104

Extensions of Jordan Bases for Invariant Subspaces of a Matrix

Rafael Bru*
Departamento Matematica Aplicada
Univ. Politecnica de Valencia
46071 Valencia, Spain

Leiba Rodman†
Department of Mathematics
College of William and Mary
Williamsburg, Virginia 23187-8795

and

Hans Schneider‡
Department of Mathematics
University of Wisconsin
Madison, Wisconsin 53706

Submitted by Daniel Hershkowitz

ABSTRACT

A characterization is obtained for the matrices A with the property that every (some) Jordan basis of every A-invariant subspace can be extended to a Jordan basis of A. These results are based on a criterion for a Jordan basis of an invariant subspace to be extendable to a Jordan basis of the whole space. The criterion involves two concepts: the constancy property and the depth property.

*Supported by Conselleria de Cultura Educacion y Ciencia de la Generalitat Valenciana, and by D.G.I.C.Y.T. of Ministerio de Educacion y Ciencia, Spain.

†Partially supported by NSF grant DMS 88-02836.

‡Partially supported by NSF grants DMS 85-21521, DMS 89-01445, ECS 87-18971.

655 Avenue of the Americas, New York, NY 10010 0024-3795/91/$3.50

1. INTRODUCTION

Let A be an $n \times n$ complex matrix considered as a linear transformation $\mathbb{C}^n \to \mathbb{C}^n$. A chain (for A) is a set of nonzero vectors

$$\{u, (A-\lambda I)u, \dots, (A-\lambda I)^{k-1}u\} \tag{1.1}$$

such that $(A-\lambda I)^k u = 0$. The complex number λ is necessarily an eigenvalue of A and $(A-\lambda I)^{k-1}u$ is an eigenvector. A Jordan basis for an invariant subspace W is a basis for W which is the union of chains. A Jordan basis of $\mathbb{C}^n$ will be called a Jordan basis for A. That is, a Jordan basis for A is a basis of the form

$$\{u_i, (A-\lambda_i I)u_i, \dots, (A-\lambda_i I)^{k_i-1}u_i;\ i=1,\dots,t\} \tag{1.2}$$

where $u_i \in \mathbb{C}^n$ and $(A-\lambda I)^{k_i}u_i = 0$. The existence of a Jordan basis for any $n \times n$ matrix A is well known and follows from the existence of the Jordan normal form of A.

Given a Jordan basis (1.2), certain A-invariant subspaces are seen immediately. Namely, for any choice of integers m_i $(i=1,\dots,t)$ such that $0 \leqslant m_i \leqslant k_i$, the subspace

$$M = \operatorname{span}\{(A-\lambda_i I)^{m_i}u_i, (A-\lambda_i I)^{m_i+1}u_i, \dots, (A-\lambda_i I)^{k_i-1}u_i;\ i=1,\dots,t\} \tag{1.3}$$

is A-invariant, i.e., $Ax \in M$ for every $x \in M$ [the equality $m_i = k_i$ for some i is interpreted as the indication that i is missing in the formula (1.3)]. The A-invariant subspaces that arise in this way, starting with any Jordan basis, are called marked in [2]. Equivalently, an A-invariant subspace M is called *marked* if there is a Jordan basis for the restriction $A_{|M}: M \to M$ which can be extended (by adjoining to it new vectors) to a Jordan basis for A in $\mathbb{C}^n$.

Generally, not every A-invariant subspace is marked (an example is given in [2]). The existence of nonmarked invariant subspaces is sometimes overlooked in linear algebra texts. In this paper we characterize those matrices A for which every invariant subspace is marked. We also characterize the matrices A with a stronger property, namely, that every A-invariant subspace is strongly marked. Let us define this notion: an A-invariant subspace M is *strongly marked* if every Jordan basis of M can be extended (by adjoining new vectors) to a Jordan basis for A in $\mathbb{C}^n$. These notions call our attention to a more general question: when can a given Jordan basis for an A-invariant subspace be extended to a Jordan basis for the whole space $\mathbb{C}^n$? We solve this problem in Section 2 in terms of the height and depth of vectors and related properties. Another characterization (in different terms) of this extendability property is given in [1].

These results are used in subsequent sections to characterize marked and strongly marked subspaces. This characterization goes as follows. (The multiplicities of a matrix A corresponding to its eigenvalue λ_0 are simply the sizes of the Jordan blocks with the eigenvalue λ_0 in the Jordan normal form of A.)

THEOREM 1.1. *Let A be an $n \times n$ matrix. Then every A-invariant subspace is marked if and only if for every eigenvalue λ_0 of A the difference between the biggest and the smallest multiplicity of A corresponding to λ_0 does not exceed* 1.

