Bull. Aust. Math. Soc. **104** (2021), 78–93
doi:10.1017/S0004972720001227

h-MINIMUM SPANNING LENGTHS AND AN EXTENSION TO BURNSIDE'S THEOREM ON IRREDUCIBILITY

W. E. LONGSTAFF

(Received 29 July 2020; accepted 15 September 2020; first published online 2 December 2020)

Abstract

We introduce the $\mathbf{h}$-minimum spanning length of a family $\mathcal{A}$ of $n \times n$ matrices over a field $\mathbb{F}$, where $\mathbf{h}$ is a p-tuple of positive integers, each no more than n. For an algebraically closed field $\mathbb{F}$, Burnside's theorem on irreducibility is essentially that the $(n, n, \ldots, n)$-minimum spanning length of $\mathcal{A}$ exists if $\mathcal{A}$ is irreducible. We show that the $\mathbf{h}$-minimum spanning length of $\mathcal{A}$ exists for every $\mathbf{h} = (h_1, h_2, \ldots, h_p)$ if $\mathcal{A}$ is an irreducible family of invertible matrices with at least three elements. The $(1, 1, \ldots, 1)$-minimum spanning length is at most $4n \log_2 2n + 8n - 3$. Several examples are given, including one giving a complete calculation of the (p, q)-minimum spanning length of the ordered pair (J^*, J), where J is the Jordan matrix.

2020 *Mathematics subject classification*: primary 15A30.

Keywords and phrases: irreducible, linear span.

1. Introduction

Let $\mathcal{S}$ be a finite nonempty set of distinct symbols. Let W be a nonempty word in $\mathcal{S}$, written in its standard form, using index notation wherever possible. Thus, $W = x_1^{s_1} x_2^{s_2} \ldots x_{t-1}^{s_{t-1}} x_t^{s_t}$, where $\{x_i : 1 \leq i \leq t\} \subseteq \mathcal{S}$, $x_1 \neq x_2 \neq \cdots \neq x_{t-1} \neq x_t$, $s_1, s_2, \ldots, s_t$ are positive integers and $t \geq 1$. We call $x_1^{s_1}, x_2^{s_2}, \ldots, x_t^{s_t}$ the *slots* of W and t the *slot length* of W. We define the *height* of W to be $\max\{s_i : 1 \leq i \leq t\}$. We say that W is a 1-*word in* $\mathcal{S}$ if it has height 1, that is, $W = x_1 x_2 \ldots x_{t-1} x_t$, with distinct adjacent factors.

Let $\mathcal{A}$ be a finite nonempty set of distinct $n \times n$ matrices over a field $\mathbb{F}$. For every $k \in \mathbb{Z}^+$, let ${}^1\mathcal{V}_k(\mathcal{A})$, or simply ${}^1\mathcal{V}_k$, be the linear span of the 1-words in $\mathcal{A}$ of length at most k. Then ${}^1\mathcal{V}_k \subseteq {}^1\mathcal{V}_{k+1}$ for every $k \geq 1$. There exists a smallest positive integer K_1 such that ${}^1\mathcal{V}_k = {}^1\mathcal{V}_{K_1}$ for every $k \geq K_1$. If ${}^1\mathcal{V}_{K_1} = M_n(\mathbb{F})$, we call K_1 the 1-*minimum spanning length of* $\mathcal{A}$, abbreviated '1-msl($\mathcal{A}$)'. So, the 1-msl of $\mathcal{A}$ is defined if and only if the 1-words in $\mathcal{A}$ span $M_n(\mathbb{F})$ and then its value is the smallest positive integer K_1 with the property that the 1-words in $\mathcal{A}$ of length at most K_1 span $M_n(\mathbb{F})$.

With $\mathcal{A}$ and $\mathbb{F}$ as above, the 'minimum spanning length of $\mathcal{A}$', abbreviated 'msl($\mathcal{A}$)', was introduced in [6]. (The related notion of 'slot length of $\mathcal{A}$' was introduced in [8].) For every integer $k \geq 1$, we let $\mathcal{V}'_k(\mathcal{A})$ denote the linear span of the nonempty words in

$\mathcal{A}$ of length at most k. If the nonempty words in $\mathcal{A}$ span $M_n(\mathbb{F})$, the minimum spanning length of $\mathcal{A}$ is the least positive integer K such that $\mathcal{V}'_K(\mathcal{A}) = M_n(\mathbb{F})$. Clearly, msl($\mathcal{A}$) exists if 1-msl($\mathcal{A}$) does and in that case msl($\mathcal{A}$) $\leq$ 1-msl($\mathcal{A}$).

Still with $\mathcal{A}$ and $\mathbb{F}$ as above, the (earlier) notion of 'length of $\mathcal{A}$' has been considered by several authors, especially since Paz's paper [11] in 1984 (see [1, 3, 4, 7, 9, 10, 12]). For every integer $k \geq 0$, we let $\mathcal{V}_k(\mathcal{A})$ denote the linear span of the words in $\mathcal{A}$ of length at most k, taking the unique word of length zero, called the empty word, to be the identity matrix. If the (possibly empty) words in $\mathcal{A}$ span $M_n(\mathbb{F})$, the length of $\mathcal{A}$ is the least nonnegative integer L such that $\mathcal{V}_L(\mathcal{A}) = M_n(\mathbb{F})$. It is easily shown that msl($\mathcal{A}$) $\leq$ length($\mathcal{A}$) + 1 [9].

In Section 2 we present some results on the '1-msl', when the underlying field $\mathbb{F}$ is algebraically closed. With this underlying field, Burnside's theorem (see [5]) states that a family has an msl if and only if it is irreducible. An irreducible family need not have a 1-msl. For example, no irreducible pair of $n \times n$ complex matrices has a 1-msl if $n \geq 4$, since there are fewer than n^2 linearly independent 1-words in the pair (see below for more details). We shall show, in Theorem 2.1, that every finite irreducible family of invertible matrices with at least three elements, over an algebraically closed field $\mathbb{F}$, has a 1-minimum spanning length and that it is no greater than $4n \log_2 2n + 8n - 3$.

In Section 3 we introduce the more general notion of the '**h**-minimum spanning length', abbreviated '**h**-msl', of a p-tuple $(A_1, A_2, \ldots, A_p)$ of distinct $n \times n$ matrices over a field $\mathbb{F}$. Here **h** is a p-tuple of positive integers, each no greater than n, and the '$(1, 1, \ldots, 1)$-msl' is the same as the '1-msl'. Our main result is Theorem 3.3, by which we show that every finite irreducible family of invertible $n \times n$ matrices, with $p \geq 3$ elements, over an algebraically closed field $\mathbb{F}$, has an **h**-minimum spanning length for every $\mathbf{h} = (h_1, h_2, \ldots, h_p)$.

We close by calculating, in Theorem 3.5, the (p, q)-minimum spanning length of the ordered pair (J^*, J), where J is the strictly upper triangular Jordan matrix, with ones on the first superdiagonal and zeros elsewhere, for every p, q.

Recall the definition of Kronecker products [2, page 98]. If $A = (a_{i,j})$ and B are square matrices, of sizes $n \times n$ and $m \times m$, respectively, over a field $\mathbb{F}$, the Kronecker product of A and B, in that order, written $A \odot B$, is defined to be the $mn \times mn$ matrix

$$\begin{bmatrix} a_{1,1}B & a_{1,2}B & \cdots & a_{1,n-1}B & a_{1,n}B \\ a_{2,1}B & a_{2,2}B & \cdots & a_{2,n-1}B & a_{2,n}B \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n,1}B & a_{n,2}B & \cdots & a_{n,n-1}B & a_{n,n}B \end{bmatrix}.$$

We will use the following properties of Kronecker products. For $\lambda \in \mathbb{F}$ and matrices $A, C \in M_m(\mathbb{F})$, $B, D \in M_n(\mathbb{F})$:

(1) $(A \odot B)(C \odot D) = (AC) \odot (BD)$;
(2) $(\lambda A) \odot B = A \odot (\lambda B) = \lambda(A \odot B)$;
(3) $(A + C) \odot B = A \odot B + C \odot B$;
(4) $A \odot (B + D) = A \odot B + A \odot D$.

https://doi.org/10.1017/S0004972720001227 Published online by Cambridge University Press

The Kronecker product $A \odot B$ as described above can be thought of as a linear transformation on $\mathbb{F}^n \oplus \mathbb{F}^n \oplus \cdots \oplus \mathbb{F}^n$, where there are m summands.

Throughout, if $\mathcal{E}$ is a set of vectors, we will denote their linear span by $\langle \mathcal{E} \rangle$. The inner product on $\mathbb{C}^n$ is denoted by $(\cdot|\cdot)$ and the standard ordered basis is denoted by $\{e_1, e_2, \ldots, e_n\}$. If $e, f \in \mathbb{C}^n$, $e \otimes f$ denotes the linear transformation on $\mathbb{C}^n$ defined by $(e \otimes f)(x) = (x|e)f$ for $x \in \mathbb{C}^n$. Thus $e \otimes f$ is conjugate linear in the first argument and linear in the second. Also, $T(e \otimes f) = e \otimes (Tf)$ and $(e \otimes f)T = (T^* e) \otimes f$ for any linear transformation T. A square matrix A is called a square zero matrix if $A^2 = 0$.

