HAMILTON AND JACOBI MEET AGAIN: QUATERNIONS AND THE EIGENVALUE PROBLEM *

NILOUFER MACKEY†

May 1993; Revised September 1993

Abstract. The algebra isomorphism between $\mathcal{M}_4(\mathcal{R})$ and $\mathcal{H}\otimes\mathcal{H}$, where $\mathcal{H}$ is the algebra of quaternions, has unexpected computational payoff: it helps construct an orthogonal similarity that 2×2 block-diagonalizes a 4×4 symmetric matrix. Replacing plane rotations with these more powerful 4×4 rotations leads to a quaternion-Jacobi method in which the 'weight' of 4 elements (in a 2×2 block) is transferred all at once onto the diagonal. Quadratic convergence sets in sooner, and the new method requires at least one fewer sweep than plane-Jacobi methods. An analogue of the sorting angle for plane rotations is developed for these 4×4 rotations.

Key words. eigenvalues, symmetric matrix, Jacobi method, quaternion, tensor product

AMS subject classifications. 65F15, 15A18, 15A21, 15A69

1. Introduction. One hundred and fifty years ago, on 16 October 1843, W. R. Hamilton carved the equations defining the algebra of quaternions on the stones of Brougham Bridge, Dublin [11], [26], [8]. Two years later, in an unrelated piece of work, C. J. G. Jacobi described an iterative method for solving the eigenproblem of an $n \times n$ symmetric matrix. This method appeared the following year in Crelle's journal [17][1].

To the student of mechanics, the names Hamilton and Jacobi are already closely linked: in 1837, Jacobi extended Hamilton's work in dynamics, giving rise to what is today known as the Hamilton-Jacobi theory[2]. Now, a century and a half later, we bring the work of these two men together in a new way: we show how Hamilton's quaternions enhance Jacobi's algorithm for solving the symmetric eigenproblem. Jacobi diagonalizes a symmetric matrix by performing a sequence of orthogonal similarity transformations. Each transformation is a plane rotation, chosen so that the induced similarity diagonalizes some 2×2 principal submatrix, moving the weight of the annihilated elements onto the diagonal. Can one explicitly specify an orthogonal transformation that diagonalizes a larger submatrix?

We show that the algebra isomorphism between $\mathcal{M}_4(\mathcal{R})$ and $\mathcal{H}\otimes\mathcal{H}$, where $\mathcal{H}$ is the algebra of quaternions, has direct and unexpected computational payoff: it leads to the construction of an orthogonal similarity to (2×2)–block diagonalize a 4×4 symmetric matrix. The quaternion-Jacobi method thus obtained produces four times as many zeroes at each step, and hence converges in fewer iterations. The price we pay for this abundance of zeroes is the cost of computing one left–right singular vector pair of a 3×3 matrix, whose entries are simple linear combinations of the entries of the 4×4 symmetric matrix being diagonalized. The quaternion-Jacobi method is at least quadratically convergent, and experimental evidence strongly suggests that it requires at least one fewer sweep than the traditional Jacobi method.

---

* An earlier version of this paper won the 1993 SIAM Student Paper Competition.

† Department of Computer Science, State University of New York at Buffalo, Amherst, NY 14260. (mackey@cs.buffalo.edu).

1 The key ideas for the method were introduced by Jacobi in an earlier paper [16]. Thanks to Sven Hammarling for pointing this out.

2 The two mathematicians first met in 1842 at the meeting of the British Association for the Advancement of Science held in Manchester, England. According to Hankins [13], Jacobi refers to Hamilton as "the illustrious Astronomer Royal of Dublin" and later, as "the Lagrange" of Ireland.

There is renewed interest in Jacobi-type methods today, because they are easily parallelizable, and compute small eigenvalues with greater accuracy than the QR method [7], [21].

2. The Quaternions. Recently, Derek Hacon [10] showed that one can use quaternions to construct an orthogonal similarity transformation which will directly reduce any given 4×4 skew-symmetric matrix to its real Schur (2×2 block diagonal) form, and thereby obtain a Jacobi-type algorithm for $n \times n$ skew-symmetric matrices. The method is intriguing, and it is natural to ask whether a 4×4 symmetric matrix can be (block) diagonalized in a similar manner. To uncover the symmetric algorithm, we start with some algebraic preliminaries.

The quaternions, $\mathcal{H}$, are a four-dimensional, associative, but non-commutative division algebra over $\mathcal{R}$, with the standard basis $\{1, i, j, k\}$. Multiplication is determined by the rules

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

which imply $jk = -kj = i, ki = -ik = j, ij = -ji = k$. The typical quaternion is

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

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

$$|q|^2 \ = q_0^2 + q_1^2 + q_2^2 + q_3^2 = q\overline{q} = \overline{q}q.$$

Thus one can compute the multiplicative inverse of any non-zero quaternion,

$$q^{-1} = \frac{\overline{q}}{(|q|)^2} \ .$$

As a vector space, $\mathcal{H}$ is identified with $\mathcal{R}^4$ via the customary isomorphism,

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

which in turn induces an isomorphism between the subspace $\mathcal{P}$ of pure quaternions and $\mathcal{R}^3$,

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

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

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

3. Tensor Products. For the convenience of the reader we include a concrete definition of tensor products of finite-dimensional algebras [1], [23].

Definition 1. *Let V and W be vector spaces over a field $\mathcal{F}$, of dimension m and n respectively. Let $\{e_r : 1 \leq r \leq m\}$ be a basis for V and $\{f_s : 1 \leq s \leq n\}$ be a basis for W. Then the tensor product $V \otimes W$ is an mn dimensional vector space over $\mathcal{F}$ with basis $\{e_r \otimes f_s \ : 1 \leq r \leq m, 1 \leq s \leq n\}$, and is equipped with a bilinear map*

$$\otimes : V \times W \longrightarrow V \otimes W$$

$$(\sum_{r=1}^{m} v_r e_r \ , \ \sum_{s=1}^{n} w_s f_s) \overset{\otimes}{\longmapsto} \sum_{r,s} v_r w_s (e_r \otimes f_s) \tag{2}$$

It is customary to denote $\otimes(v,w)$ by $v\otimes w$, where $v\in V, w\in W$.

Now let $\mathcal{A}$ and $\mathcal{B}$ be *algebras* over a field $\mathcal{F}$. The tensor product $\mathcal{A}\otimes\mathcal{B}$ over $\mathcal{F}$ can be made into an algebra that is essentially the tensor product of the underlying vector spaces, enhanced by an additional multiplicative structure:

$$*\,(a\otimes b, a'\otimes b') = aa'\otimes bb', \quad \forall\, a,a' \in A, \;\forall\, b,b' \in B \tag{3}$$

By extending the above $*$ bilinearly to all the other elements of $\mathcal{A}\otimes\mathcal{B}$, we make $\mathcal{A}\otimes\mathcal{B}$ into an algebra. We will henceforth denote $*(a\otimes b, a'\otimes b')$ by simply $(a\otimes b)(a'\otimes b')$.

