Real and Complex Hamiltonian Square Roots of Skew-Hamiltonian Matrices

Heike Faßbender* D. Steven Mackey† Niloufer Mackey‡
Hongguo Xu§

January 1999

Dedicated to Professor Ludwig Elsner
on the occasion of his 60th birthday.

Abstract

We present a constructive existence proof that every real skew-Hamiltonian matrix W has a real Hamiltonian square root. The key step in this construction shows how one may bring any such W into a real quasi-Jordan canonical form via symplectic similarity. We show further that every W has infinitely many real Hamiltonian square roots, and give a lower bound on the dimension of the set of all such square roots. Extensions to complex matrices are also presented.

This report is an updated version of the paper "Hamiltonian square roots of skew-Hamiltonian matrices" that appeared in Linear Algebra & its Applications*, v. 287, 1999, pp. 125 – 159.*

AMS subject classification. 15A21, 15A24, 15A57, 65F30, 93B10

1 Introduction

Any matrix X such that $X^2 = A$ is said to be a square root of the matrix A. For general complex matrices $A \in \mathbb{C}^{n\times n}$ there exists a well-developed although somewhat complicated theory of matrix square roots [7, 10, 17], and a number of algorithms for their effective computation [3, 14]. Similarly for the theory and computation of real square roots for real matrices [13, 17]. By contrast, *structured square root* problems, where both the matrix A

*Fachbereich 3 - Mathematik und Informatik, Zentrum für Technomathematik, Universität Bremen, D-28334 Bremen, FRG, e-mail : heike@math.uni-bremen.de

†Department of Mathematics & Computer Science, Kalamazoo College, 1200 Academy Street, Kalamazoo MI 49006, USA.

‡Department of Mathematics & Statistics, Western Michigan University, Kalamazoo, MI 49008, USA. email: nil.mackey@wmich.edu. This work was supported in part by NSF Grant CCR–9619514.

§Fakultät für Mathematik, TU Chemnitz-Zwickau, D-09107 Chemnitz, FRG. email: hongguo.xu@mathematik.tu-chemnitz.de. This work was supported in part by the Chinese National Natural Science Foundation, and by Deutsche Forschungsgemeinschaft Research Grant Me 790/7-2.

and its square root X are required to have some extra (not necessarily the same) specified structure, have been comparatively less studied. Some notable exceptions include positive (semi)definite square roots of positive (semi)definite matrices [13, 14, 16], M-matrix square roots of M-matrices [1, 14], coninvolutory square roots of coninvolutory matrices [17], and skew-symmetric square roots of symmetric matrices [18]. In this paper we investigate another such structured square root problem, that of finding real Hamiltonian square roots of real skew-Hamiltonian matrices.

A real $2n \times 2n$ matrix H of the form

$$H = \begin{bmatrix} E & F \\ G & -E^T \end{bmatrix}$$

is said to be *Hamiltonian* if $E, F, G \in \mathbb{R}^{n\times n}$, with $F^T = F$ and $G^T = G$. Equivalently, one may characterize the set $\mathcal{H}$ of all $2n \times 2n$ Hamiltonian matrices by

$$\mathcal{H} = \{H \in \mathbb{R}^{2n\times 2n} \mid (JH)^T = JH\},$$

where $J = \left[\begin{smallmatrix} 0 & I \\ -I & 0 \end{smallmatrix}\right]$ and I is the $n \times n$ identity matrix. Complementary to $\mathcal{H}$ is the set

$$\mathcal{W} = \{W \in \mathbb{R}^{2n\times 2n} \mid (JW)^T = -JW\}$$

of all *skew-Hamiltonian* matrices. Matrices in $\mathcal{W}$ are exactly those with block structure

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

where $A, B, C \in \mathbb{R}^{n\times n}$, with $B^T = -B$ and $C^T = -C$. Another useful way to look at Hamiltonian and skew-Hamiltonian matrices is from the point of view of bilinear forms. Associated with any nondegenerate bilinear form $b(x, y)$ on $\mathbb{R}^k$ one has the following sets of matrices:

$$\begin{aligned}
\mathcal{A}(b) &= \{S \in \mathbb{R}^{k\times k} \mid b(Sx, Sy) = b(x, y) \quad \forall x, y \in \mathbb{R}^k\}, \\
\mathcal{L}(b) &= \{H \in \mathbb{R}^{k\times k} \mid b(Hx, y) = -b(x, Hy) \quad \forall x, y \in \mathbb{R}^k\}, \\
\mathcal{J}(b) &= \{W \in \mathbb{R}^{k\times k} \mid b(Wx, y) = b(x, Wy) \quad \forall x, y \in \mathbb{R}^k\}.
\end{aligned}$$

These are, respectively, the automorphism group, Lie algebra, and Jordan algebra of the form b. It is now easy to see that $\mathcal{H}$ is just the Lie algebra $\mathcal{L}(b)$ and $\mathcal{W}$ the Jordan algebra $\mathcal{J}(b)$ of the bilinear form $b(x, y) = x^T Jy$ defined on $\mathbb{R}^{2n}$ by the matrix $J = \left[\begin{smallmatrix} 0 & I \\ -I & 0 \end{smallmatrix}\right]$.

The eigenproblem for Hamiltonian matrices arises in a number of important applications, and many algorithms for computing their eigenvalues and invariant subspaces have been described in the literature (see [5, 6, 19] for references). In [26], Van Loan proposed a method for calculating the eigenvalues of Hamiltonian matrices by first squaring them. Thus he was led to consider the set

$$\mathcal{H}^2 = \{N \in \mathbb{R}^{2n\times 2n} \mid N = H^2, H \in \mathcal{H}\}$$

of all *squared-Hamiltonian* matrices. The calculation

$$b(H^2x, y) = -b(Hx, Hy) = b(x, H^2y) \quad \forall x, y \in \mathbb{R}^{2n}$$

shows immediately that $\mathcal{H}^2 \subseteq \mathcal{W}$. (Indeed, the same argument shows that $\mathcal{L}^2(b) \subseteq \mathcal{J}(b)$ for any bilinear form b.) Almost all the algorithms proposed by Van Loan in [26] depend only on the skew-Hamiltonian block structure of matrices in $\mathcal{H}^2$, and hence apply equally well to every matrix in $\mathcal{W}$. It is then natural to wonder whether the sets $\mathcal{H}^2$ and $\mathcal{W}$ might actually be the same.

In this paper we show that indeed $\mathcal{H}^2 = \mathcal{W}$, or in other words, every real skew-Hamiltonian matrix has a real Hamiltonian square root. The proof occupies the next three sections: after outlining the strategy of the proof in Section 2, we focus in Sections 3 and 4 on the main technical result of this paper, a symplectic canonical form for real skew-Hamiltonian matrices. Then in Section 5 we consider the square root sets themselves: for a general $W \in \mathcal{W}$, what can be said about the size and topological nature of the set of all the real Hamiltonian square roots of W? We close in Section 6 with results on related structured square root problems involving complex Hamiltonian and skew-Hamiltonian matrices.

2 The Generic Case

We begin by giving a short proof that almost all real skew-Hamiltonian matrices (i.e., all matrices in an open dense subset of $\mathcal{W}$) have a real Hamiltonian square root. This preliminary result serves to make the general case more plausible, and at the same time allows us to introduce the basic elements and strategy of the general proof in a setting where there are no technical details to obscure the main line of the argument.

An important way to exploit the structure of Hamiltonian and skew-Hamiltonian matrices is to use only structure-preserving similarities. To that end consider the set $\mathcal{S}$ of real *symplectic* matrices defined by

$$\mathcal{S} := \{S \in \mathbb{R}^{2n\times 2n} \mid S^T J S = J\}.$$

Equivalently, $\mathcal{S}$ is the automorphism group of the bilinear form defined by J. It is well-known and easy to show from either definition that $\mathcal{S}$ forms a multiplicative group, and that symplectic similarities preserve Hamiltonian, squared-Hamiltonian and skew-Hamiltonian structure: for any $S \in \mathcal{S}$,

$$\begin{aligned} H \in \mathcal{H} &\implies S^{-1}HS \in \mathcal{H}, \\ N \in \mathcal{H}^2 &\implies S^{-1}NS \in \mathcal{H}^2, \\ W \in \mathcal{W} &\implies S^{-1}WS \in \mathcal{W}. \end{aligned}$$

The first simplifying reduction we use was introduced by Van Loan in [26]. There he showed that any skew-Hamiltonian W can be brought to block-upper-triangular form by an *orthogonal*-symplectic similarity. That is, for any $W \in \mathcal{W}$ one can explicitly compute an orthogonal-symplectic Q such that

$$Q^T W Q = \begin{bmatrix} U & R \\ 0 & U^T \end{bmatrix}, \quad \text{where} \quad U, R \in \mathbb{R}^{n\times n}. \tag{1}$$

Van Loan actually shows that one can attain an upper Hessenberg U with an orthogonal-symplectic similarity; however, this extra structure will play no role in this paper.

Now suppose we could somehow continue this reduction by (not necessarily orthogonal) symplectic similarities all the way to block-diagonal form. Then the following proposition shows that we would be done.

Proposition 1 *Suppose for $W \in \mathcal{W}$ there exists some $S \in \mathcal{S}$ such that*

$$S^{-1}WS = \begin{bmatrix} A & 0 \\ 0 & A^T \end{bmatrix}$$

with $A \in \mathbb{R}^{n\times n}$. Then W has a real Hamiltonian square root.

Proof : Every $A \in \mathbb{R}^{n\times n}$ can be expressed as a product $A = FG$ of two real symmetric[1] matrices F and G [4, 16, 22, 24]. Consequently any block-diagonal skew-Hamiltonian matrix $\left[\begin{smallmatrix} A & 0 \\ 0 & A^T \end{smallmatrix}\right]$ has a Hamiltonian square root of the form $\left[\begin{smallmatrix} 0 & F \\ G & 0 \end{smallmatrix}\right]$. Then

$$W = S\begin{bmatrix} A & 0 \\ 0 & A^T \end{bmatrix}S^{-1} = S\begin{bmatrix} 0 & F \\ G & 0 \end{bmatrix}^2 S^{-1} = \left(S\begin{bmatrix} 0 & F \\ G & 0 \end{bmatrix}S^{-1}\right)^2$$

expresses W as the square of the Hamiltonian matrix $S\left[\begin{smallmatrix} 0 & F \\ G & 0 \end{smallmatrix}\right]S^{-1}$. $\square$

Is there any reason to believe that such a symplectic block-diagonalization can be achieved for every skew-Hamiltonian matrix? In the special case of 4×4 matrices it has been shown using quaternions that every 4×4 skew-Hamiltonian *can* be block-diagonalized in the sense of Proposition 1; and this can even be done by an *orthogonal*-symplectic similarity [8]. Thus every 4×4 skew-Hamiltonian has a Hamiltonian square root. For larger matrices it is still not clear whether block-diagonalization is always possible via orthogonal-symplectic similarity, so we turn next to see what can be achieved with non-orthogonal symplectic similarities.

To continue moving forward from Van Loan's block-upper-triangular form towards block-diagonal form, we try using similarities by block-upper-triangular symplectics. One can verify directly from the definition that a block-upper-triangular matrix $\left[\begin{smallmatrix} V & X \\ 0 & Y \end{smallmatrix}\right]$ with $V, X, Y \in \mathbb{R}^{n\times n}$ is symplectic iff V is invertible, $Y = V^{-T}$, and $V^{-1}X$ is symmetric. The two simplest types of block-upper-triangular symplectics, then, are the block-diagonal symplectics $\left[\begin{smallmatrix} V & 0 \\ 0 & V^{-T} \end{smallmatrix}\right]$, and the symplectic shears $\left[\begin{smallmatrix} I & X \\ 0 & I \end{smallmatrix}\right]$ with $I \in \mathbb{R}^{n\times n}$ and $X^T = X$. We will see that one can go quite a long way using just these two special types of (non-orthogonal) symplectic matrices.[2]

Now consider the set $\mathcal{M}$ of all $2n \times 2n$ skew-Hamiltonian matrices whose eigenvalues each have multiplicity exactly two. From the Van Loan reduction (1) it is clear that any eigenvalue of a skew-Hamiltonian matrix must have *even* multiplicity, so $\mathcal{M}$ consists precisely of those matrices in $\mathcal{W}$ whose eigenvalues are of *minimal* multiplicity. Thus $\mathcal{M} \subseteq \mathcal{W}$ can be viewed as the natural skew-Hamiltonian analog of the subset of matrices in $\mathbb{R}^{2n\times 2n}$ with distinct eigenvalues; it should then not be surprising that $\mathcal{M}$ is a dense open subset of $\mathcal{W}$ with complement $\mathcal{W} \setminus \mathcal{M}$ of measure zero. In this sense, we may regard $\mathcal{M}$ as the "generic" skew-Hamiltonian matrices. The next proposition shows that the simple tools introduced so far are already sufficient to symplectically block-diagonalize any skew-Hamiltonian matrix in $\mathcal{M}$.

[1] Either one of the matrices F or G may also be chosen to be nonsingular. This extra property is not needed here, but plays an important role later in the proof of Theorem 3.

[2] Every block-upper-triangular symplectic can be uniquely expressed as the product of a block-diagonal symplectic and a symplectic shear, although we make no use of this fact here.

Proposition 2 *For any $W \in \mathcal{M}$ there exists an $S \in \mathcal{S}$ such that*

$$S^{-1}WS = \begin{bmatrix} A & 0 \\ 0 & A^T \end{bmatrix}$$

with $A \in \mathbb{R}^{n\times n}$.

Proof: A sequence of three symplectic similarities reduces any $W \in \mathcal{M}$ to block-diagonal form. First do Van Loan's reduction, constructing $S_1 \in \mathcal{S}$ so that $S_1^{-1}WS_1 = \left[\begin{smallmatrix} U & R \\ 0 & U^T \end{smallmatrix}\right]$. The assumption $W \in \mathcal{M}$ means that $U \in \mathbb{R}^{n\times n}$ has n distinct eigenvalues. Next perform a similarity by a block-diagonal symplectic $S_2 = \left[\begin{smallmatrix} V & 0 \\ 0 & V^{-T} \end{smallmatrix}\right]$, choosing $V \in \mathbb{R}^{n\times n}$ so that $V^{-1}UV = A$ is in real Jordan form. This gives

$$S_2^{-1}S_1^{-1}WS_1S_2 = \begin{bmatrix} A & K \\ 0 & A^T \end{bmatrix},$$

with $K = V^{-1}RV^{-T}$. The block-diagonalization of W is completed by similarity with a symplectic shear $S_3 = \left[\begin{smallmatrix} I & X \\ 0 & I \end{smallmatrix}\right]$. We have

$$S_3^{-1}S_2^{-1}S_1^{-1}WS_1S_2S_3 = \begin{bmatrix} A & AX - XA^T + K \\ 0 & A^T \end{bmatrix}.$$

All that remains, then, is to show that for any skew-symmetric K one can always find a *symmetric* solution X to the Sylvester equation

$$AX - XA^T = -K. \tag{2}$$

Using such a solution X in the shear S_3, we will have $S^{-1}WS = \left[\begin{smallmatrix} A & 0 \\ 0 & A^T \end{smallmatrix}\right]$ with $S = S_1S_2S_3$, and the proposition will be proved.

