

ON THE LENGTHS OF PAIRS OF COMPLEX MATRICES OF SIZE AT MOST FIVE

W.E. LONGSTAFF, A.C. NIEMEYER AND ORESTE PANAIÀ

The length of every pair $\{A, B\}$ of $n \times n$ complex matrices is at most $2n - 2$, if $n \leq 5$. That is, for $n \leq 5$, the (possibly empty) words in A, B of length at most $2n - 2$ span the unital algebra \mathcal{A} generated by A, B . For every positive integer m there exist $m \times m$ complex matrices C, D such that the length of the pair $\{C, D\}$ is $2m - 2$.

1. INTRODUCTION AND PRELIMINARIES

Let \mathbb{F} be a field and let $M_n(\mathbb{F})$ denote the set of all $n \times n$ matrices over \mathbb{F} . Let \mathcal{S} be a finite subset of $M_n(\mathbb{F})$. Let the identity matrix be defined to be the unique word in the alphabet \mathcal{S} of length zero and also call it the empty word. For every positive integer k , define a word in the alphabet \mathcal{S} to be of length k if it has k factors, counting multiplicities, so that, for example, the word $A^2BAC^2A^3$ has length 9 (assuming that $A, B, C \in \mathcal{S}$). For every natural number k let \mathcal{S}_k be the set of words in the alphabet \mathcal{S} of length at most k (including the empty word) and let \mathcal{V}_k be the subspace of $M_n(\mathbb{F})$ spanned by \mathcal{S}_k . Clearly

$$\mathbb{F}I = \mathcal{V}_0 \subseteq \mathcal{V}_1 \subseteq \mathcal{V}_2 \subseteq \cdots \subseteq \mathcal{V}_i \subseteq \mathcal{V}_{i+1} \subseteq \cdots \subseteq \mathcal{A},$$

where \mathcal{A} is the unital algebra generated by \mathcal{S} . Since \mathcal{A} is finite-dimensional, there is an integer l such that $\mathcal{V}_i = \mathcal{V}_{i+1}$. Then $\mathcal{V}_k = \mathcal{V}_l$, for every $k > l$, since

$$\mathcal{S}_k \subseteq \mathcal{S}_{k-l-1}\mathcal{S}_{l+1} \subseteq \mathcal{S}_{k-l-1}\mathcal{V}_l \subseteq \mathcal{V}_{k-1},$$

so $\mathcal{V}_k = \mathcal{V}_{k-1}$, for such k . Since $\mathcal{A} = \bigcup_{k=0}^{\infty} \mathcal{V}_k$ we then have $\mathcal{A} = \mathcal{V}_l$. Following [3], we define the length $l(\mathcal{S})$ of \mathcal{S} to be the smallest integer l for which $\mathcal{V}_l = \mathcal{A}$. Then

$$\mathbb{F}I = \mathcal{V}_0 \subset \mathcal{V}_1 \subset \mathcal{V}_2 \subset \cdots \subset \mathcal{V}_i \subset \mathcal{V}_{i+1} \subset \cdots \subset \mathcal{V}_l = \mathcal{A},$$

where ' \subset ' denotes strict inclusion. From this we get the trivial upper bound $l(\mathcal{S}) \leq d - 1$ where d is the dimension of \mathcal{A} . (Similar types of upper bounds were observed in [1, 2, 3, 4, 5].) In [4] Paz conjectures that $l(\mathcal{S}) \leq 2n - 2$, and shows this to be the case

Received 27th February, 2006

Copyright Clearance Centre, Inc. Serial-fee code: 0004-9727/06 \$A2.00+0.00.

for $n \leq 4$, by proving that $l(\mathcal{S}) \leq \lceil (n^2 + 2)/3 \rceil$, whenever $n \geq 2$. (Here ‘ $\lceil \cdot \rceil$ ’ denotes the least integer function.) Examples exhibited or referred to in [4], attributed to J. W. Carlyle, show that the inequality $l(\mathcal{S}) \leq 2n - 2$ is sharp for $n = 2, 3$ or 4 , and that $l(\mathcal{S}) = 8$ can occur when $n = 5$. Then [1, Example 2.9] shows that $l(\mathcal{S}) \geq 2n - 2$ for infinitely many values of n . Additionally, [3, Theorem 4.1 (b)] shows that $l(\mathcal{S}) \leq 2n - 2$ if \mathcal{S} is a finite set of matrices which generates $M_n(\mathbb{F})$ as a unital algebra, and which contains a matrix with distinct eigenvalues. In [3] it is shown that $l(\mathcal{S}) \leq \sqrt{2}n^{3/2} + 3n$ for every n , for every finite set of matrices \mathcal{S} which generates $M_n(\mathbb{F})$ as a unital algebra.

Here we shall be concerned with the case when $\mathbb{F} = \mathbb{C}$ and the alphabet $\mathcal{S} = \{A, B\}$ where A and B are $n \times n$ matrices. We show that $l(A, B) \leq 2n - 2$ for $n \leq 5$ and that the inequality is sharp for such n .