4. The Isomorphisms.

4.1. $\mathcal{P}\otimes\mathcal{P}$ and $\mathcal{M}_3(\mathcal{R})$. A useful isomorphism between the 9-dimensional vector spaces $\mathcal{P}\otimes\mathcal{P}$ and $\mathcal{M}_3(\mathcal{R})$ is obtained by first defining a bilinear map f on the Cartesian product $\mathcal{P}\times\mathcal{P}$:

$$f:\mathcal{P}\times\mathcal{P}\longrightarrow\mathcal{M}_3(\mathcal{R}), \quad f(p,q)=\begin{pmatrix} p_1\\ p_2\\ p_3\end{pmatrix}\begin{pmatrix} q_1 & q_2 & q_3\end{pmatrix} = pq^t.$$

We remark that $f(p,q)$ is often called the Kronecker product of p and q [15]. From the basic properties of tensor product [1], f induces a unique *linear* map

$$\psi:\mathcal{P}\otimes\mathcal{P}\longrightarrow\mathcal{M}_3(\mathcal{R})$$

with the property that

$$\psi(p\otimes q) = pq^t. \tag{4}$$

Showing that ψ is a vector space isomorphism is now a simple exercise.

4.2. $\mathcal{H}\otimes\mathcal{H}$ and $\mathcal{M}_4(\mathcal{R})$. Clearly an analogous Kronecker product gives a bilinear map from $\mathcal{H}\otimes\mathcal{H}$ to $\mathcal{M}_4(\mathcal{R})$, which then induces a linear isomorphism between $\mathcal{H}\otimes\mathcal{H}$ and $\mathcal{M}_4(\mathcal{R})$. However, this linear isomorphism does *not* preserve the multiplicative structure and hence fails to be an algebra isomorphism. (Since $\mathcal{P}\otimes\mathcal{P}$ is not an algebra, this question does not arise in §4.1).

Construct an algebra isomorphism between the 16-dimensional algebras $\mathcal{H}\otimes\mathcal{H}$ and $\mathcal{M}_4(\mathcal{R})$ as follows. First, to every ordered pair (p,q) in $\mathcal{H}\times\mathcal{H}$, associate the real 4×4 matrix that represents the linear transformation

$$\begin{aligned}\mathcal{H}&\rightarrow\mathcal{H}\\ v&\mapsto pv\overline{q}\end{aligned}$$

with respect to the standard basis. Denote this matrix by $\mu(p,q)$. Clearly this defines a bilinear map $\mu:\mathcal{H}\times\mathcal{H}\longrightarrow\mathcal{M}_4(\mathcal{R})$, which, from the basic properties of tensor product [1] induces a unique *linear* map

$$\phi:\mathcal{H}\otimes\mathcal{H}\rightarrow\mathcal{M}_4(\mathcal{R})$$

with the property that $\phi(p\otimes q)=\mu(p,q)$. It can be shown that ϕ is a bijection that preserves not only the vector space structure, but also the multiplicative structure[3].

[3] This isomorphism between $\mathcal{H}\otimes\mathcal{H}$ and $\mathcal{M}_4(\mathcal{R})$ has also been used in an entirely different context: the characterization of linear maps which preserve the Ky-Fan norms [18].

The proof is omitted, as it is largely a straightforward exercise in formal algebra [3]. Another way to prove these algebras are isomorphic is to use the fact that the Brauer group of $\mathcal{R}$ is $\mathcal{Z}/2\mathcal{Z}$ [23].

From equation (3) defining multiplication in a tensor product, we get

$$p \otimes q = (p \otimes 1)(1 \otimes q). \tag{5}$$

Applying ϕ gives $\phi(p \otimes q) = \phi(p \otimes 1)\phi(1 \otimes q)$, or in terms of matrices

$$\phi(p \otimes q) = \begin{pmatrix} p_0 & -p_1 & -p_2 & -p_3 \\ p_1 & p_0 & -p_3 & p_2 \\ p_2 & p_3 & p_0 & -p_1 \\ p_3 & -p_2 & p_1 & p_0 \end{pmatrix} \begin{pmatrix} q_0 & q_1 & q_2 & q_3 \\ -q_1 & q_0 & -q_3 & q_2 \\ -q_2 & q_3 & q_0 & -q_1 \\ -q_3 & -q_2 & q_1 & q_0 \end{pmatrix}. \tag{6}$$

Since $(p \otimes 1)(1 \otimes q) = (1 \otimes q)(p \otimes 1)$, a non-obvious fact that now follows immediately is that the matrices on the right hand side of (6) commute!

Define conjugation in $\mathcal{H}\otimes\mathcal{H}$ by

$$\overline{p \otimes q} = \overline{p} \otimes \overline{q} \qquad \forall p, q \in \mathcal{H} \tag{7}$$

and extend linearly to all of $\mathcal{H}\otimes\mathcal{H}$. Examining the matrix $\phi(p \otimes 1)$, displayed as the first factor in (6), one sees that

$$\phi(\overline{p} \otimes 1) = (\phi(p \otimes 1))^t.$$

Similarly, $\phi(1 \otimes \overline{q}) = (\phi(1 \otimes q))^t$. Thus we have

Proposition 1. *Conjugation in $\mathcal{H}\otimes\mathcal{H}$ corresponds, via ϕ, to transpose in $\mathcal{M}_4(\mathcal{R})$.*

By the usual abuse of notation, we will sometimes use $p \otimes q$ to stand for $\phi(p \otimes q)$. The reason for this indulgence is twofold: to simplify notation, and to emphasize that we will be freely moving between $\mathcal{H}\otimes\mathcal{H}$ and $\mathcal{M}_4(\mathcal{R})$, sometimes even appearing to be in both places at once!

5. Strategy. We want to translate the problem of orthogonally diagonalizing a 4×4 symmetric matrix into a corresponding problem in $\mathcal{H}\otimes\mathcal{H}$. To this end, we investigate each of the questions listed below:

- What does a symmetric matrix look like in $\mathcal{H}\otimes\mathcal{H}$?
- What does a diagonal matrix look like in $\mathcal{H}\otimes\mathcal{H}$?
- What does an orthogonal matrix correspond to in $\mathcal{H}\otimes\mathcal{H}$?
- How does one represent an orthogonal similarity in $\mathcal{H}\otimes\mathcal{H}$?

