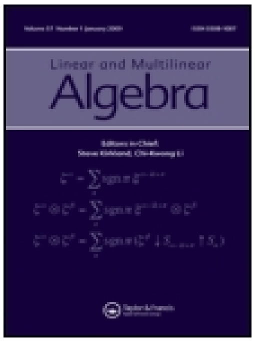

Linear and Multilinear Algebra

ISSN: 0308-1087 (Print) 1563-5139 (Online) Journal homepage: http://www.tandfonline.com/loi/glma20

An approximate, multivariable version of Specht's theorem

L. W. Marcoux , M. Mastnak & H. Radjavi

Linear and Multilinear Algebra, Vol. 55, No. 2, March 2007, 159–173

An approximate, multivariable version of Specht's theorem

L. W. MARCOUX*†, M. MASTNAK‡ and H. RADJAVI†

†Department of Pure Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1
‡Department of Mathematics, University of British Columbiam, Room 121, 1984 Mathematics Road, Vancouver, British Columbia, Canada V6T 1Z2

Communicated by C.-K. Li

(*Received 8 December 2005; in final form 4 March 2006*)

In this article we provide generalizations of Specht's theorem which states that two $n \times n$ matrices A and B are unitarily equivalent if and only if all traces of words in two non-commuting variables applied to the pairs (A, A^*) and (B, B^*) coincide. First, we obtain conditions which allow us to extend this to simultaneous similarity or unitary equivalence of families of operators, and secondly, we show that it suffices to consider a more restricted family of functions when comparing traces. Our results do not require the traces of words in (A, A^*) and (B, B^*) to coincide, but only to be close.

Keywords: Specht's theorem; Approximation; Trace; Semigroups

2000 Mathematics Subject Classifications: Primary 15A21; Secondary15A04

1. Introduction

A useful tool in determining whether two $n \times n$ complex matrices A and B are similar is to compare their Jordan canonical forms. In practice, deciding whether they are unitarily equivalent is a much more difficult problem. A theorem of Specht [5] tells us that A and B are unitarily equivalent if and only if $\mathrm{tr}(w(A, A^*)) = \mathrm{tr}(w(B, B^*))$ for all words w in two non-commuting variables. Specht's theorem was later improved by Pearcy [4], who showed that $A, B \in \mathbb{M}_n(\mathbb{C})$ are unitarily equivalent if and only if $\mathrm{tr}(w(A, A^*)) = \mathrm{tr}(w(B, B^*))$ for all words w of degree at most $2n^2$. (In a private communication, Djokovic has informed us that by combining a theorem of Razmyslov with the work of Procesi, it can be shown that in fact it suffices to

*Corresponding author. Email: LWMarcoux@uwaterloo.ca

Linear and Multilinear Algebra
ISSN 0308-1087 print/ISSN 1563-5139 online © 2007 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/03081080600693285

consider words of length at most n^2. More precisely, Razmyslov's result improves a bound in the Nagata-Higman theorem, while the work of Procesi establishes the equality of that bound and the length of words necessary to determine the unitary equivalence of two $n \times n$ matrices. The proof, however, is somewhat involved, and will not be elaborated here. We direct the interested reader to Chapter 6 of [1]).

The present work examines to what extent the Specht's theorem can be generalized. First, one can ask whether by only knowing that the traces of the words in A and A^* are *close* to the traces of the same words in B and B^* implies that A and B are close to being unitarily equivalent. In section 2, we show that if A and B are unitary matrices for which $\operatorname{tr}(A^k)$ lies within distance 1 of $\operatorname{tr}(B^k)$ for all powers $k \in \mathbb{Z}$, then A and B are unitarily equivalent. In section 3, we consider the indexed families for which the traces of words are close. Under certain natural conditions, we are able to conclude the existence of a single invertible matrix which implements the simultaneous similarity of the two families (see Corrollary 3.10). When the families are self-adjoint, the notion of similarity may be replaced by a unitary equivalence.

One may also ask whether it is sufficient to consider a more restricted class of words w in two non-commuting variables in the statement of the Specht's theorem. In section 4, we show that if $A, B \in \mathbb{M}_n(\mathbb{C})$ and if $\operatorname{tr}|p(A, A^*)| = \operatorname{tr}|p(B, B^*)|$ for all polynomials p in two non-commuting variables (here $|T| = (T^*T)^{1/2}$ for $T \in \mathbb{M}_n(\mathbb{C})$), then A is unitarily equivalent to B. This condition is shown to be equivalent to a condition involving only projection-valued polynomials.

2. The single variable unitary case

Suppose $y_1, y_2, \ldots, y_r$ are complex numbers of modulus 1. We shall say that $y_1, y_2, \ldots, y_r$ are *independent* if for each quotient f of two words (i.e. monomials) in r variables (equivalently, if f is a word in r variables and their inverses), the condition $f(y_1, y_2, \ldots, y_r) = 1$ implies that $f \equiv 1$. An equivalent formulation is that $y_1, y_2, \ldots, y_r$ are independent if $(\log(y_k)/2\pi i)_{k=1,\ldots,r}$ and 1 are linearly independent over the rational numbers.

Lemma 2.1 *If $x_1, \ldots, x_m$ are complex numbers of modulus 1, then there exist $y_1, \ldots, y_r$, independent numbers of modulus 1, functions $f_1, \ldots, f_m$, quotients of words, and torsion elements $x'_1, \ldots, x'_m$ so that $x_k = x'_k f_k(y_1, \ldots, y_r)$.*

Proof Let $z_1, \ldots, z_r$ be a maximal independent subset of $x_1, \ldots, x_m$ and abbreviate $\mathbf{z} = (z_1, \ldots, z_r)$. Note that there exist functions g_k, words in r variables and their inverses and positive integers n_k, such that $x_k^{n_k} = g_j(\mathbf{z})$. Let $n = n_1 n_2 \cdots n_m$ and choose $\mathbf{y} = (y_j)_j$ so that $y_j^n = z_j$. Now define $f_k = g_k^{n/n_k}$ and note that numbers $x'_k = x_k / f_k(\mathbf{y})$ are torsion. Indeed

$$(x'_k)^{n_k} = \frac{x_k^{n_k}}{f_k^{n_k}(\mathbf{y})} = \frac{x_k^{n_k}}{g_k^{n}(\mathbf{y})} = \frac{x_k^{n_k}}{g_k(\mathbf{z})} = \frac{x_k^{n_k}}{x_k^{n_k}} = 1. \qquad \blacksquare$$

Lemma 2.2 *Suppose that $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ are two n-tuples of complex numbers. Suppose furthermore that there exists an integer $m > 1$ for which $a_j^m = b_j^m = 1$ for all $1 \leq j \leq n$. If there does not exist a permutation π of $\{1, 2, \ldots, n\}$*

such that $a_j = b_{\pi(j)}$ *for all* $1 \le j \le n$, *then for some* $1 \le k \le (m-1)$ *we must have*

$$\left|\sum_{j=1}^{n}\left(a_j^k - b_j^k\right)\right| \ge \frac{m}{m-1}.$$

Proof Suppose, on the contrary, that for all $1 \le k < m$ we have

$$\left|\sum_{j=1}^{n}\left(a_j^k - b_j^k\right)\right| < \frac{m}{m-1}.$$

Since $(b_1, \ldots, b_n)$ is not just a permutation of $(a_1, \ldots, a_n)$, there exists some $1 \le i_0 \le n$ so that the term a_{i_0} appears more frequently in the sequence $(a_1, \ldots, a_n)$ than it does as a term in the sequence $(b_1, \ldots, b_n)$. Since our inequality is independent of permutations of the a_j's and the b_j's, we may assume without loss of generality that $i_0 = 1$ and *a fortiori* that

(i) $a_1 = a_2 = \cdots = a_{d_1}$ for some $1 \le d_1 \le n$,
(ii) $a_1 = b_1 = b_2 = \cdots = b_{d_2}$ for some $0 \le d_2 < d_1$, and
(iii) $b_j \ne a_1, j > d_2$.

