1 Introduction

Let A and B be two n-by-n complex matrices which are unitarily equivalent, that is, $B = U^*AU$ for some unitary matrix U. It is easily seen that matrices A^*A and B^*B have equal traces. Trace of the product of a matrix and its own conjugate is thus an example of a *unitary invariant*. More generally, consider the multiplicative semigroup W generated by the noncommuting variables x and y. We call an element of W a *word* and denote it by $w(x, y)$. If there is no confusion over the underlying variables one will also simply denote it by w. For example, $w(x, y) = y^2x^3y$ is such a word. In general a word w can always be written as $w = y^{i_1}x^{j_1} \dots y^{i_n}x^{j_n}$ where $i_1, j_1, \dots, i_n, j_n$ are positive integers except that i_1 or j_n may be zero. We say that $i_1, j_1, \dots, i_n, j_n$ are the *exponents* of a word w and w is of *length* n if i_1 and j_n are both different from zero. Otherwise w is said to be of length $n - 1$. If $n = 1$ and $i_1 = j_1 = 0$, we call w an *empty word*.

Suppose now we substitute x and y by A and A^* and regarding $w(A, A^*)$ as an n-by-n matrix. Same thing for $w(B, B^*)$. (If the word is empty we assign $w(A, A^*) = w(B, B^*) = I_n$.) One immediately sees that $\mathrm{tr}(w(A, A^*)) = \mathrm{tr}(w(B, B^*))$ for any word if A and B are unitarily equivalent. W. Specht [S] proved that the converse is also true. That is, the set $\{\mathrm{tr}(w(A, A^*)) : w(x, y)$ is any word in x and $y\}$ completely determines A up to unitary equivalence, and thus is a *complete* set of unitary invariants. There is a similar generalization due to N.Wiegmann [W] that considers not only two n-by-n matrices A and B but two finite sets of n-by-n matrices $\{A_1, A_2, \dots, A_t\}$ and $\{B_1, B_2, \dots, B_t\}$. In this situation we must consider words w with noncommuting variables in $x_1, y_1, x_2, y_2, \dots, x_t, y_t$. It states that there exists a unitary matrix U such that $U^*A_iU = B_i$ for $i = 1, 2, \dots, t$ *if and only if* for every word $w(x_1, y_1, x_2, y_2, \dots, x_t, y_t)$ we have $\mathrm{tr}(w(A_1, A_1^*, A_2, A_2^*, \dots, A_t, A_t^*))=$ $\mathrm{tr}(w(B_1, B_1^*, B_2, B_2^*, \dots, B_t, B_t^*))$.

The result of Specht gives an *infinte* set of complete unitary invariants for an n-by-n matrix A since one can form infinitely many words in A and A^*. Later, C. Pearcy showed in [P1] that a finite set of words would suffice. More precisely, let $\omega(k)$ denote the set of words in the variables x and y in which the sum of the exponents does not exceed k. Then A is unitarily equivalent to B if $\mathrm{tr}(w(A, A^*) = \mathrm{tr}(w(B, B^*))$ for every word w in $\omega(2n^2)$. This set contains fewer than 4^{n^2} elements. Of course, this upper bound is still far from satisfactory. In the other direction, Bhattarcharya [B] proved that for matrices whose nonzero singular values have multiplicity one, a family of about $(2n)^n$ traces would suffice. She also showed that there exist $n^2 + 1$ complex-valued continuous functions on $\mathbb{M}_n(\mathbb{C})$ which form a complete set of unitary invariants for n-by-n matrices where $\mathbb{M}_n(\mathbb{C})$ denotes the algebra by all n-by-n copm-lex matrices . It suggests that one would like to find a complete set of specific unitary invariants with the size of the set being a polynomial in n. For small n, it is easy to see three traces of words suffice for 2-by-2 matrices. Pearcy [P2] showed that a set of nine words suffices for $n = 3$ and Sibirskiĭ [Si] improved this by finding a set of seven words which suffices and forms a minimal set.

In Chapter 2 of this paper, we survey cases for $n = 2, 3$ and show that

$$\mathrm{tr}(A), \mathrm{tr}(A^2), \mathrm{tr}(A^*A)$$

form a complete set of unitary invariants for any n-by-n matrix A with rank 1. This will cover the 2-by-2 case immediately. We also show that three words are fewest possible. That is, one cannot find a set with two traces of words that is complete. For any 3-by-3 matrix A, we prove that

$$\mathrm{tr}(A), \mathrm{tr}(A^2), \mathrm{tr}(A^3), \mathrm{tr}(A^*A), \mathrm{tr}(A^{*2}A), \mathrm{tr}(A^{*2}A^2), \mathrm{tr}(A^{*2}A^2A^*A)$$

form a complete set of unitary invariants. The proof here is quite different from the computational proof given in [P2] and gives a bit more. The set given above plus $\mathrm{tr}(A^*A)^2$ is actually complete with respect to any n-by-n matrix A with rank 2. This will cover the case for 3-by-3 matrix readily.

In Chapter 3, we give the result that for another special class of matrices (matrix whose eigenvectors are not orthogonal), a set with no more than $n^4 + 1$ words suffices to determine such matrices up to unitary equivalence. We denote the algebra generated by A and A^* over the complex numbers by $\mathrm{Alg}(A, A^*)$. This is the set of all polynomial expressions $p(A, A^*)$, where $p(x, y)$ is any polynomial in the noncommuting variables x and y. The method lies in first proving $\{A^{*i}A^j : 0 \leq i, j < n\}$ spans $\mathrm{Alg}(A, A^*)$ for such an A. Then one modifies the method in [P1] somewhat to obtain that $\{\mathrm{tr}(A^{*i}A^jA^{*k}A^l) : 0 \leq i, j, k, l < n\} \cup \{\mathrm{tr}(A^n)\}$ forms a complete set of unitary invariants for A in this class.

2 Unitary equivalence for 2-by-2 and 3-by-3 Matrices

2.1 2-by-2 Matrices

First we state a basic lemma for general n-by-n matrices that will be applied repeatedly later.

Lemma 2.1 *Let A and B be two n-by-n matrices. If $\mathrm{tr}(A^i) = \mathrm{tr}(B^i)$ for $1 \le i \le n$, then A and B have the same eigenvalues counting algebraic multiplicities and hence $\mathrm{tr}(A^i) = \mathrm{tr}(B^i)$ for all integers i.*

Proof. Let $a_1, \ldots, a_n$ be eigenvalues of A and $b_1, \ldots, b_n$ be eigenvalues of B, each repeated according to its algebraic multiplicity. We define two classes of homogeneous polynomials $S_r, G_r : \mathbb{C}^n \to \mathbb{C}$ for $1 \le r \le n$ by

$$S_r(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} {x_i}^r$$

and

$$G_r(x_1, x_2, \ldots, x_n) = \sum_{\substack{1 \le i_1 < i_2 < \\ \cdots < i_r \le n}} x_{i_1} x_{i_2} \ldots x_{i_r}.$$

We also define $G_0 = 1$. Note that $S_1 = G_1$.

