamiltonian Schur decomposition of W,

$$\mathcal{T} = \mathcal{U}^T W \mathcal{U} = \begin{bmatrix} N_1 & N_2 \\ 0 & N_1^T \end{bmatrix};$$

2. use Algorithm 2.8 to compute a square root X of N_1, $X^2 = N_1$;

3. obtain one solution for the Sylvester equation $XY - YX^T = N_2$ using (4.6) and form

$$Z = \begin{bmatrix} X & Y \\ & -X^T \end{bmatrix};$$

4. obtain the Hamiltonian square root of W, $\mathcal{X} = \mathcal{U} Z \mathcal{U}^T$.

5. Numerical examples. We implemented Algorithms 4.1 and 4.2 in MATLAB 7.5.0342 (R2007b) and used the Matrix Function Toolbox by Nick Higham available in Matlab Central website `http://www.mathworks.com/matlabcentral`. To find the square root X in step 2 we used the function `sqrtm_real` of this toolbox and to solve the linear systems (4.3) in step 3 of Algorithm 4.1 we used the function `sylvsol` (the solution is always unique). In step 3 of Algorithm 4.2 the linear systems (4.6) are solved using Matlab's function `pinv` which produces the solution with the smallest norm when the system has infinitely many solutions.

Let $\bar{\mathcal{X}}$ be an approximation to a square root of W and define the *residual*

$$E = \bar{\mathcal{X}}^2 - W.$$

Then, we have $\bar{\mathcal{X}}^2 = W + E$ and, as observed by Higham [14, p. 418], the stability of an algorithm for computing a square root $\bar{\mathcal{X}}$ of W corresponds to the residual E being small relative to W. Furthermore, for $\bar{\mathcal{X}}$ computed with `sqrtm_real`, Higham gives the following error bound

$$\frac{\|E\|_F}{\|W\|_F} \le \left(1 + cn \frac{\|\bar{\mathcal{X}}\|_F^2}{\|W\|_F} \right) u,$$

where $\| \cdot \|_F$ is the Frobenius norm, c is a constant of order 1, n is the dimension of W and u is the roundoff unit. Therefore, the real Schur method is stable if

$$\alpha(\mathcal{X}) = \frac{\|\bar{\mathcal{X}}\|_F^2}{\|W\|_F}$$

is small.

$$\boxed{\textbf{ELA}}$$

We expect our structure-preserving algorithms, Algorithm 4.1 (skew-Hamiltonian square root) and Algorithm 4.2 (Hamiltonian square root) to be as accurate as Algorithm 2.8 (real Schur method) which ignores the structure. The numerical examples that follow illustrate that the three algorithms are all quite accurate when $\alpha(\mathcal{X})$ is small.

EXAMPLE 5.1. The skew-Hamiltonian matrix

$$
W = \begin{bmatrix} \boldsymbol{ee}^T & A \\ -A^T & \boldsymbol{ee}^T \end{bmatrix}, \qquad
A = \begin{bmatrix}
0 & 10^{-6} & 1 & 0 & 0 \\
-10^{-6} & 0 & 1 & 10^{-6} & 0 \\
-1 & -1 & 0 & 10^{-6} & 1 \\
0 & -10^{-6} & -10^{-6} & 0 & 1 \\
0 & 0 & -1 & -1 & 0
\end{bmatrix},
$$

where $\boldsymbol{e}$ is the vector of all ones, has one complex conjugate eigenvalue pair and 3 positive real eigenvalues (all with multiplicity 2). The relative residuals of both the skew-Hamiltonian and Hamiltonian square roots computed with Algorithm 4.1 and Algorithm 4.2 are 4×10^{-15}, the same as for the square root delivered by Algorithm 2.8.

EXAMPLE 5.2. The eigenvalues of the skew-Hamiltonian matrix

$$
W = \begin{bmatrix} A & B \\ B & A^T \end{bmatrix}, \quad
A = \begin{bmatrix}
0 & -10^{-6} & 0 & 0 \\
10^{-6} & 0 & 0 & 0 \\
0 & 0 & 0 & 10^{-6} \\
0 & 0 & -10^{-6} & 0
\end{bmatrix}, \
B = \begin{bmatrix}
0 & 1 & 2 & 3 \\
-1 & 0 & 2 & 3 \\
-2 & -2 & 0 & 3 \\
-3 & -3 & -3 & 0
\end{bmatrix},
$$

are all very close to pure imaginary (four distinct eigenvalues). The relative residuals of the square roots delivered by all the three methods are 4×10^{-16}.

If W has negative real eigenvalues there are no real square roots which are functions of W. However, all these algorithms can be applied and complex square roots will be obtained.

EXAMPLE 5.3. For random matrices A, B and C (values drawn from a uniform distribution on the unit interval), the computed square roots of the skew-hamiltonian matrix of order $2n = 50$ (several cases)

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

also have relative residuals of order at most 10^{-14}.

6. Conclusions. Based on the real skew-Hamiltonian Jordan form, we gave a clear characterization of the square roots of a real skew-Hamiltonian matrix W.