Moreover, since $\sum_{j=1}^{n}(a_j^k - b_j^k) = \sum_{j=d_2+1}^{n}(a_j^k - b_j^k)$ for each $1 \le k$, we can in turn restrict our attention to $(a_{d_2+1}, \ldots, a_n)$ and $(b_{d_2+1}, \ldots, b_n)$. If we next divide these remaining a_j's and b_j's by a_{d_2+1} and relabel the index set to run from 1 to $N := n - d_2$, then we see that we have reduced the problem to the case where:

(a) $a_1 = 1$ and $1 \notin \{b_1, \ldots, b_N\}$;
(b) $a_j^m = 1 = b_j^m$, $1 \le j \le N$, and
(c) $|\sum_{j=1}^{N}(a_j^k - b_j^k)| < m/(m-1)$, $1 \le k < m$. (Clearly this also holds for $k = 0$.)

Also note that for each $1 \le j \le N$, we have

$$\sum_{k=0}^{m-1} b_j^k = 0$$

and

$$\sum_{k=0}^{m-1} a_j^k = \begin{cases} m; a_j = 1 \\ 0; a_j \ne 1. \end{cases}$$

Hence, if $r := d_1 - d_2 > 0$ is the number of 1's among a_j's, then we have

$$\sum_{j=1}^{N}\sum_{k=1}^{m-1}\left(a_j^k - b_j^k\right) = \sum_{j=1}^{N}\sum_{k=0}^{m-1}\left(a_j^k - b_j^k\right) = rm \ge m.$$

Now compute

$$m = \sum_{k=1}^{m-1}\left(\frac{m}{m-1}\right) > \sum_{k=1}^{m-1}\left|\sum_{j=1}^{N}\left(a_j^k - b_j^k\right)\right| \ge \left|\sum_{k=1}^{m-1}\sum_{j=1}^{N}\left(a_j^k - b_j^k\right)\right| = rm \ge m,$$

a contradiction. From this the desired conclusion follows. ■

THEOREM 2.3 *If A and B are unitary matrices such that*

$$|\mathrm{tr}A^l - \mathrm{tr}B^l| \leq 1$$

for all $l \in \mathbb{Z}$, then A and B are unitarily equivalent.

Proof Without any loss of generality we can assume that $A = \mathrm{diag}(a_1, \ldots, a_n)$, $B = \mathrm{diag}(b_1, \ldots, b_n)$ for some complex numbers a_l and b_l of modulus 1. Suppose that A and B are not unitarily equivalent. Through an argument similar to the one used in the previous lemma, we may reduce the problem to the case where $a_1 = 1$ and $b_l \neq 1$ for any l. We can then use Lemma 2.1 to find $c_1, \ldots, c_r$ independent numbers from the unit circle, α_l, β_l, quotients of words and torsion elements a'_l and b'_l such that $a_l = a'_l\alpha_l(\mathbf{c})$ and $b_l = b'_l\beta_l(\mathbf{c})$ (here we abbreviate $\mathbf{c} = (c_1, \ldots, c_r)$).

For $j = 1, \ldots, r$ define $d_j = e^{2\pi i/p_j}$, where p_j are primes to be chosen as follows. First choose p_1 from primes larger than any order of b'_l. When primes $p_1, \ldots, p_{j-1}$ have been chosen then choose p_j from primes that are larger than any of the orders of

$$b'_l\beta_l(d_1, \ldots, d_{j-1}, 1, \ldots, 1).$$

Primes p_j were chosen in this manner to ensure that $b''_l := b'_l\beta_l(\mathbf{d}) \neq 1$. Indeed, if j_0 is the largest integer such that the order of x_{j_0} in β_l is non-zero, then the order of b''_l must be divisible by p_{j_0}. Now define also $a''_l := a'_l\alpha_l(\mathbf{d})$, $A_1 = \mathrm{diag}(a''_l)$ and $B_1 = \mathrm{diag}(b''_l)$. Since the matrices A_1 and B_1 are clearly of the finite order, there exists an integer m_1 such that $A_1^{m_1} = 1 = B_1^{m_1}$. Since $a''_1 = 1$ and $b''_l \neq 1$ the sequences of a''_l's and b''_l's cannot be permutations of each other and hence by Lemma 2 there exists a positive integer $l_1 < m_1$ such that $|\mathrm{tr}A_1^{l_1} - \mathrm{tr}B_1^{l_1}| \geq m_1/(m_1 - 1)$.

Now simultaneous approximation yields an integer k so that $\mathbf{c}^k = (c_j^k)$ is so close to $\mathbf{d}$ that the numbers $|\alpha_l(\mathbf{c}^k)^{l_1} - \alpha_l(\mathbf{d})^{l_1}|$ and $|\beta_l(\mathbf{c}^k)^{l_1} - \beta_l(\mathbf{d})^{l_1}|$ are smaller than $(1/2n(m_1 - 1))$ (and hence $|\mathrm{tr}A^{l_1k} - \mathrm{tr}A_1^{l_1}| + |\mathrm{tr}B^{l_1k} - \mathrm{tr}B_1^{l_1}| < 1/(m_1 - 1)$). But then

$$\begin{aligned}\left|\mathrm{tr}A^{l_1k} - \mathrm{tr}B^{l_1k}\right| &\geq \left|\mathrm{tr}A_1^{l_1} - \mathrm{tr}B_1^{l_1}\right| - \left|\mathrm{tr}A^{l_1k} - \mathrm{tr}A_1^{l_1}\right| - \left|\mathrm{tr}B^{l_1k} - \mathrm{tr}B_1^{l_1}\right| \\ &> \frac{m_1}{m_1 - 1} - \frac{1}{m_1 - 1} = 1,\end{aligned}$$

a contradiction. ■

3. Results about groups

3.1 For arbitrary matrices $A, B \in \mathbb{M}_n(\mathbb{C})$, knowing that $\mathrm{tr}(w(A, A^*))$ is close to $\mathrm{tr}(w(B, B^*)$ for all words w does not tell us very much about A and B. For example, if we fix $\varepsilon > 0$ and choose A, B so that $\|A\|, \|B\| < \varepsilon/n$, then $\|w(A, A^*)\|, \|w(B, B^*)\| < \varepsilon/n$ for all words w, and so $|\mathrm{tr}(w(A, A^*)) - \mathrm{tr}(w(B, B^*))| < 2\varepsilon$. If we let $A_0 = I_n \oplus A$ and $B_0 = I_n \oplus B$ in $\mathbb{M}_{2n}(\mathbb{C})$, then $\|A_0\| = \|B_0\| = 1$ and yet the same trace inequality holds for A_0 and B_0, showing that it is not just a matter of the norms of the original matrices being too small.

One way to avoid this problem is to require that A and B be invertible, which is what we shall now do. In fact, we are able to obtain results about the simultaneous unitary equivalence of families of matrices whose traces remain (relatively) close.

