Zentrum für Technomathematik

Fachbereich 3 – Mathematik und Informatik

Hamilton and Jacobi come full circle: Jacobi algorithms for structured Hamiltonian eigenproblems

Heike Faßbender **D. Steven Mackey**

Niloufer Mackey

Report 99–10

Berichte aus der Technomathematik

Report 99–10 **Oktober 1999**

Hamilton and Jacobi come full circle: Jacobi algorithms for structured Hamiltonian eigenproblems

Heike Faßbender* D. Steven Mackey† Niloufer Mackey‡

September 1999

Dedicated to the memory of P. J. Eberlein

Abstract

We develop Jacobi algorithms for solving the complete eigenproblem for Hamiltonian and skew-Hamiltonian matrices that are also symmetric or skew-symmetric. Based on the direct solution of 4×4, and in one case, 8×8 subproblems, these structure preserving algorithms produce a complete basis of symplectic orthogonal eigenvectors for each of the four classes of matrices under consideration. The key step in their construction is a quaternion characterization of the 4×4 symplectic orthogonal group, and the subspaces of 4×4 Hamiltonian, skew-Hamiltonian, symmetric and skew-symmetric matrices. In addition to preserving structure, the algorithms are inherently parallelizable, numerically stable, and show asymptotic quadratic convergence.

Key words. eigenvalues, eigenvectors, Hamiltonian, skew-Hamiltonian, symmetric, skew-symmetric, symplectic, Jacobi method, quaternion, tensor product, structure-preserving, parallelizable.

AMS subject classification. 65F15, 15A18, 15A21, 15A57, 15A69, 93B10, 93B40, 93B60

1 Introduction

... that in quaternion run Perpetual Circle ...

— *John Milton, Paradise Lost, Book V.*

To the student of physics, the names Hamilton and Jacobi are closely linked. In the early 1830's, Hamilton reformulated classical mechanics in order to emphasize and exploit analogies to problems in optics [26, 27]; this was eventually to play a key role in the development of quantum mechanics [29, 35]. A few years later Jacobi extended Hamilton's work in

*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. email: mackey@kzoo.edu.

‡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.

dynamics [31], resulting in what is now known as the Hamilton-Jacobi theory. Systems of differential equations arising from this formulation of mechanics are often referred to as Hamiltonian systems; linearizations of such systems lead to an important class of structured matrices — the Hamiltonian matrices.

Later in his career Hamilton made another fundamental but completely different discovery — the quaternions [28]. This was a turning point in the history of algebra, showing that non-commutative number systems can be both internally consistent and useful. Around the same time, in an unrelated development, Jacobi described a method to compute the eigenvalues and eigenvectors of symmetric matrices [32]. Recently it was shown that these two apparently unconnected discoveries of Hamilton and Jacobi have an unexpected meeting point [25, 40]. Using, in part, the connection between Hamilton's quaternions and rotations of $\mathbb{R}^3$ and $\mathbb{R}^4$, one can develop Jacobi-like algorithms for both the symmetric and skew-symmetric eigenproblems based on the solution of 4×4 subproblems.

In this paper we show that the early and later work of these contemporaries can also be tied together, so Hamilton and Jacobi now "come full circle": the quaternions are the key to developing structure-preserving Jacobi-like algorithms for special classes of Hamiltonian and skew-Hamiltonian matrices.

As is well-known, the eigenproblem for Hamiltonian matrices arises in a number of important applications, for example, in the solution of matrix Riccati equations from control theory. Hamiltonian matrices with additional structure also arise in practise: the total least squares problem with a symmetric constraint leads to a symmetric Hamiltonian eigenproblem [34]. Such matrices are also encountered in linear response theory [47] and computational chemistry. Symmetric skew-Hamiltonian matrices have applications in quantum mechanical problems with time reversal symmetry [15].

Algorithms that exploit and preserve the structure of their client matrices have several advantages over general methods. The constraint of preserving symmetry or skew-symmetry restricts the use of similarity transformations to the orthogonal variety, resulting in algorithms that are usually backward stable. Equally desirable is the invariance of eigenstructure in the presence of round-off error, as for example, when computing the stable invariant subspace of a Hamiltonian matrix: the eigenvalues in this case come in $(\lambda, -\overline{\lambda})$ pairs, and it is essential to preserve this pairing. Furthermore, better perturbation bounds and consequently improved error analyses are often possible when the client matrices are confined to a special class (see [2, 3, 7], for example). Finally, storage requirements can be appreciably lowered by simply using a truncated form of the matrix and its subsequent iterates.

The main goal of this paper is to develop a new class of Jacobi methods that preserve all the structure of their client matrices. Originally designed to work on symmetric matrices via plane rotations targeting 2×2 subproblems, Jacobi's method [32] has been adapted to several other classes of matrices (see for example, [9, 11, 17, 23, 25, 48]). Using Hamilton's quaternions, Jacobi algorithms for symmetric and skew-symmetric matrices based on 4×4 rotations were developed in [25, 40]. In the symmetric case, convergence of the method under general quasi-cyclic orderings was also established in [41].

We show here that the connection between 4×4 real matrices and quaternions can be further exploited. We begin in §4 by constructing convenient representations of the 4×4

symplectic orthogonal group as well as of Hamiltonian and skew-Hamiltonian matrices in the tensor square of the quaternion algebra. In §5 we show how to efficiently compute direct solutions to four types of 4×4 doubly-structured real eigenproblems: symmetric Hamiltonian, skew-symmetric Hamiltonian, symmetric skew-Hamiltonian, and skew-symmetric skew-Hamiltonian. This enables us to develop structure-preserving Jacobi algorithms in §7 for $2n \times 2n$ doubly-structured matrices belonging to these four classes. To provide the algorithms with an identifiable goal, canonical forms that can be realized using only symplectic orthogonal similarities are first introduced in §6.

Somewhat surprisingly, an algorithm for $2n \times 2n$ skew-symmetric skew-Hamiltonian matrices that is directly based on 4×4 subproblems *does not converge*. This is discussed in §7.3, where we also show how to get around this difficulty: we base the algorithm instead on an explicit solution of the 8×8 skew-symmetric skew-Hamiltonian eigenproblem, which can be computed with unexpected ease using generalized symplectic Givens rotations. Built using quaternions, these 4×4 Givens rotations may well be of independent interest. The existence and construction of higher dimensional analogues of such rotations is discussed in §7.3.1.

The algorithms developed here are all inherently parallelizable; since only symplectic orthogonal similarities are used, they are structure preserving as well as numerically stable. In addition to eigenvalues, our algorithms calculate a symplectic orthogonal basis for all the invariant subspaces. In §8 we present the results of numerical experiments showing that all four algorithms exhibit the asymptotic quadratic convergence typical of Jacobi methods.

2 Preliminaries

Let E, F, $G \in \mathbb{R}^{n\times n}$. 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 $F^t = F$ and $G^t = G$. Equivalently, one may characterize the set $\mathcal{H}(2n)$ of all $2n \times 2n$ Hamiltonian matrices by

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

where $J_{2n} = \left[\begin{smallmatrix} 0 & I_n \\ -I_n & 0 \end{smallmatrix}\right]$, and I_n is the $n \times n$ identity matrix. Complementary to $\mathcal{H}(2n)$ is the set

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

of all *skew-Hamiltonian* matrices. Matrices in $\mathcal{W}(2n)$ 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 $C^t = -C$.

An important way to exploit the structure of Hamiltonian and skew-Hamiltonian matrices and also preserve the symmetry of the spectrum, is to use only structure-preserving similarities. To that end consider the set $\mathrm{Sp}(2n)$ of real *symplectic* matrices defined by

$$\mathrm{Sp}(2n) = \{S \in \mathbb{R}^{2n\times 2n} \mid S^t J_{2n} S = J_{2n}\}. \tag{1}$$

It is well-known and easy to show from this definition that $\mathrm{Sp}(2n)$ forms a multiplicative group, and that symplectic similarities preserve Hamiltonian and skew-Hamiltonian structure: for any $S \in \mathrm{Sp}(2n)$, $H \in \mathcal{H}(2n) \Rightarrow S^{-1}HS \in \mathcal{H}(2n)$, and $W \in \mathcal{W}(2n) \Rightarrow S^{-1}WS \in \mathcal{W}(2n)$.

3 Structural Constraints

A variety of methods for computing the eigenvalues and invariant subspaces of Hamiltonian matrices have been described in the literature (for example see [2, 3, 8, 9, 10, 11, 36, 45, 49, 54]). In this section we examine some of the basic structural issues involved in designing Jacobi algorithms that are completely structure-preserving.

The first issue to consider is the appropriate set of similarities to use. Since orthogonal matrices are perfectly conditioned, and symplectic similarities preserve structure, symplectic orthogonal similarity transformations are ideal tools in algorithms for the numerical solution of real Hamiltonian and skew-Hamiltonian eigenproblems. The use of such transformations was first promoted by Paige and Van Loan [49], who established a Schur-like decomposition for Hamiltonian matrices that have no pure imaginary eigenvalues. In [37], Mehrmann, Xu and Lin extended this result to the general case.

Next we consider the restrictions on the pivots, if a Jacobi method is to be structure-preserving. In a cyclic Jacobi method designed for arbitrary matrices, *any* off-diagonal entry can be a pivot and thereby be targeted for annihilation. If $R(i,j)$ denotes a plane rotation corresponding to the pivot location (i,j), then $R(i,j)$ is the same as the identity matrix except for the entries r_{ii}, r_{ij}, r_{ji}, and r_{jj}, where we have

$$\begin{bmatrix} r_{ii} & r_{ij} \\ r_{ji} & r_{jj} \end{bmatrix} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}.$$

Here $|c|^2 + |s|^2 = 1$, with c, s chosen to annihilate the ijth entry of the client matrix.

Describing a structure-preserving Jacobi algorithm based on the solution of 2×2 subproblems, Byers [11] showed that such an approach presents serious difficulties when the $2n \times 2n$ matrix $H = \left[\begin{smallmatrix} E & F \\ G & -E^T \end{smallmatrix}\right]$ is Hamiltonian. Although $R(i,j)$ is always orthogonal, it is symplectic only if the pivot h_{ij} is one of the diagonal entries of the off-diagonal block F. Thus a similarity by a plane rotation $R(i,j)$ can be structure-preserving only if $j = n + i$. It is worth pointing out that if H were skew-Hamiltonian, then the difficulty would be even more acute: the diagonal entries of both $n \times n$ off-diagonal blocks are always zero! Thus we see that the goal of preserving the (skew-)Hamiltonian structure puts a severe restriction on the choice of pivots for 2×2 based algorithms. This restriction can be somewhat ameliorated by using "double plane rotations" of the form

$$S = \begin{bmatrix} R(i,j) & 0 \\ 0 & R(i,j) \end{bmatrix}$$

where $R(i,j) \in \mathbb{R}^{n\times n}$ is a plane rotation as described earlier. S is both symplectic and orthogonal, even though neither

$$\begin{bmatrix} I & 0 \\ 0 & R(i,j) \end{bmatrix} \quad \text{nor} \quad \begin{bmatrix} R(i,j) & 0 \\ 0 & I \end{bmatrix}$$

is symplectic. Using such an S allows us to choose any off-diagonal entry of the $n\times n$ block E as the pivot, but unfortunately, the off-diagonal entries of F and G remain forever inaccessible to direct attack by symplectic plane rotations. In Byers' Hamiltonian-Jacobi algorithm [11], these two kinds of symplectic-orthogonals are the only tools available. This severely constrains the pivot strategy and results in unacceptably slow convergence, as pointed out by Byers. To improve this situation without compromising structure, we have to find a way to make *every* entry part of a structured subproblem at least once every sweep. Since symplectic plane rotations cannot achieve this, one is forced to consider the next larger Hamiltonian subproblem, which is 4×4. This approach was recently used in [9] for complex Hamiltonian matrices, but the 4×4 subproblems were solved iteratively.

By contrast, in this paper we construct simple and direct ways to solve four types of real doubly-structured 4×4 eigenproblems in a completely structure-preserving way. This requires more than just an entry-by-entry view of matrices — we show that the quaternion representation of $\mathbb{R}^{4\times 4}$ provides the insight needed for this task, allowing us to develop Jacobi-like algorithms for the corresponding $2n \times 2n$ structured matrices. It is also worth pointing out that unlike the algorithm in [9], where complex arithmetic cannot be avoided even when the initial matrix is real, the computations in the methods presented here stay in the real field: the quaternions bring into focus real 4×4 symplectic orthogonal transformations that are otherwise obscured from view in the trackless jungles of $\mathbb{R}^{16}$. [1]

4 Quaternions and Structured 4×4 Matrices

... I start by looking at a 2×2 matrix.
Sometimes I look at a 4×4 matrix.
That's when things get out of control and too hard.

— *Paul Halmos*

We will see here and in the next section that many aspects of 4×4 real matrices come under our control when viewed through quaternion lenses. For the convenience of the reader we briefly review the basic properties of the real algebra of quaternions, and the relation of its tensor square to the algebra of 4×4 real matrices.

[1]For more on quaternions, elephant's trunks, and trackless jungles, see [14, p. 214] or [41, p. 152–3].

The quaternions, denoted by $\mathbb{H}$ (for Hamilton, their discoverer), form a four-dimensional vector space over $\mathbb{R}$ with basis $\{1, i, j, k\}$. The multiplication rules

$$i^2 = j^2 = k^2 = ijk = -1\,,$$

which imply $jk = -kj = i,\ ki = -ik = j,\ ij = -ji = k$, make $\mathbb{H}$ into an associative, but non-commutative division algebra over the reals. The typical quaternion is

$$q = q_0 + q_1 i + q_2 j + q_3 k, \qquad q_0, q_1, q_2, q_3 \in \mathbb{R}.$$

The real part of q is q_0 and the pure quaternion part is $q_1 i + q_2 j + q_3 k$. The *conjugate* of q is given by $\overline{q} = q_0 - q_1 i - q_2 j - q_3 k$, and the *norm* $|q|$ is defined as $|q|^2 = q_0^2 + q_1^2 + q_2^2 + q_3^2 = q\overline{q} = \overline{q}q$. The multiplicative inverse of any non-zero quaternion is hence $q^{-1} = \overline{q}/|q|^2$. As a vector space, $\mathbb{H}$ is identified with $\mathbb{R}^4$ in the usual way,

$$q_0 + q_1 i + q_2 j + q_3 k \longleftrightarrow (q_0, q_1, q_2, q_3)^t\,.$$

Similarly, the subspace $\mathbb{P}$ of pure quaternions can be identified with $\mathbb{R}^3$,

$$q_1 i + q_2 j + q_3 k \longleftrightarrow (q_1, q_2, q_3)^t.$$

Motivated by these vector space isomorphisms we will, when convenient, denote the elements $1, i, j, k$ of $\mathbb{H}$ by e_0, e_1, e_2, e_3, respectively. We will also make use of the standard decomposition,

$$\mathbb{H} \;=\; \mathrm{span}\{1\} \,\oplus\, \mathrm{span}\{i, j, k\} \;=\; \mathbb{R} \oplus \mathbb{P}. \tag{2}$$

Corresponding to any pair (p, q) of quaternions is a real linear transformation from $\mathbb{H}$ to $\mathbb{H}$ that maps a quaternion v to the quaternion $pv\overline{q}$. Let $\mu(p, q)$ denote the matrix that encodes this transformation in the standard basis $\{1, i, j, k\}$. Clearly μ defines a real bilinear map from the cartesian product $\mathbb{H} \times \mathbb{H}$ into $\mathbb{R}^{4\times 4}$. ¿From the basic properties of the tensor product [4] it follows that μ induces a unique *linear* map

$$\phi : \mathbb{H} \otimes \mathbb{H} \to \mathbb{R}^{4\times 4}$$

such that $\phi(p \otimes q) = \mu(p, q)$. It can be shown [6, 50] that ϕ is a bijection that preserves not only the vector space structure, but also the multiplicative structure, thus exhibiting the algebra isomorphism between $\mathbb{H} \otimes \mathbb{H}$ and $\mathbb{R}^{4\times 4}$.