Let \mathcal{S} be a finite set of $n \times n$ matrices over a field \mathbb{F} which generates $M_n(\mathbb{F})$ as an algebra. Following [2] we define the minimum spanning length of \mathcal{S} , denoted $\text{msl}(\mathcal{S})$, to be the smallest positive integer m such that the set of *nonempty* words of length at most m , in the alphabet \mathcal{S} , span $M_n(\mathbb{F})$. It is obvious that $l(\mathcal{S}) \leq \text{msl}(\mathcal{S})$. Almost as obvious is the fact that $\text{msl}(\mathcal{S}) \leq l(\mathcal{S}) + 1$. (Note that if the nonempty words of length $l - 1$ do not span $M_n(\mathbb{F})$ then their span is strictly included in the span of the nonempty words of length l .) We can have $l(\mathcal{S}) \neq \text{msl}(\mathcal{S})$. Indeed, as is remarked in [2], if B is the 3×3 complex strictly upper triangular elementary Jordan matrix and $A = B^*$, then $l(A, B) = 3$ and $\text{msl}(A, B) = 4$.

OTHER COMMENTS CONCERNING THE RESULTS OF [2].

The manuscript for [2] was written in ignorance of [1, 3, 4, 5]. The first author thanks Thomas J. Laffey for bringing these articles to his attention. In [2] only matrices over the complex field were considered, but some of the results apply to more general fields. Also, most of the results in [2] concerning ‘minimum spanning length’ yield results on ‘length’.

1. Consideration of the proof of [2, Theorem 2] leads to the following theorem.

THEOREM 1. *Let \mathbb{F} be a field with characteristic zero. Let $n \geq 2$ and let $B \in M_n(\mathbb{F})$ be the strictly upper triangular elementary Jordan matrix. For any matrix $A \in M_n(\mathbb{F})$ such that $\{A, B\}$ generates $M_n(\mathbb{F})$ as an algebra, we have $\text{msl}(A, B) \leq 2n - 2$.*

2. With obvious modifications [2, Example 2], shows that, if $n \geq 3$ and \mathbb{F} is a field with characteristic zero, then $\text{msl}(A, B) = 2n - 2$ where $B \in M_n(\mathbb{F})$ is the strictly upper triangular elementary Jordan matrix and $A = (B^t)^{n-1}$ (where B^t denotes the transpose of B). In fact, since it shows that every word $W = (w_{i,j})$ in A and B , including the empty word, of length at most $2n - 3$ satisfies $w_{1,n-1} = w_{2,n}$ it follows that $l(A, B) = 2n - 2$. The latter is true even when $n = 2$.

3. The proofs of [2, Propositions 3 and 4] hold when the underlying field has characteristic zero, not just when it is \mathbb{C} . Thus we have the following proposition.

PROPOSITION 1. *Let \mathbb{F} be a field with characteristic zero. Let $n \geq 2, n \neq 3$, let*

$B \in M_n(\mathbb{F})$ be the strictly upper triangular elementary Jordan matrix and let $A = B^t$. Then $l(A, B) = \text{msl}(A, B) = n$.

2. MAIN RESULTS

In the remainder of this paper the underlying field will be \mathbb{C} , the complex field. Let $n \geq 2$ and let A, B be $n \times n$ complex matrices. Let \mathcal{V} be the set of all words, including the empty word, in A and B . If $U, V \in \mathcal{V}$ and U and V are the same word we write $U \equiv V$. (So $U \equiv V$ is strictly stronger than $U = V$ where the latter means equality *as matrices*.) For each integer $k \geq 1$, totally order the words in A and B of length k using dictionary order. (So, if W_1 and W_2 are words of equal nonzero length, we say that $W_1 \preceq W_2$ if $W_1 \equiv XAV_1$ and $W_2 \equiv XBV_2$, where each of X, V_1, V_2 is a word in A, B , possibly empty.) Extend these orders to a total order on \mathcal{V} by additionally defining $W_1 \preceq W_2$ if the length of W_1 is strictly less than the length of W_2 .

In the totally ordered set \mathcal{V} , define \mathcal{B} to be the set of elements which do not belong to the span of their strict predecessors. Then \mathcal{B} is a linearly independent set of words, hence finite. Clearly $I \in \mathcal{B}$. Note that, if W is a word in \mathcal{B} of length at least 2, then every proper subword of W belongs to \mathcal{B} . For if a word U belongs to the span of words strictly less than it so do the words UV and VU , for any word V .

Now the length of the pair $\{A, B\}$ is at most $2n - 2$ if and only if \mathcal{V}_{2n-2} is the unital algebra generated by A and B if and only if \mathcal{B} does not contain a word of length $2n - 1$. If \mathcal{B} did contain a word of length $2n - 1$, we could let W be the smallest word (in the sense of the total order \preceq) of length $2n - 1$ in \mathcal{B} . Then

- (i) W has length $2n - 1$,
- (ii) W has no factors of the form A^n or B^n ,
- (iii) the number of subwords of W (including the empty subword) is at most n^2 .