2. 1-minimum spanning lengths

Let $\{A, B\}$ be an irreducible pair of complex $n \times n$ matrices. We mentioned above that $\{A, B\}$ does not have a 1-msl if $n \geq 4$. This is simply because, in that case, there are too few linearly independent 1-words in A, B. Indeed, for every integer $k \geq 1$ there are precisely two 1-words, namely, $ABAB\ldots$ and $BABA\ldots$, of length k. So, the number of 1-words of length at most k is $2k$. In the notation introduced above, for such a pair, ${}^1\mathcal{V}_k = {}^1\mathcal{V}_{k-1}$ whenever $k \geq 2n$, since any 1-word of length $k \geq 2n$ begins $(AB)^n \ldots$ or $(BA)^n \ldots$ and so belongs to ${}^1\mathcal{V}_{k-2}$, using the appropriate characteristic polynomial. Thus, ${}^1\mathcal{V}_{2n-1} = \langle \cup \{ {}^1\mathcal{V}_k : k \geq 1 \} \rangle$ is the linear span of the 1-words in $\{A, B\}$. But ${}^1\mathcal{V}_{2n-1}$ is spanned by $2(2n-1) = 4n-2$ elements and $4n - 2 < n^2$ for $n \geq 4$. On the other hand, the irreducible pair $\{J^*, J\}$ has 1-msl equal to 2 on $\mathbb{C}^2$, and the pair $\{D, J\}$ has 1-msl equal to 5 on $\mathbb{C}^3$, where

$$D = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 3 \end{bmatrix}.$$

In the following theorem, the condition of invertibility cannot be dropped since $\{A, B, C\}$ does not have a 1-msl if $C = 0$ and $\{A, B\}$ is an irreducible pair for $n \geq 4$.

THEOREM 2.1. *Let $\mathbb{F}$ be an algebraically closed field. If $n \geq 2$ and $p \geq 3$ and $\mathcal{F} = \{F_i : 1 \leq i \leq p\}$ is an irreducible family of distinct invertible $n \times n$ matrices over $\mathbb{F}$, the 1-words in $\mathcal{F}$ span $M_n(\mathbb{F})$ and the 1-minimum spanning length of $\mathcal{F}$ is at most $4n \log_2 2n + 8n - 3$.*

PROOF. Define the set $\mathcal{A} = \{A_i : 1 \leq i \leq p\}$ of $2n \times 2n$ matrices by

$$A_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \odot F_1, \quad A_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \odot F_2, \quad A_i = \begin{bmatrix} \alpha_i & -1 \\ \alpha_i^2 & -\alpha_i \end{bmatrix} \odot F_i \quad (3 \leq i \leq p),$$

where $\{\alpha_i : 3 \leq i \leq p\}$ is a set of distinct, nonzero scalars. Notice that

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}^2 = \begin{bmatrix} \alpha_i & -1 \\ \alpha_i^2 & -\alpha_i \end{bmatrix}^2 = 0 \quad (3 \leq i \leq p).$$

It follows that $A_i^2 = 0$ for $1 \leq i \leq p$.

We show that $\mathcal{A}$ is irreducible. Let $\mathcal{M} \subseteq \mathbb{F}^n \oplus \mathbb{F}^n$ be invariant under every element of $\mathcal{A}$. Define the subspaces $\mathcal{M}_1, \mathcal{M}_2$ by $\mathcal{M}_1 = \{x \in \mathbb{F}^n : (x, y) \in \mathcal{M}$ for some $y \in \mathbb{F}^n\}$

https://doi.org/10.1017/S0004972720001227 Published online by Cambridge University Press

and $\mathcal{M}_2 = \{y \in \mathbb{F}^n : (x, y) \in \mathcal{M}$ for some $x \in \mathbb{F}^n\}$. Then $\mathcal{M} \subseteq \mathcal{M}_1 \oplus \mathcal{M}_2$. Since

$$A_1A_2 = \begin{bmatrix} F_1F_2 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad A_2A_1 = \begin{bmatrix} 0 & 0 \\ 0 & F_2F_1 \end{bmatrix},$$

if $(x, y) \in \mathcal{M}$, both $(F_1F_2x, 0)$ and $(0, F_2F_1y)$ belong to $\mathcal{M}$. Hence, $F_1F_2\mathcal{M}_1 \subseteq \mathcal{M}_1$ and $F_2F_1\mathcal{M}_2 \subseteq \mathcal{M}_2$. Since F_1F_2 and F_2F_1 are both invertible, it follows that $F_1F_2\mathcal{M}_1 = \mathcal{M}_1$ and $F_2F_1\mathcal{M}_2 = \mathcal{M}_2$. Let $x \in \mathcal{M}_1$. Then $(F_1F_2)^{-1}x \in \mathcal{M}_1$, so $((F_1F_2)^{-1}x, v) \in \mathcal{M}$ for some $v \in \mathbb{F}^n$ and $A_1((F_1F_2)^{-1}x, v) = (x, 0) \in \mathcal{M}$. It follows that $\mathcal{M}_1 \oplus (0) \subseteq \mathcal{M}$. Similarly, $(0) \oplus \mathcal{M}_2 \subseteq \mathcal{M}$. Thus, $\mathcal{M} = \mathcal{M}_1 \oplus \mathcal{M}_2$. Now $(x, y) \in \mathcal{M}$ implies that $A_i(x, y) = (F_i(\alpha_i x - y), \alpha_i F_i(\alpha_i x - y)) \in \mathcal{M}$ for $3 \le i \le p$. Taking $y = 0$ gives $F_i\mathcal{M}_1 \subseteq \mathcal{M}_1 \cap \mathcal{M}_2$, and taking $x = 0$ gives $F_i\mathcal{M}_2 \subseteq \mathcal{M}_1 \cap \mathcal{M}_2$, for $3 \le i \le p$. Thus, using invertibility,

$$F_i\mathcal{M}_1 = \mathcal{M}_1 = \mathcal{M}_1 \cap \mathcal{M}_2 \quad \text{and} \quad F_i\mathcal{M}_2 = \mathcal{M}_2 = \mathcal{M}_1 \cap \mathcal{M}_2 \quad \text{for } 3 \le i \le p.$$

Thus, $\mathcal{M}_1 = \mathcal{M}_2 = \mathcal{N}$, say, and $\mathcal{M} = \mathcal{N} \oplus \mathcal{N}$, where the subspace $\mathcal{N}$ is invariant under every element of $\{F_i : 3 \le i \le p\}$. If $x \in \mathcal{N}$, then $(0, x) \in \mathcal{M}$ so $A_1(0, x) = (F_1x, 0) \in \mathcal{M}$, so $F_1x \in \mathcal{N}$. Similarly, $\mathcal{N}$ is invariant under F_2. Since $\mathcal{N}$ is invariant under every element of $\mathcal{F}$, $\mathcal{N} = (0)$ or $\mathbb{F}^n$. Thus, $\mathcal{M} = (0)$ or $\mathbb{F}^n \oplus \mathbb{F}^n$. This shows that $\mathcal{A}$ is irreducible.

Since every A_i has square zero, every nonzero nonempty word in $\mathcal{A}$ is a 1-word. By Burnside's theorem, the 1-words in $\mathcal{A}$ span $M_{2n}(\mathbb{F})$. If $W = X_1X_2 \ldots X_k$ is a 1-word in $\mathcal{A}$ and if $X_i = g_i \odot G_i$, where

$$g_i \in \left\{ \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha_j & -1 \\ \alpha_j^2 & -\alpha_j \end{bmatrix} : 3 \le j \le p \right\},$$

and $G_i \in \mathcal{F}$ for $1 \le i \le k$, then $W = (g_1g_2 \ldots g_k) \odot (G_1G_2 \ldots G_k)$, where $g_1g_2 \ldots g_k$ is a 2×2 matrix $B = [\beta_{ij}]$ and $G = G_1G_2 \ldots G_k$ is a 1-word in $\mathcal{F}$. So, $W = B \odot G$. Let L be the msl of $\mathcal{A}$. Then L is the 1-msl of $\mathcal{A}$ and every 2×2 matrix with entries in $M_n(\mathbb{F})$ can be expressed as a linear combination of words such as W of length at most L. It follows that the 1-words in $\mathcal{F}$ of length at most L span $M_n(\mathbb{F})$ and that the 1-msl of $\mathcal{F}$ satisfies 1-msl$(\mathcal{F}) \le L$.

Now $L = \text{msl}(\mathcal{A}) \le \text{length}(\mathcal{A}) + 1$ and $\text{length}(\mathcal{A}) \le 4n \log_2(2n) + 8n - 4$ (by [12, Theorem 3]), so $1 - \text{msl}(\mathcal{F}) \le L \le 4n \log_2(2n) + 8n - 3$. □

COROLLARY 2.2. *If $n \ge 2$ and $\{F, G\}$ is an irreducible pair of invertible complex $n \times n$ matrices, the* 1*-words in $\{I, F, G\}$ span $M_n(\mathbb{C})$. If*

$$A = \begin{bmatrix} 0 & F \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ G & 0 \end{bmatrix} \quad \textit{and} \quad C = \begin{bmatrix} I & -I \\ I & -I \end{bmatrix},$$

then $\text{msl}\{F, G\} \le 1 - \text{msl}\{I, F, G\} \le \text{msl}\{A, B, C\} \le 2\,\text{msl}\{F, G\} + 3$.

