

ON THE LENGTHS OF IRREDUCIBLE PAIRS OF COMPLEX MATRICES

W. E. LONGSTAFF AND PETER ROSENTHAL

(Communicated by Marius Junge)

ABSTRACT. The length of a pair of matrices is the smallest integer l such that words in the matrices with at most l factors span the unital algebra generated by the pair. Upper bounds for lengths have been much studied. If B is a rank one $n \times n$ (complex) matrix, the length of the irreducible pair $\{A, B\}$ is $2n - 2$ and the subwords of $A^{n-1}BA^{n-2}$ form a basis for $M_n(\mathbb{C})$. New examples are given of irreducible pairs of $n \times n$ matrices of length n . There exists an irreducible pair of 5×5 matrices of length 4. We begin the study of determining lower bounds for lengths.

1. INTRODUCTION AND PRELIMINARIES

Let \mathbb{C} be the complex field and let $M_n(\mathbb{C})$ denote the set of all $n \times n$ complex matrices. Let \mathcal{S} be a finite subset of $M_n(\mathbb{C})$. 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{V}_k be the subspace of $M_n(\mathbb{C})$ spanned by the set of words in the alphabet \mathcal{S} of length at most k (including the empty word). Clearly

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

where \mathcal{A} is the unital algebra generated by \mathcal{S} . Since \mathcal{A} is finite-dimensional, there is an integer l such that $\mathcal{V}_l = \mathcal{V}_{l+1}$. Then clearly $\mathcal{V}_k = \mathcal{V}_l$, for every $k > l$. Since $\mathcal{A} = \bigcup_{k=0}^{\infty} \mathcal{V}_k$, we then have $\mathcal{A} = \mathcal{V}_l$. The length $l(\mathcal{S})$ of \mathcal{S} is defined to be the smallest integer l for which $\mathcal{V}_l = \mathcal{A}$. Then

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

where ‘ \subset ’ denotes strict inclusion. If \mathcal{S} is an irreducible family of matrices (that is, its members have no common invariant subspaces except (0) and \mathbb{C}^n), then $\mathcal{A} = M_n(\mathbb{C})$ by Burnside’s Theorem. In this case, the length l of \mathcal{S} is the smallest integer such that \mathcal{V}_l contains a basis for $M_n(\mathbb{C})$.

Thanks to the Cayley-Hamilton Theorem, it is clear that if $\mathcal{S} = \{A, B\}$ where A and B are commuting matrices, then \mathcal{S} has length at most $2n - 2$. Interestingly, in [8], Paz conjectures that $l(\mathcal{S}) \leq 2n - 2$ for all \mathcal{S} . He shows this to be the case for

Received by the editors March 1, 2010.

2010 *Mathematics Subject Classification*. Primary 15A30; Secondary 47L05.

Key words and phrases. Length, words.

©2011 American Mathematical Society
Reverts to public domain 28 years from publication

$n \leq 4$, 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.) The inequality $l(\mathcal{S}) \leq 2n - 2$ has been verified and proven to be sharp for the case where \mathcal{S} is a pair of complex matrices of size at most 6 in [4], [6]. It is shown in [7] that $l(\mathcal{S}) \leq 2n - 2$ if \mathcal{S} is an irreducible set of $n \times n$ matrices which contains a matrix with n distinct eigenvalues and that $l(\mathcal{T}) \leq \sqrt{2}n^{3/2} + 3n$, for every irreducible set of $n \times n$ matrices \mathcal{T} . Some other recent results concerning length are contained in [2], [3], [5].

If e and f are nonzero vectors in \mathbb{C}^n , we will use the notation $e \otimes f$ to denote the rank one operator on \mathbb{C}^n defined by $(e \otimes f)x = (x|e)f$, $x \in \mathbb{C}^n$, where $(\cdot|\cdot)$ denotes the usual inner-product on \mathbb{C}^n .

2. WORDS OF THE FORM A^pBA^q

We first consider the phenomenon where, for a given pair of $n \times n$ matrices, the words of the form A^pBA^q , $0 \leq p, q \leq n - 1$, form a basis for $M_n(\mathbb{C})$. Possibly the simplest example of this phenomenon is when $A = J$ and $B = (J^*)^{n-1}$, where J is the $n \times n$ strictly upper triangular elementary Jordan matrix (that is, the matrix with ones on the first superdiagonal and zeroes elsewhere). It is easily verified that, for this A and B , $A^pBA^q = E_{n-p, q+1}$, where $E_{i,j}$ is the elementary matrix having a one as its (i, j) -th entry, and all its other entries zero.

