

A PROOF OF THE BOX CONJECTURE FOR COMMUTING PAIRS OF MATRICES

JOHN IRVING, TOMAŽ KOŠIR, AND MITJA MASTNAK

ABSTRACT. We prove the Box Conjecture for pairs of commuting nilpotent matrices, as formulated by Iarrobino et al [28]. This describes the Jordan type of the dense orbit in the nilpotent commutator of a given nilpotent matrix. Our main tool is the Burge correspondence between the set of all partitions and a set of binary words [15, 16]. For connection with the algebraic and geometric setup of matrices and orbits we employ some of Shayman's results on invariant subspaces of a nilpotent matrix [45, 46]. Our proof is valid over an arbitrary field.

1. INTRODUCTION

The study of commuting matrices has a long history (see e.g. [18, 19, 22, 50]). As such it is somewhat surprising that the question of which pairs of conjugacy classes can occur for pairs of commuting matrices has been considered only relatively recently. The question reduces to the case of commuting nilpotent matrices. Fix an infinite ground field F . For a given nilpotent matrix B (over F), the set \mathcal{N}_B of all nilpotent matrices commuting with B is an irreducible algebraic variety [8, Lem. 2.3]. Its intersection with one of the conjugacy classes of nilpotent matrices is a dense open subset. Let \mathcal{P} be the set of all partitions of natural numbers together with the empty partition ε . The partition $P \in \mathcal{P}$ determined by the Jordan canonical form of B is called the *Jordan type* of B . Following Panyushev [42] we denote by $\mathfrak{D}(P)$ the partition determining the Jordan type of the dense conjugacy class in \mathcal{N}_B .

The map $\mathfrak{D} : \mathcal{P} \rightarrow \mathcal{P}$ has been studied by several authors; see e.g. [9, 10, 11, 32, 33, 28, 36, 42] and also the recent survey by Khatami [34]. Map \mathfrak{D} is idempotent [36, Thm. 6] and its image is the set \mathcal{Q} of *super-distinct* partitions, i.e., those whose parts differ by at least two [10, Thm. 1.12], [42, Thm. 2.1]. The number of parts of $\mathfrak{D}(P)$ was determined by Basili [7, Prop. 2.4]. Oblak later identified the largest part of $\mathfrak{D}(P)$ [39, Thm. 13] and Khatami the smallest [33, Thm. 4.1]. Oblak also conjectured a recursive process to compute $\mathfrak{D}(P)$. Iarrobino and Khatami [27] showed that Oblak's process produces a lower bound (in dominance order), and Basili [9] proved that the algorithm in fact yields $\mathfrak{D}(P)$.

The inverse image under \mathfrak{D} is not as well understood, owing to the apparent difficulty of unwinding Oblak's process. Iarrobino, Khatami, Van Steirteghem, and Zhao [28] described $\mathfrak{D}^{-1}(Q)$ for partitions $Q \in \mathcal{Q}$ having at most two parts and they also formulated the *Box Conjecture* [28, Conj. 4.11] regarding the structure of $\mathfrak{D}^{-1}(Q)$ in general. The name of the conjecture comes from its precise statement: Given a partition $Q = (q_1, q_2, \dots, q_k) \in \mathcal{Q}$ with parts $q_1 > q_2 > \dots > q_k$ differing by at least 2, the elements of $\mathfrak{D}^{-1}(Q)$ can be arranged in a box (i.e. array) of size $q_k \times (q_{k-1} - q_k - 1) \times \dots \times (q_1 - q_2 - 1)$ such that the partition in the (i_1, i_2, \dots, i_k) -th position has exactly $\sum_{j=1}^k i_j$ parts.

Date: April 4, 2024.

2020 *Mathematics Subject Classification.* 15A27, 05A17, 15A21, 13E10, 14A25.

We prove the Box Conjecture and provide further insight into the structure of partitions in $\mathfrak{D}^{-1}(Q)$. Our principal algebraic tool is the structure theory of invariant subspaces as presented by Shayman [45, 46], but our argument hinges on a combinatorial correspondence due to Burge [15, 16] that encodes partitions as binary words.

Let $\Omega : \mathcal{P} \rightarrow \mathcal{W}$ denote the Burge encoding, where $\mathcal{W} = (\alpha^*\beta)^*\alpha$ is the set of words on the alphabet $\{\alpha, \beta\}$ that end with a single α . (The encoding will be described in detail in Section 2.) Our main result is the following:

Theorem 1. *Suppose $P \in \mathcal{P}$ has Burge code $\Omega(P) = \omega_1 \cdots \omega_n$. Then $\mathfrak{D}(P) = (q_1, q_2, \dots, q_k)$, where $q_1 > q_2 > \cdots > q_k$ is the complete list of indices q for which $\omega_q = \beta$ and $\omega_{q+1} = \alpha$.*

For instance, we will find that $P = (5, 3, 2, 2, 1)$ has Burge code $\Omega(P) = \beta\alpha\beta\alpha\alpha\beta\beta\beta\alpha$ and thus Theorem 1 gives $\mathfrak{D}(P) = (9, 3, 1)$. In general, $\Omega(P)$ can be obtained from P via a straightforward iterative process, so Theorem 1 provides another recursive algorithm for computing $\mathfrak{D}(P)$, distinct from (but related to) Oblak's.

The q_i in Theorem 1 clearly differ pairwise by at least 2, so the theorem immediately gives $\mathfrak{D}(P) \in \mathcal{Q}$ for all P . Moreover, a simple characterization of super-distinct partitions in terms of the Burge correspondence (Proposition 16) will reveal that Theorem 1 implies $P \in \mathcal{Q}$ if and only if $\mathfrak{D}(P) = P$.

We will also see that the number of copies of β in $\Omega(P)$ equals the number of parts of P (Proposition 12). The Box Conjecture then follows from Theorem 1 simply by specifying the words in \mathcal{W} that contain substring $\beta\alpha$ at prescribed positions.

Corollary 2. *Let $Q = (q_1, q_2, \dots, q_k) \in \mathcal{Q}$ and set $\delta_1 = q_k$ and $\delta_i = q_{k-i+1} - q_{k-i+2} - 1$ for $2 \leq i \leq k$. Then $\mathfrak{D}^{-1}(Q)$ is of size $\delta_1\delta_2 \cdots \delta_k$ and consists of precisely those partitions whose Burge code is of the form*

$$(1) \quad \alpha^{\delta_1-i_1} \beta^{i_1} \alpha^{\delta_2-i_2+1} \beta^{i_2} \alpha^{\delta_3-i_3+1} \beta^{i_3} \cdots \alpha^{\delta_k-i_k+1} \beta^{i_k} \alpha,$$

where $(i_1, \dots, i_k) \in [1, \delta_1] \times [1, \delta_2] \times \cdots \times [1, \delta_k]$. Moreover, the partition determined by (1) has exactly $\sum_j i_j$ parts.

We noted above that super-distinct partitions are stable under \mathfrak{D} , that is, $Q \in \mathfrak{D}^{-1}(Q)$ for all $Q \in \mathcal{Q}$. This corresponds to taking $(i_1, \dots, i_k) = (1, \dots, 1)$ in Corollary 2.

Example 3. *For $Q = (10, 7, 3) \in \mathcal{Q}$ we have $(\delta_1, \delta_2, \delta_3) = (3, 7 - 3 - 1, 10 - 7 - 1) = (3, 3, 2)$, so $\mathfrak{D}^{-1}(Q)$ contains 18 partitions indexed by coordinates in the box $[1, 3] \times [1, 3] \times [1, 2]$. These are displayed in Figure 1.*

Theorem 1 will be proved in Section 3 after we have established the necessary combinatorial framework in Section 2. Our proof does not depend on the nature of the underlying field F , except insofar as the definition of $\mathfrak{D}(P)$ requires the assumption that F is infinite. (We explain an extension to finite fields in §3.5.) In Section 4 we explore the interaction between the Burge encoding and Oblak's recursive scheme for computing $\mathfrak{D}(P)$. This leads to a new proof (independent of Basili [9]) that Oblak's algorithm does indeed output $\mathfrak{D}(P)$.

Related Work. Other closely related topics have also been studied over the last several decades. For some partial results on pairs of partitions describing commuting nilpotent matrices see [12, 40]. Irreducibility of the variety of pairs of commuting nilpotent matrices was proved by Baranovsky [6] (see also Basili [8]). In the general Lie algebra setting, the question of irreducibility of the nilpotent commuting variety was resolved by Premet [43]. Nilpotent and other commuting varieties of triples (and d -tuples, for larger d) of matrices have also been considered, see e.g., [22, 23, 24, 25, 26, 31, 37, 38, 47, 48, 49]. The connection

(i_1, i_2, i_3)	code ω	partition $\Omega^{-1}(\omega)$	# parts
(1, 1, 1)	$\alpha\alpha\beta\alpha\alpha\alpha\beta\alpha\alpha\beta\alpha$	$[10, 7, 3]$	3
(2, 1, 1)	$\alpha\beta\beta\alpha\alpha\alpha\beta\alpha\alpha\beta\alpha$	$[10, 7, 2, 1]$	4
(3, 1, 1)	$\beta\beta\beta\alpha\alpha\alpha\beta\alpha\alpha\beta\alpha$	$[10, 7, 1^3]$	5
(1, 2, 1)	$\alpha\alpha\beta\alpha\alpha\beta\beta\alpha\alpha\beta\alpha$	$[10, 5, 3, 2]$	4
(2, 2, 1)	$\alpha\beta\beta\alpha\alpha\beta\beta\alpha\alpha\beta\alpha$	$[10, 4, 3, 2, 1]$	5
(3, 2, 1)	$\beta\beta\beta\alpha\alpha\beta\beta\alpha\alpha\beta\alpha$	$[10, 4, 3, 1^3]$	6
(1, 3, 1)	$\alpha\alpha\beta\alpha\beta\beta\beta\alpha\alpha\beta\alpha$	$[10, 5, 2^2, 1]$	5
(2, 3, 1)	$\alpha\beta\beta\alpha\beta\beta\beta\alpha\alpha\beta\alpha$	$[10, 5, 2, 1^3]$	6
(3, 3, 1)	$\beta\beta\beta\alpha\beta\beta\beta\alpha\alpha\beta\alpha$	$[10, 5, 1^5]$	7
(1, 1, 2)	$\alpha\alpha\beta\alpha\alpha\alpha\beta\alpha\beta\beta\alpha$	$[9, 5, 3^2]$	4
(2, 1, 2)	$\alpha\beta\beta\alpha\alpha\alpha\beta\alpha\beta\beta\alpha$	$[9, 4^2, 2, 1]$	5
(3, 1, 2)	$\beta\beta\beta\alpha\alpha\alpha\beta\alpha\beta\beta\alpha$	$[9, 4^2, 1^3]$	6
(1, 2, 2)	$\alpha\alpha\beta\alpha\alpha\beta\beta\alpha\beta\beta\alpha$	$[9, 5, 2^3]$	5
(2, 2, 2)	$\alpha\beta\beta\alpha\alpha\beta\beta\alpha\beta\beta\alpha$	$[9, 4, 3, 2, 1^2]$	6
(3, 2, 2)	$\beta\beta\beta\alpha\alpha\beta\beta\alpha\beta\beta\alpha$	$[9, 4, 3, 1^4]$	7
(1, 3, 2)	$\alpha\alpha\beta\alpha\beta\beta\beta\alpha\beta\beta\alpha$	$[9, 5, 2^2, 1^2]$	6
(2, 3, 2)	$\alpha\beta\beta\alpha\beta\beta\beta\alpha\beta\beta\alpha$	$[9, 5, 2, 1^4]$	7
(3, 3, 2)	$\beta\beta\beta\alpha\beta\beta\beta\alpha\beta\beta\alpha$	$[9, 5, 1^6]$	8

FIGURE 1. The elements of $\mathfrak{D}^{-1}(10, 7, 3)$ and the corresponding Burge codes, indexed as in Corollary 2 by their coordinates $(i_1, i_2, i_3) \in [1, 3] \times [1, 3] \times [1, 2]$.

between commuting nilpotent matrices and Artinian algebras was studied extensively in recent years (e.g. [29]); see [1] for a survey of the topic. Finally, we note that results on commuting matrices have been applied to multiparameter eigenvalue problems [5, 35], the area of research that originated from applications to boundary value problems for elliptic partial differential equations [51].