PROOF. In the theorem, take $p = 3, F_1 = F, F_2 = G, F_3 = I, \alpha_3 = 1, A_1 = A, A_2 = B$. By the theorem, $\{A, B, C\}$ is irreducible and the 1-words in $\{I, F, G\}$ span $M_n(\mathbb{C})$ and, by its proof, 1-msl$\{I, F, G\} \le \text{msl}\{A, B, C\}$. Further, $\text{msl}\{F, G\} \le$ 1-msl$\{I, F, G\}$ as remarked earlier.

https://doi.org/10.1017/S0004972720001227 Published online by Cambridge University Press

Define 2×2 matrices a, b, c by

$$a = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad c = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix},$$

so that $A = a \odot F, B = b \odot G$ and $C = c \odot I$. Let $W = E_1 E_2 \ldots E_k$ be any word (not necessarily a 1-word) in $\{F, G\}$ with $E_i = F$ or G for $1 \le i \le k$. Let $f_i = a$ or b, according as $E_i = F$ or G. Then $X_i = f_i \odot E_i = A$ or B for each i. We prove by induction that $CX_1CX_2 \ldots CX_{k-1}CX_kC = \pm(c \odot W)$ for every $k \ge 1$.

The result is true for $k = 1$ since

$$CAC = \begin{bmatrix} F & -F \\ F & -F \end{bmatrix} = c \odot F \quad \text{and} \quad CBC = \begin{bmatrix} -G & G \\ -G & G \end{bmatrix} = -(c \odot G).$$

Let $k \ge 1$ and assume that the result is true for k. Let $W = E_1 E_2 \ldots E_k E_{k+1}$, with each $E_i = F$ or G, and define $X_i = f_i \odot E_i$, where $f_i = a$ or b, so that $X_i = A$ or B. Then

$$CX_1CX_2C \ldots X_kCX_{k+1}C = \pm(c \odot (E_1E_2 \ldots E_k))X_{k+1}C.$$

Now

$$X_{k+1}C = \begin{bmatrix} F & -F \\ 0 & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 & 0 \\ G & -G \end{bmatrix},$$

according as $X_{k+1} = A$ or $X_{k+1} = B$. Thus,

$$CX_1CX_2C \ldots X_kCX_{k+1}C = \begin{cases} \pm(c \odot (E_1E_2 \ldots E_kF)) & \text{if } X_{k+1} = A, \\ \mp(c \odot (E_1E_2 \ldots E_kG)) & \text{if } X_{k+1} = B. \end{cases}$$

This completes the proof by induction.

It now follows that if W is a word in $\{F, G\}$ of length $k \ge 1$, then $c \odot W$ is in $\mathcal{V}'_{2k+1}(A, B, C)$. Consequently, $c \odot T \in \mathcal{V}'_{2k+1}(A, B, C)$ if $T \in \mathcal{V}'_k(F, G)$. But, by definition, $\mathcal{V}'_K(F, G) = M_n(\mathbb{C})$, where $K = \mathrm{msl}(F, G)$. Hence, $c \odot T \in \mathcal{V}'_{2K+1}(A, B, C)$ for every $T \in M_n(\mathbb{C})$. Now, for any $T \in M_n(\mathbb{C})$,

$$Ac \odot T = A \begin{bmatrix} T & -T \\ T & -T \end{bmatrix} = \begin{bmatrix} FT & -FT \\ 0 & 0 \end{bmatrix} \in \mathcal{V}'_{2K+2}(A, B, C),$$

$$A \begin{bmatrix} -G & G \\ -G & G \end{bmatrix} B = \begin{bmatrix} -FTG & 0 \\ 0 & 0 \end{bmatrix} \in \mathcal{V}'_{2K+3}(A, B, C).$$

Since F, G are invertible,

$$\begin{bmatrix} T & S \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ R & Q \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} T & S \\ R & Q \end{bmatrix} \in \mathcal{V}'_{2K+3}(A, B, C)$$

for every $Q, R, S, T \in M_n(\mathbb{C})$. The first assertion follows from the preceding calculations, a similar argument gives the second and together they imply the third. So, the nonempty words in $\{A, B, C\}$ of length at most $2K + 3$ span $M_{2n}(\mathbb{C})$ and it follows that $L = \mathrm{msl}(A, B, C) \le 2K + 3$. This completes the proof. □

https://doi.org/10.1017/S0004972720001227 Published online by Cambridge University Press

The proof of the preceding theorem shows how to obtain examples of irreducible triples consisting of square zero $k \times k$ complex matrices, where $k \geq 4$ is even. It is perhaps interesting to note that three is minimal in this regard, even on odd-dimensional spaces.

PROPOSITION 2.3. *If $n \geq 3$ and Q is an irreducible, linearly independent set of square zero $n \times n$ complex matrices with q elements, then $3 \leq q \leq n^2 - 1$, where both inequalities are sharp, for every $n \geq 3$.*

PROOF. Let $n \geq 3$ and let Q be an irreducible, linearly independent set of $n \times n$ square zero matrices with q elements. Obviously, $q \geq 2$. Suppose that $q = 2$ and $Q = \{A, B\}$ with $A^2 = B^2 = 0$. Then AB maps the range of A into itself, so there exists f in range(A) such that $ABf = \lambda f$ for some scalar λ. Then $\langle\{f, Bf\}\rangle$ is invariant under both A and B. This contradicts the irreducibility of Q, so $q \geq 3$.

Since every nilpotent matrix has trace zero, every matrix in the linear span of Q has trace zero. Hence, $q \leq n^2 - 1$.

Next we show that we can have $q = n^2 - 1$. (We may take, additionally, $n = 2$ in what follows.) Let

$$Q_1 = \{e_i \otimes e_j : 1 \leq i, j \leq n,\ i \neq j\} \cup \{(e_1 + e_j) \otimes (e_1 - e_j) : 2 \leq j \leq n\}.$$

Then Q_1 has $n^2 - 1$ elements, each with square zero. Now

$$(e_1 + e_j) \otimes (e_1 - e_j) = (e_1 \otimes e_1) - (e_1 \otimes e_j) + (e_j \otimes e_1) - (e_j \otimes e_j) \quad \text{for } 2 \leq j \leq n.$$

Since $e_1 \otimes e_j$ and $e_j \otimes e_1$ both belong to Q_1 if $j \neq 1$, the linear span of $\{e_1 \otimes e_1\} \cup Q_1$ contains every $e_j \otimes e_j$ and so equals $M_n(\mathbb{C})$. Thus, Q_1 is linearly independent, since its linear span has dimension $n^2 - 1$. Let $R = (e_1 \otimes e_2)((e_1 + e_2) \otimes (e_1 - e_2))$. Since $R = (e_1 \otimes e_2) + (e_2 \otimes e_2)$ does not have trace zero, it does not belong to the linear span of Q_1. Thus, $\langle Q_1 \cup \{R\}\rangle = M_n(\mathbb{C})$ and it follows that Q_1 is irreducible.

Finally, we show that an irreducible set of square zero matrices with three elements exists in $M_n(\mathbb{C})$. Such a set is obviously linearly independent.

Case: $n = 2m$ even, $m \geq 2$. Let $\{F, G\}$ be an irreducible pair of $m \times m$ complex matrices. Put

$$A = \begin{bmatrix} 0 & F \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ G & 0 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} I & -I \\ I & -I \end{bmatrix}.$$

Clearly $A^2 = B^2 = C^2 = 0$ and, as noted in the proof of Corollary 2.2, $\{A, B, C\}$ is irreducible.

Case: $n = 3$. $\left\{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}\right\}$ is irreducible.

Case: $n = 2m + 1$ odd, $m \geq 2$. Let $E \in M_m(\mathbb{C})$ be an invertible matrix for which there exist vectors e, f in $\mathbb{C}^m$ satisfying (i) e is a cyclic vector for E, (ii) $(E^{-1}e \mid f) \neq 0$, (iii) $N \subseteq \langle f\rangle^{\perp}, N$ an invariant subspace of E, implies that $N = (0)$. (For example, we can take $E = I - J$, where J is the upper triangular Jordan matrix and $e = (1, 1, \ldots, 1)$,

https://doi.org/10.1017/S0004972720001227 Published online by Cambridge University Press

$f = (1,0,0,\dots,0)$.) Define $(2m+1)\times(2m+1)$ matrices A, B, C by

$$A = \begin{bmatrix} 0 & 0 & 0 \\ (\cdot|f) & 0 & 0 \\ E & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & E \\ 0 & 0 & (\cdot|f) \\ 0 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & (\cdot)e & 0 \\ 0 & 0 & 0 \\ 0 & (\cdot)e & 0 \end{bmatrix},$$

where $(\cdot|f) : \mathbb{C}^m \to \mathbb{C}$ is the linear map $(\cdot|f)(x) = (x|f)$ and $(\cdot)(e) : \mathbb{C} \to \mathbb{C}^m$ is the linear map $(\cdot)(e)(\lambda) = \lambda e$. Then

$$A(x,\lambda,y) = (0,(x|f),Ex), \quad B(x,\lambda,y) = (Ey,(y|f),0), \quad C(x,\lambda,y) = (\lambda e,0,\lambda e) \tag{2.1}$$

for all $x, y \in \mathbb{C}^m$, $\lambda \in \mathbb{C}$. Clearly $A^2 = B^2 = C^2 = 0$.

Let $\mathcal{M} \subseteq \mathbb{C}^m \oplus \mathbb{C} \oplus \mathbb{C}^m$ be invariant under A, B and C. The matrix C has rank one. Indeed, $C = (0,1,0) \otimes (e,0,e)$. Thus, either $(e,0,e) \in \mathcal{M}$ or $\mathcal{M} \subseteq \langle (0,1,0)\rangle^{\perp}$.