We show below (Corollaries 1 and 2) that not much can be said about A and B if $\{A^pBA^q : 0 \leq p, q \leq n - 1\}$ is a basis for $M_n(\mathbb{C})$. More precisely, under such circumstances, we can deduce only that the characteristic polynomial and minimal polynomials of A are the same and that B is nonscalar.

Proposition 1. *For every pair $\{A, B\}$ of $n \times n$ complex matrices (irreducible or not), we have $A^{n-1}BA^{n-1} \in \mathcal{V}_{2n-2}$.*

Proof. By the Cayley-Hamilton Theorem, $(A + \lambda B)^n \in \mathcal{V}_{n-1}$ for every $\lambda \in \mathbb{C}$. Now, there exist matrices $S_0, S_1, S_2, \dots, S_n$ such that, for every λ , $(A + \lambda B)^n = S_0 + S_1\lambda + S_2\lambda^2 + \dots + S_n\lambda^n$. It follows from [4, Proposition 4] that $S_i \in \mathcal{V}_{n-1}$, for every $i = 0, 1, 2, \dots, n$. In particular, $S_1 \in \mathcal{V}_{n-1}$ where S_1 is the sum of all words of length n in A, B with precisely one B factor. Clearly S_1 is a sum of n terms:

$$S_1 = A^{n-1}B + A^{n-2}BA + A^{n-3}BA^2 + \dots + A^2BA^{n-3} + ABA^{n-2} + BA^{n-1}.$$

Premultiplying by A^{n-1} , each of the first $n - 1$ terms that result belongs to \mathcal{V}_{2n-2} by the Cayley-Hamilton Theorem. Since $A^{n-1}S_1 \in \mathcal{V}_{2n-2}$, it follows that the last term that results, namely $A^{n-1}BA^{n-1}$, also belongs to \mathcal{V}_{2n-2} . \square

Remark 1. If $A, B \in M_n(\mathbb{C})$ and the set of words $\{A^pBA^q : 0 \leq p, q \leq n - 1\}$ is a basis for $M_n(\mathbb{C})$, then $\{A, B\}$ is an irreducible pair with length at most $2n - 1$ (the length of $A^{n-1}BA^{n-1}$). By the above proposition, however, $\{A^pBA^q : 0 \leq p, q \leq n - 1\} \subseteq \mathcal{V}_{2n-2}$, so the length of $\{A, B\}$ is at most $2n - 2$.

Remark 2. In what follows we need the notion of *echelon representative*.

Definition 1. Let $\rho : M_n(\mathbb{C}) \rightarrow \mathbb{C}^{n^2}$ be the mapping defined by

$$\rho(C) = (c_{n,1}, c_{n-1,1}, c_{n,2}, c_{n-2,1}, c_{n-1,2}, c_{n,3}, \dots, c_{1,n-1}, c_{2,n}, c_{1,n})$$

where $C = (c_{i,j})$. (Here we are writing the matrix as a row vector starting from the bottom left-hand corner and finishing at the top right-hand corner.) The $n \times n$ matrix C is an (i, j) -echelon representative if the first nonzero entry of the row vector $\rho(C)$ is $c_{i,j}$ in the row position corresponding to the matrix position (i, j) .

Proposition 2. *Let $n \geq 2$, let $B = (b_{i,j})$ be an $n \times n$ matrix and let J be the strictly upper triangular elementary $n \times n$ Jordan matrix. Then*

- (i) $\{J^p B J^q : 0 \leq p, q \leq n - 1\}$ is a basis for $M_n(\mathbb{C})$ if and only if $b_{n,1} \neq 0$,
- (ii) if $b_{n,1} \neq 0$, then the subwords of $J^{n-2} B J^{n-1}$, including the empty word I , form a basis for $M_n(\mathbb{C})$.

Proof. (i) Suppose that $b_{n,1} \neq 0$. The n^2 matrices $\{J^p B J^q : 0 \leq p, q \leq n - 1\}$ give a set of echelon representatives for all of the n^2 matrix positions.