Burge's correspondence provides a natural description of all Jordan types of matrices that have equal dense or dominant orbit in the nilpotent commutator. It is an example of a combinatorial description of objects that are of great interest in algebra and geometry, and have interesting applications in physics [30].

2. BURGE CODES AND THE DESCENT MAP

The goal of this section is to describe Burge's elegant bijection [15] between partitions \mathcal{P} and binary words \mathcal{W} . This is best expressed in terms of the frequency representation of partitions, which we introduce in §2.1. The correspondence itself is defined in §2.2. All results presented therein can be found (at least implicitly) in [15], but we offer a complete exposition in the interests of fixing notation and making this article self-contained. In §2.3 we define the *descent map* on \mathcal{P} . This key transformation arises naturally from Burge's construction and lies at the heart of our results. Finally, in §2.4, we describe how another bijection of Burge transforms the elements of $\mathfrak{D}^{-1}(Q)$ into partitions whose diagonal hook lengths are given by Q .

2.1. Preliminaries. Recall that a *partition* $P \in \mathcal{P}$ is a weakly decreasing sequence (p_1, \dots, p_k) of positive integers. The p_i are called the *parts* of P . The *size* of P is the sum of the parts, $|P| := \sum_i p_i$, and the *length* of P is the number of parts, $\ell(P) := k$. The *empty partition* is the unique element $\varepsilon \in \mathcal{P}$ of size (and length) 0, characterized by $|\varepsilon| = 0$ or $\ell(\varepsilon) = 0$.

Partitions can equivalently be regarded as unordered multisets of positive integers. We use the standard multiset notation $[1^{f_1}, 2^{f_2}, \dots]$ to denote the partition whose parts consist of f_1 copies of 1, f_2 copies of 2, etc.

The *2-measure* of a partition P , denoted $\mu_2(P)$, is the maximum length of a super-distinct subpartition of P . For example, $P = (8, 7, 4, 4, 3, 2, 2, 1)$ contains the subpartition $(7, 4, 1) \in \mathcal{Q}$ of length 3 and none longer, so $\mu_2(P) = 3$. This statistic was recently studied by Andrews et al [4] through the lens of q -series. Here it will play a fundamental combinatorial role. Clearly we have $0 \leq \mu_2(P) \leq \ell(P)$, with $\mu_2(P) = 0 \iff P = \varepsilon$, and $\mu_2(P) = \ell(P) \iff P \in \mathcal{Q}$.

Let \mathcal{F} be the set of all finitely supported sequences of nonnegative integers. We will work extensively with partitions via their “frequency” representations in \mathcal{F} given by the trivial correspondence $[1^{f_1} 2^{f_2} \dots] \leftrightarrow (f_1, f_2, \dots)$. This identification of \mathcal{F} and \mathcal{P} should be kept in mind throughout, but when we wish to be explicit we will write $f(P)$ for the frequency sequence of $P \in \mathcal{P}$ and $P(f)$ for the partition corresponding to $f \in \mathcal{F}$.

Convention 1. *When working with generic $f \in \mathcal{F}$, we adopt the convention that $f_0 = 0$. When working with specific $f \in \mathcal{F}$, trailing zeroes will often be omitted; e.g. $(0, 3, 2, 0, 1) = (0, 3, 2, 0, 1, 0, 0, \dots)$.*

We now introduce various nomenclature related to \mathcal{F} . The *support* of $f \in \mathcal{F}$ will be denoted $\mathcal{S}(f) := \{i \geq 1 : f_i \neq 0\}$. The size and length statistics on \mathcal{P} are extended to \mathcal{F} in the obvious way, namely

$$|f| := |P(f)| = \sum_i i f_i, \quad \text{and} \quad \ell(f) := \ell(P(f)) = \sum_i f_i.$$

Similarly we let $\mu_2(f) := \mu_2(P(f))$. Note that this is the maximum size of a subset of $\mathcal{S}(f)$ that contains no consecutive pairs $\{i, i+1\}$.

A *spread* of f is defined to be a maximal interval $[i, j] \subseteq \mathcal{S}(f)$, and a spread $[i, i] = \{i\}$ of size 1 is said to be *trivial*. Finally, we define sets

$$\begin{aligned} \mathbf{L}(f) &:= \bigcup \{i, i+2, \dots, i+2\lfloor \frac{i-i}{2} \rfloor\} \\ \mathbf{R}(f) &:= \bigcup \{j, j-2, \dots, j-2\lfloor \frac{j-i}{2} \rfloor\}, \end{aligned}$$

where the unions extend over all spreads $[i, j]$ of f . The importance of $\mathbf{L}(f)$ and $\mathbf{R}(f)$ will be apparent shortly. For now we observe only that they are of equal size, namely $\mu_2(f) = \sum \lceil \frac{j-i+1}{2} \rceil$.

Example 4. *Let $f = (2, 1, 0, 3, 2, 2, 0, 0, 1)$. The corresponding partition is $P = P(f) = [9, 6^2, 5^2, 4^3, 2, 1^2]$, and $|f| = |P| = 47$ and $\ell(f) = \ell(P) = 11$. The spreads of f are $\{1, 2\}$, $\{4, 5, 6\}$ and $\{9\}$, so $\mathbf{L}(f) = \{1, 4, 6, 9\}$, $\mathbf{R}(f) = \{2, 4, 6, 9\}$ and $\mu_2(f) = \mu_2(P) = 4$.*

2.2. The Burge Correspondence. We begin by recasting the sets $\mathbf{L}(f)$ and $\mathbf{R}(f)$, defined above, in terms of certain partial pairings of the entries of $f \in \mathcal{F}$. Observe that the elements $i_1 < i_2 < \dots < i_m$ of $\mathbf{L}(f)$ specify the pairs $(f_{i_1}, f_{i_1+1}), \dots, (f_{i_m}, f_{i_m+1})$ that result from parsing f from left-to-right and grouping consecutive entries f_i, f_{i+1} with $f_i > 0$. We call these the *forward pairs* of f . Similarly, the elements $j_1 > j_2 > \dots > j_m$ of $\mathbf{R}(f)$ determine the *backward pairs* $(f_{j_1-1}, f_{j_1}), \dots, (f_{j_m-1}, f_{j_m})$ obtained by parsing f from right-to-left and grouping consecutive entries f_j, f_{j-1} with $f_j > 0$. Here we rely on the convention $f_0 = 0$ to allow for the degenerate backward pair (f_0, f_1) , which arises if and only if $1 \in \mathbf{R}(f)$.

Example 5. *For $f = (2, 2, 1, 3, 1, 0, 4, 0, 0, 2, 1)$ we have $\mathbf{L}(f) = \{1, 3, 5, 7, 10\}$ and $\mathbf{R}(f) = \{1, 3, 5, 7, 11\}$. The forward/backward pairs of f are indicated below with arrows pointed in*

the direction of parsing.

$$\begin{array}{ccc} \overrightarrow{0} \overrightarrow{(2, 2, 1, 3, 1, 0, 4, 0, 0, 2, 1)} & & \overleftarrow{0} \overleftarrow{(2, 2, 1, 3, 1, 0, 4, 0, 0, 2, 1)} \\ \text{forward} & & \text{backward} \end{array}$$

The fictional entry $f_0 = 0$ has been prepended in red. Observe that every nonzero entry of f appears in one forward and one backward pair, while all unpaired entries of f are 0.

Let $\mathcal{A} := \{f \in \mathcal{F} : 1 \notin R(f)\}$ and $\mathcal{B} := \{f \in \mathcal{F} : 1 \in R(f)\}$, observing that these sets partition \mathcal{F} . Also make note of the equivalent characterizations $f \in \mathcal{B} \iff (f_0, f_1)$ is a backward pair $\iff 1$ is contained in a spread of f of odd size.

We now introduce two related transformations $\alpha : \mathcal{F} \rightarrow \mathcal{A}$ and $\beta : \mathcal{F} \rightarrow \mathcal{B}$, along with a mapping $\partial : \mathcal{F} \rightarrow \mathcal{F}$ that undoes them. Each of these acts as a sequence of raising/lowering operators on the forward or backward pairs of f . The precise recipes are as follows:

- $\alpha(f)$ is obtained from f by replacing (f_i, f_{i+1}) with $(f_i - 1, f_{i+1} + 1)$ for each $i \in L(f)$.
- $\beta(f) := (f_1 + 1, \alpha(f_2, f_3, \dots))$
- $\partial(f)$ is obtained from f by replacing (f_{j-1}, f_j) with $(f_{j-1} + 1, f_j - 1)$ for each $j \in R(f)$, where this is understood to mean only that f_1 is reduced by 1 in the case $j = 1$ (i.e., the fictional f_0 is ignored).

In effect, α scans f from left-to-right and “promotes” each forward pair (f_i, f_{i+1}) to $(f_i - 1, f_{i+1} + 1)$, thereby transforming the forward pairs of f into the backward pairs of $\alpha(f)$. (That is, $i \in L(f) \iff i + 1 \in R(\alpha(f))$.) In particular, for any $f \in \mathcal{F}$ we have $1 \notin R(\alpha(f))$ and hence $1 \in R(\beta(f))$, validating the claim that α and β map \mathcal{F} into \mathcal{A} and \mathcal{B} , respectively. Moreover, f can be recovered from either $\alpha(f)$ or $\beta(f)$ by scanning from right-to-left and “demoting” backward pairs. This is exactly the action of ∂ .