6. Quaternion Representations. The decomposition of $\mathcal{H}$ as the direct sum $\mathcal{R} \oplus \mathcal{P}$ induces a decomposition of $\mathcal{H}\otimes\mathcal{H}$:

$$\begin{aligned} \mathcal{H}\otimes\mathcal{H} &\cong (\mathcal{R} \oplus \mathcal{P}) \otimes (\mathcal{R} \oplus \mathcal{P}) \\ &\cong \Big\{(\mathcal{R} \otimes \mathcal{R}) \oplus (\mathcal{P} \otimes \mathcal{P})\Big\} \ \oplus \ \Big\{(\mathcal{P} \otimes \mathcal{R}) \oplus (\mathcal{R} \otimes \mathcal{P})\Big\}. \end{aligned} \tag{8}$$

Let $\mathcal{S} = (\mathcal{R} \otimes \mathcal{R}) \oplus (\mathcal{P} \otimes \mathcal{P})$, and $\mathcal{K} = (\mathcal{P} \otimes \mathcal{R}) \oplus (\mathcal{R} \otimes \mathcal{P})$. Then $\mathcal{S}$ and $\mathcal{K}$ are eigenspaces for the conjugation map on $\mathcal{H}\otimes\mathcal{H}$ corresponding to the eigenvalues 1 and -1, respectively. Since conjugation translates to transpose in $\mathcal{M}_4(\mathcal{R})$, every element of $\mathcal{S}$ represents a 4×4 symmetric matrix, and every element of $\mathcal{K}$ represents a 4×4

skew-symmetric matrix. Observing that $\mathcal{S}$ is a 10-dimensional subspace of $\mathcal{H}\otimes\mathcal{H}$ and $\mathcal{K}$ is a 6-dimensional subspace, it follows that

$$\begin{aligned} \phi(\mathcal{S}) &= \{S \in \mathcal{M}_4(\mathcal{R}) : S \text{ is symmetric}\} \\ \text{and } \phi(\mathcal{K}) &= \{K \in \mathcal{M}_4(\mathcal{R}) : K \text{ is skew-symmetric}\}. \end{aligned}$$

For notational convenience $\mathcal{S}$ will be used freely to denote either $\mathcal{S}$ or $\phi(\mathcal{S})$; context will make it clear which is intended. (Similarly for $\mathcal{K}$, S, K.) (NOTE: See remark at end of section 4.2). We have thus established :

Proposition 2. *The following are equivalent.*

1. *S is 4×4 symmetric.*
2. *$S \in (\mathcal{R}\otimes\mathcal{R}) \oplus (\mathcal{P}\otimes\mathcal{P})$.*
3. *There exists $c \in \mathcal{R}$ and pure quaternions p, q and r such that* $S = c1\otimes 1 \;+\; p\otimes i \;+\; q\otimes j \;+\; r\otimes k.$

Proposition 3 (Hacon). *The following are equivalent.*

1. *K is 4×4 skew-symmetric.*
2. *$K \in (\mathcal{P}\otimes\mathcal{R}) \oplus (\mathcal{R}\otimes\mathcal{P})$.*
3. *There exists pure quaternions p, q such that $K = p\otimes 1 \;+\; 1\otimes q$.*

Indeed, the standard basis $\{e_r \otimes e_s : 1 \le r, s \le 4\}$ for $\mathcal{H}\otimes\mathcal{H}$ gives, via ϕ, a beautiful basis for $\mathcal{M}_4(\mathcal{R})$ comprised entirely of *orthogonal* matrices, ten of them symmetric and the remaining six skew-symmetric[4]. One can use this basis, which is listed in the appendix, to calculate the quaternion representation of any 4×4 matrix. For a symmetric matrix $S = [s_{\ell m}] = c1\otimes 1 \;+\; p\otimes i \;+\; q\otimes j \;+\; r\otimes k$, the pure quaternions p, q, r and the scalar c are given by

$$c = \frac{1}{4}\,\mathrm{trace}(S), \tag{9}$$

$$p_1 = \frac{1}{4}(s_{11}+s_{22}-s_{33}-s_{44}), \quad p_2 = \frac{1}{2}(s_{23}+s_{14}), \quad p_3 = \frac{1}{2}(s_{24}-s_{13}), \tag{10}$$

$$q_1 = \frac{1}{2}(s_{23}-s_{14}), \quad q_2 = \frac{1}{4}(s_{11}-s_{22}+s_{33}-s_{44}), \quad q_3 = \frac{1}{2}(s_{34}+s_{12}), \tag{11}$$

$$r_1 = \frac{1}{2}(s_{24}+s_{13}), \quad r_2 = \frac{1}{2}(s_{34}-s_{12}), \quad r_3 = \frac{1}{4}(s_{11}-s_{22}-s_{33}+s_{44}). \tag{12}$$

The corresponding calculation for a skew-symmetric matrix $K = [k_{\ell m}] = p\otimes 1 + 1\otimes q$ is even simpler:

$$p_1 = -\frac{1}{2}(k_{12}+k_{34}), \quad p_2 = \frac{1}{2}(k_{24}-k_{13}), \quad p_3 = -\frac{1}{2}(k_{14}+k_{23}), \tag{13}$$

$$q_1 = \frac{1}{2}(k_{12}-k_{34}), \quad q_2 = \frac{1}{2}(k_{24}+k_{13}), \quad q_3 = \frac{1}{2}(k_{14}-k_{23}). \tag{14}$$

Further examination of the quaternion basis reveals the following simple characterizations of the canonical forms of interest to us:

[4] Any pair of elements from this basis either commute or anti-commute. While this property may at first appear striking, it is a natural consequence of the way multiplication is defined in a tensor product, and the fact that the basis elements $\{1, i, j, k\}$ of $\mathcal{H}$ either commute or anti-commute.

PROPOSITION 4.
(a) *A 4×4 matrix is diagonal if and only if it can be expressed as*
$$c_0 1\otimes 1 + c_1 i\otimes i + c_2 j\otimes j + c_3 k\otimes k, \quad \textit{for some } c_0, c_1, c_2, c_3 \in \mathcal{R}\ .$$
(b) (Hacon) *A 4×4 skew-symmetric matrix is in real Schur (2×2 block diagonal) form if and only if it can be expressed as* $si\otimes 1 + 1\otimes ti$, *for some* $s,t\in\mathcal{R}$.

7. Orthogonal Similarities in $\mathcal{H}\otimes\mathcal{H}$.

7.1. Rotations. Within a year of the discovery of the quaternions, Hamilton and Cayley had found a connection between orthogonal maps of $\mathcal{R}^3$ and quaternions [12], [4], [9]. About a decade later, in 1855, Cayley observed that every orthogonal map of $\mathcal{R}^4$ can be represented by pairs of quaternions [5], [9]. We now outline these results in modern terminology.