If $b_{n,1} = 0$, it is obvious that every one of the matrices in $\{J^p B J^q : 0 \leq p, q \leq n - 1\}$ has $(n, 1)$ -entry equal to zero, so then will every matrix in their span.

(ii) Let $b_{n,1} \neq 0$. Echelon representatives for the first row positions are $J^{n-1} B J^q$, $0 \leq q \leq n - 1$. The desired result is obtained by replacing these with $I, J, J^2, \dots, J^{n-1}$. □

Corollary 1. *If B is an $n \times n$ matrix, there exists an $n \times n$ matrix A such that $\{A^p B A^q : 0 \leq p, q \leq n - 1\}$ is a basis for $M_n(\mathbb{C})$ if and only if B is not a scalar multiple of the identity matrix.*

Proof. The necessity of the condition is obvious. Conversely, if B is nonscalar, there is a basis for \mathbb{C}^n such that the $(n, 1)$ -entry of its matrix with respect to this basis is nonzero. The existence of the required matrix A follows from Proposition 2(i). □

Remark 3. Suppose that $B \notin \mathbb{C}I$. The matrix A obtained in the proof of Corollary 1 has only one eigenvalue. In fact, there is a matrix A with n distinct eigenvalues such that $\{A^p B A^q : 0 \leq p, q \leq n - 1\}$ is a basis for $M_n(\mathbb{C})$. For, since B is not a scalar multiple of the identity matrix, there exists a basis for \mathbb{C}^n such that the matrix of B with respect to it has every entry nonzero [9]. By properties of similarities we may suppose at the outset that every entry of B is nonzero. Let A be an $n \times n$ diagonal matrix with distinct diagonal entries. Then each of the elementary matrices $E_{i,i}$ is a polynomial in A of degree $n - 1$ and $E_{i,i} B E_{j,j} = b_{i,j} E_{i,j}$ for every $1 \leq i, j \leq n$. This shows that every elementary matrix $E_{i,j}$ belongs to the span of $\{A^p B A^q : 0 \leq p, q \leq n - 1\}$. Since the latter set has n^2 elements, it forms a basis for $M_n(\mathbb{C})$.

Theorem 1. *If $\{A, B\}$ is an irreducible pair of $n \times n$ matrices and B has rank one, then the words $\{A^p B A^q : 0 \leq p, q \leq n - 1\}$ form a basis for $M_n(\mathbb{C})$. The subwords of $A^{n-1} B A^{n-2}$, including the empty word I , also form a basis for $M_n(\mathbb{C})$ and the length of the pair $\{A, B\}$ is $2n - 2$.*

Proof. Let $B = e \otimes f$ for nonzero vectors $e, f \in \mathbb{C}^n$. Note that $B^* = f \otimes e$ and $CB = C(e \otimes f) = e \otimes (Cf)$ and $BC = (e \otimes f)C = (C^*e) \otimes f$, for every $n \times n$ matrix C .

Note that every word in A and B with at least one B factor is equal to a scalar multiple of a word, of no greater length, with no adjacent B factors. This is because $B^2 = B(e \otimes f) = e \otimes (Bf) = e \otimes ((f|e)f) = (f|e)B$. Also, since

$$BA^p B = BA^p(e \otimes f) = e \otimes (BA^p f) = e \otimes ((A^p f|e)f) = (A^p f|e)B,$$

it follows that

- (**) every word in A and B is equal to a scalar multiple of a word, of no greater length, of the form $A^p B A^q$ with $p, q \geq 0$ or A^r with $r \geq 0$.

Consider the $n \times n$ array

$$\begin{pmatrix} B & BA & BA^2 & \dots & BA^{n-2} & BA^{n-1} \\ AB & ABA & ABA^2 & \dots & ABA^{n-2} & ABA^{n-1} \\ A^2B & A^2BA & A^2BA^2 & \dots & A^2BA^{n-2} & A^2BA^{n-1} \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ A^{n-2}B & A^{n-2}BA & A^{n-2}BA^2 & \dots & A^{n-2}BA^{n-2} & A^{n-2}BA^{n-1} \\ A^{n-1}B & A^{n-1}BA & A^{n-1}BA^2 & \dots & A^{n-1}BA^{n-2} & A^{n-1}BA^{n-1} \end{pmatrix}.$$

First notice that f must be a cyclic vector for A ; that is, $\{f, Af, A^2f, \dots, A^{n-1}f\}$ is a basis for \mathbb{C}^n . This is because the span of $\{f, Af, A^2f, \dots, A^{n-1}f\}$ is invariant under both A and B . Similarly, $\{e, A^*e, (A^*)^2e, \dots, (A^*)^{n-1}e\}$ is a basis for \mathbb{C}^n .