Lemma 6. *Let α, β, ∂ be defined on \mathcal{F} as above. Then α and β are bijections from \mathcal{F} to \mathcal{A} and \mathcal{B} , respectively, while $\partial : \mathcal{F} \rightarrow \mathcal{F}$ is 2-to-1 and restricts to α^{-1} on \mathcal{A} and β^{-1} on \mathcal{B} .*

Example 7. *Let $f = (2, 2, 1, 3, 1, 0, 4, 0, 0, 2, 1)$ as in Example 5. Then*

$$\begin{aligned} \alpha(f) &= \alpha(\overrightarrow{2, 2, 1, 3, 1, 0, 4, 0, 0, 2, 1}) = (1, 3, 0, 4, 0, 1, 3, 1, 0, 1, 2) \\ \beta(f) &= (2 + 1, \alpha(\overrightarrow{2, 1, 3, 1, 0, 4, 0, 0, 2, 1})) = (3, 1, 2, 2, 2, 0, 3, 1, 0, 1, 2) \\ \partial(f) &= \partial(\overleftarrow{2, 2, 1, 3, 1, 0, 4, 0, 0, 2, 1}) = (1, 3, 0, 4, 0, 1, 3, 0, 0, 3, 0). \end{aligned}$$

Forward/backward pairs are promoted/demoted by transferring one unit along each arrow. For the degenerate backward pair (f_0, f_1) , which corresponds to $1 \in R(f)$ (i.e., $f \in \mathcal{B}$), demotion by ∂ amounts to reducing f_1 by 1. We leave it to the reader to verify that $\partial(\alpha(f)) = \partial(\beta(f)) = \beta(\partial(f)) = f$. Note, however, that $\alpha(\partial(f)) = (0, 4, 0, 3, 1, 0, 4, 0, 0, 2, 1) \neq f$. This inequality is due to the fact that $f \notin \mathcal{A}$.

Clearly we have $0 \leq |\partial(f)| < |f|$ for all $f \neq \varepsilon$. Applying ∂ repeatedly to any $f \in \mathcal{F}$ will therefore result in $\partial^k f = \varepsilon$ for some least positive integer $k \geq 1$. This leads to the following definitions.

Definition 8. *We define the Burge chain of $f \in \mathcal{F}$ to be the sequence $\partial^0 f, \partial^1 f, \partial^2 f, \dots, \partial^k f$, where k is the smallest positive integer satisfying $\partial^k f = \varepsilon$. The Burge code of f is the binary word $\Omega(f) := \omega_1 \omega_2 \dots \omega_{k+1} \in \{\alpha, \beta\}^*$ defined by*

$$\omega_i = \begin{cases} \alpha & \text{if } \partial^{i-1} f \in \mathcal{A}, \\ \beta & \text{if } \partial^{i-1} f \in \mathcal{B}. \end{cases}$$

So $f = \varepsilon$ has trivial Burge chain ε and Burge code $\Omega(\varepsilon) = \alpha$, whereas the chains/codes of all other sequences f are of length at least 2. In fact since $\partial f = \varepsilon$ only for $f = \varepsilon$ and $f = (1)$, the chain of every $f \neq \varepsilon$ ends with $\partial^{k-1} f = (1) \in \mathcal{B}$ and $\partial^k f = \varepsilon \in \mathcal{A}$, and thus the code $\Omega(f)$ ends with $\omega_k \omega_{k+1} = \beta\alpha$.

Since $\partial|_{\mathcal{A}} = \alpha^{-1}$ and $\partial|_{\mathcal{B}} = \beta^{-1}$, the definition of $\Omega(f) = \omega_1 \cdots \omega_n$ ensures that $\partial^{i-1} f = (\omega_i \circ \partial)(\partial^{i-1} f) = \omega_i(\partial^i f)$ for all i . Thus we have

$$f = \omega_1(\partial^1 f) = \omega_1 \omega_2(\partial^2 f) = \omega_1 \omega_2 \omega_3(\partial^3 f) = \cdots = (\omega_1 \omega_2 \cdots \omega_n)(\varepsilon),$$

where the products of ω_i are to be interpreted as functional composition in the usual right-to-left order. That is, the right-to-left reading of $\Omega(f)$ specifies the unique manner by which f can be “built” from ε via iterative applications of α and β , beginning with a single application of α .

Proposition 9. *The Burge encoding $f \mapsto \Omega(f)$ is a one-one correspondence between \mathcal{F} and the set $\mathcal{W} := (\alpha^* \beta)^* \alpha$ of all finite words on $\{\alpha, \beta\}$ that end with a singleton α .*

Example 10. *Let $f = (1, 2, 1, 0, 1)$. The Burge chain of f is displayed below, along with the values of ω_i along the chain. The Burge code $\Omega(f) = \beta\alpha\beta\alpha\alpha\alpha\beta\beta\beta\alpha$ is obtained by concatenating these values.*

i	$\partial^i f$	ω_{i+1}
0	$(1, 2, 1, 0, 1)$	β
1	$(0, 3, 0, 1)$	α
2	$(1, 2, 1)$	β
3	$(0, 3)$	α
4	$(1, 2)$	α
5	$(2, 1)$	α
6	(3)	β
7	(2)	β
8	(1)	β
9	ε	α

Conversely, given the word $\omega = \alpha\beta\beta\alpha\beta\alpha \in \mathcal{W}$ we build its corresponding partition by iterating α and β as follows:

$$\varepsilon \xrightarrow{\alpha} \varepsilon \xrightarrow{\beta} (1) \xrightarrow{\alpha} (0, 1) \xrightarrow{\beta} (1, 0, 1) \xrightarrow{\beta} (2, 0, 0, 1) \xrightarrow{\alpha} (1, 1, 0, 0, 1).$$

Thus ω is the Burge code of $f = (1, 1, 0, 0, 1)$.

We now describe how size, length and 2-measure are affected by ∂ , and use this to express $|f|$, $\ell(f)$ and $\mu_2(f)$ in terms of familiar statistics on $\Omega(f)$. For convenience we hereon omit parentheses and write ∂f instead of $\partial(f)$.

Lemma 11. *For any $f \in \mathcal{F}$ we have:*

- (1) $\ell(\partial f) = \begin{cases} \ell(f) - 1 & \text{if } f \in \mathcal{B} \\ \ell(f) & \text{otherwise.} \end{cases}$
- (2) $|\partial f| = |f| - \mu_2(f)$
- (3) $\mu_2(\partial f) = \begin{cases} \mu_2(f) - 1 & \text{if } f \in \mathcal{B} \text{ and } \partial f \in \mathcal{A} \\ \mu_2(f) & \text{otherwise.} \end{cases}$

Proof. The sum $\ell(f) = \sum_i f_i$ is invariant under the demotion $(f_{j-1}, f_j) \rightsquigarrow (f_{j-1} + 1, f_j - 1)$ except when $j = 1$, in which case it is decreased by 1, whereas $|f| = \sum_i i f_i$ is decreased by 1 by every such demotion. Claims (1) and (2) follow.

For (3), first observe that if $f \in \mathcal{A}$ then ∂ maps the backward pairs of f to the forward pairs of ∂f , i.e., $i \in \mathbf{R}(f) \iff i-1 \in \mathbf{L}(\partial f)$. Hence $\mu_2(\partial f) = |\mathbf{L}(\partial f)| = |\mathbf{R}(f)| = \mu_2(f)$ in this case.

Now suppose $f \in \mathcal{B}$ and let $g = (f_2, f_3, \dots)$. Then $g \in \mathcal{A}$ and $\partial f = (f_1-1, \partial g)$. Interpreting $\mu_2(\cdot)$ as a count of backward pairs gives

$$\mu_2(f) = 1 + \mu_2(g) \quad \text{and} \quad \mu_2(\partial f) = \begin{cases} \mu_2(\partial g) & \text{if } \partial f \in \mathcal{A} \\ 1 + \mu_2(\partial g) & \text{if } \partial f \in \mathcal{B}. \end{cases}$$

But we already established that $g \in \mathcal{A}$ implies $\mu_2(g) = \mu_2(\partial g)$, so the above identities prove (3) in the case $f \in \mathcal{B}$. \square

Let $\omega = \omega_1 \omega_2 \dots \omega_n \in \{\alpha, \beta\}^*$. The *descent set* of ω , denoted $\text{Des}(\omega)$, is the set of all indices i for which $\omega_i \omega_{i+1} = \beta\alpha$. That is, $\text{Des}(\omega)$ records the positions of descents in ω relative to the ordering $\beta > \alpha$. The *descent number* and *major index* of ω are defined by $\text{des}(\omega) := |\text{Des}(\omega)|$ and $\text{maj}(\omega) := \sum_{i \in \text{Des}(\omega)} i$, respectively.

Proposition 12. *Let $f \in \mathcal{F}$ and let $\omega = \Omega(f)$. Then:*

- (1) $\ell(f) = \# \text{ occurrences of } \beta \text{ in } \omega$
- (2) $|f| = \text{maj}(\omega)$
- (3) $\mu_2(f) = \text{des}(\omega) = \# \text{ occurrences of } \beta\alpha \text{ in } \omega$

Example 13. In Example 10 we saw that $f = (1, 2, 1, 0, 1)$ has Burge code $\omega = \beta\alpha\beta\alpha\alpha\beta\beta\beta\alpha$. Note that ω contains 5 = $\ell(f)$ copies of β , and we have $\text{Des}(\omega) = \{1, 3, 9\}$, $\text{maj}(\omega) = 13 = |f|$ and $\text{des}(\omega) = 3 = \mu_2(f)$.

Proof of Proposition 12: Claims (1) and (3) are immediate from the corresponding parts of Lemma 11. To prove (2), let $\omega = \omega_1 \omega_2 \dots \omega_{k+1}$ and $\text{Des}(\omega) = \{i_1, i_2, \dots, i_m\}$, where $1 \leq i_1 < i_2 < \dots < i_m = k$. Iterating part 2 of the lemma gives $|f| = \sum_{r=0}^k \mu_2(\partial^r f)$, while iterating part 3 gives $\mu_2(\partial^r f) = m - j$ for r in the range $i_j \leq r < i_{j+1}$, where we take $i_0 = 0$ and $i_{m+1} = \infty$. Hence $|f| = \sum_{j=0}^{m-1} (m - j)(i_{j+1} - i_j) = \sum_{j=0}^{m-1} i_{j+1} = \text{maj}(\omega)$. \square

2.3. The Descent Map. We have described a three-way equivalence between partitions, frequency vectors, and binary words, namely

$$\mathcal{P} \xrightarrow{\text{frequency } f} \mathcal{F} \xrightarrow{\text{Burge code } \Omega} \mathcal{W}.$$

With this in mind we adopt the following convention.

Convention 2. *All notation defined previously for elements of \mathcal{F} and \mathcal{W} is extended to \mathcal{P} through $f(\cdot)$ and $\Omega(\cdot)$. In particular, we define $\partial P := \partial(f(P))$, $\partial^i P := \partial^i f(P)$, $\Omega(P) := \Omega(f(P))$ and then $\text{Des}(P) := \text{Des}(\Omega(P))$.*

Suppose $\text{Des}(P) = \{i_1, i_2, \dots, i_m\}$ where $i_1 < i_2 < \dots < i_m$. Then evidently $i_j - i_{j-1} \geq 2$ for all j , and Proposition 12(2) gives $\sum_j i_j = \text{maj}(\Omega(P)) = |P|$. Thus $\text{Des}(P)$ can be regarded as a super-distinct partition of size $|P|$. Moreover, if $\Omega(P) = \omega_1 \omega_2 \dots \omega_n$ then $\Omega(\partial P) = \omega_2 \dots \omega_n$, so we have $\text{Des}(\partial P) = \{i_1 - 1, i_2 - 1, \dots, i_m - 1\} \setminus \{0\}$. See, for example, Figure 2.