ELA

http://math.technion.ac.il/iic/ela

This includes those square roots which are functions of W and those which are not. Although the Jordan canonical form is the main theoretical tool in our analysis, it is not suitable for numerical computation. We have designed a method for the computation of square roots of such structured matrices. An important component of our method is the real Schur decomposition tailored for skew-Hamiltonian matrices, which has been used by others in solving problems different from ours.

Our algorithm requires considerably less floating point operations (about 70% less) than the general real Schur method due to Higham. Furthermore, in numerical experiments, our algorithm has produced results which are as accurate as those obtained with sqrtm_real.

REFERENCES

[1] R.H. Bartels and G.W. Stewart. Solution of the Matrix Equation AX+XB=C. *Communications of the ACM*, 15:820–826, 1972.

[2] P. Benner and D. Kressner. Fortran 77 Subroutines for Computing the Eigenvalues of Hamiltonian Matrices II. *ACM Transactions on Mathematical Software*, V:1–22, 2005.

[3] P. Benner, D. Kressner, and V. Mehrmann. Skew-Hamiltonian and Hamiltonian eigenvalue problems: theory, algorithms and applications. *Proceedings of the Conference on Applied Mathematics and Scientific Computing,* Brijuni (Croatia), June 23–27, 2003, Springer, Dordrecht, 3–39, 2005.

[4] A. Bunse-Gerstner, R. Byers, and V. Mehrmann. A chart of numerical methods for structured eigenvalue problems. *SIAM Journal on Matrix Analysis and Applications*, 13:419–453, 1992.

[5] A. Björck and S. Hammarling. A Schur method for the square root of a matrix. *Linear Algebra and its Applications*, 52/53:127–140, 1983.

[6] G.W. Cross and P. Lancaster. Square roots of complex matrices. *Linear and Multilinear Algebra*, 1:289–293, 1973/74.

[7] P.I. Davis and N.J. Higham. A Schur-Parlett algorithm for computing matrix functions. *SIAM Journal on Matrix Analysis and Applications*, 25:464–485, 2003.

[8] H. Faßbender, D.S. Mackey, N. Mackey, and H. Xu. Hamiltonian square roots of skew-Hamiltonian matrices. *Linear Algebra and its Applications*, 287:125–159, 1999.

[9] H. Faßbender and K. Ikramov. Several observations on symplectic, Hamiltonian, and skew Hamiltonian matrices. *Linear Algebra and its Applications*, 400:15–29,2005.

[10] G. Freiling, V. Mehrmann, and H. Xu. Existence, uniqueness, and parametrization of Lagrangian invariant subspaces. *SIAM Journal on Matrix Analysis and Applications*, 23:1045–1069, 2002.

[11] F.R. Gantmacher. *The Theory of Matrices*, Vol. I. Chelsea Publishing Co., New York, 1977.

[12] G.H. Golub and C.F. Van Loan. *Matrix Computations*, third edition. Johns Hopkins University Press, Baltimore and London, 1996.

[13] N.J. Higham. Computing real square roots of a real matrix. *Linear Algebra and its Applications*, 88/89:405–430, 1987.

[14] N.J. Higham. Stable iterations for the matrix square root. *Numerical Algorithms*, 15:227–242, 1997.

[15] N.J. Higham. *Accuracy and Stability of Numerical Algorithms*, second edition. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002.

ELA

http://math.technion.ac.il/iic/ela

[16] N.J. Higham. *Functions of Matrices, Theory and Computation.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[17] N.J. Higham, D.S. Mackey, N. Mackey, and F. Tisseur. Functions preserving matrix groups and iterations for the matrix square root. *SIAM Journal on Matrix Analysis and Applications*, 26:849–877, 2005.

[18] R.A. Horn and C.R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge, 1985.

[19] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis.* Cambridge University Press, Cambridge, 1991.

[20] P. Lancaster. *Theory of Matices.* Academic Press, New York, 1969.

[21] L. Lin and Z.Y. Liu. On the square root of an H-matrix with positive diagonal elements. *Annals of Operations Research*, 103:339–350, 2001.

[22] Z. Liu, H. Chen, and H. Cao. The computation of the principal square roots of centrosymmetric H-matrices. *Applied Mathmatics and Computation*, 175:319–329, 2006.

[23] Z. Liu, Y. Zhang and R. Ralha. Computing the square roots of matrices with central symmetry. *Applied Mathematics and Computation*, 186:715–726, 2007.

[24] D.S. Mackey, N. Mackey, and F. Tisseur. Structured tools for structured matrices. *Electronic Journal of Linear Algebra*, 10:106–145, 2003.

[25] C.C. Paige and C.F. Van Loan. A Schur decomposition for Hamiltonian matrices. *Linear Algebra and its Applications*, 41:11–32, 1981.

[26] M.I. Smith. A Schur algorithm for computing matrix pth roots. *SIAM Journal on Matrix Analysis and Applications*, 24:971–989, 2003.

[27] C. Van Loan. A symplectic method for approximating all the eigenvalues of a Hamiltonian matrix. *Linear Algebra and its Applications*, 61:233–251, 1984.