We next show that the matrices in the above array are linearly independent. Since there are n^2 of them, they will then form a basis for $M_n(\mathbb{C})$. Let $g_0, g_1, g_2, \dots, g_{n-1}$ be the dual basis for the basis $\{e, A^*e, (A^*)^2e, \dots, (A^*)^{n-1}e\}$, so that $(g_q | (A^*)^{q'} e) = \delta_{q,q'}$. Suppose that

$$\sum_{p'=0}^{n-1} \sum_{q'=0}^{n-1} \lambda_{(p',q')} A^{p'} B A^{q'} = 0$$

where $\{\lambda_{(p',q')} : 0 \leq p', q' \leq n - 1\}$ is a set of scalars. Then, applying the above double sum to g_q gives

$$\sum_{p'=0}^{n-1} \lambda_{(p',q)} A^{p'} f = 0$$

(using $A^{p'} B A^{q'} = ((A^*)^{q'} e) \otimes (A^{p'} f)$). Since $\{f, Af, A^2f, \dots, A^{n-1}f\}$ is linearly independent, $\lambda_{(p',q)} = 0$ for every $p' \in \{0, 1, 2, \dots, n-1\}$. But $q \in \{0, 1, 2, \dots, n-1\}$ was arbitrary, so $\lambda_{(p',q')} = 0$ for every p', q' .

The subwords of $A^{n-1}BA^{n-2}$ are those matrices in the first $n - 1$ columns of the above array, together with $I, A, A^2, \dots, A^{n-1}$. We prove by induction that each matrix in the final column of the array belongs to the span of the subwords of $A^{n-1}BA^{n-2}$.

The words of length n in A, B with precisely one B factor are those occurring in the main skew diagonal of the above array. Their sum belongs to \mathcal{V}_{n-1} (see the proof of Proposition 1), so

$$BA^{n-1} = - \sum_{p=0}^{n-2} A^{n-1-p} B A^p + V_{n-1}$$

for some $V_{n-1} \in \mathcal{V}_{n-1}$. By (**), it follows that BA^{n-1} belongs to the span of the subwords of $A^{n-1}BA^{n-2}$. Now, premultiplying by A ,

$$ABA^{n-1} = - \sum_{p=1}^{n-2} A^{n-p} B A^p + V_n$$

where $V_n = -A^n B + AV_{n-1} \in \mathcal{V}_n$. It follows from (**), and from what we have just proved that ABA^{n-1} belongs to the span of the subwords of $A^{n-1}BA^{n-2}$. Assume

that the words $A^{q-1}BA^{n-1}$ for $1 \leq q \leq n - 2$ belong to the span of the subwords of $A^{n-1}BA^{n-2}$. Then,

$$\begin{aligned} A^qBA^{n-1} &= -\sum_{p=0}^{n-2} A^{n-1-p+q}BA^p + A^qV_{n-1} \\ &= -\sum_{p=q}^{n-2} A^{n-1-p+q}BA^p + V_{n-1+q}, \end{aligned}$$

where $V_{n-1+q} \in \mathcal{V}_{n-1+q}$. It follows by the induction assumption and (***) that A^qBA^{n-1} belongs to the span of the subwords of $A^{n-1}BA^{n-2}$. Finally, $A^{n-1}BA^{n-1} \in \mathcal{V}_{2n-2}$ by Proposition 1, so it too belongs to the span of the subwords of $A^{n-1}BA^{n-2}$.