Let $\mathcal{GL}_n(\mathcal{R})$ denote the group of non-singular matrices. The orthogonal and special orthogonal subgroups are respectively,
$$\mathcal{O}(n) = \{R \in \mathcal{M}_n(\mathcal{R}) : R^tR = RR^t = I_n\}\ ,$$
$$\mathcal{SO}(n) = \{R \in \mathcal{O}(n) : \det R = +1\}.$$
We will refer to elements of $\mathcal{SO}(n)$ as n-dimensional rotations. Let $\mathcal{U}$ denote the set of unit quaternions,
$$\mathcal{U} = \{u \in \mathcal{H} :\ u\overline{u} = 1\},$$
which is a subgroup of $\mathcal{H}$ under multiplication.

PROPOSITION 5. *The map*
$$\begin{aligned} \mathcal{U}\times\mathcal{U} &\longrightarrow \mathcal{GL}_n(\mathcal{R}) \\ (x,y) &\ \mapsto\ \phi(x\otimes y) \end{aligned}$$
is a group homomorphism with image $\mathcal{SO}(4)$ and kernel $\{(1,1),(-1,-1)\}$.

COROLLARY 1. *The map*
$$\begin{aligned} \mathcal{U} &\longrightarrow \mathcal{GL}_n(\mathcal{R}) \\ x &\ \mapsto\ \phi(x\otimes x) \end{aligned}$$
is a group homomorphism with kernel $\{1,-1\}$ *and image the set of all matrices in $\mathcal{SO}(4)$ of the form*
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot \end{pmatrix},$$
which can be interpreted as rotations of $\mathcal{P} \cong \mathcal{R}^3$.

We resist the temptation to provide a detailed proof of these results here as they are 'well-known' [9], [6]. Instead we confine ourselves to pointing out why $\phi(x\otimes y)$ is in $\mathcal{SO}(4)$. Observe that $(x\otimes y)(\overline{x}\otimes\overline{y}) = (x\overline{x}\otimes y\overline{y}) = 1\otimes 1$. But conjugation in $\mathcal{H}\otimes\mathcal{H}$ corresponds to transpose in $\mathcal{M}_4(\mathcal{R})$, so $\phi(x\otimes y)$ must be an orthogonal matrix. The continuity of the determinant, together with the connectedness of $\mathcal{U}\times\mathcal{U}$ imply that $\det(\phi(x\otimes y))$ is positive.

To see why $\phi(x\otimes x)$ has the form stated in the corollary, recall that $\phi(x\otimes x)$ is the matrix encoding the map $q\mapsto xq\overline{x}$ in the standard basis (see §4.2). The first column of this matrix is the vector representing the quaternion $x1\overline{x} = x\overline{x} = 1$, i.e. e_0. Orthogonality of the matrix now forces the rest.

7.2. Similarities. Let $R \in \mathcal{SO}(4)$. Then by Proposition 5, there exist $x, y \in \mathcal{U}$ such that $\phi(x \otimes y) = R$. Consider an element of $\mathcal{H}\otimes\mathcal{H}$ of the form $p \otimes q$. By Proposition 1, the similarity $R\,\phi(p \otimes q)R^t$ in $\mathcal{M}_4(\mathcal{R})$ translates to

$$(x \otimes y)(p \otimes q)(\overline{x} \otimes \overline{y})$$

in $\mathcal{H}\otimes\mathcal{H}$. Using the definition of multiplication in a tensor product, this becomes

$$(xp\overline{x}) \otimes (yq\overline{y}).$$

Now comes the crucial observation: $xp\overline{x}$ is just the image of p under the rotation $\phi(x \otimes x)$, while $yq\overline{y}$ is the image of q under the rotation $\phi(y \otimes y)$. By Corollary 1, both of these rotations act in a nontrivial way only on the pure quaternion parts of p and q. Since a general element of $\mathcal{H}\otimes\mathcal{H}$ is a linear combination of elements of the form $p \otimes q$, we have

> *The effect of a* 4×4 *special orthogonal similarity on a* 4×4 *matrix can be reduced to the* independent *action of two 3-dimensional rotations.*

8. The Skew-Symmetric Algorithm. We have everything we need to outline how Hacon's method [10] for transforming a 4×4 skew-symmetric matrix K directly into its real Schur form works. By Proposition 3, there exist pure quaternions p, q such that $K = p \otimes 1 \,+\, 1 \otimes q$. Let $R = x \otimes y \,\in\, \mathcal{SO}(4)$. Then

$$RKR^t \;=\; xp\overline{x} \otimes 1 \;+\; 1 \otimes yq\overline{y}.$$

By Proposition 4b, all that remains is to find two 3-dimensional rotations: $\phi(x \otimes x)$ to rotate the 3-vector p into the i-direction and $\phi(y \otimes y)$ to rotate the 3-vector q into the i-direction. Hacon [10] shows that taking

$$x = \frac{|p| - ip}{|\,|p| - ip\,|}, \quad y = \frac{|q| - iq}{|\,|q| - iq\,|} \tag{15}$$

achieves this goal. That is,

$$RKR^t = si \otimes 1 \,+\, 1 \otimes ti = \begin{pmatrix} 0 & t-s & 0 & 0 \\ s-t & 0 & 0 & 0 \\ 0 & 0 & 0 & -s-t \\ 0 & 0 & s+t & 0 \end{pmatrix},$$

where $s = |p|$, $t = |q|$.

The matrix $R = x \otimes y$ can be computed from (15) and (6) as the product of the matrices

$$\frac{1}{|\,|p| - ip\,|} \begin{pmatrix} |p| + p_1 & 0 & -p_3 & p_2 \\ 0 & |p| + p_1 & p_2 & p_3 \\ p_3 & -p_2 & |p| + p_1 & 0 \\ -p_2 & -p_3 & 0 & |p| + p_1 \end{pmatrix} \tag{16}$$

and

$$\frac{1}{|\,|q| - iq\,|} \begin{pmatrix} |q| + q_1 & 0 & q_3 & -q_2 \\ 0 & |q| + q_1 & q_2 & q_3 \\ -q_3 & -q_2 & |q| + q_1 & 0 \\ q_2 & -q_3 & 0 & |q| + q_1 \end{pmatrix}, \tag{17}$$

where the pure quaternions p and q are computed from K via (13) and (14).