LEMMA 3.2 *Let $r, s \geq 1$ be integers. Suppose that $0 < \mu < 1$ is fixed, $\omega_1, \omega_2, \ldots, \omega_r$ and $\nu_1, \nu_2, \ldots, \nu_s$ are complex numbers of modulus one and that there exists $k_0 \in \mathbb{N}$ so that $k \geq k_0$ implies*

$$\left| \sum_{i=1}^{r} \omega_i^k - \sum_{j=1}^{s} \nu_j^k \right| \leq \mu.$$

Then $s = r$ and there exists a permutation π of $\{1, 2, \ldots, r\}$ such that $\omega_i = \nu_{\pi(i)}$, $1 \leq i \leq r$.

Proof We may assume without loss of generality that $r \geq s$. Let $0 < \varepsilon < (1 - \mu)/2$. We can find $k_1 > k_0$ so that $|1 - \omega_i^{k_1}|1 < \varepsilon/r$ for all $1 \leq i \leq r$ and $|1 - \nu_j^{k_1}|1 < \varepsilon/s$ for all $1 \leq j \leq s$. Then

$$|r - s| \leq \left| r - \sum_{i=1}^{r} \omega_i^{k_1} \right| + \left| \sum_{i=1}^{r} \omega_i^{k_1} - \sum_{j=1}^{s} \nu_j^{k_1} \right| + \left| \sum_{j=1}^{s} \nu_j^{k_1} - s \right| \tag{1}$$

and

$$\leq \varepsilon + \mu + \varepsilon < 1. \tag{2}$$

Since r and s are integers, $r = s$.

The result now follows as an easy application of Theorem 2.3. For each $k \geq k_0$, let $A_k := \mathrm{diag}(\omega_1^k, \ldots, \omega_r^k)$, $B_k := \mathrm{diag}(\nu_1^k, \ldots, \nu_r^k)$. By Theorem 2.3, A_k and B_k are unitarily equivalent and as such they have the same eigenvalues appearing with equal multiplicities. Thus there exists a permutation π_k of $\{1, 2, \ldots, r\}$ so that $\omega_i^k = (\nu_{\pi_k(i)}^k)$, $1 \leq i \leq r$. Since there are infinitely many primes bigger than k_0, but only finitely many permutations of $\{1, 2, \ldots, r\}$, we can choose two distinct primes $p, q > k_0$ so that $\pi_p = \pi_q$. Then $\omega_i^p = \nu_{\pi_p(i)}^p$ and $\omega_i^q = \nu_{\pi_p(i)}^q$ with p and q relatively prime implies $\omega_i = \nu_{\pi_p}(i)$, $1 \leq i \leq r$, completing the proof. ■

LEMMA 3.3 *Suppose $m \geq 1$, $\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_m \in \mathbb{C} \backslash \{0\}$ and that*

$$\left| \sum_{i=1}^{m} \left(\alpha_i^k - \beta_i^k \right) \right| \leq 1 \quad \textit{for all } k \in \mathbb{Z}. \tag{3}$$

Then there exists a permutation π of $\{1, 2, \ldots, m\}$ so that $\alpha_i = \beta_{\pi(i)}$, $1 \leq i \leq m$.

Proof If $|\alpha_i| = 1 = |\beta_i|$ for all $1 \leq i \leq m$, then by setting $A := \mathrm{diag}(\alpha_1, \ldots, \alpha_m)$ and $B = \mathrm{diag}(\beta_1, \ldots, \beta_m)$, we see that $|\mathrm{tr} A^k - \mathrm{tr} B^k| \leq 1$ for all $k \in \mathbb{Z}$, and so by Theorem 2.3 again, A and B are unitarily equivalent. As before, this implies that they have the same eigenvalues appearing with the same multiplicities, from which the existence of π immediately follows.

Let us next assume that some $|\alpha_i| \neq 1$ or some $|\beta_i| \neq 1$.

Observe that in the statement of the lemma we can replace each α_i by α_i^{-1} if we also replace each β_i by β_i^{-1}. Using this fact, and switching the roles of α_i and β_i if necessary, it is not hard to see that without loss of generality, we may assume that

- $|\alpha_1| \geq |\alpha_2| \geq \cdots \geq |\alpha_m|$;
- $|\beta_1| \geq |\beta_2| \geq \cdots \geq |\beta_m|$;
- $|\alpha_1| > 1$; and
- $|\alpha_1| \geq |\beta_1|$.

Our argument will proceed by induction upon the number m of terms.

Step 1 $m = 1$.

Now $|\alpha_1| > 1$ and $|\alpha_1| \geq |\beta_1|$. If $\alpha_1 \neq \beta_1$, then $\lim_{k\to\infty} |\alpha_1^k - \beta_1^k| = \infty$. In particular, there exists $k \in \mathbb{N}$ so that $|\alpha_1^k - \beta_1^k| > 1$, contradicting our assumption. Thus $\alpha_1 = \beta_1$ in this case.

Step 2 $m > 1$.

Suppose that the result holds for $m' < m$. Now consider $\alpha_1, \alpha_2, \ldots, \alpha_m,\ \beta_1, \beta_2, \ldots, \beta_m$ satisfying the inequalities (7). Fix $1 \leq r, s \leq m$ maximal with respect to the conditions $|\alpha_1| = |\alpha_2| = \cdots = |\alpha_r|$, $|\beta_1| = |\beta_2| = \cdots = |\beta_s|$.

Set $\omega_j = \alpha_j/|\alpha_1|$ and $\nu_j = \beta_j/|\alpha_1|$, $1 \leq j \leq m$. Note that $|\omega_j| = 1$, $1 \leq j \leq r$, $|\omega_j| < 1$ if $j > r$, $|\nu_j| < 1$ if $j > s$, then

$$\left|\sum_{i=1}^{r} \omega_i^k - \sum_{j=1}^{s} \nu_j^k + \left[\sum_{i=r+1}^{m} \omega_i^k - \sum_{j=s+1}^{m} \nu_j^k\right]\right| \leq \frac{1}{|\alpha_1|^k}, \quad k \in \mathbb{Z}. \tag{4}$$

Suppose $|\nu_1| < 1$ (i.e., $|\beta_1| < |\alpha_1|$) and $\varepsilon < 1/2$. Then we can choose k_0 sufficiently large so that $k \geq k_0$ implies that

(i) $1/|\alpha_1|^k < \varepsilon/4$ and
(ii) $\sum_{i=r+1}^{m} |\omega_i|^k + \sum_{j=1}^{m} |\nu_j|^k < \varepsilon/4$.

Moreover, since each $|\omega_i| = 1$, $1 \leq i \leq r$, we can find $k_1 \geq k_0$ so that $|1 - \omega_i^{k_1}| < \varepsilon/(4r)$, $1 \leq i \leq r$. From equation (4) we deduce that

$$r \leq \left|r - \sum_{i=1}^{r} \omega_i^{k_1}\right| + \sum_{i=r+1}^{m} |\omega_i|^{k_1} + \sum_{j=1}^{m} |\nu_j|^{k_1} + \left|\sum_{i=1}^{m} \omega_i^{k_1} - \sum_{j=1}^{m} \nu_j^{k_1}\right| \tag{5}$$

$$\leq \varepsilon/4 + \varepsilon/4 + \varepsilon/4 = 3\varepsilon/4 < 1, \tag{6}$$

a contradiction since $r \geq 1$.