The statement that $\mathrm{tr}(A^i)$=$\mathrm{tr}(B^i)$ for $1 \le i \le n$ is the same as $S_r(a_1, \ldots, a_n) = S_r(b_1, \ldots, b_n)$ for $1 \le r \le n$. We want to show that this implies $G_r(a_1, \ldots, a_n) = G_r(b_1, \ldots, b_n)$ for $1 \le r \le n$. In fact, we have the following so-called Newton's identities:

$$rG_r - S_1 G_{r-1} + \cdots + (-1)^r S_r G_0 = 0 \tag{2.1}$$

for $1 \le r \le n$ (cf. [Pr, p.20]). Using $S_r(a_1, \ldots, a_n) = S_r(b_1, \ldots, b_n)$ for $1 \le r \le n$ together with (2.1), one concludes that $G_r(a_1, \ldots, a_n) = G_r(b_1, \ldots, b_n)$ for $1 \le r \le n$.

Now we have

$$\begin{aligned} \prod_{i=1}^{n}(x - a_i) &= \sum_{i=0}^{n}(-1)^i G_i(a_1, ..., a_n) x^{n-i} \\ &= \sum_{i=0}^{n}(-1)^i G_i(b_1, ..., b_n) x^{n-i} = \prod_{i=1}^{n}(x - b_i). \end{aligned} \tag{2.2}$$

So $a_1, ..., a_n$ and $b_1, ..., b_n$ both represent the zeros of *the same polynomial* and hence must coincide.

To prove (2.1) we define another class of homogenous polynomials $K_r^{(m)}: \mathbb{C}^n \to \mathbb{C}$ for integers r and m such that $r < n$ and $1 \le m \le n$ by

$$K_r^{(m)}(x_1, x_2, ..., x_n) = \sum_{\substack{1 \le i_1 < i_2 < .. < i_r \le n \\ i_1, i_2, .. i_r \ne m}} x_{i_1} x_{i_2} \dots x_{i_r}.$$

We also define $K_0^{(m)} = 1$ and $K_{-1}^{(m)} = 0$.

Notice that $G_r = x_m K_{r-1}^{(m)} + K_r^{(m)}$ for every m such that $1 \le m \le n$ and every r such that $0 \le r \le n$. So we have

$$S_i G_r = \sum_{j=1}^{n} x_j^i G_r = \sum_{j=1}^{n} x_j^i (x_j K_{r-1}^{(j)} + K_r^{(j)}). \tag{2.3}$$

Substituting (2.3) back into (2.1), we have

$$\begin{aligned}
rG_r + \sum_{i=1}^{r} (-1)^i G_{r-i} S_i &= rG_r + \sum_{i=1}^{r} (-1)^i \sum_{j=1}^{n} x_j^i (x_j K_{r-i-1}^{(j)} + K_{r-i}^{(j)}) \\
&= rG_r + \sum_{j=1}^{n} \left[\sum_{i=1}^{r} (-1)^i x_j^{i+1} K_{r-i-1}^{(j)} + \sum_{i=1}^{r} (-1)^i x_j^i K_{r-i}^{(j)} \right] \\
&= rG_r - \sum_{j=1}^{n} x_j K_{r-1}^{(j)} + \sum_{j=1}^{n} \left[\sum_{i=1}^{r-1} (-1)^i x_j^{i+1} K_{r-i-1}^{(j)} + \sum_{i=2}^{r} (-1)^i x_j^i K_{r-i}^{(j)} \right] \\
&= rG_r - \sum_{j=1}^{n} x_j K_{r-1}^{(j)} + \sum_{j=1}^{n} \left[\sum_{i=1}^{r-1} (-1)^i x_j^{i+1} K_{r-i-1}^{(j)} + \sum_{i=1}^{r-1} (-1)^{i+1} x_j^{i+1} K_{r-i-1}^{(j)} \right] \\
&= rG_r - \sum_{j=1}^{n} x_j K_{r-1}^{(j)} = 0.
\end{aligned}$$

Hence the assertion is proved. □

Suppose we know beforehand that A and B have r common eigenvalues (counting algebraic multiplicities). Then it is clear that we only have to check $\mathrm{tr}(A^i)$=$\mathrm{tr}(B^i)$ for $1 \le i \le n - r$. We will use mostly Lemma 2.1 in the following form:

Corollary 2.2 *Let A and B be two n-by-n matrices both of rank r. If $\mathrm{tr}(A^i)$=$\mathrm{tr}(B^i)$ for $1 \le i \le r$, then $\mathrm{tr}(A^i)$=$\mathrm{tr}(B^i)$ for all i.*

Theorem 2.3 *Let A and B be matrices both of rank 1. If $\mathrm{tr}(A)$=$\mathrm{tr}(B)$ and $\mathrm{tr}(A^*A)$ =$\mathrm{tr}(B^*B)$, then A is unitarily equivalent to B.*

Proof. Since A and B are both of rank 1, tr(A)=tr(B) guarantees that tr(A^i)=tr(B^i) for any integer i and that A and B have equal characteristic polynomials of degree 2. So all it remains to check are words w with exponents all equal to 1. That is, we still need to check if tr$(A^*A)^i$=tr$(B^*B)^i$ for any integer i. Similarly, since A^*A and B^*B are both of rank 1, tr(A^*A)=tr(B^*B) guarantees that tr$(A^*A)^i$=tr$(B^*B)^i$ for any integer i. Hence we conclude that for every word w we have tr($w(A, A^*)$)=tr($w(B, B^*)$). So by Specht's theorem A is unitarily equivalent to B. □

Theorem 2.4 *Let A and B be 2-by-2 matrices. If* tr(A)=tr(B), tr(A^2) =tr(B^2) *and* tr(A^*A)=tr(B^*B)*, then A is unitarily equivalent to B.*

Proof. Since tr(A)=tr(B) and tr(A^2)=tr(B^2) it follows that A and B have the same eigenvalues. Let λ be one of the common eigenvalues of A and B and let $A' = A - \lambda I_2$ and $B' = B - \lambda I_2$. One can easily check that tr(A)=tr(B) and tr(A^*A)=tr(B^*B) imply tr(A')=tr(B') and tr(A'^*A')=tr(B'^*B'). Since both A' and B' are of rank 1, using Theorem 2.1 we obtain that A' is unitarily equivalent to B'. Hence A is unitarily equivalent to B as well. □

In Theorem 2.4 we use traces of three words to characterize 2-by-2 matrices up to unitary equivalence. Can we still make it better? The answer is no.

Theorem 2.5 *Traces of two words do not suffice to determine a 2-by-2 matrix up to unitary equivalence.*

Proof. Let A and B be two 2-by-2 matrices. Using the property that tr(PQ)=tr(QP) for any matrices P and Q, one needs only to consider words w of the following two types: (1) $w(x, y) = y^{i_1}x^{j_1} \dots y^{i_n}x^{j_n}$ with positive integers $i_1, j_1, \dots, i_n, j_n$ and (2) $w(x, y) = y^n$ with positive integer n.

Take w_1 and w_2 to be any two words. First suppose that neither w_1 nor w_2 takes the form $(yx)^n$ for some integer n. This means that if w_1 is of the first type then there must exist some exponent of w_1 which is larger than 1. Now we may give a counterexample of

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

That tr($w_1(B, B^*)$)=tr($w_2(B, B^*)$)=0 is trivial. Suppose that w_1 is of the first type. Then we have tr($w_1(A, A^*)$)=0 because $A^2 = 0$. Suppose that w_1 is of the second type. Then we still have tr($w_1(A, A^*)$)=0 as well. So we conclude that tr($w_1(A, A^*)$)=0. Similarly, we have tr($w_2(A, A^*)$)=0. So we have

$$\text{tr}(w_1(A, A^*)) = \text{tr}(w_2(A, A^*)) = \text{tr}(w_1(B, B^*)) = \text{tr}(w_2(B, B^*)) = 0$$

while A is NOT unitarily equivalent to B.