¿From the definition of ϕ it is easy to verify that

$$\phi(p \otimes 1) = \begin{pmatrix} p_0 & -p_1 & -p_2 & -p_3 \\ p_1 & p_0 & -p_3 & p_2 \\ p_2 & p_3 & p_0 & -p_1 \\ p_3 & -p_2 & p_1 & p_0 \end{pmatrix}, \quad \phi(1 \otimes q) = \begin{pmatrix} q_0 & q_1 & q_2 & q_3 \\ -q_1 & q_0 & -q_3 & q_2 \\ -q_2 & q_3 & q_0 & -q_1 \\ -q_3 & -q_2 & q_1 & q_0 \end{pmatrix}. \tag{3}$$

Since the tensor multiplication rule $(a \otimes b)(a' \otimes b') = (aa' \otimes bb')$ implies

$$p \otimes q = (p \otimes 1)(1 \otimes q) = (1 \otimes q)(p \otimes 1)\,, \tag{4}$$

we immediately conclude that the matrices in (3) commute, and their product is $\phi(p \otimes q)$. Next, conjugation in $\mathbb{H} \otimes \mathbb{H}$ is determined by extending the rule

$$\overline{p \otimes q} = \overline{p} \otimes \overline{q} \qquad \forall p, q \in \mathbb{H} \tag{5}$$

linearly to all of $\mathbb{H} \otimes \mathbb{H}$. Examining the matrices in (3) one observes that $\phi(\overline{p} \otimes 1) = (\phi(p \otimes 1))^t$, and $\phi(1 \otimes \overline{q}) = (\phi(1 \otimes q))^t$. Thus we see that ϕ preserves more than the algebra structure: *conjugation in* $\mathbb{H} \otimes \mathbb{H}$ *corresponds, via* ϕ, *to transpose in* $\mathbb{R}^{4\times4}$.

By the usual abuse of notation, we will sometimes use $p \otimes q$ to stand for the matrix $\phi(p \otimes q)$, both to simplify notation and to emphasize the identification of $\mathbb{H} \otimes \mathbb{H}$ with $\mathbb{R}^{4\times4}$.

4.1 Symmetrics and skew-symmetrics

The equivalence between conjugation and transpose immediately implies that the sixteen 4×4 matrices corresponding to the standard basis $\mathcal{B} = \{1 \otimes 1,\ 1 \otimes i,\ \dots,\ k \otimes j,\ k \otimes k\}$ must all be orthogonal, and either symmetric or skew-symmetric. For example, $k \otimes j$ is its own conjugate, so $\phi(k \otimes j)$ must be symmetric; furthermore, $(k \otimes j)(\overline{k} \otimes \overline{j}) = k\overline{k} \otimes j\overline{j} = 1 \otimes 1$, hence $\phi(k \otimes j)$ must also be orthogonal. Table 1 shows which basis elements are symmetric and which are skew-symmetric.

$\otimes$	1	i	j	k
1	S	K	K	K
i	K	S	S	S
j	K	S	S	S
k	K	S	S	S

Table 1: 4×4 Symmetrics(S) and Skew-symmetrics(K)

This table also highlights the decomposition of $\mathbb{H} \otimes \mathbb{H}$ induced by the natural decomposition $\mathbb{H} = \mathbb{R} \oplus \mathbb{P}$. The subset $\{1 \otimes i,\ 1 \otimes j,\ 1 \otimes k,\ i \otimes 1,\ j \otimes 1,\ k \otimes 1\}$ forms a basis for the subspace $(\mathbb{R} \otimes \mathbb{P}) \oplus (\mathbb{P} \otimes \mathbb{R})$, which corresponds to the 6-dimensional subspace of real 4×4 skew-symmetric matrices. The remaining basis elements generate the complementary subspace $(\mathbb{R} \otimes \mathbb{R}) \oplus (\mathbb{P} \otimes \mathbb{P})$, corresponding to the 10-dimensional subspace of real 4×4 symmetric matrices.

Thus, as was shown in [25, 40], we get convenient quaternion representations of real 4×4 skew-symmetric and symmetric matrices:

Proposition 1

(a) $K \in \mathbb{R}^{4\times4}$ *is skew-symmetric* $\iff \exists\ p, q \in \mathbb{P}$ *such that* $K = p \otimes 1 + 1 \otimes q$.

(b) $S \in \mathbb{R}^{4\times4}$ *is symmetric* $\iff \exists\ p, q, r \in \mathbb{P},\ c \in \mathbb{R}$ *such that* $S = c\,(1 \otimes 1) + p \otimes i + q \otimes j + r \otimes k$.

Formulas for calculating these quaternion representations from the entries of K, S are given in [40, Eqns. (9)–(14)]. These formulas are linear in the matrix entries, and straightforward to derive from the matrix representation of the basis $\mathcal{B}$ given in Appendix A.

4.2 Hamiltonians and skew-Hamiltonians

We now prove that real 4×4 Hamiltonian and skew-Hamiltonian matrices also have natural representations in $\mathbb{H} \otimes \mathbb{H}$. First, observe that the bijection on $\mathbb{R}^{2n \times 2n}$ given by $A \mapsto J_{2n}A$, where $J_{2n} = \left[\begin{smallmatrix} 0 & I_n \\ -I_n & 0 \end{smallmatrix}\right]$, maps symmetric matrices to Hamiltonian matrices, and skew-symmetric matrices to skew-Hamiltonians. Thus premultiplication by the quaternion equivalent of J_4 in $\mathbb{H} \otimes \mathbb{H}$ will transform the basis $\mathcal{B}$ into a (possibly different) quaternion basis, made up exclusively of real 4×4 Hamiltonian and skew-Hamiltonian matrices. Now a simple calculation using Eqn. (3) shows that $\phi(1 \otimes j) = J_4$. This is indeed fortuitous, as up to sign, $\mathcal{B}$ is closed under multiplication. It is easy to check that premultiplication by $1 \otimes j$ permutes the elements of $\mathcal{B}$ (again up to sign) in a simple way — the first and third columns in Table 1 are interchanged as are the second and fourth columns. Thus the Hamiltonian and skew-Hamiltonian structure of $\mathcal{B}$ can be quickly deduced directly from the properties of the quaternion tensor algebra:

$\otimes$	1	i	j	k
1	W	W	H	W
i	H	H	W	H
j	H	H	W	H
k	H	H	W	H

Table 2: 4×4 Hamiltonians(H) and Skew-Hamiltonians(W)

Alternatively, one can use Appendix A to verify that each of the matrices in $\mathcal{B}$ has the structure specified in Table 2. This readily gives us quaternion characterizations of real 4×4 Hamiltonian and skew-Hamiltonian matrices:

Proposition 2

(a) $H \in \mathbb{R}^{4\times 4}$ *is Hamiltonian* $\iff \exists\, p,\, q,\, r \in \mathbb{P},\ b \in \mathbb{R}$ *such that* $H = b\,(1 \otimes j) + p \otimes 1 + q \otimes i + r \otimes k$.

(b) $W \in \mathbb{R}^{4\times 4}$ *is skew-Hamiltonian* $\iff \exists\, p \in \mathbb{P},\ b, c, d \in \mathbb{R}$ *such that* $W = b\,(1 \otimes 1) + p \otimes j + 1 \otimes (c\,i + d\,k)$.

Linear equations determining the quaternion parameters in terms of the entries of H, W can be readily found, in a manner analogous to the symmetric and skew-symmetric cases. For a 4×4 real Hamiltonian matrix $H = [h_{\ell m}] = b\,(1 \otimes j) + p \otimes 1 + q \otimes i + r \otimes k$, the pure quaternions p, q, r and the real scalar b are given by

$$b = \tfrac{1}{4}\,(h_{13} - h_{31} + h_{24} - h_{42})\,, \tag{6}$$

$$p_1 = \tfrac{1}{2}(h_{21} - h_{12}), \quad p_2 = \tfrac{1}{4}(h_{24} - h_{13} + h_{31} - h_{42}), \quad p_3 = \tfrac{1}{2}(h_{41} - h_{14}), \tag{7}$$

$$q_1 = \tfrac{1}{2}(h_{11} + h_{22}), \quad q_2 = \tfrac{1}{2}(h_{41} + h_{14}), \quad q_3 = \tfrac{1}{4}(h_{24} - h_{13} + h_{42} - h_{31}), \tag{8}$$

$$r_1 = \tfrac{1}{4}(h_{13} + h_{31} + h_{24} + h_{42}), \quad r_2 = -\tfrac{1}{2}(h_{21} + h_{12}), \quad r_3 = \tfrac{1}{2}(h_{11} - h_{22}). \tag{9}$$

The corresponding calculation for a 4×4 real skew-Hamiltonian matrix $W = [w_{\ell m}] = b\,(1 \otimes 1) + p \otimes j + 1 \otimes (c\,i + d\,k)$ yields even simpler equations for the pure quaternion p and the real scalars b, c, d:

$$p_1 = \tfrac{1}{2}(w_{32} - w_{14}), \quad p_2 = \tfrac{1}{2}(w_{11} - w_{22}), \quad p_3 = \tfrac{1}{2}(w_{12} + w_{21}), \tag{10}$$

$$b = \tfrac{1}{2}(w_{11} + w_{22}), \quad c = \tfrac{1}{2}(w_{12} - w_{21}), \quad d = \tfrac{1}{2}(w_{14} + w_{32}). \tag{11}$$

Thus we have shown that the quaternion basis $\mathcal{B}$ *simultaneously* provides a basis for two direct sum decompositions of 4×4 real matrices,

$$\begin{aligned} \{\text{Symmetrics}\} \;&\oplus\; \{\text{Skew-symmetrics}\} \\ \{\text{Hamiltonians}\} \;&\oplus\; \{\text{Skew-Hamiltonians}\}. \end{aligned}$$

Combining Tables 1 and 2 gives us convenient representations for the four classes of doubly-structured matrices under consideration.

Proposition 3

(a) $H \in \mathbb{R}^{4\times4}$ *is symmetric Hamiltonian* $\iff \exists\, q, r \in \mathbb{P}$ *such that* $H = q \otimes i + r \otimes k$.

(b) $H \in \mathbb{R}^{4\times4}$ *is skew-symmetric Hamiltonian* $\iff \exists\, p \in \mathbb{P}, b \in \mathbb{R}$ *such that* $H = b\,(1 \otimes j) + p \otimes 1$.

(c) $W \in \mathbb{R}^{4\times4}$ *is symmetric skew-Hamiltonian* $\iff \exists\, p \in \mathbb{P}, b \in \mathbb{R}$ *such that* $W = b\,(1 \otimes 1) + p \otimes j$.

(d) $W \in \mathbb{R}^{4\times4}$ *is skew-symmetric skew-Hamiltonian* $\iff \exists\, c, d \in \mathbb{R}$ *such that* $W = 1 \otimes (c\,i + d\,k)$

The quaternion parameters for these four classes are determined by specializing Eqns. (6)–(11). The additional symmetric or skew-symmetric structure considerably simplifies the four sets of formulae, which we include here for the sake of completeness, and to make the algorithms of §5 more transparent.

Symmetric Hamiltonian :

$$q_1 = \frac{1}{2}(h_{11} + h_{22}), \quad q_2 = h_{14}, \quad q_3 = \frac{1}{2}(h_{24} - h_{13}) \tag{12}$$

$$r_1 = \frac{1}{2}(h_{13} + h_{24}), \quad r_2 = -h_{12}, \quad r_3 = \frac{1}{2}(h_{11} - h_{22}). \tag{13}$$

Skew-symmetric Hamiltonian :

$$b = \frac{1}{2}(h_{13} + h_{24}), \quad p_1 = h_{21}, \quad p_2 = \frac{1}{2}(h_{31} - h_{42}), \quad p_3 = h_{41} \tag{14}$$

Symmetric skew-Hamiltonian :

$$b = \frac{1}{2}(w_{11} + w_{22}), \quad p_1 = -w_{14}, \quad p_2 = \frac{1}{2}(w_{11} - w_{22}), \quad p_3 = w_{12}. \tag{15}$$

Skew-symmetric skew-Hamiltonian :

$$c = w_{12}, \quad d = w_{14}. \tag{16}$$

4.3 Orthogonals

The connection between quaternions and rotations of $\mathbb{R}^3$ and $\mathbb{R}^4$ goes back to Hamilton and Cayley [12, 13, 28]. Briefly put, if u and v are unit quaternions (i.e., $u\overline{u} = 1 = v\overline{v}$), then $\phi(u \otimes v)$ is an orthogonal matrix. This follows directly from the equivalence of conjugation with transpose : $\phi(u \otimes v)(\phi(u \otimes v))^t = \phi(u \otimes v)\phi(\overline{u} \otimes \overline{v}) = \phi(u\overline{u} \otimes v\overline{v}) = \phi(1 \otimes 1) = I_4$. It can further be shown that $\det \phi(u \otimes v) = +1$, by using, for example, the continuity of the determinant and the connectedness of the unit quaternions. Thus $\phi(u \otimes v)$ is a rotation of $\mathbb{R}^4$. When $u = v$, $\phi(u \otimes u)$ can be interpreted as a rotation of $\mathbb{R}^3$. The first column of this matrix is the vector representation of the quaternion $u1\overline{u} = 1$, that is, e_0. Orthogonality now forces

$$\phi(u \otimes u) = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & & & \\ 0 & & U & \\ 0 & & & \end{array}\right), \quad \text{where} \quad U \in \mathrm{SO}(3) \tag{17}$$

which can be viewed as a rotation acting on $\mathbb{P} \cong \mathbb{R}^3$.

Another important fact is the converse — every element of SO(4) can be expressed as $\phi(u \otimes v)$ for some pair of unit quaternions u, v. Similarly, every element of SO(3) can be realized as $\phi(u \otimes u)$ for some unit quaternion u. In fact, there is a useful and direct relation between the co-ordinates of a unit quaternion $u = u_0 + u_1 i + u_2 j + u_3 k$ and the geometry of the associated 3-dimensional rotation $\phi(u \otimes u)$. The angle θ of this rotation is encoded in the real part, $u_0 = \cos(\theta/2)$, while the rotational axis is along the direction (u_1, u_2, u_3) determined by the pure quaternion part. For further details, see for example [16, 40, 51].

4.4 Symplectic Orthogonals

It follows from Eqn. (1) that an orthogonal matrix is symplectic if and only if it commutes with J_{2n}. In the 4×4 case, this translates to

$$(u \otimes v)\,(1 \otimes j) \;=\; (1 \otimes j)\,(u \otimes v)$$

where u, v are unit quaternions. Equivalently, we must have $u \otimes vj = u \otimes jv$, which can be rewritten as $u \otimes (vj - jv) = 0$. Since $\mathbb{H}$ is a division algebra, it follows that $vj - jv = 0$, in other words, v must commute with j. [2] This can only happen if $v \in \operatorname{span}\{1, j\}$. Thus a 4×4 matrix R is symplectic orthogonal if and only if $R = \phi(u \otimes v)$ where u, v are unit quaternions, and $v = c + dj$ for $c, d \in \mathbb{R}$. [3]

Symplectic orthogonal matrices can also be simply characterized in purely matrix terms. It is easy to check that a $2n \times 2n$ matrix commutes with J_{2n} if and only if it has the block structure $\left[\begin{smallmatrix} U & -V \\ V & U \end{smallmatrix}\right]$. So a matrix $R \in \mathbb{R}^{2n \times 2n}$ is symplectic orthogonal, i.e. $R \in \mathrm{SpO}(2n)$, if and only if R is an orthogonal matrix of the form $\left[\begin{smallmatrix} U & -V \\ V & U \end{smallmatrix}\right]$.

[2] If X, Y are division algebras, and $x \in X, y \in Y$, then $x \otimes y = 0$ if and only if $x = 0$ or $y = 0$.

[3] This characterization gives us a double cover of the group SpO(4) of 4×4 symplectic orthogonal matrices by $\mathrm{S}^3 \times \mathrm{S}^1$. In particular, this shows that as a manifold SpO(4) is four-dimensional.

4.5 Dictionary

The following table summarizes the quaternion representations of all the 4×4 structured matrices that we will need.

For $p,\, q,\, r \in \mathbb{P}$, and $b,\, c,\, d,\, f \in \mathbb{R}$	
Symmetric Skew-Symmetric	$b\,(1 \otimes 1) \,+\, p \otimes i \,+\, q \otimes j \,+\, r \otimes k$ $p \otimes 1 \,+\, 1 \otimes q$
Diagonal 2×2 block-diagonal	$b\,(1 \otimes 1) \,+\, c\,(i \otimes i) \,+\, d\,(j \otimes j) \,+\, f\,(k \otimes k)$ Diagonal $+\, b\,(i \otimes 1) \,+\, c\,(1 \otimes i) \,+\, d\,(j \otimes k) +$ $f\,(k \otimes j)$
Hamiltonian Symmetric Hamiltonian Skew-symmetric Hamiltonian	$b\,(1 \otimes j) \,+\, p \otimes 1 \,+\, q \otimes i \,+\, r \otimes k$ $q \otimes i \,+\, r \otimes k$ $b\,(1 \otimes j) \,+\, p \otimes 1$
Skew-Hamiltonian Symmetric skew-Hamiltonian Skew-symmetric skew-Hamiltonian	$b\,(1 \otimes 1) \,+\, p \otimes j \,+\, 1 \otimes (c\,i + d\,k)$ $b\,(1 \otimes 1) \,+\, p \otimes j$ $1 \otimes (ci + dk)$
Orthogonal Symplectic orthogonal	$u \otimes v, \quad \lvert u\rvert = \lvert v\rvert = 1,\, u,\, v \in \mathbb{H}$ $u \otimes v, \quad \lvert u\rvert = \lvert v\rvert = 1,\, u \in \mathbb{H},\, v = c + d\,j$

Table 3: Quaternion Dictionary

5 Doubly-Structured 4×4 Eigenproblems

We now turn our attention to solving the four doubly-structured algebraic eigenproblems in $\mathbb{R}^{4\times4}$ by translating them into 3-dimensional geometric problems in $\mathbb{H} \otimes \mathbb{H}$. We begin this task by examining the action of a symplectic orthogonal similarity from the perspective of the quaternion tensor.

