QUATERNIONIC MATRICES: UNITARY SIMILARITY, SIMULTANEOUS TRIANGULARIZATION AND SOME TRACE IDENTITIES

DRAGOMIR Ž. ĐOKOVIĆ AND BENJAMIN H. SMITH

Abstract. We construct six unitary trace invariants for 2×2 quaternionic matrices which separate the unitary similarity classes of such matrices, and show that this set is minimal. We have discovered a curious trace identity for two unit-speed one-parameter subgroups of $\mathrm{Sp}(1)$. A modification gives an infinite family of trace identities for quaternions as well as for 2×2 complex matrices. We were not able to locate these identities in the literature.

We prove two quaternionic versions of a well known characterization of triangularizable subalgebras of matrix algebras over an algebraically closed field. Finally we consider the problem of describing the semi-algebraic set of pairs (X, Y) of quaternionic $n \times n$ matrices which are simultaneously triangularizable. Even the case $n = 2$, which we analyze in more detail, remains unsolved.

2000 Mathematics Subject Classification 15A33,15A18,16R30

1. Introduction

We denote by $\mathbb{R}, \mathbb{C}, \mathbb{H}$ the field of real numbers, complex numbers and the division ring of real quaternions, respectively. Throughout we use D to represent an element from the set $\{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$ and $M_n(D)$ denotes the algebra of $n \times n$ matrices over D. Also, we let $\mathcal{U}_n(D)$ resp. $\mathcal{L}_n(D)$ denote the space of upper resp. lower triangular matrices in $M_n(D)$. In the case where $D = \mathbb{H}$ we will omit parentheses and write $M_n, \mathcal{U}_n, \mathcal{L}_n$.

The maximal compact subgroup of the general linear group $\mathrm{GL}_n(D)$, $\mathrm{O}(n)$ in the real case, $\mathrm{U}(n)$ in the complex case and $\mathrm{Sp}(n)$ in the quaternionic case, acts on $M_n(D)$ by conjugation, i.e., $(X, A) \mapsto XAX^{-1}$, $A \in M_n(D)$. It is an old problem to determine an explicit minimal set of generators for the algebra of polynomial invariants for this action. Explicit results are known (in all three cases) for small values of n. Let us

Key words and phrases. Quaternionic matrices, unitary similarity, simultaneous triangularization, trace identities, semi-algebraic sets.

The first author was supported by the NSERC Discovery Grant A-5285, and the second by an NSERC Undergraduate Student Research Award.

mention that the algebra of real polynomial $\mathrm{Sp}(n)$-invariants is generated by the trace functions $\mathrm{Tr}\ (w(X,X^*))$ where $\mathrm{Tr}\ : M_n \to \mathbb{R}$ is the quaternionic trace (see next section for definition) and w is any word in two letters.

The first question that we consider is that of separating the orbits of the above action by means of a minimal set of polynomial invariants. The real and complex cases have been studied extensively. For instance, Pearcy shows in [12] that $A, B \in M_2(\mathbb{C})$ are unitarily similar if and only if $\mathrm{tr}\ (X)$, $\mathrm{tr}\ (X^2)$ and $\mathrm{tr}\ (XX^*)$ take the same values on A and B. He also gives a list of nine words in X and X^* whose traces distinguish the unitary orbits in $M_3(\mathbb{C})$. This is reduced to a minimal set of seven words by Sibirskiĭ [16]. As far as we know, there are no such results in the quaternionic case except for the case M_1, which is trivial.

In section 3 we extend the first of Pearcy's results to M_2, i.e., the 2×2 quaternionic matrices. In this case we show that six words suffice (see Theorem 3.5), and that our set of six words is minimal

In section 4 we consider two unit-speed one-parameter subgroups of $\mathrm{Sp}(1)$, say

$$\phi_p(s) = e^{ps}, \quad \phi_q(t) = e^{qt}$$

where p and q are pure quaternions of norm 1. We prove (Theorem 4.1) that the real part of

$$\prod_{i=1}^{k} \phi_p(s_i)\phi_q(t_i); \quad s_i, t_i \in \mathbb{R}$$

remains the same when p and q are switched.

This fact remains true for arbitrary pure quaternions p and q provided we take $k = 2$ and set $s_1 = t_1$ and $s_2 = t_2$. From here we obtain an infinite family of trace identities for quaternions (see Proposition 4.2 and it's corollaries), which we were not able to find anywhere in literature. For instance we show that

$$\mathrm{Tr}\ (x^m y^m x^n y^n) = \mathrm{Tr}\ (y^m x^m y^n x^n)$$

is valid for all quaternions x, y and nonnegative integers m, n. One can easilly convert these identities into trace identities for $M_2(\mathbb{C})$.

Section 5 is about the simultaneous triangularization of subalgebras of M_n. In [15] Radjavi and Rosenthal give several characterizations of triangularizability of unital subalgebras of finite dimensional linear operators over an algebraically closed field. By changing the equality in part (iv) of [15, Theorem 1.5.4] to an inequality we are able to extend that result to the quaternionic case (see Theorem 5.7). Next, we observe a peculiar polynomial equation which is satisfied on any triangularizable subalgebra $\mathcal{A} \subseteq M_n$, namely that Tr $([X,Y]^3) = 0$, and we investigate its significance. We introduce the concept of quasi-triangularization (generalizing the triangularization) which refers to the possibility of obtaining a simultaneous block upper triangular form with the diagonal blocks restricted to M_1 or $M_2(\mathbb{C})$. Based on our trace identity for complex matrices given in section 4, we find that a unital subalgebra $\mathcal{A} \subseteq M_n$ is quasi-triangularizable if and only if the identity Tr $([X,Y]^3) = 0$ is valid on $\mathcal{A}$.

In section 6 we explore the semi-algebraic set $\mathcal{W}_n$ of pairs of quaternionic matrices which are simultaneously triangularizable. Hence $\mathcal{W}_n$ is generated from $\mathcal{U}_n \times \mathcal{U}_n$ via the simultaneous conjugation action of the group $\mathrm{GL}_n(\mathbb{H})$. One can replace here $\mathrm{GL}_n(\mathbb{H})$ by $\mathrm{Sp}(n)$ and deduce that $\mathcal{W}_n$ is closed (in the Euclidean topology). For generic $A \in M_n$, i.e., one with n distinct eigenvalues, we find that the fibre of the first projection $\mathcal{W}_n \to M_n$ is the union of $n!$ real vector spaces, each of dimension $2n(n+1)$, with a pairwise intersection a common subspace of dimension $\geq 4n$. We deduce that the dimension of $\mathcal{W}_n$ is $2n(3n+1)$. We also construct two infinite families of polynomial equations (and some inequalities) which are satisfied on $\mathcal{W}_n$.

The problem of pairwise triangularizability in M_2 is of special interest as the first nontrivial case of the above mentioned general problem. Here, the set $\mathcal{W}_2$ can be defined geometrically as the set of quaternionic matrix pairs which share a common eigenvector. In [6] Friedland describes exactly when this occurs in the complex case, see Theorem 7.1 below. In section 7, we look at his result and attempt to extend it to the quaternionic case. We now give some details.

Let $\mathcal{P}_2$ be the algebra of real polynomial functions on $M_2 \times M_2$, and $\mathcal{P}_2'$ resp. $\mathcal{P}_2''$ the subalgebra of $\mathrm{GL}_2(\mathbb{H})$ resp. $\mathrm{Sp}(2)$-invariants. Let $\mathcal{I}_2 \subseteq \mathcal{P}_2$ be the ideal of functions that vanish on $\mathcal{W}_2$, and set $\mathcal{I}_2' = \mathcal{I}_2 \cap \mathcal{P}_2'$ and $\mathcal{I}_2'' = \mathcal{I}_2 \cap \mathcal{P}_2''$. The Zariski closure $\overline{\mathcal{W}}_2$ of $\mathcal{W}_2$ is the set of common zeros of $\mathcal{I}_2$, and the same is true for the ideal $\mathcal{I}_2''$ of $\mathcal{P}_2''$. While the codimension of $\mathcal{W}_2$ in $M_2 \times M_2$ is only 4, we can show that a minimal set of generators of $\mathcal{I}_2'$ has cardinal ≥ 92. Let $\mathcal{J}_m' \subseteq \mathcal{I}_2' \subseteq \mathcal{P}_2'$ be the ideal of $\mathcal{P}_2'$ generated by all polynomials $f \in \mathcal{I}_2'$ of (total) degree $\leq m$. We have constructed a minimal set of generators of $\mathcal{J}_m'$ for $m \leq 14$ (see Table 2 for the generators of $\mathcal{J}_9'$).

In the last section, 8, we state four open problems.

We thank the referee and R. Guralnick for their valuable comments and suggestions.