We shall describe, for $2 \leq n \leq 5$ the forms that W can have, given the constraints (i), (ii) and (iii) immediately above, and show that W having any one of these forms leads to a contradiction. The pertinent question here is:

QUESTION. Let $n \in \mathbb{Z}^+$, $n \geq 2$. Which words of length $2n - 1$ in the symbols a and b with no factor of a^n or b^n have at most n^2 subwords, including the empty subword?

PROPOSITION 2. *Let $n = 2, 3$ or 4 and let w be a word of length $2n - 1$ in the symbols a and b with no factor of a^n or b^n . If $n = 2$ or 3 , w has more than n^2 subwords. If $n = 4$, w has more than n^2 subwords unless it is $(ab)^3a$ or $(ba)^3b$ in which case it has 14 subwords.*

PROOF:

CASE $n = 2$. If $n = 2$, w must be aba or bab . Consequently, w has $6 > 2^2$ subwords (These are aba, ab, ba, a, b, e , where e is the empty subword, if w is aba .)

CASE $n = 3$. If $n = 3$, w has length 5. The words of length 5 with no factors of a^3 or b^3 are $a^2ba^2, a^2bab, a^2b^2a, aba^2b, ababa$ together with those 11 words that can be obtained from these 5 words by using symmetry, that is, by interchanging a and b or by reading them backwards or by combining the two. Now $ababa$ has 10 subwords and a^2b^2a has 13; each of a^2ba^2, a^2bab, aba^2b has 12.

CASE $n = 4$. If $n = 4$, w has length 7. If w is either $(ab)^3a$ or $(ba)^3b$ it has 14 subwords. If it is neither of these words it has at least $17 > 4^2$ subwords. Indeed, the numbers of different subwords satisfy the following table.

length	7	6	5	4	3	2	1	0
#subwords	1	2	3	≥ 3	≥ 2	≥ 3	2	1

Note that if w was $w_1w_2 \dots w_7$ where $w_i \in \{a, b\}, i = 1, 2, \dots, 7$ and it had only 2 subwords of length 4, then $w_1w_2w_3w_4$ and $w_3w_4w_5w_6$ would be the same, and so would $w_2w_3w_4w_5$ and $w_4w_5w_6w_7$ be. Consequently, w would have the form $(pq)^3p$. It would also have this form if it had no factor of a^2 or b^2 . □

COROLLARY 1. *Let $n \in \mathbb{Z}$. If $n = 2$ or 3 there are no words of length $2n - 1$ in a and b containing no factors of a^n or b^n with at most n^2 subwords. If $n = 4$ the only words of length $2n - 1$ containing no factor of a^n or b^n with at most n^2 subwords (including the empty word) are $(ab)^3a$ and $(ba)^3b$. Each of the latter has 14 subwords.*

Analysis of the case $n = 5$ is a little more difficult.

PROPOSITION 3. *Let w be a word of length 9 in the letters a and b satisfying*

- (i) w has no factor of the form a^5 or b^5 ,
- (ii) w is not of the form $(pq)^4r, p(qr)^4$ or $(pqr)^3$, where $\{p, q, r\} \subseteq \{a, b\}$.

Then w has at least 26 subwords, including the empty subword.

PROOF: If words f and g , in a and b , are the same word we write $f \equiv g$.

Let $w \equiv w_1w_2 \dots w_9$ where $\{w_1, w_2, \dots, w_9\} \subseteq \{a, b\}$. We shall show that the numbers of different subwords satisfy the following table.

length	9	8	7	6	5	4	3	2	1	0
#subwords	1	2	3	4	≥ 4	≥ 4	≥ 2	≥ 3	2	1

The sum of the numbers in the last row is at least 26, so this will prove the theorem.

Consider the 4 subwords of length 6. Call them $V_1 \equiv w_1w_2 \dots w_6, V_2 \equiv w_2w_3 \dots w_7, V_3 \equiv w_3w_4 \dots w_8, V_4 \equiv w_4w_5 \dots w_9$. By condition (i), $V_1 \not\equiv V_2 \not\equiv V_3 \not\equiv V_4$. If $V_3 \equiv V_1$ then w has the form $(pq)^4r$. If $V_4 \equiv V_2$ then w has the form $p(qr)^4$. If $V_4 \equiv V_1$ then w has the form $(pqr)^3$. Each of these contradicts condition (ii), so V_1, V_2, V_3, V_4 are distinct. It follows that all of the subwords of lengths 6, 7, 8 or 9 are distinct.