Definition 14. *For a partition P , let $P - 1$ denote the reduced partition obtained by subtracting 1 from each part of P and eliminating any resulting zeros.*

We have just seen that the *descent map* $P \mapsto \text{Des}(P)$ is a size-preserving function from \mathcal{P} to \mathcal{Q} satisfying $\text{Des}(\partial P) = \text{Des}(P) - 1$. It is easy to see that this is the unique such function.

i	$\partial^i f$	$\Omega(\partial^i f)$	$\partial^i P$	$\text{Des}(\partial^i P)$
0	(1, 1, 0, 1, 0, 0, 1)	$\alpha\beta\alpha\beta\beta\alpha\beta\alpha$	[7, 4, 2, 1]	[7, 5, 2]
1	(2, 0, 1, 0, 0, 1)	$\beta\alpha\beta\beta\alpha\beta\alpha$	[6, 3, 1 ²]	[6, 4, 1]
2	(1, 1, 0, 0, 1)	$\alpha\beta\beta\alpha\beta\alpha$	[5, 2, 1]	[5, 3]
3	(2, 0, 0, 1)	$\beta\beta\alpha\beta\alpha$	[4, 1 ²]	[4, 2]
4	(1, 0, 1)	$\beta\alpha\beta\alpha$	[3, 1]	[3, 1]
5	(0, 1)	$\alpha\beta\alpha$	[2]	[2]
6	(1)	$\beta\alpha$	[1]	[1]
7	ε	α	ε	ε

FIGURE 2. The evolution of $\text{Des}(\partial^i P)$ along a Burge chain.

Lemma 15. *Let $\mathcal{P}' \subseteq \mathcal{P}$ be a nonempty set of partitions such that for every $P \in \mathcal{P}'$, $\partial P \in \mathcal{P}'$. Suppose $\Omega : \mathcal{P}' \rightarrow \mathcal{P}$ satisfies*

- (1) $|\Omega(P)| = |P|$ and
- (2) $\Omega(\partial P) = \Omega(P) - 1$.

Then $\Omega(P) = \text{Des}(P)$ for all $P \in \mathcal{P}'$.

Proof. Condition (1) forces $\Omega(\varepsilon) = \varepsilon$, and if $\Omega(\partial P)$ has frequencies (f_1, f_2, \dots) then (1) and (2) force $\Omega(P)$ to have frequencies $(|P| - |\partial P| - \sum_i f_i, f_1, f_2, \dots)$. So $\Omega(P)$ is determined uniquely by induction on $|P|$. \square

Theorem 1 is equivalent to statement that $\mathfrak{D}(P) = \text{Des}(P)$ for all $P \in \mathcal{P}$. In Section 3 we prove this identity by showing that \mathfrak{D} satisfies the hypothesis of Lemma 15.

We now present a multi-faceted characterization of super-distinct partitions. Observe that the equivalence $P \in \mathcal{Q} \Leftrightarrow \text{Des}(P) = P$ together with Theorem 1 imply that $P \in \mathcal{Q}$ if and only if $\mathfrak{D}(P) = P$.

Proposition 16. *Let $P \in \mathcal{P}$ and $f = f(P)$. The following are equivalent:*

- (1) $P \in \mathcal{Q}$
- (2) $\ell(P) = \mu_2(P)$
- (3) $\Omega(P)$ does not contain the substring $\beta\beta$
- (4) $f_i = 1$ if $i \in \mathbf{R}(f)$ and $f_i = 0$ otherwise
- (5) $\partial f = (f_2, f_3, \dots)$
- (6) $\partial P = P - 1$
- (7) $\text{Des}(P) = P$

Proof. Equivalences (1) \Leftrightarrow (2), (1) \Leftrightarrow (4) and (5) \Leftrightarrow (6) are clear from the respective definitions, while (2) \Leftrightarrow (3) is immediate from the fact that $\ell(P)$ and $\mu_2(P)$ count occurrences of β and $\beta\alpha$ in $\Omega(P)$. (See Proposition 12.)

If f satisfies (4) then by definition of ∂ we have $(\partial f)_i = 1$ if $i + 1 \in \mathbf{R}(f)$ and $(\partial f)_i = 0$ otherwise. Thus (4) \Rightarrow (5). Conversely, suppose $\partial f = (f_2, f_3, \dots)$. Say $\mathbf{R}(f) = \{i_1, \dots, i_k\}$ with $i_1 > \dots > i_k$. Then $(f_{i_j-1} + 1, f_{i_j} - 1) = (f_{i_j}, f_{i_j+1})$ for all j . Since $f_{i_1+1} = 0$ we have $(f_{i_1-1}, f_{i_1}) = (0, 1)$, which in turn implies $f_{i_2+1} = 0$ and therefore $(f_{i_2-1}, f_{i_2}) = (0, 1)$, etc. Thus (5) \Rightarrow (4).

We know $\text{Des}(P) \in \mathcal{Q}$ for all P , so clearly (7) \Rightarrow (1). The remaining implication, (1) \Rightarrow (7), follows by applying Lemma 15 to $\Omega : \mathcal{Q} \rightarrow \mathcal{P}$ given by $\Omega(T) = T$. Note that \mathcal{Q} is ∂ -invariant and that this function satisfies the conditions of the lemma due to the fact that, as proven above, (1) implies (6). \square

2.4. Additional Remarks. Burge's work [15, 16] provides additional bijections between partitions and $\{\alpha, \beta\}$ -words. One of these [15, §2] interprets α and β as operators on \mathcal{P} that enlarge the diagonal hooks of a partition (viewed as a Young diagram or Ferrer's graph) along the horizontal and vertical axes, respectively. Together with Corollary 2, it is easy to check that this gives a correspondence between $\mathfrak{D}^{-1}(Q)$ and the set of partitions with hook lengths q_1, \dots, q_k (and hence Durfee square of size $k = \ell(Q)$). Such a mapping was sought by the proposers of the Box Conjecture, who inferred its existence enumeratively; see [28, Rmk. 4.12]. Curiously, this correspondence also answers a question raised recently by Andrews et al [2, 4], which asks for a size- and length-preserving bijection on \mathcal{P} that maps 2-measure to size of Durfee square.

It is worth noting that the aforementioned correspondence can also be defined in terms of another well-known combinatorial mapping on words. Let M be a finite multiset with elements from an ordered alphabet S , and let $\mathfrak{S}(M)$ be the set of all permutations of M — equivalently, words on S with content M . Foata's *second fundamental transformation* is a recursively-defined bijection $\phi : \mathfrak{S}(M) \rightarrow \mathfrak{S}(M)$ that maps major index to inversion number; i.e., $\text{maj}(w) = \text{inv}(\phi(w))$. When the alphabet S is of size 2, e.g. $S = \{\alpha, \beta\}$ with $\beta > \alpha$, Foata's transformation has a simple non-recursive description; see [44, Prop. 2.2]. Let $Q = (q_1, \dots, q_k) \in \mathcal{Q}$ and suppose $P \in \mathfrak{D}^{-1}(Q)$ has Burge code w given by (1). Then one finds

$$\phi(w) = \beta \alpha^{\delta_k - i_k} \beta \alpha^{\delta_{k-1} - i_{k-1}} \dots \beta \alpha^{\delta_2 - i_2} \beta \alpha^{\delta_1 - i_1} \beta^{i_1 - 1} \alpha \beta^{i_2 - 1} \alpha \dots \beta^{i_k - 1} \alpha.$$

This word describes a lattice path by reading from right-to-left and interpreting α and β as east and north steps, respectively. This path, in turn, determines a partition P' by tracing the boundary of its Young diagram. By construction we have $|P'| = \text{inv}(\phi(w)) = \text{maj}(w) = |P|$ and $\ell(P') = \ell(P) = \sum_{j=1}^k i_j$, and it is easy to check that P' has diagonal hooks of lengths q_1, \dots, q_k . Indeed $P \mapsto P'$ is precisely the bijection described in the previous paragraph.

Burge's bijections [15, 16] are in fact more refined than the versions presented here, as they track certain additional partition statistics that were not relevant for our purposes. However, some of these statistics may have interesting algebraic manifestations and be of interest in the future when working on the question of nilpotent orbits corresponding to commuting matrices. In the more general setting it is fruitful to view the associated binary words as lattice paths, as this affords geometric interpretations of the additional parameters [3, 13, 30].

3. THE BURGE CORRESPONDENCE AND INVARIANT SUBSPACES

In this section we first define the maximal commutative subalgebra \mathcal{U}_B of the *nilpotent commutator* $\mathcal{N}_B = \{A \in M_n(F); AB = BA, A \text{ nilpotent}\}$. The image W of a generic element of \mathcal{U}_B is an invariant subspace for B . It is of dimension $n - \mu_2(P)$ [8, Prop. 2.4]. Our main result of the section is the fact that if B is of Jordan type $P \in \mathcal{P}$, then the Jordan type of the restriction $B|_W$ is ∂P . Theorem 1 is an immediately corollary due to the universality of the descent map (Lemma 15).

3.1. The Nilpotent Commutator and its Maximal Nilpotent Subalgebra. Assume that $P = P(B)$ is the Jordan type of B and that $B \in M_n(F)$ is in the (upper-triangular) Jordan canonical form. Let $f(P)$ be the frequency sequence of P . So $P = [1^{f_1}, 2^{f_2}, \dots, z^{f_z}]$, where z is the size of the largest Jordan block in B . Thus $f_z \neq 0$. Recall that $\mathcal{S}(f) = \{i \in \mathbb{N}; f_i \neq 0\}$ is the support of f . We then write also $P = [i^{f_i}]_{i \in \mathcal{S}(f)}$. The Jordan canonical

form induces a cyclic decomposition of the underlying vector space

$$V = \bigoplus_{i \in \mathcal{S}(f)} \bigoplus_{k=1}^{f_i} V_{ik}$$

into cyclic subspaces V_{ik} for B . We denote by v_{ik1} the cyclic generator of V_{ik} . Then $v_{ikh} = B^{h-1}v_{ik1}$, $h = 1, 2, \dots, i$, is a Jordan chain for B spanning V_{ik} and the union of all these Jordan chains

$$\mathcal{B}_B = \bigcup_{i \in \mathcal{S}(f)} \bigcup_{k=1}^{f_i} \{v_{iki}, v_{ik,i-1}, \dots, v_{ik1}\}$$

is a Jordan basis for B . Note that $Bv_{iki} = 0$ for all i and k .

The following lemma gives a possible modification of Jordan basis \mathcal{B}_B .

Lemma 17. *Choose $s \in \mathcal{S}(f)$ and $t \in \{1, 2, \dots, f_s\}$. Let $1 \leq r \leq s-1$. Then for any scalars α_{ikj} , $i \in \mathcal{S}(f)$, $i \leq s-r$, $k \in \{1, 2, \dots, f_i\}$, $j \in \{1, 2, \dots, i\}$, vectors*

$$v'_{sth} = B^{h-1} \left(v_{str} + \sum_{i \in \mathcal{S}(f), i \leq s-r} \sum_{k=1}^{f_i} \sum_{j=1}^i \alpha_{ikj} v_{ikj} \right), \quad h = 1, 2, \dots, s-r+1,$$

form a Jordan chain for B . For $r = 1$, we obtain another Jordan basis for B on V by replacing $\{v_{st1}, v_{st2}, \dots, v_{sts}\}$ with $\{v'_{st1}, v'_{st2}, \dots, v'_{sts}\}$ in \mathcal{B}_B .