Suppose that $(e,0,e) \in \mathcal{M}$. Let $x, y \in \mathbb{C}^m$. We prove by induction that if $W = X_1X_2\dots X_k$ is a 1-word in $\{A,B\}$, so that $X_i = A$ or B, and $X_1 \neq X_2 \neq \cdots \neq X_{k-1} \neq X_k$ and $k \geq 1$, then

$$W(x,0,y) = \begin{cases} (0,(E^{k-1}y|f),E^k y) & \text{if } k \text{ is even and } W \text{ begins with } A, \\ (0,(E^{k-1}x|f),E^k x) & \text{if } k \text{ is odd and } W \text{ begins with } A, \\ (E^k x,(E^{k-1}x|f),0) & \text{if } k \text{ is even and } W \text{ begins with } B, \\ (E^k y,(E^{k-1}y|f),0) & \text{if } k \text{ is odd and } W \text{ begins with } B. \end{cases} \tag{2.2}$$

The result is true when $k = 1$, using (2.1). Let $k \geq 1$ and suppose that the result is true for k. Let $V = X_1X_2\dots X_kX_{k+1}$ be a 1-word in $\{A,B\}$. Let $W = X_2X_3\dots X_{k+1}$.

Case: $k+1$ even and V begins with A. Then k is odd and W begins with B, so

$$V(x,0,y) = A(E^k y,(E^{k-1}y|f),0) = (0,(E^k y|f),E^{k+1}y).$$

Case: $k+1$ odd and V begins with B. Then k is even and W begins with A, so

$$V(x,0,y) = B(0,(E^{k-1}y|f),E^k y) = (E^{k+1}y,(E^k y|f),0).$$

Case: $k+1$ even and V begins with B. Then k is odd and W begins with A, so

$$V(x,0,y) = B(0,(E^{k-1}x|f),E^k x) = (E^{k+1}x,(E^k x|f),0).$$

Case: $k+1$ odd and V begins with A. Then k is even and W begins with B, so

$$V(x,0,y) = A(E^k x,(E^{k-1}x|f),0) = (0,(E^k x|f),E^{k+1}x).$$

Thus, the result is true for $k+1$. We next show that the set of vectors

$$\mathcal{W} = \{(e,0,e)\} \cup \{W(e,0,e) : W \text{ is a 1-word in } \{A,B\} \text{ of length } 1 \leq k \leq m\}$$

is linearly independent. Since this set of vectors has $2m+1$ elements and is contained in $\mathcal{M}$, it spans $\mathbb{C}^m \oplus \mathbb{C} \oplus \mathbb{C}^m$ and $\mathcal{M} = \mathbb{C}^m \oplus \mathbb{C} \oplus \mathbb{C}^m$. Thus,

$$W(e,0,e) = \begin{cases} (0,(E^{k-1}e|f),E^k e) & \text{if } W \text{ begins with } A, \\ (E^k e,(E^{k-1}e|f),0) & \text{if } W \text{ begins with } B, \end{cases}$$

https://doi.org/10.1017/S0004972720001227 Published online by Cambridge University Press

where $W = X_1X_2\dots X_k$ is a 1-word in $\{A, B\}$. Then

$$\mathcal{W} = \{(e, 0, e)\} \cup \{(0, (E^{k-1}e \mid f), E^k e), (E^k e, (E^{k-1}e \mid f), 0) : 1 \le k \le m\}.$$

Suppose that

$$\alpha(e, 0, e) + \sum_{k=1}^{m} \beta_k(0, (E^{k-1}e \mid f), E^k e) + \sum_{k=1}^{m} \gamma_k(E^k e, (E^{k-1}e \mid f), 0) = 0.$$

Then $\alpha e = -\sum_{k=1}^{m} \beta_k E^k e = -\sum_{k=1}^{m} \gamma_k E^k e$, so

$$2\alpha e = -\sum_{k=1}^{m} (\beta_k + \gamma_k) E^k e \quad \text{and} \quad 2\alpha E^{-1} e = -\sum_{k=1}^{m} (\beta_k + \gamma_k) E^{k-1} e.$$

Taking the inner product with f gives

$$2\alpha(E^{-1}e \mid f) = -\sum_{k=1}^{m} (\beta_k + \gamma_k)(E^{k-1}e \mid f) = 0.$$

Thus, $\alpha = 0$ and $\sum_{k=1}^{m} \beta_k E^{k-1} e = \sum_{k=1}^{m} \gamma_k E^{k-1} e = 0$. From the latter, $\beta_k = \gamma_k = 0$, for $1 \le k \le m$, since e is a cyclic vector for E. So, $\mathcal{M} = \mathbb{C}^m \oplus \mathbb{C} \oplus \mathbb{C}^m$.

Finally, suppose that $\mathcal{M} \subseteq \langle (0, 1, 0) \rangle^{\perp}$. If $(x, \lambda, y) \in \mathcal{M}$, it follows that $\lambda = 0$. If $(x, 0, y) \in \mathcal{M}$, then, from (2.2), the matrix E has cyclic invariant subspaces $M = \langle x, Ex, E^2x, \dots, E^{m-1}x \rangle$ and $N = \langle y, Ey, E^2y, \dots, E^{m-1}y \rangle$ and they satisfy $M, N \subseteq \langle f \rangle^{\perp}$, so $M = N = (0)$ and $x = y = 0$. Thus, $\mathcal{M} = (0)$. This completes the proof. □

REMARK 2.4. Let $n \ge 2$ and let $\{A, B\}$ be an irreducible pair of $n \times n$ complex matrices. Note that the 1-words in $\{I, A, B\}$ will span $M_n(\mathbb{C})$, even if neither A nor B is invertible. For, if W is a nonempty word in $\{A, B\}$ of length t, then W is equal to a word in $\{I, A, B\}$ of length $2t - s$, where s is the slot length of W. For example, $A^2BAB^3A^2 = AIABABIBIBAIA$. It follows that 1-msl$(\{I, A, B\}) \le 2\,msl(\{A, B\}) - 1$. So, if msl$(\{A, B\}) \le 2n - 2$, then 1-msl$(\{I, A, B\}) \le 4n - 5$. We advance the conjecture that 1-msl$(\{A, B\}) \le 4n - 7$ for any irreducible pair $\{A, B\}$. In support of this we give the following proposition.

PROPOSITION 2.5. *If $n \ge 2$ and $\{A, B\}$ is an irreducible pair of $n \times n$ complex matrices with* msl$\{A, B\} \le 2n - 2$, *then* 1-msl$(\{I, A, B\}) \le 4n - 7$ *provided that*

$$\{A^{n-1}B^{n-1}, B^{n-1}A^{n-1}\} \subseteq \mathcal{V}'_{2n-3}(\{A, B\}) + \langle \mathcal{W}''_{2n-2}(\{A, B\}) \rangle, \tag{2.3}$$

where $\mathcal{V}'_{2n-3}(\{A, B\})$ denotes the linear span of the nonempty words of length at most $2n - 3$ in $\{A, B\}$ and $\mathcal{W}''_{2n-2}(\{A, B\})$ denotes the set of nonempty words in $\{A, B\}$ of length $2n - 2$, different from both $A^{n-1}B^{n-1}$ and $B^{n-1}A^{n-1}$.

PROOF. Let $\{A, B\}$ be an irreducible pair with msl$(\{A, B\}) \le 2n - 2$ satisfying condition (2.3). We need to show that the 1-words in I, A, B of length at most $4n - 7$ span $M_n(\mathbb{C})$. Since the nonempty words in $\{A, B\}$ of length at most $2n - 2$ span $M_n(\mathbb{C})$, it is enough to show that every such word W belongs to the span of the 1-words in $\{I, A, B\}$ of length at most $4n - 7$. Denote the length of W by $|W|$ and the slot length of W

https://doi.org/10.1017/S0004972720001227 Published online by Cambridge University Press

by $s(W)$ and suppose that $s(W) \geq 3$. Then $W = V_1$, where V_1 is a 1-word in $\{I, A, B\}$ of length $2|W| - s(W)$. But $2|W| - s(W) \leq 2(2n-2) - 3 = 4n - 7$, giving the result in this case.

Suppose that W has slot length 1. Then $W = A^l$ or B^l, where $l = |W|$. Each of these belongs to $\langle\{I, A, A^2, \ldots, A^{n-1}, B, B^2, \ldots, B^{n-1}\}\rangle$, so it is enough to consider the case where $|W| = A^p$ or B^p, where $1 \leq p \leq n-1$. In this case $W = V_2$, where V_2 is a 1-word in $\{I, A, B\}$ of length $2p - 1 \leq 2n - 3 \leq 4n - 7$, again giving the desired result.

Finally, suppose that W has slot length two. We have $W = A^pB^q$ or B^qA^p, with $p, q \geq 1$ and $|W| = p + q \leq 2n - 2$. It is enough to consider the case where $p, q \leq n - 1$. If $p + q \leq 2n - 3$, we observe that $W = V_3$, where V_3 is a 1-word in $\{I, A, B\}$ of length $2|W| - 2 \leq 4n - 8$, and the desired result follows. If $p + q = 2n - 2$, then $p = q = n - 1$. By (2.3), $W \in \mathcal{V}'_{2n-3}(\{A, B\}) + \langle\mathcal{W}''_{2n-2}(\{A, B\})\rangle$ and every element of the latter, as we have just shown, belongs to the span of the 1-words in $\{I, A, B\}$ of length at most $4n - 7$. This completes the proof. □

COROLLARY 2.6. *Let $n \geq 2$ and let $\{A, B\}$ be an irreducible pair of $n \times n$ complex matrices with A unicellular. If the degree of the minimum polynomial of B is less than n, then* 1-msl$(\{I, A, B\}) \leq 4n - 7$.