9. The Symmetric Algorithm. Let $S \in \mathcal{S}$. By Proposition 2, there exist $p, q, r \in \mathcal{P}$ and $c \in \mathcal{R}$ such that

$$S - c1 \otimes 1 \;=\; S - cI \;=\; p \otimes i + q \otimes j + r \otimes k \quad \in \quad \mathcal{P} \hat{\otimes} \mathcal{P}.$$

Certainly it suffices to diagonalize $\widehat{S} = S - cI$, since RSR^t is diagonal $\Leftrightarrow$ $R(S-cI)R^t$ is diagonal. Now for $R = x \otimes y \in \mathcal{SO}(4)$ we have

$$R\widehat{S}R^t \;=\; x\,p\,\overline{x} \otimes y\,i\,\overline{y} \;+\; x\,q\,\overline{x} \otimes y\,j\,\overline{y} \;+\; x\,r\,\overline{x} \otimes y\,k\,\overline{y}. \tag{18}$$

If R is to diagonalize $\widehat{S}$, then by Proposition 4, the 3-dimensional rotation $\phi(x \otimes x)$ must align the *triple of 3-vectors* $\{p, q, r\}$ along the *orthogonal* triple $\{i, j, \pm k\}$. Since $\{p, q, r\}$ is in general an oblique triple, this is clearly impossible. In the skew-symmetric case, by contrast, the problem reduces to rotating just *one* given 3-vector into a specified direction (granted, this has to be done twice, but the rotations are completely independent of one another). This in a nutshell is why the symmetric case is more complicated, and requires a new idea.

What we need is a different tensor decomposition of $\widehat{S}$,

$$a_1 \otimes b_1 + a_2 \otimes b_2 + a_3 \otimes b_3 \quad \in \quad \mathcal{P} \hat{\otimes} \mathcal{P},$$

with the property that *both* triples $\{a_1, a_2, a_3\}$ and $\{b_1, b_2, b_3\}$ are *orthogonal.* Then clearly it will be possible to independently rotate each triple into alignment with $\{i, j, \pm k\}$, thus diagonalizing $\widehat{S}$. This is where the vector space isomorphism $\psi : \mathcal{P} \hat{\otimes} \mathcal{P} \to \mathcal{M}_3(\mathcal{R})$, introduced in §4.1, and the singular value decomposition work together beautifully to produce the "right" tensor decomposition of $\widehat{S}$.

Since the canonical inclusion map embeds $\mathcal{P} \hat{\otimes} \mathcal{P}$ into $\mathcal{H} \hat{\otimes} \mathcal{H}$, every element of $\mathcal{P} \hat{\otimes} \mathcal{P}$ can be associated with a 3×3 matrix as well as a quite different 4×4 matrix[5]:

$$\begin{array}{ccc} \mathcal{P} \hat{\otimes} \mathcal{P} & \xrightarrow{\psi} & \mathcal{M}_3(\mathcal{R}) \\ \big\downarrow & & \\ \mathcal{H} \hat{\otimes} \mathcal{H} & \xrightarrow{\phi} & \mathcal{M}_4(\mathcal{R}) \end{array}$$

Successively exploiting the properties of ψ we get

$$\begin{array}{rcll} \psi(\widehat{S}) & = & \psi(p \otimes i + q \otimes j + r \otimes k) & \\ & = & pe_1^t + qe_2^t + re_3^t & \text{[linearity of } \psi \text{ and (5)]} \\ & = & \begin{pmatrix} p_1 & q_1 & r_1 \\ p_2 & q_2 & r_2 \\ p_3 & q_3 & r_3 \end{pmatrix} & \\ & = & \sigma_1 u_1 v_1^t + \sigma_2 u_2 v_2^t + \sigma_3 u_3 v_3^t & \text{[SVD decomposition]} \\ & = & \psi(\sigma_1 u_1 \otimes v_1 + \sigma_2 u_2 \otimes v_2 + \sigma_3 u_3 \otimes v_3) & \text{[linearity of } \psi \text{ and (5)]} \end{array}$$

But ψ is bijective, so we must have

$$\widehat{S} = p \otimes i \;+\; q \otimes j \;+\; r \otimes k \;=\; \sigma_1 u_1 \otimes v_1 \;+\; \sigma_2 u_2 \otimes v_2 \;+\; \sigma_3 u_3 \otimes v_3.$$

[5] For example, $\psi(i \otimes i) \;=\; e_1\, e_1^t \;=\; \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, but $\phi(i \otimes i) = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}$

Since $\{u_1, u_2, u_3\}$ and $\{v_1, v_2, v_3\}$ are orthonormal triples, the "right" tensor decomposition of a 4×4 symmetric matrix has thus been obtained from the singular value decomposition of its associated 3×3 matrix.

Although it is geometrically clear that a rotation of $\mathcal{P}$ which aligns a given orthonormal triple $\{u_1, u_2, u_3\}$ with $\{i, j, \pm k\}$ exists, trying to compute it all at once is somewhat complicated. Instead we can align these triples in stages. Begin by choosing $R = x \otimes y$ so that $\phi(x \otimes x)$ rotates any one of the left singular vectors into i, and $\phi(y \otimes y)$ rotates the *corresponding* right singular vector into i. Then the remaining singular vectors will perforce be moved into the jk-plane. Does it matter which left-right singular vector pair is rotated into i? The next section provides the surprising answer.

For now, let σ_1, σ_2 and σ_3 denote the singular values of $\psi(\widehat{S})$, not necessarily in decreasing order. And let R align u_1 and v_1 with i. Then

$$R\widehat{S}R^t = \sigma_1 i \otimes i \;+\; \sigma_2 x u_2 \overline{x} \otimes y v_2 \overline{y} \;+\; \sigma_3 x u_3 \overline{x} \otimes y v_3 \overline{y}, \tag{19}$$

where $xu_2\overline{x}, yv_2\overline{y}, xu_3\overline{x}$ and $yv_3\overline{y}$ are each linear combinations of j and k only. Thus $R\widehat{S}R^t$ is a linear combination of $i \otimes i, j \otimes j, k \otimes k, j \otimes k$ and $k \otimes j$. A quick look at the quaternion basis in the appendix reveals that $R\widehat{S}R^t$ and hence, RSR^t is already 2×2 block diagonal!

One may continue the alignment process by constructing a second orthogonal matrix $Q = z \otimes w$, where $\phi(z \otimes z)$ is a 3-dimensional rotation with axis i that aligns $xu_2\overline{x}$ with j; and $\phi(w \otimes w)$ is a 3-dimensional rotation with axis i that rotates $yv_2\overline{y}$ into j. Then $xu_3\overline{x}$ and $yv_3\overline{y}$ will necessarily be aligned with $\pm k$, and thus $QR\widehat{S}R^tQ^t$ (and hence $QRSR^tQ^t$) will be diagonal. (Alternatively, two ordinary Jacobi rotations in the $(1,2)$ and $(3,4)$ planes could be used to achieve diagonalization.)

We remark that the matrix $R = x \otimes y$ is given by the product of the matrices in (16) and (17), with the left singular vector that R aligns with i playing the role of p and the corresponding right singular vector playing the role of q. By exploiting the special structure of these matrices and noting that $|p| = 1 = |q|$ in this case, their product can be computed for the small price of 14 additions, 14 multiplications and one square root.