Now suppose that we have one of the words, say, w_1 equal to $(yx)^n$. If the sum of the exponents of w_2 is even, we may take

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

as a counterexample. It is easily seen that

$$w_1(A, A^*) = w_1(B, B^*) = w_2(A, A^*) = w_2(B, B^*) = I_2.$$

Thus we again have

$$\operatorname{tr}(w_1(A, A^*)) = \operatorname{tr}(w_1(B, B^*)) \text{ and } \operatorname{tr}(w_2(A, A^*)) = \operatorname{tr}(w_2(B, B^*))$$

while A is clearly NOT unitarily equivalent to B.

Finally if we have $w_1 = (yx)^n$ for some integer n and the sum of the exponents of w_2 is odd, we may take

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \text{ and } B = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

One can check that $\operatorname{tr}(w_1(A, A^*))=\operatorname{tr}(w_1(B, B^*))$ and $\operatorname{tr}(w_2(A, A^*))=$ $\operatorname{tr}(w_2(B, B^*))$ while A is clearly NOT unitarily equivalent to B.

Since we have exhausted all possible cases, we conclude that traces of two words do not suffice to determine any 2-by-2 matrix up to unitary equivalence. □

2.2 3-by-3 Matrices

Let A and B be two n-by-n matrices of rank 2. Then $\operatorname{tr}(A)=\operatorname{tr}(B)$ and $\operatorname{tr}(A^2)=\operatorname{tr}(B^2)$ guarantee that they have equal characteristic polynomials of degree 3 (unless $n = 2$, but this case is already solved). So we only have to check

$$\operatorname{tr}(w(A, A^*)) = \operatorname{tr}(w(B, B^*))$$

for words $w(x, y) = y^{i_1}x^{j_1} \cdots y^{i_n}x^{j_n}$ with $1 \leq i_1, j_1, \ldots, i_n, j_n \leq 2$. For the sake of convenience we denote P, Q, R, S words in W by

$$P(x, y) = yx, Q(x, y) = y^2x, R(x, y) = yx^2, S(x, y) = y^2x^2. \tag{2.4}$$

It is clear that every word we need to verify can be written as the product of P,Q,R and S. Also for w a word in W, we will often denote $w(A, A^*)$ and $w(B, B^*)$ as w_A

and w_B. Whenever we write $\text{tr}(w_A)^{1,2}$=$\text{tr}(w_B)^{1,2}$ it is to be understood that we mean $\text{tr}(w_A)^i$=$\text{tr}(w_B)^i$ for $i = 1, 2$.

Let $w^{(1)}, \ldots, w^{(8)}$ be words in W defined by

$$\begin{aligned} &w^{(1)}(x,y) = x, \quad w^{(2)}(x,y) = x^2, \quad w^{(3)}(x,y) = x^3, \quad w^{(4)}(x,y) = P, \\ &w^{(5)}(x,y) = Q, \quad w^{(6)}(x,y) = S, \quad w^{(7)}(x,y) = P^2, \quad w^{(8)}(x,y) = SP. \end{aligned} \tag{2.5}$$

We are going to show that $\text{tr}(w_A^{(i)})$=$\text{tr}(w_B^{(i)})$ for $1 \le i \le 8$ implies the unitary equivalence of A and B. This can be easily applied to arbitrary 3-by-3 matrices. There is a small difference in that $w^{(7)}$ is actually not needed in the 3-by-3 case.

Lemma 2.6 *Let A and B be two n-by-n matrices both of rank 2. Suppose* $\text{tr}\,(w_A^{(i)})$=$\text{tr}\,(w_B^{(i)})$ *for $1 \le i \le 8$ as defined in* (2.5)*. Then for every word w of length 2 we have* $\text{tr}(w_A)$=$\text{tr}(w_B)$*.*

Proof. Using $\text{tr}(A^*)$=$\overline{\text{tr}(A)}$ and $\text{tr}(AB)$=$\text{tr}(BA)$ we only need to verify

$$\begin{aligned} &\text{tr}(QP)_A = \text{tr}(QP)_B, \quad \text{tr}(RQ)_A = \text{tr}(RQ)_B, \quad \text{tr}(QQ)_A = \text{tr}(QQ)_B, \\ &\text{tr}(SS)_A = \text{tr}(SS)_B, \quad \text{tr}(SQ)_A = \text{tr}(SQ)_B. \end{aligned} \tag{2.6}$$

For any complex numbers u and v, $uA + vA^*$ and $uB + vB^*$ are both matrices with rank at most 4 if A and B both have rank 2. So $\text{tr}(uA + vA^*)^i$=$\text{tr}(uB + vB^*)^i$ for $1 \le i \le 4$ and hence $\text{tr}(uA + vA^*)^i$= $\text{tr}(uB + vB^*)^i$ for any integer i and any complex numbers u and v. One can check that $\text{tr}(w_A^{(i)})$=$\text{tr}(w_B^{(i)})$ for $1 \le i \le 7$ implies $\text{tr}(uA + vA^*)^i$= $\text{tr}(uB + vB^*)^i$ for $1 \le i \le 4$. Using the equality for $i = 5$ and comparing the coefficients of u^2v^3 one gets

$$5(\text{tr}(A^{*3}A^2) + \text{tr}(QP)_A) = 5(\text{tr}(B^{*3}B^2) + \text{tr}(QP)_B).$$

However since $\text{tr}(A^{*3}A^2)$ can be written as a linear combination of $\text{tr}(S_A)$,$\text{tr}(Q_A)^*$ and $\text{tr}(A^2)$ while $\text{tr}(S_A)$=$\text{tr}(S_B)$ and $\text{tr}(Q_A)$=$\text{tr}(Q_B)$, one gets $\text{tr}(A^{*3}A^2)$=$\text{tr}(B^{*3}B^2)$ and thus $\text{tr}(QP)_A$=$\text{tr}(QP)_B$ as well.

Next we apply $\text{tr}(uA + vA^*)^6$=$\text{tr}(uB + vB^*)^6$ and comparing the coefficients of u^3v^3 to get

$$\begin{aligned} &6(\text{tr}(SP)_A + \text{tr}(RQ)_A + \text{tr}(A^{*3}A^3)) + 2 * \text{tr}(P^3)_A \\ &\qquad = 6(\text{tr}(SP)_B + \text{tr}(RQ)_B + \text{tr}(B^{*3}B^3)) + 2 * \text{tr}(P^3)_B. \end{aligned}$$

Since $\text{tr}(P^{1,2})_A$=$\text{tr}(P^{1,2})_B$ while P_A and P_B are both matrices of rank 2, it implies that $\text{tr}(P^i)_A$=$\text{tr}(P^i)_B$ for any integer i. We also have $\text{tr}(A^{*3}A^3)$=$\text{tr}(B^{*3}B^3)$ which is trivial and $\text{tr}(SP)_A$=$\text{tr}(SP)_B$ which is given. So we deduce that $\text{tr}(RQ)_A$=$\text{tr}(RQ)_B$.