2. Preliminaries

Let $\mathbb{H} = \{a + \mathsf{i}b + \mathsf{j}c + \mathsf{k}d : a, b, c, d \in \mathbb{R}\}$ represent the skew-field of real quaternions. We shall identify the field $\mathbb{C}$ with the subfield $\{a + \mathsf{i}b : a, b \in \mathbb{R}\}$ of $\mathbb{H}$. For a quaternion $q = a + \mathsf{i}b + \mathsf{j}c + \mathsf{k}d$ we define the *norm*, *real part*, *pure part* and *conjugate* of q in the usual fashion as:

$$\begin{aligned} &|q|^2 := a^2 + b^2 + c^2 + d^2, \\ &\Re(q) := a, \\ &\mathsf{i}b + \mathsf{j}c + \mathsf{k}d \text{ and} \\ &\overline{q} := a - \mathsf{i}b - \mathsf{j}c - \mathsf{k}d, \end{aligned}$$

respectively. The *adjoint* of a matrix $A \in M_n$ is given by $A^* = \overline{A}^T$, which is also known as the conjugate-transpose of A. With this, we can define the (compact) *symplectic group*, $\mathrm{Sp}(n)$, as the collection of quaternionic unitary matrices, namely

$$\mathrm{Sp}(n) := \{U \in M_n : U^*U = I_n\},$$

where I_n is the identity matrix see e.g. [2]. Consider the equivalence relation $\sim$ defined on M_n by the conjugation action of $\mathrm{Sp}(n)$. To be precise, we have:

$$A \sim B \leftrightarrow \exists U \in \mathrm{Sp}(n), \quad A = UBU^{-1}.$$

This shall be referred to as $\mathrm{Sp}(n)$-*equivalence*. We will speak of individual $\mathrm{Sp}(n)$-equivalence class for a given matrix $A \in M_n$ and thus, denote this orbit by $\mathcal{O}_A = \{UAU^{-1} : U \in \mathrm{Sp}(n)\}$. It is well known that M_n can be embedded nicely into $M_{2n}(\mathbb{C})$. For this purpose, given $A \in M_n$ we write $A = A_1 + \mathsf{j}A_2$ with $A_1, A_2 \in M_n(\mathbb{C})$. Then

$$\chi_n : M_n \to M_{2n}(\mathbb{C}), \quad \chi_n(A) = \begin{bmatrix} A_1 & -\overline{A}_2 \\ A_2 & \overline{A}_1 \end{bmatrix},$$

is an injective homomorphism of $\mathbb{R}$-algebras. From this, the quaternionic matrices inherit various analogous properties regarding invertibility, triangularizability, canonical forms, decomposition, determinants, numerical range and more. See [18] for detailed results on quaternionic linear algebra.

Given $A \in M_n$, there exists $P \in \mathrm{GL}_n(\mathbb{H})$ such that $PAP^{-1} \in \mathcal{U}_n$ with successive diagonal entries $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ and imaginary parts $\Im(\lambda_i) \geq 0$. The sequence $(\lambda_1, \ldots, \lambda_n)$ is unique up to permutation and we refer to $\lambda_1, \ldots, \lambda_n$ as the *eigenvalues* of A. We shall use the word "eigenvalues", in the context of quaternionic matrices, exclusively in this sense. Note that the eigenvalues of the complex matrix $\chi_n(A)$ are $\lambda_1, \ldots, \lambda_n$ and $\bar{\lambda}_1, \ldots, \bar{\lambda}_n$ (counting multiplicities).

We seek to classify exactly when two 2×2 quaternionic matrices are $\mathrm{Sp}(2)$-equivalent using polynomial functions which remain constant on the equivalence classes. It is known that the algebra of polynomial invariants for complex matrices under the action of conjugation by unitary group is generated by a finite number of particular *known* trace functions on $M_n(\mathbb{C})$ for $n = 2, 3, 4$. Since M_2 embeds into $M_4(\mathbb{C})$ as seen by χ_2, it is only natural to assume a classification of this type can be extended in some way to the quaternionic case. That is, we should be capable of defining explicitly which functions separate orbits.

First, we will need to introduce the notion of trace for quaternions and matrices of such which will be preserved by χ_n above. For the general definition of the reduced trace and the reduced norm for central simple algebras we refer the reader to [13, 7].

Definition 2.1. The *quaternionic trace* of $A = [a_{ij}] \in M_n$ is,

$$\mathrm{Tr}\ (A) := \sum_{i=1}^{n} (a_{ii} + \bar{a}_{ii}) = 2\Re \left(\sum_{i=1}^{n} a_{ii} \right).$$

In particular, when $n = 1$ we have

$$\mathrm{Tr}\ (q) = q + \bar{q} = 2\Re(q)$$

It is important to notice the clear distinction between the usual trace, tr , on a matrix algebra over a field and the quaternionic analog presented above. For instance, we have $\mathrm{tr}\ I_n = n$ while $\mathrm{Tr}\ I_n = 2n$. Some properties which follow directly from our definition as well as the simple fact that $\Re(pq) = \Re(qp)$ for all quaternions p, q are as follows:

Let $A, B \in M_n$, $P \in \mathrm{GL}_n(\mathbb{H})$, $U \in \mathrm{Sp}(n)$ and w be any word on two letters, then

(1) $\mathrm{Tr}\ (A) = \mathrm{tr}\ \chi_n(A)$
(2) $\mathrm{Tr}\ (A + B) = \mathrm{Tr}\ (A) + \mathrm{Tr}\ (B)$
(3) $\mathrm{Tr}\ (\lambda A) = \lambda \mathrm{Tr}\ (A), \quad \lambda \in \mathbb{R}$
(4) $\mathrm{Tr}\ (AB) = \mathrm{Tr}\ (BA)$
(5) $\mathrm{Tr}\ (w(PAP^{-1}, PBP^{-1})) = \mathrm{Tr}\ (w(A, B))$
(6) $\mathrm{Tr}\ (w(UAU^*, UA^*U^*)) = \mathrm{Tr}\ (w(A, A^*))$

Observe, properties (2),(3) and (6) tell us that the quaternionic trace of any $\mathbb{R}$-linear combination of any words on $\{A, A^*\}$ is invariant under the action of $\mathrm{Sp}(n)$.

3. Sp(2)-equivalence of matrices

For our purposes it is essential to introduce the following six trace polynomials on M_2

$$(p_1, p_2, p_3, p_4, p_5, p_6) := \mathrm{Tr}\,(A, A^2, A^3, A^4, AA^*, A^2A^{*2}).$$

We proceed in showing that these form a minimal set of Sp(2)-invariants that separate orbits in M_2.

Let us first describe a simple canonical form for 2×2 quaternionic matrices under Sp(2)-equivalence.

Definition 3.1. Denote by $\mathcal{K}$ the set of all matrices of the form

$$A = \begin{bmatrix} \alpha & z \\ 0 & \beta \end{bmatrix}$$

such that $\alpha, \beta \in \mathbb{C}$; $\Im(\alpha), \Im(\beta) \geqslant 0$; $z = z_1 + \mathsf{j}z_3$; $z_1, z_3 \geqslant 0$; and if α, or β is real then $z_3 = 0$.

Notice that $\mathcal{K}$ is a semi-algebraic set of dimension 6. Eventually, we will see $\mathcal{K}$ intersects each Sp(2)-equivalence class at either one or two points. The following result shows that $\mathcal{K}$ meets every orbit $\mathcal{O}_A$ at least once.

Lemma 3.2. *Every* 2×2 *quaternionic matrix is* Sp(2)*-equivalent to some matrix* $A \in \mathcal{K}$.

Proof. First, by the generalization of Schur's theorem for quaternionic matrices found in [18] any matrix in M_2 is Sp(2)-equivalent to a matrix of the form

$$\begin{bmatrix} \alpha & z \\ 0 & \beta \end{bmatrix},$$

with $\alpha, \beta \in \mathbb{C}$; and $\Im(\alpha), \Im(\beta) \geqslant 0$. If we write $z = c_1 + \mathsf{j}c_2$; $c_1, c_2 \in \mathbb{C}$, we can reduce our matrix as follows,

$$\begin{bmatrix} u & 0 \\ 0 & v \end{bmatrix} \begin{bmatrix} \alpha & c_1 + \mathsf{j}c_2 \\ 0 & \beta \end{bmatrix} \begin{bmatrix} u^{-1} & 0 \\ 0 & v^{-1} \end{bmatrix} = \begin{bmatrix} \alpha & |c_1| + \mathsf{j}|c_2| \\ 0 & \beta \end{bmatrix}.$$

It suffices to choose unit complex numbers u and v such that $uc_1v^{-1} = |c_1|$ and $\bar{u}c_2v^{-1} = |c_2|$, which is always possible. If, say, β is real then we can set $u = 1$ and choose a unit quaternion v such that zv^{-1} is real nonnegative. □

Now we provide a technical result which simplifies the form of p_6 when restricted to $\mathcal{K}$. This will be useful for later computations.

Lemma 3.3. *Let* $A = \begin{bmatrix} \alpha & z \\ 0 & \beta \end{bmatrix}$ *be as in Definition 3.1. In particular* $z = z_1 + \mathsf{j} z_3$ *with* $z_1, z_3 \geqslant 0$. *Then,*

$$\frac{1}{2} p_6(A) = |\alpha|^4 + |\beta|^4 + |\alpha + \bar{\beta}|^2 |z|^2 + z_1^2 (\alpha - \bar{\alpha})(\bar{\beta} - \beta).$$

Proof. We have

$$p_6(A) = \mathrm{Tr}\,(A^2 A^{*2}) = 2\left(|\alpha|^4 + |\beta|^4 + (|\alpha|^2 + |\beta|^2)|z|^2 + 2\Re(\alpha z \bar{\beta} \bar{z})\right)$$

and

$$z \bar{\beta} \bar{z} = (\bar{\beta} z_1^2 + \beta z_3^2) + \mathsf{j} z_1 z_3 (\beta + \bar{\beta}).$$

Notice that the second term above is a pure quaternion even upon multiplication by α and so has zero real part. So

$$\begin{aligned} 2\Re(\alpha z \bar{\beta} \bar{z}) &= z_1^2 (\alpha \bar{\beta} + \bar{\alpha} \beta) + z_3^2 (\alpha \beta + \bar{\alpha} \bar{\beta}) \\ &= |z|^2 (\alpha \beta + \bar{\alpha} \bar{\beta}) + z_1^2 (\alpha \bar{\beta} + \bar{\alpha} \beta - \alpha \beta + \bar{\alpha} \bar{\beta}) \\ &= |z|^2 (\alpha \beta + \bar{\alpha} \bar{\beta}) + z_1^2 (\alpha - \bar{\alpha})(\bar{\beta} - \beta). \end{aligned}$$

□

With this formulation in place, we can classify Sp(2)-equivalence between matrices which lie in $\mathcal{K}$. In fact we shall prove that an orbit $\mathcal{O}_A$ meets $\mathcal{K}$ in two points when A has distinct eigenvalues and in a single point otherwise.