Consider the 5 subwords of length 5. We shall abuse notation and call them V_1, V_2, \dots, V_5 , where

$$V_1 \equiv w_1 w_2 \dots w_5, V_2 \equiv w_2 w_3 \dots w_6, V_3 \equiv w_3 w_4 \dots w_7, V_4 \equiv w_4 w_5 \dots w_8, V_5 \equiv w_5 w_6 \dots w_9.$$

Again, by condition (i), $V_1 \not\equiv V_2 \not\equiv V_3 \not\equiv V_4 \not\equiv V_5$. If at most 3 of these 5 subwords were distinct, there are only 7 possibilities, namely,

- (5-3i) $V_1 \equiv V_3$ and $V_2 \equiv V_5$,
- (5-3ii) $V_1 \equiv V_4$ and $V_3 \equiv V_5$,
- (5-3iii) $V_1 \equiv V_3$ and $V_2 \equiv V_4$,
- (5-3iv) $V_1 \equiv V_3 \equiv V_5$,
- (5-3v) $V_1 \equiv V_5$ and $V_2 \equiv V_4$,
- (5-3vi) $V_2 \equiv V_4$ and $V_3 \equiv V_5$,
- (5-3vii) $V_1 \equiv V_4$ and $V_2 \equiv V_5$.

If case (5-3i) or (5-3ii) held, w would be of the form p^9 . If either of (5-3iii), (5-3iv) or (5-3v) held, w would be of the form $(pq)^4 r$. If (5-3vi) held, w would be of the form $p(qr)^4$ and if (5-3vii) held it would be of the form $(pqr)^3$. Each of these contradicts condition (5-3i) or (5-3ii). Thus w has at least 4 distinct subwords of length 5.

Consider the 6 subwords of length 4, namely,

$$V_1 \equiv w_1 \dots w_4, V_2 \equiv w_2 \dots w_5, V_3 \equiv w_3 \dots w_6, V_4 \equiv w_4 \dots w_7, V_5 \equiv w_5 \dots w_8, V_6 \equiv w_6 \dots w_9.$$

Again, by condition (i), $V_1 \not\equiv V_2 \not\equiv V_3 \not\equiv V_4 \not\equiv V_5 \not\equiv V_6$. If at most 3 of these 6 subwords were distinct, there are only 15 possibilities, namely,

- (4-3i) $V_1 \equiv V_3 \equiv V_6$ and $V_2 \equiv V_4$,
- (4-3ii) $V_1 \equiv V_3 \equiv V_6$ and $V_2 \equiv V_5$,
- (4-3iii) $V_1 \equiv V_3, V_2 \equiv V_5$ and $V_4 \equiv V_6$,
- (4-3iv) $V_1 \equiv V_4 \equiv V_6$ and $V_2 \equiv V_5$,
- (4-3v) $V_1 \equiv V_4 \equiv V_6$ and $V_3 \equiv V_5$,
- (4-3vi) $V_1 \equiv V_4, V_2 \equiv V_6$ and $V_3 \equiv V_5$,
- (4-3vii) $V_1 \equiv V_5, V_2 \equiv V_4$ and $V_3 \equiv V_6$,
- (4-3viii) $V_1 \equiv V_6, V_2 \equiv V_4$ and $V_3 \equiv V_5$,
- (4-3ix) $V_1 \equiv V_3 \equiv V_5$ and $V_2 \equiv V_4$,
- (4-3x) $V_1 \equiv V_3$ and $V_2 \equiv V_4 \equiv V_6$,
- (4-3xi) $V_1 \equiv V_3 \equiv V_5$ and $V_2 \equiv V_6$,
- (4-3xii) $V_1 \equiv V_3 \equiv V_5$ and $V_4 \equiv V_6$,
- (4-3xiii) $V_1 \equiv V_5$ and $V_2 \equiv V_4 \equiv V_6$,
- (4-3xiv) $V_1 \equiv V_4 \equiv V_6$ and $V_3 \equiv V_5$,

$$(4-3xv) \quad V_1 \equiv V_4, V_2 \equiv V_5, \text{ and } V_3 \equiv V_6.$$

If either of cases (4-3i) to (4-3viii) held, w would be of the form p^9 . If either of (4-3ix) to (4-3xiii) held, w would be of the form $(pq)^4r$. If (4-3xiv) held, w would be of the form $p(qr)^4$ and if (4-3xv) held it would be of the form $(pqr)^3$. But each of these contradicts condition (i) or (ii). Thus w has at least 4 distinct subwords of length 4.

The subwords $w_1w_2w_3, w_2w_3w_4, w_3w_4w_5$ of w of length 3 cannot be the same, otherwise w would be of the form p^5qrst which contradicts condition (i). Thus w has at least two subwords of length 3.

With regard to subwords of w of length 2, note that both ab and ba must be subwords by condition (i). Also, either a^2 or b^2 must be a subword, by condition (ii). Thus w has at least 3 subwords of length 2.

The above analysis shows that the subwords of w occur as in the table given at the beginning of the proof. As remarked earlier, this completes the proof. \square

COROLLARY 2. *Let w be a word in a and b of length 9. If w has no factors of a^5 or b^5 and has 25 or fewer subwords (including the empty word) it must be one of*

$$(ab)^4a; (ba)^4b$$

$$(a^2b)^3; (b^2a)^3; a(ab)^4; b(ba)^4; (aba)^3; (bab)^3; (ab)^4b; (ba)^4a; (ab^2)^3; (ba^2)^3.$$

The words in the first row above each have 18 subwords; those in the last row have 24.