Next we consider matrices uA^*A+vA^* and uB^*B+vB^* for any complex numbers u and v. Since $uA^*A+vA^* = A^*(uA+vI_n)$, uA^*A+vA^* has rank 2 for any u and v. Similarly uB^*B+vB^* has rank 2. So $\text{tr}(uA^*A+vA^*)^{1,2}$=$\text{tr}(uB^*B+vB)^{1,2}$ implies $\text{tr}(uA^*A+vA^*)^i$=$\text{tr}(uB^*B+vB^*)^i$ for any integer i. One could see that

$$\text{tr}(A^{1,2}) = \text{tr}(B^{1,2}), \text{tr}(P^{1,2})_A = \text{tr}(P^{1,2})_B, \text{tr}(Q_A) = \text{tr}(Q_B)$$

indeed imply $\text{tr}(uA^*A+vA^*)^{1,2}$=$\text{tr}(uB^*B+vB)^{1,2}$. Using the equality for i=4 and matching the coefficients of u^2v^2 one gets

$$4\text{tr}(A^{*3}AA^*A) + 2\text{tr}(Q^2)_A = 4\text{tr}(B^{*3}BB^*B) + 2\text{tr}(Q^2)_B.$$

However we have $\text{tr}(A^{*3}AA^*A)$=$\text{tr}(B^{*3}BB^*B)$ so we infer that $\text{tr}(Q^2)_A$ =$\text{tr}(Q^2)_B$.

Now we consider $uA^{*2}A+vA$ and $uB^{*2}B+vB$ for any complex numbers u and v. For the same reason as before,

$$\text{tr}(uA^{*2}A+vA)^{1,2} = \text{tr}(uB^{*2}B+vB)^{1,2}$$

implies $\text{tr}(uA^{*2}A+vA)^i$=$\text{tr}(uB^{*2}B+vB)^i$ for any integer i. Using the equality for i=4 and comparing the coefficients of u^2v^2 one gets

$$4\text{tr}(A^{*2}AA^{*2}A^3) + 2\text{tr}(S^2)_A = 4\text{tr}(B^{*2}AB^{*2}B^3) + 2\text{tr}(S^2)_B.$$

Consider matrices uQ_A+vA^2 and uQ_B+vB^2 for any complex numbers u and v. One can check that indeed $\text{tr}(uQ_A+vA^2)^{1,2}$=$\text{tr}(uQ_B+vB^2)^{1,2}$, so we have

$$\text{tr}(uQ_A+vA^2)^i = \text{tr}(uQ_B+vB^2)^i$$

for every integer i. Using the equality for i=3 and matching the coefficients of uv^2 one gets $\text{tr}(A^{*2}AA^{*2}A^3)$=$\text{tr}(B^{*2}BB^{*2}B^3)$. Hence we conclude that $\text{tr}(S^2)_A$=$\text{tr}(S^2)_B$ as well.

Next we consider $uA^{*2}A^2+vA^*A+tA^*$ and $uB^{*2}B^2+vB^*B+tB^*$ for any complex number u, v and t. For the same reason as before,

$$\text{tr}(uA^{*2}A^2+vA^*A+tA^*)^{1,2} = \text{tr}(uB^{*2}B^2+vB^*B+tB^*)^{1,2}$$

implies $\text{tr}(uA^{*2}A^2+vA^*A+tA^*)^i$=$\text{tr}(uB^{*2}B^2+vB^*B+tB^*)^i$ for any integer i. One can check that

$$\text{tr}(A) = \text{tr}(B), \text{tr}(P_A) = \text{tr}(P_B), \text{tr}(S_A) = \text{tr}(S_B),$$
$$\text{tr}(Q_A) = \text{tr}(Q_B), \text{tr}(SP)_A = \text{tr}(SP)_B$$

indeed guarantee that $\text{tr}(uA^{*2}A^2 + vA^*A + tA^*)^{1,2}$=$\text{tr}(uB^{*2}B^2 + vB^*B + wB^*)^{1,2}$. Using the equality for i=3 and matching the coefficients of uvt one gets

$$3(\text{tr}(SQ)_A + \text{tr}(A^{*3}A^2A^*A)) = 3(\text{tr}(SQ)_B + \text{tr}(B^{*3}B^2B^*B)).$$

Since we have $\text{tr}(A^{*3}A^2A^*A)$=$\text{tr}(B^{*3}B^2B^*B)$, we deduce that $\text{tr}(SQ)_A$= $\text{tr}(SQ)_B$. □

Lemma 2.7 *Let A and B be two n-by-n matrices both of rank 2. Suppose* $\text{tr}(w_A^i)$=$\text{tr}(w_B^i)$ *for* $1 \le i \le 8$. *Let k be any integer. Assume that for every word w of length k we have* $\text{tr}(w_A^{1,2})$=$\text{tr}(w_B^{1,2})$. *If for every word w' of length $k+1$ we have* $\text{tr}(w'_A)$=$\text{tr}(w'_B)$, *then* $\text{tr}(w'_A)^2$=$\text{tr}(w'_B)^2$ *as well.*

Proof. Suppose a word w' of length $n+1$ can be written as, without loss of generality, say, Qw with w a word of length n. Since for any complex number u and v matrices $uQ_A + vw_A$ and $uQ_B + vw_B$ both have rank 2,

$$\text{tr}(uQ_A + vw_A)^{1,2} = \text{tr}(uQ_B + vw_B)^{1,2}$$

implies

$$\text{tr}(uQ_A + vw_A)^i = \text{tr}(uQ_B + vw_B)^i$$

for every integer i. One could check that

$$\text{tr}(Q_A) = \text{tr}(Q_B), \text{tr}(Q^2)_A = \text{tr}(Q^2)_B, \text{tr}(w_A) = tr(w_B),$$
$$\text{tr}(w_A)^2 = \text{tr}(w_B)^2 \text{tr}(Qw)_A = \text{tr}(Qw)_B$$

indeed guarantee that $\text{tr}(uQ_A + vw_A)^{1,2}$=$\text{tr}(uQ_B + vw_B)^{1,2}$. Using the equality for i=4 and matching the coefficients of u^2v^2 one gets

$$\text{tr}(QwQw)_A + \text{tr}(Q^2w^2)_A = \text{tr}(QwQw)_B + \text{tr}(Q^2w^2)_B.$$

Using the equality for i=3 and matching the coefficients of u^2v we get

$$\text{tr}(Q^2w)_A = \text{tr}(Q^2w)_B.$$

Now we consider $uQ_A^2 + w_A$ and $uQ_B^2 + w_B$. For the same reason as above,

$$\text{tr}(uQ_A^2 + w_A)^{1,2} = \text{tr}(uQ_B^2 + w_B)^{1,2}$$

implies

$$\text{tr}(uQ_A^2 + w_A)^i = \text{tr}(uQ_B^2 + w_B)^i$$

for every integer i. One could check that

$$\text{tr}(Q_A^2) = \text{tr}(Q_B^2), \text{tr}(Q_A^4) = \text{tr}(Q_B^4), \text{tr}(w_A) = \text{tr}(w_B),$$
$$\text{tr}(w_A)^2 = \text{tr}(w_B)^2, \text{tr}(Q^2w)_A = \text{tr}(Q^2w)_B$$

imply $\mathrm{tr}(uQ_A^2 + w_A)^{1,2}$=$\mathrm{tr}(uQ_B^2 + w_B)^{1,2}$. Using the equality for i=3 and matching the coefficients of uv^2 one gets $\mathrm{tr}(Q^2w^2)_A$=$\mathrm{tr}(Q^2w^2)_B$. And thus we deduce $\mathrm{tr}(QwQw)_A$=$\mathrm{tr}(QwQw)_B$ as well. This is just $\mathrm{tr}(w'_A)^2$=$\mathrm{tr}(w'_B)^2$. □