It follows therefore that $|\nu_1| = 1$, whence $|\omega_1| = \cdots = |\omega_r| = 1 = |\nu_1| = \cdots = |\nu_s|$. Since $|\omega_i| < 1$ if $i > r$ and $|\nu_j| < 1$ if $j > s$, we can find an integer $k_2 > 0$ so that $k \geq k_2$ implies

(a) $1/|\alpha_1|^k < \varepsilon/4$ and
(b) $\sum_{i=r+1}^{m} |\omega_i|^k + \sum_{j=s+1}^{m} |\nu_j|^k < \varepsilon/4$.

From equation (4) we see that for $k \geq k_2$,

$$\left| \sum_{i=1}^{r} \omega_i^k - \sum_{j=1}^{s} \nu_j^k \right| \leq \frac{1}{|\alpha_1|^k} + \sum_{i=r+1}^{m} |\omega_i|^k + \sum_{j=s+1}^{m} |\nu_j|^k \tag{7}$$

$$< \varepsilon/4 + \varepsilon/4 < 1/2. \tag{8}$$

By Lemma 3.2, $s = r$ and there exists a permutation π_0 of $\{1, 2, \ldots, r\}$ so that $\omega_i = \nu_{\pi_0(i)}$, $1 \leq i \leq r$. It follows that $\alpha_i = \beta_{\pi_0(i)}$, $1 \leq i \leq r$. But then equation (3) holds for $\alpha_{r+1}, \ldots, \alpha_m$, $\beta_{r+1}, \ldots, \beta_m$ in that for all $k \in \mathbb{Z}$,

$$\left| \sum_{i=r+1}^{m} \left(\alpha_i^k - \beta_i^k\right) \right| = \left| \sum_{i=1}^{r} \left(\alpha_i^k - \beta_{\pi_0(i)}^k\right) + \sum_{i=r+1}^{m} \left(\alpha_i^k - \beta_i^k\right) \right| \tag{9}$$

$$= \left| \sum_{i=1}^{m} \left(\alpha_i^k - \beta_i^k\right) \right| \leq 1. \tag{10}$$

Since $m' := m - r < m$, we may apply our induction hypothesis to obtain a permutation π_1 of $\{r+1, r+2, \ldots, m\}$ so that $\alpha_i = \beta_{\pi_1(i)}$, $r+1 \leq i \leq m$. This clearly establishes our claim. ■

As a simple consequence of the above lemma, we obtain the following.

PROPOSITION 3.4 *Suppose that $A, B \in \mathbb{M}_n(\mathbb{C})$ are two invertible matrices and that*

$$|\mathrm{tr}(A^k) - \mathrm{tr}(B^k)| \leq 1$$

for all $k \in \mathbb{Z}$. Then $\sigma(A) = \sigma(B)$, including multiplicities.

Proof We can (without loss of generality) assume that both $A = [a_{ij}]$ and $B = [b_{ij}]$ are in upper triangular form. Let $\alpha_i = a_{ii}$, $\beta_i = b_{ii}$, $1 \leq i \leq n$. Since A, B are invertible, $\alpha_i \neq 0 \neq \beta_i$ for all $1 \leq i \leq n$. Our trace condition implies that

$$\left| \sum_{i=1}^{n} \left(\alpha_i^k - \beta_i^k\right) \right| = \left| \mathrm{tr}(A^k) - \mathrm{tr}(B^k) \right|$$

$$\leq 1 \quad \text{for all } k \in \mathbb{Z}.$$

By Lemma 3.3, there exists a permutation π of $\{1, 2, \ldots, n\}$ such that $\alpha_i = \beta_{\pi(i)}$, $1 \leq i \leq n$. ■

THEOREM 3.5 *Suppose that $A, B \in \mathbb{M}_n(\mathbb{C})$ are two invertible matrices and that for all words w in two non-commuting variables we have:*

$$|\mathrm{tr}(w(A, A^*)) - \mathrm{tr}(w(B, B^*))| \leq 1,$$

and

$$|\mathrm{tr}(w(A, A^*)^{-1}) - \mathrm{tr}(w(B, B^*)^{-1})| \leq 1.$$

Then A is unitarily equivalent to B.

Proof Let w denote an arbitrary (but temporarily fixed) word in two non-commuting variables. Let $A_0 = w(A, A^*)$ and $B_0 = w(B, B^*)$. Note that any word in A_0 (resp. B_0) is just another word in A and A^* (resp. B and B^*). The conditions in the statement of the theorem therefore imply that

$$|\mathrm{tr}(A_0^k) - \mathrm{tr}(B_0)^k| \leq 1$$

for all $k \in \mathbb{Z}$, and so by Proposition 3.4, $\sigma(A_0) = \sigma(B_0)$, including multiplicities. But then $\mathrm{tr}(w(A, A^*)) = \mathrm{tr}(A_0) = \mathrm{tr}(B_0) = \mathrm{tr}(w(B, B^*))$. Since w was arbitrary, Specht's theorem implies that A and B are unitarily equivalent. ■

3.6 In [2], a semigroup of operators $\mathfrak{G} \subseteq \mathbb{M}_n(\mathbb{C})$ was defined to be *semisimple* if its linear span forms a semisimple algebra. We extend this definition slightly, namely: we shall say that a non-empty subset $\mathfrak{A} \subseteq \mathbb{M}_n(\mathbb{C})$ is *semisimple* if the algebra $\mathrm{Alg}\,\mathfrak{A}$ generated by $\mathfrak{A}$ is semisimple. When $\mathfrak{A}$ is a semigroup, these two notions coincide. Also, if $\mathfrak{A}$ is an algebra to begin with, then all definitions of semisimplicity are consistent.

We say that a family $\mathfrak{A} \subseteq \mathbb{M}_n(\mathbb{C})$ is *self-adjoint* if $T \in \mathfrak{A}$ implies that $T^* \in \mathfrak{A}$. It is readily verified that any self-adjoint family $\mathfrak{A}$ is semisimple in the above sense.

We next recall a theorem of Hladnik, Omladič, and the third author of the present work which we need. We do not state that theorem in its full generality, but rather only in the context we require.

Theorem 3.7 [2] *Suppose that $\mathfrak{G}$ and $\mathfrak{H}$ are two semisimple semigroups of invertible $n \times n$ matrices. If $\varphi\colon \mathfrak{G} \to \mathfrak{H}$ is a surjective, trace-preserving semigroup homomorphism, then there exists an invertible operator $R \in \mathbb{M}_n(\mathbb{C})$ so that*

$$\varphi(A) = R^{-1}AR \quad \textit{for all } A \in \mathfrak{G}.$$

Let us write Ad_R to denote the map $X \mapsto R^{-1}XR$. The domain of this map will be clear from the context.

Theorem 3.8 *Let $\mathfrak{G}$, $\mathfrak{H} \subseteq \mathbb{M}_n(\mathbb{C})$ be two semisimple groups of invertible matrices. If $\varphi : \mathfrak{G} \to \mathfrak{H}$ is a surjective homomorphism, and if*

$$\left|\mathrm{tr}(\varphi(A)) - \mathrm{tr}(A)\right| \leq 1 \quad \textit{for all } A \in \mathfrak{G},$$

then $\varphi = \mathrm{Ad}_R$ for some invertible operator $R \in \mathbb{M}_n(\mathbb{C})$.