5.1 Symplectic Orthogonal Similarities

A purely geometric interpretation of a similarity by an element of SO(4) was developed in [40] using the quaternion tensor square and properties of the map ϕ. Since this is a crucial step in solving 4×4 eigenproblems, we include a brief review of those ideas here for the convenience of the reader.

If x, y are unit quaternions, then the product

$$(x \otimes y)(p \otimes q)(\overline{x} \otimes \overline{y}) = (xp\overline{x}) \otimes (yq\overline{y}) \tag{18}$$

translates to an orthogonal similarity in $\mathbb{R}^{4\times4}$ acting on the matrix $\phi(p \otimes q)$. Moreover, the definition of ϕ shows that the quaternion product $xp\overline{x}$ can be viewed as the image of $p \in \mathbb{R}^4 \cong \mathbb{H}$ under the map $\phi(x \otimes x)$. By § 4.3, this merely rotates the pure quaternion part of p by the 3-dimensional rotation $\phi(x \otimes x)$. Similarly, the pure quaternion part of q is rotated by $\phi(y \otimes y) \in \mathrm{SO}(3)$. Since every element of $\mathbb{H} \otimes \mathbb{H}$ is a real linear combination of elements of the form $p \otimes q$, the effect of an orthogonal similarity by $\phi(x \otimes y)$ in the 16-dimensional space $\mathbb{R}^{4\times 4}$ can be reduced to the independent action of a pair of 3-dimensional rotations.

Preserving the Hamiltonian or skew-Hamiltonian structure restricts us to *symplectic* orthogonal similarities, so our choice of the vector y in $\phi(x \otimes y)$ is limited to the unit circle in the plane spanned by $\{1, j\}$. This means that the additional constraint of being symplectic puts no restriction on the 3-dimensional rotation $\phi(x \otimes x)$; but while we are free to choose the angle of the 3-dimensional rotation $\phi(y \otimes y)$, its axis must lie along $j = (0, 1, 0)$. Thus if $\phi(x \otimes y)$ is to be symplectic, then $\phi(y \otimes y)$ is necessarily a rotation of the $\{i, k\}$-plane.

Thus we see that the solution of 4×4 structured eigenproblems reduces to the construction of rotations of $\mathbb{P} \cong \mathbb{R}^3$ with a specified action. Indeed, the common recurring task in all the algorithms is to rotate a given pure quaternion a into alignment with either $\pm i$, $\pm j$, or $\pm k$. The following proposition, adapted from Hacon [25], gives a general formula for a unit quaternion x so that $x \otimes x$ rotates a into alignment with b, where a, b are given pure quaternions. In other words, so that $(x \otimes x)(a) = xa\overline{x}$ is a positive scalar multiple of b.

Proposition 4
Suppose $a, b \in \mathbb{P}$ are nonzero pure quaternions such that $|ba| - ba \neq 0$ (equivalently, such that $a/|a| \neq -b/|b|$), and let x be the unit quaternion

$$x = \frac{|ba| - ba}{|\,|ba| - ba\,|} = \frac{|b|\,|a| - ba}{|\,|b|\,|a| - ba\,|}\,. \tag{19}$$

Then $\phi(x \otimes x) \in \mathrm{SO}(3)$ rotates a into alignment with b.

Remarks

1. Since $(-x) \otimes (-x) = x \otimes x$, choosing $x = \frac{ba - |ba|}{|\,ba - |ba|\,|}$ will also work.
2. When a and b are already in alignment, i.e. when $a/|a| = b/|b|$, then Eqn. (19) reduces to $x = 1$ so that $\phi(x \otimes x) = I_4$.
3. Geometrically, it is clear that there are infinitely many 3-dimensional rotations that send a into alignment with b. Among these, it is natural to prefer the one which uses the smallest angle. This smallest angle rotation will have the normal N to the plane spanned by a and b as its axis. By contrast, the largest angle (180°) rotation has the angle bisector between a and b as its axis. It can be shown that when a and b are linearly independent, then $\phi(x \otimes x)$ has precisely the normal N as its axis of rotation. Thus $\phi(x \otimes x)$ is the *smallest angle* rotation that aligns a with b.

5.2 Symmetric Hamiltonian

Given a symmetric Hamiltonian matrix $H = q \otimes i + r \otimes k$, our goal is to find a symplectic orthogonal matrix $R = x \otimes y$ such that $RHR^t = (x \otimes y)(q \otimes i + r \otimes k)(\overline{x} \otimes \overline{y})$ is diagonal. ¿From Eqn. (18) and the characterization of diagonal matrices given in Table 3, we see that this would be achieved if the pure quaternion q were rotated into a multiple of i, and the pure quaternion r into a multiple of k. But q and r are affected only by x. As vectors in $\mathbb{R}^3$, q is not in general orthogonal to r, so no rotation can simultaneously align q along i, and r along $\pm k$ as desired.

We therefore exploit a different quaternion representation of symmetric Hamiltonian matrices, as was done for general symmetric matrices in [40]. The difference here is that the transformation R must be symplectic orthogonal, and not just orthogonal.

Using the vector space isomorphism $\psi : \mathbb{P} \otimes \mathbb{P} \to \mathbb{R}^{3\times 3}$, defined as the unique linear extension of the map that sends $a \otimes b$ to the rank one matrix $ab^t \in \mathbb{R}^{3\times 3}$, we get

$$\begin{array}{rcll} \psi(H) &=& \psi(q \otimes i) + \psi(r \otimes k) & \psi \text{ is linear} \\ &=& q\, e_1^t + r\, e_3^t & \text{definition of } \psi \\ &=& \begin{bmatrix} q_1 & 0 & r_1 \\ q_2 & 0 & r_2 \\ q_3 & 0 & r_3 \end{bmatrix} & \\ &=& \sigma_1 u_1 v_1^t + \sigma_2 u_2 v_2^t & \text{singular value decomposition} \\ &=& \psi(\sigma_1 u_1 \otimes v_1 + \sigma_2 u_2 \otimes v_2)\,. & \end{array}$$

Since ψ is an isomorphism, we have $H = q \otimes i + r \otimes k = \sigma_1(u_1 \otimes v_1) + \sigma_2(u_2 \otimes v_2)$, where $\sigma_1 \geq \sigma_2 \geq 0$ are the singular values of the 3×2 matrix $[\,q \quad r\,]$. This alternative tensor representation of H has the advantage that it uses two *orthonormal* pairs of vectors, the left singular vectors $\{u_1, u_2\}$, and the right singular vectors $\{v_1, v_2\}$. Furthermore, note that v_1, v_2 lie in the ik–plane. This follows directly from the geometry of the singular value decomposition, which tells us that v_1, v_2 lie in the plane perpendicular to the remaining right singular vector, v_3. Now by inspection, $\sigma_3 = 0$ and $v_3 = j = (0,\, 1,\, 0)^t$, so we have our result. Alternatively, we can replace v_1 and v_2 by general linear combinations of i, j, k in the equation $q \otimes i + r \otimes k = \sigma_1(u_1 \otimes v_1) + \sigma_2(u_2 \otimes v_2)$. Equating like terms then leads to the same conclusion after an elementary algebraic argument.

Clearly, there exists a rotation of $\mathbb{R}^3$ that simultaneously sends u_1 to i and u_2 to $\pm k$. To complete the diagonalization of H, we need a second rotation of $\mathbb{R}^3$ that sends the right singular vectors v_1, v_2 to i, $\pm k$ respectively. The symplectic constraint requires this second rotation to be *about the j axis.* Such a rotation can achieve the goal of mapping v_1 to i and v_2 to $\pm k$ if and only if v_1, v_2 already lie in the ik–plane. But, as we've shown in the preceding paragraph, this is precisely the case! Hence we conclude that given any symmetric Hamiltonian matrix $H \in \mathbb{R}^{4\times 4}$, there exists a 4×4 symplectic orthogonal matrix R such that RHR^t is diagonal.

To calculate R, first choose a unit quaternion x so that the 3-dimensional rotation $R_1 = \phi(x \otimes x)$ acting on $\mathbb{P} \cong \mathbb{R}^3$ sends u_1 to i. (If u_1 is already aligned with i, then skip this step.) ¿From Proposition 4 we have

$$x = \frac{|i||u_1| - iu_1}{|\,|i||u_1| - iu_1\,|} = \frac{1 - iu_1}{|\,1 - iu_1\,|}\,. \tag{20}$$

Now since R_1 rotates u_1 to i, it will also send u_2 to some $\widehat{u}_2$ in the jk–plane. Consequently, unless $\sigma_2 = 0$, a subsequent rotation with axis along i is needed to bring $\widehat{u}_2$ into alignment with k, but we postpone this calculation for now.

Instead, we turn our attention to the right singular vectors v_1, v_2 and the rotation R_2 needed to align them with i, $\pm k$ respectively. As we remarked earlier in this section, for the entire transformation to be symplectic, R_2 must have j as its axis. Since the two right singular vectors already lie in the ik–plane, the rotation with the smallest angle is *precisely* the one with axis along j, and we are presented with an unexpected bonus: the unique rotation that meets the symplectic condition is also the rotation that uses the minimal angle. Thus the unit quaternion y associated with $R_2 = \phi(y \otimes y)$ is again given by Proposition 4:

$$y = \frac{1 - iv_1}{|\, 1 - iv_1 \,|} \tag{21}$$

With x and y as in Eqns. (20) and (21), consider the effect of a similarity by the symplectic orthogonal matrix $\widehat{R} = \phi(x \otimes y)$ on H:

$$\begin{aligned} \widehat{R} H \widehat{R}^t &= (x \otimes y)(\sigma_1 u_1 \otimes v_1 \; + \; \sigma_2 u_2 \otimes v_2)(\overline{x} \otimes \overline{y}) \\ &= \sigma_1 (x u_1 \overline{x} \otimes y v_1 \overline{y}) \; + \; \sigma_2 (x u_2 \overline{x} \otimes y v_2 \overline{y}) \\ &= \sigma_1 (i \otimes i) \; \pm \; \sigma_2 (\widehat{u_2} \otimes k) \end{aligned}$$

where $x u_2 \overline{x} = \widehat{u_2} \in \operatorname{span}\{j, k\}$. Checking Table 3, one quickly observes that $\widehat{R} H \widehat{R}^t$ is already 2×2 block-diagonal.

Furthermore, since the singular vectors u_1, v_1 associated with the largest singular value are sent to i, we can use [40, Prop. 8] to conclude that the eigenvalues of the upper diagonal block are larger than those of the lower one. Since H is also Hamiltonian, its eigenvalues come in plus-minus pairs, so the upper diagonal block of $\widehat{R} H \widehat{R}^t$ has positive eigenvalues, while their negative partners are the eigenvalues of the lower diagonal block. Indeed, by putting $\sigma_3 = 0$ in [40, Eqns. (22)–(25)], we see that the eigenvalues of H (in decreasing order) are $\sigma_1 + \sigma_2$, $\sigma_1 - \sigma_2$, $-\sigma_1 + \sigma_2$ and $-\sigma_1 - \sigma_2$. Here $\sigma_1 \geq \sigma_2$ are the singular values of the 3×2 matrix $[\, q \quad r \,]$ associated with $H = q \otimes i + r \otimes k$.

The symplectic orthogonal matrix $\widehat{R} = \phi(x \otimes y)$ is the product of the commuting matrices $\phi(x \otimes 1)$ and $\phi(1 \otimes y)$ which can be found using Eqns. (3), (20) and (21). Writing $u_1 = (u_{11},\ u_{12},\ u_{13})^t$ and $v_1 = (v_{11},\ 0,\ v_{13})^t$, the expression for $\widehat{R}$ ultimately reduces to

$$\widehat{R} = \frac{1}{2\sqrt{d_u\, d_v}} \begin{pmatrix} d_u & 0 & -u_{13} & u_{12} \\ 0 & d_u & u_{12} & u_{13} \\ u_{13} & -u_{12} & d_u & 0 \\ -u_{12} & -u_{13} & 0 & d_u \end{pmatrix} \begin{pmatrix} d_v & 0 & v_{13} & 0 \\ 0 & d_v & 0 & v_{13} \\ -v_{13} & 0 & d_v & 0 \\ 0 & -v_{13} & 0 & d_v \end{pmatrix}, \tag{22}$$

where $d_u = 1 + u_{11}$, and $d_v = 1 + v_{11}$.

Finally, complete diagonalization is most simply achieved using a similarity by a double plane rotation

$$\widetilde{R} = \begin{pmatrix} \cos\theta & \sin\theta & & \\ -\sin\theta & \cos\theta & & \\ & & \cos\theta & \sin\theta \\ & & -\sin\theta & \cos\theta \end{pmatrix}.$$

Note that $\widetilde{R}$ is symplectic orthogonal, ensuring that $R = \widetilde{R}\widehat{R}$ is likewise symplectic orthogonal. If θ is chosen to be the sorting angle [41], then the eigenvalues of H appear in decreasing order on the main diagonal.

The pseudo-MATLAB code of Algorithm 1 condenses the computational results of this section. Organized for clarity, this code is not optimal – for example, the singular vector pair corresponding to the largest singular value of a 3×2 matrix is determined by a call to the `svd` routine. Note that the reduced SVD is computed, which means that the right singular vectors v_1 and v_2 have only two components. A Jacobi algorithm for $2n \times 2n$ symmetric Hamiltonian matrices based on Algorithm 1 is developed in §7.2. There we show why complete diagonalization of the target 4×4 symmetric Hamiltonian subproblem is necessary, unlike the case for general symmetric matrices, for which (2×2)–block-diagonalization of the 4×4 target suffices [40].

Algorithm 1 (4×4 SYMMETRIC HAMILTONIAN)
INPUT: 4×4 *real symmetric Hamiltonian matrix* H
OUTPUT: 4×4 *real symplectic orthogonal matrix* R *such that* RHR^t *is diagonal.*

```
Begin
  q = [.5*(H(1,1) + H(2,2)); H(1,4); .5*(H(2,4)-H(1,3))];
  r = [.5*(H(2,4) + H(1,3)); -H(1,2); .5*(H(1,1)-H(2,2))];
  [U, D, V]= svd([q r]);
  u = U(:,1); du = 1 + u(1);
  v = V(:,1); dv = 1 + v(1);
  R_hat1 = (1/sqrt(2*du))*[du 0 -u(3) u(2); 0 du u(2) u(3);...
                        u(3) -u(2) du 0;-u(2) -u(3) 0 du];
  R_hat2 = (1/sqrt(2*dv))*[dv 0 v(2) 0; 0 dv 0 v(2);...
                        -v(2) 0 dv 0;0 -v(2) 0 dv];
  R = R_hat1*R_hat2; H = R*H*R^t;    %H is now block diagonal
  tau = (H(2,2)- H(1,1))/(2*H(1,2));
  t = 1/(abs(tau) + sqrt(1 + tau*tau));
  if tau < 0 then t = -t endif;
  c = 1/sqrt(1 + t*t); s = t*c;
  if (H(1,1) >= H(2,2))
     then R_tilde = [c s 0 0; -s c 0 0; 0 0 c s; 0 0 -s c];    % use small angle
     else R_tilde = [s c 0 0; c -s 0 0; 0 0 s c; 0 0 c -s];    % use large angle
  endif
  R = R_tilde*R;    %H = R*H*R^t is diagonal, R is symplectic orthogonal
End.
```

5.3 Skew-symmetric Hamiltonian

If the Hamiltonian matrix $H \in \mathbb{R}^{4\times 4}$ is skew-symmetric, then we can write $H = b\,(1 \otimes j) + p \otimes 1$, where $b \in \mathbb{R}$ and $p \in \mathbb{P}$ (see Table 3). Since $(1 \otimes j) = J_4$, it is immediately obvious that similarity by a symplectic orthogonal has no effect on the first term of H. A quick check of the quaternion basis in Appendix A also makes the choice of action equally clear: rotate p into a multiple of j. ¿From Proposition 4 it is clear that if

$$x = \frac{|p| - jp}{|\,|p| - jp\,|} \quad \text{and} \quad R = x \otimes 1\,, \tag{23}$$

then the symplectic orthogonal similarity RHR^t reduces H to the form $b\,(1 \otimes j) + |p|\,(j \otimes 1)$. In matrix form, if $p = p_1 i + p_2 j + p_3 k$, then

$$R = x \otimes 1 = \frac{1}{\sqrt{2\,|p|(|p| + p_2)}} \begin{pmatrix} |p| + p_2 & p_3 & 0 & -p_1 \\ -p_3 & |p| + p_2 & -p_1 & 0 \\ 0 & p_1 & |p| + p_2 & p_3 \\ p_1 & 0 & -p_3 & |p| + p_2 \end{pmatrix} \tag{24}$$

gives

$$RHR^t = \begin{pmatrix} 0 & 0 & -|p| + b & 0 \\ 0 & 0 & 0 & |p| + b \\ |p| - b & 0 & 0 & 0 \\ 0 & -|p| - b & 0 & 0 \end{pmatrix} \tag{25}$$

The values of the parameters b, p_1, p_2, p_3 in terms of the entries of H are specified in Eqn. (14), giving us the following algorithm:

Algorithm 2 (4×4 SKEW-SYMMETRIC HAMILTONIAN)
INPUT: 4×4 *real skew-symmetric Hamiltonian matrix* H
OUTPUT: 4×4 *real symplectic orthogonal matrix* R *such that* RHR^t *is in the canonical form specified in Eqn. (25).*

```
Begin
  R = eye(4);
  p = [H(2,1); .5*(H(3,1) - H(4,2)); H(4,1); ];
  np = norm(p); dp = np + p(2);
  R = (1/sqrt(2*np*dp))*[dp p(3) 0 -p(1); -p(3) dp -p(1) 0;...
                          0 p(1) dp p(3); p(1) 0 -p(3) dp];
End.
```

5.4 Symmetric skew-Hamiltonian

In $\mathbb{H} \otimes \mathbb{H}$, a 4×4 symmetric skew-Hamiltonian is of the form $W = b\,(1 \otimes 1) + p \otimes j$, and a symplectic orthogonal similarity that diagonalizes W can be immediately constructed: rotate the pure quaternion p to j using $R = x \otimes 1$ as given in Eqns. (23) and (24). Then $RWR^t = b\,(1 \otimes 1) + |p|(j \otimes j)$ is readily seen to be of the form $\mathrm{diag}(b + |p|,\ b - |p|,\ b + |p|,\ b - |p|)$. Except for the specification of p, which is given by Eqn. (15), the procedure here is identical to Algorithm 2 given in the previous section.