Theorem 3.4. *If* $A = \begin{bmatrix} \alpha & z \\ 0 & \beta \end{bmatrix}$ *and* $B = \begin{bmatrix} \gamma & w \\ 0 & \delta \end{bmatrix}$ *belong to* $\mathcal{K}$, *then*

$$A \sim B \iff z = w \ \&\ \{\alpha, \beta\} = \{\gamma, \delta\}.$$

Proof. When $A \sim B$ we get from $p_k(A) = p_k(B)$, $k \in \{1, 2, 3, 4\}$, that the sets of eigenvalues, $\{\alpha, \beta, \bar{\alpha}, \bar{\beta}\}$ and $\{\gamma, \delta, \bar{\gamma}, \bar{\delta}\}$, for $\chi_2(A), \chi_2(B)$ are the same. So we get that $\{\alpha, \beta\} = \{\gamma, \delta\}$ as well. Also, from $p_5(A) = p_5(B)$ which is $|\alpha|^2 + |\beta|^2 + |z|^2 = |\gamma|^2 + |\delta|^2 + |w|^2$, we can see that $|z|^2 = |w|^2$. Recall that $z = z_1 + \mathsf{j} z_3$ and $w = w_1 + \mathsf{j} w_3$, where z_1, z_3, w_1, w_3 are real and nonnegative. Now, $p_6(A) = p_6(B)$ and Lemma 3.3 above show that $z_1^2 = w_1^2$, and since both z_1 and w_1 are nonnegative we have $z_1 = w_1$. It follows that also $z = w$.

To prove the converse, we may assume that $A \neq B$. Thus $\delta = \alpha, \gamma = \beta$ and $\alpha \neq \beta$. So, we know A is Sp(2)-equivalent to a matrix $A' = \begin{bmatrix} \beta & w' \\ 0 & \alpha \end{bmatrix} \in \mathcal{K}$, since Schur's theorem allows us to place the eigenvalues in any order along the diagonal. Hence, by the first part of the proof and the hypothesis we have $w' = z = w$ and so $A \sim A' = B$. □

With Lemma 3.2 along with Theorem 3.4, we have reached the promised canonical form for 2×2 quaternionic matrices under Sp(2)-equivalence. It is unique up to permutation of the diagonal entries.

Now we may begin looking to find polynomial invariants which separate the Sp(2)-equivalence classes. Also, it is ideal for computational purposes to obtain the least number of these polynomials which do the job.

Theorem 3.5. *Two matrices $A, B \in M_2$ are* Sp(2)*-equivalent if and only if the following six equations hold:*

$$\begin{aligned} &\text{Tr }\left(A^i\right) = \text{Tr }\left(B^i\right), \quad i \in \{1,2,3,4\}; \\ &\text{Tr }\left(AA^*\right) = \text{Tr }\left(BB^*\right); \\ &\text{Tr }\left(A^2 A^{*^2}\right) = \text{Tr }\left(B^2 B^{*^2}\right), \end{aligned}$$

i.e., $p_k(A) = p_k(B)$ for $1 \leq k \leq 6$. Moreover, this is a minimal set of invariants with the mentioned property.

Proof. First, if $A \sim B$, our trace equations are trivially satisfied as $\text{Tr } w(A, A^*) = \text{Tr } w(B, B^*)$ for all words w on two letters.

Conversely, suppose the given set of traces match for A and B. By Lemma 3.2 we may assume that

$$A = \begin{bmatrix} \alpha & z \\ 0 & \beta \end{bmatrix}, \; B = \begin{bmatrix} \gamma & w \\ 0 & \delta \end{bmatrix} \in \mathcal{K}.$$

We know the first four invariants of A, B uniquely determine the sets $\{\alpha, \beta\}$, $\{\gamma, \delta\}$ respectively and since these invariants are the same, we have that these sets are equal. It remains to show that $z = w$ which can be done using p_5 and p_6. We have from $p_5(A) = p_5(B)$ that

$$|\alpha|^2 + |\beta|^2 + |z|^2 = |\gamma|^2 + |\delta|^2 + |w|^2.$$

As $\{\alpha, \beta\} = \{\gamma, \delta\}$, we have $|z| = |w|$. Similarly, from $p_6(A) = p_6(B)$ and Lemma 3.3 we get

$$\begin{aligned} &|\alpha|^4 + |\beta|^4 + |\alpha + \bar{\beta}|^2 |z|^2 + z_1^2 (\alpha - \bar{\alpha})(\bar{\beta} - \beta) \\ &\quad = |\gamma|^4 + |\delta|^4 + |\gamma + \bar{\delta}|^2 |z|^2 + w_1^2 (\gamma - \bar{\gamma})(\bar{\delta} - \delta), \end{aligned}$$

and so $z_1^2 = w_1^2$. From our description of $\mathcal{K}$ we know $z_1, w_1 \geq 0$. This implies that $z_1 = w_1$ and thus, $z = w$. Now, it follows from Theorem 3.4 that $A \sim B$.

Finally, to prove minimality, we provide pairs of matrices, each of which agree on all but one invariant from our list. In Table 1 the matrices in the k-th row have distinct values of p_k only.

Table 1: Examples for minimality

1.	$\begin{bmatrix} \sqrt{3}-\mathsf{i} & 0 \\ 0 & -\sqrt{3}+\mathsf{i} \end{bmatrix}$	$\begin{bmatrix} -\sqrt{3}+\mathsf{i} & 0 \\ 0 & -\sqrt{3}-\mathsf{i} \end{bmatrix}$
2.	$\begin{bmatrix} 2+\mathsf{i} & 1 \\ 0 & -2-\mathsf{i} \end{bmatrix}$	$\begin{bmatrix} 1+2\mathsf{i} & -1 \\ 0 & -1-2\mathsf{i} \end{bmatrix}$
3.	$\begin{bmatrix} -1+2\mathsf{i} & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 \\ 0 & 1+2\mathsf{i} \end{bmatrix}$
4.	$\begin{bmatrix} 0 & \sqrt{6}+\mathsf{j} \\ 0 & \sqrt{2}\mathsf{i} \end{bmatrix}$	$\begin{bmatrix} \mathsf{i} & \sqrt{3}+2\mathsf{j} \\ 0 & -\mathsf{i} \end{bmatrix}$
5.	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$
6.	$\begin{bmatrix} \mathsf{i} & 1 \\ 0 & \mathsf{i} \end{bmatrix}$	$\begin{bmatrix} \mathsf{i} & \mathsf{j} \\ 0 & \mathsf{i} \end{bmatrix}$

Thus, we have shown that our set of invariants is minimal. □

Let us point out that the complex analog of this theorem is not valid. The reason is that, in the complex case, the polynomial invariants do not separate the orbits.

4. Some trace identities for quaternions

It is well known that every one parameter subgroup of Sp(1) has the form

$$\phi_p(s) = e^{sp} := \sum_{i\geq 0} \frac{1}{i!} s^i p^i, \quad s \in \mathbb{R},$$

for a unique pure quaternion p. Note that $\phi_{\lambda p}(s) = \phi_p(\lambda s)$ for all real λ and

$$u\phi_p(s)u^{-1} = \phi_{upu^{-1}}(s) \tag{4.1}$$

for any nonzero quaternion u.

Theorem 4.1. *If p and q are pure unit quaternions, then*

$$\operatorname{Tr}\left(\prod_{i=1}^{k} \phi_p(s_i)\phi_q(t_i)\right) = \operatorname{Tr}\left(\prod_{i=1}^{k} \phi_q(s_i)\phi_p(t_i)\right). \tag{4.2}$$

is valid for any real numbers $s_1, \ldots, s_k$ and $t_1, \ldots, t_k$.

Proof. Since $|p| = |q| = 1$ there exists a 180°-rotation, in the 3-dimensional space of pure quaternions, which interchanges p and q. Consequently, there exists a pure unit quaternion u such that $upu^{-1} = q$ and $uqu^{-1} = p$ (if $p + q \neq 0$ one can choose u in the direction of $p + q$, and otherwise just to be orthogonal to p.) It suffices now to observe that conjugation by u interchanges the two products in (4.2). □

The identity (4.2) is not valid for arbitrary pure quaternions p and q. Let us look at the special cases where

$$s_1 = t_1, s_2 = t_2, \ldots, s_k = t_k. \tag{4.3}$$

Proposition 4.2. *Let p and q be arbitrary pure quaternions, then*

(a) *For every $s, t \in \mathbb{R}$, we have*

$$\mathrm{Tr}\ (\phi_p(s)\phi_q(s)\phi_p(t)\phi_q(t)) = \mathrm{Tr}\ (\phi_q(s)\phi_p(s)\phi_q(t)\phi_p(t));$$

(b) *When at least two of $r, s, t \in \mathbb{R}$ are equal, we have*

$$\begin{aligned} &\mathrm{Tr}\ (\phi_p(r)\phi_q(r)\phi_p(s)\phi_q(s)\phi_p(t)\phi_q(t)) \\ &\quad = \mathrm{Tr}\ (\phi_q(r)\phi_p(r)\phi_q(s)\phi_p(s)\phi_q(t)\phi_p(t)). \end{aligned}$$

Proof. Let us prove (a). We can write $p = \lambda p_0$ and $q = \mu q_0$, where p_0 and q_0 are pure unit quaternions and $\lambda, \mu \geq 0$. From (4.1) we may assume that $p_0 = \mathsf{i}$ and $q_0 = \mathsf{i}\cos\rho + \mathsf{j}\sin\rho$. Then we have

$$\phi_p(s) = \phi_{\lambda p_0}(s) = \phi_{p_0}(\lambda s) = \cos\lambda s + \mathsf{i}\sin\lambda s,$$

$$\phi_q(s) = \cos\mu s + (\mathsf{i}\cos\rho + \mathsf{j}\sin\rho)\sin\mu s.$$

Next we get that

$$\begin{aligned} \phi_p(s)\phi_q(s) = &\cos\lambda s\cos\mu s - \sin\lambda s\sin\mu s\cos\rho \\ &+ \mathsf{i}(\cos\lambda s\sin\mu s\cos\rho + \sin\lambda s\cos\mu s) \\ &+ \mathsf{j}\cos\lambda s\sin\mu s\sin\rho \\ &+ \mathsf{k}\sin\lambda s\sin\mu s\sin\rho. \end{aligned}$$

From here, we compute the product $\phi_p(s_1)\phi_q(s_1)\phi_p(s_2)\phi_q(s_2)$ and verify (we did it using Maple) that its real part remains the same when p and q are switched. Hence, proving part (a).