PROOF: By Proposition 3, w must be of one of the forms $(pq)^4r, p(qr)^4$ or $(pqr)^3$, where $\{p, q, r\} \subseteq \{a, b\}$. Bearing in mind the fact that w has no factor of a^5 or b^5 gives that it must be one of the 12 words listed. To verify that the actual number of subwords is as claimed, one need only consider, by symmetry, $(ab)^4a; (a^2b)^3; a(ab)^4; (aba)^3$. \square

PROPOSITION 4. *Let $n \geq 2$. For all complex $n \times n$ matrices A and B , the matrix $(AB)^{n-1}A$ belongs to \mathcal{V}_{2n-3} .*

PROOF: First, assume that B is invertible. Let $W \equiv (AB)^{n-1}A$. Then, using the Cayley–Hamilton Theorem,

$$WB \equiv (AB)^n = \lambda_0 I + \sum_{k=1}^{n-1} \lambda_k (AB)^k,$$

for some scalars $\lambda_0, \lambda_1, \dots, \lambda_{n-1}$. Hence

$$W = \lambda_0 B^{-1} + \sum_{k=1}^{n-1} \lambda_k (AB)^{k-1} A.$$

Since B^{-1} can be written as a polynomial in B of degree $n - 1$, the desired result follows.

Finally, assume that B is not invertible. Choose a scalar λ such that $B - \lambda I$ is invertible. By the above, $(A(B - \lambda I))^{n-1}A$ belongs to the span of words in $A, B - \lambda I$ of

length at most $2n - 3$ together with the identity, and hence to \mathcal{V}_{2n-3} . But

$$(A(B - \lambda I))^{n-1}A - (AB)^{n-1}A$$

belongs to \mathcal{V}_{2n-4} , so the proof is complete. □

PROPOSITION 5. *If $n = 2, 3$ or 4 , then $l(A, B) \leq 2n - 2$, for all $n \times n$ complex matrices A and B , that is, the unital algebra generated by A, B is the span of the set of all, possibly empty, words in A, B of length at most $2n - 2$.*

PROOF: Define \mathcal{B} as before, after totally-ordering the set of all words in A and B . If $l(A, B) > 2n - 2$, \mathcal{B} would contain a word W of length $2n - 1$. This word would have no factor of A^n or B^n and have at most n^2 subwords. By Corollary 1, there is no such word if $n = 2$ or 3 . If $n = 4$, again by Corollary 1, W could only be $(AB)^3A$ or $(BA)^3B$. But these words cannot belong to \mathcal{B} by Proposition 4. □

The remainder of this paper is devoted to showing that n can also be taken to be 5 in Proposition 5.

THEOREM 2. *For all 5×5 complex matrices A and B , $l(A, B) \leq 8$.*

PROOF: Suppose that $l(A, B) > 8$ and let W be the smallest word of length 9 belonging to \mathcal{B} . Then W contains no factors of A^5 or B^5 and has at most 25 subwords (including the empty subword). By Corollary 2 it must be one of the words

$$(AB)^4A; (BA)^4B; (A^2B)^3; (B^2A)^3; A(AB)^4; B(BA)^4; \\ (ABA)^3; (BAB)^3; (AB)^4B; (BA)^4A; (AB^2)^3; (BA^2)^3.$$

By Proposition 4, W cannot be either $(AB)^4A$ or $(BA)^4B$. Now each of the remaining 10 possible words has 24 subwords. We complete the proof by showing that W being any one of these 10 possible words leads to a contradiction. Notice that

- (a) Since W is the smallest word of length 9 belonging to \mathcal{B} , if V is a word of length 9 satisfying $V \preceq W, V \neq W$ then $V \in \mathcal{V}_8$,
- (b) Since $M_5(\mathbb{C})$ has dimension 25, every proper subword of any element of \mathcal{B} is a subword of W .

By the Cayley–Hamilton Theorem both

$$E = \frac{(A + B)^5 + (A - B)^5}{2} \quad \text{and} \quad F = \frac{(A + iB)^5 + (A - iB)^5}{2}$$

belong to \mathcal{V}_4 . Therefore, so do

$$(I) \quad Y = \frac{E - F}{2} = A^3B^2 + A^2BAB + A^2B^2A + ABA^2B + ABABA \\ + AB^2A^2 + BA^3B + BA^2BA + BABA^2 + B^2A^3.$$

and, interchanging A and B ,

$$(II) \quad Z = A^2B^3 + ABAB^2 + AB^2AB + AB^3A + BA^2B^2 + BABAB + BAB^2A + B^2A^2B + B^2ABA + B^3A^2$$

(i) Suppose that W was the word $(AB)^4B$. In the expression for Y the first 4 words on the right hand side are less than the fifth $ABABA$. Thus, if X is any of these first 4 words $(X)BAB^2 \preceq (ABABA)BAB^2 \equiv W$ and so $(X)BAB^2 \in \mathcal{V}_8$, by the remark (a) immediately above. Thus

$$(Y)BAB^2 = (ABABA + AB^2A^2 + BA^3B + BA^2BA + BABA^2 + B^2A^3)BAB^2 + V \in \mathcal{V}_8$$

where $V \in \mathcal{V}_8$. Hence

$$W + AB^2A^2BAB^2 + BA^3B^2AB^2 + BA^2BABAB^2 + BABA^2BAB^2 + B^2A^3BAB^2 \in \mathcal{V}_8.$$