Proof Fix $A \in \mathfrak{G}$ and set $B = \varphi(A)$. Since the trace condition holds for all members of the group $\mathfrak{G}$, we have

$$\begin{aligned}\left|\mathrm{tr}(B^k) - \mathrm{tr}(A^k)\right| &= \left|\mathrm{tr}(\varphi(A^k)) - \mathrm{tr}(A^k)\right| \\ &\leq 1 \quad \text{for all } k \in \mathbb{Z}.\end{aligned}$$

By Proposition 3.4, $\sigma(A) = \sigma(B)$ including multiplicities. But then $\mathrm{tr}(B) = \mathrm{tr}(\varphi(A)) = \mathrm{tr}(A)$. Since $A \in \mathfrak{G}$ was arbitrary, φ is a trace preserving surjective homomorphism among the semisimple groups of $\mathbb{M}_n(\mathbb{C})$. It follows from Theorem 3.7 above that $\varphi = \mathrm{Ad}_R$ for some invertible operator $R \in \mathbb{M}_n(\mathbb{C})$. ■

Recall that if $A \in \mathbb{M}_n(\mathbb{C})$, then the *absolute value* of A is the element $|A| = (A^*A)^{1/2}$.

Corollary 3.9 *If $\mathfrak{G}$, $\mathfrak{H}$ are self-adjoint subgroups of the invertible group of $\mathbb{M}_n(\mathbb{C})$ and $\varphi : \mathfrak{G} \to \mathfrak{H}$ is a surjective *-homomorphism satisfying*

$$\left|\operatorname{tr}(\varphi(A)) - \operatorname{tr}(A)\right| \le 1 \quad \textit{for all } A \in \mathfrak{G},$$

then $\varphi = \mathrm{Ad}_U$ for some unitary operator $U \in \mathbb{M}_n(\mathbb{C})$.

Proof By Theorem 3.8, $\varphi = \mathrm{Ad}_R$ for some invertible operator R. A standard argument shows that if a *-homomorphism is implemented by a similarity, then it is implemented by the unitary part of the polar decomposition of that similarity.

We include the argument for completeness: for $A \in \mathfrak{G}$, $\varphi(A) = R^{-1}AR$, while $R^{-1}A^*R = \varphi(A^*) = \varphi(A)^* = (R^{-1}AR)^* = R^*A^*(R^{-1})^*$. Thus $A^*(RR^*) = (RR^*)A^*$ for all $A \in \mathfrak{G}$, whence $(RR^*)A = A(RR^*)$ for all $A \in \mathfrak{G}$. But then $|R^*|A = A|R^*|$ for all $A \in \mathfrak{G}$. Write the polar decomposition $R^* = U|R^*|$ where U is unitary. Then $R = |R^*|U^*$ and $R^{-1} = U|R^*|^{-1}$, and so

$$\begin{aligned}\varphi(A) = R^{-1}AR &= U|R^*|^{-1}A|R^*|U^* \\ &= U|R^*|^{-1}\,|R^*|AU^* \\ &= UAU^*\end{aligned}$$

for all $A \in \mathfrak{G}$. ■

We can now rephrase some of these results as multivariable versions of Specht's theorem.

Corollary 3.10 *Suppose that $\mathcal{A} = \{A_\alpha\}_{\alpha\in\Lambda}$ and $\mathcal{B} = \{B_\alpha\}_{\alpha\in\Lambda}$ are two semisimple, inverse-closed families of invertible operators in $\mathbb{M}_n(\mathbb{C})$. If*

$$\left|\operatorname{tr}(w(\mathcal{A})) - \operatorname{tr}(w(\mathcal{B}))\right| \le 1$$

for all finite words w in $|\mathcal{A}|$ non-commuting variables, then there exists $R \in \mathbb{M}_n(\mathbb{C})$ invertible such that

$$A_\alpha = R^{-1}B_\alpha R \quad \textit{for all } \alpha \in \Lambda.$$

Note We first recall that the semisimplicity of $\mathcal{B}$ implies that $\mathrm{alg}(\mathcal{B})$ is similar to a C^*-algebra. It is readily verified that there is no loss of generality in invoking that similarity at the outset and assuming *a priori* that $\mathrm{alg}(\mathcal{B})$ is a C^*-algebra, as we shall proceed below.

Proof Fix q, a finite word in $|\mathcal{A}|$ variables, and let $U_q := q(\mathcal{A})$, $V_q := q(\mathcal{B})$. Note that U_q and V_q are invertible operators. If $k \in \mathbb{Z}$, then U_q^k and V_q^k represent the same words in $\mathcal{A}$ and $\mathcal{B}$ respectively. By our hypothesis,

$$\left|\operatorname{tr}(U_q^k) - \operatorname{tr}(V_q^k)\right| \le 1 \quad \text{for all } k \in \mathbb{Z}.$$

From Proposition 3.4, $\sigma(U_q) = \sigma(V_q)$, including multiplicities and so $\mathrm{tr}(U_q) = \mathrm{tr}(V_q)$.

Let $\mathcal{S}_{\mathcal{A}}$ (resp. $\mathcal{S}_{\mathcal{B}}$) denote the multiplicative semigroup generated by $\mathcal{A}$ (resp. by $\mathcal{B}$). Let

$$\begin{aligned} \varphi : \mathcal{S}_{\mathcal{A}} &\to \mathcal{S}_{\mathcal{B}} \\ r(\mathcal{A}) &\mapsto r(\mathcal{B}) \end{aligned}$$

where r is an arbitrary word in $|\mathcal{A}|$ variables. We claim that φ is well defined, when it is a semigroup homomorphism.

Indeed, suppose that $r_1(\mathcal{A}) = r_2(\mathcal{A})$ for words r_1, r_2. Then $r_3 := r_1 r_2^{-1}$ is simply another word in $|\mathcal{A}|$-variables, and $r_3(\mathcal{A}) = I$. As such, if w is any other word, then the argument of the first paragraph shows that

$$\sigma(w(\mathcal{B})) = \sigma(w(\mathcal{A})) = \sigma(r_3(\mathcal{A})w(\mathcal{A})) = \sigma(r_3(\mathcal{B})w(\mathcal{B})),$$

including multiplicities. In particular, therefore,

$$\mathrm{tr}(w(\mathcal{B})) = \mathrm{tr}(w(\mathcal{A})) = \mathrm{tr}(r_3(\mathcal{A})w(\mathcal{A})) = \mathrm{tr}(r_3(\mathcal{B})w(\mathcal{B}))$$

for all words w. By linearity, it follows that

$$\mathrm{tr}((r_3(\mathcal{B}) - I)Q) = 0$$

for all $Q \in \mathrm{span}\, \mathcal{S}_{\mathcal{B}} = \mathrm{alg}(\mathcal{B})$. But $\mathcal{B}$ is semisimple, and so as pointed out above, we may assume that $\mathrm{alg}(\mathcal{B})$ is a C^*-algebra. But then $(r_3(\mathcal{B}) - I) \in \mathrm{alg}(\mathcal{B})$ implies that $(r_3(\mathcal{B}) - I)^* \in \mathrm{alg}(\mathcal{B})$ and therefore that

$$\mathrm{tr}((r_3(\mathcal{B}) - I)(r_3(\mathcal{B}) - I)^*) = 0.$$

Since the trace is faithful on $\mathbb{M}_n(\mathbb{C})$, it follows that $r_3(\mathcal{B}) - I = 0$, or that $r_3(\mathcal{B}) = I$. From this we get $r_1(\mathcal{B}) = r_2(\mathcal{B})$. In particular, φ is a well-defined semigroup homomorphism.