PROOF. By [6, Theorem 2], msl$(\{A, B\}) \leq 2n - 2$. The condition (2.3) of the proposition obviously holds, so the result follows. □

EXAMPLE 2.7. Over the complex field, consider the pair $\{A, B\}$, where $B = J$, the upper-triangular $n \times n$ Jordan matrix, $n \geq 3$ and $A = e_1 \otimes e_n$. Then $A^2 = 0$ and $\{A, B\}$ is irreducible. (The nonzero invariant subspaces of J are the subspaces $\langle\{e_1, e_2, \ldots, e_k\}\rangle$ for $1 \leq k \leq n$.) By the preceding corollary, 1-msl$(\{I, A, B\}) \leq 4n - 7$. We show that, in fact, 1-msl$(\{I, A, B\}) = 4n - 7$. For this it is enough to show that $(We_n|e_2) = (We_{n-1}|e_1)$ for every 1-word W in $\{I, A, B\}$ of length at most $4n - 8$. Deleting the I factors in W, there remains a word in $\{A, B\}$ which is equal to a scalar multiple of one of the following types of words: $I, A, B^pA, AB^q, B^pAB^q$, where $1 \leq p, q \leq n - 1$. Here we have used the fact that A has rank one, so for any matrix X we have $AXA = \lambda A$ for some scalar λ. Each of I, A, B^pA, AB^q satisfies the aforementioned condition. The only time $(B^pAB^qe_n|e_2) \neq (B^pAB^qe_{n-1}|e_1)$ is when $p + q = 2n - 3$. If $W = \gamma J^pBJ^q$ for some scalar γ and $p + q = 2n - 3$, then the length of W as a 1-word in $\{I, A, B\}$ would be at least $2(2n - 2) - 3 = 4n - 7$. Thus, $(We_n|e_2) = (We_{n-1}|e_1)$ always holds.

3. *h*-minimum spanning lengths

We close by introducing, and briefly investigating, the more general notion of '**h**-minimum spanning length', where **h** is a p-tuple of positive integers.

Let $p \in \mathbb{Z}^+$ and let $\{m_i : 1 \leq i \leq p\}$ be a set of positive integers. Let $\mathbf{m} = (m_1, m_2, \ldots, m_p)$ and let $\mathcal{H}(\mathbf{m})$ be the set of integer p-tuples $(h_1, h_2, \ldots, h_p)$ satisfying $0 \leq h_i \leq m_i$, $1 \leq i \leq p$. The set $\mathcal{H}(\mathbf{m})$ is partially ordered by defining $\mathbf{r} \leq \mathbf{s}$ if $r_i \leq s_i$ for $1 \leq i \leq p$, where $\mathbf{r} = (r_1, r_2, \ldots, r_p)$ and $\mathbf{s} = (s_1, s_2, \ldots, s_p)$. With this partial order,

https://doi.org/10.1017/S0004972720001227 Published online by Cambridge University Press

$\mathcal{H}(\mathbf{m})$ is a lattice with least element $\mathbf{0} = (0, 0, \ldots, 0)$ and greatest element $\mathbf{m}$. The element $(1, 1, \ldots, 1)$ is denoted by $\mathbf{1}$.

Let $\mathcal{S} = \{a_1, a_2, \ldots, a_p\}$ be a set of distinct symbols and let $\mathbf{a} = (a_1, a_2, \ldots, a_p)$. Let W be a nonempty word in $\mathcal{S}$, written in its standard form, using index notation wherever possible. Thus, $W = x_1^{s_1} x_2^{s_2} \ldots x_{t-1}^{s_{t-1}} x_t^{s_t}$, where $\{x_i : 1 \leq i \leq t\} \subseteq \mathcal{S}$, $x_1 \neq x_2 \neq \cdots \neq x_{t-1} \neq x_t$, $s_1, s_2, \ldots, s_t$ are positive integers and $t \geq 1$. If $a \in \mathcal{S}$, we define the *height of a in W* to be 0 if $a \notin \{x_1, x_2, \ldots, x_t\}$ and $\max\{s_i : a = x_i, 1 \leq i \leq t\}$ otherwise. Define the *height* of W to be $\max\{s_i : 1 \leq i \leq t\}$. For example, if $\mathcal{S} = \{a, b, c, d, e, f, g\}$ and $V = ab^3c^5b^4d^2a^3$, the heights of a, b, c, d, e, f, g in V are, respectively, $3, 4, 5, 2, 0, 0, 0$. Returning to generalities, with W as above, the *height profile* of W in $\mathbf{a}$ is the p-tuple $(h_1, h_2, \ldots, h_p)$, where h_i is the height of a_i in W. For example, the height profile of the word V above, in (a, b, c, d, e, f, g), is $(3, 4, 5, 2, 0, 0, 0)$.

DEFINITION 3.1. Let $\mathcal{A} = \{A_1, A_2, \ldots, A_p\}$ be a set of distinct $n \times n$ matrices over a field $\mathbb{F}$, where $n, p \geq 1$. Let $\mathbf{A}$ be the p-tuple $\mathbf{A} = (A_1, A_2, \ldots, A_p)$. Put $\mathbf{n} = (n, n, \ldots, n)$ and, for every $\mathbf{h} \in \mathcal{H}(\mathbf{n})$, $\mathbf{h} \geq \mathbf{1}$, call a nonempty word in $\mathcal{A}$ an $\mathbf{h}$-word of $\mathbf{A}$ if its height profile is at most $\mathbf{h}$ (in the poset $(\mathcal{H}(\mathbf{n}), \leq)$). For $k \in \mathbb{Z}^+$, let $\mathcal{V}'_k(\mathbf{A}; \mathbf{h})$, or simply $\mathcal{V}'_k(\mathbf{h})$, be the linear span of the $\mathbf{h}$-words of $\mathbf{A}$ of length at most k. Then $\mathcal{V}'_k(\mathbf{h}) \subseteq \mathcal{V}'_{k+1}(\mathbf{h})$ for every $k \geq 1$. There exists a smallest positive integer $K_{\mathbf{h}}$ such that $\mathcal{V}'_k(\mathbf{h}) = \mathcal{V}'_{K_{\mathbf{h}}}(\mathbf{h})$ for every $k \geq K_{\mathbf{h}}$. If $\mathcal{V}'_{K_{\mathbf{h}}}(\mathbf{h}) = M_n(\mathbb{F})$, we call $K_{\mathbf{h}}$ the $\mathbf{h}$-*minimum spanning length of* $\mathbf{A}$, abbreviated '$\mathbf{h}$-msl($\mathbf{A}$)'.

REMARK 3.2. We make some remarks on the preceding definition.

(1) If $W = x_1^{s_1} x_2^{s_2} \ldots x_{t-1}^{s_{t-1}} x_t^{s_t}$, then W is an $\mathbf{h}$-word, where $\mathbf{h} = (h_1, h_2, \ldots, h_p)$ if a_1 has height at most h_1 in W, a_2 has height at most h_2 in $W, \ldots, a_p$ has height at most h_p in W. Any, but not all, of these heights in W may be zero.
(2) The $\mathbf{h}$-msl of $\mathbf{A}$ is defined if and only if the $\mathbf{h}$-words of $\mathbf{A}$ (in $\mathcal{A}$) span $M_n(\mathbb{F})$ and then its value is the smallest positive integer $K_{\mathbf{h}}$ with the property that the $\mathbf{h}$-words of $\mathbf{A}$ (in $\mathcal{A}$) of length at most $K_{\mathbf{h}}$ span $M_n(\mathbb{F})$.
(3) In the earlier notation and terminology, '1-word $\equiv$ $\mathbf{1}$-word', '${}^1\mathcal{V}_k(\mathcal{A}) \equiv \mathcal{V}'_k(\mathcal{A}; \mathbf{1})$'.
(4) In defining the $\mathbf{h}$-msl of $\mathbf{A}$, the order of the elements of $\mathcal{A}$ matters. For example, if V is the word in (a, b, c, d, e, f, g) given by $V = ab^3c^5b^4d^2a^3$, the height profile of V in (d, e, g, b, f, c, a) is $(2, 0, 0, 4, 0, 5, 3)$. In general, if $\mathbf{A} = (A_1, A_2, \ldots, A_p)$ and σ is a permutation of $\{1, 2, \ldots, p\}$, the height profile of a word W in $\mathbf{A}^\sigma = (A_{\sigma(1)}, A_{\sigma(2)}, \ldots, A_{\sigma(p)})$ is $\mathbf{h}^\sigma = (h_{\sigma(1)}, h_{\sigma(2)}, \ldots, h_{\sigma(p)})$ if $(h_1, h_2, \ldots, h_p)$ is its height profile in $\mathbf{A}$. Consequently, $\mathbf{h}$-msl($\mathbf{A}$) exists if and only if $\mathbf{h}^\sigma(\mathbf{A}^\sigma)$ does, in which case they are equal. In the special cases where $\mathbf{h} = (h, \ldots, h)$, where $1 \leq h \leq n$, the set of $\mathbf{h}$-words is independent of the order in which the elements of $\mathcal{A}$ are taken. We define the set of h-words in $\mathcal{A}$ as those words in $\mathcal{A}$ in which the height of each element of $\mathcal{A}$ is at most h. We also define the *h-minimum spanning length* of $\mathcal{A}$ to be $\mathbf{h}$-msl($\mathbf{A}$), where $\mathcal{A}$ is any p-tuple enumerating $\mathcal{A}$.
(5) If $\mathbf{r}, \mathbf{s} \in \mathcal{H}(\mathbf{n})$ and $\mathbf{1} \leq \mathbf{r} \leq \mathbf{s}$, then, for any given $\mathbf{A}$, $\mathcal{V}'_k(\mathbf{r}) \subseteq \mathcal{V}'_k(\mathbf{s})$ for every positive integer k. Thus, if $\mathbf{r}$-msl($\mathbf{A}$) exists, so does $\mathbf{s}$-msl($\mathbf{A}$) and, moreover,

https://doi.org/10.1017/S0004972720001227 Published online by Cambridge University Press

s-msl(**A**) ≤ **r**-msl(**A**). In particular, if the 1-msl exists, then the **h**-msl exists, for $\mathbf{1} \leq \mathbf{h} \leq \mathbf{n}$, and 1-msl≤ **h**-msl.