The major cost of computing R lies in finding a left–right singular vector pair of a 3×3 matrix. Note that the orthogonality of R does not depend on the accuracy of the singular vector pair used to construct it — as long as the vectors are unit, R will be orthogonal. During the early iterations of the quaternion-Jacobi method, it may perhaps not be necessary to compute the vectors with hair-splitting accuracy, since the annihilated elements will shortly be resurrected. Bear in mind also that the complete SVD is *not* needed — only one left–right singular vector pair is called for. We shall see in the next section that the pair corresponding to the largest singular value is the one to compute. What then is the best way to carry out this task? Among the many schemes to be evaluated are several iterative techniques and one direct method due to Bojanczyk and Lutoborski [2], that gives closed form formulae for the eigenvectors of a 3×3 symmetric matrix.

A matrix representation of Q can be obtained analogously. However, we shall see in §12 that in practise block diagonalization suffices, so Q need never be computed.

10. Eigenvalues from Singular Values. Assume that $R, Q \in \mathcal{SO}(4)$ are chosen as described at the end of §9, that is,

$$R \text{ rotates } u_1, v_1 \text{ into } i, \tag{20}$$

(21) $$QR \text{ rotates } u_1,\, v_1 \text{ into } i, \text{ and } u_2,\, v_2 \text{ into } j.$$

Equation (19) suggests there is a connection between the singular values of T and the eigenvalues of S:

PROPOSITION 6. *If $\sigma_1 \geq \sigma_2 \geq \sigma_3$ are the singular values of $T = [p\ q\ r]$, then the eigenvalues of $S - cI = p \otimes i\, +\, q \otimes j\, +\, r \otimes k$ are*

$$\sigma_1 + \sigma_2 + \tau\sigma_3,\quad \sigma_1 - \sigma_2 - \tau\sigma_3,\quad -\sigma_1 + \sigma_2 - \tau\sigma_3,\quad -\sigma_1 - \sigma_2 + \tau\sigma_3$$

where $\tau = \operatorname{sign}(\det(T))$. (If $\det(T) = 0$, set $\tau = 0$).

Proof. Let $\sigma_1 u_1 v_1^t + \sigma_2 u_2 v_2^t + \sigma_3 u_3 v_3^t$ be the SVD of T. The sign of $\det(T)$ determines whether T is orientation preserving or reversing. Since $T(v_\ell) = \sigma_\ell u_\ell$ for $1 \leq \ell \leq 3$, the triples $\{u_1, u_2, u_3\}$ and $\{v_1, v_2, v_3\}$ will have the *same* "handedness" when $\tau > 0$. That is, either both can be rotated into $\{i, j, k\}$ or both can be rotated into $\{i, j, -k\}$. On the other hand, $\tau < 0$ means that $\{u_1, u_2, u_3\}$ and $\{v_1, v_2, v_3\}$ have *opposite* handedness, i.e. one triple can be rotated into $\{i, j, k\}$ and the other into $\{i, j, -k\}$. We have seen that

$$\begin{aligned} S - cI &= p \otimes i + q \otimes j + r \otimes k \\ &= \sigma_1 u_1 \otimes v_1 + \sigma_2 u_2 \otimes v_2 + \sigma_3 u_3 \otimes v_3, \end{aligned}$$

so choosing rotations $R, Q \in \mathcal{SO}(4)$ as described in (20)–(21) gives

$$\begin{aligned} QR(S - cI)R^t Q^t &= \begin{cases} \sigma_1 i \otimes i + \sigma_2 j \otimes j + \sigma_3 k \otimes k, & \text{if } \tau > 0 \\ \sigma_1 i \otimes i + \sigma_2 j \otimes j - \sigma_3 k \otimes k, & \text{if } \tau < 0 \end{cases} \\ &= \operatorname{diag}(\sigma_1 + \sigma_2 + \tau\sigma_3, \sigma_1 - \sigma_2 - \tau\sigma_3, -\sigma_1 + \sigma_2 - \tau\sigma_3, \\ &\qquad -\sigma_1 - \sigma_2 + \tau\sigma_3). \end{aligned}$$

The case $\tau = 0$ is even simpler. In this situation $\sigma_3 = 0$, so

$$\begin{aligned} QR(S - cI)R^t Q^t &= \sigma_1 i \otimes i + \sigma_2 j \otimes j \\ &= \operatorname{diag}(\sigma_1 + \sigma_2, \sigma_1 - \sigma_2, -\sigma_1 + \sigma_2, -\sigma_1 - \sigma_2), \end{aligned}$$

which establishes the result. $\square$

The alert reader may notice that, in the above proof, the hypothesis $\sigma_1 \geq \sigma_2 \geq \sigma_3$ is used only in the case when $\det(T) = 0$, and then only to conclude that σ_3 must be zero. When T is non-singular, the expressions for the eigenvalues of $S - cI$ in terms of the singular values of T remain valid under any permutation of the σ_i's. That is, the ordering of the σ_i's has no effect on the *set* of eigenvalues of $S - cI$ (as expected), but it does affect the order in which these eigenvalues appear on the diagonal of $QR(S - cI)R^t Q^t$. It is to this issue that we turn next.

11. Sorting Similarities. For notational convenience, denote the eigenvalues of $S - cI$ by λ_ℓ, $1 \leq \ell \leq 4$, where

(22) $$\lambda_1 = \sigma_1 + \sigma_2 + \tau\sigma_3,$$
(23) $$\lambda_2 = \sigma_1 - \sigma_2 - \tau\sigma_3,$$
(24) $$\lambda_3 = -\sigma_1 + \sigma_2 - \tau\sigma_3,$$
(25) $$\lambda_4 = -\sigma_1 - \sigma_2 + \tau\sigma_3.$$

Note that in the course of proving the previous proposition we showed that

$$QR(S-cI)R^tQ^t = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3, \lambda_4).$$

We now ask, what conditions on R and Q will ensure that the eigenvalues of $S-cI$ appear in non-increasing order on the diagonal? This question is particularly relevant because Mascarenhas [20] has recently shown that choosing the sorting angle for the plane rotations used in the Jacobi method can lead to significant improvement in performance when orderings with higher than quadratic rates of convergence are used. (The sorting angle eventually becomes the usual (small) angle in the plane-Jacobi methods.) If one could construct a 4×4 analogue of the 2×2 sorting similarity, one would expect a similar improvement in performance. Happily, the construction is unexpectedly simple. But first, a definition.