Neither of A^3, A^2B belongs to \mathcal{B} , by remark (b) immediately above (each has A^2 as a proper subword), so both belong to \mathcal{V}_2 (because of the way that \mathcal{B} is defined). Since each of

$$AB^2A^2BAB^2, BA^3B^2AB^2, BA^2BABAB^2, BABA^2BAB^2, B^2A^3BAB^2$$

has A^2B as a factor, each of these 5 words belong to \mathcal{V}_8 . It follows that $W \in \mathcal{V}_8$. This is a contradiction. Thus W cannot be $(AB)^4B$.

The proof that W cannot be $B(BA)^4$ is similar to the one just given.

(ii) Suppose that W was the word $B(BA)^4$. In the expression (I) for Y , if X is any of the first 4 words on the right hand side then $B^2(X)BA \preceq B^2(ABABA)BA \equiv W$ and so $B^2(X)BA \in \mathcal{V}_8$. This gives

$$B^2(Y)BA = B^2(ABABA + AB^2A^2 + BA^3B + BA^2BA + BABA^2 + B^2A^3)BA + V \in \mathcal{V}_8$$

where $V \in \mathcal{V}_8$. Hence

$$W + B^2AB^2A^2BA + B^3A^3B^2A + B^3A^2BABA + B^3ABA^2BA + B^4A^3BA \in \mathcal{V}_8.$$

Again, neither A^3 nor A^2B belongs to \mathcal{B} so both belong to \mathcal{V}_2 . Since each of

$$B^2AB^2A^2BA, B^3A^3B^2A, B^3A^2BABA, B^3ABA^2BA, B^4A^3BA$$

has A^2B as a factor, each of these 5 words belong to \mathcal{V}_8 . It follows that $W \in \mathcal{V}_8$. This is a contradiction. Thus W cannot be $B(BA)^4$.

(iii) Suppose that $W \equiv (BA)^4A$. If X is any of the first 8 words on the right hand side in the expression (I) for Y , then $BABA(X) \preceq BABA(BABA^2) \equiv W$ and so $BABA(X) \in \mathcal{V}_8$. This gives

$$BABA(Y) = BABA(BABA^2 + B^2A^3) + V_1 \in \mathcal{V}_8$$

where $V_1 \in \mathcal{V}_8$. Hence $W + BABAB^2A^3 \in \mathcal{V}_8$. If $A^3 \notin \mathcal{B}$ then $A^3 \in \mathcal{V}_2$ and so $BABAB^2A^3 \in \mathcal{V}_8$. This gives $W \in \mathcal{V}_8$, which is a contradiction. Thus $A^3 \in \mathcal{B}$, and the elements of \mathcal{B} are precisely A^3 together with all the subwords of W . It follows that neither A^2B nor AB^2 belongs to \mathcal{B} and that $AB^2 = \alpha A^3 + \beta ABA + V_2$ for some scalars α, β and some element $V_2 \in \mathcal{V}_2$. Hence

$$BABAB^2A^3 = BAB(\alpha A^3 + \beta ABA + V_2)A^3 = \alpha BABA^6 + \beta BABABA^4 + BABV_2A^3.$$

But $BABA^6 \in \mathcal{V}_8$ by the Cayley–Hamilton Theorem, $BABABA^4 \in \mathcal{V}_8$ because $A^4 \in \mathcal{V}_3$ (since $A^4 \notin \mathcal{B}$) and clearly $BABV_2A^3 \in \mathcal{V}_8$. Thus $BABAB^2A^3 \in \mathcal{V}_8$, so $W \in \mathcal{V}_8$. Again this is a contradiction.

(iv) Suppose that $W \equiv A(AB)^4$. If X is any of the first 5 words on the right hand side in the expression (II) for Z , then $A^2(X)AB \preceq A^2(BABAB)AB \equiv W$ and so $A^2(X)AB \in \mathcal{V}_8$. This gives

$$A^2(Z)AB = A^2(BABAB + BAB^2A + B^2A^2B + B^2ABA + B^3A^2)AB + V_1 \in \mathcal{V}_8$$

where $V_1 \in \mathcal{V}_8$. Hence

$$W + A^2BAB^2A^2B + A^2B^2A^2BAB + A^2B^2ABA^2B + A^2B^3A^3B \in \mathcal{V}_8.$$

None of $A^4, A^3B, A^2B^2, ABA^2, AB^2$ can belong to \mathcal{B} since each has a proper subword which is not a subword of W . In particular, $A^4, A^3B \in \mathcal{V}_3$ and $AB^2 = \alpha A^3 + \beta A^2B + \gamma ABA + V_2$ for some scalars α, β, γ and some $V_2 \in \mathcal{V}_2$. Thus

$$\begin{aligned} A^2B(AB^2)A^2B &= A^2B(\alpha A^3 + \beta A^2B + \gamma ABA + V_2)A^2B \\ &= \alpha A^2BA^5B + \beta (A^2B)^3 + \gamma A^2BABA^3B + A^2BV_2A^2B. \end{aligned}$$