5.5 Skew-symmetric skew-Hamiltonian

Once again, this case can be handled with ease in the quaternion tensor algebra, where a 4×4 skew-symmetric skew-Hamiltonian can be expressed as $W = 1 \otimes (c\,i + dk)$. Clearly there exists a rotation of $\mathbb{P} \cong \mathbb{R}^3$ with axis j that aligns the vector $c\,i + dk$ along i. Hence we can construct a symplectic orthogonal matrix $R = 1 \otimes x$ such that RWR^t is a real scalar multiple of $1 \otimes i$. ¿From Proposition 4 we see that the unit quaternion x is given by

$$x = \frac{|\,c\,i + dk\,| - i(c\,i + dk)}{|\,|(c\,i + dk)| - i(c\,i + dk)\,|} = \frac{\sqrt{c^2 + d^2} + c + dj}{|\,\sqrt{c^2 + d^2} + c + dj\,|}.$$

Writing $x = x_0 + x_2 j$ and using Eqn. (3), we have

$$R = 1 \otimes x = \begin{pmatrix} x_0 & 0 & x_2 & 0 \\ 0 & x_0 & 0 & x_2 \\ -x_2 & 0 & x_0 & 0 \\ 0 & -x_2 & 0 & x_0 \end{pmatrix}.$$

Working in $\mathbb{H} \otimes \mathbb{H}$, then using the matrix representation of $1 \otimes i$ from Appendix A shows that

$$\begin{aligned} RWR^t &= (1 \otimes x)\,(1 \otimes (ci + dk))\,(1 \otimes \overline{x}) \\ &= |\,ci + dk\,|(1 \otimes i) \\ &= \begin{pmatrix} 0 & \sqrt{c^2 + d^2} & 0 & 0 \\ -\sqrt{c^2 + d^2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\sqrt{c^2 + d^2} \\ 0 & 0 & \sqrt{c^2 + d^2} & 0 \end{pmatrix}. \end{aligned} \tag{26}$$

We do not present pseudo-code for this algorithm, for reasons that will become clear in §7.3.

6 Symplectic Orthogonal Canonical Forms

The results for 4×4 matrices provide a foundation on which to build structure-preserving Jacobi algorithms for the corresponding four classes of $2n \times 2n$ doubly structured matrices. Equally important, they suggest which canonical forms might be achievable in the $2n \times 2n$ case, when similarity transformations are restricted to the symplectic orthogonal group $\mathrm{SpO}(2n)$. The next theorem gives such Jacobi algorithms a goal to aim at, and confirms the canonical forms suggested by the 4×4 case. We use $D \in \mathbb{R}^{n \times n}$ to denote a diagonal matrix, and $B \in \mathbb{R}^{n \times n}$ a block-diagonal matrix which is the direct sum of 1×1 zero blocks and 2×2 blocks of the form $\left[\begin{smallmatrix} 0 & b \\ -b & 0 \end{smallmatrix}\right]$.

Theorem 1

(a) *For any symmetric Hamiltonian $H \in \mathbb{R}^{2n \times 2n}$ there exists a symplectic orthogonal $S \in \mathrm{SpO}(2n)$ such that*

$$S^t H S = \begin{pmatrix} D & 0 \\ 0 & -D \end{pmatrix}.$$

(b) *For any skew-symmetric Hamiltonian* $H \in \mathbb{R}^{2n \times 2n}$ *there exists a symplectic orthogonal* $S \in \mathrm{SpO}(2n)$ *such that*

$$S^t H S = \begin{pmatrix} 0 & -D \\ D & 0 \end{pmatrix}.$$

(c) *For any symmetric skew-Hamiltonian* $W \in \mathbb{R}^{2n \times 2n}$ *there exists a symplectic orthogonal* $S \in \mathrm{SpO}(2n)$ *such that*

$$S^t W S = \begin{pmatrix} D & 0 \\ 0 & D \end{pmatrix}.$$

(d) *For any skew-symmetric skew-Hamiltonian* $W \in \mathbb{R}^{2n \times 2n}$ *there exists a symplectic orthogonal* $S \in \mathrm{SpO}(2n)$ *such that*

$$S^t W S = \begin{pmatrix} B & 0 \\ 0 & -B \end{pmatrix}.$$

Some parts of this theorem can be found scattered in the literature, with various unconnected proofs. See for example [8, 9, 30, 43]. It is possible, however, to develop all four of the above canonical form results simultaneously in a unified framework [39].

7 Structure Preserving Jacobi Algorithms

7.1 Embeddings, Sweep patterns

Based on the solutions of 4×4 problems given in §5 and the canonical form results of §6, Jacobi algorithms for $2n \times 2n$ matrices can be developed once we determine how to design a sweep consisting of a *complete* set of 4×4 structured subproblems. By this we mean that every element of the $2n \times 2n$ matrix should be part of a target submatrix at least once during the sweep, so no region of the large matrix is exempt from direct participation in the structure preserving drive towards the canonical form.

It turns out that such sweeps are not difficult to construct. First note that the location of a structured subproblem is controlled by the position of an off-diagonal element h_{ij} chosen from the $n \times n$ upper diagonal block. This element uniquely determines a 4×4 principal submatrix located in rows and columns $i, j, n+i$ and $n+j$ as shown below:

$$\left[\begin{array}{cc|cc} h_{ii} & h_{ij} & h_{i,n+i} & h_{i,n+j} \\ h_{ji} & h_{jj} & h_{j,n+i} & h_{j,n+j} \\ \hline h_{n+i,i} & h_{n+i,j} & h_{n+i,n+i} & h_{n+i,n+j} \\ h_{n+j,i} & h_{n+j,j} & h_{n+j,n+i} & h_{n+j,n+j} \end{array}\right] \tag{27}$$

Submatrices of this type inherit both structures from the parent matrix H — Hamiltonian or skew-Hamiltonian together with symmetry or skew-symmetry. One can also immediately see that *any* complete Jacobi sweep of the $n \times n$ upper diagonal block consisting of 2×2 principal submatrices generates a corresponding complete sweep of the $2n \times 2n$ matrix comprised entirely of 4×4 structured subproblems. This is illustrated in Fig. 1 using an 8×8 matrix and a simple ordering: as the entry represented by $\diamond$ traces a row-cyclic pattern in the $n \times n$ upper diagonal block, it determines a sequence of structured 4×4 target matrices of the type given in Eqn. (27). The other fifteen entries of the targets are denoted by a heavy bullet. Observe that every entry of the 8×8 matrix has the opportunity to be part of a target submatrix during the course of the sweep.

Once a target 4×4 submatrix of H has been identified, a 4×4 symplectic orthogonal matrix R is constructed by the appropriate algorithm from §5. Embedding R into the $2n \times 2n$ identity matrix in exactly the same manner that the 4×4 target was extracted from H yields a $2n \times 2n$ symplectic orthogonal. Only four rows and four columns of H change under the resulting similarity; the accumulated symplectic orthogonal can be updated by rewriting four columns if the update is performed by postmultiplication. We remark that a one-sided Jacobi implementation [18, 46, 52] would require only column updates to both H as well as the accumulated symplectic orthogonal matrix.

Much work has been done on investigating various Jacobi orderings in the 2×2 setting, epsecially those that are parallelizable. See for example [19, 38, 42, 44, 53]. Every 2×2 based parallel ordering is applicable to our situation, yielding a 4×4 based parallel ordering as a natural extension. Again we emphasize that these 4×4 based sweeps are complete, reaching every part of the matrix, unlike the sweep strategy used in [11].

Finally we note that since the double structure of the $2n \times 2n$ matrix is always preserved, both storage requirements and flop counts can be lowered by roughly a factor of four. For example, a $2n \times 2n$ symmetric Hamiltonian matrix needs only $n^2 + n$ storage locations, whereas a $2n \times 2n$ skew-symmetric skew-Hamiltonian matrix can be stored using $n^2 - n$ locations. Since the entire $2n \times 2n$ matrix does not have to be recomputed, the preservation of structure also leads to a corresponding savings in flops.

7.2 $2n \times 2n$ Symmetric Hamiltonians

In §5.2 we remarked that if the $2n \times 2n$ matrix $H = \left[\begin{smallmatrix} E & F \\ G & -E^t \end{smallmatrix}\right]$ is symmetric Hamiltonian, then to achieve a convergent Jacobi algorithm it is not sufficient to merely (2×2)-block-diagonalize the target 4×4 submatrices. The reason lies in their location in H, which is best understood by examining Fig. 1. For example, looking at the second or third iteration of the row-cyclic sweep makes it clear that block-diagonalizing the target submatrices could at best result in zeroing out the $n \times n$ off-diagonal blocks F and G, thereby block-diagonalizing H. Also note that while Fig. 1 illustrates just the row-cyclic ordering, the constraint of using structured subproblems means that a different ordering will only permute the sequence of target submatrices, not choose a different set of target submatrices. Thus under any complete structured sweep where the target submatrices are only (2×2)–block diagonalized, the best we can expect to achieve is to block-diagonalize H into two $n \times n$ diagonal blocks.

But, this expectation is much too sanguine. Again, a careful look at the second and third iterations in Fig. 1 reveals why. By block-diagonalizing the target, the $(1, 3)$ entry in

$$
\left(\begin{array}{cccc|cccc}
\bullet & \diamond & \cdot & \cdot & \bullet & \bullet & \cdot & \cdot \\
\bullet & \bullet & \cdot & \cdot & \bullet & \bullet & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\hline
\bullet & \bullet & \cdot & \cdot & \bullet & \bullet & \cdot & \cdot \\
\bullet & \bullet & \cdot & \cdot & \bullet & \bullet & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
\rightsquigarrow
\left(\begin{array}{cccc|cccc}
\bullet & \cdot & \diamond & \cdot & \bullet & \cdot & \bullet & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\bullet & \cdot & \bullet & \cdot & \bullet & \cdot & \bullet & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\hline
\bullet & \cdot & \bullet & \cdot & \bullet & \cdot & \bullet & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\bullet & \cdot & \bullet & \cdot & \bullet & \cdot & \bullet & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
\rightsquigarrow
$$

$$
\rightsquigarrow
\left(\begin{array}{cccc|cccc}
\bullet & \cdot & \cdot & \diamond & \bullet & \cdot & \cdot & \bullet \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \bullet & \bullet & \cdot & \cdot & \bullet \\
\hline
\bullet & \cdot & \cdot & \bullet & \bullet & \cdot & \cdot & \bullet \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \bullet & \bullet & \cdot & \cdot & \bullet
\end{array}\right)
\rightsquigarrow
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \bullet & \diamond & \cdot & \cdot & \bullet & \bullet & \cdot \\
\cdot & \bullet & \bullet & \cdot & \cdot & \bullet & \bullet & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \bullet & \bullet & \cdot & \cdot & \bullet & \bullet & \cdot \\
\cdot & \bullet & \bullet & \cdot & \cdot & \bullet & \bullet & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
\rightsquigarrow
$$

$$
\rightsquigarrow
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \bullet & \cdot & \diamond & \cdot & \bullet & \cdot & \bullet \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \bullet & \cdot & \bullet & \cdot & \bullet & \cdot & \bullet \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \bullet & \cdot & \bullet & \cdot & \bullet & \cdot & \bullet \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \bullet & \cdot & \bullet & \cdot & \bullet & \cdot & \bullet
\end{array}\right)
\rightsquigarrow
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \bullet & \diamond & \cdot & \cdot & \bullet & \bullet \\
\cdot & \cdot & \bullet & \bullet & \cdot & \cdot & \bullet & \bullet \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \bullet & \bullet & \cdot & \cdot & \bullet & \bullet \\
\cdot & \cdot & \bullet & \bullet & \cdot & \cdot & \bullet & \bullet
\end{array}\right)
$$

Figure 1: Row Cyclic Structured Sweep

the second iteration (marked $\diamond$), is allowed to grow in absolute value, due to the transfer of some "weight" from the target entries of F and G that are killed. (Since the similarity transformations are orthogonal, norms of both the target submatrix as well as H are preserved.) During the next iteration, this entry is not part of the target submatrix. But since it lies in an affected row, there is liable to be some "leakage" of its weight back into the $(5, 3)$ and $(8, 3)$ entries, that is, into one of the $n \times n$ off-diagonal blocks that we hope to zero out eventually. When a row-cyclic Jacobi algorithm based on (2×2)–block diagonalization of the 4×4 target submatrices is implemented on a symmetric Hamiltonian matrix, this is exactly what happens — the norm of the $n \times n$ off-diagonal blocks rises and falls, and non-convergence is the result.

While it is true that the heuristic argument we have presented seems to depends on a particular order in which the target submatrices are encountered, the same problem arises with any complete sweep. In any iteration, the affected rows and columns contain "large" entries that are not part of the current target submatrix. These entries, which are located in the $n \times n$ diagonal blocks E and $-E^t$ of H, interact with their "partners" in the off-diagonal $n \times n$ blocks, with the result that the norm of F (and hence the norm of G) fluctuates without converging to zero.

This problem can be circumvented by *diagonalizing* the target 4×4 submatrix. Since the targets are always principal submatrices, this moves the weight of the target entirely onto the main diagonal of H, from which it can never leave. Diagonal elements are never part of affected rows or columns without also being part of a target submatrix. And when this happens, additional weight is transferred onto them, rather than moved out from them. Thus no subsequent iterations, no matter what the ordering, can extract weight from the main diagonal and deposit it elsewhere in the matrix. The diagonal is in effect a "safe haven" for the norm of H; once off-diagonal weight arrives at the safe haven, it never gets expelled by later iterations. It is now reasonable to hope that the corresponding Jacobi algorithm converges; this is borne out by numerical experimentation. Formal proofs of convergence fashioned along the lines of the theorems in [41], which are valid under mild hypotheses for all quasi-cyclic orderings, would be the subject of a future paper.

We end this section with a comment on why (2×2)–block-diagonalization of the target 4×4 submatrices *does* suffice for the quaternion-Jacobi algorithms for symmetric matrices developed in [40]. There the only structure being preserved is symmetry of the client matrix, and the most convenient structured subproblems to use are those in rows and columns $i, i+1, j, j+1$. For illustrative purposes, a row-cyclic sweep consisting of such 4×4 targets is shown in Fig. 2 for an 8×8 matrix. By block-diagonalizing the target submatrix, weight is always transferred onto 2×2 blocks on the *main diagonal*, and these areas are immune to "leakage" for the same reason mentioned earlier — the only time they are involved in an update is when weight is being transferred onto them. Other cyclic orderings involve these same subproblems, just visited in a different sequence during the sweep. So for these algorithms, the set of 2×2 blocks along the main diagonal forms a "safe haven", and convergence is once again a reasonable expectation. Numerical experiments bear this out [40]; it is even possible to show formally that algorithms based on (2×2)–block-diagonalization with sweep patterns analogous to those in Fig. 2 do converge to (2×2)–block-diagonal form. However, if the Hamiltonian or skew-Hamiltonian structure of a client matrix is also to be preserved, then such sweeps cannot be used.