In solving (2) we can make use of the following well-known and fundamental fact about Sylvester equations of the form $AX - XB = Y$, where $A \in \mathbb{R}^{k\times k}$, $B \in \mathbb{R}^{\ell\times\ell}$ and $Y \in \mathbb{R}^{k\times\ell}$: whenever the spectra of A and B are disjoint, then the equation $AX - XB = Y$ has a unique solution $X \in \mathbb{R}^{k\times\ell}$ for any $Y \in \mathbb{R}^{k\times\ell}$ (see Proposition 3 in §3.1). To bring this result into play to solve (2), partition A, X, and K into blocks compatible with the block-diagonal structure of A. Since A is the real Jordan form of a matrix with distinct eigenvalues, we can write $A = A_{11} \oplus A_{22} \oplus \ldots \oplus A_{mm}$ where each A_{ii} is 1×1 or 2×2, and any 2×2 diagonal block A_{ii} has the form $\left[\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right]$ with $b \neq 0$.

With X and K partitioned conformally with A, observe that the ij^{th} block of $AX - XA^T$ depends only on the ij^{th} block of X :

$$(AX - XA^T)_{ij} = A_{ii}X_{ij} - X_{ij}A_{jj}^T, \qquad i, j = 1, 2, \ldots, m.$$

Thus the equation $AX - XA^T = -K$ decomposes blockwise into m^2 *independent* subproblems

$$A_{ii}X_{ij} - X_{ij}A_{jj}^T = -K_{ij}, \qquad i, j = 1, 2, \ldots, m. \tag{3}$$

Among the diagonal-block subproblems ($i = j$) there are two cases to consider. Whenever A_{ii} is 1×1, equation (3) collapses to a scalar equation and any real X_{ii} is a solution. When

A_{ii} is 2×2, equation (3) can be solved by a simple computation. Let $A_{ii} = \left[\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right]$ with $b \neq 0$, $X_{ii} = \left[\begin{smallmatrix} x & y \\ y & z \end{smallmatrix}\right]$, and $-K_{ii} = \left[\begin{smallmatrix} 0 & -k \\ k & 0 \end{smallmatrix}\right]$. Then

$$A_{ii}X_{ii} - X_{ii}A_{ii}^T = b\begin{bmatrix} 0 & -x-z \\ x+z & 0 \end{bmatrix},$$

so $x = k/b$ together with $y = z = 0$ provides one of many possible symmetric solutions X_{ii}. For each of the off-diagonal-block subproblems ($i \neq j$) the above fundamental fact guarantees the existence of a unique solution X_{ij} to (3). Taking transpose of both sides of (3) shows that these blockwise solutions satisfy $X_{ji} = X_{ij}^T$, and thus fit together compatibly to form a symmetric solution X for (2). □

REMARKS

1. This proof highlights the importance of the solvability of various types of Sylvester equations, especially

$$AX - XA^T = Y\,, \tag{4}$$

   for the symplectic block-diagonalization problem. By extending the argument used above, we will see in the next section how to *characterize* the set of matrices A for which (4) has a symmetric solution X for every skew-symmetric Y. As one might guess from Proposition 2, among such A's are all matrices with distinct eigenvalues. The counterexample $A = I$ suggests that multiple eigenvalues cause difficulties, but that is not always the case. It turns out that the problem is not multiple eigenvalues per se, but rather multiple Jordan blocks (see Proposition 5 in §3.1).

2. In light of the first remark, we can now see that the second step of the above reduction to block-diagonal form (similarity by the block-diagonal symplectic S_2) is unnecessary. For matrices in $\mathcal{M}$, one can always go directly from the Van Loan reduced form $\left[\begin{smallmatrix} U & R \\ 0 & U^T \end{smallmatrix}\right]$ to block-diagonal form $\left[\begin{smallmatrix} U & 0 \\ 0 & U^T \end{smallmatrix}\right]$ via similarity by some symplectic shear $\left[\begin{smallmatrix} I & Z \\ 0 & I \end{smallmatrix}\right]$. Such a similarity leads to the equation $UZ - ZU^T + R = 0$ where U has distinct eigenvalues, so it will have a symmetric solution Z for any skew-symmetric R. With V and X defined as in the proof of Proposition 2, the matrix $Z = VXV^T$ is one such solution, although there are many others (in fact there is a whole n-dimensional affine subspace of symmetric solutions).

Now that we know that $\mathcal{H}^2$ contains an open dense subset of W, it is natural to consider trying the usual kind of analytic argument to complete the proof that $\mathcal{H}^2 = \mathcal{W}$. That is, we could approximate an arbitrary skew-Hamiltonian W by a sequence $W_i \longrightarrow W$ with $W_i \in \mathcal{M}$, pick Hamiltonian square roots H_i for each W_i by Proposition 2, and then try to show that the set $\{H_i\}$ has some limit point H. Any such H would be a Hamiltonian square root of W. Now even though each $W_i \in \mathcal{M}$ has infinitely many Hamiltonian square roots, it is not immediately evident that one can always choose the H_i so as to guarantee the existence of any limit points at all for $\{H_i\}$. Instead of pursuing this analytic line of attack, we will continue with a more algebraic approach, showing that the symplectic block-diagonalization result of Proposition 2 can be extended to all of $\mathcal{W}$. Indeed we will prove the following canonical form result, which may itself be of some independent interest.

Theorem 1 *Every real skew-Hamiltonian matrix can be brought into "skew-Hamiltonian Jordan form" via symplectic similarity. That is, for any $W \in \mathcal{W}$ there exists an $S \in \mathcal{S}$ such that*

$$S^{-1}WS = \begin{bmatrix} A & 0 \\ 0 & A^T \end{bmatrix},$$

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

As an immediate corollary we then have

Theorem 2 *Every real skew-Hamiltonian matrix has a real Hamiltonian square root. In other words, $\mathcal{H}^2 = \mathcal{W}$.*

The main goal of the next two sections of the paper is to prove Theorem 1. Unfortunately, the potential presence of nontrivial Jordan structure in the general skew-Hamiltonian matrix introduces difficulties which cannot be handled using only symplectic shears, although they still have an important role to play. We begin with some technical results concerning the detection and manipulation of Jordan structure, and further results about Sylvester equations.

3 Auxiliary Results

3.1 Sylvester Equations

In the proof of Proposition 2 we have seen that the effect of similarity by a symplectic shear on a block-upper-triangular skew-Hamiltonian matrix is simply to replace the $(1,2)$ block K by the expression $AX - XA^T + K$:

$$\begin{bmatrix} I & X \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} A & K \\ 0 & A^T \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & AX - XA^T + K \\ 0 & A^T \end{bmatrix}.$$

It is important for the symplectic reduction of general skew-Hamiltonian matrices to structured Jordan form to find out how far it is possible to simplify various types of such matrix expressions. These simplification questions can be concisely expressed in terms of the corresponding "Sylvester operators", so let us introduce the following notation. Suppose $A \in F^{k\times k}$ and $B \in F^{\ell\times\ell}$ are fixed but arbitrary square matrices with entries in the field F. Then we denote by $\mathrm{Syl}(A,B)$ the linear Sylvester operator

$$\begin{aligned} \mathrm{Syl}(A,B)\ : F^{k\times\ell} &\longrightarrow F^{k\times\ell} \\ X &\mapsto AX - XB. \end{aligned}$$

In this section we characterize the range of several types of such operators, beginning with a well-known result referred to earlier in the proof of Proposition 2. Knowing the range of an operator $\mathrm{Syl}(A,B)$ enables us immediately to see how much it is possible to simplify the "Sylvester expression" $AX - XB + Y$ for an arbitrary Y.

Although the proof of Theorem 1 involves only real Sylvester operators and the simplification of their associated Sylvester expressions, it is often convenient to prove results first for complex operators and then derive the real case from the complex case. Much of this section follows that pattern. In order to proceed directly to our main result in §4, the proofs of Propositions 4, 5 and 6 will be deferred to an appendix. We use F here to denote $\mathbb{C}$ or $\mathbb{R}$.

Proposition 3 *Let $A \in F^{k \times k}$ and $B \in F^{\ell \times \ell}$. Then the operator* $\mathrm{Syl}(A, B)$ *is nonsingular iff the spectra $\lambda(A)$ and $\lambda(B)$ are disjoint subsets of $\mathbb{C}$.*

Proof: Proofs of this result for $F = \mathbb{C}$ can be found in many places, e.g. [2, 10, 16, 17, 22, 23]. When A and B are real, $\mathcal{T} = \mathrm{Syl}(A, B)$ may be viewed either as a real operator $\mathcal{T}_R$ or as a complex operator $\mathcal{T}_C$. Since $\mathcal{T}_C$ is the "complexification" of $\mathcal{T}_R$ [12, 15, 22], we have $\dim_{\mathbb{C}}(\ker \mathcal{T}_C) = \dim_{\mathbb{R}}(\ker \mathcal{T}_R)$. Thus $\mathcal{T}_R$ is nonsingular iff $\mathcal{T}_C$ is nonsingular. $\Box$

Next we characterize the range of one of the simplest types of singular Sylvester operator with $\lambda(A) = \lambda(B)$. Let N_k denote the $k \times k$ nilpotent matrix $\begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}$, and $M_k(\lambda) = \lambda I_k + N_k$ denote the $k \times k$ Jordan block corresponding to the eigenvalue λ. We also need the notion of the m^{th} *antidiagonal* of a matrix $Y \in F^{k \times \ell}$, by which is meant the set of all entries Y_{ij} such that $i + j - 1 = m$. Note that a $k \times \ell$ matrix Y has a total of $k + \ell - 1$ antidiagonals. Since the collection of the m^{th} antidiagonals of $Y \in F^{k \times \ell}$ with $m = k + \ell - 1\,,\ k + \ell - 2\,,\ \ldots\,,\ k + \ell - d$ plays a particularly important role in the following result, we refer to this collection as the *last* d antidiagonals of Y.

Proposition 4 *Consider the operator* $\mathrm{Syl}(A, B^T)$*, where $A = M_k(\lambda)$ and $B = M_\ell(\lambda)$ are Jordan blocks corresponding to the same eigenvalue $\lambda \in F$. Let $d = \min(k, \ell)$. Then the range of* $\mathrm{Syl}(A, B^T)$ *consists of all $Y \in F^{k \times \ell}$ such that the sum of the entries along each of the last d antidiagonals of Y is zero. Thus* $\dim_F(\mathrm{range}\,\mathrm{Syl}(A, B^T)) = k\ell - d$.

Many of the Sylvester expressions $AX - XB + Y$ arising in the proof of Theorem 1 require simplification with a *symmetric* X, not just with an arbitrary unstructured X. This is because shears $\left[\begin{smallmatrix} I & X \\ 0 & I \end{smallmatrix}\right]$ are symplectic iff X is symmetric. We address this situation in the next proposition. But first a little more notation: let F-Sym(n) and F-Skew(n) denote the sets of all matrices $X \in F^{n \times n}$ such that $X^T = X$ and $X^T = -X$, respectively. Also recall that a matrix $A \in F^{n \times n}$ is said to be *nonderogatory* [17] if the complex Jordan form of A has exactly one Jordan block for each eigenvalue.

Proposition 5 *For $A \in F^{n \times n}$, consider the operator* $\mathrm{Syl}(A, A^T)$ *with domain and codomain restricted to* F-Sym(n) *and* F-Skew(n)*, respectively. That is, consider*

$$\begin{aligned} \mathcal{T}_A : \text{ F-Sym}(n) &\longrightarrow \text{F-Skew}(n) \\ X &\mapsto AX - XA^T \,. \end{aligned}$$

Then $\mathcal{T}_A$ is onto $\iff$ A is nonderogatory.

The final result we need is the real analog of Proposition 4 for complex-conjugate eigenvalue pairs. Our goal is to characterize the range of real operators $\mathrm{Syl}(A, B^T) : \mathbb{R}^{2k\times 2\ell} \to \mathbb{R}^{2k\times 2\ell}$ in completely real terms, when A and B are real Jordan blocks corresponding to the same complex-conjugate eigenvalue pair. To achieve this, we need some preliminary definitions and simple facts about real 2×2 matrices and their relations to complex 2×2 matrices.

Consider the *centralizer* $\mathcal{C}_2$ and *anticentralizer* $\mathcal{A}_2$ of $J = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$ defined by

$$\begin{aligned} \mathcal{C}_2 &= \{X \in \mathbb{R}^{2\times 2} \mid JX = XJ\} = \{\left[\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right] \mid a, b \in \mathbb{R}\} \\ \text{and} \quad \mathcal{A}_2 &= \{X \in \mathbb{R}^{2\times 2} \mid JX = -XJ\} = \{\left[\begin{smallmatrix} c & d \\ d & -c \end{smallmatrix}\right] \mid c, d \in \mathbb{R}\}. \end{aligned}$$

Then the following lemma is straightforward to prove.

Lemma 1

1. $\mathbb{R}^{2\times 2} = \mathcal{C}_2 \oplus \mathcal{A}_2$.
2. *Every matrix in $\mathcal{C}_2$ is diagonalized by similarity with the unitary matrix $\Phi_2 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 & 1 \\ -i & i \end{smallmatrix}\right]$. That is, $\Phi_2^H \left[\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right] \Phi_2 = \left[\begin{smallmatrix} a+ib & 0 \\ 0 & a-ib \end{smallmatrix}\right]$.*
3. *Every matrix in $\mathcal{A}_2$ is "anti-diagonalized" by similarity with Φ_2. That is, $\Phi_2^H \left[\begin{smallmatrix} c & d \\ d & -c \end{smallmatrix}\right] \Phi_2 = \left[\begin{smallmatrix} 0 & c+id \\ c-id & 0 \end{smallmatrix}\right]$.*
4. *The set $\mathcal{U}_2 = \{\left[\begin{smallmatrix} u & v \\ \bar{v} & \bar{u} \end{smallmatrix}\right] \mid u, v \in \mathbb{C}\}$ is a (real) subalgebra of $\mathbb{C}^{2\times 2}$, and the map*
$$\begin{aligned} \mathbb{R}^{2\times 2} &\xrightarrow{\cong} \mathcal{U}_2 \\ X &\mapsto \Phi_2^H X \Phi_2 \end{aligned}$$
is an algebra isomorphism.

With these facts in hand we can now prove the following proposition.

Proposition 6 *Consider the operator* $\mathrm{Syl}(A, B^T)$*, where $A \in \mathbb{R}^{2k\times 2k}$ and $B \in \mathbb{R}^{2\ell\times 2\ell}$ are both real Jordan blocks corresponding to the same complex-conjugate eigenvalue pair $a \pm ib$. That is, both A and B are of the form*

$$\begin{bmatrix} \Lambda & I_2 & & \\ & \ddots & \ddots & \\ & & \ddots & I_2 \\ & & & \Lambda \end{bmatrix}, \quad \text{where } \Lambda = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \text{ with } b \neq 0. \tag{5}$$

The range of $\mathrm{Syl}(A, B^T)$ *can be characterized as follows. Let $d = \min(k, \ell)$. Partition $Y \in \mathbb{R}^{2k\times 2\ell}$ into blocks $Y_{ij} \in \mathbb{R}^{2\times 2}$ so that $Y = \begin{bmatrix} Y_{11} & \cdots & Y_{1\ell} \\ \vdots & & \vdots \\ Y_{k1} & \cdots & Y_{k\ell} \end{bmatrix}$. Let the set of all 2×2 blocks Y_{ij} such that $i+j-1 = m$ be called the m^{th} block-antidiagonal of Y. Then the range of* $\mathrm{Syl}(A, B^T)$ *consists of all $Y \in \mathbb{R}^{2k\times 2\ell}$ such that the sum of the $\mathcal{A}_2$-components of the blocks along each of the last d block-antidiagonals of Y is $\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$.*

From the above characterizations of the ranges of Sylvester operators, we can now see exactly how much it is possible to simplify the four types of Sylvester expression $AX-XB+Y$ appearing in the proof of Theorem 1. Each simplification result is an immediate consequence of the indicated proposition.