Since $ABA^2 = \delta A^2BA + V_3$ for some scalar δ and some $V_3 \in \mathcal{V}_3$,

$$(A^2B)^3 \equiv A(ABA^2)BA^2B = A(\delta A^2BA + V_3)BA^2B = \delta A^3BABA^2B + AV_3BA^2B \in \mathcal{V}_8,$$

since $A^3B \in \mathcal{V}_3$. Using the fact that $A^3B \in \mathcal{V}_3$ once again together with the Cayley–Hamilton Theorem gives that $A^2BAB^2A^2B \in \mathcal{V}_8$.

Finally note that

$$\begin{aligned} A^2B^2A^2BAB &\equiv A(AB^2)A^2BAB = A(\alpha A^3 + \beta A^2B + \gamma ABA + V_2)A^2BAB \\ &= \alpha A^6BAB + \beta A^3BA^2BAB + \gamma A^2BA^3BAB + AV_2A^2BAB \in \mathcal{V}_8 \end{aligned}$$

and

$$\begin{aligned} A^2B^2ABA^2B &\equiv A(AB^2)AB^2ABA^2B = A(\alpha A^3 + \beta A^2B + \gamma ABA + V_2)ABA^2B \\ &= \alpha A^5BA^2B + \beta A^3BABA^2B + \gamma (A^2B)^3 + AV_2ABA^2B \in \mathcal{V}_8, \end{aligned}$$

using the Cayley–Hamilton Theorem and the facts that $A^3B \in \mathcal{V}_3, (A^2B)^3 \in \mathcal{V}_8$. This now gives $W \in \mathcal{V}_8$ which is a contradiction. Thus W cannot be of the form $A(AB)^4$.

(v) Suppose that W was the word $(A^2B)^3$. In the expression (I) for Y , the first 7 words on the right hand side are less than the eighth BA^2BA . If X is any one of these seven, $A^2(X)AB \preceq A^2(BA^2BA)AB \equiv W$, so $A^2(X)AB \in \mathcal{V}_8$. Thus

$$A^2(Y)AB = A^2(BA^2BA + BABA^2 + B^2A^3)AB + V \in \mathcal{V}_8$$

where $V \in \mathcal{V}_8$. Hence $W + A^2BABA^3B + A^2B^2A^4B \in \mathcal{V}_8$. If $A^3 \notin \mathcal{B}$ then $A^3 \in \mathcal{V}_2$, so both A^2BABA^3B and $A^2B^2A^4B$ belong to \mathcal{V}_8 , and it follows that $W \in \mathcal{V}_8$. This is a contradiction. Thus $A^3 \in \mathcal{B}$ and \mathcal{B} consists precisely of A^3 together with all the subwords of W . In particular, $A^4, A^3B \notin \mathcal{B}$ so $A^4, A^3B \in \mathcal{V}_3$. From the former it follows that $A^2B^2A^4B \in \mathcal{V}_8$ and from the latter that $A^2BABA^3B \in \mathcal{V}_8$. It then follows that $W \in \mathcal{V}_8$. Again, this is a contradiction. Thus W cannot be $(A^2B)^3$.

The proof that W cannot be $(ABA)^3$ is similar to the above.

(vi) Suppose that $W \equiv (ABA)^3$. Note that if X is any one of the first seven words on the right hand side in the expression (I) for Y , then

$$A(X)ABA \preceq A(BA^2BA)ABA \equiv W,$$

so $A(X)ABA \in \mathcal{V}_8$. It follows that $W + ABABA^3BA + AB^2A^4BA \in \mathcal{V}_8$. Again, A^3 must belong to \mathcal{B} and $A^4, A^3B \notin \mathcal{B}$, so $A^4, A^3B \in \mathcal{V}_3$. From the former it follows that $AB^2A^4BA \in \mathcal{V}_8$ and from the latter that $ABABA^3BA \in \mathcal{V}_8$. It then follows that $W \in \mathcal{V}_8$. Again, this is a contradiction. Thus W cannot be $(ABA)^3$.

(vii) Suppose that $W \equiv (BA^2)^3$. We use the same Y and note that

$$BA^2(X)A \preceq BA^2(BA^2BA)A \equiv W,$$

so $BA^2(X)A \in \mathcal{V}_8$, if X is any one of the first seven words on the right hand side. This gives $W + BA^2BABA^3 + BA^2B^2A^4 \in \mathcal{V}_8$ and again, $A^3 \in \mathcal{B}$ and \mathcal{B} consists precisely of A^3 together with all the subwords of W . In particular, $A^4 \notin \mathcal{B}$ so $A^4 \in \mathcal{V}_3$ and $BA^2B^2A^4 \in \mathcal{V}_8$. Also (using the fact that $AB^2 \notin \mathcal{B}$) $BAB = \alpha A^3 + \beta A^2B + \gamma ABA + \delta BA^2 + V$ for some scalars $\alpha, \beta, \gamma, \delta$ and some $V \in \mathcal{V}_2$. Then

$$\begin{aligned} BA^2(BAB)A^3 &= BA^2(\alpha A^3 + \beta A^2B + \gamma ABA + \delta BA^2 + V)A^3 \\ &= \alpha BA^8 + \beta BA^4BA^3 + \gamma BA^3BA^4 + \delta BA^2BA^5 + BA^2VA^3 \in \mathcal{V}_8 \end{aligned}$$