THEOREM 1.2. *Let A be an $n \times n$ matrix. Then every A-invariant subspace is strongly marked if and only if for every eigenvalue λ_0 of A all multiplicities of A corresponding to λ_0 are equal.*

To illustrate these results consider the following example. Let

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

According to Theorems 1.1 and 1.2, every A-invariant subspace is marked, but there are A-invariant subspaces which are not strongly marked. For example, $K(A)$ is not strongly marked. Here and elsewhere in this paper $K(A)$ stands for the kernel (null space) of the matrix A. Indeed, a Jordan basis for $A_{|K(A)}$ given by $(\alpha_1, 0, \beta_2, 0, \gamma_1)^T, (\alpha_2, 0, \beta_2, 0, \gamma_2)^T, (\alpha_3, 0, \beta_3, 0, \gamma_3)^T$ (here $\alpha_j, \beta_j, \gamma_j, \in \mathbb{C}$) can be extended to a Jordan basis for A if and only if there are two zeros among the numbers $\gamma_1, \gamma_2, \gamma_3$. An easy (but somewhat tedious) analysis shows that the following is a complete list of all A-invariant subspaces which are not strongly marked: $K(A)$; all 2-dimensional A-invariant subspaces spanned by eigenvectors, with the exception of $\mathrm{Span}\{(1,0,0,0,0)^T, (0,0,1,0,0)^T\}$; all 4-dimensional A-invariant subspaces containing $K(A)$.

As a corollary we recover the following result from [2] (Theorem 2.9.2). In fact the conclusion of Theorem 2.9.2 of [2] is weaker in the sense that only the marked property of every A-invariant subspace is asserted there.

COROLLARY 1.2. *Let A be an $n \times n$ matrix such that for every eigenvalue λ of A at least one of the following holds:*

(a) *the geometric multiplicity (i.e., the dimension of $K(A-\lambda I)$) is equal to the algebraic multiplicity;*
(b) $\dim K(A-\lambda I)=1$.

Then every A-invariant subspace is strongly marked.

The proofs of Theorems 1.1 and 1.2 will be given in Sections 3 and 4, respectively.

We conclude the introduction by remarking that it is sufficient to prove Theorems 1.1 and 1.2 (and Theorem 2.1 stated below) for the case when A has a single eigenvalue λ_0 (without loss of generality it can be assumed that $\lambda_0 = 0$). This follows readily from the well-known fact that every A-invariant subspace M can be written as

$$M = \left[M \cap R_{\lambda_1}(A)\right] + \cdots + \left[M \cap R_{\lambda_r}(A)\right],$$

where $\lambda_1, \dots, \lambda_r$ are all the distinct eigenvalues of A and

$$R_{\lambda_j}(A) = K(A - \lambda_j I)^n$$

is the root subspace of A corresponding to λ_j. Thus, it will be assumed in Sections 2, 3, and 4 that A is nilpotent: $A^n = 0$.

2. HEIGHT AND DEPTH

Let A be an $n \times n$ nilpotent complex matrix.

For a given $x \in \mathbb{C}^n$ let the *height* of x [notation: $\mathrm{ht}(x)$] be the minimal nonnegative integer k such that $A^k x = 0$ (as usual, we assume $A^0 = I$; thus zero is the only vector of height zero). For $x \neq 0$, the *depth* of x [notation: $\mathrm{dpth}(x)$] is by definition the maximal nonnegative integer k such that

$x = A^k y$ for some y. Note the following easily verified properties: For complex numbers $\alpha_1, \dots, \alpha_s$, and vectors $x_1, \dots, x_s$ we have

$$\mathrm{ht}\left(\sum_{i=1}^{s} \alpha_i x_i\right) \leqslant \max\{\mathrm{ht}(x_i) : i = 1, \dots, s\}, \tag{i}$$

$$\mathrm{dpth}\left(\sum_{i=1}^{s} \alpha_i x_i\right) \geqslant \min\{\mathrm{dpth}(x_i) : i = 1, \dots, s\}, \tag{ii}$$

and the strict inequality

$$\mathrm{dpth}(x) \neq \mathrm{dpth}(y) \quad \Rightarrow \quad \mathrm{dpth}(x+y) = \min\{\mathrm{dpth}(x), \mathrm{dpth}(y)\}, \tag{iii}$$

provided all vectors in (ii) and (iii) are nonzero. Also, for $0 \neq u \in \mathbb{C}^n$, we have

$$\mathrm{ht}(Au) = \mathrm{ht}(u) - 1, \tag{iv}$$

$$\mathrm{dpth}(Au) > \mathrm{dpth}(u), \tag{v}$$

provided that $Au \neq 0$.