Proposition 7

(a) *Suppose $A \in \mathbb{R}^{k\times k}$ and $B \in \mathbb{R}^{\ell\times\ell}$ have disjoint spectra. Then for any $Y \in \mathbb{R}^{k\times\ell}$ there exists (a unique) $X \in \mathbb{R}^{k\times\ell}$ such that $AX - XB + Y = 0$. (Proposition 3)*

(b) *Suppose $A \in \mathbb{R}^{k\times k}$ is a real Jordan block corresponding to either a real eigenvalue or a complex-conjugate eigenvalue pair. Then for any skew-symmetric $Y \in \mathbb{R}^{k\times k}$ there exist (infinitely many) symmetric $X \in \mathbb{R}^{k\times k}$ such that $AX - XA^T + Y = 0$. (Proposition 5)*

(c) *Suppose $A \in \mathbb{R}^{k\times k}$ and $B \in \mathbb{R}^{\ell\times\ell}$ are Jordan blocks corresponding to the same real eigenvalue. Let $d = \min(k,\ell)$. Then for any $Y \in \mathbb{R}^{k\times\ell}$ there exist (infinitely many) $X \in \mathbb{R}^{k\times\ell}$ such that $AX - XB^T + Y$ is zero everywhere except possibly in the last d entries of the bottom row. (Proposition 4)*

(d) *Suppose $A \in \mathbb{R}^{2k\times 2k}$ and $B \in \mathbb{R}^{2\ell\times 2\ell}$ are real Jordan blocks corresponding to the same complex-conjugate eigenvalue pair. Let $d = \min(k,\ell)$. Then for any $Y \in \mathbb{R}^{2k\times 2\ell}$ there exist (infinitely many) $X \in \mathbb{R}^{2k\times 2\ell}$ such that $AX - XB^T + Y$ is zero everywhere except possibly in the last d (2×2)–blocks of the bottom row. These last d (2×2)–blocks are all elements of $\mathcal{A}_2$. (Proposition 6)*

3.2 Jordan Structure

An important step in the proof of Theorem 1 concerns certain block-diagonal matrices B and perturbations $\widetilde{B} = B + C_p$ that differ from each other only in a single column of blocks. We need to compare the maximum Jordan block size of such pairs B and $\widetilde{B}$. The results in this section address this question.

For a matrix $A \in \mathbb{C}^{n\times n}$ with eigenvalue λ, it is well-known [16] that the Jordan structure of A corresponding to λ can be deduced from the ranks of the powers of $A - \lambda I$. That is, if $r_k = \operatorname{rank}(A-\lambda I)^k$, then the number and sizes of all the Jordan blocks associated with λ are completely determined by the sequence of numbers $r_0, r_1, \cdots, r_n$. For our purposes, we need only the following basic result.

Proposition 8 *Suppose $A \in \mathbb{C}^{n\times n}$ is a matrix with exactly one eigenvalue λ. Letting $r_k = \operatorname{rank}(A-\lambda I)^k$, there is an integer s with $0 < s \le n$ such that $n = r_0 > r_1 > \cdots > r_s = r_{s+1} = \cdots = r_n = 0$. The largest Jordan block of A has size $s \times s$.*

Next consider matrices of the form

$$\begin{bmatrix} B_1 & & & & F & & & \\ & B_2 & & & * & & & \\ & & \ddots & & \vdots & & & \\ & & & B_{p-1} & * & & & \\ & & & & B_p & & & \\ & & & & * & B_{p+1} & & \\ & & & & \vdots & & \ddots & \\ & & & & * & & & B_q \end{bmatrix} = B + C_p, \tag{6}$$

where $B = \mathrm{diag}(B_1, B_2, \cdots, B_q)$, C_p is zero everywhere except possibly in the off-diagonal blocks of the p^{th} column of blocks, and $*$ stands for an arbitrary matrix of the appropriate size. Observe that if we fix the sizes of the diagonal blocks and the column p, then the set of matrices of the form (6) is closed under multiplication. We have the following two results for certain special matrices of this form.

Proposition 9 *Suppose $A \in \mathbb{C}^{n\times n}$ is a matrix of the form (6) satisfying the following conditions:*

1. *Each $B_k \in \mathbb{C}^{n_k\times n_k}$ with $k \neq p$ is a Jordan block $M_{n_k}(\lambda)$ corresponding to the same eigenvalue λ; $B_p \in \mathbb{C}^{n_p\times n_p}$ is the transpose of a Jordan block corresponding to λ, i.e. $B_p = M^T_{n_p}(\lambda)$.*
2. *B_1 is the largest block on the diagonal of A, so that $n_1 \geq n_k$ for all k.*
3. *The non-zero off-diagonal blocks are* not *in the first column, i.e. $p > 1$. The topmost block of the p^{th} column of blocks, $F \in \mathbb{C}^{n_1\times n_p}$, is of the form $\left[\begin{smallmatrix} 0 \\ f \end{smallmatrix}\right]$ where $f = [f_1 \cdots f_{n_p}] \in \mathbb{C}^{1\times n_p}$ is nonzero.*

Then λ is the only eigenvalue of A, and in the Jordan canonical form of A there is at least one Jordan block with size bigger than $n_1 \times n_1$. Hence the largest Jordan block of A is strictly bigger than the largest Jordan block of B.

Proof: That λ is the only eigenvalue of A follows immediately from partitioning A into block-upper-triangular form $A = \left[\begin{smallmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{smallmatrix}\right]$ where $A_{11} = \mathrm{diag}(B_1, B_2 \cdots B_{p-1})$. In order to establish the claim about Jordan block size, it suffices (by Proposition 8) to show that $\mathrm{rank}(A - \lambda I)^{n_1} > 0$, or equivalently that $(A - \lambda I)^{n_1} \neq 0$. Thus we consider powers of

$$A - \lambda I = \mathrm{diag}(N_{n_1}, N_{n_2} \cdots N^T_{n_p} \cdots N_{n_q}) + C_p.$$

All powers $(A-\lambda I)^k$ are of the form (6), and for convenience we designate the topmost block in the p^{th} column of blocks of $(A - \lambda I)^k$ by $F^{(k)}$. Then it is easy to see inductively that

$$(A - \lambda I)^k = \begin{bmatrix} N^k_{n_1} & & & & \\ & \ddots & & & \\ & & (N^T_{n_p})^k & & \\ & & & \ddots & \\ & & & & N^k_{n_q} \end{bmatrix} + \begin{bmatrix} 0 & & F^{(k)} & & \\ & \ddots & \vdots & & \\ & & 0 & & \\ & & \vdots & \ddots & \\ & & * & & 0 \end{bmatrix},$$

where $F^{(k)}$ satisfies the recurrence

$$F^{(k)} = N_{n_1}^{k-1} F^{(1)} + F^{(k-1)} N_{n_p}^T , \quad F^{(1)} = F.$$

From this recurrence we deduce that

$$F^{(n_1)} = \begin{bmatrix} f_1 & \cdots & f_{n_p} \\ \vdots & \cdot^{\cdot^{\cdot}} & \\ f_{n_p} & & \\ & 0_{(n_1-n_p)\times n_p} & \end{bmatrix}$$

is a Hankel matrix with f as the top row. Since f is non-zero, so is $F^{(n_1)}$ and hence also $(A - \lambda I)^{n_1}$. $\square$

To complete this section we establish a real analog of Proposition 9 for matrices of the form (6) where each B_k is a real Jordan block corresponding to the same complex-conjugate eigenvalue pair. We employ the same sort of strategy as in the proof of Proposition 6; first convert the real Jordan blocks to complex Jordan blocks by an appropriate similarity, then apply Proposition 9 to the resulting complex matrix, and finally translate back into completely real terms.

Proposition 10 *Suppose $L \in \mathbb{R}^{2n\times 2n}$ is a matrix of the form (6) satisfying the following conditions:*

1. *Each $B_k \in \mathbb{R}^{2n_k\times 2n_k}$ with $k \neq p$ is a real Jordan block corresponding to the complex-conjugate eigenvalue pair $(\lambda, \overline{\lambda}) = (a+ib, a-ib)$; $B_p \in \mathbb{R}^{2n_p\times 2n_p}$ is the transpose of a real Jordan block corresponding to $(\lambda, \overline{\lambda})$. In other words, B_p^T and B_k with $k \neq p$ have the form (5) as described in Proposition 6.*

2. *$B_1 \in \mathbb{R}^{2n_1\times 2n_1}$ is the largest block on the diagonal of L, so $n_1 \geq n_k$ for all k.*

3. *The non-zero off-diagonal blocks are* not *in the first column, i.e. $p > 1$. When each block in the p^{th} column of C_p is partitioned into (2×2) sub-blocks, then every such (2×2) sub-block is an element of $\mathcal{A}_2$. The topmost block of the p^{th} column of blocks, $F \in \mathbb{R}^{2n_1\times 2n_p}$, has the form $\left[\begin{smallmatrix} 0 \\ g \end{smallmatrix}\right]$ where $g = [g_1 \cdots g_{n_p}] \in \mathbb{R}^{2\times 2n_p}$ is nonzero and $g_i \in \mathcal{A}_2$ for $1 \leq i \leq n_p$.*

Then λ and $\overline{\lambda}$ are the only eigenvalues of L, and in the real Jordan canonical form of L there is at least one real Jordan block with size bigger than $2n_1 \times 2n_1$. Hence the largest real Jordan block of L is strictly bigger than the largest real Jordan block of B.

Proof: Recall the unitary matrix $\Psi_{2n} = \Phi_{2n} P_{2n}$ defined in the proof of Proposition 6. From the discussion there of the effect of similarity by Ψ_{2n} on real matrices, we see that $\widehat{L} = \Psi_{2n}^H \, L \, \Psi_{2n}$ will be of the form $\left[\begin{smallmatrix} U & V \\ \overline{V} & \overline{U} \end{smallmatrix}\right]$ with $U, V \in \mathbb{C}^{n\times n}$. More specifically,

$$U = \mathrm{diag}(M_{n_1}, M_{n_2} \cdots \overline{M}_{n_p}^T \cdots M_{n_q})$$

where $M_{n_k} = M_{n_k}(\lambda)$ denotes the $n_k \times n_k$ Jordan block for $\lambda = a + ib$. The matrix V, partitioned conformally with U, has non-zero entries only in the off-diagonal blocks of the p^{th} column of blocks, i.e.

$$V = \begin{bmatrix} 0 & & & \widehat{F} & & \\ & 0 & & * & & \\ & & \ddots & \vdots & & \\ & & & 0 & & \\ & & & \vdots & \ddots & \\ & & & * & & 0 \end{bmatrix}.$$

The topmost block $\widehat{F} \in \mathbb{C}^{n_1 \times n_p}$ of this p^{th} column has the form $\left[\begin{smallmatrix} 0 \\ \widehat{f} \end{smallmatrix}\right]$ where $\widehat{f} = [\widehat{f}_1 \cdots \widehat{f}_{n_p}] \in \mathbb{C}^{1 \times n_p}$ is nonzero.

A final permutation similarity shifts these non-zero blocks of V into U, and thus block-diagonalizes L. Letting $P = \left[\begin{smallmatrix} C & S \\ S & C \end{smallmatrix}\right]$, with

$$\begin{aligned} C &= \operatorname{diag}(I_{n_1} \cdots I_{n_{p-1}}, 0_{n_p}, I_{n_{p+1}} \cdots I_{n_q}) \quad \text{and} \\ S &= \operatorname{diag}(0_{n_1} \cdots 0_{n_{p-1}}, I_{n_p}, 0_{n_{p+1}} \cdots 0_{n_q}), \end{aligned}$$

we have $P^T \widehat{L} P = \left[\begin{smallmatrix} A & 0 \\ 0 & \overline{A} \end{smallmatrix}\right]$, where

$$A = \begin{bmatrix} M_{n_1} & & & \widehat{F} & & \\ & M_{n_2} & & * & & \\ & & \ddots & \vdots & & \\ & & & M_{n_p}^T & & \\ & & & \vdots & \ddots & \\ & & & * & & M_{n_q} \end{bmatrix}$$

is exactly the type of matrix considered in Proposition 9. Thus A and $\overline{A}$ have only the eigenvalues λ and $\overline{\lambda}$, respectively, and each has at least one Jordan block bigger than $n_1 \times n_1$. Consequently L has only the eigenvalues λ and $\overline{\lambda}$, and at least one real Jordan block bigger than $2n_1 \times 2n_1$. $\square$

4 Skew-Hamiltonian Jordan Form

The results of §3 provide us with all the technical tools needed to show that every real skew-Hamiltonian matrix can be symplectically brought into structured real Jordan form. This is the content of Theorem 1, which we recall now and prove.

Theorem 1 *For any $W \in \mathcal{W}$ there exists an $S \in \mathcal{S}$ such that*

$$S^{-1} W S = \begin{bmatrix} A & 0 \\ 0 & A^T \end{bmatrix},$$

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

Proof: Begin as in Proposition 2 with Van Loan's reduction, constructing an orthogonal-symplectic $S_1 \in \mathcal{S}$ so that $S_1^{-1} W S_1 = \left[\begin{smallmatrix} U & R \\ 0 & U^T \end{smallmatrix}\right]$. Next perform a similarity with a block-diagonal symplectic $S_2 = \left[\begin{smallmatrix} V & 0 \\ 0 & V^{-T} \end{smallmatrix}\right]$, where $V \in \mathbb{R}^{n\times n}$ is chosen so that $V^{-1}UV = D$ is in real Jordan form. Then we have $S_2^{-1} S_1^{-1} W S_2 S_1 = \left[\begin{smallmatrix} D & K \\ 0 & D^T \end{smallmatrix}\right]$, where $K = V^{-1} R V^{-T}$. We will assume that the real Jordan blocks of D corresponding to the same real eigenvalue (or to the same complex-conjugate eigenvalue pair) have all been grouped together into blocks B_i; that is, we can write $D = \operatorname{diag}(B_1, B_2 \cdots B_\ell)$, where

(i) the spectrum of each B_i is either a single real number or a single complex-conjugate pair, and

(ii) distinct blocks B_i and B_j have disjoint spectra.

With a symplectic shear $\left[\begin{smallmatrix} I & X \\ 0 & I \end{smallmatrix}\right]$ we can now block-diagonalize K. Recall that the effect of similarity by a symplectic shear on a block-upper-triangular skew-Hamiltonian matrix $\left[\begin{smallmatrix} D & K \\ 0 & D^T \end{smallmatrix}\right]$ is simply to replace K by the expression $DX - XD^T + K$. Thus we wish to find a symmetric X so that $DX - XD^T + K$ is a block-diagonal matrix conformal with D. To build such an X, start by partitioning K and X into blocks conformal with the direct-sum decomposition $D = B_1 \oplus B_2 \oplus \cdots \oplus B_\ell$. The block-diagonal nature of D means that $DX - XD^T + K$ may be handled blockwise:

$$(DX - XD^T + K)_{ij} = B_i X_{ij} - X_{ij} B_j^T + K_{ij}\ , \qquad 1 \le i, j \le \ell.$$

Since B_i and B_j have disjoint spectra for any $i \neq j$, we know from Proposition 7a that there is a unique $\widetilde{X}_{ij}$ such that $B_i \widetilde{X}_{ij} - \widetilde{X}_{ij} B_j^T + K_{ij} = 0$. Transposing this equation and invoking the skew-symmetry of K (i.e. $K_{ij}^T = -K_{ji}$) shows that $\widetilde{X}_{ji} = \widetilde{X}_{ij}^T$. Letting $\widetilde{X}_{ii} = 0$ for $1 \le i \le \ell$, we see that the blocks $\widetilde{X}_{ij}$ fit together to form a symmetric matrix $\widetilde{X}$. The corresponding symplectic shear $S_3 = \left[\begin{smallmatrix} I & \widetilde{X} \\ 0 & I \end{smallmatrix}\right]$ gives us

$$S_3^{-1} S_2^{-1} S_1^{-1} W S_1 S_2 S_3 = \begin{bmatrix} D & K_{\text{diag}} \\ 0 & D^T \end{bmatrix}.$$

where $K_{\text{diag}} = \operatorname{diag}(K_{11}, K_{22} \cdots K_{\ell\ell})$. The problem of symplectically block-diagonalizing an arbitrary real skew-Hamiltonian is thus reduced to that of symplectically block-diagonalizing "degenerate" skew-Hamiltonian matrices $\left[\begin{smallmatrix} B_i & K_{ii} \\ 0 & B_i^T \end{smallmatrix}\right]$, that is skew-Hamiltonian matrices whose spectrum consists either of a single real number (type 1) or a single complex-conjugate pair (type 2).