By (**), it now follows that words of length $2n - 3$ or less cannot span $M_n(\mathbb{C})$, so the length of the pair $\{A, B\}$ is $2n - 2$. □

Remark 4. By considering adjoints, the subwords of $A^{n-2}BA^{n-1}$ also form a basis for $M_n(\mathbb{C})$ if B has rank one and $\{A, B\}$ is an irreducible pair. We noted above in Proposition 2(ii) that the subwords of $A^{n-2}BA^{n-1}$ also form a basis if A is the strictly upper triangular elementary Jordan matrix and $b_{n,1} \neq 0$. In the former case, when B has rank one, the length of the pair $\{A, B\}$ is $2n - 2$. In the latter case, when $b_{n,1} \neq 0$, the length of the pair need not be $2n - 2$. A simple example of this, for the case $n = 3$, is where $B = J^* + (J^*)^2$. (Another example is Example 1 below.) Here the length of $\{A, B\}$ is 3, not 4. The matrices $B, AB, BA, I, ABA, BA^2, A, A^2$ and $BA^2 - AB^2 + AB - I$ are echelon representatives for the 9 matrix positions. Since there are only 7 words of length at most 2, it follows that the length of this pair is 3.

Corollary 2. *If A is an $n \times n$ matrix, there exists an $n \times n$ matrix B such that $\{A^pBA^q : 0 \leq p, q \leq n - 1\}$ is a basis for $M_n(\mathbb{C})$ if and only if the characteristic polynomial and minimal polynomials of A are the same.*

Proof. First, suppose that there exists such a B . Then $\{I, A, A^2, A^3, \dots, A^{n-1}\}$ is linearly independent because $\sum_{k=1}^n \lambda_k A^{k-1} = 0$ implies $\sum_{k=1}^n \lambda_k BA^{k-1} = 0$, and this implies that $\lambda_k = 0$ for $k = 1, 2, 3, \dots, n$.

Conversely, suppose that the characteristic polynomial and minimal polynomials of A are the same. Then A is cyclic and its lattice of invariant subspaces is finite and distributive [1]. Thus there is a vector f not belonging to any proper invariant subspace of A , and there is a vector e not belonging to any proper invariant subspace of A^* . Let $B = e \otimes f$. Then $\{A, B\}$ is an irreducible pair and B has rank one. The result now follows from Theorem 1. □

3. LOWER BOUNDS ON LENGTH

All of the investigations concerning the lengths of families of matrices undertaken so far have been concerned with finding the length or an upper bound for the length. We consider below the problem of finding lower bounds for the length. Of course, if $\{A, B\}$ is an irreducible pair of matrices and k is an integer, then $l(\{A, B\}) \geq k + 1$ if and only if the words in A, B of length at most k do not contain a basis for $M_n(\mathbb{C})$. We begin with an example.

Example 1. There exists an irreducible pair of 5×5 matrices with length 4.

Let J be the strictly upper triangular 5×5 elementary Jordan matrix and let B be the matrix given by

$$B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ -1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then B together with J forms an irreducible pair of length 4. We show that the words of length at most 4 span $M_5(\mathbb{C})$.

Let X be the matrix defined by $X = \frac{1}{3}(B^2JB - JB) + J^3B + J^2B$. Then

$$X = \begin{pmatrix} -1 & -1 & 2 & 1 & 1 \\ 0 & 0 & 16/3 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We show that \mathcal{V}_4 (the span of words of length at most 4) contains an (i, j) -echelon representative, for every i and j satisfying $1 \leq i, j \leq 5$.

The matrices J^pBJ^q with $0 \leq p, q \leq 3$ and $p + q \leq 3$ are echelon representatives for the strictly lower triangular positions. The matrices I, B^2, B^3, B^4, X give echelon representatives for the positions on the main diagonal (after performing the necessary row operations on their images under the row map ρ). The matrices J, B^2J, B^3J, JB^2 give echelon representatives for the positions on the first super diagonal, i.e., $(1, 2), (2, 3), (3, 4), (4, 5)$ (again after performing the necessary row operations on their images under the row map ρ). The matrices J^2, B^2J^2, J^2B^2 give echelon representatives for the positions on the second super diagonal, i.e., $(1, 3), (2, 4), (3, 5)$ (again after performing the necessary row operations on their images under the row map ρ). The matrix J^4 is an echelon representative for the position $(1, 5)$.

Finally, the matrices J^3 and Y are echelon representatives for the positions $(1, 4)$ and $(2, 5)$, respectively, where

$$Y = 528J - 479J^2 - 1050BJ - 80JB + 160J^2B - 80JBJB - 32JB^2 - 186B^3J \\ + 708B^2J + 90BJB^2 + 11JB^2J - 45J^2B^2 + 240JB^2J + 80B^2JB.$$

Indeed, $Y = -45E_{2,5}$, where $E_{2,5}$ is the usual elementary matrix.