We address the question when a given Jordan basis B (1.2) for W can be extended to a Jordan basis for the whole space $\mathbb{C}^n$, i.e., when there is a Jordan basis T in $\mathbb{C}^n$ such that $B \subseteq T$ (as sets of vectors). The answer is based on two notions that we call the constancy property and the depth property.

We say that a nonzero vector x has the *constancy property* (CP) if either $Ax = 0$ or $Ax \neq 0$ and

$$\mathrm{dpth}(Ax) = \mathrm{dpth}(x) + 1.$$

A set S of nonzero vectors is said to have the CP if every vector in S has the CP. In particular, the notion of the constancy property can be applied to a chain $S = \{x, Ax, \dots, A^{k-1}x\}$; thus, this chain has CP if and only if

$$\mathrm{dpth}(A^{i-1}x) = \mathrm{dpth}(x) + i - 1, \qquad i = 1, \dots, k. \tag{2.1}$$

As by (iv) $\mathrm{ht}(A^i x) = k - i$ $(0 \leqslant i \leqslant k-1)$, these equalities can be rewritten in the form

$$\mathrm{dpth}(x) + \mathrm{ht}(x) = \mathrm{dpth}(A^i x) + ht(A^i x), \qquad i = 0, \dots, k-1. \quad (2.2)$$

Also, if $\mathrm{dpth}(A^{k-1}x) = k-1$, then necessarily $\mathrm{dpth}(x) = 0$ and (2.2) holds, and thus the chain $\{x, Ax, \dots, A^{k-1}x\}$ has the CP.

In what follows we use the notation $\langle q \rangle$ for the set $\{1, \dots, q\}$.

We say that a linearly independent set of vectors $\{x^i : i \in \langle q \rangle\}$ has the DP (the *depth property*) if $w = \sum_{i \in \langle q \rangle} \alpha_i x^i$, $w \neq 0$, implies that

$$\mathrm{dpth}(w) = \min\{\mathrm{dpth}(x^i) : i \in \langle q \rangle \text{ and } \alpha_i \neq 0\}. \quad (2.3)$$

The two properties CP and DP do not imply each other, as examples will presently show. First, note that every chain $\{x, Ax, \dots, A^{k-1}x\}$ is linearly independent, by a standard argument. It follows from (iii) and (v) that every chain has the DP. An example of a chain without the CP (but with DP) is furnished by $\{u, Au\}$ where

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and $u = (0,1,0,1)^T$. The following example shows a linearly independent set of chains without the DP. Let

$$B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$u = (1,0,1)^T$, and $v = (0,0,1)^T$. Then each of the (singleton) chains $\{u\}$ and $\{v\}$ has the CP, the set $\{u, v\}$ is linearly independent, but $\{u, v\}$ does not have the DP. Indeed, $\mathrm{dpth}(u) = \mathrm{dpth}(v) = 0$, but for a nonzero vector $w = \alpha u + \beta v$ we have $\mathrm{dpth}(w) = 0$ if $\alpha + \beta \neq 0$ and $\mathrm{dpth}(w) = 1$ if $\alpha + \beta = 0$.

The main result of this paper is the following (which holds without the assumption that A is nilpotent, though the proof is given only for nilpotent A; see the end of Section 1).

THEOREM 2.1. *Let A be a complex $n \times n$ matrix. Let B be a Jordan basis for an A-invariant subspace W. Then B can be extended to a Jordan basis for*

A in $\mathbb{C}^n$ if and only if B has the CP and the DP.

Proof. "If": Let

$$C = \{u_i, Au_i, \dots, A^{k_i - 1}u_i : i = 1, \dots, t\}$$

be a Jordan basis for $\mathbb{C}^n$. It is enough to prove that C has the CP and the DP, for then any subset of C has the CP and the DP.

Let w be a nonzero vector in $\mathbb{C}^n$. Then w can be written uniquely as

$$w = \sum_{i=1}^{t} \sum_{j=0}^{k_i - 1} \alpha_{ij} A^j u_i, \qquad \alpha_{ij} \in \mathbb{C}, \quad j = 0, \dots, k_i - 1, \ i = 1, \dots, t.$$

Then, for $p \geqslant 0$, $A^p w$ has the unique representation

$$A^p w = \sum_{i=1}^{t} \sum_{j=0}^{k_i - 1} \alpha_{i,j-p} A^j u_i, \qquad j = 0, \dots, k_i - 1, \quad i = 1, \dots, t,$$

where $a_{ij} = 0$ whenever $j < 0$, $i = 1, \dots, t$. If $Aw \neq 0$, then it follows easily that

$$\operatorname{dpth}(w) = \min\{ j : \text{at least one of } \alpha_{ij}, \ i = 1, \dots, t, \text{ is nonzero}\}. \quad (2.4)$$

In particular,

$$\operatorname{dpth}(A^j u_i) = j, \qquad j = 0, \dots, k_i - 1, \quad i = 1, \dots, t. \qquad (2.5)$$

Thus, by (2.1), C has the CP, and, by applying (2.5) to (2.4) we see that C has the DP.