Definition 2. *Let $A \in \mathcal{M}_n(\mathcal{R})$ have real eigenvalues. Let $B \in \mathcal{M}_n(\mathcal{R})$ block diagonalize A, i.e., $BAB^{-1} = \mathrm{diag}(\widetilde{A}_1, \widetilde{A}_2, \ldots, \widetilde{A}_m)$ where each $\widetilde{A}_\ell$ is some $n_\ell \times n_\ell$ matrix. If every eigenvalue of $\widetilde{A}_\ell$ is greater than or equal to every eigenvalue of $\widetilde{A}_{\ell+1}$, for $1 \le \ell \le m-1 \le n-1$, then BAB^{-1} is called a sorting similarity.*

Proposition 7. *Suppose $\sigma_1 \ge \sigma_2 \ge \sigma_3$. If R and Q are chosen according to (20)-(21), then QR is a sorting similarity for S.*

Proof. $\lambda_1 \ge \lambda_2 \Leftrightarrow \sigma_1+\sigma_2+\tau\sigma_3 \ge \sigma_1-\sigma_2-\tau\sigma_3 \Leftrightarrow \sigma_2 \ge -\tau\sigma_3$, which follows from $\sigma_2 \ge \sigma_3$. Next, $\lambda_2 \ge \lambda_3 \Leftrightarrow \sigma_1-\sigma_2-\tau\sigma_3 \ge -\sigma_1+\sigma_2-\tau\sigma_3 \Leftrightarrow \sigma_1 \ge \sigma_2$. Finally, $\lambda_3 \ge \lambda_4 \Leftrightarrow -\sigma_1+\sigma_2-\tau\sigma_3 \ge -\sigma_1-\sigma_2+\tau\sigma_3 \Leftrightarrow \sigma_2 \ge \tau\sigma_3$, which again follows from $\sigma_2 \ge \sigma_3$. □

Proposition 8. *Suppose $\sigma_1 \ge \sigma_2 \ge \sigma_3$. If R is chosen according to (20), then R is a sorting similarity for S. That is, the eigenvalues of the upper 2×2 block of RSR^t are larger than those of the lower 2×2 block.*

Proof. We know from §9 that

$$R(S-cI)R^t = \sigma_1 i\otimes i + c_{jj} j\otimes j + c_{kk} k\otimes k + c_{jk} j\otimes k + c_{kj} k\otimes j,$$

where the $c_{\ell m}$'s are real constants. Hence by the discussion following (19), $R(S-cI)R^t$ is 2×2 block diagonal. Expressing this in matrix form we have

$$R(S-cI)R^t = \begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix},$$

where

$$S_1 = \begin{pmatrix} \sigma_1 + c_{jj} + c_{kk} & -c_{jk} + c_{kj} \\ -c_{jk} + c_{kj} & \sigma_1 - c_{jj} - c_{kk} \end{pmatrix}$$

and

$$S_2 = \begin{pmatrix} -\sigma_1 + c_{jj} - c_{kk} & c_{jk} + c_{kj} \\ c_{jk} + c_{kj} & -\sigma_1 - c_{jj} + c_{kk} \end{pmatrix}.$$

The 2×2 block diagonal form implies that the eigenvalues of S_1 must be two of the eigenvalues $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ of $S-cI$. From Proposition 6 and 7, the sum of the two largest eigenvalues of $S-cI$ is $\lambda_1+\lambda_2 = 2\sigma_1$; of course the sum of any other pair $\lambda_i+\lambda_j$ can be no bigger than this. But we also have $\mathrm{trace}(S_1) = \Sigma(\text{eigenvalues of } S_1) = 2\sigma_1$. Therefore λ_1 and λ_2 must be eigenvalues of S_1. Consequently λ_3, λ_4 must be eigenvalues of S_2, so by Proposition 7 we are done. □

12. Preliminary Experimental Results. In practise, one can partition a $2n \times 2n$ symmetric matrix A into contiguous 2×2 blocks, denoted by $A_{\ell m}$, with $1 \leq \ell, m \leq n$. Then annihilate each off-diagonal block $A_{\ell m}$ with $\ell \neq m$ according to some order, thereby obtaining a Jacobi-type method. Complete diagonalization of the target 4×4 submatrix is unnecessary (hence wasteful): block diagonalization incurs no penalty in the form of additional sweeps or additional iterations per sweep. This is because the weight of the annihilated (pivot) block moves onto the two corresponding 2×2 diagonal blocks, and stays there! In other words, if we define

$$\mathrm{BlockOff}^2(A) = \|A\|_F^2 \; - \; \sum_1^n \|A_{\ell\ell}\|_F^2,$$

then just as in traditional Jacobi methods,

$$\mathrm{BlockOff}^2(A^{(t+1)}) = \mathrm{BlockOff}^2(A^{(t)}) \; - \; 2\|A_{\ell m}^{(t)}\|_F^2.$$

Here $A^{(t)}$ denotes the matrix A after t block annihilations, and ℓm denotes the position of the t^{th} pivot block. Thus $\{\mathrm{BlockOff}^2(A^{(t)})\}$ is a decreasing sequence.

It is observed experimentally that the sequence $\{A^{(t)}\}$ converges to a $(2{\times}2)$- block-diagonal matrix, which can then be diagonalized with negligible cost. We present below the results of a MATLAB implementation of Jacobi and quaternion-Jacobi methods on a random 64×64 symmetric matrix with entries uniformly distributed in $[-1, 1]$. This behavior seems to be typical, but more experimentation is needed.

$n = 64$	*Odd-Even Ordering; Sorting Similarity*			
	Jacobi		*quaternion-Jacobi*	
Sweep	Number of 2×2 Rotations	BlockOff	Number of 4×4 Rotations	BlockOff
0		$1.800e+01$		$1.800e+01$
1	2016	$1.067e+01$	496	$1.032e+01$
2	2016	$2.883e+00$	496	$2.609e+00$
3	2016	$4.633e-01$	496	$3.301e-01$
4	2016	$1.840e-02$	496	$5.025e-03$
5	2016	$2.877e-05$	496	$1.261e-06$
6	2016	$8.111e-11$	490	$6.702e-14$
7	541	$1.218e-13$		

$n = 64$	*Row-Cyclic Ordering; Sorting Similarity*			
	Jacobi		*quaternion-Jacobi*	
Sweep	Number of 2×2 Rotations	BlockOff	Number of 4×4 Rotations	BlockOff
0		$1.800e+01$		$1.800e+01$
1	2016	$1.120e+01$	496	$1.066e+01$
2	2016	$3.664e+00$	496	$3.171e+00$
3	2016	$8.714e-01$	496	$6.249e-01$
4	2016	$1.208e-01$	496	$7.069e-02$
5	2016	$3.321e-03$	496	$1.332e-03$
6	1976	$4.324e-06$	468	$5.580e-07$
7	1279	$8.445e-12$	158	$1.529e-13$
8	1270	$1.230e-13$		

The same matrix is used in both tables; the first one compares the two methods when the odd-even ordering is used, the second when the row-cyclic order is used.