Remark 5. No irreducible pair $\{A, B\}$ of 5×5 matrices can have length less than 4, since there are only 15 words in A and B of length 3 or less. In general, for any pair of matrices, of whatever size, the number of words of length k or less is $2^{k+1} - 1$. For every positive integer n let κ_n be the smallest natural number k satisfying $2^{k+1} - 1 \geq n^2$. Then $\kappa_1 = 0, \kappa_2 = 2, \kappa_3 = 3, \kappa_4 = \kappa_5 = 4$. If $k < \kappa_n$, no irreducible pair of size n matrices can have length k . The pair of matrices in the above example shows that an irreducible pair of matrices of size 5 can have length κ_5 . The pair $\{J, J^*\}$ where J is the strictly upper triangular elementary Jordan matrix shows that there are irreducible pairs of size $n \in \{2, 3, 4\}$ of length κ_n [6]. Every pair of 1×1 matrices is irreducible and has length κ_1 . This raises the following question.

Question 1. For every positive integer n , is there an irreducible pair of $n \times n$ complex matrices of length κ_n ?

Note that $\kappa_n \leq n$ for every $n \geq 1$. For every $n \geq 2$ there exists a pair of $n \times n$ matrices of length n , namely, $\{J, J^*\}$ where J is the strictly upper triangular elementary Jordan matrix [5].

Theorem 2. For $n \geq 2$ and $1 \leq p \leq n - 1$, the length of the irreducible pair $\{J, (J^*)^p\}$ is greater than or equal to $n + p - 1$.

Proof. We may assume that $n \geq 3$ [5]. We show that $(We_n|e_2) = (We_{n-1}|e_1)$, for every word W in $\{J, (J^*)^p\}$ of length less than $n + p - 1$. Let W be a nonempty word of length less than $n + p - 1$. Then each of $(We_n|e_2)$ and $(We_{n-1}|e_1)$ is either 0 or 1, since each of J and $(J^*)^p$ maps each e_i into the set $\{e_1, e_2, \dots, e_n, 0\}$.

Suppose that $(We_n|e_2) = 1$. Then $We_n = e_2$. Suppose that W has s J factors and t $(J^*)^p$ factors. Then $s + t \leq n + p - 2$ and $We_n = e_{n-t+sp}$. Thus $n - t + sp = 2$. But $n - t + sp \geq n - n - p + 2 + s + sp = 2 + s + sp - p$, so $s + sp - p \leq 0$. Thus $s = 0$ and $W = J^t$. Since $J^t e_n = e_2$, $t = n - 2$. Then $(We_{n-1}|e_1) = (J^{n-2}e_{n-1}|e_1) = 1$.

Suppose that $(We_{n-1}|e_1) = 1$. Then $We_{n-1} = e_1$. Again, suppose that W has s J factors and t $(J^*)^p$ factors. Then $s + t \leq n + p - 2$ and $We_{n-1} = e_{n-1-t+sp}$. Thus $n - 1 - t + sp = 1$. But $n - 1 - t + sp \geq n - 1 - n - p + 2 + s + sp = 1 + s + sp - p$, so $s + sp - p \leq 0$. Again, $s = 0$ and $W = J^t$. Since $J^t e_{n-1} = e_1$, $t = n - 2$. Then $(We_n|e_2) = (J^{n-2}e_n|e_2) = 1$. □

Remark 6. The pair $\{J, (J^*)^p\}$ need not have length $n + p - 1$. The simplest example is when $n = 4$ and $p = 2$, that is, the pair $\{J, (J^*)^2\}$ on \mathbb{C}^4 . Here $n + p - 1 = 5$. Every word $V = (v_{i,j})$ of length 5 or less, including the empty word I , satisfies $v_{1,1} - v_{2,2} + v_{3,3} - v_{4,4} = 0$, so these words cannot span $M_4(\mathbb{C})$. The length of this pair is, in fact, 6.

Example 2. We give another example of irreducible pairs of $n \times n$ matrices of length $n \geq 3$. If $B = E_{n,1} + \text{diag}(a_1, a_2, a_3, \dots, a_n)$ where $a_1 = a_n = 0$ and $a_2, a_3, a_4, \dots, a_{n-1}$ are distinct and nonzero, then the length of the irreducible pair $\{J, B\}$ is n . Here $E_{n,1}$ is the usual elementary matrix with $(n, 1)$ -entry equal to 1.