Proof. Use the fact that \mathcal{B}_B is a Jordan basis for B and the restrictions imposed on r and i . \square

It is well-known (see e.g. [50, §1.3]) that any matrix A commuting with B (i.e., a matrix A in the commutator \mathcal{C}_B of B) has two-step block structure: First with respect to the sizes of Jordan blocks into $A = [A_{ij}]_{i,j \in \mathcal{S}(f)}$, and then with respect to the frequency of each size to

$A_{ij} = [A_{ij}^{kl}]_{k,l=1}^{f_i, f_j}$. Each block A_{ij}^{kl} is an upper-triangular Toeplitz matrix either of the form

$$(2) \quad \begin{bmatrix} 0 & \cdots & 0 & a_1^{kl} & a_2^{kl} & \cdots & a_i^{kl} \\ \vdots & & & 0 & a_1^{kl} & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & a_2^{kl} \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & a_1^{kl} \end{bmatrix}$$

if $i < j$, or of the form

$$(3) \quad \begin{bmatrix} a_1^{kl} & a_2^{kl} & \cdots & a_j^{kl} \\ 0 & a_1^{kl} & \cdots & \vdots \\ \vdots & \ddots & \ddots & a_2^{kl} \\ \vdots & & 0 & a_1^{kl} \\ \vdots & & & 0 \\ 0 & \cdots & \cdots & 0 \end{bmatrix}$$

if $i > j$. When $i = j$ the initial columns of zeros in the first form are omitted (or equivalently the trailing rows of zeros in the second form are omitted). The parameters a_h^{kl} in different blocks are all distinct. We omitted the indices i and j to simplify notation.

For each $i \in \mathcal{S}(f)$, observe that all the blocks in A_{ii} are $i \times i$ square blocks. We choose the diagonal element of each of the blocks in A_{ii} and we form new matrix $A_i^D = [a_1^{kl}]_{k,l=1}^{f_i}$. Next, we point out that for each A the blocks A_i^D , $i \in \mathcal{S}(f)$, form the semisimple part of A considered as an element of algebra \mathcal{C}_B [11, Lem. 2.2]. Basili [8, Lem. 2.3] showed that an element $A \in \mathcal{C}_B$ is nilpotent (i.e., it belongs to the nilpotent commutator \mathcal{N}_B) if and only if its blocks A_i^D are nilpotent for all $i \in \mathcal{S}(f)$ and that for each $A \in \mathcal{N}_B$ it is possible to find a Jordan basis \mathcal{B}_B for B such that the corresponding matrices A_i^D are all lower-triangular. Thus each matrix in the nilpotent commutator \mathcal{N}_B is similar to a matrix in a maximal nilpotent subalgebra \mathcal{U}_B of \mathcal{N}_B defined by

$$\mathcal{U}_B = \{A \in \mathcal{N}_B; \text{ blocks } A_i^D \text{ are strictly lower-triangular for all } i \in \mathcal{S}(f)\}.$$

Let us now consider an example.

Example 18. Assume that B has Jordan form determined by partition $P = [4^2, 3, 2^2]$, so that $f(P) = (0, 2, 1, 2)$. A matrix A in \mathcal{U}_B is of the form

$$\begin{bmatrix} 0 & a_{11} & a_{12} & a_{13} & 0 & a_{21} & a_{22} & a_{23} & b_1 & b_2 & b_3 & c_{11} & c_{12} & c_{21} & c_{22} \\ 0 & 0 & a_{11} & a_{12} & 0 & 0 & a_{21} & a_{22} & 0 & b_1 & b_3 & 0 & c_{11} & 0 & c_{21} \\ 0 & 0 & 0 & a_{11} & 0 & 0 & 0 & a_{21} & 0 & 0 & b_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{30} & a_{31} & a_{12} & a_{33} & 0 & a_{41} & a_{42} & a_{43} & d_1 & d_2 & d_3 & e_{11} & e_{12} & e_{21} & e_{22} \\ 0 & a_{30} & a_{31} & a_{32} & 0 & 0 & a_{41} & a_{42} & 0 & d_1 & d_2 & 0 & e_{11} & 0 & e_{21} \\ 0 & 0 & a_{30} & a_{31} & 0 & 0 & 0 & a_{41} & 0 & 0 & d_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{30} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & f_1 & f_2 & f_3 & 0 & g_1 & g_2 & g_3 & 0 & h_1 & h_2 & i_{11} & i_{12} & i_{21} & i_{22} \\ 0 & 0 & f_1 & f_2 & 0 & 0 & g_1 & g_2 & 0 & 0 & h_1 & 0 & i_{11} & 0 & i_{21} \\ 0 & 0 & 0 & f_1 & 0 & 0 & 0 & g_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & j_{11} & j_{12} & 0 & 0 & k_{11} & k_{12} & 0 & l_{11} & l_{12} & 0 & m_{11} & 0 & m_{21} \\ 0 & 0 & 0 & j_{11} & 0 & 0 & 0 & k_{11} & 0 & 0 & l_{11} & 0 & 0 & 0 & 0 \\ 0 & 0 & j_{21} & j_{22} & 0 & 0 & k_{21} & k_{22} & 0 & l_{21} & l_{22} & m_{30} & m_{31} & 0 & m_{41} \\ 0 & 0 & 0 & j_{21} & 0 & 0 & 0 & k_{21} & 0 & 0 & l_{21} & 0 & m_{30} & 0 & 0 \end{bmatrix}.$$

Then we have

$$A_4^D = \begin{bmatrix} 0 & 0 \\ a_{30} & 0 \end{bmatrix}, \quad A_3^D = [0] \quad \text{and} \quad A_2^D = \begin{bmatrix} 0 & 0 \\ m_{30} & 0 \end{bmatrix}.$$

3.2. Invariant Subspaces. The structure of invariant subspaces of a nilpotent matrix in Jordan form was described by Shayman [45, 46]. The following proposition is an easy consequence of Shayman's results.

Proposition 19. *Each invariant subspace of B is equal to the image of an element of \mathcal{C}_B .*

Proof. Suppose that W is an invariant subspace of B of dimension d . By [45, Prop. 6] it is equal to the span of columns of an $n \times d$ matrix \tilde{A} of rank d in Toeplitz block form. (Note a misprint in the formulation of [45, Prop. 6] – instead of M it should be S .) Let us describe Shayman's result in further details. A partition $P' = (p'_1, p'_2, \dots, p'_l)$ of d is *compatible with* partition $P = (p_1, p_2, \dots, p_l)$ of n if $0 \leq p'_i \leq p_i$ for all i . Here P is the Jordan type of B (written with possible repetitions to conform more closely with Shayman's notation). Then $\tilde{A} = [\tilde{A}_{ij}]_{i,j=1}^{p_i, p'_j}$ is a block matrix and each block A_{ij} is in the upper-triangular Toeplitz form (2) if $p_i \leq p'_j$ or (3) if $p_i \geq p'_j$. Next, we complete the matrix \tilde{A} to an $n \times n$ matrix $A = [A_{ij}]_{i,j=1}^{p_i, p_j}$

by adding columns of zeros. If $p'_j < p_j$ then we append $p_j - p'_j$ columns of zeros on the left-side of \tilde{A}_{ij} to obtain A_{ij} for each i . If $p'_j = p_j$ then $A_{ij} = \tilde{A}_{ij}$ for each i . Observe now that $W = \text{Im } A = \text{Im } \tilde{A}$ and that $A \in \mathcal{C}_B$ by construction of A . \square

Remark 20. Note that the above proof is valid over any field F . Matrix \tilde{A} is a full rank solution of the linear matrix equation $BX = XJ$, where B is in the Jordan canonical form corresponding to P and J in the Jordan canonical form corresponding to P' .

The statement of Proposition 19 holds more generally even for matrices that are not nilpotent. An invariant subspace of a matrix has a direct sum decomposition with respect to the primary decomposition (sometimes called also rational canonical form) of the underlying vector space. See e.g. [20, Thm. 2.1.5] for a proof over complex numbers and [21, Thm. 4.3] for a proof over general F . Note that when the field is algebraically closed the primary decomposition coincides with the spectral decomposition [21, Cor. 4.4].

3.3. Elementary Cyclic Column Operations. Now, consider the image $W = \text{Im } A$ of a generic element A of \mathcal{U}_B . Since A and B commute it follows easily that W is an invariant subspace of B . To determine the Jordan type of the restriction $B|_W$ we use [45, Lem. 4]. Let us briefly outline the construction.

We denote by \mathcal{G}_B the subgroup of invertible matrices in the commutator \mathcal{C}_B that have all A_i^D blocks lower-triangular, i.e.,

$$\mathcal{G}_B = \{A \in \mathcal{C}_B \cap \text{GL}_n(F); \text{ blocks } A_i^D \text{ are lower-triangular for all } i \in \mathcal{S}(f)\}.$$

For each $A \in \mathcal{U}_B$ and $Z \in \mathcal{G}_B$ it follows that $\text{Im } A = \text{Im } AZ$. Next we introduce certain *elementary cyclic column operations* on matrices from \mathcal{U}_B partitioned into A_{ij}^{kl} blocks as per the cyclic structure of B . These are operations on columns of a matrix in \mathcal{U}_B that generalize the standard elementary column operations on a matrix.

- (i) Fix $s \in \mathcal{S}(f)$ and $t \in \{1, \dots, f_s\}$. Multiply each of the elements in blocks A_{is}^{kt} , $i \in \mathcal{S}(f)$, $k = 1, \dots, f_i$, by a nonzero scalar α .
- (ii) Fix $s, s' \in \mathcal{S}(f)$, $t \in \{1, \dots, f_s\}$ and $t' \in \{1, \dots, f_{s'}\}$. Take the first a columns of a block A_{is}^{kt} , multiply each of them by a nonzero scalar α and add them consecutively to the last a columns of block $A_{is'}^{kt'}$ for all $i \in \mathcal{S}(f)$ and $k = 1, \dots, f_i$. Here we assume that $a \leq \min\{s, s'\}$, and if $(s, t) = (s', t')$ then $a < s$.

These elementary cyclic column operations can be realized by matrix multiplication of elements in \mathcal{U}_B by certain *elementary matrices* that belong to \mathcal{G}_B .

- (i) Blocks Z_{ij}^{kl} of an elementary matrix $Z(s, t, \alpha)$ are given by:
 - $Z_{ij}^{kl} = 0$ if $(i, k) \neq (j, l)$,
 - $Z_{ii}^{kk} = I$ if $(i, k) \neq (s, t)$,
 - $Z_{ss}^{tt} = \alpha I$.
- (ii) If $(s, t) \neq (s', t')$ then blocks Z_{ij}^{kl} of an elementary matrix $Z(s, t, s', t', \alpha)$ are given by:
 - $Z_{ij}^{kl} = 0$ if $(i, k) \neq (s, t)$ or $(j, l) \neq (s', t')$,
 - $Z_{ii}^{kk} = I$,
 - in $Z_{ss'}^{tt'}$ the entries on the a -th diagonal (counting from the top-right corner of the block) are equal to α and other entries are all equal to 0.
- If $(s, t) = (s', t')$ then blocks Z_{ij}^{kl} of an elementary matrix $Z(s, t, s, t, \alpha)$ are given by:
 - $Z_{ij}^{kl} = 0$ if $(i, k) \neq (s, t)$ or $(j, i) \neq (s, t)$,
 - $Z_{ii}^{kk} = I$ if $(i, k) \neq (s, t)$,

- in Z_{ss}^{tt} the entries on the a -th diagonal (counting from the top-right corner of the block) are equal to α , the diagonal entries are equal to 1 and all other entries are equal to 0.