(6) For every $k \geq 1$, $\mathcal{V}_k'(\mathbf{n})$ is the linear span of all the nonempty words in $\mathcal{A}$ of length at most k (by the Cayley–Hamilton theorem). Thus, the **n**-msl exists if and only if the msl exists, in which case they are equal. More precisely, if m_i is the degree of the minimum polynomial of A_i (as in the preceding definition), for $1 \leq i \leq p$, and $\mathbf{m} = (m_1, m_2, \ldots, m_p)$, then, for $\mathbf{A} = (A_1, A_2, \ldots, A_p)$, $\mathcal{V}_k'(\mathbf{n}) = \mathcal{V}_k'(\mathbf{m})$ for every $k \in \mathbb{Z}^+$, so the msl exists if and only if the **m**-msl exists, in which case they are equal. Thus, if the 1-msl exists, then the h-msl exists for $1 \leq h \leq n$ and

$$\mathrm{msl}(\mathbf{A}) \leq (n-1)\text{-msl}(\mathbf{A}) \leq \cdots \leq h\text{-msl}(\mathbf{A}) \leq \cdots \leq 2\text{-msl}(\mathbf{A}) \leq 1\text{-msl}(\mathbf{A}).$$

THEOREM 3.3. *Let $\mathbb{F}$ be an algebraically closed field. Let $n, p \in \mathbb{Z}^+$ with $n \geq 2,\ p \geq 3$ and let $\boldsymbol{h} = (h_1, h_2, \ldots, h_p)$ be a p-tuple of integers satisfying $1 \leq h_i \leq n, 1 \leq i \leq p$. If $\boldsymbol{F} = (F_1, F_2, \ldots, F_p)$, where $\mathcal{F} = \{F_i : 1 \leq i \leq p\}$ is an irreducible family of distinct invertible $n \times n$ matrices over $\mathbb{F}$, the $\boldsymbol{h}$-words of $\boldsymbol{F}$ in $\mathcal{F}$ span $M_n(\mathbb{F})$, so the $\boldsymbol{h}$-minimum spanning length of $\boldsymbol{F}$ exists.*

PROOF. This follows from Theorem 2.1 and the preceding remarks. □

In the remainder of this paper, the underlying field will be the complex field $\mathbb{C}$. In Example 2.7, we showed that $1\text{-msl}(\{I, A, B\}) = 4n - 7$, where $A = e_1 \otimes e_n$ and $B = J$. By the next result, $2\text{-msl}(\{I, A, B\}) \leq 3n - 3$, since $\mathrm{msl}\{A, B\} = 2n - 2$ [6, Example 2].

PROPOSITION 3.4. *Let $\mathcal{E} = \{E_1, E_2, \ldots, E_p\}$ be a finite, irreducible family of distinct complex $n \times n$ matrices with $n \geq 2$ and suppose that $I \notin \mathcal{E}$. Then, for $1 \leq h \leq n$,*

$$h\text{-msl}(\{I\} \cup \mathcal{E}) \leq \Big(1 + \frac{1}{h}\Big)\mathrm{msl}(\mathcal{E}).$$

PROOF. Let $K = \mathrm{msl}(\mathcal{E})$. Let $W = E_1^{p_1} E_2^{p_2} \ldots E_t^{p_t}$ be a nonempty word in $\mathcal{E}$ such that $E_i \in \mathcal{E}, p_i \in \mathbb{Z}^+,\ 1 \leq i \leq t$ and $E_1 \neq E_2 \neq \cdots \neq E_t$. For each i, let $p_i = hm_i + r_i$, where $m_i \in \mathbf{N}$ and $0 \leq r_i \leq h - 1$. For $1 \leq i \leq t$, define ${}^1W_i$ by

$${}^1W_i = \begin{cases} ((E_i)^h I)^{m_i - 1} E_i^h & \text{if } r_i = 0, \\ ((E_i)^h I)^{m_i} E_i^{r_i} & \text{if } r_i \neq 0. \end{cases}$$

As a word in $\{I\} \cup \mathcal{E}$, each ${}^1W_i$ has height at most h and the length of ${}^1W_i$ is at most $p_i + m_i$. Now $W = {}^1W_1\,{}^1W_2 \ldots {}^1W_t$, where the right-hand side is a word in $\{I\} \cup \mathcal{E}$ of height at most h and length at most $\sum_{i=1}^{t}(p_i + m_i)$. Since each $p_i \geq hm_i$, we have written W as a word in $\{I\} \cup \mathcal{E}$ of height at most h and length at most $(1 + 1/h)$ length(W), where length(W) is the length of W as a word in $\mathcal{E}$. Since the words in $\mathcal{E}$ of length at most K span $M_n(\mathbb{C})$, it follows that the words in $\{I\} \cup \mathcal{E}$ of height at most h and length at most $(1 + 1/h)K$ also span $M_n(\mathbb{C})$. Hence, the result follows. □

We finish by describing all of the possible **h**-minimum spanning lengths for certain ordered pairs (A, B). In the following, for any real number x, $\lceil x \rceil$ denotes the smallest integer greater than or equal to x.

THEOREM 3.5. *Let $n \geq 4$ and let J be the strictly upper-triangular $n \times n$ (complex) Jordan matrix. If $p, q \in \mathbb{Z}^+$, the (p, q)-minimum spanning length of the ordered pair (J^*, J) exists if and only if $p, q \geq 2$, in which case its value is $n - 2 + 2\lceil (n-2)/(r-1) \rceil$, where $r = \min\{p, q\}$.*

PROOF. As we have mentioned earlier, no ordered pair has a $(1, 1)$-msl. For any ordered pair (A, B), the adjoint of a (p, q)-word of length k in (A, B) is a (q, p)-word of length k in (B^*, A^*). Thus, $(\mathcal{V}'_k((A, B); (p, q)))^* = \mathcal{V}'_k((B^*, A^*); (q, p))$ and the (p, q)-msl of (A, B) exists if and only if the (q, p)-msl of (B^*, A^*) exists, in which case they are equal.

Put $A = J^*$ and $B = J$ and consider only the ordered pair (A, B) in the remainder of the proof. Then $A^* = B$ and, by what we have just noted, for any positive integers p, q, the (p, q)-msl (of (A, B)) exists if and only if the (q, p)-msl does, in which case they are equal. If $p \geq 2$, the $(p, 1)$-msl of (A, B) does not exist since $(We_n|e_1) = 0$ for every $(p, 1)$-word. For, suppose that $We_n \neq 0$ or e_{n-1}. Then W must equal $(AB)^m$ for some $m \geq 1$, since it must end in a B and begin with an A, so $We_n = e_n$. Since $We_n = 0$ or e_{n-1} or e_n, we have $(We_n|e_1) = 0$. It follows that the $(p, 1)$-words cannot span $M_n(\mathbb{C})$. This shows that the (p, q)-msl exists only if $p, q \geq 2$.

We continue by proving by induction that, for any $p \geq 2$,

$$A^{(m-1)(p-1)+2} = A^2(BA^p)^{(m-1)} \quad \text{for all } m \geq 2. \tag{3.1}$$

This is true for $m = 2$ since $A = ABA$ gives $A^{p+1} = A(ABA)A^{p-1} = A^2(BA^p)$. Let $m \geq 2$ and suppose that the result holds for m. Then

$$A^{r(p-1)+2} = A^{p-1}A^2(BA^p)^{r-1} = A^{p+1}(BA^p)^{r-1} = A^2(BA^p)^r.$$

We next show that, for every $p \geq 2$,

$$A^k = A^{x_k}(BA^p)^{(r_k-1)} \quad \text{for every } k \geq p+1, \tag{3.2}$$

where $x_k = k - (r_k - 1)(p - 1)$ and $r_k = \lceil (k-1)/(p-1) \rceil$. Let $k \geq p + 1$. Then $r_k - 1 < (k-1)/(p-1) \leq r_k$, so $(r_k - 1)(p - 1) < k - 1$, and $0 < x_k - 1 \leq p - 1$, so $2 \leq x_k \leq p$. Then $k = x_k + (r_k - 1)(p - 1)$, so

$$A^k = A^{x_k-2}A^{(r_k-1)(p-1)+2} = A^{x_k-2}A^2(BA^p)^{(r_k-1)} = A^{x_k}(BA^p)^{r_k-1}$$

using (3.1). By symmetry, since $BAB = B$, we get a similar equality for B, namely,

$$B^k = B^{y_k}(AB^q)^{(s_k-1)} \quad \text{for every } q \geq 2 \text{ and } k \geq p+1, \tag{3.3}$$

where $y_k = k - (s_k - 1)(q - 1)$ and $s_k = \lceil (k-1)/(q-1) \rceil$.