We say that two words w and w' are *cyclically equivalent* if w' can be obtained from a cyclic permutation of w. For example, $PQRS$, $QRSP$, $RSPQ$ and $SPQR$ are all cyclically equivalent. Note that for two cyclically equivalent words w and w', $\mathrm{tr}(w_A)$=$\mathrm{tr}(w'_A)$ and $\mathrm{tr}(w_B)$=$\mathrm{tr}(w'_B)$. Thus when we check if $\mathrm{tr}(w_A)$=$\mathrm{tr}(w_B)$ we can always freely change w to any word w' that is cyclically equivalent to w.

Theorem 2.8 *Let A and B be two n-by-n matrices both of rank* 2. *If* $\mathrm{tr}(w_A^{(i)})$=$\mathrm{tr}(w_B^{(i)})$ *for* $1 \leq i \leq 8$, *then A is unitarily equivalent to B.*

Proof. We proceed by induction on the length n of words. From Lemmas 2.6 and 2.7 we see that for words w of length 1 or 2, one has $\mathrm{tr}(w_A)^{1,2}$=$\mathrm{tr}(w_B)^{1,2}$. This proves our assertion for $n = 1$ and $n = 2$.

Now suppose for $n = k, k+1$, every word w of block n satisfies $\mathrm{tr}(w_A)^{1,2}$=$\mathrm{tr}(w_B)^{1,2}$. Let w' be any word of length $k+2$. If w' is *not* any of the forms $(SP)^n$, $(PS)^n$, $(QR)^n$ or $(RQ)^n$, then we have w cyclically equivalent to one of the following 12 cases:

(1) $w = PPK$ for some word K of length k

Consider for any complex numbers u and v matrices $uP_A + vK_A$ and $uP_B + vK_B$. Note that $uP_A + vK_A$ and $uP_B + vK_B$ are both of rank 2. Also,

$$\mathrm{tr}(uP_A + vK_A)^{1,2} = \mathrm{tr}(uP_B + vK_B)^{1,2}$$

implies

$$\mathrm{tr}(uP_A + vK_A)^i = \mathrm{tr}(uP_B + vK_B)^i$$

for every integer i. From $\mathrm{tr}(P_A)$=$\mathrm{tr}(P_B)$ we see that the coefficients of u of both sides are equal. Since K is of length k, by induction hypothesis we have $\mathrm{tr}(K_A)$=$\mathrm{tr}(K_B)$ and so the coefficients of v of both sides are equal. From $\mathrm{tr}(P_A)^2$=$\mathrm{tr}(P_B)^2$ we see that the coefficients of u^2 of both sides are equal. Since PK is of length $k+1$, by induction hypothesis we have $\mathrm{tr}(PK)_A$=$\mathrm{tr}(PK)_B$ and so the coefficients of uv of both sides are equal. Finally, since K is of length k, by the induction hypothesis we have $\mathrm{tr}(K_A)^2$=$\mathrm{tr}(K_B)^2$ and so the coefficients of v^2 of both sides are equal. Hence we conclude that $\mathrm{tr}(uP_A + vK_A)^{1,2}$ =$\mathrm{tr}(uP_B + vK_B)^{1,2}$. Using the equality for i=3 and matching the coefficients of uv^2 we get $\mathrm{tr}(PPK)_A$=$\mathrm{tr}(PPK)_B$.

(2) $w = PRK$ for some word K of length k:

Consider for any complex numbers u and v the matrices $uP_A + vAK_A$ and $uP_B + vBK_B$. One could check indeed that $\mathrm{tr}(uP_A+vAK_A)^{1,2}$ $=\mathrm{tr}(uP_B+vBK_B)^{1,2}$ and this implies $\mathrm{tr}(uP_A + vAK_A)^i$ $=\mathrm{tr}(uP_B + vBK_B)^i$ for every integer i. Using the equality for $i = 3$ and matching the coefficients of u^2v one gets $\mathrm{tr}(PRK)_A$=$\mathrm{tr}(PRK)_B$.

(3) $w = RPK$ for some word K of length k:

Matching the coefficients of uvt in $\mathrm{tr}(uP_A + vR_A + tK_A)^3$=$\mathrm{tr}(uP_B + vR_B + tK_B)^3$ one gets $\mathrm{tr}(PRK)_A$+$\mathrm{tr}(RPK_A)$=$\mathrm{tr}(PRK)_B$+$\mathrm{tr}(RPK_B)$. Thus from (2) we deduce $\mathrm{tr}(RPK_A)$=$\mathrm{tr}(RPK_B)$.

(4) $w = PQK$ for some word K of length k:

Note that $\mathrm{tr}(PQK)^*_A$=$\mathrm{tr}(K^*RP)_A$=$\mathrm{tr}(RPK^*)_A$. Applying (3) we obtain $\mathrm{tr}(PQK)_A$ $=\mathrm{tr}(PQK)_B$.

(5) $w = QPK$ for some word K of length k:

Matching the coefficients of uvt in $\mathrm{tr}(uQ_A + vP_A + tK_A)^3$=$\mathrm{tr}(uQ_B + vP_B + tK_B)^3$, one gets $\mathrm{tr}(QPK)_A$+$\mathrm{tr}(PQK_A)$=$\mathrm{tr}(QPK)_B$+$\mathrm{tr}(PQK_B)$. Thus from (2) we deduce $\mathrm{tr}(QPK_A)$=$\mathrm{tr}(QPK_B)$.

(6) $w = QQK$ for some word K of length k:

Consider for any complex numbers u and v the matrices uQ_A+vK_A and uQ_B+vK_B. One could check indeed that $\mathrm{tr}(uQ_A + vK_A)^{1,2}$ $=\mathrm{tr}(uQ_B + vK_B)^{1,2}$ and this implies $\mathrm{tr}(uQ_A + vK_A)^i$ $=\mathrm{tr}(uQ_B + vK_B)^i$ for every integer i. Using the equality for $i = 3$ and matching the coefficients of u^2v one gets $\mathrm{tr}(QQK)_A$=$\mathrm{tr}(QQK)_B$.

(7) $w = QSK$ for some word K of length k:

Consider for any complex numbers u and v the matrices $uQ_A + vAK_A$ and $uQ_B + vBK_B$. One could check indeed that $\mathrm{tr}(uQ_A+vAK_A)^{1,2}$ $=\mathrm{tr}(uQ_B+vBK_B)^{1,2}$ and this implies $\mathrm{tr}(uQ_A + vAK_A)^i$ $=\mathrm{tr}(uQ_B + vBK_B)^i$ for every integer i. Using the equality for $i = 3$ and matching the coefficients of u^2v one gets $\mathrm{tr}(QSK)_A$=$\mathrm{tr}(QSK)_B$.