We are now in a position to apply Theorem 3.8 to conclude that $\varphi = \mathrm{Ad}_R$ for some invertible matrix $R \in \mathbb{M}_n(\mathbb{C})$, from which the result is easily obtained. ■

Corollary 3.11 *Suppose that $\mathcal{A} = \{A_\alpha\}_{\alpha \in \Lambda}$ and $\mathcal{B} = \{B_\alpha\}_{\alpha \in \Lambda}$ are two self-adjoint, inverse-closed families of invertible operators in $\mathbb{M}_n(\mathbb{C})$. If*

$$\left| \mathrm{tr}(w(\mathcal{A})) - \mathrm{tr}(w(\mathcal{B})) \right| \leq 1$$

for all finite words w in $|\mathcal{A}|$ non-commuting variables, then there exists $Z \in \mathbb{M}_n(\mathbb{C})$ unitary such that

$$A_\alpha = Z^{-1} B_\alpha Z \quad \textit{for all } \alpha \in \Lambda.$$

Proof Note that $\mathcal{A}$ and $\mathcal{B}$ self-adjoint automatically implies that these families are semisimple. By Corollary 3.10, we can find $R \in \mathbb{M}_n(\mathbb{C})$ an invertible operator so that $A_\alpha = R^{-1} B_\alpha R$ for all $\alpha \in \Lambda$. As in the proof of Corollary 3.9, we find that the self-adjointness of the families $\mathcal{A}$ and $\mathcal{B}$ implies that the unitary part Z of the polar decomposition of R implements the simultaneous unitary equivalence of $\mathcal{A}$ and $\mathcal{B}$. ∎

4. The projection condition

Recall that for a matrix $A \in \mathbb{M}_n(\mathbb{C})$, $|A|$ denotes the positive square root of A^*A. Also, $C^*(A)$ denotes the C^*-algebra generated by A, that is, the smallest norm-closed, unitary self-adjoint subalgebra of $\mathbb{M}_n(\mathbb{C})$ which contains A. If we use $\mathbb{C}[X, Y]$ to denote the set of polynomials in two non-commuting variables X and Y with complex coefficients, then, in the finite-dimensional setting, $C^*(A)$ is easily seen to coincide with the set $\{p(A, A^*) : p \in \mathbb{C}[X, Y]\}$.

Definition 4.1 Let $A, B \in \mathbb{M}_n(\mathbb{C})$.

We shall say that A and B satisfy the 'projection condition' (we abbreviate this to the PC) if, for any polynomial $p \in \mathbb{C}[X, Y]$ in two non-commuting variables x and y for which $p(A, A^*)$ is a projection, it follows that $p(B, B^*)$ is a projection of the same trace.

We shall say that A and B satisfy the 'absolute value condition' (we abbreviate this to the AVC) if, for any polynomial $p \in \mathbb{C}[X, Y]$ in two non-commuting variables x and y, $|p(A, A^*)|$ is unitarily equivalent to $|p(B, B^*)|$.

It is worth making a few observations. First we remark that there is an apparent asymmetry in our definition of the projection condition. However, as the next proposition demonstrates, the PC and the AVC are equivalent for pairs A and B of $n \times n$ matrices. Since the AVC is easily seen to be a symmetric relation, it follows that the PC is also symmetric. Secondly, it is clear that the trace condition in the definition of the projection condition can be replaced with the condition that $p(A, A^*)$ and $p(B, B^*)$ are projections of equal rank, or are unitarily equivalent projections. Finally, if $\mathrm{tr}(w(A, A^*)) = \mathrm{tr}(w(B, B^*))$ for all words w in two non-commuting variables, then by Specht's theorem, A is unitarily equivalent to B and so A and B satisfy the PC.

Our goal in this section is to prove the converse of this result, namely: if A and B satisfy the projection condition, then they are unitarily equivalent.

Proposition 4.2 *Suppose $A, B \in \mathbb{M}_n(\mathbb{C})$. The following are equivalent*:

(i) *A and B satisfy the PC and*
(ii) *A and B satisfy the AVC.*

Proof Suppose first that they satisfy the AVC. Let $p \in \mathbb{C}[X, Y]$ and suppose $P := p(A, A^*)$ is a projection. Let $Q := p(B, B^*)$. Then $0 = |PP^* - P^*P|$ implies that $0 = |QQ^* - Q^*Q|$, and hence Q is normal. Also, $0 = |P^2 - P|$ implies $0 = |Q^2 - Q|$, and so Q is a projection. But then the AVC implies that $P = |P| \simeq |Q| = Q$, and so P and Q are clearly projections with the same trace. Thus A and B satisfy the PC.

Suppose next that A and B satisfy the PC. Let $p \in \mathbb{C}[X, Y]$ be any polynomial in two non-commuting variables. Set $P_A = p(A, A^*)^* p(A, A^*)$ and $P_B = p(B, B^*)^* p(B, B^*)$. It suffices to prove that P_A is unitarily equivalent to P_B.

Now P_A and P_B are positive matrices, and so we can find distinct non-negative real numbers $a_1, a_2, \dots, a_{\kappa_A}$ and distinct non-negative real numbers $b_1, b_2, \dots, b_{\kappa_B}$ so that $P_A \simeq \oplus_{j=1}^{\kappa_A} a_j I_{l_j}$ and $P_B \simeq \oplus_{i=1}^{\kappa_B} b_i I_{m_i}$ for some $l_j, m_i \geq 1$ satisfying $\sum_{j=1}^{\kappa_A} l_j = n = \sum_{i=1}^{\kappa_B} m_i$.

Suppose that there exists $1 \leq i \leq \kappa_B$ so that $b_i \notin \{a_1, \dots, a_{\kappa_A}\}$. Without loss of generality, we may assume that $b_1 \notin \{a_1, \dots, a_{\kappa_A}\}$. For each $1 \leq j \leq \kappa_A$, consider the polynomials

$$q_j(z) = \left(\prod_{1 \leq r \neq j \leq \kappa_A} \frac{z - a_r}{a_j - a_r} \right),$$

and $q_j'(z) = q_j(z)((z - b_1)/(a_j - b_1))$.

Then $q_j'(a_j) = 1, q_j'(a_r) = 0, 1 \leq r \neq j \leq \kappa_A$, and $q_j'(b_1) = 0$.

As such, $\sum_{j=1}^{\kappa_A} q_j'(P_A) = I$, and clearly $\sum_{j=1}^{\kappa_A} q_j'(P_A)$ is a polynomial in A and A^*. It follows from the projection condition that $\sum_{j=1}^{\kappa_A} q_j'(P_B) = I$. But,

$$\sum_{j=1}^{\kappa_A} q_j'(P_B) \simeq \oplus_{i=1}^{\kappa_B} \left(\sum_{j=1}^{\kappa_A} q_j'(b_i) \right) I_{m_i}.$$

Since $\sum_{j=1}^{\kappa_A} q_j'(b_1) = 0$, we get that $\sum_{j=1}^{\kappa_A} q_j'(P_B) \neq I$, a contradiction. From this, we conclude that $\sigma(P_B) \subseteq \sigma(P_A)$, i.e.; $b_i = a_{j(i)}$ for some $1 \leq j(i) \leq \kappa_A$, $1 \leq i \leq \kappa_B$. Recalling that the b_i's are distinct, we see that $a_{j(i)} \neq a_{j(i')}$ if $i \neq i'$.