The formulae (3.2) and (3.3) can be used to convert any word in A, B into a (p, q)-word in (A, B) (if it needs converting). So, since $\{A, B\}$ is irreducible, it follows that the (p, q)-words of (A, B) span $M_n(\mathbb{C})$ and the (p, q)-minimum spanning length of (A, B) exists for every $p, q \geq 2$. Let $2 \leq p \leq q \leq n - 1$ and suppose that $p \neq n - 1$. We show that the (p, q)-minimum spanning length of (A, B) is $n - 2 + 2R$, where $R = \lceil (n-2)/(p-1) \rceil$.

https://doi.org/10.1017/S0004972720001227 Published online by Cambridge University Press

In [6, Proposition 3], it is shown that the minimum spanning length of $\{A, B\}$ is at most n. This is done by exhibiting a basis for $M_n(\mathbb{C})$, consisting of words of length at most n. We will describe this basis here for the convenience of the reader. In its description, Δ_t, for $0 \le t \le n-1$, denotes the set of matrix positions $\{(i, j) : j - i = t\}$ and $\mathcal{D}_t$ denotes the subspace of $n \times n$ matrices with zero elements off the diagonal Δ_t.

$\mathcal{D}_0$: a basis for $\mathcal{D}_0$ is

$$\{AB, A^2B^2, \ldots, A^mB^m\} \cup \{BA, B^2A^2, \ldots, B^mA^m\} \qquad (n = 2m \text{ even}),$$
$$\{AB, A^2B^2, \ldots, A^mB^m\} \cup \{BA, B^2A^2, \ldots, B^mA^m\} \cup \{BA^2B\}\ (n = 2m+1 \text{ odd}).$$

$\mathcal{D}_{-t}$, $1 \le t \le n-2$: a basis for $\mathcal{D}_{-t}$ is

$$\{A^{t+1}B, A^{t+2}B^2, \ldots, A^{t+k}B^k\} \cup \{BA^{t+1}, B^2A^{t+2}, \ldots, B^kA^{t+k}\} \quad (n - t = 2k \text{ even}),$$
$$\{A^t\} \cup \{A^{t+1}B, A^{t+2}B^2, \ldots, A^{t+k}B^k\} \cup \{BA^{t+1}, B^2A^{t+2}, \ldots, B^kA^{t+k}\}\ (n - t = 2k+1 \text{ odd}).$$

$\mathcal{D}_{-(n-1)}$: a basis for $\mathcal{D}_{-(n-1)}$ is $\{A^{n-1}\}$.

$\mathcal{D}_t$, $1 \le t \le n-1$: a basis for $\mathcal{D}_t$ is $\{T^* : T$ a basis element of $\mathcal{D}_{-t}$ described above$\}$.

It is not too difficult to show that if each of these basis elements is converted into a (p, q)-word in (A, B) using (3.2) and (3.3) when necessary, the conversions all have length at most $n - 2 + 2R$ (noting that (a) $x \le y$ implies that $\lceil x\rceil \le \lceil y\rceil$ and (b) $\lceil x\rceil + \lceil y\rceil \le \lceil x + y\rceil + 1$ for real numbers x and y). Indeed, when A^x ($x \ge p + 1$) is converted into a (p, q)-word using (3.2), the length of the conversion is $x + 2(R_x - 1)$, where $R_x = \lceil (x-1)/(p-1)\rceil$ and, when B^y ($y \ge q + 1$) is converted using (3.3), the length of the conversion is $y + 2(S_y - 1)$, where $S_y = \lceil (y-1)/(q-1)\rceil$. So, when A^xB^y or B^yA^x is converted into a (p, q)-word, the length of conversion is L, where

$$\begin{aligned} L &= x + y + 2(R_x + S_y - 2) = x + y + 2\Big(\Big\lceil \frac{x-1}{p-1}\Big\rceil + \Big\lceil \frac{y-1}{q-1}\Big\rceil - 2\Big) \\ &\le x + y + 2\Big(\Big\lceil \frac{x-1}{p-1}\Big\rceil + \Big\lceil \frac{y-1}{p-1}\Big\rceil - 2\Big) \le x + y + 2\Big(\Big\lceil \frac{x-1}{p-1} + \frac{y-1}{p-1}\Big\rceil + 1 - 2\Big) \\ &\le x + y + 2\Big(\Big\lceil \frac{x+y-2}{p-1}\Big\rceil - 1\Big) \le n - 2 + 2R \quad \text{if } x + y \le n. \end{aligned}$$

It follows that the (p, q)-words (of (A, B)) of length at most $n - 2 + 2R$ span $M_n(\mathbb{C})$, so the (p, q)-msl is at most $n - 2 + 2R$.

There is a (p, q)-word mapping e_1 to e_n. Indeed, since $A^{n-1}e_1 = e_n$, the conversion of A^{n-1} to a (p, q)-word, namely, $A^z(BA^p)^{R-1}$, where $z = (n-1) - (R-1)(p-1)$, has this property (A^{n-1} needs to be converted since $n - 1 \ge p + 1$). Let X be a (p, q)-word of minimum length satisfying $Xe_1 = e_n$. We show that $X = BA^z(BA^p)^{R-1}$ (with z as above). Clearly such an X begins and ends in A, so that $X = A^{u_{t+1}}B^{v_t} \ldots A^{u_2}B^{v_1}A^{u_1}$, where $1 \le u_i \le p$ for $1 \le i \le t + 1$ and where $1 \le v_j \le q$ for $1 \le j \le t$. Define f_k for $1 \le k \le t + 1$ by $f_k = A^{u_k}B^{v_{k-1}} \ldots B^{v_1}A^{u_1}e_1$. Then $f_k = e_{\alpha_k}$, where $\alpha_k = 1 + \sum_{s=1}^{k} u_s - \sum_{s=1}^{k-1} v_s$, $1 \le k \le t$. For $1 \le i < j \le t + 1$, let $X_{i,j}$ be the segment of X defined by $X_{i,j} = A^{u_j}B^{v_{j-1}} \ldots A^{u_{i+1}}B^{v_i}$.

Then $X_{i,j} f_i = f_j$ and, by the minimality of length of X, $X_{i,j}$ is a (p, q)-word of minimum length, beginning with A and ending with B, mapping f_i to f_j. (If Y is a (p, q)-word, beginning with A and ending with B, mapping f_i to f_j with shorter length than $X_{i,j}$, then $X' = X_{j,t+1} Y X_{1,i} A^{u_1}$ is a (p, q)-word satisfying $X' e_1 = e_n$ of shorter length that X.)

Next we show that $v_i = 1$ for $1 \le i \le t$. Suppose that $v_i \ge 2$, where $1 \le i \le t$. Since $f_i \ne e_n = f_{t+1}$, we have $\alpha_i < n = \alpha_{t+1}$, so there exists an integer $j > i$ such that $\alpha_j > \alpha_i$. Let j be the smallest such integer. Then $\alpha_{j-1} - \alpha_i \le 0$ and $1 \le \alpha_j - \alpha_i = (\alpha_j - \alpha_{j-1}) + (\alpha_{j-1} - \alpha_i) = u_j - v_{j-1} + (\alpha_{j-1} - \alpha_i) \le u_j - v_{j-1} \le p - 1$, so $2 \le \alpha_j - \alpha_i + 1 \le p$. Thus, $Y = A^{\alpha_j - \alpha_i + 1} B$ is a (p, q)-word, beginning with A and ending with B, satisfying $Y f_i = f_j$. The length of Y is $\alpha_j - \alpha_i + 2 = \sum_{s=i+1}^{j} u_s - \sum_{s=i}^{j-1} v_s + 2$ and this is strictly less than the length of the segment $X_{i,j}$, which is $\sum_{s=i+1}^{j} u_s + \sum_{s=i}^{j-1} v_s$, since $v_i \ge 2$. This contradicts the minimality of the length of the segment $X_{i-1,j}$.

Next we show that $u_i = p$ for all $1 \le i \le t$. Suppose that $u_i \le p - 1$ with $1 \le i \le t$. Again, since $f_i \ne e_n = f_{t+1}$, $\alpha_i < n = \alpha_{t+1}$, there exists an integer $j > i$ such that $\alpha_j > \alpha_i$. Let j be the smallest such integer. Then, since $\alpha_{j-1} - \alpha_i \le 0$, once again $1 \le \alpha_j - \alpha_i \le u_j - v_{j-1} \le p - 1$. Thus, $Z = A^{\alpha_j - \alpha_i} B A^{u_i + 1} B$ is a (p, q)-word, beginning with A and ending with B, satisfying $Z f_{i-1} = f_j$. The length of Z is $\alpha_j - \alpha_i + u_i + 3 = \sum_{s=i}^{j} u_s - \sum_{s=i}^{j-1} v_s + 3$ and this is strictly less than the length of the segment $X_{i-1,j}$, which is $\sum_{s=i}^{j} u_s + \sum_{s=i-1}^{j-1} v_s$. (All the v_s equal 1.) This contradicts the minimality of the length of the segment $X_{i,j}$.

We have shown that $X = A^x (BA^p)^t$, where $2 \le x \le p$. (Obviously $x = u_{t+1} > 1$.) Since $Xe_1 = e_n$, it follows that $x + t(p - 1) = n - 1$, so $(x - 1) + t(p - 1) = n - 2$ and $(x - 1)/(p - 1) + t = (n - 2)/(p - 1)$, where $0 < (x - 1)/(p - 1) \le 1$. If $(n - 2)/(p - 1) \in \mathbb{Z}$, then $R = \lceil (n - 2)/(p - 1) \rceil = (n - 2)/(p - 1)$ and $(x - 1)/(p - 1) \in \mathbb{Z}$, so $x = p$. Then $t = R - 1$ and $x = (n - 1) - (R - 1)(p - 1)$. On the other hand, if $(n - 2)/(p - 1) \notin \mathbb{Z}$, then $0 < (x - 1)/(p - 1) < 1$, so $t - 1 < (n - 2)/(p - 1) < t$, so once again $t = R$ and $x = (n - 1) - (R - 1)(p - 1)$. Thus, X is precisely the conversion of A^{n-1} into a (p, q)-word.