(8) $w = RRK$ for some word K of length k:

Consider for any complex numbers u and v the matrices uR_A+vK_A and uR_B+vK_B. One could check indeed that $\mathrm{tr}(uR_A + vK_A)^{1,2}$ $=\mathrm{tr}(uR_B + vK_B)^{1,2}$ and this implies $\mathrm{tr}(uR_A + vK_A)^i$ $=\mathrm{tr}(uR_B + vK_B)^i$ for every integer i. Using the equality for $i = 3$

and matching the coefficients of u^2v one gets $\text{tr}(RRK)_A$=$\text{tr}(RRK)_B$.

(9) $w = SQK$ for some word K of length k:

Matching the coefficients of uvt in $\text{tr}(uQ_A + vS_A + tK_A)^3$=$\text{tr}(uQ_B + vS_B + tK_B)^3$ one gets $\text{tr}(QSK)_A$+$\text{tr}(SQK_A)$=$\text{tr}(QSK)_B$+$\text{tr}(SQK_B)$. Thus from (7) we deduce $\text{tr}(SQK_A)$=$\text{tr}(SQK_B)$.

(10) $w = RSK$ for some word K of length k:
Note that $\text{tr}(RSK)^*_A$=$\text{tr}(K^*SQ)_A$=$\text{tr}(SQK^*)_A$. Applying (9) we obtain $\text{tr}(RSK)_A$= $\text{tr}(RSK)_B$.

(11) $w = SRK$ for some word K of length k:

Matching the coefficients of uvt in $\text{tr}(uS_A + vR_A + tK_A)^3$=$\text{tr}(uS_B + vR_B + tK_B)^3$ one gets $\text{tr}(RSK)_A$+$\text{tr}(SRK_A)$=$\text{tr}(RSK)_B$+$\text{tr}(SRK_B)$, Thus from (10) we deduce $\text{tr}(SRK_A)$=$\text{tr}(SRK_B)$.

(12) $W = SSK$ for some word K of length k

Consider for any complex numbers u and v the matrices uS_A+vK_A and uS_B+vK_B. One could check indeed that $\text{tr}(uS_A + vK_A)^{1,2}$ =$\text{tr}(uS_B + vK_B)^{1,2}$ and this implies $\text{tr}(uS_A + vK_A)^i$ =$\text{tr}(uS_B + vK_B)^i$ for every integer i. Using the equality for $i = 3$ and matching the coefficients of u^2v one gets $\text{tr}(SSK)_A$=$\text{tr}(SSK)_B$.

For $w' = (PS)^n$, we showed in Lemma 2.6 that $\text{tr}(PS)_A$=$\text{tr}(PS)_B$ and $\text{tr}(PSPS)_A$ =$\text{tr}(PSPS)_B$. Since $(PS)_A$ and $(PS)_B$ are both of rank 2, this implies that $\text{tr}(PS)^n_A$= $\text{tr}(PS)^n_B$ for all integers n. Similarly, we have $\text{tr}(QR)^n_A$=$\text{tr}(QR)^n_B$ for all integers n. So we have proved that $\text{tr}(w'_A)$=$\text{tr}(w'_B)$ for every word w' of length $k + 2$. Finally, from Lemma 2.7 and the induction hypothesis that $\text{tr}(w_A)^{1,2}$=$\text{tr}(w_B)^{1,2}$ for every word w of length $k + 1$ we conclude that $\text{tr}(w'_A)^2$=$\text{tr}(w'_B)^2$ for every w' of length $k + 2$. Thus we conclude by induction that $\text{tr}(w_A)$=$\text{tr}(w_B)$ for any w of length $n, n \geq 1$. So A is unitarily equivalent to B. □

Now let us come back to the 3-by-3 cases. First we show that $w^{(7)}$ is redundant.

Lemma 2.9 *Suppose A and B are two 3-by-3 matrices. If $\text{tr}(w^{(i)}_A)$=$\text{tr}(w^{(i)}_B)$ for $1 \leq i \leq 6$, then $\text{tr}(w^{(7)}_A)$=$\text{tr}(w^{(7)}_B)$*

Proof. For any complex numbers u and v consider matrices $uA + vA^*$ and $uB + vB^*$. One could check that tr (w^i_A)=$\text{tr}(w^i_B)$ for $1 \leq i \leq 5$ implies

$$\text{tr}(uA + vA^*)^{1,2,3} = \text{tr}(uB + vB^*)^{1,2,3}.$$

Hence we have $\mathrm{tr}(uA + vA^*)^i$=$\mathrm{tr}(uB + vB^*)^i$ also for every integer i. Using the equality for $i = 4$ and matching the coefficients of u^2v^2 one gets

$$\mathrm{tr}(w_A^{(6)}) + \mathrm{tr}(w_A^{(7)}) = \mathrm{tr}(w_B^{(6)}) + \mathrm{tr}(w_B^{(7)}).$$

From tr $(w_A^{(6)})$=$\mathrm{tr}(w_B^{(6)})$ we then deduce tr $(w_A^{(7)})$=$\mathrm{tr}(w_B^{(7)})$ □

Lemma 2.10 *Suppose A and B are two 3-by-3 matrices and* tr $(w_A^{(i)})$=$\mathrm{tr}(w_B^{(i)})$ *for* $1 \le i \le 6$ *and* $i = 8$. *Then, for any complex number* λ, $\mathrm{tr}(w_{(A-\lambda I_3)}^{(i)})$=$\mathrm{tr}(w_{(B-\lambda I_3)}^{(i)})$ *for* $1 \le i \le 6$ *and* $i = 8$ *as well.*

Proof. The only case we really need to verify is $\mathrm{tr}(w_{(A-\lambda I_3)}^{(8)})$=$\mathrm{tr}(w_{(B-\lambda I_3)}^{(8)})$. This in turn is equivalent to checking that if $\mathrm{tr}(A^{*2}AA^*A)$=$\mathrm{tr}(B^{*2}BB^*B)$. Using

$$\mathrm{tr}(uA + vA^*)^7 = \mathrm{tr}(uB + vB^*)^7$$

and matching the coefficients of u^3v^4 one gets indeed

$$\mathrm{tr}(A^{*2}AA^*A) = \mathrm{tr}(B^{*2}BB^*B).$$

□

Theorem 2.11 *If A and B are two 3-by-3 matrices such that* $\mathrm{tr}(w_A^{(i)})$ =$\mathrm{tr}(w_B^{(i)})$ *for* $1 \le i \le 6$ *and* $i = 8$, *then A is unitarily equivalent to B.*

Proof. First from $\mathrm{tr}(A)^{1,2,3}$=$\mathrm{tr}(B)^{1,2,3}$ we conclude that A and B have the same eigenvalues. Subtracting a common eigenvalue λ we get $A' = A - \lambda I_3$ and $B' = B - \lambda I_3$ both having rank 2. From Lemma 2.5 we have tr $(w_{A'}^{(i)})$=$\mathrm{tr}(w_{B'}^{(i)})$ for $1 \le i \le 6$ and $i = 8$. From Lemma 2.4 we see $\mathrm{tr}(w_{A'}^{(7)})$=$\mathrm{tr}(w_B^{(7)})$ as well. From Theorem 2.1 we conclude that A' is unitarily equivalent to B'. Thus A is also unitarily equivalent to B. □

3 Matrices with Eigenvectors Not Orthogonal

We prove in this chapter the most general result in this paper, namely, for any matrix whose eigenvectors are not orthogonal a set with $n^4 + 1$ words suffices to determine it up to unitary equivalence. For example,

$$A = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}, \; b \neq 0$$

is such a matrix.