The proof of the part (b) is similar and we omit the details. □

A discrete version of the last proposition can be extended to arbitrary quaternions x, y by considering their polar decompositions. Thus, we have the following corollary.

Corollary 4.3. *Let m, n, r be nonnegative integers. Then*

$$\text{Tr }(x^m y^m x^n y^n) = \text{Tr }(y^m x^m y^n x^n)$$

is valid for all $x, y \in \mathbb{H}$. If m, n, r are not distinct, then

$$\text{Tr }(x^m y^m x^n y^n x^r y^r) = \text{Tr }(y^m x^m y^n x^n y^r x^r)$$

is also valid.

As a result of this identity for quaternions we obtain another important corollary about $M_2(\mathbb{C})$. Take note the trace below is the standard trace, as denoted by the lowercase tr.

Corollary 4.4. *Let m, n, r be nonnegative integers. Then*

$$\text{tr }(x^m y^m x^n y^n) = \text{tr }(y^m x^m y^n x^n)$$

is valid for all $x, y \in M_2(\mathbb{C})$. If m, n, r are not distinct , then

$$\text{tr }(x^m y^m x^n y^n x^r y^r) = \text{tr }(y^m x^m y^n x^n y^r x^r)$$

is also valid.

Proof. We have a direct decomposition $M_2(\mathbb{C}) = \chi_1(\mathbb{H}) \oplus \mathrm{i}\chi_1(\mathbb{H})$. Since the left hand side is a complex analytic polynomial (in 8 indeterminates) which we know from Corollary 4.3 vanishes on $\chi_1(\mathbb{H})$, it follows that this polynomial must be identically zero and our identity holds. $\square$

The referee supplied a simple proof of these two corollaries for any algebra with trace. For simplicity we sketch his argument in the setting of Corollary 4.3. By expanding products, one can easily check that the identity

$$\text{Tr }((a+bx)(c+dy)(e+fx)(g+hy)) = \text{Tr }((c+dy)(a+bx)(g+hy)(e+fx))$$

holds for all real scalars a, b, c, d, e, f, g, h and quaternions x, y. This implies the first identity since $x^m = u + vx$ etc for some real scalars u, v. The second identity can be deduced from the first by using the fact that the elements of $\mathbb{H}$ are quadratic over $\mathbb{R}$.

For convenience we shall use the standard notation for the commutator. That is $[A, B] := AB - BA$. Notice that

$$\text{Tr }\left([A,B]^3\right) = 3\text{Tr }(A^2B^2AB - B^2A^2BA) = -3\text{Tr }\left(AB^2A[A,B]\right), \tag{4.4}$$

as this will be useful for further results.

5. Triangularizable subalgebras of M_n

A set $\mathcal{S} \subseteq M_n(D)$ is *triangularizable* if there is a $P \in \mathrm{GL}_n(D)$ such that $P\mathcal{S}P^{-1} \subseteq \mathcal{U}_n(D)$. In the recent book [15] one can find several characterizations of triangularizable complex subalgebras of $M_n(\mathbb{C})$. Some of these results can be easily transferred to the quaternionic case while others are no longer valid. In particular, it is trivial to see that the next proposition does not hold for the quaternionic case. First, we introduce the notion of permutable functions.

Definition 5.1. Let $\mathcal{S} \subseteq M_n(D)$ be a collection of matrices. For a function f on $M_n(D)$, taking values in the center of D, we say f is *permutable on* $\mathcal{S}$ if

$$f(A_1A_2 \dots A_k) = f(A_{\sigma(1)}A_{\sigma(2)} \dots A_{\sigma(k)})$$

for all $A_1, A_2, \dots, A_k \in \mathcal{S}$ and all permutations σ of $\{1, 2, \dots, k\}$.

One can find in [14] the following characterization of triangularizable subalgebras of $M_n(\mathbb{C})$. See [8] for generalization to other fields.

Proposition 5.2. *Let $\mathcal{A} \subseteq M_n(\mathbb{C})$ be a unital complex subalgebra. Then $\mathcal{A}$ is triangularizable if and only if trace is permutable on $\mathcal{A}$.*

The analogous assertion for (real) subalgebras of M_n is invalid because Tr is not permutable on $\mathcal{U}_n$. This is due to the lack of commutativity of $\mathbb{H}$.

To prove our result giving a characterization of triangularizable subalgebras of M_n we shall use the concept of quaternionic representations of real algebras.

Definition 5.3. Let $\mathcal{A}$ be an associative unital $\mathbb{R}$-algebra. A *quaternionic representation* of $\mathcal{A}$ is a $\mathbb{R}$-algebra homomorphism

$$\rho : \mathcal{A} \to \mathrm{End}_{\mathbb{H}}(\mathcal{V}),$$

where $\mathcal{V}$ is a right quaternionic vector space. We also say that $\mathcal{V}$ is a *quaternionic (left) $\mathcal{A}$-module*. We say that ρ is *irreducible* if $\mathcal{V}$ is non-zero and has no proper non-zero $\mathcal{A}$-invariant quaternionic subspaces.

Let us briefly outline the basic facts regarding quaternionic representations of finite-dimensional unital $\mathbb{R}$-algebras $\mathcal{A}$. If $\mathcal{R}$ is the radical of $\mathcal{A}$, then $\mathcal{A}/\mathcal{R} \cong \mathcal{A}_1 \times \cdots \times \mathcal{A}_s$, where each $\mathcal{A}_i$ is a simple algebra. Thus, each $\mathcal{A}_k$ is isomorphic to one of the algebras $M_r(\mathbb{R})$, $M_r(\mathbb{C})$ or $M_r(\mathbb{H})$, for some integer $r \geq 1$. In each of these three cases, the right quaternionic space $\mathbb{H}^r$ (column vectors) is an irreducible quaternionic $\mathcal{A}_k$-module, and also an irreducible quaternionic $\mathcal{A}$-module. In this way we obtain all irreducible quaternionic $\mathcal{A}$-modules (up to isomorphism).

Moreover, the $\mathcal{A}$-modules arising for different values of k are pairwise non-isomorphic.

Remark 5.4. Let us also mention another useful fact: There exists a subalgebra $\mathcal{B} \subseteq \mathcal{A}$ such that $\mathcal{A} = \mathcal{R} \oplus \mathcal{B}$ see [13, Wedderburn–Malcev Theorem, p.209].

The following proposition plays a crucial role in the sequel.

Proposition 5.5. *Let $\mathcal{A} \subseteq M_n$ be a unital subalgebra and $\rho : \mathcal{A} \to M_r$ an irreducible quaternionic representation. Let $p(x, y)$ be a polynomial, in two non-commuting variables x and y, with real coefficients. If the inequality*

$$\mathrm{Tr}\ (p(x, y)) \leq 0$$

is satisfied for all $x, y \in \mathcal{A}$, then it is also satisfied for all $x, y \in \rho(\mathcal{A})$. The same assertion remains valid if the inequality sign is replaced by equality.

Proof. If $\mathcal{R}$ is the radical of $\mathcal{A}$, then $\mathcal{A}/\mathcal{R} \cong \mathcal{A}_1 \times \cdots \times \mathcal{A}_s$, where each $\mathcal{A}_i$ is a simple algebra. Let W_k be the unique (up to isomorphism) irreducible quaternionic module of $\mathcal{A}_k$. Then $W_1, \ldots, W_s$ are representatives of the isomorphism classes of irreducible quaternionic $\mathcal{A}$-modules and let $\rho_1, \ldots, \rho_s$ be their corresponding representations. Let $0 = V_0 \subset V_1 \subset \cdots \subset V_m = \mathbb{H}^n$ be a Jordan–Hölder series of $\mathbb{H}^n$ (viewed as a quaternionic $\mathcal{A}$-module). Denote by n_k the number of indices $i \in \{1, 2, \ldots, m\}$ such that $V_i/V_{i-1} \cong W_k$. As $\mathbb{H}^n$ is a faithful $\mathcal{A}$-module, we have $n_k \geq 1$ for each k. For any $x, y \in \mathcal{A}$, we have

$$\mathrm{Tr}\ (p(x, y)) = \sum_{i=1}^{s} n_i \mathrm{Tr}\ (p(\rho_i(x), \rho_i(y))).$$

By the above remark, for a fixed $k \in \{1, 2, \ldots, s\}$ and any $x_k, y_k \in \rho_k(\mathcal{A})$ there exist $x, y \in \mathcal{A}$ such that $\rho_k(x) = x_k$, $\rho_k(y) = y_k$, while $\rho_l(x) = \rho_l(y) = 0$ for $l \neq k$. For such x, y we have

$$\mathrm{Tr}\ (p(x, y)) = n_k \mathrm{Tr}\ (p(x_k, y_k)).$$

As $n_k \geq 1$ and $\mathrm{Tr}\ (p(x, y)) \leq 0$, we conclude that $\mathrm{Tr}\ (p(x_k, y_k)) \leq 0$. Since the representation ρ is equivalent to some ρ_k, The first assertion is proved.

The second assertion is a consequence of the first. □

We shall also need the following easy lemma.

Lemma 5.6. *Let $\mathcal{A} = M_r(D)$ where $D \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$.*

(1) *If* $\mathrm{Tr}\ ([A, B]^2) \leq 0$ *holds for all $A, B \in \mathcal{A}$ then $r = 1$.*

(2) Tr $([A,B]^3)=0$ *holds true in* $\mathcal{A}$ *if and only if either* $r=1$ *or* $r=2$ *and* $D\in\{\mathbb{R},\mathbb{C}\}$.