7.3 $2n \times 2n$ Skew-symmetric skew-Hamiltonians

We now come to the interesting case of the skew-symmetric skew-Hamiltonians. As shown in §5.5, the quaternion tensor square makes the 4×4 canonical form problem for this class completely straightforward to solve. However, while the 4×4 based Jacobi algorithms for the other three classes converge, the 4×4 based algorithm for this class does not!

Examining Eqn. (26), which gives the 4×4 canonical form, and Fig. 1, which shows the location of structured subproblems, we see why. The most such an algorithm can hope to accomplish is to zero out the $n \times n$ off-diagonal blocks. But as discussed in §7.2, "leakage" of

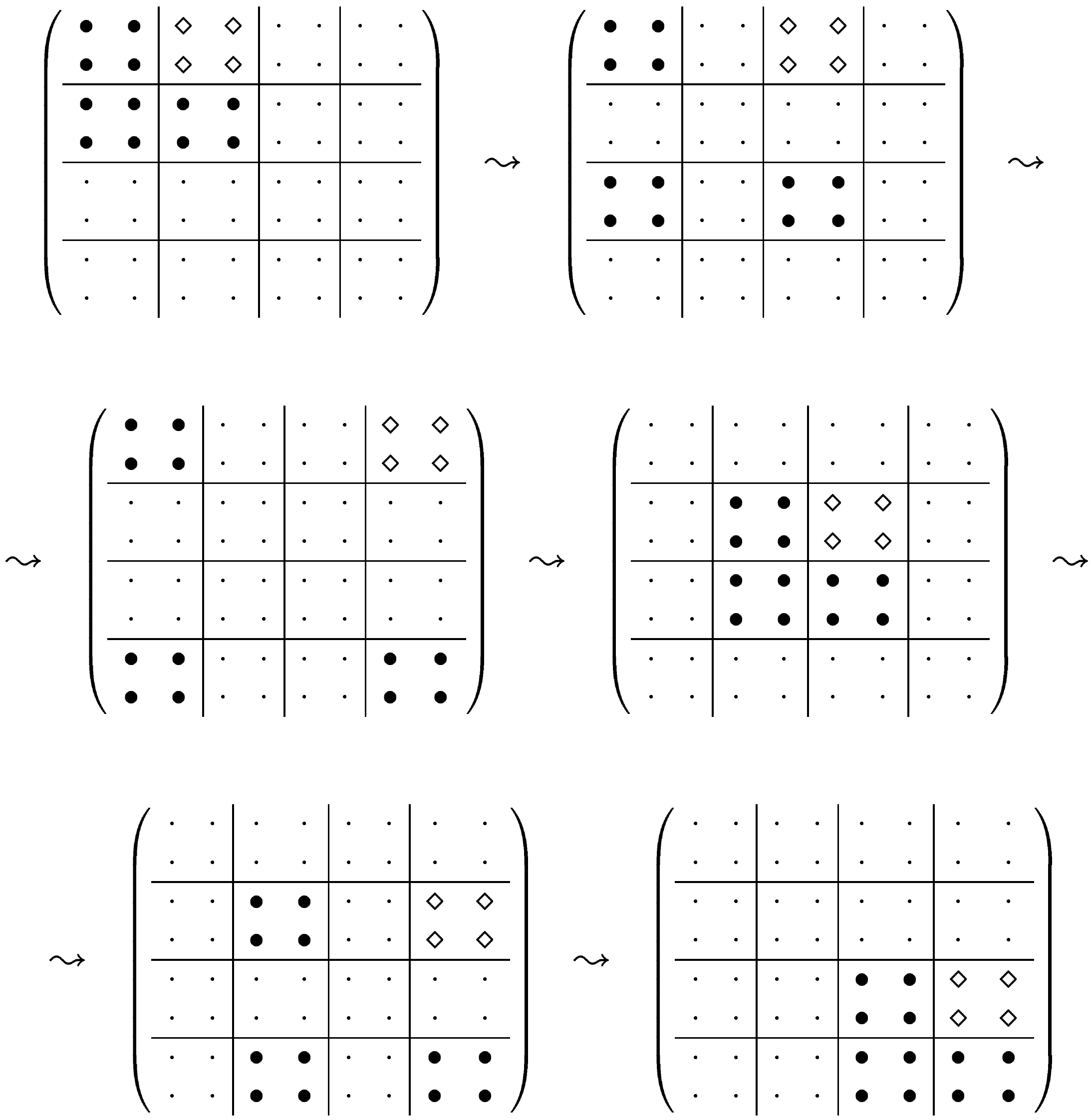

Figure 2: Row Cyclic Sweep for Symmetric or Skew-symmetric matrices that are **not** Hamiltonian or skew-Hamiltonian

weight from the $n \times n$ diagonal blocks to the $n \times n$ off-diagonal blocks will prevent this from happening. For the symmetric Hamiltonian problem, the solution was simply to diagonalize the target. In this case, however, the main diagonal is always zero, so we literally have no "safe haven" where the norm of the target submatrix can be sheltered, and the 4×4 based Jacobi algorithm for skew-symmetric skew-Hamiltonians is doomed to failure.

This seems like an insurmountable difficulty until one realizes that it is possible to explicitly solve even *larger* skew-symmetric skew-Hamiltonian subproblems in *closed form*. We now describe a direct solution to the 8×8 canonical form problem.

Let W be an arbitrary 8×8 skew-symmetric skew-Hamiltonian matrix. Then $W =$

$\left[\begin{smallmatrix} A & C \\ C & -A \end{smallmatrix}\right]$ with skew-symmetric blocks $A, C \in \mathbb{R}^{4\times 4}$. We can first reduce W to $\widetilde{W} = \left[\begin{smallmatrix} \widetilde{A} & 0 \\ 0 & -\widetilde{A} \end{smallmatrix}\right]$, where $\widetilde{A} \in \mathbb{R}^{4\times 4}$ is both skew-symmetric and tridiagonal, by a *finite* sequence of symplectic orthogonal similarities in one of two ways:

1. Use the finite reduction procedure for general skew-Hamiltonian matrices described by Van Loan in [54, §4]. As illustrated in Appendix B, this method involves four symplectic 8×8 Householder matrices (each is a direct sum of two "ordinary" 4×4 Householder matrices) and three symplectic Givens transformations (each being a 2×2 rotation appropriately embedded in an 8×8 identity matrix.)

2. A simpler method, developed in §7.3.1 and §7.3.2 that uses three symplectic 4×4 Givens transformations and one symplectic 2×2 Givens.

It is well-known that any skew-symmetric matrix can be put into real Schur form by a rotational similarity; for a 4×4 skew-symmetric matrix K this similarity can, using $\mathbb{H}\otimes\mathbb{H}$, be simply and explicitly expressed in terms of the entries of K as described in [25, §3] and [40, §8]. In our case the computation of $R \in \mathrm{SO}(4)$ such that $R\widetilde{A}R^t$ is in real Schur form is even simpler, since $K = \widetilde{A}$ is also tridiagonal. Thus we can put $\widetilde{W}$ into the canonical form specified in Theorem 1(d) by similarity with a symplectic "double quaternion rotation" of the form $S = \left[\begin{smallmatrix} R & 0 \\ 0 & R \end{smallmatrix}\right] \in \mathrm{SpO}(8)$:

$$S\widetilde{W}S^t = \begin{bmatrix} R\widetilde{A}R^t & 0 \\ 0 & -R\widetilde{A}R^t \end{bmatrix} = \begin{bmatrix} B & 0 \\ 0 & -B \end{bmatrix} \in \mathbb{R}^{8\times 8}.$$

Pseudo-code for the compete solution of the 8×8 problem using symplectic 4×4 Giviens can be found in §7.3.3.

Jacobi algorithms for the $2n \times 2n$ problem based on this solution of the 8×8 problem are now straightforward to describe. Let $X = \left[\begin{smallmatrix} A & C \\ C & -A \end{smallmatrix}\right] \in \mathbb{R}^{2n\times 2n}$ be skew-symmetric skew-Hamiltonian; the $n \times n$ blocks A and C are skew-symmetric. For simplicity let us suppose that n is even. First partition X into contiguous 2×2 blocks. Next consider any complete 4×4 based sweep of *just* A using these 2×2 blocks, such as the row-cyclic sweep illustrated in Fig. 2. This sweep now generates a corresponding complete structured sweep of the $2n \times 2n$ matrix X comprised entirely of 8×8 structured subproblems. These 8×8 structured subproblems are embedded in X exactly as in Fig. 1, but with each $\bullet$, $\diamond$, and $\cdot$ now representing a 2×2 block rather than just a single entry of X. By using 8×8 subproblems embedded in this way, the 2×2 diagonal blocks of X become a safe haven for the norm of target submatrices, and the resulting 8×8 based Jacobi algorithms can reasonably be expected to converge. They do indeed converge; numerical experiments illustrating this can be found in §8.

When n is odd, neither A nor C can be completely partitioned into just 2×2 blocks, so a modification of the above procedure is needed. Partition the $n \times n$ submatrix A as much as possible into 2×2 blocks, leaving 2×1 and 1×2 blocks along the rightmost and lower edges, as in Fig. 3. Then use these blocks to build complete sweeps of A, such as the row-cyclic sweep shown in Fig. 3; with C partitioned in the same way as A, this generates corresponding

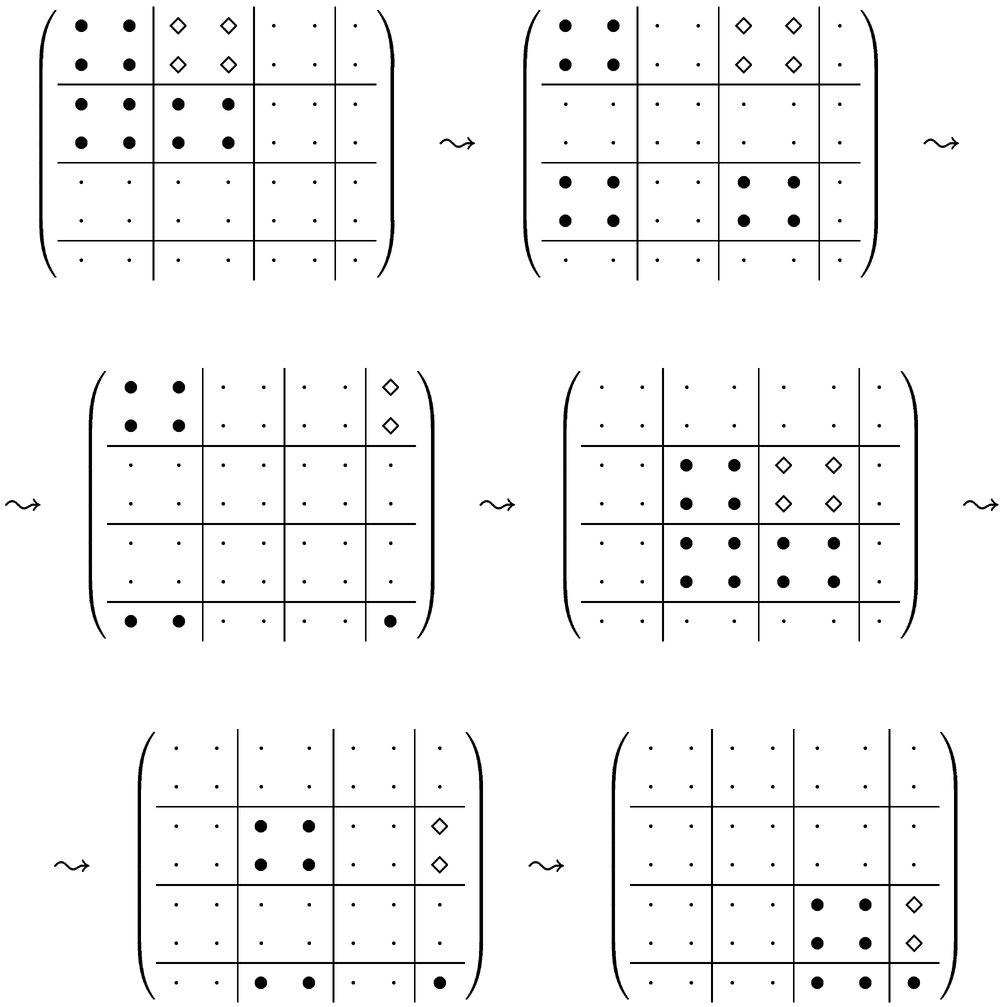

Figure 3: Row Cyclic Sweep in block A when n is odd

complete structured sweeps of X. Most of the structured subproblems in these sweeps of X are still 8×8, but some around the edges are 6×6, so we also need to be able to explicitly solve 6×6 subproblems.

As one might expect, these 6×6 subproblems can be solved using the same strategy that worked for the 8×8 problem, with some small modifications of the details. Let $W \in \mathbb{R}^{6\times 6}$ be skew-symmetric skew-Hamiltonian. By a specialization of the reduction procedure described in §7.3.2 for the 8×8 case, W can be reduced to $\widetilde{W} = \left[\begin{smallmatrix} \widetilde{A} & 0 \\ 0 & -\widetilde{A} \end{smallmatrix}\right]$ using just *one* 4×4 symplectic Givens followed by one 2×2 symplectic Givens. Here $\widetilde{A} \in \mathbb{R}^{3\times 3}$ is skew-symmetric and tridiagonal. Next we need to find $R \in \mathrm{SO}(3)$ so that $R\widetilde{A}R^t$ is in the real Schur form $B = \left[\begin{smallmatrix} 0 & -b & 0 \\ b & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right]$. Then similarity by the symplectic "double 3×3 rotation" $S = \left[\begin{smallmatrix} R & 0 \\ 0 & R \end{smallmatrix}\right] \in \mathrm{SpO}(6)$ gives $S\widetilde{W}S^t = \left[\begin{smallmatrix} B & 0 \\ 0 & -B \end{smallmatrix}\right] \in \mathbb{R}^{6\times 6}$, thus solving the 6×6 structured subproblem in a manner compatible with the sweep patterns described above.

The 3×3 rotation R needed to bring $\widetilde{A}$ into the real Schur form B can be conveniently computed using $\mathbb{H} \otimes \mathbb{H}$. Recall from §4.3 that every 3×3 rotation U can be represented as $u \otimes u$ for some unit quaternion u, where $\phi(u \otimes u)$ is the 4×4 matrix $\left[\begin{smallmatrix} 1 & 0 \\ 0 & U \end{smallmatrix}\right]$. Thus for 3×3 rotations to act naturally on 3×3 skew-symmetric matrices inside $\mathbb{H} \otimes \mathbb{H}$, we should embed 3×3 skew-symmetric matrices in an analogous way into the lower right corner of 4×4 matrices. The following result, a specialization of Proposition 1, shows the quaternion form of such matrices.

Proposition 5
Suppose $M \in \mathbb{R}^{3\times 3}$ is skew-symmetric, and $K(M)$ is the 4×4 skew-symmetric matrix obtained by embedding M in the lower right corner of the 4×4 zero matrix. That is,

$$K(M) = \left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & & & \\ 0 & & M & \\ 0 & & & \end{array}\right).$$

Then $K(M) = p \otimes 1 + 1 \otimes p$ for some $p \in \mathbb{P}$.

Using this result and the techniques described in [25] and [40], it is now straightforward to find the desired $R \in SO(3)$. First determine $q \in \mathbb{P}$ so that $K(\widetilde{A}) = q \otimes 1 + 1 \otimes q$. Next use Proposition 4 to find the unit quaternion u representing the rotation of $\mathbb{P} \cong \mathbb{R}^3$ that moves q into alignment with $k \in \mathbb{P}$. Then $\phi(u \otimes u) = \left[\begin{smallmatrix} 1 & 0 \\ 0 & R \end{smallmatrix}\right]$ gives the desired $R \in SO(3)$. Details of the computation can be found in the psuedo-code given in §7.3.3.

7.3.1 On Givens Transformations

Givens transformations, 2×2 rotations G embedded as principal submatrices of an $n \times n$ identity matrix, are a well-known and much-used tool in numerical linear algebra. Their primary use is to introduce zeroes into specified entries of a given matrix by premultiplication [24]. We may summarize their essential properties as follows:

1. $G \in SO(2)$
2. G is chosen so as to rotate a given nonzero vector $p \in \mathbb{R}^2$ into alignment with $e_1 = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$. In other words, $G \cdot p = \|p\| \cdot e_1 = \left[\begin{smallmatrix} \|p\| \\ 0 \end{smallmatrix}\right]$.
3. G can be *simply* computed from the vector p in a *uniform* manner for all nonzero $p \in \mathbb{R}^2$. Indeed, if $p = \left[\begin{smallmatrix} p_1 \\ p_2 \end{smallmatrix}\right]$, then

$$G = \frac{1}{\|p\|} \cdot \begin{bmatrix} p_1 & p_2 \\ -p_2 & p_1 \end{bmatrix}. \tag{28}$$

Embedding G as a principal submatrix into rows (and columns) r and s of the $n \times n$ identity always gives an $n \times n$ orthogonal matrix $G(r,s,n)$, but not necessarily a symplectic one. The matrix $G(r,s,n)$ will be symplectic orthogonal if and only if $n = 2m$ and $s = r + m$. In this case $G(r,s,n)$ is sometimes referred to as a symplectic Givens [8, 54].