Next, for each $1 \leq j \leq \kappa_A$, $q_j(P_A) \simeq I_{l_j}$ and hence $q_j(P_B)$ must be a projection of the same rank. But

$$q_j(P_B) \simeq \oplus_{i=1}^{\kappa_B} q_j(b_i) I_{m_i} \simeq \oplus_{i=1}^{\kappa_B} q_j(a_{j(i)}) I_{m_i}.$$

Thus there exists a unique i_0 so that $b_{i_0} = a_{j(i_0)} = a_j$, and $m_{i_0} = l_j$. It follows that the multiplicity of a_j as an eigenvalue of P_A is the same as its multiplicity as an eigenvalue of P_B. Hence P_A is unitarily equivalent to P_B, and so, as stated, A and B satisfy the absolute value condition. ■

Lemma 4.3 *Suppose $A, B \in \mathbb{M}_n(\mathbb{C})$ satisfy the PC. Then A is similar to B.*

Proof By Proposition 4.2, A and B satisfy the AVC as well. Given $T \in \mathbb{M}_n(\mathbb{C})$, a complete set of similarity invariants for T is given by $\{\text{nul}\,(T - \lambda I)^k : \lambda \in \mathbb{C}, 1 \leq k \leq n\}$. Since $\ker(T - \lambda I)^k = \ker|(T - \lambda I)^k|$, and since $|(A - \lambda I)^k|$ is unitarily equivalent to $|(B - \lambda I)^k|$ for each $\lambda \in \mathbb{C}$ and $1 \leq k \leq n$, we see that A and B share the same similarity invariants, and hence the same Jordan form. In particular, A is similar to B. ■

As an immediate consequence, we observe that if A and B satisfy the absolute value condition (or equivalently the projection condition), then A and B have the same spectrum occurring with the same multiplicities.

LEMMA 4.4 *Suppose* $A, B \in \mathbb{M}_n(\mathbb{C})$ *satisfy the PC. Let* $P_1, P_2, \ldots, P_m$ *denote the minimal central projections of* $C^*(A)$. *Choose polynomials* $p_1, p_2, \ldots, p_m \in \mathbb{C}[X, Y]$ *so that* $P_i = p_i(A, A^*)$ *for each* $1 \le i \le m$. *Then* $Q_i := p_i(B, B^*)$ *are the minimal central projections of* $C^*(B)$. *Moreover,* P_i *is unitarily equivalent to* Q_i *for each* $1 \le i \le m$.

Proof By definition of the PC, Q_i is a projection of the same rank as P_i for each i. Moreover, $P_iP_j = p_i(A, A^*)p_j(A, A^*) = \delta_{i,j}P_i$ (where $\delta_{i,j}$ denotes the Kronecker delta) and hence $Q_iQ_j \simeq P_iP_j = 0$ if $i \ne j$. That is, the Q_i's form a family of pairwise orthogonal projections.

Let $r \in \mathbb{C}[X, Y]$ be any polynomial. If we set $R = r(A, A^*)$ and $S = r(B, B^*)$, then, since A and B satisfy the AVC as well, $|P_iR - RP_i| = 0$, which implies $|Q_iS - SQ_i| = 0$, and so we see that Q_i's are central in $C^*(B)$.

By symmetry, the minimality of the P_i's as central projections for $C^*(A)$ implies that the Q_i's are central projections for $C^*(B)$. ■

LEMMA 4.5 *Suppose* $A, B \in \mathbb{M}_n(\mathbb{C})$ *satisfy the PC, and that* $C^*(A)$ *contains no central projections other than* 0 *and I. Then the same holds for* $C^*(B)$, *and furthermore, A is unitarily equivalent to B.*

Proof By Lemma 4.2, we may assume that A and B satisfy the AVC as well. By Lemma 4.4, $C^*(B)$ has no proper central projections. We consider $R_{1,1}, R_{2,2}, \ldots, R_{k,k}$, a maximal set of minimal projections in $C^*(A)$, and choose polynomials $r_{j,j} \in \mathbb{C}[X, Y]$ so that $R_{j,j} = r_{j,j}(A, A^*)$. Then $T_{j,j} := r_{j,j}(B, B^*)$ is a projection of the same (constant) rank m in $C^*(B)$, by the projection condition. By symmetry, any proper subprojection of the $T_{j,j}$'s would be carried to a proper subprojection of the $R_{j,j}$'s, and so the minimality of the $R_{j,j}$'s implies that of the $T_{j,j}$'s. Without loss of generality, we may assume that $R_{j,j} = T_{j,j}$ for each $1 \le j \le k$.

For $i = 1, 2, \ldots, k-1$, fix a polynomial $r_{i,i+1} \in \mathbb{C}[X, Y]$ so that $R_{i,i+1}(A, A^*)$ satisfies $R_{i,i}\, R_{i,i+1}\, R_{i+1,i+1} = R_{i,i+1}$ and $R_{i,i+1}\, R^*_{i,i+1} = R_{i,i}$. Let $T_{i,i+1} := r_{i,i+1}(B, B^*)$. Then $|T_{i,i}\, T_{i,t+1}\, T_{i+1,i+1} - T_{i,i+1}|$ is unitarily equivalent to $|R_{i,i}\, R_{i,t+1}\, R_{i+1,i+1} - R_{i,i+1}| = 0$, and so $T_{i,i}\, T_{i,t+1}\, T_{i+1,i+1} = T_{i,i+1} = R_{i,i}\, T_{i,i+1}\, R_{i+1,i+1}$. Since $|T_{i,i+1}|$ is unitarily equivalent to $R_{i,i+1}$, again, without loss of generality we may assume that $T_{i,i+1} = R_{i,i+1}$ for all $1 \le i \le k-1$. For $i < j$, define $R_{i,j} = R_{i,i+1}\, R_{i+1,i+2} \cdots R_{j-1,j}$ and for $j < i$, define $R_{i,j} = R^*_{j,i}$. To complete the proof, we shall show that $R_{i,i}AR_{j,j} = R_{i,i}BR_{j,j}$ for all i and j.

Consider, for $1 \le j \le k$, $X_{i,j} = R_{j,i}(R_{i,i}AR_{j,j}) = a_{i,j}R_{j,j}$. Then $Y_{i,j} = T_{j,i}(T_{i,i}BT_{j,j}) = b_{i,j}T_{j,j} = b_{i,j}R_{j,j}$. But $X_{i,j}$ and $Y_{i,j}$ satisfy the AVC – or equivalently the PC – and so by Lemma 4.3, $X_{i,j}$ is similar to $Y_{i,j}$. Hence their spectra agree, which means that $a_{i,j} = b_{i,j}$ for each pair i and j. That is, $A = B$. ■

THEOREM 4.6 *Suppose* $A, B \in \mathbb{M}_n(\mathbb{C})$ *satisfy the projection condition. Then A is unitarily equivalent to B.*

Proof By Lemma 4.4, we may choose a set of minimal central projections $P_1, P_2, \ldots, P_m$ for $C^*(A)$ and $Q_1, Q_2, \ldots, Q_m$ for $C^*(B)$ such that P_i is unitarily equivalent to Q_i for each i. Since $P_iP_j - P_jP_i = 0$, and hence is a projection, it follows that $Q_iQ_j - Q_jQ_i$ is again a projection with the same trace – namely zero. Thus $Q_iQ_j = Q_jQ_i$. Similarly, $I - \sum_{i=1}^{m} P_i = 0$ is the zero projection, and hence so is $I - \sum_{i=1}^{m} Q_i$. It follows that there must exist a unitary $U \in \mathbb{M}_n(\mathbb{C})$ such that $U^* P_i U = Q_i$, $1 \le i \le m$. As such, without loss of generality, we may assume that $P_i = Q_i$, $1 \le i \le m$.