Choose a generic matrix $A \in \mathcal{U}_B$ partitioned into blocks A_{ij}^{kl} . All blocks with fixed pair of indices (i, k) form a *row of blocks* of A and all blocks with fixed pair of indices (j, l) form a *column of blocks* of A . A block $A_{ss'}^{tt'}$ is called a *row-pivot* of A if its rank is maximal among all ranks of blocks in its row of blocks, i.e.,

$$\text{rank}\left(A_{ss'}^{tt'}\right) = \max_{j \in S(f), l=1, \dots, f_j} \left\{ \text{rank}\left(A_{sj}^{tl}\right) \right\}.$$

A block $A_{ss'}^{tt'}$ is called a *column-pivot* of A if its rank is maximal among all ranks of blocks in its column of blocks, i.e.,

$$\text{rank}\left(A_{ss'}^{tt'}\right) = \max_{i \in S(f), k=1, \dots, f_i} \left\{ \text{rank}\left(A_{is'}^{kt'}\right) \right\}.$$

Since we perform elementary cyclic operation on columns the row-pivots are more important for us. So we will use the term *pivot* instead of row-pivot in the rest of the paper.

We fix pairs (s, t) and (s', t') such that $A_{ss'}^{tt'}$ is a pivot and write $r = \text{rank}\left(A_{ss'}^{tt'}\right)$. Then the same arguments as those in the proof of [45, Lem. 4] show that A can be transformed by elementary cyclic column operations into a matrix \hat{A} with blocks \hat{A}_{sj}^{tl} all 0 for $(j, l) \neq (s', t')$ and $\hat{A}_{ss'}^{tt'}$ has entries on the r -th diagonal (counting from the top-right corner) equal to 1 and all the other entries equal to 0.

Let us consider an example illustrating the transformation $A \rightarrow \hat{A}$.

Example 21. Continue with B corresponding to $P = (4^2, 3, 2^2)$. Observe that for a generic matrix A in \mathcal{U}_B block A_{44}^{21} is a pivot of rank equal to 4 (i.e., generically $m_{30} \neq 0$). Using elementary cyclic column operations on the second row of blocks in A we obtain that \hat{A} is of the form

$$\begin{bmatrix} 0 & a_{11} & a_{12} & a_{13} & 0 & a_{21} & a_{22} & a_{23} & b_1 & b_2 & b_3 & c_{11} & c_{12} & c_{21} & c_{22} \\ 0 & 0 & a_{11} & a_{12} & 0 & 0 & a_{21} & a_{22} & 0 & b_1 & b_3 & 0 & c_{11} & 0 & c_{21} \\ 0 & 0 & 0 & a_{11} & 0 & 0 & 0 & a_{21} & 0 & 0 & b_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & f_1 & f_2 & f_3 & 0 & g_1 & g_2 & g_3 & 0 & h_1 & h_2 & i_{11} & i_{12} & i_{21} & i_{22} \\ 0 & 0 & f_1 & f_2 & 0 & 0 & g_1 & g_2 & 0 & 0 & h_1 & 0 & i_{11} & 0 & i_{21} \\ 0 & 0 & 0 & f_1 & 0 & 0 & 0 & g_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & j_{11} & j_{12} & 0 & 0 & k_{11} & k_{12} & 0 & l_{11} & l_{12} & 0 & m_{11} & 0 & m_{21} \\ 0 & 0 & 0 & j_{11} & 0 & 0 & 0 & k_{11} & 0 & 0 & l_{11} & 0 & 0 & 0 & 0 \\ 0 & 0 & j_{21} & j_{22} & 0 & 0 & k_{21} & k_{22} & 0 & l_{21} & l_{22} & m_{30} & m_{31} & 0 & m_{41} \\ 0 & 0 & 0 & j_{21} & 0 & 0 & 0 & k_{21} & 0 & 0 & l_{21} & 0 & m_{30} & 0 & 0 \end{bmatrix}.$$

Here and later we use the letters only to indicate which entries are nonzero and which entries are equal. The exact values of an entry in A and the corresponding entry in \hat{A} need not be equal.

Lemma 22. *Suppose that A is a generic matrix in \mathcal{U}_B , P the Jordan type of B , $f = f(P)$ its frequency sequence and $z = \max\{i; i \in \mathcal{S}(f)\}$. Then the pivots of A are the following blocks:*

- (a) $A_{ii}^{k,k-1}$, $k = 2, 3, \dots, f_i$, for any $i \in \mathcal{S}(f)$ with $f_i \geq 2$,
- (b) $A_{z,z-1}^{1l}$, $l = 1, 2, \dots, f_{z-1}$, if $f_{z-1} \neq 0$,
- (c) $A_{zz}^{1f_z}$,
- (d) A_{ij}^{kl} , $j \in \mathcal{S}(f)$, $j > i$, and $l = 1, 2, \dots, f_j$, for any $i \in \mathcal{S}(f)$ and $k = 1, 2, \dots, f_i$.

Proof. Item (a) follows from the fact that A_i^D is strictly lower-triangular and so $\text{rank}(A_{ii}^{k,k-1}) = i$, maximal possible for the row of columns. Items (b) and (c) follow from the fact that A_z^D is strictly lower-triangular and so we have $\text{rank}(A_{z,z-1}^{1l}) = \text{rank}(A_{zz}^{1f_z}) = z - 1$, which is maximal for the first row of columns. For item (d) observe that for any $j \in \mathcal{S}(f)$ with $j > i$ and any l we have $\text{rank}(A_{ij}^{kl}) = i$, which is maximal possible in the row of blocks with $i \in \mathcal{S}(f)$ and $k = 1, 2, \dots, f_i$. \square

3.4. Jordan Type of the Restriction to a Generic Image. The following theorem links the algebraic and combinatorial setups. Our main result Theorem 1 is then an immediate consequence.

Theorem 23. *Suppose that B is a nilpotent matrix and that $P = P(B)$ is its Jordan type. Suppose that A is a generic nilpotent matrix commuting with B and that $W = \text{Im } A$ is its image. Then the Jordan type of the restriction $B|_W$ is given by ∂P .*

Proof. We may assume that B is in Jordan canonical form and that $A \in \mathcal{U}_B$. Suppose that $f = f(P)$ is the frequency sequence of P and that z is the largest element of $\mathcal{R}(f)$. In other words, the largest blocks of B are of size z . Suppose first that $f_z \geq 2$. By Lemma 22 blocks $A_{zz}^{k,k-1}$, $k = 2, 3, \dots, f_z$, are pivots of A . Their rank is equal to z . After performing elementary cyclic column operations consecutively on the second, third, \dots , f_z -th row of blocks of A , we obtain a matrix \hat{A} with only nonzero blocks \hat{A}_{zz}^{kl} for $k = 2, 3, \dots, f_z$, those for $l = k - 1$ and then $\hat{A}_{zz}^{k,k-1} = I$. Assume now that $f_{z-1} = 0$. Then by Lemma 22(c) block $A_{zz}^{1f_z}$ is a pivot. After elementary cyclic column operations on the first row of blocks we get the only nonzero block $\hat{A}_{zz}^{1f_z}$ in this row of blocks is equal to nilpotent Jordan block of size z . We obtain $f_z - 1$ columns of blocks of rank z and 1 column of blocks (i.e., the f_z -th column of blocks) of rank $z - 1$. We are done with the case $f_{z-1} = 0$.

Assume next that $f_{z-1} \neq 0$ and f_z arbitrary. By Lemma 22 it follows then that $A_{z,z-1}^{1f_{z-1}}$ and $A_{z-1,z}^{f_z 1}$ are pivots, and furthermore, if $f_{z-1} \geq 2$ then also $A_{z-1,z-1}^{k,k-1}$, $k = 2, 3, \dots, f_{z-1}$, are pivots of A . After performing elementary cyclic column operations with these pivots, we obtain that all the blocks \hat{A}_{ij}^{kl} with $i = z, z - 1$ and $j \leq z - 2$ are zero. Furthermore, each block of columns corresponding to these blocks has exactly one pivot. The rank of pivot blocks is equal to z for the first $f_z - 1$ columns of blocks and $z - 1$ for the other $f_{z-1} + 1$ columns of blocks. Observe that these pivot blocks are also column-pivots for their corresponding column of blocks. These follows from the assumption that blocks A_i^D are strictly lower-triangular in A .

It remains to consider the case $f_z = 1$ and $f_{z-1} = 0$. Then there is only one pivot in the columns corresponding to Jordan chains of sizes z and $z - 1$ of B , the block A_{zz}^{11} . We perform elementary cyclic column operations on the first row of blocks, so that all its blocks except the first one are zero, and the first one is equal to the nilpotent Jordan block of size z . So,

we do not have in this case any columns of blocks of rank z and only 1 column of blocks of rank $z - 1$. Hence, we are done with the columns corresponding to Jordan chains of size z and $z - 1$ of B .

Next, we apply Lemma 17 consecutively on each column of blocks and replace Jordan chains corresponding to blocks of sizes z and $z - 1$ with the chains given by \hat{A} . These gives $f_z - 1$ Jordan chains of length z and $f_{z-1} + 1$ Jordan chains of length $z - 1$ for $B|_W$. These are exactly the frequencies of z and $z - 1$ in $\partial(P)$.

Let us pause the proof and consider again our example.

Example 24. Suppose that B is corresponding to $P = [4^2, 3, 2^2]$. Then 4 is the maximal element in $R(f)$ and we have $f_4 = 2$ and $f_3 = 1$. Observe that A_{44}^{21} , A_{43}^{11} and A_{34}^{21} are pivots of A . After performing elementary cyclic column operations with these pivots we get that \hat{A} is of the form

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & j_{11} & j_{12} & 0 & 0 & k_{11} & k_{12} & 0 & l_{11} & l_{12} & 0 & m_{11} & 0 & m_{21} \\ 0 & 0 & 0 & j_{11} & 0 & 0 & 0 & k_{11} & 0 & 0 & l_{11} & 0 & 0 & 0 & 0 \\ 0 & 0 & j_{21} & j_{22} & 0 & 0 & k_{21} & k_{22} & 0 & l_{21} & l_{22} & m_{30} & m_{31} & 0 & m_{41} \\ 0 & 0 & 0 & j_{21} & 0 & 0 & 0 & k_{21} & 0 & 0 & l_{21} & 0 & m_{30} & 0 & 0 \end{bmatrix}.$$

A Jordan chain of B corresponding to the block of size 4 in \hat{A} is

$$B^{h-1} \left(v_{421} + \sum_{i=1}^2 \sum_{j=1}^2 j_{ij} v_{2ij} \right), h = 1, 2, 3, 4,$$

and two Jordan chains of B corresponding to blocks of size 3 in \hat{A} are

$$B^{h-1} \left(v_{412} + \sum_{i=1}^2 \sum_{j=1}^2 l_{ij} v_{2ij} \right), h = 1, 2, 3,$$

and

$$B^{h-1} \left(v_{311} + \sum_{i=1}^2 \sum_{j=1}^2 k_{ij} v_{2ij} \right), h = 1, 2, 3.$$

The frequencies of the restriction $B|_W$ are $f_4 = 1$ and $f_3 = 2$.