Proof. To prove (1), suppose $r=2$ then the matrix pair

$$A=\begin{bmatrix}1&0\\0&0\end{bmatrix},\quad B=\begin{bmatrix}0&1\\-1&0\end{bmatrix}$$

has Tr $([A,B]^2)=4>0$. If $r\geq 3$, we may extend the matrices A,B with rows and columns of zeros to see that inequality does not hold. Hence, we must have that $r=1$.

Next we prove (2). If $r=1$ or $r=2$ and $D\in\{\mathbb{R},\mathbb{C}\}$ then Corollaries 4.3 and 4.4 give Tr $([A,B]^3)=0$ on $\mathcal{A}$.

To prove the converse, we proceed by contradiction. If $r=2$ and $D=\mathbb{H}$ then

$$A=\begin{bmatrix}0&1\\0&\mathsf{j}\end{bmatrix},\quad B=\begin{bmatrix}0&1\\\mathsf{i}&0\end{bmatrix}$$

satisfy Tr $([A,B]^3)=-4\neq 0$. Similarly, for $r\geq 3$ we see our equality is invalid even for $D=\mathbb{R}$: Observe that

$$A=\begin{bmatrix}0&0&0\\0&1&0\\1&0&0\end{bmatrix},\quad B=\begin{bmatrix}1&1&0\\0&0&1\\0&0&0\end{bmatrix}$$

have Tr $([A,B]^3)=-6\neq 0$. Thus, by our argument in part (1) we are done. □

The following theorem is a quaternionic version of Theorem 1.5.4 from [15]. The only change in the wording appears in part (4) where we have replaced their equality with an inequality.

Theorem 5.7. *For a unital subalgebra* $\mathcal{A}\subseteq M_n$*, the following are equivalent:*

(1) $\mathcal{A}$ *is triangularizable.*
(2) *If* $A,B\in\mathcal{A}$ *are nilpotent then so is* $A+B$.
(3) *If at least one of* $A,B\in\mathcal{A}$ *is nilpotent then so is* AB.
(4) Tr $([A,B]^2)\leq 0$ *for all* $A,B\in\mathcal{A}$.

Proof. First, if we assume (1) holds then it is trivial to see that (2),(3), and (4) are all satisfied.

Conversely, suppose at least one of (2),(3) or (4) holds. As in the proof of Proposition 5.5, choose a Jordan–Hölder series $0=V_1\subset V_2\subset\cdots\subset V_m=\mathbb{H}^n$ for the quaternionic $\mathcal{A}$-module $\mathbb{H}^n$. Let n_k be the quaternionic dimension of the quotient $W_k=V_k/V_{k-1}$. It suffices to show that each $n_k=1$.

There is a $Q \in \mathrm{GL}_n(\mathbb{H})$ such that $Q\mathcal{A}Q^{-1}$ consists of block upper triangular matrices with successive diagonal blocks square of size n_k, $k = 1, \dots, m$. For any $X \in \mathcal{A}$ let $\rho_k(X)$ denote the k-th diagonal block of size n_k of the matrix QXQ^{-1}. Each $\rho_k(\mathcal{A})$ is a simple real algebra isomorphic to $M_{n_k}(D_k)$ with $D_k \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$. By Noether–Skolem Theorem we may assume that Q is chosen so that each $\rho_k(\mathcal{A}) = M_{n_k}(D_k)$.

If (2) or (3) holds then $n_k = 1$ because otherwise the pair of nilpotent matrices $\left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]$ will add and multiply to a matrix which is not nilpotent. If (4) holds, then Lemma 5.6 gives that $n_k = 1$. □

Remark 5.8. The equivalence of (1),(2) and (3) was also mentioned by Kermani at the recent ILAS Conference [10].

Our next objective is to characterize subalgebras of M_n that satisfy the identity $\mathrm{Tr}\,([X, Y]^3) = 0$. For that purpose we need the concept of quasi-triangularizability which we now define

Let us define a $\{1, 2\}$-*sequence* as a sequence $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_m)$ of integers from {1,2}. We say that m is its *length* and $|\sigma| := \sigma_1 + \cdots + \sigma_m$ is its *size*. Assuming that $|\sigma| = n$, we denote by M_σ the subalgebra of M_n consisting of all block triangular matrices

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1,m-1} & A_{1m} \\ 0 & A_{22} & & A_{2,m-1} & A_{2m} \\ \vdots & & & & \\ 0 & 0 & & 0 & A_{mm} \end{bmatrix}$$

with the diagonal blocks $A_{ii} \in M_{\sigma_i}$ subject to the additional condition that $A_{ii} \in M_2(\mathbb{C})$ whenever $\sigma_i = 2$, and all $A_{i,j}, i < j$, arbitrary quaternionic matrices of appropriate sizes.

We can now define the quasi-triangularizable sets of quaternionic matrices.

Definition 5.9. A collection $\mathcal{S}$ of $n \times n$ quaternionic matrices is *quasi-triangularizable* (denoted by q.t.) if $P\mathcal{S}P^{-1} \subseteq M_\sigma$ for some $P \in \mathrm{GL}_n(\mathbb{H})$ and some $\{1, 2\}$-sequence σ of size n. (If all σ_i can be taken to be 1, then $\mathcal{S}$ is triangularizable.)

Theorem 5.10. *A unital subalgebra $\mathcal{A} \subseteq M_n$ is q.t. if and only if* $\mathrm{Tr}\,([A, B]^3) = 0$ *for all* $A, B \in \mathcal{A}$.

Proof. If $\mathcal{A} \subseteq M_n$ is q.t., then for $A, B \in \mathcal{A}$ there is a $P \in \mathrm{GL}_n(\mathbb{H})$ and a {1,2}-sequence σ with $|\sigma| = n$ such that $P\mathcal{A}P^{-1} \subseteq M_\sigma$. For $A, B \in \mathcal{A}$ with diagonal blocks, denoted $A_1, \dots, A_k$ and $B_1, \dots, B_k$ for PAP^{-1}, PBP^{-1} respectively, we see that all satisfy the trace identity proven in Corollary 4.4. In particular the identity holds for integers $(m,n)=(2,1)$. By Lemma 5.6 and the identity (4.4), we have

$$\mathrm{Tr}\,([A,B]^3)=\sum_{i=1}^{k}\mathrm{Tr}\,([A_i,B_i]^3)=0.$$

Conversely, suppose that any $A,B\in\mathcal{A}$ satisfy the given identity. If $\mathcal{R}$ is the radical of $\mathcal{A}$ then we know that $\mathcal{A}/\mathcal{R}=\mathcal{A}_1\times\cdots\times\mathcal{A}_s$ where the $\mathcal{A}_i$'s are simple $\mathbb{R}$-algebras. That is, for each $i\in\{1,\ldots,s\}$ we have that $\mathcal{A}_i\cong M_r(D_i)$, for some positive integer r and $D_i\in\{\mathbb{R},\mathbb{C},\mathbb{H}\}$. Proposition 5.5 guarantees that our trace identity remains true on each $\mathcal{A}_i$. Thus, Lemma 5.6 implies that the possibilities for r_i and D_i can be reduced to exactly one of $r_i=1$ and $D_i\in\{\mathbb{R},\mathbb{C},\mathbb{H}\}$ or $r_i=2$ and $D_i\in\{\mathbb{R},\mathbb{C}\}$.

Next, as in the proof of Proposition 5.5, we choose a Jordan–Hölder series $0=V_0\subset V_1\subset\ldots\subset V_m=\mathbb{H}^n$ for the quaternionic $\mathcal{A}$-module $\mathbb{H}^n$. From the fact just proved above it follows that each irreducible quotient V_i/V_{i-1} has quaternionic dimension $\sigma_i=1$ or 2. Consequently, there exists $Q\in\mathrm{GL}_n(\mathbb{H})$ such that $Q\mathcal{A}Q^{-1}$ is contained in the subalgebra of M_n consisting of the block upper triangular matrices whose successive diagonal blocks have sizes given by the {1,2}-sequence $\sigma=(\sigma_1,\ldots,\sigma_m)$ of size n. For any $X\in\mathcal{A}$ let $\rho_i(X)$ denote the i-th diagonal block of the matrix QXQ^{-1}.

Assume that $\sigma_i=2$. Then $\rho_i(\mathcal{A})$ is a unital subalgebra of M_2 isomorphic to $M_2(\mathbb{R})$ or $M_2(\mathbb{C})$. By Noether–Skolem theorem see [13, p.230], there exists a matrix $P_i\in\mathrm{GL}_2(\mathbb{H})$ such that the subalgebra $P_i\rho_i(\mathcal{A})P_i^{-1}$ is exactly equal to $M_2(\mathbb{R})$ or $M_2(\mathbb{C})$, respectively. If $\sigma_i=1$ we just set $P_i=[1]$. Let $P\in\mathrm{GL}_n(\mathbb{H})$ be the block diagonal matrix with successive diagonal blocks $P_1,\ldots,P_m$. Then we have $PQ\mathcal{A}Q^{-1}P^{-1}\subseteq M_\sigma$. Therefore, $\mathcal{A}$ is quasi-triangularizable. □

The referee remarks that the theory of polynomial identities (see e.g. [4]) is relevant for the last theorem. The algebras $\mathbb{H}$, $M_2(\mathbb{R})$ and $M_2(\mathbb{C})$ all satisfy the same polynomial identities over $\mathbb{R}$ because they are all central simple of dimension 4 over their respective centers, which need not be $\mathbb{R}$. Thus a unital subalgebra $\mathcal{A}$ is quasi-triangularizable if and only if $\mathcal{A}/\mathcal{R}$ satisfies all the polynomial identities of $M_2(\mathbb{R})$.

6. Simultaneously triangularizable matrix pairs

The pairs of matrices over a field that are Simultaneously triangularizable have been studied for a long time, see e.g the book [15] and its references. Most of the known results deal with the problem of characterizing such matrix pairs. On the other hand the set of all such matrix pairs does not have a simple description, apart from a particular case which will be mentioned in the next section. In this section we make several observations concerning this problem for quaternionic matrices.