(since $A^4 \notin \mathcal{B}$ implies that $A^4 \in \mathcal{V}_3$). Again it follows that $W \in \mathcal{V}_8$. This is a contradiction so W cannot be $(BA^2)^3$.

The proofs that W cannot be $(AB^2)^3, (BAB)^3$ or $(B^2A)^3$ are similar to one another.

(viii) Suppose that $W \equiv (AB^2)^3$. In the expression (II) for Z the first 6 words on the right hand side are less than the seventh BAB^2A . If X is any one of these six, $AB(X)B^2 \preceq AB(BAB^2A)B^2 \equiv W$, so $AB(X)B^2 \in \mathcal{V}_8$. Thus

$$AB(Z)B^2 = AB(BAB^2A + B^2A^2B + B^2ABA + B^3A^2)B^2 + V \in \mathcal{V}_8$$

where $V \in \mathcal{V}_8$. Hence

$$W + AB^3A^2B^3 + AB^3ABAB^2 + AB^4A^2B^2 \in \mathcal{V}_8.$$

Since AB^2A is a subword of W it belongs to \mathcal{B} . In fact AB^2A is the smallest word of length 4 in A and B which belongs to \mathcal{B} . Indeed, none of $A^4, A^3B, A^2BA, A^2B^2, ABA^2$ can belong to \mathcal{B} because each has A^2 as a proper subword, and A^2 is not a subword of W (see remark (b) above). Also, $ABAB \notin \mathcal{B}$ since ABA is not a subword of W . Thus, A^2B^2 and $ABAB$ both belong to \mathcal{V}_4 . It follows that

$$AB^3A^2B^3, AB^3ABAB^2, AB^4A^2B^2 \in \mathcal{V}_8$$

and that $W \in \mathcal{V}_8$. This is a contradiction so W cannot be $(AB^2)^3$.

(ix) Suppose that $W \equiv (BAB)^3$. Using Z as above, we have $(X)B^2AB \preceq (BAB^2A)B^2AB \equiv W$, so $(X)B^2AB \in \mathcal{V}_8$, for any of the first 6 words X on the right hand side of the expression (II) for Z . This gives

$$W + B^2A^2B^3AB + B^2ABAB^2AB + B^3A^2B^2AB \in \mathcal{V}_8.$$

By exactly the same argument as in the proof of case (viii) immediately above, both A^2B^2 and $ABAB$ belong to \mathcal{V}_4 . It follows that

$$B^2A^2B^3AB, B^2ABAB^2AB, B^3A^2B^2AB \in \mathcal{V}_8$$

and that $W \in \mathcal{V}_8$. Again, this is a contradiction so W cannot be $(BAB)^3$.

(x) Suppose that $W \equiv (B^2A)^3$. Using Z as above, we have $B(X)B^2A \preceq B(BAB^2A)B^2A \equiv W$, so $B(X)B^2A \in \mathcal{V}_8$, for any of the first 6 words X on the right hand side of the expression (II) for Z . This gives

$$W + B^3A^2B^3A + B^3ABAB^2A + B^4A^2B^2A \in \mathcal{V}_8.$$

Once again, both A^2B^2 and $ABAB$ belong to \mathcal{V}_4 so

$$B^3A^2B^3A, B^3ABAB^2A, B^4A^2B^2A \in \mathcal{V}_8$$

and $W \in \mathcal{V}_8$. This is a contradiction so W cannot be $(B^2A)^3$.

This completes the proof of the theorem. □

REFERENCES

- [1] A. Freedman, R. Gupta and R. Guralnick, 'Shirshov's theorem and representations of semigroups', *Pacific J. Math., Special Issue* (1997), 159-176.
- [2] W.E. Longstaff, 'Burnside's Theorem: irreducible pairs of transformations', *Linear Algebra Appl.* **382** (2004), 247-269.
- [3] C.J. Pappacena, 'An upper bound for the length of a finite-dimensional algebra', *J. Algebra* **197** (1997), 535-545.
- [4] A. Paz, 'An application of the Cayley-Hamilton theorem to matrix polynomials in several variables', *Linear and Multilinear Algebra* **15** (1984), 161-170.
- [5] C. Procesi, *Rings with polynomial identities* (Marcel Dekker, New York, 1973).

School of Mathematics and Statistics
The University of Western Australia
35 Stirling Highway
Crawley WA 6009
Australia
e-mail: longstaf@maths.uwa.edu.au
alice@maths.uwa.edu.au
oreste@maths.uwa.edu.au