13. Convergence. Sorting similarities lead to a globally convergent quaternion-Jacobi method [19]:

THEOREM 1. *Let A_0 be an $n \times n$ symmetric matrix with distinct eigenvalues $\lambda_1 > \lambda_2 > \cdots > \lambda_n$, and let $A_{i+1} = R_i A_i R_i^t$ be the sequence of iterates of the quaternion-Jacobi method, under any quasicyclic ordering*[6]. *If each R_i effects a sorting, diagonalizing similarity on its target 4×4 submatrix, then the sequence $\{A_i\}$ converges to* $\mathrm{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n]$. *Further, the rate of convergence is asymptotically quadratic.*

The proof of this theorem is quite long and will be published in a later paper. It contains results which encompass extremely general Jacobi-type methods and grew out of an attempt to extend the ideas found in Mascarenhas' thesis [20]. Among the tools used is a majorization result due to Schur [15], and an SVD perturbation theorem due to Wedin [24], [25], [27]. Note that the quadratic rate of convergence applies to any quasicyclic ordering.

14. Extension to Normal Matrices. Combining the two algorithms leads to a quaternion-Jacobi method for real normal matrices. First, write the matrix A as the sum of a symmetric matrix S and a skew-symmetric matrix K. Since A is normal, S and K commute. Diagonalize S using the symmetric quaternion-Jacobi algorithm with sorting transformations at every iteration. This gives an orthogonal matrix U such that $UAU^t = D_s + UKU^t$, with the entries of the diagonal matrix D_s appearing in non-increasing order. The skew-symmetric matrix $B_k = UKU^t$ is block-diagonal: since B_k commutes with D_s, the only non-zero off-diagonal elements of B_k occur in diagonal blocks corresponding to equal eigenvalues of D_s. Now each diagonal block of B_k can be independently transformed into its real Schur form using Hacon's skew-symmetric algorithm, without affecting D_s.

As is well-known, the plane rotations which are so effective in diagonalizing 2×2 symmetric matrices leave all 2×2 skew-symmetric matrices unscathed[7]. The remarkable correspondence between $\mathcal{H} \otimes \mathcal{H}$ and $\mathcal{M}_4(\mathcal{R})$ gives us, for the first time, a Jacobi method which works in a similar manner for *both* symmetric and skew-symmetric matrices, exploiting and preserving their special structure while lending a unity to the eigenproblem for these two classes that has hitherto been lacking.

15. Acknowledgements. It is a pleasure to thank Patricia Eberlein, Nick Higham, Steve Mackey and Don Schack for stimulating discussions, and Cleve Moler and the MATLAB team for providing such a convenient environment in which to test matrix algorithms.

[6] Quasicyclic orderings were introduced in [14].

[7] Pardekooper [22] deserves mention for first addressing the skew-symmetric case.

APPENDIX

The Quaternion Basis for $\mathcal{M}_4(\mathcal{R})$

$$
\underset{1\otimes 1}{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}
\quad
\underset{1\otimes i}{\begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}}
\quad
\underset{1\otimes j}{\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}}
\quad
\underset{1\otimes k}{\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}}
$$

$$
\underset{i\otimes 1}{\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}}
\quad
\underset{i\otimes i}{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}}
\quad
\underset{i\otimes j}{\begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}}
\quad
\underset{i\otimes k}{\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}}
$$

$$
\underset{j\otimes 1}{\begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}}
\quad
\underset{j\otimes i}{\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}}
\quad
\underset{j\otimes j}{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}}
\quad
\underset{j\otimes k}{\begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}}
$$

$$
\underset{k\otimes 1}{\begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}}
\quad
\underset{k\otimes i}{\begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}}
\quad
\underset{k\otimes j}{\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}}
\quad
\underset{k\otimes k}{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}
$$

REFERENCES

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

[2] A. W. Bojanczyk and A. Lutoborski, *Computation of the Euler angles of a symmetric 3×3 matrix*, SIAM J. of Matrix Anal. Appl., 12 (1991), pp. 41–48.

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

[4] 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 Philosophical Magazine, vol. XXVI, 1845.

[5] ———, *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 Journal für die Reine und Angewandte Mathematik, 1855.

[6] H. S. M. Coxeter, *Quaternions and reflections*, Amer. Math. Monthly, 53 (1946), pp. 136–146.

[7] J. Demmel and K. Veselić, *Jacobi's method is more accurate than QR*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 1204–1245.

[8] R. Dimitric and B. Goldsmith, *The mathematical tourist*, Mathematical Intelligencer, 11 (1989), pp. 29–30.

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

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

[11] W. R. Hamilton, in The Collected Mathematical Papers, H. Halberstam and R. E. Ingram, eds., vol. 3, Cambridge University Press, 1967, pp. xx–xvi. Letter to his son Archibald Hamilton.

[12] ———, *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, 1967, pp. 355–362. Originally published in Proc. Royal. Irish Acad. vol III, 1847, communicated November 1844.

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

[14] P. Henrici, *On the speed of convergence of cyclic and quasicyclic Jacobi methods for computing eigenvalues of Hermitian matrices*, J. Soc. Indust. Appl. Math, 6 (1958), pp. 144–161.

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

[16] C. G. J. Jacobi, *Über eine neue Auflösungart der bei Methode der kleinsten Quadrate vorkommenden lineären Gleichungen*, Astronomische Nachrichten, 22 (1845). Translated by G. W. Stewart, University of Maryland, Tech. Report TR-92-42.

[17] ———, *Über ein leichtes Verfahren die in der Theorie der Säcularstörungen vorkommenden Gleichungen numerisch aufzulösen*, Journal für die Reine und Angewandte Mathematik, 30 (1846), pp. 51–94.

[18] C. R. Johnson, T. Laffey, and C.-K. Li, *Linear transformations on $M_n(R)$ that preserve the Ky Fan k-norm and a remarkable special case when $(n,k) = (4,2)$*, Linear and Multilinear Algebra, 23 (1988), pp. 285–298.

[19] N. Mackey, *Global convergence of the quaternion-Jacobi method.* In Preparation.

[20] W. F. Mascarenhas, *On the convergence of the Jacobi method for arbitrary orderings*, PhD thesis, M.I.T., 1991.

[21] R. Mathias, *Accurate eigensystem computations by Jacobi methods.* Preprint, June 1993.

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

[23] R. S. Pierce, *Associative Algebras*, Springer, 1982.

[24] G. W. Stewart, *Perturbation theory for the singular value decomposition*, Tech. Report TR-90-124, University of Maryland, 1990.

[25] G. W. Stewart and J. Sun, *Matrix Perturbation Theory*, Academic Press, 1990.

[26] B. L. van der Waerden, *Hamilton's discovery of quaternions*, Mathematics Magazine, 49 (1976), pp. 227–234.

[27] P.-A. Wedin, *Perturbation bounds in connection with singular value decomposition*, BIT, 12 (1972), pp. 99–111.