Now $P_i A P_i$ and $P_i B P_i$ satisfy the PC, and so $A_i := P_i A P_i|_{P_i \mathbb{C}^n}$ and $B_i = P_i B P_i|_{P_i \mathbb{C}^n}$ are also readily seen to satisfy the PC.

But $C^*(A_i)$ and $C^*(B_i)$ contain no proper central projections, and so by Lemma 4.5, A_i is unitarily equivalent to B_i, $1 \le i \le k$, say $V_i^* A_i V_i = B_i$ for some V_i unitary in $\mathcal{B}(P_i \mathbb{C}^n)$.

Letting $V = \oplus_{i=1}^{m} V_i$, $V^* A V = B$ and we are done. ■

Corollary 4.7 *Suppose $A, B \in \mathbb{M}_n(\mathbb{C})$, and that*

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

for all polynomials p in two non-commuting variables. Then A is unitarily equivalent to B.

Proof Fix a polynomial p in two non-commuting variables. Let $H_p = |p(A, A^*)|^2$ and $K_p = |p(B, B^*)|^2$. Since any word in H_p and its adjoint is really just a power of H_p, it is easily seen that $H_p^k = q_k(A, A^*)^* q_k(A, A^*)$ for some polynomial q and that $K_p^k = q_k(B, B^*)^* q_k(B, B^*)$.

In particular, $H_p^k = |H_p^k|$ and $K_p^k = |K_p^k|$. Our assumption therefore implies that $\operatorname{tr} H_p^k = \operatorname{tr} K_p^k$ for all $k \ge 1$, whence $\operatorname{tr} w(H_p, H_p^*) = \operatorname{tr} w(K_p, K_p^*)$ for all words w in two variables. It follows from Specht's theorem that H_p and K_p are unitarily equivalent. But then $H_p^{1/2} = |p(A, A^*)|$ is unitarily equivalent to $K_p^{1/2} = |p(B, B^*)|$. That is, A and B satisfy the AVC.

By Theorems 4.2 and 4.6, A and B are unitarily equivalent. ■

Example 4.8 We point out that in Corollary 4.7, it is not sufficient to consider absolute values of words (as opposed to absolute values of polynomials) in A and A^*. For example, if $A = I$ and U is any unitary other than I in $\mathbb{M}_n(\mathbb{C})$, then for all words w, $|w(A, A^*)| = |w(B, B^*)| = I$, and so their traces agree, despite the fact that A and U are not unitarily equivalent.

On the other hand, consideration of dimensions of kernels as in Lemma 4.3 shows that if we begin with two nilpotent matrices A and B, then the unitary equivalence of absolute values of words in A and A^* with the corresponding words in B and B^* implies the similarity of A and B.

Closing Remarks 4.9 We have been asked by C.-K. Li, and also by the referee, whether or not the results of this work can be extended to real matrices, this is indeed the case that follows from the following arguments.

It is relatively well known that if two matrices $A, B \in \mathbb{M}_n(\mathbb{R})$ are similar when regarded as elements of $\mathbb{M}_n(\mathbb{C})$ (i.e., there exists $S \in \mathbb{M}_n(\mathbb{C})$ invertible such that $S^{-1} A S = B$), then there exists $R \in \mathbb{M}_n(\mathbb{R})$ invertible such that $R^{-1} A R = B$. Since the argument is short, we include it here for completeness.

Let $S = [s_{ij}]$, $S_1 = [\operatorname{Re}(s_{ij})]$, and $S_2 = [\operatorname{Im}(s_{ij})]$, so that $S = S_1 + i S_2$, and $S_1, S_2 \in \mathbb{M}_n(\mathbb{R})$. The equation $S^{-1} A S = B$ implies that $A(S_1 + i S_2) = (S_1 + i S_2) B$. By comparing the real and complex components of these matrices, we find that $A S_1 = S_1 B$ and $A S_2 = S_2 B$. Thus, for $z \in \mathbb{C}$, we have $A(S_1 + z S_2) = (S_1 + z S_2) B$. Observe that the complex-valued function $p(z) := \det(S_1 + z S_2), z \in \mathbb{C}$ is a polynomial of degree at most n. Furthermore, since $p(i) = \det(S) \neq 0$, p is a non-trivial polynomial, and hence has at most n roots. From this it immediately follows that p has at most n real roots, and therefore for most $x \in \mathbb{R}$, $S_1 + xS_2$ is invertible. Choosing such an x_0, we see that $R = S_1 + x_0 S_2$ is an invertible element of $\mathbb{M}_n(\mathbb{R})$ such that $R^{-1}AR = B$.

Using this, one can obtain version (for matrices over $\mathbb{R}$) of Theorems 3.7 and 3.8, and Corollary 3.10.

The results we have for simultaneous unitary equivalence deal with self-adjoint families of matrices. These may also be extended to matrices over $\mathbb{R}$ as follows.

Suppose that $A, B \in \mathbb{M}_n(\mathbb{R})$, and suppose also that there exists $U \in \mathbb{M}_n(\mathbb{C})$ so that $U^*AU = B$. Then $U^*A^*U = B^*$ is immediate. (Here, of course, A^* agrees with the transpose of A.) Then, treating U as a similarity, it follows from the the above paragraphs that there exists $S \in \mathbb{M}_n(\mathbb{R})$ so that $S^{-1}AS = B$ and $S^{-1}A^*S = B^*$.

Thus $AS = SB$ and $A^*S = SB^*$, from which we also obtain $S^*A = BS^*$. Thus $S^*SB = S^*(AS) = BS^*S$, so that B commutes with S^*S. Since $|S|$ is a limit of polynomials in S^*S, we may conclude that $B|S| = |S|B$. Let $S = V|S|$ denote the polar decomposition of S. Since S is invertible, V is unitary and

$$(AV)|S| = A(|V|S) = AS = SB = (V|S|)B = V(B|S|) = (VB)|S|.$$

Since $|S|$ is invertible, $AV = VB$. There remains only to show that $V \in \mathbb{M}_n(\mathbb{R})$. On the other hand, $S^*S \in \mathbb{M}_n(\mathbb{R})$, and since $|S|$ is a limit of polynomials in S^*S, $|S| \in \mathbb{M}_n(\mathbb{R})$ as well. Finally, $V = S|S|^{-1} \in \mathbb{M}_n(\mathbb{R})$ since both terms are.

This shows that the $\mathbb{M}_n(\mathbb{R})$ versions of Corollaries 3.9 and 3.11, and all of the results of section 4 also hold.

Acknowledgement

Research supported in part by NSERC (Canada).

References

[1] Drensky, V. and Formanek, E., 2004, *Polynomial Identity Rings*, Advanced Courses in Mathematics, (CRM Barcelona: Birkäuser Verlag, Basel).

[2] Hladnik, M., Omladič, M. and Radjavi, H., 2003, Trace-preserving homomorphisms of semigroups. *Journal of Functional Analysis*, **204**, 269–292.

[3] Pearcy, C., 1962, A complete set of unitary invariants for operators generating finite W^*-algebras of Type I. *Pacific Journal of Mathematics*, **12**, 1405–1416.

[4] Specht, W., 1940, Zur theorie der matrizen, II. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, **50**, 19–23.