Hence if B is a subset of C, then B has the CP and the DP.

"Only if": We suppose that $W \neq \mathbb{C}^n$, for otherwise of course there is nothing to prove. We consider two cases. In each case we construct a subspace W' which properly contains W and a Jordan basis $B' \supseteq B$ for W' such that B' has the CP and the DP.

We say that a chain $\{u, Au, \dots, A^{k-1}u\}$ is maximal if it is not contained (set theoretically) in a larger chain; in other words, a chain $\{u, Au, \dots, A^{k-1}u\}$ is maximal if $\operatorname{dpth}(u) = 0$.

Case I: Some chain of B is not maximal. Suppose that $S = \{u, \dots, A^{h-1}u\}$ is a chain of B, and that $\operatorname{dpth}(u) = d > 0$. Let $y \in \mathbb{C}^n$ satisfy $A^d y = u$.

Then $\mathrm{dpth}(y)=0$. Let $S'=\{y,\dots,A^{d+h-1}y\}$. We let B' consist of the chains of B with S replaced by S', and we let $W'=\mathrm{span}(B')$.

Claim 1.1. *B' is linearly independent.* Otherwise, there exists a nontrivial linear relation on B', and since this cannot be a nontrivial linear relation on B, it must involve an element of form $A^r y$, where $r<d$. We choose the minimal such r. Multiplying this linear relation by A^{d-r}, we obtain a linear relation on the elements of B, which is nontrivial, since it involves $A^d y = u$. But this is impossible, since B is linearly independent.

Claim 1.1. *The chain $S'=\{y,\dots,A^{d+h-1}y\}$ has the CP.* Otherwise, by (2.1), $\mathrm{dpth}(A^{d+h-1)}y) > d+h-1$, and there is a $y'\in\mathbb{C}^n$ such that

$$A^{d+h}y' = A^{d+h-1}y = A^{h-1}u.$$

But then

$$\mathrm{dpth}(A^{h-1}u) - \mathrm{dpth}(u) \geqslant d+h-d = h,$$

which is impossible by (2.1), since S has the CP. Hence $\mathrm{dpth}(A^{d+h-1}y) = d+h-1$, and hence S' has the CP.

Claim 1.3. *B' has the DP.* Recall that every chain has the DP. Suppose B' does not. Since B and S' have the DP, it is easily shown using (iii) that there exists a $w\in W$ and an

$$x = \sum_{s\in\{r,\dots,d-1\}} \gamma_s A^s y, \qquad \gamma_s\in\mathbb{C}, \quad \gamma_r\neq 0, \tag{2.6}$$

where $0\leqslant r<d$, such that $w\neq 0$, $x\neq 0$,

$$\mathrm{dpth}(w) = \mathrm{dpth}(x) = r, \tag{2.7}$$

and, for $v=w+x$,

$$\mathrm{dpth}(v) > r. \tag{2.8}$$

We then obtain

$$\mathrm{dpth}(A^{d-r}v) \geqslant \mathrm{dpth}(v) + d - r > d. \tag{2.9}$$

But this is impossible, for $A^{d-r}v$ is a linear combination of nonzero elements of B one of which is $A^d y = u$, and $\operatorname{dpth}(u) = d$. This proves the claim, and completes the proof of case I.

Case II: Every chain of B is maximal.

Claim II.1. There exists $v \in \mathbb{C}^n$, $v \notin W$, with $\operatorname{ht}(v) = 1$ *(i.e., v is an eigenvector of A).* Let $u \in \mathbb{C}^n$, $u \notin W$. Let $\operatorname{ht}(u) = h$. If $A^{h-1}u \notin W$, the claim is true. Otherwise there exists a least r, $0 < r < h-1$, such that $A^r u \in W$. Thus $A^r u$ is a linear combination of B, and, since B has the DP, it is a linear combination of elements of B whose depth is at least 1. Thus there exists a $w \in W$ such that $Aw = A^r u$. Let $x = w - A^{r-1}u$. Then $x \notin W$ and $Ax = 0$, which proves the claim.