To conclude the proof of Theorem 23, observe that the blocks \hat{A}_{ij}^{kl} are 0 for $i = z, z - 1$ and $j \in \mathcal{S}(f)$ with $j \leq z - 2$. We restrict both B and \hat{A} to the span V' of all Jordan chains of B of length at most $z - 2$ and we repeat the above arguments by replacing the role of B

by $B|_{V'}$ and the role of $f(P)$ by $f' = (f_1, f_2, \dots, f_{z-2}, 0, 0)$. The process concludes when all the elements of $R(f)$ are used. \square

Example 25. In Example 24 we determined the frequencies of 4 and 3 of the restriction $B|_W$. It remains to consider the remaining columns of blocks. As done in the proof we restrict to the span V' of chains of length 2 (which are the only chains of lengths less than 3). The corresponding 4×4 matrix has now two pivots - blocks A_{22}^{21} and A_{22}^{12} . After elementary cyclic column operations we obtain

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

So, the restriction $B|_W$ has one chain of length 2 and one chain of length 1. All together, the frequency vector of $B|_W$ is equal to $(1, 1, 2, 1) = \partial(f(P))$.

Corollary 26 (Theorem 1). $\mathfrak{D}(P) = \text{Des}(P)$ for all $P \in \mathcal{P}$.

Proof. Theorem 23 gives $\mathfrak{D}(\partial P) = \mathfrak{D}(P) - 1$ since the Jordan type of the restriction of a nilpotent matrix of Jordan type T to its own image is equal to $T - 1$. Obviously $|\mathfrak{D}(P)| = |P|$, so Lemma 15 identifies \mathfrak{D} as the descent map. \square

3.5. Dominant Jordan type and Finite Fields. While our proof of Theorem 1 does not require any assumptions about the nature of the underlying field F , the definition of $\mathfrak{D}(P)$ depends on the assumption that F is infinite. Some care is needed in extending this definition to finite F . (Britnell and Wildon [14, Prop. 4.12] showed that, in general, the pairs of conjugacy classes that occur for pairs of commuting matrices are field dependent.)

Theorem 23 relies on Burge's correspondence that is completely combinatorial and the results based on Shayman's ideas that are described in §3.1–3.3. A careful reading of the latter reveals that only solutions of linear matrix equations are used and elementary column operations are performed. Moreover, in the proof of Theorem 23, we use only the fact that all the free entries of an element in \mathcal{U}_B are nonzero, and for this no assumption on the field is required. So, over an arbitrary field (finite or infinite), we see that for any matrix B of Jordan type P there is an $A \in \mathcal{N}_B$ such that the Jordan type of $B|_{\text{Im } A}$ is equal to ∂P . For instance, matrix A with only nonzero entries equal to 1 on the first nonzero diagonal in pivots (as described in Lemma 22) satisfies the condition. Together with Lemma 15, this implies that $\text{Des}(P)$ occurs as the Jordan type of an element in \mathcal{N}_B over any field F .

Over infinite fields it is known that $\mathfrak{D}(P)$ dominates any partition R corresponding to an orbit having nonempty intersection with \mathcal{N}_B , i.e., $R \preceq \mathfrak{D}(P)$ in the dominance order of partitions [34, Prop. 3.2]. Suppose now that the field F is finite, and take E to be any infinite field extending F , e.g., the algebraic closure of F . Then, over E , Theorem 1 identifies $\mathfrak{D}(P) = \text{Des}(P)$ as the unique dominant partition among Jordan types in \mathcal{N}_B . But over F , we know that $\text{Des}(P)$ occurs as the Jordan type of an element in \mathcal{N}_B . Hence $\text{Des}(P)$ is also the unique dominant partition among Jordan types in \mathcal{N}_B over F .

In summary: Regardless of the ground field, the set of Jordan types in \mathcal{N}_B always has a maximum element — namely $\text{Des}(P)$, where P is the Jordan type of B . The following therefore serves as an alternative definition of $\mathfrak{D}(P)$ over any field (finite or infinite), under which Theorem 1 is valid in all cases.

Definition 27. Suppose that $P \in \mathcal{P}$ is the Jordan type of a nilpotent matrix B . Then $\mathfrak{D}(P)$ is the maximum partition in the dominance order that occurs as a Jordan type of a matrix in \mathcal{N}_B .

3.6. Additional Remarks. Partition $\mathfrak{D}(P)$ dominates all partitions that occur as a Jordan type of elements in \mathcal{N}_B . In particular, a partition $Q \in \mathcal{Q}$ dominates all the partitions in $\mathfrak{D}^{-1}(Q)$. Observe that the partition with the maximal number of parts in $\mathfrak{D}^{-1}(Q)$ need not be dominated by all the partitions in $\mathfrak{D}^{-1}(Q)$. For instance, the partition with 8 parts in Example 3 (see Figure 1) is not related in dominance order to any of partitions that are of the form $[9, 4^2, \dots]$ in the figure. This shows that the dominance order on partitions does not relate as naturally to orbits in \mathcal{N}_B as it does to orbits of all nilpotent matrices. Compare with the Gerstenhaber-Hesselink Theorem over complex numbers [17, Thm. 6.2.5].

Observe also that if $Q = (q_1, q_2, \dots, q_r) \in \mathcal{Q}$ then the partition with the maximal number of parts in $\mathfrak{D}^{-1}(Q)$ is equal to $[q_2 + 2, q_3 + 2, \dots, q_r + 2, 1^{q_1 - 2r + 2}]$. Thus, the answer to the question of Oblak in [41, Rmk. 3, p. 612] is affirmative.

Finally, we remark that there is a natural symmetry mapping on partitions in $\mathfrak{D}^{-1}(Q)$ obtained by swapping i_j with $\delta_j - i_j + 1$ for some or all j . Furthermore, if Q and R are two distinct partitions in \mathcal{Q} with equal set (not a sequence) of δ_i ; say, σ is the permutation on δ_i of Q so that $\delta_{\sigma(i)}$ form the sequence of δ_j for R . Then the bijection from $\mathfrak{D}^{-1}(Q)$ to $\mathfrak{D}^{-1}(R)$ sending partition in position $(i_j)_j$ to the partition in position $(i_{\sigma(j)})_j$ respects the number of parts. This is the bijection mentioned in [28, Rmk. 1.2] from $\mathfrak{D}^{-1}(u, u - r)$ to $\mathfrak{D}^{-1}(u, r - 1)$ for appropriate u and r .

4. THE OBLAK PROCESS AND BASILI'S THEOREM

In this section we describe how Oblak's recursive computation of $\mathfrak{D}(P)$ can be understood through its interaction with the Burge operator ∂ . In particular, we will see how Oblak's conjecture (i.e., correctness of the algorithm) follows from the universality of the descent map (Lemma 15). Although the conjecture was recently settled by Basili [9], we believe our approach offers significant additional insight into the nature of the algorithm.

4.1. Evaluation, Annihilation, and the Oblak Process. Oblak's algorithm was formulated in the context of Gansner-Saks theory, which connects the chain structure of a finite poset to the Jordan type of a generic nilpotent element in its incidence algebra. By building on this combinatorial framework, Oblak conjectured that $\mathfrak{D}(P)$ could be computed through a greedy recursive decomposition of a certain digraph associated with P . We refer the reader to [34] for a survey of these developments.

We shall instead recast Oblak's algorithm entirely in terms of partitions, or rather their frequency representations in \mathcal{F} . The translation is straightforward and leads naturally to the following definitions.

Definition 28. For $i \geq 1$ define $\text{val}_i : \mathcal{F} \rightarrow \mathbb{N}$ and $\text{ann}_i : \mathcal{F} \rightarrow \mathcal{F}$ by

$$\text{val}_i(f) := if_i + (i + 1)f_{i+1} + 2 \sum_{j>i+1} f_j$$

and

$$\text{ann}_i(f) := (f_1, f_2, \dots, f_{i-1}, f_{i+2}, f_{i+3}, \dots).$$

For technical reasons (to be seen later), we extend these definitions to $i = 0$ as follows:

$$\text{val}_0(f) := \text{val}_1(f) = f_1 + 2(f_2 + f_3 + \dots)$$

$$\text{ann}_0(f) := \text{ann}_1(f) = (f_3, f_4, \dots).$$

We refer to val_i and ann_i as *evaluation* and *annihilation* at i , respectively. The following relation is clear from the definitions.

Lemma 29. For $f \in \mathcal{F}$ and $i \in \mathbb{N}$ we have $|\text{ann}_i(f)| = |f| - \text{val}_i(f)$.

For example, $f = (3, 0, 2, 1, 3, 0, 1)$ gives $|f| = 35$, $\text{val}_3(f) = 18$, $\text{ann}_3(f) = (3, 0, 3, 0, 1)$, and $|\text{ann}_3(f)| = 17$.

Definition 30. Let $f \in \mathcal{F}$. We say index i is maximal for f if $\text{val}_i(f)$ is nonzero and is the maximum possible evaluation of f . That is, $\text{val}_i(f) = \max\{\text{val}_j(f) : j \in \mathbb{N}\} \neq 0$.

Clearly f admits a maximal index if and only if $f \neq \varepsilon$. The nonzero condition ensures that if i is maximal for f then $|\text{ann}_i(f)| < |f|$.

With these definitions in hand, Oblak's iterative algorithm can be described as follows:

Definition 31. The Oblak process takes input $f \in \mathcal{F}$ and produces multiset output $[q_1, \dots, q_k] \in \mathcal{P}$ according to the following procedure:

```

0   INPUT  $f \in \mathcal{F}$ 
1    $f^0 \leftarrow f$ 
2    $k \leftarrow 1$ 
3   WHILE  $f^{k-1} \neq \varepsilon$ 
4       choose a maximal index  $i_k$  for  $f^{k-1}$ 
5        $q_k \leftarrow \text{val}_{i_k}(f^{k-1})$ 
6        $f^k \leftarrow \text{ann}_{i_k}(f^{k-1})$ 
7        $k \leftarrow k + 1$ 
8   END WHILE
9   OUTPUT  $[q_1, \dots, q_k]$ 

```

The table below shows the Oblak process applied to input $f = (3, 3, 2, 0, 3, 1, 0, 0, 2)$, corresponding to partition $P(f) = [1^3, 2^3, 3^2, 5^3, 6, 9^2]$. The output is $[q_1, q_2, q_3, q_4] = [25, 17, 10, 2]$.

k	f^{k-1}	$(\text{val}_{0/1}, \text{val}_2, \text{val}_3, \dots)$	i_k	q_k
1	$(3, 3, 2, 0, 3, 1, 0, 0, 2)$	$(25, 24, 18, 21, 25, 10, 4, 18, 18, 0, \dots)$	5	25
2	$(3, 3, 2, 0, 0, 0, 2)$	$(17, 16, 10, 4, 4, 14, 14, 0, \dots)$	1	17
3	$(2, 0, 0, 0, 2)$	$(6, 4, 4, 10, 10, 0, \dots)$	4	10
4	(2)	$(2, 0, \dots)$	1	2
5	ε			

Alternatively, this same input could be processed by making different choices of i_1 and i_3 :

k	f^{k-1}	$(\text{val}_{0/1}, \text{val}_2, \text{val}_3, \dots)$	i_k	q_k
1	$(3, 3, 2, 0, 3, 1, 0, 0, 2)$	$(25, 24, 18, 21, 25, 10, 4, 18, 18, 0, \dots)$	1	25
2	$(2, 0, 3, 1, 0, 0, 2)$	$(14, 15, 17, 8, 4, 14, 14, 0, \dots)$	3	17
3	$(2, 0, 0, 0, 2)$	$(6, 4, 4, 10, 10, 0, \dots)$	5	10
4	(2)	$(2, 0, \dots)$	1	2
5	ε			

Note that these distinct computations yield the same result. It was shown by Khatami that this is always the case.