Our aim is to use the convenient representation of 4×4 rotations in $\mathbb{H} \otimes \mathbb{H}$ to find analogues of Givens rotations that would play the same role for vectors $p \in \mathbb{R}^4$ that the usual Givens rotations do for vectors in $\mathbb{R}^2$. Such rotations might reasonably be referred to as "4×4 Givens" rotations. There are two natural choices:

$$G_4 = \overline{v} \otimes 1 = (v \otimes 1)^t \quad \text{and} \quad \widetilde{G}_4 = 1 \otimes v, \quad \text{where} \quad v = p/\|p\|. \tag{29}$$

¿From the characterization given in §4.3, it is clear that G_4 and $\widetilde{G}_4$ are indeed 4×4 rotations; since $G_4 p = (\overline{v} \otimes 1)(p) = (\overline{p}p1)/\|p\| = \|p\|$ (and similarly, that $\widetilde{G}_4 p = \|p\|$), we see that the analog of property (2) holds. Finally, the matrix formulas

$$G_4 = \frac{1}{\|p\|} \cdot \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \\ -p_2 & p_1 & p_4 & -p_3 \\ -p_3 & -p_4 & p_1 & p_2 \\ -p_4 & p_3 & -p_2 & p_1 \end{bmatrix} \quad \text{and} \quad \widetilde{G}_4 = \frac{1}{\|p\|} \cdot \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \\ -p_2 & p_1 & -p_4 & p_3 \\ -p_3 & p_4 & p_1 & -p_2 \\ -p_4 & -p_3 & p_2 & p_1 \end{bmatrix} \tag{30}$$

show that G_4 and $\widetilde{G}_4$ are no harder to compute than ordinary 2×2 Givens rotations, their rows being (up to a scaling) just signed permutations of the coordinates of p. Since all three "Givens properties" hold for G_4 and $\widetilde{G}_4$, we can justifiably call them both 4×4 Givens rotations.

Eqns. (29) and (30) together with the characterization of symplectic orthogonals in §4.4 imply that of these two choices G_4 and $\widetilde{G}_4$, only G_4 is symplectic. Hence only G_4 is useful for the purposes of this paper. In fact G_4 is the 4×4 symplectic Givens used in §7.3.2 and 7.3.3 to reduce 8×8 skew-symmetric skew-Hamiltonian matrices to block-diagonal form $\left[\begin{smallmatrix} \widetilde{A} & 0 \\ 0 & -\widetilde{A} \end{smallmatrix}\right]$.

Having seen that there are simple 4×4 analogs of the classical Givens rotations, it is natural to ask whether $n \times n$ analogs exist for any other n. There do exist such analogs; here is an example for $n = 8$, adapted from the discussion in [20].

$$G_8 = \frac{1}{\|p\|} \cdot \left[\begin{array}{cccc|cccc} p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 & p_8 \\ -p_2 & p_1 & p_4 & -p_3 & p_6 & -p_5 & -p_8 & p_7 \\ -p_3 & -p_4 & p_1 & p_2 & p_7 & p_8 & -p_5 & -p_6 \\ -p_4 & p_3 & -p_2 & p_1 & p_8 & -p_7 & p_6 & -p_5 \\ \hline -p_5 & -p_6 & -p_7 & -p_8 & p_1 & p_2 & p_3 & p_4 \\ -p_6 & p_5 & -p_8 & p_7 & -p_2 & p_1 & -p_4 & p_3 \\ -p_7 & p_8 & p_5 & -p_6 & -p_3 & p_4 & p_1 & -p_2 \\ -p_8 & -p_7 & p_6 & p_5 & -p_4 & -p_3 & p_2 & p_1 \end{array}\right]$$

Note that G_8 is orthogonal but not symplectic; we do not know if there are any symplectic 8×8 Givens rotations.

We do know, however, that this is the end of the line; only for $n = 2, 4,$ and 8 do there exist matrices that can reasonably be called $n \times n$ Givens rotations. The reason is essentially topological, and is connected with the classical question of the existence of vector fields on spheres. Let us briefly see why this is so, by re-examining the three essential properties of Givens rotations. ¿From the first property, that $G \in \mathrm{SO}(n)$, we know that the n rows of G

form an orthonormal set of vectors in $\mathbb{R}^n$. But then from property (2), that $G \cdot p$ should be $\|p\| \cdot e_1$, it follows that the last $n-1$ rows of G form an orthonormal basis for $p^\perp$. Thus the first row of G must be a scalar multiple of p. However, the only scalar multiple of p consistent with the condition $G \cdot p = \|p\| e_1$ is $p/\|p\|$; hence the first row of G must be $p/\|p\|$. Now the third condition's call for the entries of G to be simply computed from p in a uniform manner can be interpreted to mean that each entry of G should (at least) be a *continuous* function of the coordinates of p, defined for all nonzero $p \in \mathbb{R}^n$. Thus the last $n-1$ rows of G define a continuously-varying orthonormal basis for $p^\perp$, equivalently a set of $n-1$ orthonormal tangent vector fields on the unit sphere S^{n-1} in $\mathbb{R}^n$. But it is well-known [1, 5, 21, 33] that this is only possible if $n = 2, 4,$ or 8. Hence these are the only dimensions for which Givens rotations are possible.

7.3.2 Block-diagonalization of 8×8 subproblems via 4×4 symplectic Givens

In this section we see how the symplectic 4×4 Givens rotation described in §7.3.1 can be used to reduce 8×8 skew-symmetric skew-Hamiltonian matrices $\left[\begin{smallmatrix} A & C \\ C & -A \end{smallmatrix}\right]$ to block-diagonal form $\left[\begin{smallmatrix} \widetilde{A} & 0 \\ 0 & -\widetilde{A} \end{smallmatrix}\right]$. The four steps of this reduction are illustrated in Fig. 4 on a *general* (not necessarily skew-symmetric) 8×8 skew-Hamiltonian matrix W. For purposes of comparison, Van Loan's reduction described in [54] has been diagrammed in a similar fashion in Appendix B.

Let us consider the first step, displayed in Eqn.(31), in some detail. The four entries denoted by $\times$ in the first column of W form a vector $p \in \mathbb{R}^4$. This p is the "target vector" for the first step in the reduction.

$$\left(\begin{array}{cccc|cccc} \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\ \times & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\ \times & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \hline 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \times & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \times & \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & \cdot \end{array}\right) \xrightarrow{4\times 4 \text{ Givens}} \left(\begin{array}{cccc|cccc} \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\ \bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \hline 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & \cdot \end{array}\right) \tag{31}$$

Now construct the symplectic 4×4 Givens rotation G_4 for this vector p as in Eqn.(30) and embed it as a principal submatrix of the 8×8 identity in rows (and columns) $2, 3, 6$ and 7, the rows from which the entries of p were extracted, thus forming an 8×8 symplectic orthogonal matrix P. The result of the similarity PWP^t is to zero out three of the $\times$ entries, moving all of the norm of p into the entry labelled $\bullet$; preservation of the skew-Hamiltonian structure forces three additional zeroes to appear.

Subsequent steps of the reduction are depicted analogously in Fig 4. Entries marked $\times$ define a target vector, a 4×4 or 2×2 symplectic Givens is constructed from this target vector, and then appropriately embedded in I_8 to give an 8×8 symplectic orthogonal matrix $\widetilde{P}$. New zero entries produced by similarity with $\widetilde{P}$ are then added to the diagram; the entry receiving the norm of the target vector is denoted by $\bullet$.

Now suppose we applied this same four-step reduction to an 8×8 skew-Hamiltonian matrix that was also skew-symmetric. Since only symplectic orthogonal similarities are

$$\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\times & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
\times & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\times & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\times & \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
\xrightarrow{4\times 4 \text{ Givens}}
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & \cdot \\
0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)$$

$$\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\times & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
\times & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & \cdot \\
\times & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
\times & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
\xrightarrow{4\times 4 \text{ Givens}}
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & 0 & \cdot & \bullet & 0 & 0 \\
0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)$$

$$\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
0 & \times & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
0 & \times & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & 0 & \cdot & \bullet & 0 & 0 \\
0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \times & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \times & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
\xrightarrow{4\times 4 \text{ Givens}}
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
0 & \bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & 0 & \cdot & \bullet & 0 & 0 \\
0 & 0 & 0 & 0 & \cdot & \cdot & \bullet & 0 \\
0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)$$

$$\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
0 & \bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
0 & 0 & \times & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & 0 & \cdot & \bullet & 0 & 0 \\
0 & 0 & 0 & 0 & \cdot & \cdot & \bullet & 0 \\
0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & \times & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
\xrightarrow{2\times 2 \text{ Givens}}
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
0 & \bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
0 & 0 & \bullet & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & 0 & \cdot & \bullet & 0 & 0 \\
0 & 0 & 0 & 0 & \cdot & \cdot & \bullet & 0 \\
0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \bullet \\
0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)$$

Figure 4: Reduction of 8×8 skew-Hamiltonian using 4×4 Symplectic Givens

used, the final result will also be skew-symmetric. Thus we will have attained the desired block-diagonal form

$$\begin{pmatrix} \widetilde{A} & 0 \\ 0 & -\widetilde{A} \end{pmatrix} = \left(\begin{array}{cccc|cccc} 0 & \bullet & 0 & 0 & & & & \\ \bullet & 0 & \bullet & 0 & & & & \\ 0 & \bullet & 0 & \bullet & & & & \\ 0 & 0 & \bullet & 0 & & & & \\ \hline & & & & 0 & \bullet & 0 & 0 \\ & & & & \bullet & 0 & \bullet & 0 \\ & & & & 0 & \bullet & 0 & \bullet \\ & & & & 0 & 0 & \bullet & 0 \end{array}\right),$$

where $\widetilde{A}$ is not only skew-symmetric but also tridiagonal.

7.3.3 Algorithm to reduce 8×8 and 6×6 subproblems to structured Schur form

We now present pseudo-code for the finite step reduction of 8×8 and 6×6 skew-symmetric skew-Hamiltonian matrices to the canonical form specified in Theorem 1(d). We will use the following notation. If A is an $n \times n$ matrix, and v is a vector whose co-ordinates $v_1, v_2, \ldots, v_k$ are positive integers between 1 and n, then $A(v,i)$ denotes the vector $[A(v_1,i), \ldots, A(v_k,i)]$. Similarly, $A(v,v)$ denotes the principal submatrix whose diagonal elements are $A(v_1,v_1)$, $A(v_2,v_2), \ldots, A(v_k,v_k)$. Givens$(4,p)$ denotes the 4×4 symplectic Givens rotation G_4 given in Eqn. (30), while Givens(2, p) denotes the classical 2×2 Givens given in Eqn. (28). Finally, norm(p) denotes the 2-norm of the vector p. We reiterate that the code is organized for clarity, rather than optimal performance.

Algorithm 3 (8×8 Skew-symmetric skew-Hamiltonian)
INPUT: 8×8 *real skew-symmetric skew Hamiltonian matrix* W
OUTPUT: 8×8 *real symplectic orthogonal matrix* S *such that* SWS^t *is in the canonical form* $\left[\begin{smallmatrix} B & 0 \\ 0 & -B \end{smallmatrix}\right]$, *where* $B \in \mathbb{R}^{4 \times 4}$ *is a skew-symmetric matrix in real Schur form.*

```
Begin
  S = eye(8);

  % first column
  target = [2 4 6 8]; p = W(target, 1);
  R = eye(8); R(target, target) = Givens(4, p);
  W = R*W*Rᵗ; S = R;

  target = [2 3 6 7]; p = W(target, 1);
  R = eye(8); R(target, target) = Givens(4, p);
  W = R*W*Rᵗ; S = R*S;

  % second column
  target = [3 4 7 8]; p = W(target, 2);
  R = eye(8); R(target, target) = Givens(4, p);
```

```
  W = R*W*R^t; S = R*S;

  % third column
  target = [4 8]; p = W(target, 3);
  R = eye(8); R(target, target) = Givens(2,p);
  W = R*W*R^t; S = R*S;

  % W is now block-tridiagonal; finish with a double
  % 4x4 quaternion rotation, computed using [40, Eqns.  16, 17].
  p = [ .05*(W(2,1) + W(4,3))   0   0.5*W(3,2) ];
  np = norm(p); dp = np + p(1);
  P = (1/sqrt(2*np*dp))*[dp 0 -p(3) 0; 0 dp 0 p(3);...
                         p(3) 0 dp 0; 0 -p(3) 0 dp];
  q = [ .05*(W(4,3) - W(2,1))   0   0.5*W(3,2) ];
  nq = norm(q); dq = nq + q(1);
  Q = (1/sqrt(2*nq*dq))*[dq 0 q(3) 0; 0 dq 0 q(3);...
                         -q(3) 0 dq 0; 0 -q(3) 0 dq];
  R = P*Q; R = [R zeros(4); zeros(4) R];
  S = R*S;
End.
```

Algorithm 4 (6×6 SKEW-SYMMETRIC SKEW-HAMILTONIAN)
INPUT: 6×6 *real skew-symmetric skew Hamiltonian matrix* W
OUTPUT: 6×6 *real symplectic orthogonal matrix* S *such that* SWS^t *is in the canonical form* $\left[\begin{smallmatrix} B & 0 \\ 0 & -B \end{smallmatrix}\right]$, *where* $B \in \mathbb{R}^{3\times 3}$ *is a skew-symmetric matrix in real Schur form.*

```
Begin
  S = eye(6);

  % first column
  target = [2 3 5 6]; p = W(target, 1);
  R = eye(6); R(target, target) = Givens(4, p);
  W = R*W*R^t; S = R;

  % second column
  target = [3 6]; p = W(target, 2);
  R = eye(6); R(target, target) = Givens(2, p);
  W = R*W*R^t; S = R*S;

  % W is now block-tridiagonal; finish with a double
  % 3x3 quaternion rotation.
  q = [ .05*(W(3,2)   0   0.5*W(2,1) ];
  x = q(1)/(norm(q) + q(3));
  c = (1 - x*x)/(1 + x*x); s = 2*x/(1 + x*x);
  R = [c 0 -s; 0 1 0; s 0 c];
```

```
  R = [R zeros(3); zeros(3) R];
  S = R*S;
End.
```

8 Numerical Results

We present a brief set of numerical experiments to demonstrate the effectiveness of our algorithms. All computations were done using MATLAB[4] Version 5.3.0 on a Sun Sparc10 with IEEE double-precision arithmetic and machine precision $\epsilon = 2.2204 \times 10^{-16}$. As a stopping criterion we choose $\mathrm{Off}(H)/\|H\|_F < tol$ where $\mathrm{Off}(H)$ is the appropriate off-diagonal norm for the case under consideration, $\|H\|_F$ is the Frobenius norm of H, and $tol = \epsilon\|H\|_F$.

For each of the four doubly structured classes, and for each $n = 20, 25, \ldots, 100$, the algorithms were run on 100 random matrices of size $2n \times 2n$ with entries normally distributed. For the sake of brevity, and since the remaining cases are similar, only the average results when $2n = 50, 100, 150, 200$, and the ordering is row-cyclic are reported in Tables 4 - 7.

The computational results can be summarized as follows:

- The methods always converged, and the off-diagonal norm always decreased monotonically. The convergence rate was initially linear, but asymptotically quadratic. This is shown in Fig. 5 using a sample 200×200 matrix from each of the four classes.

- Fig. 6 suggests that the number of sweeps needed for convergence depends only on n, the matrix size, and increases roughly as $\log n$. The standard deviation of the average number of sweeps was consistently very low — between 0 and 0.5. This leads to an *a priori* stopping criterion, which is an important consideration on parallel architectures — a stopping criterion that depends on global knowledge of the matrix elements is too expensive to implement.

- The eigenvalues $\{\lambda_i^{jac}\}_1^{2n}$ computed by our algorithms were quite accurate. As the matrices are always either symmetric or skew-symmetric, the eigenvalues have condition number equal to 1, are all real or all pure imaginary and can be easily sorted and compared with the sorted eigenvalues computed by MATLAB's `eig` function. The maximum relative error, $\mathrm{releig} = \max_j |\lambda_j^{eig} - \lambda_j^{jac}|/|\lambda_j^{eig}|$ was of the order of 10^{-13} as shown in the last column of Tables 4 - 7.