Let us start with the definition.

Definition 6.1. We denote by $\mathcal{W}_n$, the set of all matrix pairs $(A, B) \in M_n \times M_n$ such that A and B are simultaneously triangularizable.

The problem of describing $\mathcal{W}_n$ is apparently very hard (see next section for the case $n = 2$). An easy observation is that this set is semi-algebraic. Indeed, the group $G_n := \mathrm{GL}_n(\mathbb{H})$ is a real algebraic group and the map

$$G_n \times \mathcal{U}_n \times \mathcal{U}_n \to M_n \times M_n \tag{6.1}$$

which sends (g, x, y) to (gxg^{-1}, gyg^{-1}) is regular. Now observe that $\mathcal{W}_n$ is the set theoretic image of this map, and it is well known that the image of a regular map is a semi-algebraic set.

We proceed to show that $\mathcal{W}_n$ is a closed set. We claim that the image of $\mathrm{Sp}(n) \times \mathcal{U}_n \times \mathcal{U}_n$ under the map (6.1) is the whole set $\mathcal{W}_n$. Indeed, let $(x, y) \in \mathcal{W}_n$ and choose $g \in G_n$ and $a, b \in \mathcal{U}_n$ such that $x = gag^{-1}$ and $y = gbg^{-1}$. Let us write $g = ut$ where $u \in \mathrm{Sp}(n)$ and $t \in G_n \cap \mathcal{U}_n$. Then we have $x = ucu^{-1}$, $y = udu^{-1}$ with $c = tat^{-1}$ and $d = tbt^{-1}$ in $\mathcal{U}_n$. This proves our claim.

Since $\mathrm{Sp}(n)$ is a compact group and $\mathcal{U}_n \times \mathcal{U}_n$ is a closed set, we infer that $\mathcal{W}_n$ is closed in the ordinary (Euclidean) topology. Apparently, this is not true for the Zariski topology (see Problem 8.3).

Let $\mathcal{P}_n$ denote the algebra of real polynomial functions on $M_n \times M_n$. Denote by $\mathcal{P}'_n$ the subalgebra of G_n-invariant functions, i.e., functions $f \in \mathcal{P}_n$ such that

$$f(gxg^{-1}, gyg^{-1}) = f(x, y); \quad \forall g \in G_n; \forall x, y \in M_n.$$

Similarly, let $\mathcal{P}''_n$ denote the subalgebra of $\mathcal{P}''_n$ consisting of $\mathrm{Sp}(n)$-invariant functions,

Since $\mathcal{W}_n$ is G_n-invariant, its Zariski closure $\overline{\mathcal{W}}_n$ is also G_n-invariant. Let $\mathcal{I}_n \subseteq \mathcal{P}_n$ be the ideal consisting of all functions $f \in \mathcal{P}_n$ that vanish on $\mathcal{W}_n$, and set $\mathcal{I}'_n = \mathcal{I}_n \cap \mathcal{P}'_n$ and $\mathcal{I}''_n = \mathcal{I}_n \cap \mathcal{P}''_n$. By the definition of $\overline{\mathcal{W}}_n$ we have

$$\overline{\mathcal{W}}_n = \{(x, y) \in M_n \times M_n : f(x, y) = 0, \forall f \in \mathcal{I}_n\}.$$

By using the fact that $\mathrm{Sp}(n)$ is a compact group, one can easily show that also

$$\overline{\mathcal{W}}_n = \{(x, y) \in M_n \times M_n : f(x, y) = 0, \forall f \in \mathcal{I}'_n\}.$$

The algebra $\mathcal{P}_n$ is bigraded: We assign to the $4n^2$ coordinate functions of the matrix x the bidegree (1,0), and to the coordinate functions of y the bidegree (0,1). The subalgebras $\mathcal{P}'_n$ and $\mathcal{P}''_n$ inherit the bigradation from $\mathcal{P}_n$. The ideals $\mathcal{I}_n$, $\mathcal{I}'_n$ and $\mathcal{I}''_n$ are also bigraded.

We shall now exhibit two infinite families of concrete polynomials that belong to $\mathcal{I}'_n$. Let us first state an obvious result about matrices with purely imaginary eigenvalues.

Lemma 6.2. *If all eigenvalues of $A \in M_n$ are purely imaginary, then* $\mathrm{Tr}\,(A^{2k-1}) = 0$, $\mathrm{Tr}\,(A^{4k-2}) \leq 0$ *and* $\mathrm{Tr}\,(A^{4k}) \geq 0$ *for all integers $k \geq 1$.*

It is clear that, for $(X,Y) \in \mathcal{W}_n$, all eigenvalues of $[X,Y]$ are purely imaginary. Hence, we obtain as a simple corollary from above our first family of polynomial equations (and inequalities) that are satisfied on $\mathcal{W}_n$.

Corollary 6.3. *If $(X,Y) \in \mathcal{W}_n$ then*

$$\mathrm{Tr}\,([X,Y]^{2k-1}) = 0, \quad \mathrm{Tr}\,([X,Y]^{4k-2}) \leq 0, \quad \mathrm{Tr}\,([X,Y]^{4k}) \geq 0$$

are valid for all integers $k \geq 1$.

We can use the results of section 4 to obtain our second family.

Corollary 6.4. *If $(X,Y) \in \mathcal{W}_n$ then*

$$\mathrm{Tr}\,(X^k Y^k X^m Y^m - Y^k X^k Y^m X^m) = 0$$

for all integers $k, m \geq 1$.

To get more insight into the structure of the set $\mathcal{W}_n$, we shall analyze the generic fibres of the first projection map $\pi_1 : \mathcal{W}_n \to M_n$. As any matrix $A \in M_n$ is triangularizable, π_1 is surjective. We denote by $\mathcal{F}_A$ the fibre of π_1 over A, i.e.,

$$\mathcal{F}_A = \pi_1^{-1}(A) = \{(A,B) : (A,B) \in \mathcal{W}_n\}.$$

We say that a matrix $A \in M_n$ is *generic* if it has n distinct eigenvalues. The set of all generic matrices is an open dense subset of M_n. We shall now describe the generic fibres of π_1, i.e., the fibres $\mathcal{F}_A$ with A generic.

For convenience, let us identify the symmetric group S_n with the group of $n \times n$ permutation matrices.

Proposition 6.5. *For generic $A \in M_n$, the fibre $\mathcal{F}_A$ is the union of $n!$ real vector spaces, each of dimension $2n(n+1)$. Any two of these spaces intersect in a common vector subspace of dimension $\geq 4n$.*

Proof. Let $\lambda_1, \dots, \lambda_n$ be the distinct eigenvalues of A. If $P \in G_n$ then

$$\mathcal{F}_{PAP^{-1}} = \pi_1^{-1}(PAP^{-1}) = P\pi_1^{-1}(A)P^{-1} = P\mathcal{F}_A P^{-1}.$$

Hence, without any loss of generality, we may assume that A is a diagonal matrix $A = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Then it suffices to prove that

$$\mathcal{F}_A = \bigcup_{P \in S_n} \mathcal{F}_{A,P}$$

where

$$\mathcal{F}_{A,P} = \{A\} \times P\mathcal{U}_n P^{-1}.$$

Let $P \in S_n$. Since $P^{-1}\mathcal{F}_{A,P}P = \{P^{-1}AP\} \times \mathcal{U}_n \subseteq \mathcal{U}_n \times \mathcal{U}_n$, we have $\mathcal{F}_{A,P} \subseteq \mathcal{W}_n$. It follows that $\mathcal{F}_{A,P} \subseteq \mathcal{F}_A$ for all $P \in S_n$.

Conversely, let $(A, B) \in \mathcal{F}_A$. Choose $Q \in G_n$ such that

$$(QAQ^{-1}, QBQ^{-1}) \in \mathcal{U}_n \times \mathcal{U}_n.$$

Since QAQ^{-1} has n distinct eigenvalues $\lambda_1, \dots, \lambda_n$ and $QAQ^{-1} \in \mathcal{U}_n$, there is an invertible upper triangular matrix R such that $RQAQ^{-1}R^{-1}$ is a diagonal matrix with diagonal entries $\lambda_1, \dots, \lambda_n$ in some order. Hence, $RQAQ^{-1}R^{-1} = P^{-1}AP$ for some $P \in S_n$.

It follows that $S := PRQ$ commutes with A and so S is a diagonal matrix. Now $RQBQ^{-1}R^{-1} \in \mathcal{U}_n$ implies that $B \in S^{-1}P\mathcal{U}_n P^{-1}S = P\mathcal{U}_n P^{-1}$, i.e., $(A, B) \in \mathcal{F}_{A,P}$. This concludes the proof of the first assertion.

The second assertion follows from the assertion that, for each $P \in S_n$, $P\mathcal{U}_n P^{-1}, P \in S_n$, contains the space of diagonal matrices. □

We show next that $\mathcal{W}_n$ is the image of a smooth map defined on a suitable vector bundle. The group $T_n = G_n \cap \mathcal{U}_n$ acts on $\mathcal{U}_n \times \mathcal{U}_n$ by simultaneous conjugation $(t, x, y) \mapsto (txt^{-1}, tyt^{-1})$, where $t \in T_n$ and $x, y \in \mathcal{U}_n$. There is also the right action of T_n on G_n by right multiplication. By using these two actions one can construct a vector bundle

$$G_n \times_{T_n} (\mathcal{U}_n \times \mathcal{U}_n)$$

with base the homogeneous space G_n/T_n and fibre $\mathcal{U}_n \times \mathcal{U}_n$. For more details about this construction we refer the reader to [1, p.46].

The map (6.1) induces a smooth map from the above vector bundle to $M_n \times M_n$. Since $\mathcal{W}_n$ is the image of this induced map, we have

$$\dim \mathcal{W}_n \leq \dim (G_n \times_{T_n} (\mathcal{U}_n \times \mathcal{U}_n)) = 4n^2 + 2n(n+1) = 2n(3n+1).$$

We shall see next that the equality sign holds here.