Up to this point, the proof of Theorem 1 is essentially the same as the proof of Proposition 2. In the generic class $\mathcal{M}$ of skew-Hamiltonians considered in Proposition 2, however, the B_i were only 1×1 or 2×2 blocks, and the corresponding degenerate subproblems could be handled directly and explicitly in an elementary manner. It is in block-diagonalizing larger degenerate subproblems that the chief technical difficulty of the general case lies. Symplectic shears alone cannot in general be sufficient for this task, because the real Jordan structure of $\left[\begin{smallmatrix} B_i & K_{ii} \\ 0 & B_i^T \end{smallmatrix}\right]$ may differ from that of $\left[\begin{smallmatrix} B_i & 0 \\ 0 & B_i^T \end{smallmatrix}\right]$. This is where the results of §3 come into play.

To complete the proof of Theorem 1 we describe an iterative procedure, terminating in a *finite* number of steps, which brings any degenerate skew-Hamiltonian matrix into structured skew-Hamiltonian Jordan form. We concentrate on the type 2 case, matrices with a single complex-conjugate eigenvalue pair $(\lambda, \overline{\lambda})$. It is easy to see that the following argument will also work for the type 1 case, simply by replacing Proposition 7d with Proposition 7c, and Proposition 10 with Proposition 9. Let us suppose, then, that $\left[\begin{smallmatrix} B & K \\ 0 & B^T \end{smallmatrix}\right]$ is a degenerate skew-Hamiltonian matrix where $B = \mathrm{diag}(A_1, A_2 \cdots A_q)$ is in real Jordan form. Each $A_k \in \mathbb{R}^{2n_k \times 2n_k}$ is a real Jordan block of the form (5), and A_1 is the largest such block.

We begin with the termination case for this procedure. When B has only one real Jordan block, then block-diagonalization can be achieved in one step. By Proposition 7b there exists a symmetric X such that $BX - XB^T + K = 0$. Thus similarity by the symplectic shear $\left[\begin{smallmatrix} I & X \\ 0 & I \end{smallmatrix}\right]$ using this X transforms $\left[\begin{smallmatrix} B & K \\ 0 & B^T \end{smallmatrix}\right]$ into $\left[\begin{smallmatrix} B & 0 \\ 0 & B^T \end{smallmatrix}\right]$, and we are done.

Now suppose that B has more than one Jordan block. We define a two-step reduction process to simplify $\left[\begin{smallmatrix} B & K \\ 0 & B^T \end{smallmatrix}\right]$, not necessarily all the way to block-diagonal form, but at least bringing it closer to structured Jordan form.

STEP 1: "Simplification of K"
Here we simplify K as much as possible using only a symplectic shear $\left[\begin{smallmatrix} I & X \\ 0 & I \end{smallmatrix}\right]$. Begin by partitioning K and X conformally with the real Jordan decomposition of B. Now simplify K blockwise, replacing each block K_{ij} by the expression $A_i X_{ij} - X_{ij} A_j^T + K_{ij} = Y_{ij}$, where X_{ij} is chosen to produce a Y_{ij} with as many zeroes as possible. By Proposition 7b each block K_{ii} on the diagonal of K can be zeroed out completely. In general the off-diagonal blocks K_{ij} $(i \neq j)$ cannot be completely zeroed out in this way, but Proposition 7d shows what we can be sure of achieving. For blocks K_{ij} above the diagonal $(i < j)$ choose X_{ij} so that all entries are zeroed out except possibly for the last d (2×2)–blocks of the bottom row. In other words, each Y_{ij} with $i < j$ has the form $Y_{ij} = \left[\begin{smallmatrix} 0 \\ g \end{smallmatrix}\right]$ where $g = [g_1, \cdots g_{n_j}] \in \mathbb{R}^{2 \times 2n_j}$ and $g_i \in \mathcal{A}_2$ for $1 \leq i \leq n_j$. For blocks K_{ji} below the diagonal $(j > i)$ we choose $X_{ji} = X_{ij}^T$ so that X will be symmetric and $Y_{ji} = -Y_{ij}^T$. This zeroes out all entries of each K_{ji} $(j > i)$ except possibly for the bottom d (2×2)–blocks of the last column. Note that these (2×2)–blocks are also elements of $\mathcal{A}_2$, since any $g_i \in \mathcal{A}_2$ is symmetric. Thus we simplify $\left[\begin{smallmatrix} B & K \\ 0 & B^T \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} B & Y \\ 0 & B^T \end{smallmatrix}\right]$ via similarity by the symplectic shear $\left[\begin{smallmatrix} I & X \\ 0 & I \end{smallmatrix}\right]$, where Y has the form

$$Y = \begin{bmatrix} 0 & Y_{12} & \cdots & \cdots & Y_{1q} \\ -Y_{12}^T & 0 & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & 0 & Y_{q-1,q} \\ -Y_{1q}^T & \cdots & \cdots & -Y_{q-1,q}^T & 0 \end{bmatrix}.$$

STEP 2: "Transfer of Jordan structure"
If all the blocks $Y_{12} \cdots Y_{1q}$ in the first row of Y (and hence also all the blocks in the first column of Y) are zero, then we can deflate to a smaller degenerate

skew-Hamiltonian $\left[\begin{smallmatrix} \widetilde{B} & \widetilde{K} \\ 0 & \widetilde{B}^T \end{smallmatrix}\right]$, where $\widetilde{B} = \mathrm{diag}(A_2 \cdots A_q)$ and

$$\widetilde{K} = \begin{bmatrix} 0 & Y_{23} & \cdots & Y_{2q} \\ -Y_{23}^T & 0 & & \vdots \\ \vdots & & \ddots & \vdots \\ -Y_{2q}^T & \cdots & \cdots & 0 \end{bmatrix}.$$

Otherwise there is some block Y_{1p} in the first row of Y that is non-zero. By a permutation-like symplectic similarity on $\left[\begin{smallmatrix} B & Y \\ 0 & B^T \end{smallmatrix}\right]$ we can shift Y_{1p}, indeed the whole p^{th} column of blocks, from Y into B. Let $Q = \left[\begin{smallmatrix} C & -S \\ S & C \end{smallmatrix}\right]$, where

$$\begin{aligned} C &= \mathrm{diag}(I_{2n_1} \cdots I_{2n_{p-1}}, 0_{2n_p}, I_{2n_{p+1}} \cdots I_{2n_q}) \quad \text{and} \\ S &= \mathrm{diag}(0_{2n_1} \cdots 0_{2n_{p-1}}, I_{2n_p}, 0_{2n_{p+1}} \cdots 0_{2n_q}) . \end{aligned}$$

Then we have $Q^T \left[\begin{smallmatrix} B & Y \\ 0 & B^T \end{smallmatrix}\right] Q = \left[\begin{smallmatrix} L & \tilde{Y} \\ 0 & L^T \end{smallmatrix}\right]$, where

$$L = \begin{bmatrix} A_1 & & & Y_{1p} & & \\ & A_2 & & Y_{2p} & & \\ & & \ddots & \vdots & & \\ & & & A_p^T & & \\ & & & \vdots & \ddots & \\ & & & -Y_{pq}^T & & A_q \end{bmatrix}$$

is exactly the type of matrix considered in Proposition 10. Consequently L has only the eigenvalues λ and $\overline{\lambda}$, and the largest real Jordan block of L is strictly bigger than the largest real Jordan block of B. Roughly speaking, similarity by Q has the effect of "transferring some Jordan structure" from Y into B. To complete Step 2, perform a similarity with the block-diagonal symplectic $T = \left[\begin{smallmatrix} Z & 0 \\ 0 & Z^{-T} \end{smallmatrix}\right]$, choosing Z so that $Z^{-1}LZ = \widetilde{B}$ is in real Jordan form with the largest real Jordan block in the $(1,1)$ position.

The result of this two-step reduction process, then, is a matrix

$$\begin{bmatrix} \widetilde{B} & \widetilde{K} \\ 0 & \widetilde{B}^T \end{bmatrix} = T^{-1} \begin{bmatrix} L & \widetilde{Y} \\ 0 & L^T \end{bmatrix} T$$

of the same form as the input $\left[\begin{smallmatrix} B & K \\ 0 & B^T \end{smallmatrix}\right]$ to the two-step reduction process, but with the crucial difference that the largest real Jordan block of $\widetilde{B}$ is strictly bigger than the largest real Jordan block of B.

Now repeat this two-step reduction process on $\left[\begin{smallmatrix} \widetilde{B} & \widetilde{K} \\ 0 & \widetilde{B}^T \end{smallmatrix}\right]$. After finitely many iterations we can either deflate to a smaller degenerate skew-Hamiltonian, or we reach a stage where the largest real Jordan block has grown in size to fill all of $\widetilde{B}$. On any deflated problem we again iterate the two-step reduction; after finitely many iterations we can either deflate once more, or the largest real Jordan block will have grown to fill all of $\widetilde{B}$. Only finitely many such deflations can occur, and ultimately we must reach the termination case, a $\widetilde{B}$ with only one real Jordan block. Block-diagonalization is achieved in one final step as described above.

Thus we have shown that there exists a symplectic S such that $S^{-1}WS = \left[\begin{smallmatrix} A & 0 \\ 0 & A^T \end{smallmatrix}\right]$, with $A \in \mathbb{R}^{n\times n}$ in real Jordan form. But any matrix A is similar to its transpose, so $\left[\begin{smallmatrix} A & 0 \\ 0 & A \end{smallmatrix}\right]$ must be the (usual) real Jordan canonical form of W. The uniqueness of this Jordan form then immediately implies the uniqueness of A, up to a permutation of Jordan blocks. $\square$

5 Infinitely Many Square Roots

With the completion of the proof of Theorem 1, we know that every real skew-Hamiltonian matrix W has at least one real Hamiltonian square root. Let us next consider the set $\sqrt[\mathcal{H}]{W} = \{H \in \mathcal{H} \mid H^2 = W\}$ of all the Hamiltonian square roots of W, and what can be said about the size and topological nature of this set for various $W \in \mathcal{W}$. A closer look at the proof of Theorem 1 shows that there are infinitely many distinct symplectic similarities bringing any given $W \in \mathcal{W}$ into structured real Jordan form. Hence it is quite reasonable to expect that every $W \in \mathcal{W}$ actually has infinitely many distinct Hamiltonian square roots. Indeed, by sharpening the previous arguments we can obtain a uniform lower bound on the size of the $\sqrt[\mathcal{H}]{W}$ sets. First we need one more result about Sylvester operators, strengthening a theorem of Taussky and Zassenhaus [25]. The proof will be deferred to the appendix.

Proposition 11 *Let $A \in \mathbb{R}^{n\times n}$, and consider (as in Proposition 5) the restricted domain Sylvester operator*

$$\begin{aligned} \mathcal{T}_A \ &: \mathbb{R}\text{-Sym}(n) \longrightarrow \mathbb{R}\text{-Skew}(n) \\ &X \mapsto AX - XA^T . \end{aligned}$$

Denote the set of all nonsingular matrices in $\ker \mathcal{T}_A$ *by* $\mathrm{Inv}(\ker \mathcal{T}_A)$. *Then for any $A \in \mathbb{R}^{n\times n}$,* $\mathrm{Inv}(\ker \mathcal{T}_A)$ *is a dense open submanifold of* $\ker \mathcal{T}_A$ *; thus* $\dim \mathrm{Inv}(\ker \mathcal{T}_A) = \dim \ker \mathcal{T}_A$.

With this result in hand we can now establish the following lower bound on the size of Hamiltonian square root sets.

Theorem 3 *Every $2n \times 2n$ real skew-Hamiltonian matrix W has at least a $2n$-parameter family of real Hamiltonian square roots.*

Proof: Pick any fixed $S \in \mathcal{S}$ such that $S^{-1}WS = \left[\begin{smallmatrix} A & 0 \\ 0 & A^T \end{smallmatrix}\right]$ is block-diagonal, and factor A as a product $A = FG$ of $n \times n$ symmetric matrices F and G. Without loss of generality we may also assume that G is non-singular [4, 22, 24]. Since in general there are many such factorizations of A, let us introduce the set

$$\mathcal{G} = \{G \in \mathbb{R}\text{-Sym}(n) \mid G \text{ is nonsingular, and } F = AG^{-1} \text{ is symmetric}\}.$$

Then as we have previously seen, $H_G = S \left[\begin{smallmatrix} 0 & F \\ G & 0 \end{smallmatrix}\right] S^{-1}$ is a Hamiltonian square root of W for any $G \in \mathcal{G}$.

To construct even more elements of $\sqrt[\mathcal{H}]{W}$ from H_G, consider symplectic shears $T_X = \left[\begin{smallmatrix} I & X \\ 0 & I \end{smallmatrix}\right]$ such that

$$T_X^{-1} \left[\begin{smallmatrix} A & 0 \\ 0 & A^T \end{smallmatrix}\right] T_X = \left[\begin{smallmatrix} A & 0 \\ 0 & A^T \end{smallmatrix}\right] . \tag{7}$$

Defining $\mathcal{X} = \{X \in \mathbb{R}\text{-Sym}(n) \mid T_X \text{ satisfies (7)}\}$, it is easy to see that $\mathcal{X}$ is just the subspace $\ker \mathcal{T}_A$, where $\mathcal{T}_A$ is the operator considered in Proposition 11. Now insert a similarity by any such T_X into H_G to define

$$H_{G,X} = S\, T_X^{-1} \left[\begin{smallmatrix} 0 & F \\ G & 0 \end{smallmatrix}\right] T_X\, S^{-1}. \tag{8}$$

Clearly $H_{G,X}$ is Hamiltonian and $H_{G,X}^2 = W$ for every $(G, X) \in \mathcal{G} \times \mathcal{X}$. To see that these matrices $H_{G,X}$ are all distinct, let (G_1, X_1) and (G_2, X_2) be ordered pairs from $\mathcal{G} \times \mathcal{X}$. Then

$$\begin{aligned} H_{G_1,X_1} = H_{G_2,X_2} \quad &\Longrightarrow \quad S\, T_{X_1}^{-1} \left[\begin{smallmatrix} 0 & F_1 \\ G_1 & 0 \end{smallmatrix}\right] T_{X_1} S^{-1} = S\, T_{X_2}^{-1} \left[\begin{smallmatrix} 0 & F_2 \\ G_2 & 0 \end{smallmatrix}\right] T_{X_2} S^{-1} \\ &\Longrightarrow \quad T_{X_1}^{-1} \left[\begin{smallmatrix} 0 & F_1 \\ G_1 & 0 \end{smallmatrix}\right] T_{X_1} = T_{X_2}^{-1} \left[\begin{smallmatrix} 0 & F_2 \\ G_2 & 0 \end{smallmatrix}\right] T_{X_2} \\ &\Longrightarrow \quad \left[\begin{smallmatrix} -X_1 G_1 & F_1 - X_1 G_1 X_1 \\ G_1 & G_1 X_1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} -X_2 G_2 & F_2 - X_2 G_2 X_2 \\ G_2 & G_2 X_2 \end{smallmatrix}\right] \\ &\Longrightarrow \quad G_1 = G_2 \text{ and } X_1 = X_2, \text{ since } G_1 = G_2 \text{ is nonsingular} . \end{aligned}$$

Thus H_{G_1,X_1} and H_{G_2,X_2} are distinct whenever the pairs (G_1, X_1) and (G_2, X_2) are distinct, so $\sqrt[\mathcal{H}]{W}$ contains a family $\{H_{G,X}\}$ parametrized by $\mathcal{G} \times \mathcal{X}$. All that remains is to bound the sizes of $\mathcal{G}$ and $\mathcal{X}$.