- The computed symplectic orthogonal transformations S from which the eigenvectors/invariant subspaces can be obtained were both symplectic as well as orthogonal to within 2.3×10^{-14}, as shown in columns 2 and 3 of Tables 4 - 7. Column 4 records a check on the block structure of $S = \left[\begin{smallmatrix} U & V \\ -V & U \end{smallmatrix}\right]$. Here $\mathrm{block} = \|\{S(1:n, 1:n) - S(n+1:2n, n+1:2n)\}\| + \|\{S(1:n, n+1:2n) + S(n+1:2n, 1:n)\}\|$; both terms in this sum had about the same size. Eigenvectors computed by MATLAB's `eig` function cannot be directly compared to the symplectic bases obtained by our algorithms.

[4]MATLAB is a trademark of The MathWorks, Inc.

For example, every eigenspace of any skew-Hamiltonian matrix is at least 2 dimensional, so MATLAB's eigenvectors are extremely unlikely to yield a symplectic basis for comparison.

2n	reloff	$\|S^TJS-J\|$	$\|S^TS-I\|$	block	releig
50	$1.13*10^{-15}$	$1.93*10^{-14}$	$1.96*10^{-14}$	$2.08*10^{-15}$	$2.00*10^{-14}$
100	$6.72*10^{-16}$	$4.17*10^{-14}$	$4.20*10^{-14}$	$3.17*10^{-15}$	$4.24*10^{-14}$
150	$3.27*10^{-15}$	$6.53*10^{-14}$	$6.57*10^{-14}$	$4.03*10^{-15}$	$6.57*10^{-14}$
200	$7.72*10^{-15}$	$8.89*10^{-14}$	$8.94*10^{-14}$	$4.71*10^{-15}$	$8.87*10^{-14}$

Table 4: $2n \times 2n$ symmetric Hamiltonian matrices

2n	reloff	$\|S^TJS-J\|$	$\|S^TS-I\|$	block	releig
50	$6.11*10^{-16}$	$6.63*10^{-15}$	$6.83*10^{-15}$	$1.64*10^{-15}$	$7.86*10^{-15}$
100	$4.27*10^{-15}$	$1.14*10^{-14}$	$1.17*10^{-14}$	$2.47*10^{-15}$	$1.39*10^{-14}$
150	$1.26*10^{-15}$	$1.80*10^{-14}$	$1.82*10^{-14}$	$3.14*10^{-15}$	$9.66*10^{-15}$
200	$1.71*10^{-15}$	$2.24*10^{-14}$	$2.28*10^{-14}$	$3.69*10^{-15}$	$1.48*10^{-14}$

Table 5: $2n \times 2n$ skew-symmetric Hamiltonian matrices

2n	reloff	$\|S^TJS-J\|$	$\|S^TS-I\|$	block	releig
50	$5.43*10^{-16}$	$6.69*10^{-15}$	$6.89*10^{-15}$	$1.63*10^{-15}$	$5.08*10^{-14}$
100	$4.54*10^{-15}$	$1.18*10^{-14}$	$1.21*10^{-14}$	$2.47*10^{-15}$	$4.81*10^{-14}$
150	$1.03*10^{-15}$	$1.77*10^{-14}$	$1.80*10^{-14}$	$3.13*10^{-15}$	$1.19*10^{-13}$
200	$1.73*10^{-15}$	$2.23*10^{-14}$	$2.26*10^{-14}$	$3.68*10^{-15}$	$2.21*10^{-13}$

Table 6: $2n \times 2n$ symmetric skew-Hamiltonian matrices

2n	reloff	$\|S^TJS-J\|$	$\|S^TS-I\|$	block	releig
50	$1.07*10^{-15}$	$8.37*10^{-15}$	$8.69*10^{-15}$	$2.20*10^{-15}$	$6.93*10^{-15}$
100	$4.17*10^{-15}$	$1.55*10^{-14}$	$1.59*10^{-14}$	$3.48*10^{-15}$	$1.53*10^{-14}$
150	$2.19*10^{-15}$	$2.27*10^{-14}$	$2.32*10^{-14}$	$4.36*10^{-15}$	$2.01*10^{-14}$
200	$1.06*10^{-14}$	$2.98*10^{-14}$	$3.04*10^{-14}$	$5.18*10^{-15}$	$3.47*10^{-14}$

Table 7: $2n \times 2n$ skew-symmetric skew-Hamiltonian matrices

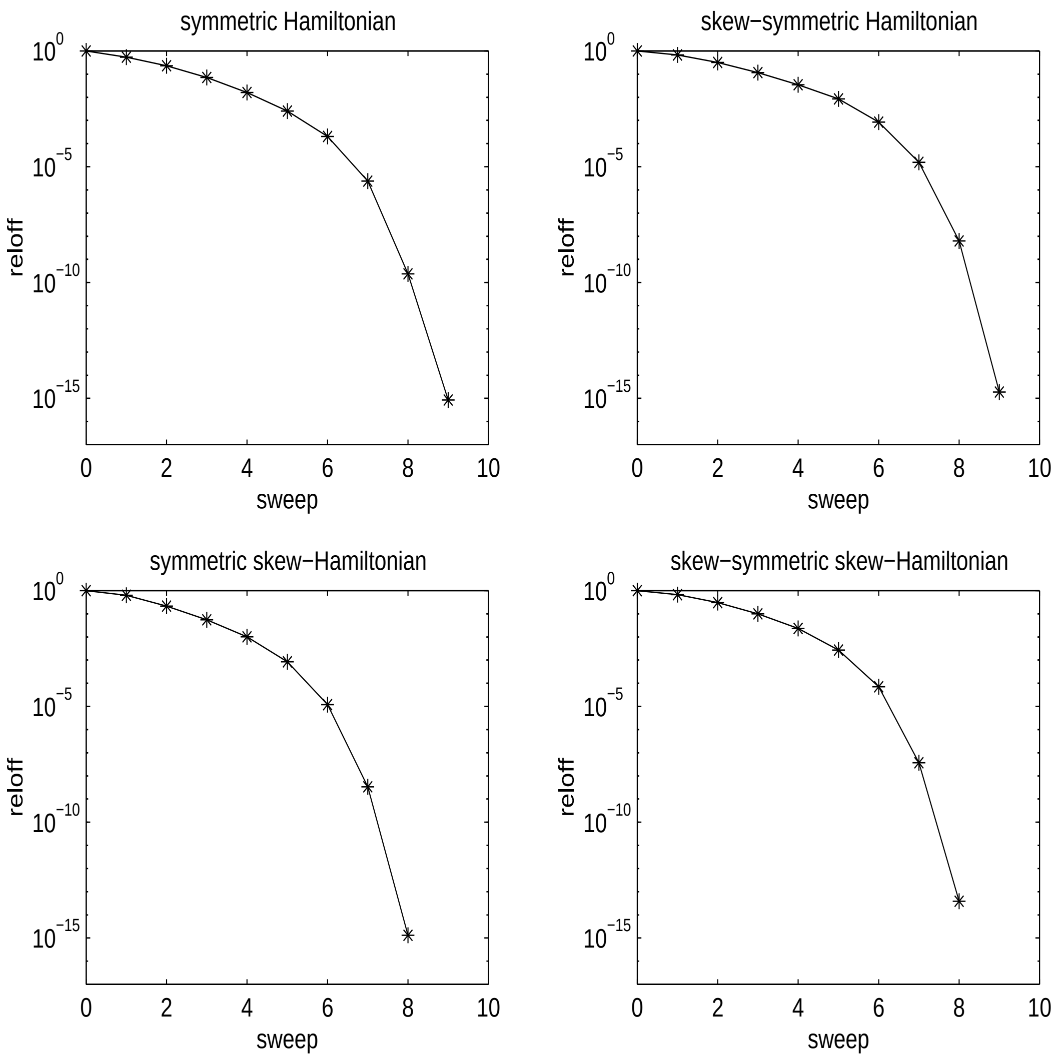

Figure 5: Typical convergence behavior of 200×200 matrices

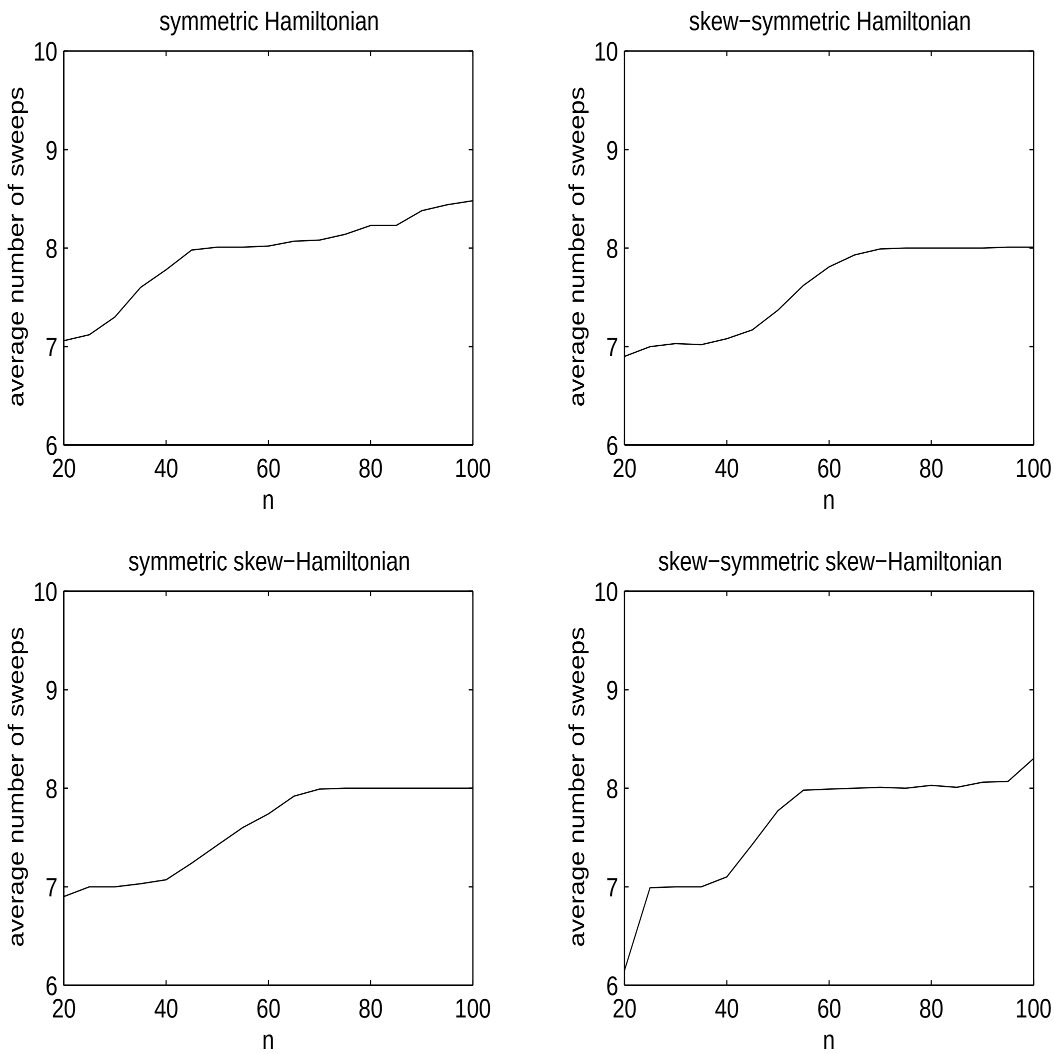

Figure 6: Average number of sweeps for convergence for $2n \times 2n$ matrices

9 $2n \times 2n$ Hamiltonians and skew-Hamiltonians

Since we have a quaternion representation of 4×4 Hamiltonian and skew-Hamiltonian matrices, it is natural to apply the techniques of this paper to these more general eigenproblems. Unfortunately, there are difficulties with this plan; we briefly discuss them here.

If W is a 4×4 skew-Hamiltonian matrix, then we can write (see Table 3 in §4.5)

$$W = b\,(1 \otimes 1) + p \otimes j + 1 \otimes q\,,$$

where $b \in \mathbb{R}$, $p \in \mathbb{P}$ and $q = c\,i + dk \in \mathbb{P}$. We can put W into (2×2)–block-diagonal form by constructing a symplectic orthogonal similarity $S = \phi(x \otimes y)$ as follows. Using Proposition 4, choose $\phi(x \otimes x)$ so that p is aligned with j. Next, since q lies in the ik-plane, we can choose a rotation $\phi(y \otimes y)$ with axis j that aligns q with i. Now SWS^t is a linear combination of just $1 \otimes 1$, $j \otimes j$ and $1 \otimes i$. The quaternion basis shows this to be a (2×2)–block-diagonal skew-Hamiltonian matrix, with its $(1,2)$ and $(2,1)$ entries being equal and of opposite sign:

$$SWS^t = \begin{bmatrix} b+|p| & |q| & 0 & 0 \\ -|q| & b-|p| & 0 & 0 \\ 0 & 0 & b+|p| & -|q| \\ 0 & 0 & |q| & b-|p| \end{bmatrix}$$

The Jacobi algorithm for a $2n \times 2n$ skew-Hamiltonian matrix W based on this direct solution can at best be expected to put W into block-diagonal form $\left[\begin{smallmatrix} A & 0 \\ 0 & A^t \end{smallmatrix}\right]$, where the blocks are $n \times n$. But the argument in §7.2 shows that this is unlikely to be achieved; experimentation confirms this for $n \geq 4$. However, there is a theoretical question that is independent of the failure of this Jacobi algorithm for general n: is it possible to block-diagonalize any real skew-Hamiltonian matrix by a symplectic *orthogonal* similarity? In [22], it is shown that any $2n \times 2n$ real skew-Hamiltonian can always be symplectically block-diagonalized, but it is still an open question whether this can also be achieved by a symplectic orthogonal similarity. Note that the quaternions settle the question when $n = 2$, by explicitly providing such a symplectic orthogonal matrix.

The Hamiltonian case presents other difficulties. An examination of the quaternion representation of a 4×4 Hamiltonian matrix shows that this eigenproblem cannot be solved solely by the techniques used in this paper. Designing a structure-preserving Jacobi algorithm for this class along the lines of the other algorithms in this paper thus remains under investigation.

10 Concluding Summary

We have shown that several classes of real 4×4 matrices can be naturally embedded in the quaternion tensor algebra, $\mathbb{H} \otimes \mathbb{H}$. These include Hamiltonian, skew-Hamiltonian and symplectic orthogonal matrices, in addition to the previously known embeddings of 4×4 symmetric, skew-symmetric and orthogonal matrices.

The advantage of the quaternion representation is twofold. Algebraic eigenproblems in $\mathbb{R}^{4\times 4}$ can be converted into geometric problems in $\mathbb{R}^3$, and the complicated action of 4×4

structure-preserving similarities reduced to the action of easily understood three-dimensional rotations. This insight leads to direct solutions of the 4×4 eigenproblems for four classes of doubly structured matrices: symmetric or skew-symmetic Hamiltonian, and symmetric or skew-symmetic skew-Hamiltonian. Structure-preserving Jacobi algorithms for $2n \times 2n$ matrices based on these 4×4 solutions converge in practise for three out of the four classes. For the class of $2n \times 2n$ skew-symmetric skew-Hamiltonian matrices, the 4×4 based algorithm does not converge. However, we show in this case that the 8×8 eigenproblem can also be directly solved; the resulting Jacobi algorithm then does converge in practise.

We have discussed the extension of the classical Givens rotations to higher dimensions, and shown how to construct such rotations in four dimensions. We have also shown how 4×4 symplectic Givens can be built, and used them in our solution to the 8×8 skew-symmetric skew-Hamiltonian eigenproblem.

The algorithms described in this paper are completely structure-preserving, inherently parallelizable and asympotically quadratically convergent; symplectic orthogonal bases for all the invariants subspaces are computed.