Corollary 6.6. $\dim \mathcal{W}_n = 2n(3n+1)$.

Proof. Since the generic matrices form an open submanifold of M_n and each generic fibre has dimension $2n(n+1)$, we conclude that

$$\dim \mathcal{W}_n \geq 4n^2 + 2n(n+1) = 2n(3n+1).$$

Hence, equality holds. □

We conclude that $\overline{\mathcal{W}}_n$ has codimension $2n(n-1)$ in $M_n \times M_n$. Consequently, $\mathcal{I}'_n$ must have at last $2n(n-1)$ generators. In the next section we shall see that this bound is too low when $n = 2$.

7. Matrix pairs in M_2 with a common eigenvector

In this section we shall consider the special case $n = 2$. The set $\mathcal{W}_2$ can be described also as the set of all ordered pairs $A, B \in M_2$ such that A and B share a common eigenvector. For complex matrices, this special case has been fully resolved (see e.g. [11, 9, 6]). Let us recall the result.

Theorem 7.1. *For a pair of matrices $A, B \in M_2(\mathbb{C})$ the following are equivalent:*
(a) A and B are simultaneously triangularizable,
(b) $[A,B]^2 = 0$,
(c) $\mathrm{tr}\ ([A,B]^2) = 0$,
(d) $\mathrm{tr}\ (A^2B^2 - (AB)^2) = 0$,
(e) $(2\mathrm{tr}\ (A^2) - (\mathrm{tr}\ (A))^2(2\mathrm{tr}\ (B^2) - (\mathrm{tr}\ (B))^2) =$
$$(2\mathrm{tr}\ (AB) - \mathrm{tr}\ (A)\mathrm{tr}\ (B))^2.$$

Clearly, this result is much stronger than what Proposition 5.2 gives in this case.

We continue with an easy lemma of independent interest.

Lemma 7.2. *Both eigenvalues of the matrix $A \in M_2$ are purely imaginary if and only if*
(1) $\mathrm{Tr}\ (A) = \mathrm{Tr}\ (A^3) = 0$,
(2) $\mathrm{Tr}\ (A^2) \leq 0$ *and*
(3) $2\mathrm{Tr}\ (A^4) \leq (\mathrm{Tr}\ (A^2))^2 \leq 4\mathrm{Tr}\ (A^4)$.

Proof. Necessity of (1) and (2) follows directly from Lemma 6.2. For (3), we may assume $A = \begin{bmatrix} i\alpha & * \\ 0 & i\beta \end{bmatrix}$ with $\alpha, \beta \geq 0$. Then $\mathrm{Tr}\ A^2 = -2(\alpha^2 + \beta^2)$ and $\mathrm{Tr}\ A^4 = 2(\alpha^4 + \beta^4)$. It is clear from this that (3) is satisfied.

Suppose now that the conditions (1-3) hold and let λ_1, λ_2 be the eigenvalues of A. If

$$f(z) = z^4 - e_1 z^3 + e_2 z^2 - e_3 z + e_4,$$

is the characteristic polynomial for $\chi_2(A)$ then e_1, e_2, e_3, e_4 are elementary symmetric functions of the eigenvalues $\lambda_1, \lambda_2, \bar{\lambda}_1, \bar{\lambda}_2$ of $\chi_2(A)$. By (1) and Newton's identities, we have

$$e_1 = e_3 = 0, \quad e_2 = -\frac{1}{2}\mathrm{Tr}\ A^2, \quad e_4 = \frac{1}{8}\left((\mathrm{Tr}\ A^2)^2 - 2\mathrm{Tr}\ A^4\right).$$

So we have that

$$f(z) = z^4 - \frac{1}{2}(\mathrm{Tr}\ A^2)z^2 + \frac{1}{8}\left((\mathrm{Tr}\ A^2)^2 - 2\mathrm{Tr}\ A^4\right).$$

The inequalities of the lemma show that this quadratic polynomial in z^2 has two real roots, both ≤ 0. Hence the eigenvalues are indeed purely imaginary. □

As in the previous section, we obtain the following corollary.

Corollary 7.3. *Let $(A, B) \in \mathcal{W}_2$ and let $\mathcal{A} \subseteq M_2$ be the unital subalgebra generated by A and B. Then for all $X, Y \in \mathcal{A}$ we have*

(1) $\mathrm{Tr}\ ([X, Y]^3) = 0$,
(2) $\mathrm{Tr}\ ([X, Y]^2) \leq 0$, *and*
(3) $2\mathrm{Tr}\ ([X, Y]^4) \leq (\mathrm{Tr}\ ([X, Y]^2))^2 \leq 4\mathrm{Tr}\ ([X, Y]^4)$.

Proof. It suffices to observe that $\mathcal{A} \times \mathcal{A} \subseteq \mathcal{W}_2$. □

It is known that the algebra $\mathcal{P}'_2$ (see the previous section for the definition) has a minimal set of bihomogeneous generators (MSG) of cardinality 32 (see [5, 3]). In the remainder of this section we shall summarize the results that we obtained while trying to construct an MSG of the ideal $\mathcal{I}'_2 \subseteq \mathcal{P}'_2$. In our computations we used the generators constructed in [3].

Let $\mathcal{I}'_2(k, l)$ denote the subspace of $\mathcal{I}'_2$ consisting of homogeneous functions of bidegree (k, l) and let $d_{k,l}$ be its dimension. Let $\mathcal{I}'_2(s)$ be the sum of the $\mathcal{I}'_2(k, s-k)$ for $k = 0, 1, \ldots, s$ and set $d_s = \dim \mathcal{I}'_2(s)$. Since $\mathcal{W}_2$ is invariant under the switching map $(x, y) \mapsto (y, x)$, we have $d_{l,k} = d_{k,l}$ for all k and l. We have computed the dimensions $d_{k,l}$ for $k + l \leq 15$, as seen in Figure 1.

$l \swarrow$ $\searrow k$

0
0 0
0 0 0
0 0 0 0
0 0 0 0 0
0 0 0 0 ⋱
0 1 0
2 2
4 6 4
8 14 14 8
13 29 37 29 13
20 52 80 80 52 20
30 86 150 182 150 86 30
42 132 256 351 351 256 132 42
57 193 406 617 702 617 406 193 57
76 270 608 1001 1264 1264 1001 608 270 76
⋱

FIGURE 1. The dimensions $d_{k,l}$; $k, l \geq 0$; $k + l \leq 15$.

The sequence, seen isolated in Figure 1,

$$(d_{k,3})_{k\geq 0} = 0, 0, 0, 1, 2, 48, 13, 20, 30, 42, 57, 76, \ldots$$

is apparently the same as the sequence A061866 in the On-Line Encyclopedia of Integer Sequences [17]. The latter sequence $(a_k)_{k\geq 0}$ has the following definition: The integer a_k is the number of integer triples (x, y, z) such that $1 \leq x < y < z \leq k$ and $x + y + z \equiv 0 \pmod 3$. The middle "vertical" sequence

$$(d_{k,k})_{k\geq 0} = 0, 0, 0, 1, 6, 37, 180, 698, \ldots$$

is not recorded in this encyclopedia.

In principle one can construct an MSG of $\mathcal{I}'_2$ by the following routine procedure. Denote by $\mathcal{J}'_m$ the ideal of $\mathcal{P}'_2$ generated by the subspaces $\mathcal{I}'_2(k)$ for $k \leq m$. Define the subspaces $\mathcal{J}'_m(k, l)$ and $\mathcal{J}'_m(s)$ similarly to $\mathcal{I}'_2(k, l)$ and $\mathcal{I}'_2(s)$. Clearly we have that $0 = \mathcal{J}'_0 \subseteq \mathcal{J}'_1 \subseteq \cdots$ and $\mathcal{J}'_m \subseteq \mathcal{I}'_2$ for all m. Since $d_m = 0$ for $m < 6$, we also have $\mathcal{J}'_m = 0$ for $m < 6$. Since $d_6 = d_{3,3} = 1$, $\mathcal{J}'_6$ is generated by a single polynomial $f_1 \in \mathcal{J}'_2(3, 3)$, see Table 2 and formula (4.4). Since $\dim \mathcal{J}'_6(3, 4) =$

$\dim \mathcal{J}'_6(4,3) = 1$, while $d_{3,4} = d_{4,3} = 2$, the ideal $\mathcal{J}'_7$ is generated by f_1 and two new generators: $f_2 \in \mathcal{I}'_2(3,4)$ and $f_3 \in \mathcal{I}'_2(4,3)$. Similarly, $\mathcal{J}'_8$ is generated by f_1, f_2, f_3 and four new generators:

$$f_4 \in \mathcal{I}'_2(3,5); \quad f_5, f_6 \in \mathcal{I}'_2(4,4); \quad f_7 \in \mathcal{I}'_2(5,3).$$

An MSG for $\mathcal{J}'_9$ consists of $f_1, \ldots, f_7$, and ten new generators:

$$f_8, f_9 \in \mathcal{I}'_2(3,6); \quad f_{10}, f_{11}, f_{12} \in \mathcal{I}'_2(4,5);$$
$$f_{13}, f_{14}, f_{15} \in \mathcal{I}'_2(5,4); \quad f_{16}, f_{17} \in \mathcal{I}'_2(6,3).$$

To obtain an MSG for $\mathcal{J}'_{10}$, one has to add to this MSG of $\mathcal{J}'_9$ additional 19 generators $f_{18}, \ldots, f_{36}$. For $\mathcal{J}'_{11}$ we need additional 22 generators.

By Hilbert's Basis Theorem we know that this procedure must terminate and so $\mathcal{J}'_m = \mathcal{I}'_2$ holds for sufficiently large m. However we do not know the value of m. Our computations suggest that $\mathcal{J}'_{13} = \mathcal{I}'_2$.

We were able to find the first 92 generators using this procedure and hence, compute the ideal $\mathcal{J}'_m$ for $m \leq 14$. In Table 2, we give our minimal set of generators for $\mathcal{J}'_9$.