To prove that the length of the pair is no greater than n , it suffices to show that each of the n^2 matrix positions has an echelon representative which is a linear combinations of words of length at most n . Note that $B^p = \text{diag}(0, a_1^p, a_2^p, \dots, a_{n-1}^p, 0)$ for $2 \leq p \leq n - 1$. The words $\{J^p B J^q : 0 \leq p, q, p + q \leq n - 1\}$ are echelon representatives for all of the positions $\{(i, j) : i \geq j\}$. For $1 \leq p \leq n - 2$, the words $\{J^p, B^2 J^p, B^3 J^p, \dots, B^{n-p} J^p\}$ give echelon representatives for all of the positions $\{(i, j) : i - j = -p\}$ in the following way. Let $1 \leq p \leq n - 2$ and consider performing echelon row reduction on the vectors $\{\rho(J^p), \rho(B^2 J^p), \rho(B^3 J^p), \dots, \rho(B^{n-p} J^p)\}$ where ρ is the mapping $\rho : M_n(\mathbb{C}) \rightarrow \mathbb{C}^n$ described earlier. The $(n - p) \times (n - p)$ determinant

$$(*) \quad \det \begin{pmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 0 & a_1^2 & a_2^2 & \dots & a_{n-p-2}^2 & a_{n-p-1}^2 \\ 0 & a_1^3 & a_2^3 & \dots & a_{n-p-2}^3 & a_{n-p-1}^3 \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & a_1^{n-p} & a_2^{n-p} & \dots & a_{n-p-2}^{n-p} & a_{n-p-1}^{n-p} \end{pmatrix}$$

equals

$$\det \begin{pmatrix} a_1^2 & a_2^2 & \cdots & a_{n-p-2}^2 & a_{n-p-1}^2 \\ a_1^3 & a_2^3 & \cdots & a_{n-p-2}^3 & a_{n-p-1}^3 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ a_1^{n-p} & a_2^{n-p} & \cdots & a_{n-p-2}^{n-p} & a_{n-p-1}^{n-p} \end{pmatrix},$$

which in turn equals

$$a_1^2 a_2^2 \cdots a_{n-p-1}^2 \det \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ a_1 & a_2 & \cdots & a_{n-p-2} & a_{n-p-1} \\ a_1^2 & a_2^2 & \cdots & a_{n-p-2}^2 & a_{n-p-1}^2 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ a_1^{n-p-2} & a_2^{n-p-2} & \cdots & a_{n-p-2}^{n-p-2} & a_{n-p-1}^{n-p-2} \end{pmatrix}.$$

The latter $(n-p-1) \times (n-p-1)$ determinant is the classic alternant of Vandermonde. Its value is $\prod_{1 \leq i < j \leq n-p-1} (a_j - a_i)$. Since the a_i for $1 \leq i \leq n-2$ are all distinct and nonzero, it follows that the determinant $(*)$ is nonzero. So, when echelon row reduction is performed on the $n-p$ row vectors $\{\rho(J^p), \rho(B^2 J^p), \rho(B^3 J^p), \dots, \rho(B^{n-p} J^p)\}$, the matrices which give rise to the resulting $n-p$ row vectors in echelon are echelon representatives for the matrix positions $\{(i, j) : i - j = -p\}$. Each of these echelon representatives is a linear combination of words in J, B of length at most n .

Finally, J^{n-1} is an echelon representative for the position $(1, n)$.

To prove that the length of the pair $\{J, B\}$ is no less than n , we must show that the words of length $n-1$ or less do not span $M_n(\mathbb{C})$. We show that every such word has equal $(1, 1)$ - and (n, n) -entries, so every matrix in their span has this property.

It is easily proved by induction that, for every $1 \leq k \leq n-1$ and every word W of length k , $We_n \in \langle e_{n-k}, e_{n-k+1}, \dots, e_{n-1} \rangle$ and $We_1 \in \langle e_{n-k+1}, e_{n-k+2}, \dots, e_n \rangle$. It follows that $(We_n|e_n) = 0 = (We_1|e_1)$. So W has equal $(1, 1)$ - and (n, n) -entries. The empty word I also has equal $(1, 1)$ - and (n, n) -entries. This completes the proof.