Proposition 32 (Khatami [32, Thm. 2.5]). *The output of the Oblak process is independent of the choices of maximal indices made at each step.*

This invariance permits the following definition.

Definition 33. For $f \in \mathcal{F}$, let $\text{Obk}(f)$ denote the output of the Oblak process when applied to input f .

Let q_1, \dots, q_k and f^0, \dots, f^k be as defined within the process. Then Lemma 29 gives $|f^{j+1}| = |f^j| - q_{j+1}$ for all j and thus $\sum_{i=1}^k q_i = |f^0| - |f^k| = |f^0| = |P|$. It is also not difficult to see inductively that successive q_i always differ by at least 2 [32, Prop. 2.7]. Therefore $P \mapsto \text{Obk}(f(P))$ is a size-preserving function from \mathcal{P} to \mathcal{Q} . Oblak conjectured that this function is precisely \mathfrak{D} , and the conjecture was settled more than a decade later by Basili.

Theorem 34 (Basili [9]). $\mathfrak{D}(P) = \text{Obk}(f(P))$ for all $P \in \mathcal{P}$.

We will now embark on an independent proof of Theorem 34 that relies on the interaction between the Oblak process and Burge encoding. Proposition 32 will be deduced along the way.

4.2. Oblak meets Burge. As a motivating example, consider the following tables which show the Oblak process being applied to both $f = (3, 0, 1, 1, 0, 0, 0, 1) \in \mathcal{F}$ and its image $\partial f = (2, 0, 2, 0, 0, 0, 1)$ under the Burge operator ∂ .

k	f^{k-1}	$(\text{val}_{0/1}, \text{val}_2, \text{val}_3, \dots)$	i_k	q_k
1	$(3, 0, 1, 1, 0, 0, 0, 1)$	$(9, 7, 9, 6, 2, 2, 8, 8, 0, \dots)$	0	9
2	$(1, 1, 0, 0, 0, 1)$	$(5, 4, 2, 2, 6, 6, 0, \dots)$	5	6
3	$(1, 1)$	$(3, 2, 0, \dots)$	1	3
4	ε			

k	$(\partial f)^{k-1}$	$(\text{val}_{0/1}, \text{val}_2, \text{val}_3, \dots)$	i_k	q_k
1	$(2, 0, 2, 0, 0, 0, 1)$	$(8, 8, 8, 2, 2, 7, 7, 0, \dots)$	0	8
2	$(2, 0, 0, 0, 1)$	$(4, 2, 2, 5, 5, 0, \dots)$	5	5
3	(2)	$(2, 0, \dots)$	1	2
4	ε			

Notice that the outputs $\text{Obk}(f) = [9, 6, 3]$ and $\text{Obk}(\partial f) = [8, 5, 2]$ satisfy

$$\text{Obk}(\partial f) = \text{Obk}(f) - 1.$$

Our aim is to prove that this relation holds true for every $f \in \mathcal{F}$. For then Lemma 15 and Theorem 1 would give $\text{Obk}(f(P)) = \text{Des}(P) = \mathfrak{D}(P)$, thus proving Theorem 34.

The key lies in a more detailed analysis of the above example. In particular, we note that the maximal indices $(i_1, i_2, i_3) = (0, 5, 1)$ chosen when processing f remain valid maximal choices when processing ∂f . Moreover for each k we have

$$(4) \quad \text{val}_{i_k}((\partial f)^{k-1}) = \text{val}_{i_k}(f^{k-1}) - 1$$

and

$$(5) \quad (\partial f)^{k-1} = \partial(f^{k-1}).$$

The situation is neatly summarized by the following commutative diagram, in which f^0, \dots, f^3 appear along the top row and $(\partial f)^0, \dots, (\partial f)^3$ along the bottom.

$$\begin{array}{ccccccc}
(3, 0, 1, 1, 0, 0, 0, 1) & \xrightarrow[\text{val}_1=9]{\text{ann}_1} & (1, 1, 0, 0, 0, 1) & \xrightarrow[\text{val}_5=6]{\text{ann}_5} & (1, 1) & \xrightarrow[\text{val}_1=3]{\text{ann}_1} & \varepsilon \\
\partial \downarrow & & \partial \downarrow & & \partial \downarrow & & \partial \downarrow \\
(2, 0, 2, 0, 0, 0, 1) & \xrightarrow[\text{val}_1=8]{\text{ann}_1} & (2, 0, 0, 0, 1) & \xrightarrow[\text{val}_5=5]{\text{ann}_5} & (2) & \xrightarrow[\text{val}_1=2]{\text{ann}_1} & \varepsilon
\end{array}$$

We could instead choose maximal indices $(i_1, i_2, i_3) = (3, 5, 1)$ when computing $\text{Obk}(f)$. These, too, remain maximal in the computation of $\text{Obk}(\partial f)$, and both (4) and (5) continue

to hold. This is captured by the diagram below.

$$\begin{array}{ccccccc}
 (3, 0, 1, 1, 0, 0, 0, 1) & \xrightarrow[\text{val}_3=9]{\text{ann}_3} & (3, 0, 0, 0, 0, 0, 1) & \xrightarrow[\text{val}_5=6]{\text{ann}_5} & (3) & \xrightarrow[\text{val}_1=3]{\text{ann}_1} & \varepsilon \\
 \partial \downarrow & & \partial \downarrow & & \partial \downarrow & & \partial \downarrow \\
 (2, 0, 2, 0, 0, 0, 0, 1) & \xrightarrow[\text{val}_3=8]{\text{ann}_3} & (2, 0, 0, 0, 0, 1) & \xrightarrow[\text{val}_5=5]{\text{ann}_5} & (2) & \xrightarrow[\text{val}_1=2]{\text{ann}_1} & \varepsilon
 \end{array}$$

While these observations are suggestive of a general principal, some caution is required. The reader should confirm that $(i_1, i_2, i_3) = (1, 6, 1)$ is a maximal index sequence for f , but *not* for ∂f .

Based on the foregoing discussion we turn our attention to the following questions:

- (a) If i is maximal for $f \in \mathcal{F}$, under what conditions is i maximal for ∂f ?
- (b) For which i do we have $\text{val}_i(\partial f) = \text{val}_i(f) - 1$ and $\text{ann}_i(\partial f) = \partial(\text{ann}_i(f))$?

A thorough understanding of second question will allow us to address the first, so we begin with (b). Our answer will be framed in terms of the following notions.

Definition 35. Let $f \in \mathcal{F}$. We say $i \geq 1$ is right admissible in f if $f_i > 0$ and either $f_{i+1} > 0$ or $f_{i-1} = f_{i+1} = 0$. We say $i \geq 0$ is left admissible in f if $f_{i+1} > 0$ and either $f_i > 0$ or $f_i = f_{i+2} = 0$.

In other words, i is right admissible if $f_i \neq 0$ and i is not the right end of a nontrivial spread, and i is left admissible if $f_{i+1} \neq 0$ and $i + 1$ is not the left end of a nontrivial spread. For example, the admissible indices for two elements of \mathcal{F} are shown below.

f	right	left
$(3, 0, 2, 0, 0, 1, 2, 1, 0, 0, 1)$	$1, 3, 6, 7, 11$	$0, 2, 6, 7, 10$
$(1, 2, 1, 0, 2, 1, 0, 1, 0, 2)$	$1, 2, 5, 8, 10$	$1, 2, 5, 7, 9$

At first glance it seems that there is an asymmetry between right and left admissibility. However, this is merely a result of notational convention. Both notions really concern the pair (f_i, f_{i+1}) , which we choose to identify with index i .

In fact the only distinction in left vs. right admissibility lies in the treatment of trivial spreads. If f has no trivial spreads, then the sets of left and right admissible indices coincide, whereas if $\{i\}$ is a trivial spread then i is right admissible (but not left) and $i - 1$ is left admissible (but not right). In particular, note that 0 is never right admissible and is left admissible if and only if $f_1 > 0$ and $f_2 = 0$ (i.e., $\{1\}$ is a trivial spread).

Proposition 36. Let $f \in \mathcal{F}$, $i \geq 0$. If i is either left admissible for f or right admissible for ∂f , then $\text{val}_i(\partial f) = \text{val}_i(f) - 1$ and the following diagram commutes:

$$(*) \quad \begin{array}{ccc}
 f & \xrightarrow{\text{ann}_i} & g \\
 \partial \downarrow & & \downarrow \partial \\
 \partial f & \xrightarrow{\text{ann}_i} & \partial g
 \end{array}$$

Example 37. Let $f = (3, 0, 2, 1, 0, 1, 2, 1)$, which has $\partial f = (2, 0, 3, 0, 1, 0, 3)$. Then $\{0, 3, 6, 7\}$ is the set of left admissible indices in f , and $\{1, 3, 5, 7\}$ is the set of right admissible indices in ∂f . The proposition guarantees that $\text{val}_i(\partial f) = \text{val}_i(f) - 1$ and diagram $(*)$ commutes whenever $i \in \{0, 1, 3, 5, 6, 7\}$.

For example, at $i = 1$ we have $\text{val}_1(f) = 17$, $\text{val}_1(\partial f) = 16$, $g = \text{ann}_1(f) = (2, 1, 0, 1, 2, 1)$ and $\partial g = (3, 0, 1, 0, 3) = \text{ann}_1(\partial f)$. It turns out that $(*)$ also commutes when $i = 4$, but it

fails to commute when $i \in \{2, 8\}$. For instance, when $i = 8$, we get

$$\partial(\text{ann}_8(f)) = \partial(3, 0, 2, 1, 0, 1, 2) = (2, 0, 3, 0, 0, 2, 1),$$

as compared to $\text{ann}_8(\partial f) = (2, 0, 3, 0, 1, 0, 3)$.

The proof of Proposition 36 is elementary but technical, involving the same essential argument applied to various special cases. The core idea is contained in the proof of the following Lemma.

Lemma 38. *Let $f \in \mathcal{F}$ and suppose $i \geq 1$ and $i + 1 \in R(f)$. