NORTH-HOLLAND

Quaternions and Matrices of Quaternions*

Fuzhen Zhang†
Department of Mathematical Science
Nova Southeastern University
Fort Lauderdale, Florida 33314

Dedicated to Robert C. Thompson

Submitted by Richard A. Brualdi

ABSTRACT

We give a brief survey on quaternions and matrices of quaternions, present new proofs for certain known results, and discuss the quaternionic analogues of complex matrices. The methods of converting a quaternion matrix to a pair of complex matrices and homotopy theory are emphasized. © *Elsevier Science Inc., 1997*

1. INTRODUCTION

The family of quaternions plays a role in quantum physics [1, 18, 19]. It often appears in mathematics as an algebraic system—a skew field or noncommutative division algebra [7]. While matrices over commutative rings have gained much attention [6], the literature on matrices with quaternion entries, though dating back to 1936 [39], is fragmentary. Renewed interest has been witnessed recently [3, 4, 9, 11–13, 31, 36, 35, 47, 48]. In hopes that it will be useful to a wide audience, a concise survey on matrices of quaternion entries with new proofs for some known results is presented here.

* Part of this paper was presented at the AMS-MAA joint meeting, San Antonio, January 1993, under the title "Everything about the matrices of quaternions."

† The main work was done while the author was a student of R. C. Thompson's at the University of California, Santa Barbara. E-mail: zhang@alpha.nova.edu.

LINEAR ALGEBRA AND ITS APPLICATIONS 251:21–57 (1997)

655 Avenue of the Americas, New York, NY 10010

0024-3795/97/$17.00
SSDI 0024-3795(95)00543-Z

The main obstacles in the study of quaternion matrices, as expected come from the noncommutative multiplication of quaternions. One will find that working on a quaternion matrix problem is often equivalent to dealing with a pair of complex matrices. Homotopy theory has also been applied in the study.

We begin by considering the following three basic questions whose answers are well known in the case of complex matrices. Suppose that A and B are $n \times n$ matrices with quaternion entries.

QUESTION 1. If $AB = I$, the $n \times n$ identity matrix, is it true that $BA = I$?

QUESTION 2. Does every quaternion matrix A have an eigenvalue?

QUESTION 3. Is the numerical range of a quaternion matrix A always convex, as is the classical numerical range of a complex matrix?

After a moment's consideration, one will realize that answering these questions is unexpectedly difficult.

Recall from elementary linear algebra that there are several ways to answer the first question when A and B are complex matrices, one of which is to utilize the fact that if $\det A \neq 0$ then A is invertible; consequently B is the inverse of A and $BA = I$. This approach apparently does not apply in our case, because the determinant of a quaternion matrix makes no sense at this point. We will give an affirmative answer to the question in Section 4 by converting each of the quaternion matrices A and B into a pair of complex matrices.

For Question 2, since the quaternions do not commute, it is necessary to treat the linear systems $Ax = \lambda x$ and $Ax = x\lambda$ separately. Rewrite the first one as $(\lambda I - A)x = 0$. In the complex case, the fact that $\det(\lambda I - A) = 0$ has a solution guarantees an eigenvalue for A. This idea does not go through for quaternion matrices. To avoid determinants, a topological approach is needed. As one will see in Section 5, the latter system has been well studied, while the former one is not easy to handle, and few results have been obtained.

The answer to Question 3 is negative. The quaternionic numerical range of a matrix is not convex in general, even for a complex normal matrix. In contrast, the classical numerical range of a complex normal matrix, as is well known, is the convex hull of the eigenvalues of the matrix. We will give in Section 9 a complete characterization of the quaternionic numerical range of a normal matrix with quaternion entries and discuss the convexity of the upper complex plane part of the quaternionic numerical range of a matrix with quaternion entries.

With terminologies of complex matrices similarly defined for quaternion matrices, we will consider the above problems further, investigate other aspects such as similarity, rank, determinant, and canonical forms, and present certain new features that are hard to foresee and to prove. Some results on quaternion matrices are analogous to those on complex matrices, some are not.

We will focus only on matrices of quaternions. There is no attempt to cover everything related to quaternions and review the historical development of quaternions. For other aspects such as quaternion algebra and analysis, see, e.g., [14], [20], [28], [31], [37], and [38].

2. QUATERNIONS AND EQUIVALENCE CLASSES

As usual, let $\mathbb{C}$ and $\mathbb{R}$ denote the fields of the complex and real numbers respectively. Let $\mathbb{Q}$ be a four-dimensional vector space over $\mathbb{R}$ with an ordered basis, denoted by $\mathbf{e}$, $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$.

A *real quaternion*, simply called *quaternion*, is a vector

$$x = x_0\mathbf{e} + x_i\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k} \in \mathbb{Q}$$

with real coefficients x_0, x_1, x_2, x_3.

Besides the addition and the scalar multiplication of the vector space $\mathbb{Q}$ over $\mathbb{R}$, the product of any two of the quaternions $\mathbf{e}, \mathbf{i}, \mathbf{j}, \mathbf{k}$ is defined by the requirement that $\mathbf{e}$ act as a identity and by the table

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1,$$

$$\mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \qquad \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \qquad \mathbf{ki} = -\mathbf{ik} = \mathbf{j}.$$

If a and b are any (real) scalars, while $\mathbf{u}, \mathbf{v}$ are any two of $\mathbf{e}, \mathbf{i}, \mathbf{j}, \mathbf{k}$, then the product $(a\mathbf{u})(b\mathbf{v})$ is defined as $(ab)(\mathbf{uv})$. These rules, along with the distribution law, determine the product of any two quaternions.

Real numbers and complex numbers can be thought of as quaternions in the natural way. Thus $x_0\mathbf{e} + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$ is simply written as $x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$.

For any $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k} \in \mathbb{Q}$, we define $\operatorname{Re} x = x_0$, the *real part* of x; $\operatorname{Co} x = x_1 + x_2\mathbf{i}$, the *complex part* of x; $\operatorname{Im} x = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$, the *imaginary part* of x; $\bar{x} = x^* = x_0 - x_1\mathbf{i} - x_2\mathbf{j} - x_3\mathbf{k}$, the *conjugate* of x; and $|x| = \sqrt{x^*x} = \sqrt{x_0^2 + x_1^2 + x_2^2 + x_3^2}$, the *norm* of x. x is said to be a *unit quaternion* if its norm is 1.

The following facts then follow from the definitions.

THEOREM 2.1. *Let x, y, and z be quaternions. Then*

1. $x^*x = xx^*$, *i.e.*, $|x| = |x^*|$;
2. $|\cdot|$ *is a norm on* $\mathbb{Q}$, *i.e.*,

$$|x| = 0 \quad \textit{if and only if} \quad x = 0;$$

$$|x + y| \leqslant |x| + |y|;$$

$$|xy| = |yx| = |x|\,|y|;$$

3. $|x|^2 + |y|^2 = \frac{1}{2}(|x + y|^2 + |x - y|^2)$;
4. $x = |x|u$ *for some unit quaternion* u;
5. $\mathbf{j}c = \bar{c}\mathbf{j}$ *or* $\mathbf{j}c\mathbf{j}^* = \bar{c}$ *for any complex number* c;
6. $x^*\mathbf{i}x = (x_0^2 + x_1^2 - x_2^2 - x_3^2)\mathbf{i} + 2(-x_0x_3 + x_1x_2)\mathbf{j} + 2(x_0x_2 + x_1x_3)\mathbf{k}$ *if* $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$;
7. $x = \frac{1}{2}(x + x^*) + \frac{1}{2}(x + \mathbf{i}x^*\mathbf{i}) + \frac{1}{2}(x + \mathbf{j}x^*\mathbf{j}) + \frac{1}{2}(x + \mathbf{k}x^*\mathbf{k})$, *and* $x^* = -\frac{1}{2}(x + \mathbf{i}x\mathbf{i} + \mathbf{j}x\mathbf{j} + \mathbf{k}x\mathbf{k})$;
8. $x^2 = |\operatorname{Re} x|^2 - |\operatorname{Im} x|^2 + 2\operatorname{Re} x \operatorname{Im} x$.
9. $(xy)^* = y^*x^*$;
10. $(xy)z = x(yz)$;
11. $(x + y)^2 \neq x^2 + 2xy + y^2$ *in general*;
12. $x^* = x$ *if and only if* $x \in \mathbb{R}$;
13. $ax = xa$ *for every* $x \in \mathbb{Q}$, *if and only if* $a \in \mathbb{R}$;
14. $x^*/|x|^2$ *is the inverse of* x *if* $x \neq 0$, *and if* x^{-1} *denotes the inverse of* x, *then* $|x^{-1}| = 1/|x|$;
15. $x^2 = -1$ *has infinitely many of quaternions* x *as solutions*;
16. x *and* x^* *are solutions of* $t^2 - 2(\operatorname{Re} x)t + |x|^2 = 0$;
17. *every quaternion* q *can be uniquely expressed as* $q = c_1 + c_2\mathbf{j}$, *where* c_1 *and* c_2 *are complex numbers.*

Two quaternions x and y are said to be *similar* if there exists a nonzero quaternion u such that $u^{-1}xu = y$; this is written as $x \sim y$. Obviously, x and y are similar if and only if there is a unit quaternion v such that $v^{-1}xv = y$, and two similar quaternions have the same norm. It is routine to check that $\sim$ is an equivalence relation on the quaternions. We denote by $[x]$ the equivalence class containing x.

LEMMA 2.1. *If* $q = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$, *then* q *and* $q_0 + \sqrt{q_1^2 + q_2^2 + q_3^2}\,\mathbf{i}$ *are similar, namely,* $q \in [q_0 + \sqrt{q_1^2 + q_2^2 + q_3^2}\,\mathbf{i}]$.

Proof. Consider the equation of quaternions

$$qx = x\left(q_0 + \sqrt{q_1^2 + q_2^2 + q_3^2}\,\mathbf{i}\right). \tag{1}$$

It is easy to verify that $x = (\sqrt{q_1^2 + q_2^2 + q_3^2} + q_1) - q_3\mathbf{j} + q_2\mathbf{k}$ is a solution to Equation (1) if $q_2^2 + q_3^2 \neq 0$. For the case where q is a complex number, Theorem 2.1, part 5, may be used. ■

One can also prove the lemma by changing (1) to a linear homogeneous system of real variables.

It is readily seen that $[x]$ contains a single element if and only if $x \in \mathbb{R}$. If $x \notin \mathbb{R}$, then $[x]$ contains infinitely many quaternions, among which there are only two complex numbers that are a conjugate pair; moreover $x \sim x^*$ for every quaternion x.

This lemma yields the following theorem.

THEOREM 2.2 (Brenner, 1951; Au-Yeung, 1984). *Let* $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$ *and* $y = y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}$ *be quaternions. Then* x *and* y *are similar if and only if* $x_0 = y_0$ *and* $x_1^2 + x_2^2 + x_3^2 = y_1^2 + y_2^2 + y_3^2$, *i.e.,* $\operatorname{Re} x = \operatorname{Re} y$ *and* $|\operatorname{Im} x| = |\operatorname{Im} y|$.

REMARK 2.1. An alternative way [24] to define quaternions is to consider the subset of the ring $M_2(\mathbb{C})$ of 2×2 matrices with complex number entries:

$$\mathbb{Q}' = \left\{ \begin{pmatrix} a_1 & a_2 \\ -\overline{a_2} & \overline{a_1} \end{pmatrix} \middle| a_1, a_2 \in \mathbb{C} \right\}.$$

$\mathbb{Q}'$ is a subring of $M_2(\mathbb{C})$ under the operations of $M_2(\mathbb{C})$. $\mathbb{Q}$ and $\mathbb{Q}'$ are essentially the same. In fact, let

$$\mathscr{M} : q = a_1 + a_2\mathbf{j} \in \mathbb{Q} \to q' = \begin{pmatrix} a_1 & a_2 \\ -\overline{a_2} & \overline{a_1} \end{pmatrix} \in \mathbb{Q}'.$$

Then $\mathscr{M}$ is bijective and preserves the operations. Furthermore, $|q|^2 = \det q'$, and the eigenvalues of q' are $\operatorname{Re} q \pm |\operatorname{Im} q|\mathbf{i}$.

3. THE FUNDAMENTAL THEOREM OF ALGEBRA FOR QUATERNIONS

The fundamental theorem of algebra over the quaternion skew field has gained little attention. It is of less importance than that over the complex number field, though some problems may reduce to it.

Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial in x over $\mathbb{Q}$ and $a_n \neq 0$. It is natural to ask whether $f(x) = 0$ always has a solution and how many solutions $f(x) = 0$ has (as an example, $x^2 + 1 = 0$ has infinitely many solutions in $\mathbb{Q}$). More generally, let $F(x)$ be a polynomial in x with general terms $a_0 x a_1 x \cdots x a_n$, where the a_i's are quaternions. Does $F(x) = 0$ have a solution in $\mathbb{Q}$?

This question was partially answered in [16].

THEOREM 3.1 (Eilenberg and Niven, 1944). *Let*

$$f(x) = a_0 x a_1 x \cdots x a_n + \phi(x), \tag{2}$$

where $a_0, a_1, \ldots, a_n$ are nonzero quaternions, x is a quaternion indeterminant, and $\phi(x)$ is a sum of a finite number of similar monomials $b_0 x b_1 x \cdots x b_k$, $k < n$. Then $f(x) = 0$ has at least one quaternion solution.

The proof in [16] is a topological one and is accomplished by showing that $f(x)$ and $g(x) = x^n$ are homotopic mappings and the latter has degree n.

Note that $f(x)$ has only one term with "degree" n. As a consequence, we have the following result which was first shown in [32].

COROLLARY 3.1 (Niven, 1941). *Let a_t be quaternions, $a_n \neq 0$ and $n \geqslant 1$. Then $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$ has at least one solution in $\mathbb{Q}$.*

Another result worth mentioning is the discussion of the solutions of the equation $xa + bx = c$ over an algebraic division ring [26].

THEOREM 3.2 (Johnson, 1944). *If $a, b, c \in \mathbb{Q}$, and a and b are not similar, then $xa + bx = c$ has a unique solution.*

4. QUATERNION MATRICES AND THEIR ADJOINTS

As has been noticed, there is no theory of eigenvalues, similarity, and triangular forms for matrices with entries in a general ring; and a field can be thought of as a "biggest" algebraic system in which the classical theory of eigenvalues, etc., can be carried through. The question of the extent to which the properties of matrices over fields can be generalized if the commutativity of the fields is removed may be of interest and of potential applications to different research fields. The rest of the paper is devoted to this topic.

Let $M_{m\times n}(\mathbb{Q})$, simply $M_n(\mathbb{Q})$ when $m = n$, denote the collection of all $m \times n$ matrices with quaternion entries. Besides the ordinary addition and multiplication, the *left* [*right*] *scalar multiplication* is defined as follows: For $A = (a_{st}) \in M_{m\times n}(\mathbb{Q})$, $q \in \mathbb{Q}$,

$$qA = (qa_{st}) \qquad [Aq = (a_{st}q)].$$

It is easy to see that for $A \in M_{m\times n}(\mathbb{Q})$, $B_{n\times o}(\mathbb{Q})$, and $p, q \in \mathbb{Q}$

$$(qA)B = Q(AB),$$

$$(Aq)B = A(qB),$$

$$(pq)A = p(qA).$$

Moreover $M_{m\times n}(\mathbb{Q})$ is a left [right] vector space over $\mathbb{Q}$.

All operators for complex matrices can be performed except the ones such as

$$(qA)B = A(qB)$$

in which commutativity is involved.

Just as for complex matrices, we associate to $A = (a_{st}) \in M_{m\times n}(\mathbb{Q})$ with $\overline{A} = (\overline{a_{st}}) = (a_{st}^*)$, the *conjugate* of A; $A^T = (a_{ts}) \in M_{n\times m}(\mathbb{Q})$, the *transpose* of A; and $A^* = (\overline{A})^T \in M_{n\times m}(\mathbb{Q})$, the *conjugate transpose* of A.

A square matrix $A \in M_n(\mathbb{Q})$ is said to be *normal* if $AA^* = A^*A$; *Hermitian* if $A^* = A$; *unitary* if $A^*A = I$, the identity matrix; and *invertible* if $AB = BA = I$ for some $B \in M_n(\mathbb{Q})$.

Just as with complex matrices, one can define *elementary row (column) operations* for quaternion matrices and the corresponding *elementary quaternion matrices*. It is easy to see that applying an elementary row (column) operation to A is equivalent to multiplying A by the corresponding elementary quaternion matrix from left (right) and that any square quaternion matrix can be brought to a diagonal matrix by elementary quaternion matrices.

A list of facts follows immediately, some of which are unexpected.

THEOREM 4.1. *Let $A \in M_{m\times n}(\mathbb{Q})$, $B \in M_{n\times o}(\mathbb{Q})$. Then*

1. $(\overline{A})^T = \overline{(A^T)}$;
2. $(AB)^* = B^*A^*$;
3. $\overline{AB} \neq \overline{A}\,\overline{B}$ *in general*;
4. $(AB)^T \neq B^TA^T$ *in general*;
5. $(AB)^{-1} = B^{-1}A^{-1}$ *if A and B are invertible*;
6. $(A^*)^{-1} = (A^{-1})^*$ *if A is invertible*;
7. $(\overline{A})^{-1} \neq \overline{A^{-1}}$ *in general*;
8. $(A^T)^{-1} \neq (A^{-1})^T$ *in general*.

For part 7 or 8, take $A = \begin{pmatrix} \mathbf{i} & \mathbf{k} \\ 0 & \mathbf{j} \end{pmatrix}$, $A^{-1} = \begin{pmatrix} -\mathbf{i} & -1 \\ 0 & -\mathbf{j} \end{pmatrix}$.

One of the effective approaches to studying matrices of quaternions may be by means of converting a matrix of quaternions into a pair of complex matrices. This was first noticed by Lee [29]. We use his idea to show a proposition below which has appeared in [13, 30] with different proofs.

PROPOSITION 4.1. *Let $A, B \in M_n(\mathbb{Q})$. If $AB = I$, then $BA = I$.*

Proof. First note that the proposition is true for complex matrices.

Let $A = A_1 + A_2\mathbf{j}$, $B = B_1 + B_2\mathbf{j}$, where A_1, A_2, B_1, and B_2 are

$n \times n$ complex matrices. Then

$$
\begin{aligned}
AB = I &\Rightarrow \left(A_1 B_1 - A_2 \overline{B_2}\right) + \left(A_1 B_2 + A_2 \overline{B_1}\right)\mathbf{j} = I \\
&\Rightarrow (A_1, A_2)\begin{pmatrix} B_1 & B_2 \\ -\overline{B_2} & \overline{B_1} \end{pmatrix} = (I, 0) \\
&\Rightarrow \begin{pmatrix} A_1 & A_2 \\ -\overline{A_2} & \overline{A_1} \end{pmatrix}\begin{pmatrix} B_1 & B_2 \\ -\overline{B_2} & \overline{B_1} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \\
&\Rightarrow \begin{pmatrix} B_1 & B_2 \\ -\overline{B_2} & \overline{B_1} \end{pmatrix}\begin{pmatrix} A_1 & A_2 \\ -\overline{A_2} & \overline{A_1} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \\
&\Rightarrow B_1 A_1 - B_2 \overline{A_2} = I, \quad B_1 A_2 + B_2 \overline{A_1} = 0 \\
&\Rightarrow \left(B_1 A_1 - B_2 \overline{A_2}\right) + \left(B_1 A_2 + B_2 \overline{A_1}\right)\mathbf{j} = I \\
&\Rightarrow BA = I.
\end{aligned}
$$

■

For $A = A_1 + A_2\mathbf{j} \in M_n(\mathbb{Q})$, we shall call the $2n \times 2n$ complex matrix

$$\begin{pmatrix} A_1 & A_2 \\ -\overline{A_2} & \overline{A_1} \end{pmatrix},$$

uniquely determined by A, the *complex adjoint matrix* or *adjoint* of the quaternion matrix A, symbolized χ_A. Note that if A is a complex matrx, then

$$\chi_A = \begin{pmatrix} A & 0 \\ 0 & \overline{A} \end{pmatrix}.$$

The adjoint of a quaternion matrix has been exploited in the study of its quaternionic numerical range [3, 4, 27, 36]; this we will enlarge on later.

If

$$P = \begin{pmatrix} 1 & q \\ 0 & 1 \end{pmatrix},$$

for instance, is a 2×2 elementary quaternion matrix, where $q = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k} = (q_0 + q_1\mathbf{i}) + (q_2 + q_3\mathbf{i})\mathbf{j}$, then

$$\chi_P = \begin{pmatrix} 1 & q_0 + q_1\mathbf{i} & 0 & q_2 + q_3\mathbf{i} \\ 0 & 1 & 0 & 0 \\ 0 & -q_2 + q_3\mathbf{i} & 1 & q_0 - q_1\mathbf{i} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Notice that $\det(\chi_P) = |\chi_P| = 1$. More generally,

$$|\chi_A| = |\chi_B| \qquad \text{when} \quad A = \begin{pmatrix} 1 & \alpha \\ 0 & B \end{pmatrix}.$$

THEOREM 4.2 (Lee, 1949). *Let $A, B \in M_n(\mathbb{Q})$. Then*

1. $\chi_{I_n} = I_{2n}$;
2. $\chi_{AB} = \chi_A \chi_B$;
3. $\chi_{A+B} = \chi_A + \chi_B$;
4. $\chi_{A^*} = (\chi_A)^*$;
5. $\chi_{A^{-1}} = (\chi_A)^{-1}$ *if A^{-1} exists*;
6. χ_A *is unitary, Hermitian, or normal if and only if A is unitary, Hermitian, or normal, respectively*;
7. *if J_λ is a Jordan block with λ on the diagonal, then*

$$\chi_{J_\lambda} = \begin{pmatrix} J_\lambda & 0 \\ 0 & \overline{J_\lambda} \end{pmatrix}.$$

The proof of Theorem 4.2 is by direct verification.

PROPOSITION 4.2. *Let A and B be any two $n \times n$ complex matrices. Then*

$$\begin{vmatrix} A & B \\ -\overline{B} & \overline{A} \end{vmatrix} \geqslant 0. \tag{3}$$

In particular,

1. $|\chi_C| \geqslant 0$ *for any* $C \in M_n(\mathbb{Q})$;
2. $|I + \overline{A}A| \geqslant 0$;
3. $|A\overline{A} + B\overline{B}| \geqslant 0$, *whenever A and B commute.*

Proof. It is known [22, p. 253] that for every $n \times n$ complex matrix A the nonreal eigenvalues of $\overline{A}A$ occur in conjugate pairs, and every negative eigenvalue of $\overline{A}A$ has even algebraic multiplicity. It follows that $|I + \overline{A}A| \geqslant 0$, that is,

$$\begin{vmatrix} I & A \\ -\overline{A} & I \end{vmatrix} \geqslant 0,$$

for any complex matrix A.

Now we assume, without loss of generality (using continuity), that A is invertible. Since

$$\begin{pmatrix} A^{-1} & 0 \\ 0 & (\overline{A})^{-1} \end{pmatrix} \begin{pmatrix} A & B \\ -\overline{B} & \overline{A} \end{pmatrix} = \begin{pmatrix} I & A^{-1}B \\ -\overline{A^{-1}B} & I \end{pmatrix},$$

we get

$$\begin{vmatrix} A & B \\ -\overline{B} & \overline{A} \end{vmatrix} \geqslant 0.$$

When A and B commute,

$$\begin{vmatrix} A & \overline{B} \\ -B & \overline{A} \end{vmatrix} = |A\overline{A} + B\overline{B}| \geqslant 0. \qquad \blacksquare$$

THEOREM 4.3. *Let* $A \in M_n(\mathbb{Q})$. *Then the following are equivalent:*

1. *A is invertible;*
2. $Ax = 0$ *has a unique solution* 0.
3. $|\chi_A| \neq 0$, *i.e.,* χ_A *is invertible.*
4. *A has no zero eigenvalue. More precisely, if* $Ax = \lambda x$ *or* $Ax = x\lambda$ *for some quaternion* λ *and some quaternion vector* $x \neq 0$, *then* $\lambda \neq 0$.
5. *A is the product of elementary quaternion matrices.*

Proof. $1 \Rightarrow 2$: This is trivial.

$2 \Rightarrow 3$: Let $A = A_1 + A_2\mathbf{j}$, $x = x_1 + x_2\mathbf{j}$, where A_1 and A_2 are complex matrices, x_1 and x_2 are complex column vectors. Then

$$\begin{aligned} Ax &= (A_1 + A_2\mathbf{j})(x_1 + x_2\mathbf{j}) \\ &= A_1x_1 + A_1x_2\mathbf{j} + A_2\mathbf{j}x_1 + A_2\mathbf{j}x_2\mathbf{j} \\ &= \left(A_1x_1 - A_2\overline{x_2}\right) + \left(A_1x_2 + A_2\overline{x_1}\right)\mathbf{j}. \end{aligned}$$

Thus $Ax = 0$ is equivalent to

$$A_1x_1 - A_2\overline{x_2} = 0$$

and

$$A_1x_2 + A_2\overline{x_1} = 0.$$

Rewriting the above equations as

$$A_1x_1 + A_2\left(-\overline{x_2}\right) = 0$$

and

$$\left(-\overline{A_2}\right)x_1 + \overline{A_1}\left(-\overline{x_2}\right) = 0,$$

we have that $Ax = 0$ if and only if

$$\begin{pmatrix} A_1 & A_2 \\ -\overline{A_2} & \overline{A_1} \end{pmatrix}\begin{pmatrix} x_1 \\ -\overline{x_2} \end{pmatrix} = 0,$$

that is, $\chi_A(x_1, -\overline{x_2})^T = 0$. It is immediate that $Ax = 0$ has a unique solution if and only if $\chi_A(x_1, -\overline{x_2})^T = 0$ does as well; namely, $|\chi_A| \neq 0$.

$3 \Leftrightarrow 4$: This can be seen from $2 \Rightarrow 3$.

$3 \Rightarrow 1$: Suppose that χ_A is invertible and that

$$\begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}\begin{pmatrix} A_1 & A_2 \\ -\overline{A_2} & \overline{A_1} \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

Let $B = B_1 + B_2\mathbf{j}$. It is easy to check that $BA = I$ and A is invertible by Proposition 4.1.

$1 \Leftrightarrow 5$: This is because if A is invertible, then A can be brought to the identity matrix by elementary row and column operations. ∎

REMARK 4.1. Let A be an $n \times n$ quaternion matrix. Applying $\mathscr{M}$ in the Remark 2.1 to each entry of A, one obtains a complex matrix, another representation of A, denoted by $\mathscr{M}_A$, which has been used in the studies of quaternionic numerical ranges [27, 36] and of similarity [39]. It is not difficult to see that there exists a permutation matrix P such that $P^T \chi_A P = \mathscr{M}_A$.

5. EIGENVALUES

We now turn attention to the eigenvalues of quaternion matrices. Since left and right scalar multiplications are different, we need to treat $Ax = \lambda x$ and $Ax = x\lambda$ separately.

A quaternion λ is said to be a *left* (*right*) *eigenvalue* provided that $Ax = \lambda x$ ($Ax = x\lambda$). The set $\{\lambda \in \mathbb{Q} | Ax = \lambda_x \text{ for some } x \neq 0\}$ is called the *left spectrum* of A, denoted by $\sigma_l(A)$. The *right spectrum* is similarly defined and is denoted by $\sigma_r(A)$.

EXAMPLE 5.1. Let

$$A = \begin{pmatrix} 0 & \mathbf{i} \\ \mathbf{j} & 0 \end{pmatrix}.$$

Then the left eigenvalues of A are 1 and $\mathbf{i}$, whereas the right eigenvalues of A are 1 and all the quaternions in $[\mathbf{i}]$.

EXAMPLE 5.2. Let

$$A = \begin{pmatrix} 0 & \mathbf{i} \\ \mathbf{j} & 0 \end{pmatrix}.$$

Then A has two left eigenvalues $\pm 1/(\sqrt{2})(\mathbf{i} + \mathbf{j})$, and infinitely many right eigenvalues which are the quaternions satisfying $\lambda^4 + 1 = 0$. Note that the $\sigma_r(A)$ is not discrete and that $\sigma_l(A) \cap \sigma_r(A) = \phi$.

EXAMPLE 5.3. Let

$$A = \begin{pmatrix} 0 & \mathbf{i} \\ -\mathbf{i} & 0 \end{pmatrix}.$$

Then $\mathbf{k}$ is a left eigenvalue but not a right one. Note that A is Hermitian.

THEOREM 5.1. *Let $A \in M_n(\mathbb{C})$ be an upper triangular matrix. Then a quaternion λ is a left eigenvalue of A if and only if λ is a diagonal entry.*

The proof is basically the same as in the complex case. The situation of general matrices is much more complicated. Generally speaking, there is no very close relation between left and right eigenvalues. For real matrices, however, we have the following theorem.

THEOREM 5.2. *If A is a real $n \times n$ matrix, then the left and right eigenvalues of A coincide; that is, $\sigma_l(A) = \sigma_r(A)$.*

Proof. Let λ be a left eigenvalue of A, i.e., $Ax = \lambda x$ for some $x \neq 0$. For any quaternion $q \neq 0$, we have

$$(qAq^{-1})qx = (q\lambda q^{-1})qx$$

and

$$Aqx = (q\lambda q^{-1})qx,$$

since A is real. Taking $0 \neq q \in \mathbb{Q}$ such that $q\lambda q^{-1}$ is a complex number and writing $qx = y = y_1 + y_2\mathbf{j}$, we have $Ay_1 = y_1 q\lambda q^{-1}$ and $Ay_2 = y_2 q\lambda q^{-1}$. It follows that λ is a right eigenvalue of A. Similarly one can prove that every right eigenvalue is also a left eigenvalue. ∎

For any given matrix $A \in M_n(\mathbb{Q})$, does there always exist $\lambda \in \mathbb{Q}$ and nonzero column vector x of quaternions such that $Ax = \lambda x$? This is of course a very basic question. It was raised in [14, p. 217] and was later proved by Wood [41]. Wood's proof, adopted below, is a purely topological one.

THEOREM 5.3 (Wood, 1985). *Every $n \times n$ quaternion matrix has at least one left eigenvalue in* $\mathbb{Q}$.

Proof. Write $Ax = \lambda x$ as $(\lambda I - A)x = 0$, and assume that $\lambda I - A$ is invertible for all $\lambda \in \mathbb{Q}$.

Consider the general linear group $\mathrm{GL}(n, \mathbb{Q})$, the collection of all invertible $n \times n$ matrices of quaternions.

Let

$$f(t, \lambda) := f_t(\lambda) = t\lambda I - A, \qquad 0 \leqslant t \leqslant 1, \quad |\lambda| = 1,$$

and

$$g(t, \lambda) := g_t(\lambda) = \lambda I - tA, \qquad 0 \leqslant t \leqslant 1, \quad |\lambda| = 1.$$

Then f and g are homotopies on $\mathrm{GL}(n, \mathbb{Q})$. Note that

$$f_0(\lambda) = -A,$$

$$f_1(\lambda) = \lambda I - A = g_1(\lambda),$$

and

$$g_0(\lambda) = \lambda I.$$

Thus g_0 is homotopic to f_0. On the other hand, f_0 and g_0, viewed as the maps from the 3-sphere S^3 into $\mathrm{GL}(n, \mathbb{Q})$, correspond to integers 0 and n, respectively, in $\pi_3\mathrm{GL}(n, \mathbb{Q})$, the third homotopy group (isomorphic to the integers [53]) of $\mathrm{GL}(n, \mathbb{Q})$. This is a contradiction. ■

An elementary proof for the cases of 2×2 and 3×3 matrices has been recently obtained [34].

By contrast, the right eigenvalues have been well studied, and this sort of eigenvalues are more useful. We shall simply call the right eigenvalues *eigenvalues* for the rest of this paper.

LEMMA 5.1. *If $A \in M_{m \times n}(\mathbb{Q})$, $m < n$, then $Ax = 0$ has a nonzero solution.*

Proof. Let $A = A_1 + A_2\mathbf{j}$, $x = x_1 + x_2\mathbf{j}$. Then $Ax = 0$ becomes

$$\begin{pmatrix} A_1 & A_2 \\ -\overline{A_2} & \overline{A_1} \end{pmatrix} \begin{pmatrix} x_1 \\ -\overline{x_2} \end{pmatrix} = 0,$$

which has nonzero complex solutions, since $2m < 2n$. ∎

Using this lemma repeatedly, we have the following

LEMMA 5.2. *Let u_1 be a unit column vector of n quaternion components. Then there exist $n - 1$ unit column vectors $u_2, \dots, u_n$ of n quaternion components such that $\{u_1, u_2, \dots, u_n\}$ is an orthogonal set, i.e., $u_s^* u_t = 0$, $s \neq t$.*

This lemma may also be interpreted as follows: if u_1 is a unit vector, then one can construct an $n \times n$ unitary matrix U with u_1 as its first column.

Since $Ax = x\lambda \Rightarrow A(xq) = (Ax)q = x\lambda q = (xq)(q^{-1}\lambda q)$, it follows that if λ is an eigenvalue of A, then so is $q^{-1}\lambda q$ for any nonzero quaternion q. Thus if λ is a nonreal eigenvalue of A, so is any element in $[\lambda]$. Therefore A has finite eigenvalues if and only if all eigenvalues of A are real.

THEOREM 5.4 (Brenner, 1951; Lee, 1949). *Any $n \times n$ quaternion matrix A has exactly n (right) eigenvalues which are complex numbers with nonnegative imaginary parts.*

Those eigenvalues are said to be the *standard eigenvalues* of A.

Proof. As before, we write A and x as $A = A_1 + A_2\mathbf{j}$, $x = x_1 + x_2\mathbf{j}$, where A_1, A_2 are $n \times n$ complex matrices, and x_1, x_2 are complex column vectors. Then $Ax = x\lambda$ is equivalent to

$$\begin{pmatrix} A_1 & A_2 \\ -\overline{A_2} & \overline{A_1} \end{pmatrix} \begin{pmatrix} x_1 \\ -\overline{x_2} \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ -\overline{x_2} \end{pmatrix} \tag{4}$$

or

$$\begin{pmatrix} A_1 & A_2 \\ -\overline{A_2} & \overline{A_1} \end{pmatrix} \begin{pmatrix} x_2 \\ \overline{x_1} \end{pmatrix} = \overline{\lambda} \begin{pmatrix} x_2 \\ \overline{x_1} \end{pmatrix}, \tag{5}$$

where λ is a complex number.

Since

$$\chi_A = \begin{pmatrix} A_1 & A_2 \\ -\overline{A_2} & \overline{A_1} \end{pmatrix}$$

is a $2n \times 2n$ complex matrix, it has exactly $2n$ complex eigenvalues (including multiplicity).

Notice that if a complex matrix X is similar to its conjugate, then the nonreal eigenvalues of X occur in conjugate pairs (this can be seen by considering $|\lambda I - X| = |\lambda I - \overline{X}| = \overline{|\lambda I - X|}$ for real λ). χ_A is similar to $\overline{\chi_A}$, so the nonreal eigenvalues of χ_A appear in conjugate pairs (with the same multiplicity). With real eigenvalues, we show by induction that every real eigenvalue of χ_A occurs an even number of times.

It is trivial when $n = 1$. Let $n \geqslant 2$.

Suppose $Ax = x\lambda = \lambda x$, where λ is real and $x \neq 0$ is a unit vector. Let $u_2, \dots, u_n$ be such that $U = (x, u_2, \dots, u_n)$ is a unitary matrix, and let

$$U^*AU = \begin{pmatrix} \lambda & \alpha \\ 0 & B \end{pmatrix},$$

where B is an $(n-1) \times (n-1)$ quaternion matrix and α is a quaternion row vector of $n-1$ components. It is easy to see that there is a $2n \times 2n$ invertible matrix T such that

$$T^{-1}\chi_{U^*}\chi_A\chi_U T = \left((\chi_U T)^{-1} \right)\chi_A(\chi_U T) = \begin{pmatrix} \chi_\lambda & \chi_\alpha \\ 0 & \chi_B \end{pmatrix}.$$

By induction, each real eigenvalue of χ_B appears an even number of times. Therefore χ_A has exactly $2n$ eigenvalues symmetrically located on the complex plane, and A has exactly n complex eigenvalues on the upper half complex plane (including the real axis). The conclusion on the real eigenvalues can also be observed by a continuity argument with $A - t\mathbf{i}I$ substituting for A. ■

COROLLARY 5.1 (Lee, 1949). *Let A and B be $n \times n$ complex matrices. Then every real eigenvalue (if any) of the matrix*

$$\begin{pmatrix} A & B \\ -\overline{B} & \overline{A} \end{pmatrix}$$

appears an even number of times, and the complex eigenvalues of that matrix appear in conjugate pairs.

Note that Proposition 4.2 follows from Corollary 5.1 immediately. The structure of the Jordan canonical form of the block matrix in the above corollary will be given in the next section.

The following illustrates why this idea does not come through for the (left eigenvalue) system $Ax = \lambda x$.

Let $\lambda = \lambda_1 + \lambda_2 \mathbf{j}$. Then

$$\begin{aligned} Ax &= (A_1 + A_2\mathbf{j})(x_1 + x_2\mathbf{j}) \\ &= A_1 x_1 + A_1 x_2 \mathbf{j} + A_2 \mathbf{j} x_1 + A_2 \mathbf{j} x_2 \mathbf{j} \\ &= \left(A_1 x_1 - A_2 \overline{x_2}\right) + \left(A_1 x_2 + A_2 \overline{x_1}\right)\mathbf{j}, \end{aligned}$$

and

$$\lambda x = \left(\lambda_1 x_1 - \lambda_2 \overline{x_2}\right) + \left(\lambda_1 x_2 + \lambda_2 \overline{x_1}\right)\mathbf{j}.$$

Thus the quaternion system $Ax = \lambda x$ is equivalent to the complex system

$$\begin{pmatrix} A_1 & A_2 \\ -\overline{A_2} & \overline{A_1} \end{pmatrix} \begin{pmatrix} x_1 \\ \overline{x_2} \end{pmatrix} = \begin{pmatrix} \lambda_1 I & \lambda_2 I \\ \overline{\lambda_2} I & \overline{\lambda_1} I \end{pmatrix} \begin{pmatrix} x_1 \\ \overline{x_2} \end{pmatrix}$$

which is virtually different from (4) and (5).

The above equation can be rewritten as

$$\begin{pmatrix} A_1 - \lambda_1 I & A_2 - \lambda_2 I \\ -\overline{A_2} + \overline{\lambda_2} I & \overline{A_1} - \overline{\lambda_1} I \end{pmatrix} \begin{pmatrix} x_1 \\ -\overline{x_2} \end{pmatrix} = 0.$$

Therefore the square quaternion matrix $A = A_1 + A_2\mathbf{j}$ has a left quaternion eigenvalue if and only if there exist complex numbers λ_1 and λ_2 such that

$$\begin{vmatrix} A_1 - \lambda_1 I & A_2 - \lambda_2 I \\ -\overline{A_2} + \overline{\lambda_2} I & \overline{A_1} - \overline{\lambda_1} I \end{vmatrix} = 0. \tag{6}$$

Note that, by Proposition 4.2, the left-hand side of the above identity is nonnegative. Can this function, regarded as a function on $\mathbb{R}^4$, vanish? It is unknown if this approach can go through without using Wood's result.

COROLLARY 5.2. *Let A be an $n \times n$ matrix with quarternion entries. Then A has exactly n (right) eigenvalues up to equivalent classes.*

QUESTION 5.1. How many left eigenvalues does a square quaternion matrix have?

QUESTION 5.2. Is there an elementary proof for the existence of the left eigenvalues?

QUESTION 5.3. Does there exist a matrix B such that $\chi_A = \overline{B}B$?

QUESTION 5.4. What conditions can be imposed on A and B when

$$\begin{vmatrix} A & B \\ -\overline{B} & \overline{A} \end{vmatrix} = 0?$$

QUESTION 5.5. Investigate $\sigma_l(A)$.

REMARK 5.1. The term "left eigenvalue" has appeared in literature with different meanings. Over a general skew field, it is defined [14, 15] to be the element λ such that $xA = \lambda x$ for some $x \neq 0$. As noticed in [15], a matrix may have no left eigenvalue in this sense. In the quaternion case, however, right and left eigenvalues make no difference in the study, since $xA = \lambda x$ if and only if $A^* x^* = x^* \lambda^*$. The elements λ satisfying $Ax = \lambda x$ for some $x \neq 0$ have been called *singular eigenvalues* of A [14, 41], and are the same as the left eigenvalues in the present paper.

6. CANONICAL FORMS

There are three sorts of canonical forms which are of fundamental importance in linear algebra: rational forms, Schur forms, and Jordan forms. Every square matrix can be reduced under similarity to a direct sum of block matrices over a (commutative) field, and to the Schur and Jordan forms over an algebraically closed field. We discuss the Schur and Jordan canonical forms for quaternion matrices. For the rational canonical form for a matrix over a general skew field, see [15].

THEOREM 6.1 (Brenner, 1949). *If $A \in M_n(\mathbb{Q})$, then there exists a unitary matrix U such that U^*AU is in upper triangular form.*

This can be shown by the mathematical induction and Lemmas 5.1 and 5.2.

As noticed earlier, every diagonal entry of a triangular matrix is a left eigenvalue. The following says that they are also right eigenvalues.

THEOREM 6.2 (Brenner, 1951). *If $A \in M_n(\mathbb{Q})$ is in triangular form, then every diagonal element is a (right) eigenvalue of A. Conversely, every (right) eigenvalue of A is similar to a diagonal entry of A.*

Proof. Let A be an $n \times n$ upper triangular quaternion matrix with diagonal entries $\lambda_1, \lambda_2, \dots, \lambda_n$. If $n = 1$, there is nothing to show.

Suppose it is true for $n - 1$. Let A be of order n and be partitioned as

$$A = \begin{pmatrix} \lambda_1 & \alpha \\ 0 & A_1 \end{pmatrix}.$$

We see that λ_1 is an eigenvalue of A, since

$$A(x_1, 0, \dots, 0)^T = (x_1, 0, \dots, 0)^T \lambda_1, \qquad x_1 \neq 0.$$

A_1 is of order $n - 1$, and has eigenvalues $\lambda_2, \dots, \lambda_n$ by induction. All we need to show is that the eigenvalues of A_1 are eigenvalues of A.

Suppose λ is one of $\lambda_2, \dots, \lambda_n$. We may assume that λ and λ_1 are not similar, and let $A_1 y = y\lambda$, $y \neq 0$. Applying Theorem 3.2, we have a quaternion x such that

$$\lambda_1 x + \alpha y = x\lambda.$$

Then

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda_1 & \alpha \\ 0 & A_1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda_1 x + \alpha y \\ A_1 y \end{pmatrix} = \begin{pmatrix} x\lambda \\ y\lambda \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}\lambda.$$

So λ is an eigenvalue of A. Conversely, if λ is an eigenvalue of A, then the complex number contained in $[\lambda]$ is an eigenvalue of χ_A. On the other hand, the eigenvalues of χ_A are also the eigenvalues of A. The conclusion follows. ∎

It turns out that if a matrix A of quaternions is brought to a triangular form T, then the only eigenvalues of A are the diagonal elements of T and the quaternions similar to them.

The following is the generalization of the Schur canonical form of complex matrices to quaternion matrices, which has been used in the study of numerical ranges of quaternion matrices [3, 4, 36, 48].

THEOREM 6.3 (Brenner, 1951; Lee, 1949). *Let A be an $n \times n$ matrix of quaternions. Then there exists a unitary matrix U such that U^*AU is an upper triangular matrix with diagonal entries $h_1 + k_1\mathbf{i}, \dots, h_n + k_n\mathbf{i}$, where the $h_t + k_t\mathbf{i}$'s are the standard eigenvalues of A, i.e., $k_t \geqslant 0$, $t = 1, 2, \dots, n$.*

For the proof, A is first brought to an upper triangular matrix with quaternions on the main diagonal by a unitary matrix, then each diagonal quaternion is changed to a complex number with nonnegative imaginary part by a pair of elementary quaternion matrices.

Note that Corollary 5.1 follows from Theorem 6.3 immediately.

COROLLARY 6.1. *Let A be an $n \times n$ quaternion matrix with standard eigenvalues $h_1 + k_1\mathbf{i}, \dots, h_n + k_n\mathbf{i}$. Then*

$$\sigma_r(A) = [h_1 + k_1\mathbf{i}] \cup \cdots \cup [h_n + k_n\mathbf{i}].$$

COROLLARY 6.2. *A is normal if and only if there exists a unitary matrix U such that*

$$U^*AU = \operatorname{diag}\{h_1 + k_1\mathbf{i}, \dots, h_n + k_n\mathbf{i}\}$$

and A is Hermitian if and only if $k_1 = \cdots = k_n = 0$.

We now consider the Jordan canonical form of an $n \times n$ matrix with quaternion entries. Let $A \in M_n(\mathbb{Q})$. If it is possible for A to have the Jordan form J with the standard eigenvalues of A on the diagonal of J; that is, $S^{-1}AS = J$ for some quaternion invertible matrix S, then

$$\chi_{S^{-1}}\chi_A\chi_S = \chi_J = \begin{pmatrix} J & 0 \\ 0 & \bar{J} \end{pmatrix}.$$

This means that χ_J is the Jordan form of χ_A. Conversely, let $\hat{J}$ be the Jordan form of χ_A, and let P be an invertible matrix such that $P^{-1}\chi_A P = \hat{J}$. If it can be shown that P and $\hat{J}$ take the forms

$$\begin{pmatrix} P_1 & P_2 \\ -\overline{P_2} & \overline{P_1} \end{pmatrix} \text{ and } \begin{pmatrix} J_1 & 0 \\ 0 & \overline{J_1} \end{pmatrix},$$

respectively, then A is similar to a complex $n \times n$ Jordan form J_1 by a direct verification. Following this line, a thorough analysis of the eigenvalues and the corresponding eigenvectors of χ_A result in the existence of P and J with the desired form [40].

THEOREM 6.4 (Wiegmann, 1954). *Every $n \times n$ quaternion matrix A has the Jordan canonical form with the standard eigenvalues of A on the diagonal.*

COROLLARY 6.3. *For any $n \times n$ complex matrices A and B, the block matrix*

$$\begin{pmatrix} A & B \\ -\overline{B} & \overline{A} \end{pmatrix}$$

has the Jordan canonical form

$$\begin{pmatrix} J & 0 \\ 0 & \overline{J} \end{pmatrix},$$

where J is a Jordan form of some $n \times n$ complex matrix. Consequently, all the Jordan blocks are paired.

REMARK 6.1. As we have noted, many terms for describing complex matrices can similarly be defined for matrices of quaternions, and the parallel conclusions may be derived for the quaternion cases. For instance, an $n \times n$ Hermitian matrix A is said to be *positive (semi-)definite* if $x^*Ax > (\geqslant) 0$ for all nonzero column vectors x of n quaternion components. It is seen that a Hermitian matrix A of quaternion is positive semidefinite if and only if A has only nonnegative eigenvalues, and if and only if χ_A is positive semidefinite.

7. RANK, SIMILARITY, AND DECOMPOSITIONS

Many aspects of quaternion matrices, such as rank, linear independence, similarity, characteristic matrix, Gram matrix, and determinantal expansion theorem, have been discussed in [11–13], [14, 15], [39], [40], and [42–46].

One can define, in the usual sense, the left and right linear independence over $\mathbb{Q}$ for a set of vectors of quaternions. Note that two linearly dependent vectors over $\mathbb{Q}$ may be linearly independent over $\mathbb{C}$. One can also easily find an example of two vectors which are left linearly dependent but right linearly independent. The Gram-Schmidt process (see, e.g., [22, p. 15]) is still effective.

The *rank* of a quaternion matrix A is defined to be the maximum number of columns of A which are *right* linearly independent. It is easy to see that for any invertible matrices P and Q of suitable sizes, A and PAQ have the same rank. Thus the rank of A is equal to the number of positive *singular values* of A (see Theorem 7.2). If a matrix A is of rank r, then r is also the maximum number of rows of A that are left linearly independent, and A is nonsingular (or invertible, meaning $BA = AB = I$ for some B) if and only if A is of (full) rank n.

Let $\mathbb{Q}_c^n$ denote the collection of column vectors with n components of quaternions. $\mathbb{Q}_c^n$ is a right vector space over $\mathbb{Q}$ under the addition and the right scalar multiplication. If A is an $m \times n$ quaternion matrix, then the solutions of $Ax = 0$ form a subspace of $\mathbb{Q}_c^n$, and the subspace has dimension r if and only if A has rank $n - r$.

We now give the polar and the singular-value decompositions for quaternion matrices.

The proof of the following theorem is based on [40].

THEOREM 7.1 (Polar decomposition). *Let $A \in M_n(\mathbb{Q})$. Then there exist a quaternion unitary matrix U and a quaternion positive semidefinite Hermitian matrix H such that $A = HU$. Moreover H and U are unique when A is of rank n.*

Proof. Let X be an $n \times n$ quaternion unitary matrix such that $X^*AA^*X = D$ is a diagonal matrix with the squares of the singular values of A on the diagonal. Let $\chi_A = KY$ be a polar decomposition of the complex matrix χ_A, where K is a $2n \times 2n$ positive semidefinite Hermitian (complex) matrix and Y is a $2n \times 2n$ unitary (complex) matrix. We show that $K = \chi_H$ for some positive semidefinite Hermitian matrix H, and $Y = \chi_U$ for some unitary quaternion matrix U.

First note that

$$\chi_A \chi_{A^*} = K^2, \qquad \chi_{X^*} \chi_A \chi_{A^*} \chi_X = \chi_D.$$

Thus

$$\chi_{X^*} K^2 \chi_X = \chi_D$$

and therefore

$$\chi_{X^*} K \chi_X = \chi_{D^{1/2}},$$

so

$$K = (\chi_{X^*})^{-1} \chi_{D^{1/2}} (\chi_X)^{-1} = \chi_{XD^{1/2}X^*}.$$

Take $H = XD^{1/2}X^*$. Then $\chi_A = \chi_H Y$. We assume that A is nonsingular (a little more work is needed for the singular case). Then H is nonsingular and

$$Y = (\chi_H)^{-1} \chi_A = \chi_{H^{-1}} \chi_A = \chi_{H^{-1}A}.$$

Set $U = H^{-1}A$. Then $A = HU$, and H and U are as desired. ∎

The following theorem follows immediately.

THEOREM 7.2 (Singular-value decomposition). *Let $A \in M_{m \times n}(\mathbb{Q})$ be of rank r. Then there exist unitary quaternion matrices $U \in M_m(\mathbb{Q})$ and $V \in M_n(\mathbb{Q})$ such that*

$$UAV = \begin{pmatrix} D_r & 0 \\ 0 & 0 \end{pmatrix}$$

where $D_r = \operatorname{diag}\{d_1, \dots, d_r\}$ and the d's are the positive singular values of A.

It is always desirable to convert a quaternion matrix problem into a complex one. The next two theorems, following from Theorem 7.2 and Theorem 6.4, are in this direction.

THEOREM 7.3 (Wolf, 1936). *The rank of a quaternion matrix A is r if and only if A has r nonzero singular values, and if and only if the rank of its complex adjoint χ_A is $2r$.*

It is readily seen that A, A^*, AA^*, and A^*A are all of the same rank.

Two square matrices of quaternions A and B are said to be *similar* if there exists an invertible quaternion matrix S of the same size such that $S^{-1}AS = B$. It is immediate that similar quaternion matrices have the same (right) eigenvalues. This is not true for left eigenvalues. As one sees below, even the trace is not preserved under similarity. Recall that the trace of a square matrix is the sum of the main diagonal entries of the matrix, which is equal to the sum of the eigenvalues of the matrix in the complex case.

EXAMPLE 7.1. Let

$$A = \begin{pmatrix} \mathbf{i} & 0 \\ 0 & \mathbf{i} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} \mathbf{j} & 0 \\ 0 & \mathbf{j} \end{pmatrix}.$$

Then A and B are similar, but they have different traces and different left eigenvalues.

EXAMPLE 7.2. Let

$$A = \begin{pmatrix} \mathbf{i} & \mathbf{i} \\ 0 & -\mathbf{i} \end{pmatrix}, \qquad U = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\mathbf{j} \\ -\frac{1}{\sqrt{2}}\mathbf{j} & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad \text{and} \quad B = U^*AU.$$

Then U is unitary and the main diagonal entries of B are $\mathbf{i} = \frac{1}{2}\mathbf{k}$ and $-\mathbf{i} - \frac{1}{2}\mathbf{k}$. However, $\operatorname{tr} A = 0$, $\operatorname{tr} B = -\mathbf{k}$. Note that $\operatorname{tr} A$ and $\operatorname{tr} B$ have the same real part, but they are neither necessarily equal nor even similar (as quaternions).

EXAMPLE 7.3. Let

$$A = \begin{pmatrix} 1 & \mathbf{i} \\ \mathbf{j} & \mathbf{k} \end{pmatrix}.$$

Then the column vectors of A are left linearly dependent and right linearly independent. A is of rank 2, and is invertible.

It is easy to show that an $n \times n$ quaternion matrix A is diagonalizable if and only if A has n right linearly independent eigenvectors belonging to right eigenvalues. The following example says that the eigenvectors belonging to different eigenavlues are not necessarily (right or left) linearly independent.

EXAMPLE 7.4. Let

$$A = \begin{pmatrix} \mathbf{i} & 1 \\ 0 & \mathbf{j} \end{pmatrix}.$$

Then the eigenvectors $(1, 0)^T$ and $(\mathbf{i} + \mathbf{j}, 0)^T$ belonging to the (right) eigenvalues $\mathbf{i}$ and $\mathbf{j}$, respectively, are not linearly independent. In addition, A is not diagonalizable. Note that if an $n \times n$ matrix has n distinct right eigenvalues, no pair of which are similar, then the matrix is diagonalizable.

THEOREM 7.4 (Wolf, 1936). *Let A and B be $n \times n$ quaternion matrices. Then A and B are similar if and only if χ_A and χ_B are similar.*

COROLLARY 7.1. *$A \in M_n(\mathbb{Q})$ is diagonalizable if and only if χ_A is diagonalizable.*

COROLLARY 7.2. *A complex matrix is diagonalizable over $\mathbb{Q}$ if and only if it is diagonalizable over $\mathbb{C}$.*

QUESTION 7.1. Suppose $A \in M_n(\mathbb{Q})$ has n distinct left eigenvalues, any two of which are not similar. Is A diagonalizable?

REMARK 7.1. Some parallel properties of normal quaternion matrices to those of complex normal matrices are obtained in [40]. For example, two normal quaternion matrices A and B are commutative if and only if they can be diagonalized by the same unitary transformation.

8. DETERMINANTS AND THE CAYLEY-HAMILTON THEOREM

We now discuss the determinants of quaternion matrices. For this purpose, we first define the determinant of a square quaternion matrix A to be that of its complex adjoint χ_A, then derive some results such as the Cayley-

Hamilton theorem, which have been obtained by several authors in the last decade through extremely difficult approaches.

Let A be an $n \times n$ quaternion matrix, and let χ_A be the complex adjoint matrix of A. We define the *q-determinant* of A to be $|\chi_A|$, simply called the *determinant* of A, denoted by $|A|_q$, i.e., $|A|_q = |\chi_A|$ by definition.

It is immediate that $|A|_q = |A||\bar{A}| = |\det A|^2$ when A is a complex matrix.

We have the following results on determinants of quaternion matrices.

THEOREM 8.1. *Let A and B be $n \times n$ quaternion matrices. Then:*

1. *A is invertible $\Leftrightarrow |A|_q \neq 0$.*
2. *$|AB|_q = |A|_q|B|_q$, consequently $|A^{-1}|_q = |A|_q^{-1}$, if A^{-1} exists.*
3. *$|PAQ|_q = |A|_q$, for any elementary matrices P and Q.*
4. *$|A|_q = \prod_{t=1}^{n}|\lambda_t|^2 \geqslant 0$, where the λ_t's are the standard eigenvalues of A.*

5. Cayley-Hamilton theorem: *Let A be a square matrix of quaternions, and $F_A(\lambda) = |\lambda I_{2n} - \chi_A|$, called the* characteristic polynomial *of A, where λ is a complex indeterminant. Then $F_A(A) = 0$, and $F_A(\lambda_0) = 0$ if and only if λ_0 is an eigenvalue.*

6. *If A and B are similar, then $|A|_q = |B|_q$ and $F_A(\lambda) = F_B(\lambda)$.*

7. Hadamard theorem: *Let $A = (a_{st})$ be positive semidefinite Hermitian. Then $|A|_q \leqslant \prod_{t=1}^{n} q_{tt}^2$, with equality if and only if A is real diagonal. Furthermore, if A is partitioned as*

$$A = \begin{pmatrix} B & C \\ C^* & D \end{pmatrix},$$

then $|A|_q \leqslant |B|_q|D|_q$, with equality if and only if $C = 0$ or $|B|_q|D|_q = 0$.

Proof. 1: See Theorem 4.3.

2: This is because $\chi_{AB} = \chi_A \chi_B$.

3: It is sufficient to observe that $|\chi_P| = 1$ when P is an elementary quaternion matrix.

4: $|A|_q \geqslant 0$ was proved in Proposition 4.2 (notice that the proof did not involve quaternion matrices).

We know from Theorem 6.3 that for any $n \times n$ quaternion matrix A there exists a unitary matrix U such that $A = U^*DU$, where D is an upper triangular matrix with the standard eigenvalues $\lambda_1, \ldots, \lambda_n$ of A on the diagonal. It is easy to compute $|A|_q = |\chi_A| = |\chi_D| = \prod_{t=1}^{n}|\lambda_t|^2$.

5: It follows from Corollary 5.1 that $F_A(\lambda)$ is a real coefficient polynomial. And $\chi_{f(A)} = f(\chi_A)$ for any real coefficient polynomial f.
Noting that $F_A(\chi_A) = 0$ by the Cayley-Hamilton theorem for complex matrices, we have $\chi_{F_A(A)} = F_A(\chi_A) = 0$, hence $F_A(A) = 0$, as desired.

The second conclusion of this part follows from the fact that every quaternion is similar to a complex number and that $Ax = x\lambda$ is equivalent to $\chi_A(x_1, -\overline{x_2})^T = \lambda(x_1, -\overline{x_2})^T$, where $x = x_1 + x_2\mathbf{j}$ and λ is a complex number.

6: This is due to the fact that χ_A and χ_B are similar.

7: If A is positive semidefinite Hermitian, we write $A = G^*G$. Then $\chi_A = (\chi_G)^*\chi_G$; therefore χ_A is a complex positive semidefinite Hermitian matrix (the converse is also true). By the Hadamard theorem for complex positive semidefinite Hermitian matrices, we have

$$|A|_q = |\chi_A| = \begin{vmatrix} A_1 & A_2 \\ A_2^* & \overline{A_1} \end{vmatrix} \leqslant \prod_{t=1}^{n} a_{tt}^2,$$

with equality if and only if A is (real) diagonal.

The latter conclusion follows from that

$$\chi_A = P^T \begin{pmatrix} \chi_B & \chi_C \\ (\chi_C)^* & \chi_D \end{pmatrix} P,$$

where P is a product of some permutation matrices. ■

From the point of view of how the theory of determinants is algebraically developed for skew fields, it may be worthwhile defining determinants of quaternion matrices by their entries. Several versions of the definition of determinants for quaternion matrices have appeared [11–13, 30, 42–46]. The definition introduced below, of a *double determinant*, is the one that gives us many properties resembling ordinary ones.

Let S_n be the symmetric group on $\{1, 2, \ldots, n\}$. For any $\sigma \in S_n$, we write σ as the product of disjoint cycles:

$$\sigma = (n_1 i_2 i_3 \cdots i_s)(n_2 j_2 j_3 \cdots j_t) \cdots (n_r k_2 k_3 \cdots k_l),$$

with the convention that each n_t, $t = 1, 2, \ldots, r$, is the largest number in its cycle and that $n = n_1 > n_2 > \cdots > n_r \geqslant 1$. For example, if σ is the identity, then $\sigma = (n)(n-1)\cdots(2)(1)$.

Let $\epsilon(\sigma)$ be the sign of permutation σ. For $A \in M_n(\mathbb{Q})$, we associate to A the quaternion

$$\det A = \sum_{\sigma \in S_n} \epsilon(\sigma) a_{n_1 i_2} a_{i_2 i_3} \cdots a_{i_{s-1} i_s} a_{i_s n_1} a_{n_2 j_2} a_{j_2 j_3} \cdots a_{j_{t-1} j_t} a_{j_t n_2}$$

$$\times \cdots a_{n_r k_2} a_{k_2 k_3} \cdots a_{k_{l-1} k_l} a_{k_l n_r}$$

$$= \sum_{\sigma \in S_n} \epsilon(\sigma) a_\sigma .$$

It is obvious that $\det A = 1$ if A is the identity matrix, and $\det A$ is the same as the usual determinant when all a_{st}'s commute. $\det A$ so defined carries few properties of the determinants of complex matrices. For instance, the column vectors of

$$A = \begin{pmatrix} \mathbf{i} & \mathbf{k} \\ \mathbf{j} & 1 \end{pmatrix}$$

are right linearly independent, but $\det A = \mathbf{i} - \mathbf{jk} = 0$.

Denote $|A|_d = \det(A^*A)$, called the *double determinant* of A. Based on this definition, Chen [11–13] and Xie [42–46] obtained the results in the previous theorem in their series of work, with extraordinary effort.

We next draw the conclusion that $|A|_q$ and $|A|_d$ are indeed the same. As we have seen, $|A|_q = \prod_{t=1}^n |\lambda_t|^2$. To show that $|A|_d$ is the same quantity, we need a result of Chen [13, Theorem 4 and 5].

LEMMA 8.1 (Chen, 1991). *For any $n \times n$ quaternion matrices A and B,*

$$|A|_d = |AT|_d = |TA|_d, \qquad |AB|_d = |A|_d |B|_d,$$

where T is any matrix obtained from the $n \times n$ identity matrix by adding a quaternion scalar multiple of one row to another. In particular, $|P^{-1}|_d = |P|_d^{-1}$ when P is an invertible matrix.

It follows from the definition and the above lemma that if D is an upper triangular matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$, then $|D|_d = \prod_{t=1}^n |\lambda_t|^2$.

Applying Theorem 6.3, we have

THEOREM 8.2. *For any quaternion matrix* A, $|A|_q = |A|_d$.

Noticing that A^*A is a positive definite Hermitian matrix for every $A \in M_n(\mathbb{Q})$ and using Theorem 8.1, part 7, one obtains

THEOREM 8.3 (Chen, 1991; Xie, 1979). *Let A be an $n \times n$ quaternion matrix. Then*

$$|A|_d \leqslant \prod_{j=1}^{n} \sum_{i=1}^{n} \overline{a_{ij}} a_{ij},$$

and equality holds if and only if the column vectors of A are mutually orthogonal.

9. NUMERICAL RANGES OF MATRICES WITH QUATERNION ENTRIES

Numerical ranges of complex matrices have been a very popular topic in linear algebra (see, e.g., [22, 23]). The purpose of this section is to study the numerical range of a quaternion matrix. Particular attention is paid to the convexity of the quaternionic numerical range of a matrix lying in the closed upper half complex plane.

Let H be a Hilbert space over $\mathbb{C}$, and let A be an operator on H. The classical *complex numerical range* of the operator A, or the *field of values* of A [21], comprises all complex numbers $(A(x), x)$ as x runs over the unit sphere $\{x \mid x \in H, \|x\| = 1\}$ in H. If A is an $n \times n$ complex matrix, the *complex numerical range* or the *field of values* [23] of A is the subset $\{x^*Ax \mid x \in \mathbb{C}^n, x^*x = 1\}$ of the complex plane.

A celebrated result known as Toplitz-Hausdorff theorem ensures the conversity of the complex numerical range of an operator (a complex matrix).

The quaternionic numerical range of an operator A on a (left) quaternionic Hilbert space over $\mathbb{Q}$ [37] or a square matrix with quaternion entries [23, p. 86] is similarly defined, and denoted by $W_q(A)$. To avoid ambiguity, we designate by $W_c(A)$ the classical numerical range, i.e., the complex numerical range of a square matrix A with complex entries.

The study of the convexity of $W_q(A)$ as a subset of $\mathbb{Q}$ was begun by Kippenhahn [37]. Since $W_q(A)$ is a subset of the real four-dimensional space $\mathbb{Q}$, it is not easy to visualize or calculate. The first natural question is what the part of $W_q(A)$ lying in $\mathbb{R}$ or $\mathbb{C}$ looks like; that is, what is $W_q(A) \cap \mathbb{R}$ or $W_q(A) \cap \mathbb{C}$? This problem has been investigated in [27, 25, 3–5, 36, 35, 47, 48]. We briefly state the existing results, then reduce the convexity of $W_q(A)$

to that of its intersection with the complex plane, and further reduce the general matrix case to 2-by-2 matrices. For normal quaternion matrices, we will give explicit characterizations of $W_q(A) \cap \mathbb{R}$ and $W_q(A) \cap \mathbb{C}$.

Let $B(A) = W_q(A) \cap \mathbb{C}$. We call $B(A)$, as Kippenhahn did, the *bild* (or representative) of A, and the subset of $B(A)$ lying in the closed upper half plane the *upper bild* of A, denoted by $B_+(A)$. Note that a complex number $c \in B(A)$ if and only if $\bar{c} \in B(A)$, by Theorem 2.1, part 5. Thus $B(A) = B_+(A) \cup B_-(A)$, where $B_-(A)$ is the reflection of $B_+(A)$ about the x-axis in the complex plane. As we will see shortly, $B(A)$ is not convex in general, while $W_c(\chi_A)$, symmetrically suited about the x-axis since χ_A and $\overline{\chi_A}$ are similar, is convex, Note that if $x \in W_q(A)$, then any quaternion similar to x is contained in $W_q(A)$. In particular, $x^* \in W_q(A)$.

A comparison of $B(A)$ and $W_c(\chi_A)$ is given below.

THEOREM 9.1 (Kippenhahn, 1951). *Let A be a quaternion matrix. Then*:

1. $B(A) \subseteq W_c(\chi_A)$.
2. *The upper envelope of $B(A)$ coincides with the upper envelope of $W_c(\chi_A)$. More precisely, suppose $h_l + k_l\mathbf{i}$, $h_r + k_r\mathbf{i} \in B_+(A)$ with the properties that $h_l \leqslant h_r$, $k_l, k_r \geqslant 0$, and $h_l \leqslant \operatorname{Re} z \leqslant h_r$ for all $z \in B_+(A)$. Then the portions of $B(A)$ and $W_c(\chi_A)$ above the line segment joining $h_l + k_l\mathbf{i}$ and $h_r + k_r\mathbf{i}$ coincide.*

To prove part 1, let $c = u^*Au \in B(A)$. Then $\chi_c = \chi_{u^*}\chi_A\chi_u$. Write $u = u_1 + u_2\mathbf{j}$. It is easy to check that $c = \hat{u}^*\chi_A\hat{u} \in W_c(\chi_A)$, where $\hat{u} = (u_1, -\overline{u_2})^T$ and $\hat{u}$ is a unit vector when u is a unit vector. For (2), let $c = x^*\chi_A x \in W_c(\chi_A)$, where $x = (x_1, x_2, \dots, x_{2n})^T \in \mathbb{C}^{2n}$. Set $u = (x_1 - \overline{x_2}\mathbf{j}, x_3 - \overline{x_4}\mathbf{j}, \dots, x_{2n-1} - \overline{x_{2n}}\mathbf{j})^T$. The conclusion then follows from a careful analysis [36] of the quaternion $q = u^*Au$.

With this theorem one can see that the numerical radius [25] and the spectral norm [10] of a quaternion matrix A, defined in the usual sense, are equal to those of the complex adjoint χ_A of A, respectively.

Kippenhahn asserted in [27] that the $B(A)$ coincided with $W_c(\chi_A)$, and consequently $B(A)$ was convex. But his proof of this part is false [25], as the following example due to Au-Yeung [5] illustrates.

EXAMPLE 8.1. Take

$$A = \begin{pmatrix} \mathbf{i} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then $\mathbf{i}, -\mathbf{i} \in B(A) \subset W_q(A)$, but $0 = \frac{1}{2}\mathbf{i} + \frac{1}{2}(-\mathbf{i}) \notin B(A)$. Thus $B(A)$, and therefore $W_q(A)$, is not convex. Notice that A is a complex matrix and its complex numerical range is convex. Some properties of quaternionic numerical ranges are not parallel to those of complex numerical ranges.

The connectedness of some subsets has been observed.

THEOREM 9.2 (Jamison, 1972). *For any operator A on a quaternionic Hilbert space, $W_q(A) \cap \mathbb{R}$ is either empty or a closed interval of $\mathbb{R}$.*

The proof by Jamison is a computational one. Considering an operator on a left Hilbert space H over $\mathbb{Q}$, Au-Yeung [5] showed that the set $\{u \in H \mid (Au, u) = 0,\ \|u\| = 1\}$ is connected when A is skew-Hermitian (i.e., $A^* = -A$) and the set $\{u \in H \mid (Au, u) = a,\ \|u\| = 1\}$ is connected when a is real and A is Hermitian. Based on these observations, he gave another proof of Jamison's result and showed the following theorem.

For convenience, we denote the projections of $W_q(A)$ on $\mathbb{R}$ and $\mathbb{C}$ by

$$W_q(A:\mathbb{R}) = \{\operatorname{Re} q \mid q \in W_q(A)\}$$

and

$$W_q(A:\mathbb{C}) = \{\operatorname{Co} q \mid q \in W_q(A)\}$$

respectively.

THEOREM 9.3 (Au-Yeung, 1984). *Let A be an operator on a quaternionic Hilbert space. Then*

1. *$W_q(A)$ is convex if and only if $W_q(A) \cap \mathbb{R} = W_q(A:\mathbb{R})$;*
2. *$W_q(A)$ is convex if and only if $W_q(A) \cap \mathbb{C} = W_q(A:\mathbb{C})$;*
3. $\operatorname{conv}(W_q(A)) = \{a + p \mid a \in \mathbb{R}, \operatorname{Re} p = 0$ *and* $a + |p|\mathbf{i} = \operatorname{Co} q$ *for some* $q \in W_q(A)\}$, *where* conv *means "convex hull of"*;
4. *if $a + p_1, a + p_2 \in W_q(A)$ where $a \in \mathbb{R}$,* $\operatorname{Re} p_1 = \operatorname{Re} p_2 = 0$, *then for any p with* $\operatorname{Re} p = 0$, *one has $a + p \in W_q(A)$ whenever $|p|$ lies between $|p_1|$ and $|p_2|$.*

The last part of the previous theorem ensures that the intersection of $B_+(A)$ with any vertical line is an either empty set or a line segment. The same is true for horizontal lines [4].

The projection of $W_q(A)$ on $\mathbb{C}$ is always convex [3], since

$$W_q(A:\mathbb{C}) = W_c\left(\begin{pmatrix} A_1 & -A_2 \\ \overline{A_2} & \overline{A_1} \end{pmatrix}\right),$$

as can be seen by verifying that

$$\mathrm{Co}(u^*Au) = (u_1^*, u_2^*)\begin{pmatrix} A_1 & -A_2 \\ \overline{A_2} & \overline{A_1} \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

where $A = A_1 + A_2\mathbf{j}$ and $u = u_1 + u_2\mathbf{j}$, as before.

Note that if $q = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k} \in W_q(A)$, then

$$q_0 + \sqrt{q_1^2 + q_2^2 + q_3^2}\,\mathbf{i} \in W_q(A) \cap \mathbb{C}$$

and

$$q_0 - \sqrt{q_1^2 + q_2^2 + q_3^2}\,\mathbf{i} \in W_q(A) \cap \mathbb{C}.$$

Thus $q_0 \in W_q(A) \cap \mathbb{R}$ if $W_q(A) \cap \mathbb{C}$ is convex. It turns out, by Theorem 9.3, part 1, that $W_q(A)$ is convex if and only if $B(A) = W_q(A) \cap \mathbb{C}$ is convex.

It is well known that the complex numerical range of a complex normal matrix is the convex hull of its eigenvalues. The quaternionic analogue has been studied recently [3, 4, 36, 47, 48].

As seen in Example 8.1, the quaternionic numerical range of a normal matrix is not a convex hull of eigenvalues, and not even convex in general. Yet a somewhat hidden convexity that can be described in terms of eigenvalues has been observed by So, Thompson, and Zhang [36, 47, 48].

THEOREM 9.4 (So, Thompson, and Zhang, 1994). *Let A be a normal matrix of quaternions with standard eigenavlues $h_1 + k_1\mathbf{i}, \dots, h_n + k_n\mathbf{i}$. Then $B_+(A)$ is the convex hull of these eigenvalues of A, $(l, 0)$, and $(r, 0)$, where*

$$l = \min_{1 \leqslant s < t \leqslant n} \left\{ \frac{h_s k_t + h_t k_s}{k_s + k_t} \right\}$$

and

$$r = \max_{1 \leqslant s < t \leqslant n} \left\{ \frac{h_s k_t + h_t k_s}{k_s + k_t} \right\},$$

under the convention that $(h_s k_t + h_t k_s)/(k_s + k_t)$ *means a pair* $\{h_s, h_t\}$ *when* $k_s = k_t = 0$. *Consequently* $B_+(A)$ *is convex.*

The main idea of the proof in [36] or [47] is to convert the upper-bild investigation into an extremal problem of eigenvalues with side conditions, then employ Langrange multipliers. Au-Yeung [3] later shortened the proof by showing that the upper bild coincides with the union of certain consecutive triangles.

Just as for the case of complex matrices [23, p. 18], the convexity of the upper bild of a quaternion matrix is easily reduced to that of 2×2 cases.

THEOREM 9.5. *If the upper bild of any* 2×2 *quaternion matrix is convex, then the upper bild of any* $n \times n$ *quaternion matrix A is convex.*

By splitting a general 2×2 matrix into Hermitian and skew-Hermitian parts and using a pair of suitable unitary matrices on A [35], one may further reduce the convexity for a 2×2 matrix to that for the matrix

$$G = \begin{pmatrix} k_1 \mathbf{i} & \alpha + \beta \mathbf{j} \\ -\alpha + \beta \mathbf{j} & 1 + k_2 \mathbf{i} \end{pmatrix},$$

where α, β, k_1, and k_2 are nonnegative numbers.

Unlike the normal case in which an upper bild is generally a polygon, the boundary of the upper bild of G is a curve, not necessarily formed by line segments. It is readily seen from Kippenhahn's theorem that the upper bounding curve (envelope) of $B_+(G)$ is concave downward. To prove the convexity, it must be shown that the lower bounding curve is concave upward. This is extremely difficult. With the help of Mathematica computations and in a great effort over years, So and Thompson have accomplished the proof in a very recent manuscript [35] (65 pages long).

THEOREM 9.6 (So and Thompson, 1995). *Let A be an* $n \times n$ *quaternion matrix. Then* $B_+(A)$ *is a convex subset of the complex plane* $\mathbb{C}$.

Many questions regarding a quaternionic numerical range may be asked even though the main problem for the finite-dimensional case has been settled.

QUESTION 9.1. Is there a short and conceptual proof for Theorem 9.6?

QUESTION 9.2. How are $B(A)$ and $W_c(\chi_A)$ related? What is the area ratio of these two sets?

QUESTION 9.3. Is there a simple proof (without using Theorem 9.6) for the statement that the quaternionic numerical range of the diagonal matrix formed by the diagonal entries (or eigenvalues) of A is contained in $W_q(A)$?

QUESTION 9.4. Investigate $W_q(A)$ and $B_+(A)$ when A is a linear operator on a left quaternionic Hilbert space of infinite dimension.

The author wishes to express his deep thanks to Professor R. C. Thompson and Professor W. So for their encouragement and helpful discussions, and Professor L. C. Chen for sending him some preprints. Thanks also go to the referee for many suggestions.

REFERENCES

1 S. L. Adler, *Quaternionic Quantum Mechanics and Quantum Fields*, Oxford U.P., New York, 1994.
2 T. M. Apostol, *Mathematical Analysis*, 2nd ed., Addison-Wesley, 1977.
3 Y. H. Au-Yeung, A short proof of a theorem on the range of a normal quaternionic matrix, *Linear and Multilinear Algebra*, to appear.
4 Y. H. Au-Yeung, 1994. On the eigenvalues and numerical range of a quaternionic matrix, Preprint.
5 Y. H. Au-Yeung, On the convexity of numerical range in quaternionic Hilbert spaces, *Linear and Multilinear Algebra* 16:93–100 (1984).
6 W. C. Brown, *Matrices over Commutative Rings*. Marcell Dekker, 1992.
7 G. Birkhoff and S. MacLane, *A Survey of Modern Algebra*, 4th ed. Macmillan, 1977.
8 J. L. Brenner, Matrices of quaternions, *Pacific J. Math.* 1:329–335 (1951).
9 A. Bunse-Gerstner, R. Byers, and V. Mehrmann, A quaternion QR algorithm, *Numer. Math.* 55:83–95 (1989).
10 C. G. Cao, The numerical radius of real quaternion matrix. *J. Xinjiang Univ.* 7, No. 2 (1990); also see MR 89k:15031.
11 L. X. Chen, Definition of determinant and Cramer solutions over the quaternion field, *Acta Math. Sinica N. S.* 7(2):171–180 (1991).
12 L. X. Chen, The extension of Cayley-Hamilton theorem over the quaternion field, *Chinese Sci. Bull.* 17:1291–1293 (1991).
13 L. X. Chen, Inverse matrix and properties of double determinant over quaternion field, *Sci. China Ser. A* 34(5):528–540 (1991).
14 P. M. Cohn, *Skew Field Constructions*, London Math. Soc. Lecture Note Ser. 27, Cambridge U.P., 1977.
15 P. M. Cohn, The similarity reduction of matrices over a skew field, *Math. Z.* 132:151–163 (1973).

16 S. Eilenberg and I. Niven, The "fundamental theorem of algebra" for quaternions, *Bull. Amer. Math. Soc.* 50:246–248 (1944).

17 S. Eilenberg and N. Steenrod, *Foundations of Algebraic Topology*, Princeton U.P., 1952.

18 D. Finkelstein, J. M. Jauch, S. Schiminovich, and D. Speiser, Foundations of quaternion quantum mechanics, *J. Math. Phys.*, 1962, pp. 3207–3220.

19 D. Finkelstein, J. M. Jauch, and D. Speiser, Notes on quaternion quantum mechanics, *CERN* 59(9):59–17 (1959).

20 K. Guerlebeck and W. Sproessig, *Quaternionic Analysis*, Akademie Verlag, Berlin, 1989.

21 P. R. Halmos, *A Hilbert Space Problem Book*, Van Nostrand, 1967.

22 R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge U.P., 1985.

23 R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge U.P., 1994.

24 N. Jacobson, *Basic Algebra I*, W. H. Freeman, 1974.

25 J. E. Jamison, Numerical Range and Numerical Radius in Quaternionic Hilbert Spaces, Doctoral Dissertation, Univ. of Missouri, 1972.

26 R. E. Johnson, On the equation $\chi\alpha = \gamma\chi + \beta$ over an algebraic division ring, *Bull. Amer. Math. Soc.* 50:202–207 (1944).

27 R. Von Kippenhahn, Über der Wertevorrat einer Matrix, *Math. Nachr.* 6(3/4):193–228 (1951).

28 A. Krieg, *Modular Forms on Half-Spaces of Quaternions*, Lecture Notes in Math. 1143, Springer-Verlag, 1985.

29 H. C. Lee, Eigenvalues of canonical forms of matrices with quaternion coefficients, *Proc. R.I.A.* 52, Sect. A (1949).

30 M. L. Mehta, Determinants of quaternion matrices, *J. Math. Phys. Sci.* 8, No. 6 (1974).

31 N. Mackey, Hamilton and Jacobi meet again: Quaternions and the eigenvalue problem, *SIAM J. Matrix Anal. Appl.*, to appear.

32 I. Niven, Equations in quaternions, *Amer. Math. Monthly* 48:654–661 (1995).

33 N. Steenrod, *The Topology of Fibre Bundles*, Princeton, U.P., 1951.

34 W. So, Left eigenvalues of quaternionic matrices, a Private communication, 1994.

35 W. So and R. C. Thompson, Convexity of the upper complex plane part of the numerical range of a quaternionic matrix, Manuscript, 1995.

36 W. So, R. C. Thompson, and F. Zhang, Numerical ranges of matrices with quaternion entries, *Linear and Multilinear Algebra* 37:175–195 (1994).

37 K. Viswanath, Normal operators on quaternionic Hilbert spaces, *Trans. Amer. Math. Soc.* 162 (1971).

38 K. Viswanath, Contributions to Linear Quaternionic Hilbert Analysis, Thesis, Indian Statistical Inst., Calcutta, 1968.

39 L. A. Wolf, Similarity of matrices in which the elements are real quaternions, *Bull. Amer. Math. Soc.* 42:737–743 (1936).

40 N. A. Wiegmann, Some theorems on matrices with real quaternion elements, *Canad. J. Math.* 7:191–201 (1955).

41 R. M. W. Wood, Quaternionic eigenvalues, *Bull. London Math. Soc.* 17:137–138 (1985).

42 B. J. Xie, An extension of Hadamard's theorem over the skew field of quaternions, *Sci. Sinica*, Special Issue I on Math., pp. 88–93 (1970); also see MR 83g:15014.

43 B. J. Xie, Applications of characteristic roots and standard forms of matrices over a skew field, *Acta Math. Sinica* 23(4):522–533 (1980); also see MR 82m:15021.

44 B. J. Xie, Existence theorems for the canonical diagonal form and the weak canonical form of characteristic matrices over a skew field, *Acta Math. Sinica* 23(3):398–410 (1980); also see MR 82k:15015.

45 B. J. Xie, An expansion theorem for determinants of self-adjoint quaternion matrices and its applications, *Acta Math. Sinica* 23(5):668–683 (1980); also see MR 83a:15007.

46 B. J. Xie, Several theorems concerning self-conjugate matrices and determinants of quaternions, *Acta Sci. Natur. Univ. Jilin*, No. 1, pp. 1–7 (1982); also see MR 84c:15022.

47 F. Zhang, Permanent Inequalities and Quaternion Matrices, Ph.D. Dissertation, Univ. of California at Santa Barbara, 1993.

48 F. Zhang, On Numerical Range of Normal Matrices of Quaternions, *J. Math. Physical Sciences*, 29(6):235–251 (1995).

Received 29 August 1994; final manuscript accepted 13 June 1995