Theorem 3.1 *Let A be an n-by-n matrix such that none of its eigenvectors are orthogonal. Then $\{A^{*i}A^j : 0 \leq i, j < n\}$ is a linearly independent set and spans* $\mathrm{Alg}(A, A^*)$.

Proof. Suppose that

$$f(x) = (x - \lambda_1)^{p_1}(x - \lambda_2)^{p_2} \ldots (x - \lambda_m)^{p_m}$$

is the characteristic polynomial of A with $\lambda_1, \lambda_2, ..., \lambda_m$ the distinct eigenvalues of A. Since none of the eigenvectors of A are orthogonal, we have in particular $\ker(A - \lambda_i I_n) = 1$ for all i, on $1 \leq i \leq m$. This implies that f is also the minimal polynomial of A. Suppose on the contrary that $\{A^{*i}A^j : 0 \leq i, j < n\}$ is linearly dependent. Then there exist polynomials $g_0, g_1, \ldots, g_{n-1}$ of degree $< n$, not all zero, such that

$$g_{n-1}(A^*)A^{n-1} + \cdots + g_0(A^*) = 0. \tag{3.1}$$

Let h_{n-1}=gcd$(\overline{f}, g_{n-1})$ and write $g_{n-1} = qh_{n-1}$ with q also a polynomial. Then $q(A^*)$ is invertible with its inverse also a polynomial in A^*. Thus we can multiply both sides of (3.1) with $j(A^*)^{-1}$ and obtain

$$h_{n-1}(A^*)A^{n-1} + \cdots + h_0(A^*) = 0 \tag{3.2}$$

with $h_0, h_1, \ldots, h_{n-1}$ still polynomials of degree $< n$. Since h_{n-1} divides $\overline{f}$, there is some r such that $h_{n-1}(x)$ divides $(x - \overline{\lambda_1})^{p_1} \cdots (x - \overline{\lambda_r})^{p_r - 1} \cdots (x - \overline{\lambda_m})^{p_m}$. We multiply both sides of (3.2) with $(A^* - \overline{\lambda_1})^{p_1} \cdots (A^* - \overline{\lambda_r})^{p_r-1} \cdots (A^* - \overline{\lambda_m})^{p_m}$ and obtain for some nonzero polynomial h the relation

$$(A^* - \overline{\lambda_1})^{p_1} \cdots (A^* - \overline{\lambda_r})^{p_r-1} \cdots (A^* - \overline{\lambda_m})^{p_m} h(A) = 0. \tag{3.3}$$

Similarly, take l=gcd(f, h) and denote $h = lu$. Then $u(A)$ is invertible with its inverse also a polynomial in A. So we multiply both sides of (3.3) with $u(A)^{-1}$ and obtain

$$(A^* - \overline{\lambda_1})^{p_1} \cdots (A^* - \overline{\lambda_r})^{p_r-1} \cdots (A^* - \overline{\lambda_m})^{p_m} l(A) = 0. \tag{3.4}$$

Since l divides f and is of degree $< n$, there is some s such that l divides $(x - \lambda_1)^{p_1} \cdots (x - \lambda_s)^{p_s - 1} \cdots (x - \lambda_m)^{p_m}$. So we multiply both sides of (3.4) with $(A - \lambda_1)^{p_1} \cdots (A - \lambda_s)^{p_s - 1} \cdots (A - \lambda_m)^{p_m}$ and obtain

$$(A^* - \overline{\lambda_1})^{p_1} \cdots (A^* - \overline{\lambda_r})^{p_r - 1} \cdots (A^* - \overline{\lambda_m})^{p_m} \\ (A - \lambda_1)^{p_1} \cdots (A - \lambda_s)^{p_s - 1} \cdots (A - \lambda_m)^{p_m} = 0. \quad (3.5)$$

Since f is the minimal polynomial of A, there exists some nonzero v in $\mathbb{C}^n$ which is not in $\ker(A - \lambda_r)^{p_r - 1}$ but in $\ker(A - \lambda_r)^{p_r}$. So $(A - \lambda_r)^{p_r - 1} v$ is nonzero and belongs in $\ker(A - \lambda_r)$. Let v_r in $\mathbb{C}^n$ be an eigenvector of A corresponding to the eigenvalue λ_r. Then for some nonzero complex number a we have $(A - \lambda_r)^{p_r - 1} v = a v_r$. Similarly, for some nonzero w in $\mathbb{C}^n$ and nonzero complex number b, we have $(A - \lambda_s)^{p_s - 1} w = b v_s$, where v_s is an eigenvector of A corresponding to the eigenvalue λ_s. From (3.5), we must have

$$\left\langle (A - \lambda_1)^{p_1} ... (A - \lambda_r)^{p_r - 1} ... (A - \lambda_m)^{p_m} v, (A - \lambda_1)^{p_1} ... (A - \lambda_s)^{p_s - 1} ... (A - \lambda_m)^{p_m} w \right\rangle \\ = 0.$$

This in turn implies

$$\left\langle \left(\prod_{1 \le i \le m; i \ne r} (A - \lambda_i)^{p_i} \right) v_r, \left(\prod_{1 \le i \le m; i \ne s} (A - \lambda_i)^{p_i} \right) v_s \right\rangle = 0$$

and hence

$$\langle v_r, v_s \rangle = 0.$$

So we find a pair of eigenvectors of A which are orthogonal, contradicting to our assumption on A. Thus the set $\{A^{*i} A^j : 0 \le i, j < n\}$ must be linearly independent. Since this set contains n^2 elements and $\mathrm{Alg}(A, A^*)$ is of dimension at most n^2, this implies that $\{A^{*i} A^j : 0 \le i, j < n\}$ spans $\mathrm{Alg}(A, A^*)$. $\square$

Suppose now that A and B are two n-by-n matrices with none of the eigenvectors of A orthogonal. By Theorem 3.1, for any integers p and q we can write $A^p A^{*q}$ as a linear combination of $A^{*i} A^j, 0 \le i, j < n$. The next theorem shows that if

$$\mathrm{tr}(A^n) = \mathrm{tr}(B^n) \text{ and } \mathrm{tr}(A^{*i} A^j A^{*k} A^l) = \mathrm{tr}(B^{*i} B^j B^{*k} B^l),\ 0 \le i, j, k, l < n,$$

then we can also write $B^p B^{*q}$ as a linear combination of $\{B^{*i} B^j : 0 \le i, j < n\}$ *with the same coefficients.*

Theorem 3.2 *Let A and B be two n-by-n matrices and p and q be some integers. Suppose that there exists a set of complex numbers $\{a_{ij} : 0 \le i, j < n\}$ such that $A^p A^{*q} = \sum_{i,j=0}^{n-1} a_{ij} A^{*i} A^j$. If*

$$tr(A^n) = tr(B^n) \text{ and } tr(A^{*i} A^j A^{*k} A^l) = tr(B^{*i} B^j B^{*k} B^l),\ 0 \le i, j, k, l < n,$$

then we also have $B^pB^{*q} = \sum_{i,j=0}^{n-1} a_{ij}B^{*i}B^j$.