Finally, we show that the (p, q)-msl is equal to $n - 2 + 2R$ by showing that every (p, q)-word V of length at most $n - 3 + 2R$ satisfies $(Ve_1 | e_{n-1}) = (Ve_2 | e_n)$. Such words cannot span $M_n(\mathbb{C})$. We do this by showing that every (p, q)-word W satisfying $(We_1 | e_{n-1}) \ne (We_2 | e_n)$ has length at least $n - 2 + 2R$. Now $We_j \in \{0, e_1, e_2, \ldots, e_n\}$ for every j, so such a word W satisfies either $\{(We_1 | e_{n-1}) = 1,\ (We_2 | e_n) = 0\}$ or $\{(We_1 | e_{n-1}) = 0,\ (We_2 | e_n) = 1\}$. Let W_1 be a (p, q)-word satisfying $W_1 e_1 = e_{n-1}$ and $W_1 e_2 \ne e_n$ and suppose that W_1 has minimum length. Such a word exists. Indeed, $Xe_1 = e_n$, so $Xe_2 = 0$ and $BXe_1 = e_{n-1}$, $BXe_2 = 0$. We will show that $W_1 = BX$.

Let $W_1 = S_k S_{k-1} \ldots S_2 S_1$, where each S_i is of one of the forms A^x or B^y for some x, y and where adjacent S_i have different forms ($S_i S_{i+1} = A^x B^y$ or $B^y A^x$). Note that $k \ge 3$ because $p \le n - 2$. There must exist j, $1 \le j \le k$, such that $S_j S_{j-1} \ldots S_2 S_1 e_1 = e_n$, otherwise $W_1 e_2 = e_n$. Let j be the smallest such positive integer. Then, since j is minimal, S_j must be A^x for some x. Then $S_j S_{j-1} \ldots S_1 e_2 = 0$ and so $BS_j S_{j-1} \ldots S_1 e_1 = e_{n-1}$ and

$BS_jS_{j-1}\dots S_1e_2 = 0$. By the minimality of the length of W_1, $W_1 = BS_jS_{j-1}\dots S_2S_1$. Thus, $W_1 = BX_1$, where $X_1 = S_jS_{j-1}\dots S_1$ is a (p,q)-word mapping e_1 to e_n. By the minimality of the length of W_1, the length of X_1 is minimal, so $X_1 = X$ and $W_1 = BX$.

Let W_2 be a (p,q)-word satisfying $W_2e_1 \neq e_{n-1}$ and $W_2e_2 = e_n$ and suppose that W_2 has minimum length. Such a word exists. Indeed, XB satisfies these conditions. We will show that $W_2 = XB$. Let $W_2 = T_mT_{m-1}\dots T_2T_1$, where each T_j is of one of the forms A^x or B^y for some x, y and where adjacent T_j have different forms ($T_jT_{j+1} = A^xB^y$ or B^yA^x). There must exist l, $1 \leq l \leq m$, such that $T_lT_{l-1}\dots T_2T_1e_2 = e_1$, otherwise $W_2e_1 = e_{n-1}$. Let l be the smallest such positive integer. Then T_l must be B^y for some y, so T_{l+1} is A^x for some x. Then $T_mT_{m-1}\dots T_{l+1}B$ is a (p,q)-word satisfying $T_mT_{m-1}\dots T_{l+1}Be_2 = e_n$ and $T_mT_{m-1}\dots T_{l+1}Be_1 = 0$. By the minimality of the length of W_2, $W_2 = T_mT_{m-1}\dots T_{l+1}B$. So, $W_2 = X_2B$, where $X_2 = T_mT_{m-1}\dots T_{l+1}$ is a (p,q)-word mapping e_1 to e_n. By the minimality of the length of W_2, the length of X_2 is also minimal, so $X_2 = X$ and $W_2 = XB$.

Finally, if $2 \leq q \leq p \leq n-1$ and $q \neq n-1$, the (p,q)-msl of (A,B) is equal to the (q,p)-msl, which is $n - 2 + 2\lceil (n-2)/(q-1)\rceil$, by what has been proven above. This completes the proof. □

REMARK 3.6. With A and B as in the preceding theorem, the theorem and symmetry ((p,q)-msl$(A,B) = (q,p)$-msl(A,B)) show that $n \leq (p,q)$-msl$(A,B) \leq 3n-6$. Similarly, the set of values that the (p,q)-msl can take is $\{(p, n-1)\text{-msl}: 2 \leq p \leq n-1\}$. For example, take $n = 11$.

value of p :	2	3	4	5	6	7	8	9	10
value of $(n-2)/(p-1)$:	9	9/2	3	9/4	9/5	9/6	9/7	9/8	1
value of $\lceil (n-2)/(p-1)\rceil$:	9	5	3	3	2	2	2	2	1
value of $(p, 10)$-msl :	27	19	15	15	13	13	13	13	11

EXAMPLE 3.7. Let $n \geq 2$ and consider the ordered pair (C, D), where $C = e_1 \otimes e_n$ and $D = J$. It is shown in [6] that the minimum spanning length of $\{C, D\}$ is $2n-2$. Thus, $(n-1, n-1)$-msl$(C,D) = 2n-2$. We show that the (p,q)-msl exists if and only if $q = n-1$ and then it equals $2n-2$, whatever the value of p. Let $1 \leq p \leq n-1$. Since $\{D^xCD^y : 1 \leq x, y \leq n-1, x + y \neq 2n-2\} \cup \{D^{n-1}\}$ is a basis for $M_n(\mathbb{C})$, the $(1, n-1)$-words span $M_n(\mathbb{C})$ and the $(1, n-1)$-msl exists and is at most $2n-2$. Since $(1, n-1)$-msl $\geq (n-1, n-1)$-msl $= 2n-2$, we have $(1, n-1)$-msl $= 2n-2$. Now it follows that $(p, n-1)$-msl $= 2n-2$ for every p with $2 \leq p \leq n-1$.

Next suppose that $q \neq n-1$ and $1 \leq p \leq n-1$. We show that the (p,q)-words do not span $M_n(\mathbb{C})$ by showing that every such word W satisfies $(We_n|e_1) = 0$. We may suppose that $W \neq 0$. Since $C^2 = 0$, W must be a $(1,q)$-word. Moreover, since for any matrix X, $CXC = \lambda C$, where $\lambda = (Xe_n|e_1)$, W must be a scalar multiple of a word of one of the following forms: $C, D^r, CD^r, D^rC, D^rCD^s$, with $1 \leq r, s \leq n-2$. It is easily checked that a matrix Z, of any of the latter forms, satisfies $(Ze_n|e_1) = 0$.

https://doi.org/10.1017/S0004972720001227 Published online by Cambridge University Press

Acknowledgement

The author thanks the referee for helpful comments.

References

[1] A. E. Guterman, T. Laffey, O. V. Markova and H. Šmigoc, 'A resolution of Paz's conjecture in the presence of a nonderogatory matrix', *Linear Algebra Appl.* **543** (2018), 234–250.
[2] P. R. Halmos, *Finite-Dimensional Vector Spaces*, Undergraduate Texts in Mathematics, 2nd edn, reprint (Springer, New York, 1974).
[3] T. Laffey, O. V. Markova and H. Šmigoc, 'The effect of assuming the identity as a generator on the length of the matrix algebra', *Linear Algebra Appl.* **498** (2016), 378–393.
[4] M. S. Lambrou and W. E. Longstaff, 'On the lengths of pairs of complex matrices of size six', *Bull. Aust. Math. Soc.* **80** (2009), 177–201.
[5] V. Lomonosov and P. Rosenthal, 'The simplest proof of Burnside's theorem on matrix algebras', *Linear Algebra Appl.* **383** (2004), 45–47.
[6] W. E. Longstaff, 'Burnside's theorem: irreducible pairs of transformations', *Linear Algebra Appl.* **382** (2004), 247–269.
[7] W. E. Longstaff, 'Irreducible families of complex matrices containing a rank one matrix', *Bull. Aust. Math. Soc.* **102**(2) (2020), 226–236.
[8] W. E. Longstaff, 'The slot length of a family of matrices', *Bull. Aust. Math. Soc.*, to appear.
[9] W. E. Longstaff, A. C. Niemeyer and O. Panaia, 'On the length of pairs of complex matrices of size at most five', *Bull. Aust. Math. Soc.* **73** (2006), 461–472.
[10] W. E. Longstaff and P. Rosenthal, 'On the lengths of irreducible pairs of complex matrices', *Proc. Amer. Math. Soc.* **139**(11) (2011), 3769–3777.
[11] A. Paz, 'An application of the Cayley–Hamilton theorem to matrix polynomials in several variables', *Linear Multilinear Algebra A* **15** (1984), 161–170.
[12] Y. Shitov, 'An improved bound for the lengths of matrix algebras' (English summary), *Algebra Number Theory* **13**(6) (2019), 1501–1507.

W. E. LONGSTAFF,
11 Tussock Crescent, Elanora, Queensland 4221, Australia
e-mail: bill.longstaff@alumni.utoronto.ca

https://doi.org/10.1017/S0004972720001227 Published online by Cambridge University Press