APPENDIX A — The Quaternion Basis for $\mathbb{R}^{4\times 4}$

$$
\begin{array}{c|ccc}
\underset{1\otimes 1}{\begin{pmatrix} 1&0&0&0\\ 0&1&0&0\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix}} &
\underset{1\otimes i}{\begin{pmatrix} 0&1&0&0\\ -1&0&0&0\\ 0&0&0&-1\\ 0&0&1&0 \end{pmatrix}} &
\underset{1\otimes j}{\begin{pmatrix} 0&0&1&0\\ 0&0&0&1\\ -1&0&0&0\\ 0&-1&0&0 \end{pmatrix}} &
\underset{1\otimes k}{\begin{pmatrix} 0&0&0&1\\ 0&0&-1&0\\ 0&1&0&0\\ -1&0&0&0 \end{pmatrix}} \\
\hline
\underset{i\otimes 1}{\begin{pmatrix} 0&-1&0&0\\ 1&0&0&0\\ 0&0&0&-1\\ 0&0&1&0 \end{pmatrix}} &
\underset{i\otimes i}{\begin{pmatrix} 1&0&0&0\\ 0&1&0&0\\ 0&0&-1&0\\ 0&0&0&-1 \end{pmatrix}} &
\underset{i\otimes j}{\begin{pmatrix} 0&0&0&-1\\ 0&0&1&0\\ 0&1&0&0\\ -1&0&0&0 \end{pmatrix}} &
\underset{i\otimes k}{\begin{pmatrix} 0&0&1&0\\ 0&0&0&1\\ 1&0&0&0\\ 0&1&0&0 \end{pmatrix}} \\
\underset{j\otimes 1}{\begin{pmatrix} 0&0&-1&0\\ 0&0&0&1\\ 1&0&0&0\\ 0&-1&0&0 \end{pmatrix}} &
\underset{j\otimes i}{\begin{pmatrix} 0&0&0&1\\ 0&0&1&0\\ 0&1&0&0\\ 1&0&0&0 \end{pmatrix}} &
\underset{j\otimes j}{\begin{pmatrix} 1&0&0&0\\ 0&-1&0&0\\ 0&0&1&0\\ 0&0&0&-1 \end{pmatrix}} &
\underset{j\otimes k}{\begin{pmatrix} 0&-1&0&0\\ -1&0&0&0\\ 0&0&0&1\\ 0&0&1&0 \end{pmatrix}} \\
\underset{k\otimes 1}{\begin{pmatrix} 0&0&0&-1\\ 0&0&-1&0\\ 0&1&0&0\\ 1&0&0&0 \end{pmatrix}} &
\underset{k\otimes i}{\begin{pmatrix} 0&0&-1&0\\ 0&0&0&1\\ -1&0&0&0\\ 0&1&0&0 \end{pmatrix}} &
\underset{k\otimes j}{\begin{pmatrix} 0&1&0&0\\ 1&0&0&0\\ 0&0&0&1\\ 0&0&1&0 \end{pmatrix}} &
\underset{k\otimes k}{\begin{pmatrix} 1&0&0&0\\ 0&-1&0&0\\ 0&0&-1&0\\ 0&0&0&1 \end{pmatrix}}
\end{array}
$$

The following figures diagram Van Loan's reduction procedure [54] for a general 8×8 skew-Hamiltonian matrix, illustrated in the same style as in §7.3.2. Here the label *House* indicates similarity by an appropriate "double Householder" matrix of the form $\left[\begin{smallmatrix} H & 0 \\ 0 & H \end{smallmatrix}\right]$, where H is an ordinary 4×4 Householder matrix chosen to act on the target vector indicated by $\times$'s. Note that this doubling of H in $\left[\begin{smallmatrix} H & 0 \\ 0 & H \end{smallmatrix}\right]$ is necessary to ensure that the similarity is symplectic as well as orthogonal.

$$
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\times & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\times & \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
\times & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
\xrightarrow{\text{House}}
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & \cdot & 0 & 0 & \cdot & \cdot & \cdot & \cdot \\
\bullet & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
$$

$$
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\times & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & \cdot & 0 & 0 & \cdot & \cdot & \cdot & \cdot \\
\times & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
\xrightarrow{2\times 2 \text{ Givens}}
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
$$

$$
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\times & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
\times & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
\times & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
\xrightarrow{\text{House}}
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & 0 & \cdot & \bullet & 0 & 0 \\
0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
$$

$$
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & 0 & \cdot & \bullet & 0 & 0 \\
0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \times & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \times & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
\xrightarrow{\text{House}}
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & 0 & \cdot & \bullet & 0 & 0 \\
0 & 0 & \cdot & 0 & \cdot & \cdot & \cdot & \cdot \\
0 & \bullet & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
$$

$$
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
0 & \times & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & 0 & \cdot & \bullet & 0 & 0 \\
0 & 0 & \cdot & 0 & \cdot & \cdot & \cdot & \cdot \\
0 & \times & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
\xrightarrow{2\times 2 \text{ Givens}}
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
0 & \bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & 0 & \cdot & \bullet & 0 & 0 \\
0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
$$

$$
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
0 & \times & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
0 & \times & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & 0 & \cdot & \bullet & 0 & 0 \\
0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
\xrightarrow{\text{House}}
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
0 & \bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & 0 & \cdot & \bullet & 0 & 0 \\
0 & 0 & 0 & 0 & \cdot & \cdot & \bullet & 0 \\
0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & \cdot & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
$$

$$
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
0 & \bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
0 & 0 & \times & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & 0 & \cdot & \bullet & 0 & 0 \\
0 & 0 & 0 & 0 & \cdot & \cdot & \bullet & 0 \\
0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & \times & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
\xrightarrow{2\times 2 \text{ Givens}}
\left(\begin{array}{cccc|cccc}
\cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\
\bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \cdot \\
0 & \bullet & \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\
0 & 0 & \bullet & \cdot & \cdot & \cdot & \cdot & 0 \\
\hline
0 & 0 & 0 & 0 & \cdot & \bullet & 0 & 0 \\
0 & 0 & 0 & 0 & \cdot & \cdot & \bullet & 0 \\
0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \bullet \\
0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot
\end{array}\right)
$$

11 Acknowledgements

The second and third authors thank Volker Mehrmann and the numerical linear algebra group at Technische Universität Chemnitz for their interest in this work, and their hospitality and support during a recent visit. Thanks are also due to Nick Higham for his interest in and comments on this paper.

References

[1] J. F. Adams, *Vector fields on spheres*, Bulletin of the American Mathematical Society, 68 (1962), pp. 39–41.

[2] P. Benner, V. Mehrmann, and H. Xu, *A new method for computing the stable invariant subspace of a real Hamiltonian matrix*, J. Comput. Appl. Math., 86 (1997), pp. 17–43.

[3] ——, *A numerically stable, structure preserving method for computing the eigenvalues of real Hamiltonian or symplectic pencils*, Numer. Math., 78 (1998), pp. 329–358.

[4] G. Birkhoff and S. MacLane, *A Survey of Modern Algebra*, Macmillan, New York, 1977.

[5] R. Bott and J. Milnor, *On the parallelizability of the spheres*, Bulletin of the Amer. Math. Society, 64 (1958), pp. 87–89.

[6] T. Bröcker and T. tom Dieck, *Representations of Compact Lie Groups*, Springer-Verlag, New York, 1985.

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

[8] A. Bunse-Gerstner, R. Byers, and V. Mehrmann, *A chart of numerical methods for structured eigenvalue problems.*, SIAM J. Matrix Anal. Applic., 13 (1992), pp. 419–453.

[9] A. Bunse-Gerstner and H. Fassbender, *A Jacobi-like method for solving algebraic Riccati equations on parallel computers*, IEEE Trans. Automat. Contr., 42 (1997), pp. 1071–1084.

[10] R. Byers, *A Hamiltonian-QR algorithm*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 212–229.

[11] ——, *A Hamiltonian-Jacobi algorithm*, IEEE Trans. Automat. Contr., AC-35 (1990), pp. 566–570.

[12] A. Cayley, *On certain results relating to quaternions*, in The Collected Mathematical Papers, vol. 1, Johnson Reprint Company, New York, 1963, pp. 123–126. Originally published in The Philos. Mag., **26** (1845), pp. 141–145.

[13] ——, *Recherches ultérieures sur les déterminants gauches*, in The Collected Mathematical Papers, vol. 2, Johnson Reprint Company, New York, 1963, pp. 202–215. Originally published in J. Reine Angew. Math., **50** (1855), pp. 299-313.

[14] M. J. Crowe, *A History of Vector Analysis*, University of Notre Dame Press, Notre Dame, 1967.

[15] J. Dongarra, J. R. Gabriel, D. D. Kölling, and J. H. Wilkinson, *The eigenvalue problem for Hermitian matrices with time reversal symmetry*, Lin. Alg. Appl., 60 (1984), pp. 27–42.

[16] H. D. Ebbinghaus, H. Hermes, F. Hirzebruch, M. Keocher, K. Mainzer, J. Neukirch, A. Prestel, and R. Remmert, *Numbers*, Springer-Verlag, New York, 1991.

[17] P. J. Eberlein, *A Jacobi-like method for the automatic computation of eigenvalues and eigenvectors of an arbitrary matrix*, J. Soc. Indust. Appl. Math, 10 (1962), pp. 74–88.

[18] ——, *On one-sided Jacobi methods for parallel computation*, SIAM J. Alg. Disc. Meth., 8 (1987), pp. 790–796.

[19] P. J. Eberlein and H. Park, *Efficient implementation of Jacobi algorithms and Jacobi sets on distributed memory architectures*, J. Paral. Dist. Comput., 8 (1990), pp. 358–366.

[20] B. Eckmann, *Stetige Lösungen linearer Gleichungssysteme*, Commentarii Math. Helvetica, 15 (1942), pp. 318–339.

[21] ——, *Topology, algebra, analysis—relations and missing links*, Notices of the American Mathematical Society, 46 (1999), pp. 520–527.

[22] H. Fassbender, D. S. Mackey, N. Mackey, and H. Xu, *Hamiltonian square roots of skew-Hamiltonian matrices*, Linear Algebra Appl., 287 (1999), pp. 125–159.

[23] H. H. Goldstine and L. P. Horwitz, *A procedure for the diagonalization of normal matrices*, J. Assoc. Comp. Mach., 6 (1959), pp. 176–195.

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

[25] D. Hacon, *Jacobi's method for skew-symmetric matrices*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 619–628.

[26] W. R. Hamilton, *On a general method in dynamics*, in The Mathematical Papers of Sir William Rowan Hamilton, A. W. Conway and A.J.McConnell, eds., vol. 2, Cambridge University Press, Cambridge, 1940, pp. 103–161. Originally published in *Phil. Trans. Royal Society of London* (part II for 1834), pp. 247–308; communicated April 1834.

[27] ——, *Second essay on a general method in dynamics*, in The Mathematical Papers of Sir William Rowan Hamilton, A. W. Conway and A. McConnell, eds., vol. 2, Cambridge University Press, Cambridge, 1940, pp. 162–211. Originally published in *Phil. Trans. Royal Society of London* (part I for 1835), pp. 95–144; communicated October 1834.

[28] ——, *On quaternions, or a new system of imaginaries in algebra; with some geometrical illustrations*, in The Collected Mathematical Papers, H. Halberstam and R. E. Ingram, eds., vol. 3, Cambridge University Press, Cambridge, 1967, pp. 355–362. Originally published in *Proc. Royal Irish Academy* **3** (1847), pp. 1–16; communicated November 1844.

[29] T. L. Hankins, *Sir William Rowan Hamilton*, Johns Hopkins University Press, Baltimore, 1980.

[30] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1990.

[31] C. G. J. Jacobi, *Über die Reduction der Integration der partiellen Differentialgleichungen erster Ordnung zwischen irgend einer Zahl Variabeln auf die Integration eines einzigen Systemes gewöhnlicher Differentialgleichungen*, J. Reine Angew. Math., 17 (1837), pp. 97–162.

[32] ——, *Über ein leichtes Verfahren die in der Theorie der Säcularstörungen vorkommenden Gleichungen numerisch aufzulösen*, J. Reine Angew. Math., 30 (1846), pp. 51–94.

[33] M. Kervaire, *Non-parallelizability of the n-sphere for $n > 7$*, Proceedings of the National Academy of Sciences, 44 (1958), pp. 280–283.

[34] P. Lancaster and L. Rodman, *Algebraic Riccati Equations*, Oxford University Press, Oxford, 1995.

[35] C. Lanczos, *William Rowan Hamilton—an appreciation*, American Scientist, 55 (1967), pp. 129–143.

[36] A. J. Laub, *A Schur method for solving algebraic Riccati equations*, IEEE Trans. Auto. Control, AC–24 (1979), pp. 913–921.

[37] W.-W. Lin, V. Mehrmann, and H. Xu, *Canonical forms for Hamiltonian and symplectic matrices and pencils*, Linear Algebra Appl., (to appear).

[38] F. T. Luk and H. Park, *On parallel Jacobi orderings*, SIAM J. Sci. Stat. Comp., 10 (1989), pp. 18–26.

[39] D. S. Mackey and N. Mackey, *Symplectic orthogonal canonical forms for doubly structured matrices.* In preparation.

[40] N. Mackey, *Hamilton and Jacobi meet again: Quaternions and the eigenvalue problem*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 421–435.

[41] ——, *Quaternions and the eigenproblem: A new Jacobi algorithm and its convergence*, PhD thesis, State University of New York at Buffalo, 1995.

[42] M. Mantharam and P. J. Eberlein, *New Jacobi-sets for parallel computations*, Parallel Computing, 19 (1993), pp. 437–454.

[43] M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Prindle, Weber & Schmidt, Inc., Boston, 1964.

[44] W. F. Mascarenhas, *On the convergence of the Jacobi method for arbitrary orderings*, PhD thesis, Massachusetts Institute of Technology, Cambridge, 1991.

[45] V. Mehrmann, *The Autonomous Linear Quadratic Control Problem: Theory and Numerical Solution*, Springer-Verlag, Berlin/Heidelberg, 1991.

[46] J. C. Nash, *A one-sided transformation method for the singular value decomposition and algebraic eigenproblem*, Computing J., 18 (1977), pp. 74–76.

[47] J. Olson, H. J. A. Jensen, and P. Jørgensen, *Solution of large matrix equations which occur in response theory*, J. Comp. Physics, 74 (1988), pp. 265–282.

[48] M. Paardekooper, *An eigenvalue algorithm for skew-symmetric matrices*, Numer. Math., 17 (1971), pp. 189–202.

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

[50] R. S. Pierce, *Associative Algebras*, Springer-Verlag, New York, 1982.

[51] E. G. Rees, *Notes on Geometry*, Springer Verlag, Berlin, 1983.

[52] A. H. Sameh, *On Jacobi and Jacobi-like algorithms for a parallel computer*, Math. of Comp., 25 (1971), pp. 579–590.

[53] G. M. Shroff, *Parallel Jacobi-like Algorithms for the Algebraic Eigenvalue Problem*, PhD thesis, Rensselaer Polytechnic Institute, 1990.

[54] C. Van Loan, *A symplectic method for approximating all the eigenvalues of a Hamiltonian matrix*, Linear Algebra Appl., 61 (1984), pp. 233–251.

Berichte aus der Technomathematik

ISSN 1435-7968

`http://www.math.uni-bremen.de/zetem/berichte.html`

— Vertrieb durch den Autor —

Reports **Stand: 6. Oktober 1999**

98–01. Peter Benner, Heike Faßbender:
An Implicitly Restarted Symplectic Lanczos Method for the Symplectic Eigenvalue Problem, Juli 1998.

98–02. Heike Faßbender:
Sliding Window Schemes for Discrete Least-Squares Approximation by Trigonometric Polynomials, Juli 1998.

98–03. Peter Benner, Maribel Castillo, Enrique S. Quintana-Ortí:
Parallel Partial Stabilizing Algorithms for Large Linear Control Systems, Juli 1998.

98–04. Peter Benner:
Computational Methods for Linear–Quadratic Optimization, August 1998.

98–05. Peter Benner, Ralph Byers, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Solving Algebraic Riccati Equations on Parallel Computers Using Newton's Method with Exact Line Search, August 1998.

98–06. Lars Grüne, Fabian Wirth:
On the rate of convergence of infinite horizon discounted optimal value functions, November 1998.

98–07. Peter Benner, Volker Mehrmann, Hongguo Xu:
A Note on the Numerical Solution of Complex Hamiltonian and Skew-Hamiltonian Eigenvalue Problems, November 1998.

98–08. Eberhard Bänsch, Burkhard Höhn:
Numerical simulation of a silicon floating zone with a free capillary surface, Dezember 1998.

99–01. Heike Faßbender:
The Parameterized SR Algorithm for Symplectic (Butterfly) Matrices, Februar 1999.

99–02. Heike Faßbender:
Error Analysis of the symplectic Lanczos Method for the symplectic Eigenvalue Problem, März 1999.

99–03. Eberhard Bänsch, Alfred Schmidt:
Simulation of dendritic crystal growth with thermal convection, März 1999.

99–04. Eberhard Bänsch:
Finite element discretization of the Navier-Stokes equations with a free capillary surface, März 1999.

99–05. Peter Benner:
Mathematik in der Berufspraxis, Juli 1999.

99–06. Andrew D.B. Paice, Fabian R. Wirth:
Robustness of nonlinear systems and their domains of attraction, August 1999.

99–07. Peter Benner, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Balanced Truncation Model Reduction of Large-Scale Dense Systems on Parallel Computers, September 1999.

99–08. Ronald Stöver:
Collocation methods for solving linear differential-algebraic boundary value problems, September 1999.

99–09. Huseyin Akcay:
Modelling with Orthonormal Basis Functions, September 1999.

99–10. Heike Faßbender, D. Steven Mackey, Niloufer Mackey:
Hamilton and Jacobi come full circle: Jacobi algorithms for structured Hamiltonian eigenproblems , Oktober 1999.