We now choose a chain $S = \{u, \dots, A^{h-1}u\}$ of maximal length such that $A^{h-1}u = v$ is not in W. Let $B' = B \cup S$. Then it is easy to prove that $W \cap \operatorname{span}(S) = 0$, and it follows that B' is a basis for $W' = W \oplus \operatorname{span}(S)$.

Claim II.2. The chain S has the CP. Since S is a maximal chain beginning at u, clearly $\operatorname{dpth}(u) = 0$. Suppose S does not have the CP. Then, by (2.1), $\operatorname{dpth}(A^{h-1}u) > h-1$. Hence there is a $w \in \mathbb{C}^n$ such that $A^h w = A^{h-1}u$. Thus the chain $\{w, \dots, A^{h-1}u\}$ has greater length than S, contrary to the assumption that S is a maximal chain whose last element is not in W.

Claim II.3. B′ has the DP. Suppose B' does not have the DP. Since B and the chain S have the DP, there must exist

$$v = w + x, \qquad w \in W, \quad x \in \operatorname{span}(S),$$

such that

$$\operatorname{dpth}(v) > \min\{\operatorname{dpth}(w), \operatorname{dpth}(x)\}. \tag{2.10}$$

By (iii), we then have

$$\operatorname{dpth}(w) = \operatorname{dpth}(x) = d, \qquad \text{say}, \tag{2.11}$$

when $0 \leqslant d < h$. By (2.11), we have

$$x = \sum_{r \in \{d, \dots, h-1\}} \gamma_r A^r u, \qquad \gamma_r \in \mathbb{C}, \quad \gamma_d \neq 0.$$

By (2.10), there is a $z \in \mathbb{C}^n$ such that $A^{d+1}z = v$. Then

$$A^h z = A^{h-d-1}v = A^{h-d-1}w + A^{h-d-1}x = A^{h-d-1}w + \gamma_d A^{h-1}u.$$

Since $A^{h-d-1}w \in W$ and $\gamma_d A^{h-1}u \neq 0$, it follows that there is a chain of length $h+1$ which ends outside W, contrary to our assumption on S. This proves our claim, and completes the proof of case II.

Thus in either case, we have constructed an invariant subspace W' with $\dim(W') > \dim(W)$ and a Jordan basis $B' \supseteq B$ with the DP such that B' has the CP. By repeating this argument we obtain a Jordan basis for $\mathbb{C}^n$, which is an extension of B. ■

Another necessary condition for extendability of B to Jordan basis if $\mathbb{C}^n$ can be given in terms of multiplicities, as follows. We write the list of all multiplicities (including repetitions, if necessary) in a nonincreasing order: $\lambda_1 \geqslant \cdots \geqslant \lambda_q$. A sequence of positive integers $\beta_1, \dots, \beta_p$ will be called a sublist of multiplicities if $p \leqslant q$ and there is a one-to-one map $\zeta : \{1, \dots, p\} \to \{1, \dots, q\}$ such that $\beta_i = \lambda_{\zeta(i)}$ for $i = 1, \dots, p$. By the *index* of a chain $S = \{x, \dots, A^{k-1}x\}$, denoted $\operatorname{ind}(S)$, we mean $\operatorname{dpth}(A^{k-1}x) + 1$.

By (2.4), it is easy to see that if a Jordan basis B of W is extendable to a Jordan basis of $\mathbb{C}^n$, then the numbers $\operatorname{ind}(S_i)$, where $S_1, \dots, S_r$ are the chains in B, form a sublist of multiplicities. The following example shows that a Jordan basis B with the CP and for which $\operatorname{ind}(S_i)$, $i \in \langle r \rangle$, form a sublist of multiplicities need not be extendable to a Jordan basis of $\mathbb{C}^n$.

Example 2.1. Let

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Let $u = (1,0,1,0)^T$, $v = (0,0,1,0)^T$. Then $\{u, v\}$ forms a Jordan basis B of the subspace $W = \operatorname{span}\{(1,0,0,0)^T, (0,0,1,0)^T\}$. By Theorem 2.1 the basis B cannot be extended to a Jordan basis in $\mathbb{C}^4$, since $\operatorname{dpth}(u) = \operatorname{dpth}(v) = 0$, while $\operatorname{dpth}(u - v) = 1$. However, the basis B has the CP and $\{1, 1\}$ is a sublist of multiplicities.