Proposition 3. *If $n \geq 2$ and $\{J, B\}$ is an irreducible pair of $n \times n$ matrices with B strictly lower triangular and J the usual elementary Jordan matrix, the length of the pair $\{J, B\}$ is at least n .*

Proof. The result is obvious if $n = 2$. Suppose that $n \geq 3$. We show that every word W in J, B of length at most $n-1$ satisfies $(We_n|e_2) = (We_{n-1}|e_1)$. Then every matrix in their span will also satisfy the same condition, so the span cannot be $M_n(\mathbb{C})$.

For $i = 1, 2, \dots, n$ define subspaces \mathcal{X}_i by $\mathcal{X}_i = \langle e_i, e_{i+1}, \dots, e_n \rangle$ and put $\mathcal{X}_0 = (0)$ and $\mathcal{X}_{n+1} = \mathbb{C}^n$. Then, for $1 \leq i \leq n$ we have $J\mathcal{X}_i \subseteq \mathcal{X}_{i-1}$ and $B\mathcal{X}_i \subseteq \mathcal{X}_{i+1}$. Let $0 \leq k \leq n-1$ and let W be a word of length k . If $k = 0$, then $W = I$ and certainly $(We_n|e_2) = (We_{n-1}|e_1)$. Suppose that $k \neq 0$. If W has no B factors and if it has k J factors, then clearly $(We_n|e_2) = (We_{n-1}|e_1)$. If W has one B factor and $k-1$ J factors, then $We_n \in \mathcal{X}_{n-k+2}$ and $We_{n-1} \in \mathcal{X}_{n-k+1}$. In general, if W has $s \geq 1$ B factors and $k-1-s$ J factors, then $We_n \in \mathcal{X}_{n-k+2s}$ and $We_{n-1} \in \mathcal{X}_{n-k-1+2s}$. Since $n-k+2s \geq 3$ and $n-k-1+2s \geq 2$, we have $(We_n|e_2) = (We_{n-1}|e_1) = 0$. \square

There are undoubtedly many other results that can be obtained about lower bounds for lengths of various sets of matrices.

REFERENCES

1. L. Brickman and P. A. Fillmore, 'The invariant subspace lattice of a linear transformation', *Canad. J. Math.* **19** (1967), 810-822. MR0213378 (35:4242)
2. D. Constantine and M. Darnall, 'Lengths of finite dimensional representations of PBW algebras', *Linear Algebra Appl.* **395** (2005), 175-181. MR2112883 (2005i:17009)
3. A. Freedman, R. Gupta and R. Guralnick, 'Shirshov's theorem and representations of semi-groups', *Pacific J. Math., Special Issue* (1997), 159-176. MR1610851 (99b:20098)
4. M. S. Lambrou and W. E. Longstaff, 'On the lengths of pairs of complex matrices of size 6', *Bull. Aust. Math. Soc.* **80** (2009), 177-201. MR2540352 (2010i:15047)
5. W. E. Longstaff, 'Burnside's Theorem: Irreducible pairs of transformations', *Lin. Alg. & Applic.* **382** (2004), 247-269. MR2050111 (2005b:47159)
6. W. E. Longstaff, A. Niemeyer and Oreste Panaia, 'On the lengths of pairs of complex matrices of size at most 5', *Bull. Aust. Math. Soc.* **73** (2006), 461-472. MR2230653 (2007c:15024)
7. Christopher J. Pappacena, 'An upper bound for the length of a finite-dimensional algebra', *J. Algebra* **197** (1997), 535-545. MR1483779 (98j:16013)
8. A. Paz, 'An application of the Cayley-Hamilton theorem to matrix polynomials in several variables', *Lin. Mult. Algebra* **15** (1984), 161-170. MR740668 (85h:15019)
9. H. Radjavi and P. Rosenthal, 'Matrices for operators and generators of $\mathcal{B}(\mathcal{H})$ ', *J. London Math. Soc. (2)* **2** (1970), 557-560. MR0265978 (42:887)

SCHOOL OF MATHEMATICS & STATISTICS, THE UNIVERSITY OF WESTERN AUSTRALIA, 35 STIRLING HIGHWAY, CRAWLEY, WA 6009, AUSTRALIA
E-mail address: longstaf@maths.uwa.edu.au

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO, TORONTO, ONTARIO M5S 2E4, CANADA
E-mail address: rosent@math.toronto.edu