Since $\mathcal{X} = \ker \mathcal{T}_A$, we have $\dim \mathcal{X} \geq n$ just from consideration of the dimensions of the domain and codomain of $\mathcal{T}_A$. By contrast the set $\mathcal{G}$ is not a subspace, so a lower bound on its dimension requires a bit more discussion. Observe that for any nonsingular symmetric G,

$$AG^{-1} = (AG^{-1})^T \iff AG^{-1} - G^{-1}A^T = 0 \iff G^{-1} \in \ker \mathcal{T}_A .$$

Thus $G \in \mathcal{G}$ iff $G^{-1} \in \mathrm{Inv}(\ker \mathcal{T}_A)$. Now by Proposition 11 $\mathrm{Inv}(\ker \mathcal{T}_A)$ is a submanifold with the same dimension as $\ker \mathcal{T}_A$. But matrix inversion is a diffeomorphism of $GL_n(\mathbb{R})$ which maps $\mathcal{G}$ bijectively to $\mathrm{Inv}(\ker \mathcal{T}_A)$, so $\mathcal{G}$ must also be a submanifold with $\dim \mathcal{G} = \dim \mathrm{Inv}(\ker \mathcal{T}_A) = \dim \ker \mathcal{T}_A$. Putting this all together, we have

$$\dim \sqrt[\mathcal{H}]{W} \;\geq\; \dim(\mathcal{G} \times \mathcal{X}) \;=\; \dim \mathcal{G} + \dim \mathcal{X} \;=\; 2 \dim(\ker \mathcal{T}_A) \;\geq\; 2n . \quad \square$$

The family $\{H_{G,X}\}$ of Hamiltonian square roots constructed in Theorem 3 does not always have dimension $2n$. In fact, since $\dim \ker \mathcal{T}_A$ can be much larger than n, it is possible for $\dim \sqrt[\mathcal{H}]{W}$ to be *much* larger than $2n$. The most extreme example of this occurs for $W = I_{2n}$, where $\dim \ker \mathcal{T}_{I_n} = \frac{1}{2}(n^2 + n)$ so that $\dim \sqrt[\mathcal{H}]{I_{2n}} \geq n^2 + n$. However, it is much more typical that the lower bound $\dim \sqrt[\mathcal{H}]{W} = 2n$ is actually attained. The next proposition makes this precise, using a standard argument from differential topology to show that $\sqrt[\mathcal{H}]{W}$ is exactly $2n$–dimensional for all but a measure zero subset of exceptional cases.

Proposition 12 *For almost all $2n \times 2n$ real skew-Hamiltonian matrices W, i.e. for all but a measure zero subset of $\mathcal{W}$, the set $\sqrt[\mathcal{H}]{W}$ is a smooth $2n$–dimensional submanifold of $\mathcal{H}$.*

Proof: Consider the squaring map

$$\begin{aligned} f : \mathcal{H} &\longrightarrow \mathcal{W} \\ H &\mapsto H^2. \end{aligned}$$

Clearly f is smooth, and the preimages $f^{-1}(W)$ are exactly the square root sets $\sqrt[\mathcal{H}]{W}$. Now for smooth maps, the Preimage Theorem [11] says that any nonempty preimage of a regular value is a smooth *submanifold* of the domain, and the dimension of this submanifold is the difference of the dimensions of the domain and codomain.[3] But the squaring map is onto (by Theorem 2), so every preimage is nonempty. And by Sard's Theorem [11], almost every point in the codomain of a smooth map is a regular value. Thus we see that for almost every skew-Hamiltonian matrix W, the set $f^{-1}(W) = \sqrt[\mathcal{H}]{W}$ is a submanifold of $\mathcal{H}$ with

$$\dim \sqrt[\mathcal{H}]{W} \;=\; \dim \mathcal{H} - \dim \mathcal{W} \;=\; (2n^2+n) - (2n^2-n) \;=\; 2n. \quad \square$$

REMARKS

1. Sard's theorem is completely nonconstructive, and in general gives no information about which values of a smooth map are regular. However, the squaring map f is simple enough that it is possible to explicitly characterize its regular values, and thus give an explicit sufficient condition for $\sqrt[\mathcal{H}]{W}$ to be a $2n$–dimensional submanifold. The first step toward achieving this is to describe the set of regular *points* of f, i.e. to find those $H \in \mathcal{H}$ where the Fréchet derivative $(df)_H$ is onto. But $(df)_H$ is precisely the map $\mathrm{Syl}(H, -H) : \mathcal{H} \to \mathcal{W}$, so we are back to the problem of deciding when certain Sylvester operators are onto. By an appropriate change of coordinates (as in the Appendix) one can transform this problem to the equivalent question of the surjectivity of $\mathrm{Syl}(H, H^T) : \mathbb{R}\text{-Sym} \to \mathbb{R}\text{-Skew}$, exactly the situation considered in Proposition 5. Thus we may conclude that $H \in \mathcal{H}$ is a regular point of f iff H is nonderogatory. Now the regular *values* of f are by definition the matrices $W \in \mathcal{W}$ such that every $H \in f^{-1}(W)$ is a regular point, so one might guess the regular values to be exactly those $W \in \mathcal{W}$ that are squares of nonderogatory $H \in \mathcal{H}$, equivalently those $W \in \mathcal{W}$ with the minimal number (two) of Jordan blocks for each eigenvalue. This is almost but not quite correct. It is possible to show that $W \in \mathcal{W}$ is a regular value of f iff W has exactly two Jordan blocks for each eigenvalue, and the multiplicity of the eigenvalue zero is not two.

2. The squaring map f is not just smooth, it's a quadratic polynomial map in the entries of H. Thus every preimage $f^{-1}(W) = \sqrt[\mathcal{H}]{W}$ is an *algebraic variety* with $\dim \geq 2n$. Proposition 12 says that most of these varieties are actually smooth submanifolds with $\dim = 2n$.

3. The existence of the family $\{H_{G,X} \mid (G,X) \in \mathcal{G} \times \mathcal{X}\}$ contained in any $\sqrt[\mathcal{H}]{W}$ enables us to conclude that every $\sqrt[\mathcal{H}]{W}$ is *unbounded.* To see this, observe that $\mathcal{G}$ is unbounded, since $G \in \mathcal{G} \Rightarrow kG \in \mathcal{G}$ for all $k \neq 0$. But similarity by any fixed $S \in \mathcal{S}$ is a nonsingular linear operator on $\mathcal{H}$, so $\{H_{G,0}\} = S\{ \left[\begin{smallmatrix} 0 & AG^{-1} \\ G & 0 \end{smallmatrix}\right] \mid G \in \mathcal{G}\} S^{-1}$ must also be unbounded.

[3]The Preimage Theorem is also known as the Regular Value Theorem.

4. Matrices in $\sqrt[\mathcal{H}]{W}$ need not be symplectically similar, or even have the same Jordan form. Examples of both situations can already be seen in the 2×2 case. $J = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$ and $-J$ are elements of $\sqrt[\mathcal{H}]{-I}$ that are not symplectically similar, although they do have the same Jordan form. Square root sets $\sqrt[\mathcal{H}]{W}$ can contain matrices with distinct Jordan forms only if W is singular; $N_2 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$ are both elements of $\sqrt[\mathcal{H}]{\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]}$. However, not every singular W exhibits this behavior; every element of $\sqrt[\mathcal{H}]{\left[\begin{smallmatrix} N_2 & 0 \\ 0 & N_2^T \end{smallmatrix}\right]}$ must have Jordan form N_4.

5. Using the results of this section it is possible to construct explicit $2n$-parameter families of Hamiltonian square roots, at least for small n. As an illustration, consider $W = \left[\begin{smallmatrix} N_2 & 0 \\ 0 & N_2^T \end{smallmatrix}\right]$. One easily finds for $A = N_2 = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$ that
$$\mathcal{X} = \ker \mathcal{T}_A = \left\{ \begin{bmatrix} c & d \\ d & 0 \end{bmatrix} \right\}, \quad \text{and} \quad \mathcal{G} = \left\{ \frac{1}{b^2} \begin{bmatrix} 0 & b \\ b & -a \end{bmatrix} \; : \; b \neq 0 \right\}.$$
Assembling these ingredients as in the proof of Theorem 3 yields the 4-parameter family
$$H_{G,X} = H(a,b,c,d) = \frac{1}{b^2} \begin{pmatrix} -bd & ad - bc & b^3 - 2bcd + ad^2 & -bd^2 \\ 0 & -bd & -bd^2 & 0 \\ 0 & b & bd & 0 \\ b & -a & bc - ad & bd \end{pmatrix}, \quad b \neq 0.$$
A direct computation shows that $[H(a,b,c,d)]^2 = W$ for every $a, b, c, d \in \mathbb{R}$ with $b \neq 0$.

6. It is not difficult to explicitly calculate $\sqrt[\mathcal{H}]{W}$ for any 2×2 skew-Hamiltonian matrix $W \in \mathcal{W}_{2\times2} = \{ \left[\begin{smallmatrix} k & 0 \\ 0 & k \end{smallmatrix}\right] \mid k \in \mathbb{R} \}$, and also to see how these square root sets fit together to partition the 3−dimensional space $\mathcal{H}_{2\times2}$ of all 2×2 Hamiltonian matrices. This is shown in Figure 1, which also nicely illustrates the first four remarks. In this figure we have identified $\mathcal{H}_{2\times2}$ with $\mathbb{R}^3$ using the isometry
$$\begin{aligned} \mathcal{H}_{2\times2} &\longrightarrow \mathbb{R}^3 \\ \left[\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right] &\mapsto \tfrac{\sqrt{2}}{2}(2a, b+c, b-c). \end{aligned}$$
The cone is exactly the set of all 2×2 nilpotent matrices, i.e. $\sqrt[\mathcal{H}]{\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]}$. Each $\sqrt[\mathcal{H}]{kI}$ with $k < 0$ is a two-sheeted hyperboloid intersecting the αJ−axis at $\pm\sqrt{|k|} \cdot J$. By contrast every $\sqrt[\mathcal{H}]{kI}$ with $k > 0$ is a hyperboloid of one sheet.
It is also interesting to note the relation of real similarity and symplectic similarity classes in $\mathcal{H}_{2\times2}$ to these square root sets. Although it is not true for larger Hamiltonian matrices, in $\mathcal{H}_{2\times2}$ every $\sqrt[\mathcal{H}]{W}$ is a finite union of similarity classes. For example, every hyperboloidal $\sqrt[\mathcal{H}]{W}$ is just the intersection of some real similarity class in $\mathbb{R}^{2\times2}$ with $\mathcal{H}_{2\times2}$. Every one-sheeted hyperboloid is also a symplectic similarity class; on the other hand, each sheet of a two-sheeted hyperboloid is a distinct symplectic similarity class. By contrast, the cone C of 2×2 nilpotents is the union of three symplectic similarity classes — the zero matrix $\mathbf{0}$ together with the two connected components C_1 and C_2 of $C \setminus \mathbf{0}$. Although matrices in C_1 are not symplectically similar to those in C_2, they are real similar, so that $C_1 \cup C_2$ constitutes a single real similarity class in $\mathcal{H}_{2\times2}$. Thus C is the union of two real similarity classes.

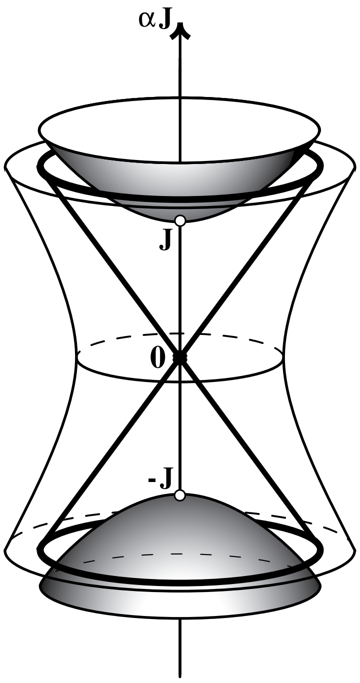

Figure 1: 2×2 Hamiltonian square root sets.

7. With the trivial exception of $0 \in \sqrt[\mathcal{H}]{0}$, *none* of the Hamiltonian square roots of *any* skew-Hamiltonian W is a polynomial in W. The basic reason for this is the eigenvalue structure of real Hamiltonian matrices; whenever $\lambda \in \mathbb{C}$ is an eigenvalue of $H \in \mathcal{H}$, then so is $-\lambda$ [5]. But distinct eigenvalues of a polynomial square root can not have the same square, so the only way for $H \in \mathcal{H}$ to be a polynomial square root of W is to have only the eigenvalue zero, i.e. H and W must be nilpotent. Now it is easy to see from the Jordan form that we can only have $H = p(W)$ and $H^2 = W$ if all of the Jordan blocks are 1×1, that is $H = W = 0$.

8. The characterization of the set $\mathcal{G}$ given in Theorem 3 and Proposition 11 constitutes an alternate proof of the two-symmetrics factorization theorem for real $n \times n$ matrices. An important feature of this proof is that it goes beyond the mere existence of the factorization to provide systematic (although not complete) information about the set of all such factorizations.

6 Complex Structured Square Roots

The notion of Hamiltonian and skew-Hamiltonian structure extends to complex matrices, so it is natural to consider the question of the existence of structured square roots in these complex classes. In this section we survey the various possibilities, beginning with some definitions and simple properties.

Recall from §1 and §2 that $\mathcal{S}$, $\mathcal{H}$, and $\mathcal{W}$ can be viewed as the automorphism group, Lie algebra, and Jordan algebra, respectively, of the bilinear form $b(x, y) = x^T J y$ defined on $\mathbb{R}^{2n}$

by the $2n \times 2n$ matrix $J = \left[\begin{smallmatrix} 0 & I \\ -I & 0 \end{smallmatrix}\right]$. This real-bilinear form $b(x,y)$ has unique extensions both to a complex-*sesquilinear* form, and to a complex-*bilinear* form on $\mathbb{C}^{2n}$. To each of these forms on $\mathbb{C}^{2n}$ there is an associated automorphism group, Lie algebra, and Jordan algebra of complex $2n \times 2n$ matrices. Thus there are two natural but distinct ways of extending the notions of real symplectic, Hamiltonian, and skew-Hamiltonian structure to complex matrices, leading to the following definitions:

$$\begin{aligned} \mathcal{S}_{\mathbb{C}}^{\natural} &= \{S \in \mathbb{C}^{2n\times 2n} \mid S^{\natural} J S = J\}\,, \\ \mathcal{H}_{\mathbb{C}}^{\natural} &= \{H \in \mathbb{C}^{2n\times 2n} \mid (JH)^{\natural} = JH\}\,, \\ \mathcal{W}_{\mathbb{C}}^{\natural} &= \{W \in \mathbb{C}^{2n\times 2n} \mid (JW)^{\natural} = -JW\}\,, \end{aligned}$$

where $(\ )^{\natural}$ denotes either transpose $(\ )^T$ or conjugate-transpose $(\ )^*$. The complex-bilinear extension of b leads to $\natural$ being T, while the sesquilinear extension of b results in $\natural$ being $*$. We remark that the use of conjugate-transpose to define complex symplectic, Hamiltonian, and skew-Hamiltonian matrices seems to be standard in control theory [6, 19], but using transpose appears to be more typical in the study of Lie groups, representation theory, and dynamical systems [9, 21, 27]. The terms J-orthogonal, J-symmetric and J-skew-symmetric for $\mathcal{S}_{\mathbb{C}}^T$, $\mathcal{H}_{\mathbb{C}}^T$ and $\mathcal{W}_{\mathbb{C}}^T$, and J-unitary, J-Hermitian and J-skew-Hermitian for $\mathcal{S}_{\mathbb{C}}^*$, $\mathcal{H}_{\mathbb{C}}^*$ and $\mathcal{W}_{\mathbb{C}}^*$ are also commonly used [5].