Proof. From $A^pA^{*q} = \sum_{i,j=0}^{n-1} a_{ij}A^{*i}A^j$, we have

$$\mathrm{tr}\left((A^pA^{*q} - \sum_{i,j=0}^{n-1} a_{ij}A^{*i}A^j)^*(A^pA^{*q} - \sum_{i,j=0}^{n-1} a_{ij}A^{*i}A^j)\right) = 0.$$

This is the same as

$$\mathrm{tr}(A^qA^{*p}A^pA^{*q}) - \mathrm{tr}(\sum_{i,j=0}^{n-1} a_{ij}A^qA^{*p}A^{*i}A^j) \\ - \mathrm{tr}(\sum_{i,j=0}^{n-1} a_{ij}A^{*i}A^jA^qA^{*p}) - \mathrm{tr}(\sum_{i,j,k,l=0}^{n-1} a_{ij}a_{kl}A^{*j}A^iA^{*k}A^l) = 0. \quad (3.6)$$

Since $\mathrm{tr}(A^i) = \mathrm{tr}(B^i), 1 \le i \le n$, A and B have equal characteristic polynomials and thus $\mathrm{tr}(A^{*i}A^jA^{*k}A^l) = \mathrm{tr}(B^{*i}B^jB^{*k}B^l), 0 \le i,j,k,l < n$ implies $\mathrm{tr}(A^{*i}A^jA^{*k}A^l) = \mathrm{tr}(B^{*i}B^jB^{*k}B^l)$ for all nonnegative integers i, j, k and l. Substituting this back into (3.6), we obtain

$$\mathrm{tr}(B^qB^{*p}B^pB^{*q}) - \mathrm{tr}(\sum_{i,j=0}^{n-1} a_{ij}B^qB^{*p}B^{*i}B^j) \\ - \mathrm{tr}(\sum_{i,j=0}^{n-1} a_{ij}B^{*i}B^jB^qB^{*p}) - \mathrm{tr}(\sum_{i,j,k,l=0}^{n-1} a_{ij}a_{kl}B^{*j}B^iB^{*k}B^l) = 0 \quad (3.7)$$

as well. So we have

$$\mathrm{tr}(B^pB^{*q} - \sum_{i,j=0}^{n-1} a_{ij}B^{*i}B^j)^*(B^pB^{*q} - \sum_{i,j=0}^{n-1} a_{ij}B^{*i}B^j) = 0.$$

Thus $B^pB^{*q} - \sum_{i,j=0}^{n-1} a_{ij}B^{*i}B^j = 0$. □

Now we are ready to prove the main result.

Theorem 3.3 *Let A and B be two n-by-n matrices. Assume that no pair of the eigenvectors of A are orthogonal to each other. If* $\mathrm{tr}(A^{*i}A^jA^{*k}A^l) = \mathrm{tr}(B^{*i}B^jB^{*k}B^l)$ *for* $0 \le i,j,k,l < n$ *and* $\mathrm{tr}(A^n) = \mathrm{tr}(B^n)$, *then A is unitarily equivalent to B.*

Proof. We are going to show that for every word $w(x, y)$, $\mathrm{tr}(w(A, A^*)) = \mathrm{tr}(w(B, B^*))$ and then apply Specht's theorem to conclude that A is unitarily equivalent to B. Using the property that $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, it suffices to consider only words w of the form $w(x, y) = y^{i_1}x^{j_1}\cdots y^{i_n}x^{j_n}$.

We proceed by induction on the length n. For n=1, this is already assumed. Suppose that the assertion is true for $n = k$. Consider the case $n = k + 1$. Then

$$\begin{aligned}\mathrm{tr}(w(A, A^*)) &= \mathrm{tr}(A^{*i_1}A^{j_1}\cdots A^{*i_k}(A^{j_k}A^{*i_{k+1}})A^{j_{k+1}}) \\ &= \mathrm{tr}\left(\sum_{p,q=0}^{n-1} a_{pq}A^{*i_1}A^{j_1}\cdots A^{*(i_k+p)}A^q\right) = \sum_{p,q=0}^{n-1} a_{pq}\mathrm{tr}\left(A^{*i_1}A^{j_1}\cdots A^{*(i_k+p)}A^q\right). \quad (3.8)\end{aligned}$$

Since $A^{*i_1}A^{j_1}\cdots A^{*(i_k+p)}A^q$ is of length k, by the induction hypothesis we have

$$\mathrm{tr}(A^{*i_1}A^{j_1}\cdots A^{*(i_k+p)}A^q) = \mathrm{tr}(B^{*i_1}B^{j_1}\cdots B^{*(i_k+p)}B^q). \quad (3.9)$$

Substituting this back into (3.8), we have

$$\begin{aligned}\mathrm{tr}(w(A, A^*)) &= \sum_{p,q=0}^{n-1} a_{pq}\mathrm{tr}\left(A^{*i_1}A^{j_1}\cdots A^{*(i_k+p)}A^q\right) \\ &= \sum_{p,q=0}^{n-1} a_{pq}\mathrm{tr}\left(B^{*i_1}B^{j_1}\cdots B^{*(i_k+p)}B^q\right) \\ &= \mathrm{tr}\left(\sum_{p,q=0}^{n-1} a_{pq}B^{*i_1}B^{j_1}\cdots B^{*(i_k+p)}B^q\right) \\ &= \mathrm{tr}(B^{*i_1}B^{j_1}\cdots B^{*i_k}(B^{j_k}B^{*i_{k+1}})B^{j_{k+1}}) = \mathrm{tr}(w(B, B^*)).\end{aligned}$$

Thus by the mathematical induction we conclude that $\mathrm{tr}(w(A, A^*))$=$\mathrm{tr}(w(B, B^*))$ for every word w. So A is unitarily equivalent to B. □

References

[B] Bhattarcharya, On the unitary invariants of an $n \times n$ matrix, Ph.D. Thesis, Indian Statis. Inst., New Delhi, 1987.

[P1] C. Pearcy, A complete set of unitary invariants for operators generating finite W^*-algebras of type I, *Pacific J. Math.* 12:1405–1416 (1962).

[P2] C. Pearcy, A complete set of unitary invariants for 3×3 complex matrices, *Trans. Amer. Math. Soc.* 104:425–429 (1962).

[Pr] V. V. Praslov, *Problems and Theorems in Linear Algebra*, Amer. Math. Soc., Providence, 1994.

[Sh] H. Shapiro, A survey of canonical forms and invariants for unitary similarity, *Linear Algebra Appl.* 147:101–167 (1991).

[Si] K. S. Sibirskiĭ, Unitary and orthogonal invariants of matrices, *Soviet Math. Dokl.* 8:36–40 (1967) [English transl. of *Dokl. Akad. Nauk SSSR* 172:40–43 (1967)].

[S] W. Specht, Zur Theorie der Matrizen, II, *Jahresber. Deutsch. Math.-Verein.* 50:19–23 (1940).

[W] N.Wiegmann, Necessary and sufficient conditions for unitary similarity, *J. Austral. Math. Soc.* 2:122–126 (1961/62).