3. PROOF OF THEOREM 1.1

Let A be an $n \times n$ nilpotent matrix. We start with the following:

Proposition 3.1. *Suppose that every A-invariant subspace is marked. Then the lengths of any two maximal chains (in a Jordan basis of A) differ by less than two.*

Proof. Arguing by contradiction, assume that

$$u, Au, \ldots, A^{r-3}u \tag{3.1}$$

is a maximal chain in a Jordan basis of A, and

$$v, Av, \ldots, A^{r-1}v$$

is a not necessarily maximal chain in the same Jordan basis of A. Put $z = u + Av$. Note that $A^{r-2}z = A^{r-1}v$. Further note that $\mathrm{dpth}(z) = 0$ [indeed, if $Ay = z$ for some y, then $u = A(y - v)$, which is a contradiction with the maximality in (3.1)]. We have $\mathrm{ht}(A^{r-2}z) = 1$, $\mathrm{dpth}(A^{r-2}z) \geqslant r-1$, $\mathrm{ht}(z) = r-1$, and $\mathrm{dpth}(z) = 0$; so the chain $z, Az, \ldots, A^{r-2}z$ does not have the CP and hence by Theorem 2.1 cannot be extended to a Jordan basis for $\mathbb{C}^n$. Observe that every Jordan basis for $\mathrm{span}\{z, \ldots, A^{r-2}z\}$ has the form

$$\{w, Aw, \ldots, A^{r-2}w\},$$

where

$$w = \sum_{j=0}^{r-2} \alpha_j A^j z, \qquad \alpha_j \in \mathbb{C}, \quad \alpha_0 \neq 0.$$

We see that $\mathrm{dpth}(z) = 0$ and $\mathrm{dpth}(A^{r-2}w) = \mathrm{dpth}(\alpha_0 A^{r-2}w) = \mathrm{dpth}(\alpha_0 A^{r-2}z) \geqslant r-1$. Thus, the chain $\{A^j w\}_{j=0}^{r-2}$ does not have the CP, and by Theorem 2.1 it cannot be extended to a Jordan basis for A. Therefore, the A-invariant subspace $\mathrm{span}\{z, \ldots, A^{r-2}z\}$ is not marked. ∎

PROPOSITION 3.2. *Let A be an $n \times n$ nilpotent matrix with sizes of all Jordan blocks equal to q or $q-1$. Then all chains have the CP.*

Proof. Let w be a vector in $\mathbb{C}^n$ such that $Aw \neq 0$. We shall show that

$$\mathrm{ht}(w) + \mathrm{dpth}(w) = \mathrm{ht}(Aw) + \mathrm{dpth}(Aw). \tag{3.2}$$

It follows from our assumptions that we may write

$$w = u + v,$$

where u and v are linear combinations of vectors in Jordan chains of lengths, respectively, q and $q-1$. Suppose that

$$\mathrm{ht}(u)=h, \qquad \mathrm{ht}(v)=k.$$

Then $0 \leqslant h \leqslant q$, $0 \leqslant k \leqslant q-1$. If $h=0$ (i.e. $u=0$) or $k=0$ (i.e. $v=0$), then we are basically in the situation [as far as (3.2) is concerned] when all the multiplicities of A are equal. But in this case (3.2) follows easily [see also the equivalence (1) $\Leftrightarrow$ (2) in Theorem 1.2′ of Section 4]. So suppose that $h, k \geqslant 1$. Note that we cannot have $h=k=1$, for then $Aw=0$, contrary to assumption. So either $h>1$ or $k>1$. It is easily checked that

$$\mathrm{ht}(w)=\max\{h,k\},$$

$$\mathrm{dpth}(w)=\min\{q-h,q-1-k\},$$

$$\mathrm{ht}(Aw)=\max\{h-1,k-1\},$$

$$\mathrm{dpth}(Aw)=\min\{q-h+1,q-k\}.$$

If $h>k$, it follows that

$$\mathrm{ht}(w)+\mathrm{dpth}(w)=\mathrm{ht}(Aw)=\mathrm{dpth}(Aw)=q,$$

while if $h \leqslant k$

$$\mathrm{ht}(w)+\mathrm{dpth}(w)=\mathrm{ht}(Aw)+\mathrm{dpth}(Aw)=q-1.$$

In either case, (3.2) holds and the proposition follows. ■

PROPOSITION 3.3. *Let A be an $n\times n$ nilpotent matrix with sizes of all Jordan blocks equal to q or $q-1$. Let M be an A-invariant subspace of $\mathbb{C}^n$, and let B be a Jordan basis for M with a maximal number of eigenvectors of depth $q-1$. Then B has the DP.*

Proof. Let B be the Jordan basis

$$\{g_i, Ag_i,\dots,A^{k_i-1}g_i,\ i=1,\dots,t\},$$

and let

$$k = \max\{k_i : i = 1,\dots,t\}.$$

Let B_h, $h = 1,\dots,k$, be the subset of B consisting of vectors of height h or less, viz.

$$B_h = \{A^{k_i - j}g_i : j \in \langle h\rangle,\, i \in \langle t\rangle\}.$$

We shall prove by induction that B_h, $h = 1,\dots,k$, has the DP.