Characterizations of the block structure of matrices in $\mathcal{H}_{\mathbb{C}}^{\natural}$ and $\mathcal{W}_{\mathbb{C}}^{\natural}$ analogous to the ones given for $\mathcal{H}$ and $\mathcal{W}$ in §1 are easily obtained. For example, $W \in \mathcal{W}_{\mathbb{C}}^*$ iff $W = \left[\begin{smallmatrix} A & B \\ C & A^* \end{smallmatrix}\right]$, where $A \in \mathbb{C}^{n\times n}$ is arbitrary but $B, C \in \mathbb{C}^{n\times n}$ are both skew-Hermitian. The following simple properties of these classes of complex matrices are easy to check.

Lemma 2

(a) $\mathcal{H} \subset \mathcal{H}_{\mathbb{C}}^{\natural}$, *and* $\mathcal{W} \subset \mathcal{W}_{\mathbb{C}}^{\natural}$.

(b) $\mathcal{H}^2 \subseteq \mathcal{W}$, $\mathcal{W}^2 \subseteq \mathcal{W}$, *and* $(\mathcal{H}_{\mathbb{C}}^{\natural})^2 \subseteq \mathcal{W}_{\mathbb{C}}^{\natural}$, $(\mathcal{W}_{\mathbb{C}}^{\natural})^2 \subseteq \mathcal{W}_{\mathbb{C}}^{\natural}$.

(c) $S \in \mathcal{S}_{\mathbb{C}}^{\natural}$, $H \in \mathcal{H}_{\mathbb{C}}^{\natural}$, $W \in \mathcal{W}_{\mathbb{C}}^{\natural} \implies S^{-1}HS \in \mathcal{H}_{\mathbb{C}}^{\natural}$ *and* $S^{-1}WS \in \mathcal{W}_{\mathbb{C}}^{\natural}$.

(d) $\mathcal{H}_{\mathbb{C}}^T$ *and* $\mathcal{W}_{\mathbb{C}}^T$ *are complex subspaces of* $\mathbb{C}^{2n\times 2n}$, *while* $\mathcal{H}_{\mathbb{C}}^*$ *and* $\mathcal{W}_{\mathbb{C}}^*$ *are only* real *subspaces. But we have* $i \cdot \mathcal{H}_{\mathbb{C}}^* = \mathcal{W}_{\mathbb{C}}^*$, *and* $i \cdot \mathcal{W}_{\mathbb{C}}^* = \mathcal{H}_{\mathbb{C}}^*$.

From part (b) of this lemma it is clear that it only makes sense to look for structured square roots (i.e., square roots in $\mathcal{H}, \mathcal{W}, \mathcal{H}_{\mathbb{C}}^{\natural}$, or $\mathcal{W}_{\mathbb{C}}^{\natural}$) of matrices in $\mathcal{W}$ and $\mathcal{W}_{\mathbb{C}}^{\natural}$. Propositions 13 and 14 settle the existence question for all possible cases.

Proposition 13 *The following table summarizes the existence of real and complex Hamiltonian and skew-Hamiltonian square roots when* $\natural$ *is taken to be conjugate-transpose.*

	$\sqrt{W} \in \mathcal{H}$?	$\sqrt{W} \in \mathcal{H}_{\mathbb{C}}^*$?	$\sqrt{W} \in \mathcal{W}$?	$\sqrt{W} \in \mathcal{W}_{\mathbb{C}}^*$?
$W \in \mathcal{W}$	(a) Always	(b) Always	(c) Sometimes	(d) Always
$W \in \mathcal{W}_{\mathbb{C}}^*$	(e) Sometimes	(f) Sometimes	(g) Sometimes	(h) Sometimes

Proof:

(a) This is Theorem 2.

(b) Trivially true, since $\mathcal{H} \subset \mathcal{H}^*_{\mathbb{C}}$.

(c) The following three conditions on matrices in $\mathcal{W}$ are equivalent:

(i) $W \in \mathcal{W}$ has a $U \in \mathcal{W}$ such that $U^2 = W$.

(ii) $W \in \mathcal{W}$ is (real) symplectically similar to *some* block-diagonal matrix $\left[\begin{smallmatrix} A & 0 \\ 0 & A^T \end{smallmatrix}\right]$ such that $A \in \mathbb{R}^{n \times n}$ has a *real* square root.[4]

(iii) For *every* block-diagonal matrix $\left[\begin{smallmatrix} A & 0 \\ 0 & A^T \end{smallmatrix}\right]$ that is (real) symplectically similar to $W \in \mathcal{W}$, the matrix $A \in \mathbb{R}^{n \times n}$ has a *real* square root.

(<u>i ⇒ ii</u>): By Theorem 1 we may symplectically block-diagonalize U, so that $U = S \left[\begin{smallmatrix} B & 0 \\ 0 & B^T \end{smallmatrix}\right] S^{-1}$ with $B \in \mathbb{R}^{n \times n}$. Then

$$W = U^2 = S \left[\begin{smallmatrix} B^2 & 0 \\ 0 & B^{2T} \end{smallmatrix}\right] S^{-1} = S \left[\begin{smallmatrix} A & 0 \\ 0 & A^T \end{smallmatrix}\right] S^{-1}, \text{ with } A = B^2.$$

(<u>ii ⇒ iii</u>): Suppose W is symplectically similar to $\left[\begin{smallmatrix} A & 0 \\ 0 & A^T \end{smallmatrix}\right]$ with $A = B^2$ for some $B \in \mathbb{R}^{n \times n}$, and $\left[\begin{smallmatrix} \widehat{A} & 0 \\ 0 & \widehat{A}^T \end{smallmatrix}\right]$ is any other block-diagonal matrix in $\mathcal{W}$ symplectically similar to W. Now both $\left[\begin{smallmatrix} A & 0 \\ 0 & A^T \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} \widehat{A} & 0 \\ 0 & \widehat{A}^T \end{smallmatrix}\right]$ can be brought into "skew-Hamiltonian Jordan form" via similarity with block-diagonal symplectics of the form $\left[\begin{smallmatrix} V & 0 \\ 0 & V^{-T} \end{smallmatrix}\right]$. But since the skew-Hamiltonian Jordan form of W is essentially unique, we can conclude that A and $\widehat{A}$ must be (real) similar to each other. Thus $\widehat{A}$ also has a real square root.

(<u>iii ⇒ i</u>): By Theorem 1 we know that $W = S \left[\begin{smallmatrix} A & 0 \\ 0 & A^T \end{smallmatrix}\right] S^{-1}$ for some $S \in \mathcal{S}$ and $A \in \mathbb{R}^{n \times n}$, and from condition (iii) we have $A = B^2$ for some $B \in \mathbb{R}^{n \times n}$. Thus $W = U^2$ for $U = S \left[\begin{smallmatrix} B & 0 \\ 0 & B^T \end{smallmatrix}\right] S^{-1} \in \mathcal{W}$. □

Condition (iii) implies that $W = \left[\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$ is a matrix in $\mathcal{W}$ with no square root in $\mathcal{W}$.

(d) There are two simple ways to see this:

- First construct some $H \in \sqrt[\mathcal{H}]{W}$ as in Proposition 1. That is, let $H = S \left[\begin{smallmatrix} 0 & F \\ G & 0 \end{smallmatrix}\right] S^{-1} \in \mathcal{H}$ be such that $H^2 = W$. Then $\widetilde{W} = S \left[\begin{smallmatrix} 0 & iF \\ -iG & 0 \end{smallmatrix}\right] S^{-1} \in \mathcal{W}^*_{\mathbb{C}}$, and $\widetilde{W}^2 = W$.
- Alternately, use Lemma 2d. Pick any $\widetilde{H} \in \sqrt[\mathcal{H}]{-W}$. Then $\widetilde{W} = i \cdot \widetilde{H}$ is in $\mathcal{W}^*_{\mathbb{C}}$ and $\widetilde{W}^2 = W$.

(e) Clearly this is only true for $W \in \mathcal{W} \subset \mathcal{W}^*_{\mathbb{C}}$.

---

[4]Conditions for the existence of real square roots may be found in [13] and [17].

(f) There are many matrices in $\mathcal{W}_{\mathbb{C}}^{*}$ with no square root in $\mathcal{H}_{\mathbb{C}}^{*}$. This fact can be established by examining the possible arrangements in the complex plane of the eigenvalues of matrices in $\mathcal{W}_{\mathbb{C}}^{*}$ and $\mathcal{H}_{\mathbb{C}}^{*}$. It is well known that the spectra of matrices in $\mathcal{H}_{\mathbb{C}}^{*}$ possess a reflection symmetry [5, 6]; whenever $\lambda \in \mathbb{C}$ is an eigenvalue of $H \in \mathcal{H}_{\mathbb{C}}^{*}$, then so is $-\overline{\lambda}$, and both have the same multiplicity (indeed even the same Jordan structure). A general $H \in \mathcal{H}_{\mathbb{C}}^{*}$, then, has an even number of eigenvalues (counting multiplicity) grouped in pairs symmetric with respect to the imaginary axis, and the rest of its eigenvalues distributed arbitrarily on the imaginary axis. An analogous description for matrices in $\mathcal{W}_{\mathbb{C}}^{*}$ may also be given. Any $W \in \mathcal{W}_{\mathbb{C}}^{*}$ can be expressed as $W = iH$ for some $H \in \mathcal{H}_{\mathbb{C}}^{*}$, so $\lambda(W) = i\lambda(H)$. Thus a general $W \in \mathcal{W}_{\mathbb{C}}^{*}$ has an even number of its eigenvalues grouped in complex-conjugate pairs, with the remaining ones spread out on the real axis without restriction. Now consider the square of any $H \in \mathcal{H}_{\mathbb{C}}^{*}$. From the above we see that the eigenvalues of H^2 are just like those of matrices in $\mathcal{W}_{\mathbb{C}}^{*}$ (i.e. either real or in complex-conjugate pairs) except for one additional restriction — any *positive* eigenvalue of H^2 must have *even* multiplicity. Thus the matrix $Z = \left[\begin{smallmatrix} 1 & 2i \\ -2i & 1 \end{smallmatrix}\right] \in \mathcal{W}_{\mathbb{C}}^{*}$, with simple eigenvalues 3 and -1, cannot be the square of any $H \in \mathcal{H}_{\mathbb{C}}^{*}$.

(g) Clearly a square root in $\mathcal{W}$ can exist only if W is real and the condition described in part (c) is satisfied.

(h) The same $Z = \left[\begin{smallmatrix} 1 & 2i \\ -2i & 1 \end{smallmatrix}\right]$ as used in part (f) also provides an example of a matrix in $\mathcal{W}_{\mathbb{C}}^{*}$ with no square root in $\mathcal{W}_{\mathbb{C}}^{*}$. To see why this is so, consider the square of a general $W \in \mathcal{W}_{\mathbb{C}}^{*}$. Since W can be written as $W = iH$ for some $H \in \mathcal{H}_{\mathbb{C}}^{*}$, we have $W^2 = -H^2$; thus any *negative* eigenvalue of W^2 must have *even* multiplicity. Consequently Z, with a simple eigenvalue at -1, cannot be the square of any $W \in \mathcal{W}_{\mathbb{C}}^{*}$. $\square$

The alert reader will have noticed that no complex analog of Theorem 1 for $\mathcal{W}_{\mathbb{C}}^{*}$ played any part in the discussion of Proposition 13. The reason is simple: no such result is true for general matrices in $\mathcal{W}_{\mathbb{C}}^{*}$. Consider, for example, the matrix $U = \left[\begin{smallmatrix} 1 & i \\ 0 & 1 \end{smallmatrix}\right] \in \mathcal{W}_{\mathbb{C}}^{*}$. Since U has Jordan form $\left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$, it cannot be brought in to the form $\left[\begin{smallmatrix} A & 0 \\ 0 & A^* \end{smallmatrix}\right]$ by any similarity, let alone by similarity with some matrix from $\mathcal{S}_{\mathbb{C}}^{*}$.

By contrast, the complex analog of Theorem 1 does hold for the class $\mathcal{W}_{\mathbb{C}}^{T}$. Indeed, with only minor changes the very same proof given in this paper for *real* skew-Hamiltonian matrices also yields the following theorem. The existence of this result for $\mathcal{W}_{\mathbb{C}}^{T}$ but not for $\mathcal{W}_{\mathbb{C}}^{*}$ accounts for much of the difference between Propositions 13 and 14.

Theorem 4 *For any $W \in \mathcal{W}_{\mathbb{C}}^{T}$ there exists an $S \in \mathcal{S}_{\mathbb{C}}^{T}$ such that*

$$S^{-1}WS = \begin{bmatrix} A & 0 \\ 0 & A^T \end{bmatrix},$$

where $A \in \mathbb{C}^{n \times n}$ is in Jordan canonical form. The matrix A is unique up to a permutation of Jordan blocks.

Proposition 14 *The following table summarizes the existence of real and complex Hamiltonian and skew-Hamiltonian square roots when $\natural$ is taken to be transpose.*

	$\sqrt{W} \in \mathcal{H}$?	$\sqrt{W} \in \mathcal{H}_{\mathbb{C}}^{T}$?	$\sqrt{W} \in \mathcal{W}$?	$\sqrt{W} \in \mathcal{W}_{\mathbb{C}}^{T}$?
$W \in \mathcal{W}$	(a) Always	(b) Always	(c) Sometimes	(d) Sometimes
$W \in \mathcal{W}_{\mathbb{C}}^{T}$	(e) Sometimes	(f) Always	(g) Sometimes	(h) Sometimes

Proof:

(a,c) These are the same as parts (a) and (c) of Proposition 13.

(b) Trivially true, since $\mathcal{H} \subset \mathcal{H}_{\mathbb{C}}^{T}$.

(d) This case is covered by the discussion in part (h) below.

(e) Clearly this is only true for $W \in \mathcal{W} \subset \mathcal{W}_{\mathbb{C}}^{T}$.