Table 2: An MSG of the ideal $\mathcal{J}'_9$

$f_1 = \mathrm{Tr}\ (xy^2x[x,y])$	$f_2 = \mathrm{Tr}\ (xy^3x[x,y])$
$f_3 = \mathrm{Tr}\ (yx^3y[x,y])$	$f_4 = \mathrm{Tr}\ (y^2x^2y^2[x,y])$
$f_5 = \mathrm{Tr}\ (xy^3x[x^2,y])$	$f_6 = \mathrm{Tr}\ ([x,y]\,[[x^2,y],[x,y^2]])$
$f_7 = \mathrm{Tr}\ (x^2y^2x^2[x,y])$	$f_8 = \mathrm{Tr}\ (yxy^3xy[x,y])$
$f_9 = \mathrm{Tr}\ ([[x,y],y]^3)$	$f_{10} = \mathrm{Tr}\ (y^2x^3y^2[x,y])$
$f_{11} = \mathrm{Tr}\ ([x,y][x,y^2][x^2,y^2])$	$f_{12} = \mathrm{Tr}\ ([[x,y],x]\,[[x,y],y]^2)$
$f_{13} = \mathrm{Tr}\ (x^2y^3x^2[x,y])$	$f_{14} = \mathrm{Tr}\ ([x,y][x^2,y][x^2,y^2])$
$f_{15} = \mathrm{Tr}\ ([[x,y],y]\,[[x,y],x]^2)$	$f_{16} = \mathrm{Tr}\ (xyx^3yx[x,y])$
$f_{17} = \mathrm{Tr}\ ([[x,y],x]^3)$	

By using our MSG for $\mathcal{J}'_{11}$, we can show that an MSG for $\mathcal{J}'_{12}$ requires additional 28 generators and $\mathcal{J}'_{13}$ requires 6. We find it surprising that an MSG of $\mathcal{I}'_2$ is so large (it has at least 92 generators). The number of generators of the given bidegree (bidegree multiplicity) contained in an MSG of $\mathcal{I}'_2$ is shown in Figure 2 for all bidegrees (k,l) with $k+l \leq 14$. The top entry corresponds to the generator f_1 of bidegree (3,3)

							l		k								
						↙		1		↘							6
							1		1								7
						1		2		1							8
					2		3		3		2						9
				1		5		7		5		1					10
			1		4		6		6		4		1				11
		1		4		5		8		5		4		1			12
	0		1		1		1		1		1		1		0		13
0		0		0		0		0		0		0		0		0	14

FIGURE 2. Bidegree multiplicities of an MSG of $\mathcal{I}_2'$.

Let us now describe one of the methods that we used to compute these generators, along with an example. We begin with the trace functions that we know vanish on $\mathcal{W}_2$ (see Corollaries 6.3, and 6.4 for available families). Notice that all these traces have bidegree of the form (k,k) and thus do not provide us with all the generators. We make the simple observation that $\mathcal{W}_2$ is invariant under the substitution $(x,y) \mapsto (x+\alpha, y+\beta)$ where $\alpha, \beta \in \mathbb{R}$.

Consider the partial derivation operators $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ on the polynomial algebra in two noncommuting indeterminates x and y. For instance

$$\frac{\partial}{\partial x}(xyxy^2) = yxy^2 + xy^3.$$

For a given noncommutative polynomial $p(x,y)$, with $\text{Tr}\,(p(x,y))$ in $\mathcal{I}_2'(k,l)$, we obtain that

$$0 = \text{Tr}\,(p(x+\alpha, y+\beta)) = \text{Tr}\,\left(\sum_{i,j=0}^{k,l} p_{i,j}(x,y)\alpha^{k-i}\beta^{l-j}\right)$$

which implies that $\text{Tr}\,(p_{i,j}(x,y)) \in \mathcal{I}_2'(i,j)$ for all i,j.

We claim that if $\text{Tr}\,(p(x,y)) \in \mathcal{I}_2'$ then also

$$\text{Tr}\,\left(\frac{\partial}{\partial x}p(x,y)\right),\ \text{Tr}\,\left(\frac{\partial}{\partial y}p(x,y)\right) \in \mathcal{I}_2'.$$

This follows from our observation above, along with the fact that $\frac{\partial}{\partial x}p(x,y)$ is equal to the coefficient of α in the expansion of $p(x+\alpha,y)$, and similarily for the other derivitave.

With this, we give explicit computation of f_{13}.

Example 7.4. Consider $p(x, y) = x^3y^3x^2y^2 - y^3x^3y^2x^2$. By Corollary 4.3 we have that Tr $(p(x, y)) \in \mathcal{I}_2'(5, 5)$. Then, we can find an element from $\mathcal{I}_2'(5, 4)$ by computing

$$\frac{\partial}{\partial y}p(x, y) = 3x^3y^2x^2y^2 + 2x^3y^3x^2y - 3y^2x^3y^2x^2 - 2y^3x^3yx^2,$$

and observing that the trace of this element is equal to

$$-2\mathrm{Tr}\ \left(x^2y^3x^2[x, y]\right).$$

Thus, we obtain the generator $f_{13} \in \mathcal{I}_2'(5, 4)$.

We have verified using Maple that the Jacobian matrix of the generators f_1, f_2, f_3, f_6 generically has rank 4. This shows that these generators are algebraically independent and agrees with the fact that $\mathcal{W}_2$ has codimension 4.

8. Some open problems

We conclude with the list of four open problems related to the topics discussed in this paper. The first problem, suggested by Lemma 6.2, is about complex numbers.

Problem 8.1. Let $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ and set $\tau_k := \Re(\lambda_1^k + \cdots + \lambda_n^k)$ for $k = 1, 2, 3, \dots$. Characterize the sequences $(\lambda_1, \dots, \lambda_n)$ for which $\tau_{2k-1} = 0$, $\tau_{4k-2} \le 0$ and $\tau_{4k} \ge 0$ for all integers $k \ge 1$.

We warn the reader that the conditions imposed on the τ_k's do not imply that all λ_i's are purely imaginary. Replacing some of the numbers λ_i with $\bar{\lambda}_i$ does not affect the conditions of the problem. Hence, without any loss of generality one may assume that all $\Im(\lambda_i) \ge 0$.

The problem we discussed in sections 6 and 7 remains open.

Problem 8.2. Find a finite set of polynomial equations and inequalities that define $\mathcal{W}_2$ as a semi-algebraic set.

Problem 8.3. Describe the Zariski closure $\overline{\mathcal{W}}_2$ and compute an MSG for the ideals $\mathcal{I}_2$, $\mathcal{I}_2'$ and $\mathcal{I}_2''$. In particular, is it true that the pair $\left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ belongs to $\overline{\mathcal{W}}_2$?

Note that this pair does not belong to $\mathcal{W}_2$. We have verified that all 92 generators of $\mathcal{J}_{14}'$ vanish on it.

Finally, Figure 1 suggests the following problem.

Problem 8.4. (a) Prove that the sequence $(d_{k,3})_{k\ge 0}$ is identical to the sequence A061866. Also construct a bijection from the set of integer triples (x, y, z), used in the definition of A061866, to a suitable basis of $\mathcal{I}_2'(k, 3)$.

(b) Identify the sequence $(d_{k,k})_{k\geq 0}$. For instance, find the generating function or an explicit formula for the $d_{k,k}$.

References

[1] G. Bredon, Introduction to Compact Transformation Groups, Academic Press, New York, 1972.

[2] C. Chevalley, Theory of Lie Groups, Princeton University Press, Princeton, 1946.

[3] D.Ž. Đoković, Poincaré series of some pure and mixed trace algebras of two generic matrices, J. Algebra 309 (2007), 654–671. (For an extended version of the paper see arXiv:math.AC/0609262v1 9 Sep 2006.)

[4] V. Drensky and E. Formanek, *Polynomial Identity Rings*, Advanced Courses in Mathematics CRM Barcelona, Birkhäuser Verlag, Basel–Boston–Berlin, 2004.

[5] V. Drensky and L. Sadikova, Generators of invariants of two 4×4 matrices, arXiv:math.RA/0503146 v2 23 Mar 2006.

[6] S. Friedland, Simultaneous similarity of matrices, Advances in Math. 50 (1983), 189–265.

[7] P. Gille and T. Szamuely, Central Simple Algebras and Galois Cohomology, Cambridge Univ. Press, 2006.

[8] L. Grunenfelder, R. Guralnick, T. Košir and H. Radjavi, Permutability of characters on algebras, Pacific J. Math. 178 (1997), 63–70.

[9] R.M. Guralnick, A note on pairs of matrices of rank one, Linear and Multilinear Algebra 8 (1979/80), 97–99.

[10] H.M. Kermani, Triangularization of matrices over quaternions, Talk given at the 13-th ILAS Conference, Amsterdam, July 18-21, 2006.

[11] T.J. Laffey, Simultaneous triangularization of matrices – low rank cases and the nonderogatory case, Linear and Multilinear Algebra 6 (1978), 269–305.

[12] C. Pearcy, A complete set of unitary invariants for operators generating finite W*-algebras of type I, Pacific J. Math. 12 (1962), 1405–1414

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

[14] H. Radjavi, A trace condition equvalent to simultaneous triangularizability, Canad. J. Math. 38 (1986), 376–386.

[15] H. Radjavi and P. Rosenthal, Simultaneous Tridiagonalization, Universitext. Springer-Verlag, New York, 2000.

[16] K.S. Sibirskiĭ, A minimal polynomial basis of unitary invariants of a square matrix of the third order, Matematicheskie Zametki, 3 (1969), 291–296.

[17] N.J.A. Sloane, The On-Line Encyclopedia of Integer Sequences, World Wide Web URL www.research.att.com/∼njas/sequences/.

[18] F. Zhang, Quaternions and matrices of quaternions, Linear Algebra Appl. 251 (1997), 21–57.

Department of Pure Mathematics, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada

E-mail address: djokovic@uwaterloo.ca

E-mail address: bh2smith@uwaterloo.ca