We first consider B_1. In view of our assumptions on multiplicities, each vector in B_1 has depth $q-2$ or $q-1$ (since it is an eigenvector of A). Consider the linear combination of B_1:

$$0 \neq w = \sum_{i=1}^{p} \alpha_i A^{k_i - 1}g_i + \sum_{i=p+1}^{t} \alpha_i A^{k_i - 1}g_i, \tag{3.3}$$

where we may assume that

$$\mathrm{dpth}(A^{k_i - 1}g_i) = q-2, \qquad i = 1,\dots,p,$$

$$\mathrm{dpth}(A^{k_i - 1}g_i) = q-1, \qquad i = p+1,\dots,t.$$

Now suppose that B_1 does not have the DP. Then we may find a vector w of form (3.3) such that at least one of the coefficients α_i, $1 \leqslant i \leqslant p$, is nonzero and $\mathrm{dpth}(w) = q-1$. But then

$$v = \sum_{i=1}^{p} \alpha_i A^{k_i - 1}g_i \tag{3.4}$$

also satisfies $\mathrm{dpth}(v) = q-1$ by (iii). Let s, $1 \leqslant s \leqslant p$, be an index for which $\alpha_s \neq 0$ in (3.4) and such that k_s is minimal among k_i for which $\alpha_i \neq 0$ in (3.4). Suppose, without loss of generality, that $\alpha_i \neq 0$, $i = 1,\dots,s$ and $\alpha_i = 0$, $i = s+1,\dots,p$. Let

$$u = \sum_{i=1}^{s} \alpha_i A^{k_s - 1}g_i,$$

and in B replace the Jordan chain

$$\{A^j g_s : j = 0, \dots, k_s - 1\}$$

by the Jordan chain

$$\{A^j u : j = 0, \dots, k_s - 1\}.$$

The result is a Jordan basis for M for which the number of eigenvectors of depth $q-1$ is $t-p+1$. But, since B has $t-p$ such eigenvectors, this contradicts our assumption on B. Hence B_1 has the DP.

Now assume inductively that $1 < h \leqslant k$ and that B_{h-1} has the DP. To prove that B_h has the DP, we consider

$$0 \neq w = \sum_{i=1}^{t} \sum_{j=k_i'}^{k_i - 1} \alpha_{ij} A^j g_i, \tag{3.5}$$

where $k_i' = \max\{0, k_i - h\}$, $i = 1, \dots, t$. We must prove that

$$\operatorname{dpth}(w) = \min\{\operatorname{dpth}(A^j g_i) : \alpha_{ij} \neq 0,\ j = k_i', \dots, k_i - 1,\ i = 1, \dots, t\}. \tag{3.6}$$

If $\alpha_{ij} = 0$ whenever $j = k_i - h$, then w is a linear combination of elements of B_{h-1}, and (3.6) follows from our inductive assumption. So assume that $\alpha_{sj} \neq 0$ for some $j = k_s - h$ and $1 \leqslant s \leqslant t$. Note that it follows from our assumption on multiplicities that

$$\operatorname{dpth}(A^{k_i - 1} g_i) \geqslant q - 2, \qquad i = 1, \dots, s$$

(since the above vectors are eigenvectors), and

$$\operatorname{dpth}(A^{k_s - h} g_s) \leqslant q - 2,$$

since $h \geqslant 2$ (and hence this vector is not an eigenvector). Thus to prove (3.6) it is enough to prove

$$\operatorname{dpth}(w) = \min\{\operatorname{dpth}(A^j g_i) : \alpha_{ij} \neq 0,\ j = k_i', \dots, k_i - 2,\ i = 1, \dots, t\}. \tag{3.7}$$

To prove (3.7) we note that

$$Aw = \sum_{i=1}^{t} \sum_{j=k_i'}^{k_i-2} \alpha_{ij} A^{j+1} g_i$$

(and thus $Aw \neq 0$). Since $A^{j+1}g_i \in B_{h-1}$, $j = k_i', \dots, k_i - 2$, $i = 1, \dots, t$, our inductive assumption yields

$$\operatorname{dpth}(Aw) = \min\{\operatorname{dpth}(A^{j+1}g_i) : \alpha_{ij} \neq 0,\ j = k_i', \dots, k_i - 2,\ i = 1, \dots, t\}. \tag{3.8}$$