(f) The argument used in Proposition 1 and Theorem 2 to show that every *real* skew-Hamiltonian matrix has a *real* Hamiltonian square root works equally well here to show that every $W \in \mathcal{W}_{\mathbb{C}}^{T}$ has a square root in $\mathcal{H}_{\mathbb{C}}^{T}$. By Theorem 4, there is an $S \in \mathcal{S}_{\mathbb{C}}^{T}$ such that $S^{-1}WS = \left[\begin{smallmatrix} A & 0 \\ 0 & A^T \end{smallmatrix}\right]$, where $A \in \mathbb{C}^{n\times n}$. But any *complex* A can be factored as the product $A = FG$ of two complex-*symmetric* matrices $F, G \in \mathbb{C}^{n\times n}$ [4, 24], so that $\left[\begin{smallmatrix} 0 & F \\ G & 0 \end{smallmatrix}\right] \in \mathcal{H}_{\mathbb{C}}^{T}$ and $\left[\begin{smallmatrix} 0 & F \\ G & 0 \end{smallmatrix}\right]^2 = \left[\begin{smallmatrix} A & 0 \\ 0 & A^T \end{smallmatrix}\right]$. Thus $S\left[\begin{smallmatrix} 0 & F \\ G & 0 \end{smallmatrix}\right]S^{-1} \in \mathcal{H}_{\mathbb{C}}^{T}$ is a square root of W.

(g) Clearly a square root in $\mathcal{W}$ can exist only if W is real and the conditions described in part(c) of Proposition 13 are satisfied.

(h) Using Theorem 4, the argument in Proposition 13c can be trivially modified to show that the following three conditions on matrices in $\mathcal{W}_{\mathbb{C}}^{T}$ are equivalent:

  (i) $W \in \mathcal{W}_{\mathbb{C}}^{T}$ has a $U \in \mathcal{W}_{\mathbb{C}}^{T}$ such that $U^2 = W$.

  (ii) $W \in \mathcal{W}_{\mathbb{C}}^{T}$ is similar (via a matrix in $\mathcal{S}_{\mathbb{C}}^{T}$) to *some* block-diagonal matrix $\left[\begin{smallmatrix} A & 0 \\ 0 & A^T \end{smallmatrix}\right]$ such that $A \in \mathbb{C}^{n\times n}$ has a (complex) square root.[5]

  (iii) For *every* block-diagonal matrix $\left[\begin{smallmatrix} A & 0 \\ 0 & A^T \end{smallmatrix}\right]$ that is similar to $W \in \mathcal{W}_{\mathbb{C}}^{T}$ via some matrix in $\mathcal{S}_{\mathbb{C}}^{T}$, the matrix $A \in \mathbb{C}^{n\times n}$ has a (complex) square root.

The equivalence of these conditions implies that the 4×4 matrix $W = \begin{bmatrix} N_2 & 0 \\ 0 & {N_2}^T \end{bmatrix} \in \mathcal{W}$ (seen earlier in remarks 4 and 5 of §5) has no square root in $\mathcal{W}_{\mathbb{C}}^{T}$, although it does have infinitely many square roots in $\mathcal{W}_{\mathbb{C}}^{*}$.

7 Conclusions

This paper has addressed the theoretical aspects of the Hamiltonian/skew-Hamiltonian structured square root problem. We have settled the existence question — every real skew-Hamiltonian matrix has a real Hamiltonian square root. Furthermore, we have shown that for any $2n \times 2n$ real skew-Hamiltonian W, the set $\sqrt[\mathcal{H}]{W}$ of all real Hamiltonian square roots of W is an unbounded algebraic variety with dimension at least $2n$. In fact $\sqrt[\mathcal{H}]{W}$ is a smooth

[5]Conditions for the existence of complex square roots may be found in [7] and [17].

manifold of dimension exactly $2n$ for almost every W. The existence question for various types of complex structured square roots of complex Hamiltonian and skew-Hamiltonian matrices has also been resolved.

We emphasize the main technical result of this paper, which may be of significant independent interest: every real skew-Hamiltonian matrix may be brought into structured real Jordan canonical form via real symplectic similarity. It is natural to ask whether there is an analogous structured canonical form for real Hamiltonian matrices. This question has been recently settled [20], but the canonical form is considerably more complicated and is usually not block triangular.

Finally, the problem of finding good numerical methods to compute Hamiltonian square roots for general skew-Hamiltonian matrices remains open. Clearly a Schur-like method involving Van Loan's reduction, Sylvester equations, and matrix inversion can be developed (see Remark 2 following Proposition 2, and the proof of Theorem 3), but such a method can only be applied in the generic case. Alternatively, one might consider iterative methods as in [14]. Unfortunately, all current iterative methods compute only square roots that are polynomials in the original matrix, and no nonzero Hamiltonian square root of any $W \in \mathcal{W}$ is a polynomial in W. Consequently the outlook for finding any structure-preserving matrix iteration to compute Hamiltonian square roots appears less than promising. We are currently exploring ways to overcome these difficulties.

Acknowledgements We thank Nick Higham, Volker Mehrmann, Charlie Van Loan, and Yieh-Hei Wan for helpful discussions and comments on various aspects of this paper.

8 Appendix

In this appendix we present proofs for Propositions 4, 5 and 6 from §3.1 and for Proposition 11 from §5. In each of these proofs we make use of the following well-known technique of "change of coordinates" between Sylvester operators [17]. Let $A \in F^{k\times k}$ and $B \in F^{\ell\times\ell}$ be fixed but arbitrary. Then for any invertible matrices $U \in F^{k\times k}$ and $Z \in F^{\ell\times\ell}$ we have:

$$\begin{aligned} AX - XB = Y \quad &\Longleftrightarrow \quad U(AX - XB)Z = UYZ \\ &\Longleftrightarrow \quad (UAU^{-1})(UXZ) - (UXZ)(Z^{-1}BZ) = UYZ\,. \end{aligned}$$

In other words the following diagram commutes:

$$\begin{tikzcd} F^{k\times\ell} \arrow[rr, "\mathrm{Syl}(A,B)"] \arrow[d, "X\mapsto UXZ"', "\cong"] & & F^{k\times\ell} \arrow[d, "\cong"', "Y\mapsto UYZ"] \\ F^{k\times\ell} \arrow[rr, "\mathrm{Syl}(UAU^{-1},\, Z^{-1}BZ)"] & & F^{k\times\ell}\,. \end{tikzcd} \tag{9}$$

Thus the kernels and ranges of $\mathrm{Syl}(A,B)$ and $\mathrm{Syl}(UAU^{-1}, Z^{-1}BZ)$ are simply related:

$$\ker \mathrm{Syl}(UAU^{-1}, Z^{-1}BZ) = U(\ker \mathrm{Syl}(A,B))Z \tag{10}$$

$$\operatorname{range} \mathrm{Syl}(UAU^{-1}, Z^{-1}BZ) = U(\operatorname{range} \mathrm{Syl}(A,B))Z\,. \tag{11}$$

In particular, $\mathrm{Syl}(A,B)$ is onto iff $\mathrm{Syl}(UAU^{-1}, Z^{-1}BZ)$ is onto.

<u>Proof of Proposition 4:</u>
With $A = M_k(\lambda)$ and $B = M_\ell(\lambda)$ both Jordan blocks corresponding to $\lambda \in F$, first observe that $\mathrm{Syl}(A, B^T)$ and $\mathcal{N}_{k\ell} = \mathrm{Syl}(N_k, N_\ell^T)$ are the same operator, since for every $X \in F^{k\times\ell}$ we have

$$AX - XB^T = (\lambda I_k + N_k)X - X(\lambda I_\ell + N_\ell^T) = N_k X - X N_\ell^T.$$

To find the range of $\mathcal{N}_{k\ell}$, we use the fundamental relationship $\mathrm{range}(\mathcal{N}_{k\ell}) = (\ker \mathcal{N}_{k\ell}^*)^\perp$; here $\mathcal{N}_{k\ell}^*$ denotes the adjoint of $\mathcal{N}_{k\ell}$ with respect to the standard inner product on $F^{k\times\ell}$ defined by $\langle X, Y\rangle = \mathrm{trace}(XY^H)$. The computation

$$\langle LX, Y\rangle = \mathrm{trace}(LXY^H) = \mathrm{trace}(XY^H L) = \langle X, L^H Y\rangle$$

shows that the adjoint of the left-multiplication operator $\mathcal{L}\colon X \mapsto LX$ is $\mathcal{L}^*\colon X \mapsto L^H X$. Similarly one sees that the adjoint of the right-multiplication operator $\mathcal{R}: X \mapsto XR$ is $\mathcal{R}^*\colon X \mapsto XR^H$. Together these imply that $\mathcal{N}_{k\ell}^* = \mathrm{Syl}(N_k^T, N_\ell)$.

Now $\ker \mathrm{Syl}(N_k, N_\ell)$ is well-known [10, 17]; it is just the set of all Toeplitz matrices of the form $[0\;\;T]$ when $k \le \ell$, or $\left[\begin{smallmatrix} T \\ 0 \end{smallmatrix}\right]$ when $k \ge \ell$, where $T \in F^{d\times d}$ with $d = \min(k,\ell)$ is upper triangular. From this known result we can obtain $\ker \mathrm{Syl}(N_k^T, N_\ell)$ by a change of coordinates as in the discussion of (9) above. Letting $E_k = [e_k\, e_{k-1} \cdots e_2\, e_1]$ denote the $k\times k$ "exchange" matrix, we have $N_k^T = E_k N_k E_k^{-1}$, so the following diagram commutes:

$$\begin{tikzcd} F^{k\times\ell} \arrow[r, "\mathrm{Syl}(N_k,N_\ell)"] \arrow[d, "X\mapsto E_kX"', "\cong"] & F^{k\times\ell} \arrow[d, "\cong"', "Y\mapsto E_kY"] \\ F^{k\times\ell} \arrow[r, "\mathrm{Syl}(N_k^T,N_\ell)"] & F^{k\times\ell}\,. \end{tikzcd} \tag{12}$$

Thus $\ker \mathcal{N}_{k\ell}^* = \ker \mathrm{Syl}(N_k^T, N_\ell) = E_k \cdot \ker \mathrm{Syl}(N_k, N_\ell)$ is the set of all $k\times\ell$ Hankel (i.e., constant along anti-diagonals) matrices which are zero everywhere except possibly along the last d antidiagonals.

Finally, consider the Hankel matrices $H_i \in F^{k\times\ell}$ where H_i has ones along the $(k+\ell-i)^{th}$ antidiagonal and zeroes everywhere else. Clearly $\{H_i \mid 1 \le i \le d\}$ is a basis for $\ker \mathcal{N}_{k\ell}^*$. For a matrix to be orthogonal to H_i, the sum of its entries along the $(k+\ell-i)^{th}$ antidiagonal must be zero. Thus

$$(\ker \mathcal{N}_{k\ell}^*)^\perp = \{Y \in F^{k\times\ell} \mid \text{sum of entries along each of the last } d \text{ antidiagonals is zero}\}. \ \square$$

<u>Proof of Proposition 5:</u>
For Sylvester operators $\mathcal{T}_A\colon X \mapsto AX - XA^T$, the appropriate coordinate changes are given as follows. Let $U \in F^{n\times n}$ be any invertible matrix, and $\widehat{A} = UAU^{-1}$. Then the following

diagram commutes:

$$
\begin{array}{ccc}
\text{F-Sym}(n) & \xrightarrow{\mathcal{T}_A} & \text{F-Skew}(n) \\
{\scriptstyle X\mapsto UXU^T}\Big\downarrow \cong & & \cong \Big\downarrow {\scriptstyle Y\mapsto UYU^T} \\
\text{F-Sym}(n) & \xrightarrow{\mathcal{T}_{\widehat{A}}} & \text{F-Skew}(n)\,.
\end{array}
\tag{13}
$$

Thus $\mathcal{T}_A$ is onto iff $\mathcal{T}_{\widehat{A}}$ is onto. Without loss of generality, then, we may assume that A is in any convenient normal form in $F^{n\times n}$.

Beginning with the complex case ($F = \mathbb{C}$), suppose that A is in Jordan canonical form, writing $A = A_{11} \oplus A_{22} \oplus \ldots \oplus A_{mm}$ where each A_{ii} is an $n_i \times n_i$ Jordan block $M_{n_i}(\lambda_i)$. Partition $X \in \mathbb{C}\text{-Sym}(n)$ into blocks X_{ij} conformally with A. The block-diagonal nature of A means that the operator $\mathcal{T}_A$ may be treated blockwise:

$$(\mathcal{T}_A X)_{ij} = (AX - XA^T)_{ij} = A_{ii}X_{ij} - X_{ij}A_{jj}^T\,,$$

so the ij^{th}-block of $\mathcal{T}_A X$ depends only on the ij^{th}-block of X. Consequently we may define the block operators $\mathcal{T}_{ij}\colon X_{ij} \mapsto A_{ii}X_{ij} - X_{ij}A_{jj}^T$ for $1 \le i, j \le m$, and observe that $\mathcal{T}_A$ is onto iff every $\mathcal{T}_{ij}$ is onto.

It is important to note a subtle difference between $\mathcal{T}_{ij}$ with $i \neq j$ and $\mathcal{T}_{ii}$. Because X is symmetric, the diagonal-block operators $\mathcal{T}_{ii}$ must be regarded as maps $\mathbb{C}\text{-Sym}(n_i) \to \mathbb{C}\text{-Skew}(n_i)$, whereas the off-diagonal-block operators $\mathcal{T}_{ij}$ $(i \neq j)$ are maps $\mathbb{C}^{n_i\times n_j} \to \mathbb{C}^{n_i\times n_j}$. These differences in domain and codomain are crucial to correctly judging whether each $\mathcal{T}_{ij}$ (and hence $\mathcal{T}_A$) is onto or not.

By Proposition 3, the off-diagonal-block operators $\mathcal{T}_{ij} = \mathrm{Syl}(A_{ii}, A_{jj}^T)$ are onto iff $\lambda_i \neq \lambda_j$. By contrast, the diagonal-block operators $\mathcal{T}_{ii}$ are always onto. To see this, first observe that the sum of the entries along any antidiagonal of a skew-symmetric matrix is zero. Then by Proposition 4 we have $\mathbb{C}\text{-Skew}(n_i) \subseteq \operatorname{range}\mathrm{Syl}(A_{ii}, A_{ii}^T)$; that is, for any $Y \in \mathbb{C}\text{-Skew}(n_i)$ there is some $Z \in \mathbb{C}^{n_i\times n_i}$ such that

$$\mathrm{Syl}(A_{ii}, A_{ii}^T)(Z) = A_{ii}Z - ZA_{ii}^T = Y. \tag{14}$$

This Z, however, may not be in the domain of $\mathcal{T}_{ii}$. But taking transpose of both sides of (14) shows that $\mathrm{Syl}(A_{ii}, A_{ii}^T)(Z^T) = Y$, so with $U = \frac{1}{2}(Z + Z^T)$ we have $\mathrm{Syl}(A_{ii}, A_{ii}^T)(U) = Y$. Since $U \in \mathbb{C}\text{-Sym}(n_i)$, we have $\mathcal{T}_{ii}(U) = Y$, showing that $\mathcal{T}_{ii}$ is onto.

Thus we conclude that $\mathcal{T}_A$ is onto $\iff$ $\mathcal{T}_{ij}$ is onto for all i, j $\iff$ $\lambda_i \neq \lambda_j$ for all $i \neq j$, and the $F = \mathbb{C}$ case is proved.

The real case ($F = \mathbb{R}$) follows almost immediately from the complex case. For a real matrix A there are two operators, $\mathcal{T}_A : \mathbb{R}\text{-Sym}(n) \to \mathbb{R}\text{-Skew}(n)$ and $\mathcal{T}_A^{\mathbb{C}} : \mathbb{C}\text{-Sym}(n) \to \mathbb{C}\text{-Skew}(n)$, defined by the same formula $X \mapsto AX - XA^T$. The operator $\mathcal{T}_A^{\mathbb{C}}$ is the complexification of $\mathcal{T}_A$, so $\dim_{\mathbb{R}}(\operatorname{range}\mathcal{T}_A) = \dim_{\mathbb{C}}(\operatorname{range}\mathcal{T}_A^{\mathbb{C}})$. Hence $\mathcal{T}_A$ is onto $\iff$ $\mathcal{T}_A^{\mathbb{C}}$ is onto. $\square$

Proof of Proposition 6:

The strategy here is to apply a change of coordinates to the real operator $\mathrm{Syl}(A, B^T)$ so that A and B are brought into *complex* Jordan form, then use Propositions 3 and 4 to find the range of the resulting complex operator, and finally translate back into completely real terms using a relation analogous to (11). We begin by recalling the similarities that convert real Jordan blocks into complex Jordan form.

Consider the block-diagonal $\Phi_{2k} = \mathrm{diag}(\Phi_2, \Phi_2, \cdots, \Phi_2)_{2k\times 2k}$, where $\Phi_2 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 & 1 \\ -i & i \end{smallmatrix}\right]$ is as in Lemma 1, and the permutation $P_{2k} = [e_1\, e_3\, e_5\, \cdots\, e_{2k-1} \mid e_2\, e_4\, \cdots\, e_{2k}]$.[6] It is now straightforward to check that

$$P_{2k}^T(\Phi_{2k}^H\, A\, \Phi_{2k})P_{2k} = \widehat{A} = \begin{bmatrix} M_k & 0 \\ 0 & \overline{M}_k \end{bmatrix},$$

where $M_k = M_k(\lambda)$ is the $k\times k$ Jordan block for $\lambda = a+ib$. Similarly we have $P_{2\ell}^T(\Phi_{2\ell}^H\, B\, \Phi_{2\ell})P_{2\ell} = \widehat{B} = \mathrm{diag}(M_\ell, \overline{M}_\ell)$, so that

$$P_{2\ell}^T(\Phi_{2\ell}^H\, B^T\, \Phi_{2\ell})P_{2\ell} = \widehat{B}^H = \begin{bmatrix} \overline{M}_\ell^T & 0 \\ 0 & M_\ell^T \end{bmatrix}.$$

It is important to note that $\lambda = a + ib$ appears in the $(1,1)$ block in $\widehat{A}$ but in the $(2,2)$ block in $\widehat{B}^H$.

More generally, let us consider the action of the "similarities" $\Phi_{2k}^H\, X\, \Phi_{2\ell}$ and $P_{2k}^T\, Y\, P_{2\ell}$ on arbitrary $2k\times 2\ell$ matrices X and Y. With X partitioned into 2×2 blocks X_{ij} $(1 \le i \le k, 1 \le j \le \ell)$, it is easy to see that the ij^{th} 2×2 block of $\Phi_{2k}^H\, X\, \Phi_{2\ell}$ is just $\Phi_2^H\, X_{ij}\, \Phi_2$. Thus any *real* block X_{ij} will be transformed into a 2×2 complex matrix $\left[\begin{smallmatrix} u & v \\ \bar{v} & \bar{u} \end{smallmatrix}\right] \in \mathcal{U}_2$ (see Lemma 1). Partition Y in the same way into 2×2 blocks $Y_{ij} = \left[\begin{smallmatrix} u_{ij} & v_{ij} \\ w_{ij} & z_{ij} \end{smallmatrix}\right]$. Then $P_{2k}^T\, Y\, P_{2\ell}$ is the block matrix $\left[\begin{smallmatrix} U & V \\ W & Z \end{smallmatrix}\right]$, where the $k\times\ell$ blocks U, V, W and Z are assembled entrywise from the Y_{ij} blocks via

$$\begin{aligned} U_{ij} &= (Y_{ij})_{11} = u_{ij} \qquad & V_{ij} &= (Y_{ij})_{12} = v_{ij} \\ W_{ij} &= (Y_{ij})_{21} = w_{ij} \qquad & Z_{ij} &= (Y_{ij})_{22} = z_{ij}\,. \end{aligned}$$

Now put these two maps together, defining $\Psi_{2k} = \Phi_{2k}P_{2k}$ and letting $\mathcal{U}_{2k\times 2\ell} \subseteq \mathbb{C}^{2k\times 2\ell}$ denote the set of all complex matrices of the form $\left[\begin{smallmatrix} U & V \\ \overline{V} & \overline{U} \end{smallmatrix}\right]$, where $U, V \in \mathbb{C}^{k\times\ell}$ are arbitrary. We see that the map

$$\begin{gathered} \mathbb{R}^{2k\times 2\ell} \xrightarrow{\cong} \mathcal{U}_{2k\times 2\ell} \\ X \mapsto \Psi_{2k}^H\, X\, \Psi_{2\ell} \end{gathered}$$

is a (real) linear isomorphism. Hence the change in coordinates in $\mathrm{Syl}(A, B^T)$ which takes A to $\widehat{A}$ and B^T to $\widehat{B}^H$ gives us the commutative diagram:

$$\begin{tikzcd} \mathbb{R}^{2k\times 2\ell} \arrow[r, "\mathrm{Syl}(A,B^T)"] \arrow[d, "X\mapsto\Psi_{2k}^H X \Psi_{2\ell}"', "\cong"] & \mathbb{R}^{2k\times 2\ell} \arrow[d, "Y\mapsto\Psi_{2k}^H Y \Psi_{2\ell}", "\cong"'] \\ \mathcal{U}_{2k\times 2\ell} \arrow[r, "\mathrm{Syl}(\widehat{A},\widehat{B}^H)"] & \mathcal{U}_{2k\times 2\ell}\,. \end{tikzcd} \tag{15}$$

[6]Note that the inverse of P_{2k} is the "perfect shuffle" permutation, known to American magicians as the "faro shuffle" and to English magicians as the "weave shuffle".

Thus

$$\operatorname{range}\operatorname{Syl}(A,\ B^T) = \Psi_{2k}(\operatorname{range}\operatorname{Syl}(\widehat{A},\ \widehat{B}^H))\Psi_{2\ell}^H. \tag{16}$$

Next we compute the range of $\operatorname{Syl}(\widehat{A},\ \widehat{B}^H)$. The block-diagonal nature of $\widehat{A}$ and $\widehat{B}^H$ means that $\operatorname{Syl}(\widehat{A},\ \widehat{B}^H)$ may be treated blockwise, just as $\mathcal{T}_A$ was in the proof of Proposition 5. We have

$$\begin{aligned}
\operatorname{Syl}(\widehat{A},\ \widehat{B}^H)\begin{bmatrix} U & V \\ \overline{V} & \overline{U} \end{bmatrix} &= \begin{bmatrix} M_k & 0 \\ 0 & \overline{M}_k \end{bmatrix}\begin{bmatrix} U & V \\ \overline{V} & \overline{U} \end{bmatrix} - \begin{bmatrix} U & V \\ \overline{V} & \overline{U} \end{bmatrix}\begin{bmatrix} \overline{M}_\ell^T & 0 \\ 0 & M_\ell^T \end{bmatrix} \\
&= \begin{bmatrix} M_k U - U\overline{M}_\ell^T & M_k V - V M_\ell^T \\ \overline{M}_k\overline{V} - \overline{V}\,\overline{M}_\ell^T & \overline{M}_k\overline{U} - \overline{U} M_\ell^T \end{bmatrix} \\
&= \begin{bmatrix} \mathcal{T}_{11}(U) & \mathcal{T}_{12}(V) \\ \mathcal{T}_{21}(\overline{V}) & \mathcal{T}_{22}(\overline{U}) \end{bmatrix}.
\end{aligned}$$

Observe that since $\mathcal{T}_{22}(\overline{U}) = \overline{\mathcal{T}_{11}(U)}$ and $\mathcal{T}_{21}(\overline{V}) = \overline{\mathcal{T}_{12}(V)}$, it suffices to find the ranges of $\mathcal{T}_{11}$ and $\mathcal{T}_{12}$. By Proposition 3, $\mathcal{T}_{11}$ is onto; Proposition 4 gives us the range of $\mathcal{T}_{12}$. Thus we see that $\operatorname{range}\operatorname{Syl}(\widehat{A},\ \widehat{B}^H)$ is the set of all $\left[\begin{smallmatrix} U & V \\ \overline{V} & \overline{U} \end{smallmatrix}\right] \in \mathcal{U}_{2k\times 2\ell}$ such that the sum of the entries along each of the last d antidiagonals of $V \in \mathbb{C}^{k\times\ell}$ is zero.

Transforming this result using (16) and Lemma 1 gives us the desired characterization of $\operatorname{range}\operatorname{Syl}(A,\ B^T)$. $\square$

<u>Proof of Proposition 11</u>:
Let $\mathcal{T}_A^{\mathbb{C}}$ denote the complexification of $\mathcal{T}_A$, that is the map

$$\begin{aligned} \mathcal{T}_A^{\mathbb{C}}:\ \mathbb{C}\text{-Sym}(n) &\longrightarrow \mathbb{C}\text{-Skew}(n) \\ X &\mapsto AX - XA^T. \end{aligned}$$

We show first that $\ker\mathcal{T}_A^{\mathbb{C}}$ contains at least one invertible matrix, and from this deduce that $\ker\mathcal{T}_A$ must also have at least one invertible element. Then the desired conclusion will follow from basic properties of algebraic varieties.

To find an invertible element of $\ker\mathcal{T}_A^{\mathbb{C}}$, begin as in the proof of Proposition 5. Change coordinates from $\mathcal{T}_A^{\mathbb{C}}$ to $\mathcal{T}_{\widehat{A}} : \mathbb{C}\text{-Sym}(n) \to \mathbb{C}\text{-Skew}(n)$, where $\widehat{A}$ is the Jordan form of A, and then treat $\mathcal{T}_{\widehat{A}}$ blockwise. The diagonal-block operators $\mathcal{T}_{ii}$ are just $\operatorname{Syl}(N_{n_i}, N_{n_i}^T)$ restricted to $\mathbb{C}\text{-Sym}(n_i)$, so an argument like the one used to compute $\ker\mathcal{N}_{k\ell}^*$ in the proof of Proposition 4 (see diagram 12) shows that $\ker\mathcal{T}_{ii} = \ker\operatorname{Syl}(N_{n_i}, N_{n_i})\cdot E_{n_i}$, a certain set of Hankel matrices. For our purposes it suffices to observe that the invertible matrix E_{n_i} is in $\ker\mathcal{T}_{ii}$; thus $E = \operatorname{diag}(E_{n_1}, E_{n_2}, \cdots, E_{n_m})$ is an invertible element of $\ker\mathcal{T}_{\widehat{A}}$. Transforming E back into $\ker\mathcal{T}_A^{\mathbb{C}}$ via (10) yields an invertible element of $\ker\mathcal{T}_A^{\mathbb{C}}$.

Now let $\{M_1 \dots M_k\}$ be any fixed basis for $\ker\mathcal{T}_A$, and consider the polynomial

$$p(x_1, x_2, \dots x_k) = \det(x_1M_1 + x_2M_2 + \cdots + x_kM_k).$$

With $(x_1, \dots x_k) \in \mathbb{R}^k$, this polynomial p distinguishes the singular and invertible elements of $\ker\mathcal{T}_A$. But $\mathcal{T}_A^{\mathbb{C}}$ is the complexification of $\mathcal{T}_A$, so the real matrices $\{M_1 \dots M_k\}$ also form a basis for $\ker \mathcal{T}_A^{\mathbb{C}}$; thus with $(x_1, \ldots x_k) \in \mathbb{C}^k$, the same polynomial p distinguishes the singular from the invertible elements of $\ker \mathcal{T}_A^{\mathbb{C}}$. Now the coefficients of p are real (since each M_i is real), and for any polynomial p with real coefficients the following are equivalent:

(i) $p \equiv 0$ as a formal polynomial, i.e all the coefficients of p are zero,

(ii) $p \equiv 0$ as a function $\mathbb{R}^k \to \mathbb{R}$,

(iii) $p \equiv 0$ as a function $\mathbb{C}^k \to \mathbb{C}$.

The existence of an invertible element in $\ker \mathcal{T}_A^{\mathbb{C}}$ means that $p \not\equiv 0$ as a complex function $\mathbb{C}^k \to \mathbb{C}$. Therefore $p \not\equiv 0$ as a real function $\mathbb{R}^k \to \mathbb{R}$ either, so there must be some invertible element in $\ker \mathcal{T}_A$. Consequently the zero set of p in $\mathbb{R}^k$ (equivalently the set of singular matrices in $\ker \mathcal{T}_A$) is a *proper* algebraic subset, and hence a closed, nowhere dense set of measure zero. Thus $\mathrm{Inv}(\ker \mathcal{T}_A)$, the complement of the singular matrices in $\ker \mathcal{T}_A$, is open and dense in $\ker \mathcal{T}_A$. $\square$

References

[1] G. Alefeld and N. Schneider. On square roots of M-matrices. *Linear Algebra Appl.*, 42:119–132, 1982.

[2] R. Bhatia and P. Rosenthal. How and why to solve the operator equation $AX - XB = Y$. *Bull. London Math. Soc.*, 29:1–21, 1997.

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

[4] A.J. Bosch. Note on the factorization of a square matrix into two Hermitian or symmetric matrices. *SIAM Review*, 29:463–468, 1987.

[5] A. Bunse–Gerstner, R. Byers and V. Mehrmann. A chart of numerical methods for structured eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 13:419–453, 1992.

[6] A. Bunse–Gerstner and H. Faßbender. A Jacobi-like method for solving algebraic Riccati equations on parallel computers. *IEEE Trans. Automat. Contr.*, 42:1071–1084, 1997.

[7] G.W. Cross and P. Lancaster. Square roots of complex matrices. *Linear and Multilinear Algebra*, 1:289–293, 1974.

[8] H. Faßbender, D.S. Mackey and N. Mackey. Structure-preserving Jacobi methods for doubly-structured eigenproblems. *In preparation.*

[9] W. Fulton and J. Harris. *Representation Theory: A First Course.* Springer-Verlag, New York, 1991.

[10] F.R. Gantmacher. *The Theory of Matrices, vol. 1.* Chelsea, New York, 1977.

[11] V. Guillemin and A. Pollack. *Differential Topology.* Prentice-Hall, Englewood Cliffs, 1974.

[12] P.R. Halmos. *Finite-Dimensional Vector Spaces.* Springer-Verlag, New York, 1974.

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

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

[15] M. Hirsch and S. Smale. *Differential Equations, Dynamical Systems, and Linear Algebra.* Academic Press, San Diego. 1974.

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

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

[18] T. Laffey. Personal communication. June 1998.

[19] V. Mehrmann. *The Autonomous Linear Quadratic Control Problem: Theory and Numerical Solution.* Lecture Notes in Control and Information Sciences Vol.163, Springer-Verlag, Berlin/Heidelberg, 1991.

[20] V. Mehrmann and H. Xu. Canonical forms for Hamiltonian and symplectic matrices and pencils. Technical report SFB393/98-07, Technische Universität Chemnitz, Germany, March 1998.

[21] K.R. Meyer and G.R. Hall. *Introduction to Hamiltonian Dynamical Systems and the N-body Problem.* Springer-Verlag, New York, 1992.

[22] V.V. Prasolov. *Problems and Theorems in Linear Algebra.* American Math. Soc., Providence, 1994.

[23] G.W. Stewart and Ji-Guang Sun. *Matrix Perturbation Theory.* Academic Press, San Diego, 1990.

[24] O. Taussky. The role of symmetric matrices in the study of general matrices. *Linear Algebra Appl.*, 5:147–154, 1972.

[25] O. Taussky and H. Zassenhaus. On the similarity transformation between a matrix and its transpose. *Pacific J. Math.*, 7:893–896, 1959.

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

[27] V.S. Varadarajan. *Lie Groups, Lie Algebras, and Their Representations.* Springer-Verlag, New York, 1984.