By Proposition 3.2, every chain has CP. Hence,

$$\operatorname{dpth}(Aw) = \operatorname{dpth}(w) + 1,$$

$$\operatorname{dpth}(A^{j+1}g_i) = \operatorname{dpth}(A^j g_i) + 1, \qquad j = k_i', \dots, k_i - 2, \quad i, \dots, t.$$

Hence (3.7) now follows from (3.8), and thus B_h has the DP. By induction, we obtain that B_k has the DP, and since $B_k = B$, the result follows. ■

Proof of Theorem 1.1. We may assume that A is nilpotent. If the difference between the biggest and the smallest multiplicity of A is at least 2, then by Proposition 3.1 not every A-invariant subspace is marked. Conversely, assume that the multiplicities of A are equal to q and $q-1$, for some $q \geqslant 2$. In view of Propositions 3.2 and 3.3, every A-invariant subspace M has a Jordan basis with the DP that also has the CP. The theorem now follows from Theorem 2.1. ■

We can augment Theorem 1.1 by the following statement.

THEOREM 3.4. *Assume A is a nilpotent. Then every A-invariant subspace is marked if and only if there is q such that the index of every vector in $\mathbb{C}^n$ is either q or $q-1$.*

Theorem 1.1 was contained in an unpublished manuscript by the authors dated July 1988. A related result (in the framework of solutions of Riccati equations) was obtained independently in [3].

4. PROOF OF THEOREM 1.2

We will actually prove a more informative result.

THEOREM 1.2′. *The following are equivalent for a nilpotent matrix* A:

(1) *All multiplicities are equal (to* q*).*
(2) *For all* $x \in \mathbb{C}^n \setminus \{0\}$, $\mathrm{ht}(x)+\mathrm{dpth}(x)=q$.
(3) *All invariant subspaces are strongly marked.*

Proof. (1) $\Rightarrow$ (2): Let C be a Jordan basis in $\mathbb{C}^n$. Then every element of C of height h has depth $q-h$, and (2) follows because C has the DP (see the proof of the "if" part of Theorem 2.1).

(2) $\Rightarrow$ (3): Clearly, (2) implies that every chain has the CP. Let W be an invariant subspace for A. Since all Jordan bases for W contain the same number of eigenvectors, and by (2), all eigenvectors have the same depth $q-1$, it follows that every Jordan basis for W satisfies the hypotheses of Proposition 3.3. Hence every Jordan basis for W has the DP. We now obtain (3) by Theorem 2.1.

(3) $\Rightarrow$ (1): Suppose (1) is false, and x and y generate Jordan chains of lengths q and r respectively, where $r<q$. Let $u=A^{r-1}y-A^{q-1}x$ and $v=A^{q-1}x$. Then $\mathrm{dpth}(u)=\mathrm{dpth}(v)=q-1$, but $\mathrm{dpth}(u+v)=r-1$. Hence the Jordan basis $\{u,v\}$ for the invariant subspace span $\{u,v\}$ does not have the DP. By Theorem 2.1, span$\{u,v\}$ is not strongly marked. ■

We now give a characterization of condition (1) in Theorem 1.2′ in terms of the Weyr characteristic. Recall that the Weyr characteristic of a matrix X corresponding to eigenvalue λ is the vector $(w_1, w_2, \dots, w_d)$, where

$$w_j = \dim K(X-\lambda I)^j - \dim K(X-\lambda I)^{j-1}, \qquad j=1,2,\dots,d,$$

and d is the largest multiplicity of X corresponding to λ.

PROPOSITION 4.2. *The following statements are equivalent for a nilpotent matrix* A *(we denote by* d *the largest multiplicity of* A*):*

(i) $K(A^{d-1}) \subset R(A)$;
(ii) *the Weyr characteristic of* A *is* $(w_1,\dots,w_d)$, *where* $w_1=w_2=\dots=w_d$;
(iii) *all multiplicities of* A *equal* d.

Here $R(A)$ denotes the range of A. Proposition 4.2 can be easily proved by inspecting the Jordan form of A.

We thank D. Hershkowitz for reading the manuscript and suggesting several improvements in the exposition.

REFERENCES

1 R. Bru and M. Lopez-Pellicer, Extensions of algebraic Jordan basis, *Glas. Mat.* 20(40):289–292 (1985).
2 I. Gohberg, P. Lancaster and L. Rodman, Invariant subspaces of matrices with applications, Wiley, 1986.
3 A. Pastor and V. Hernandez, The class of Hermitian and nonnegative definite solutions of the algebraic. Riccati equation (preprint), presented at the Conference on Riccati Equations in Systems, Signals and Control (S. Bittanti, A. J. Laub, J. C. Willems, eds.), Como, Italy, 1989.

Received 16 November 1989; final manuscript accepted 29 May 1990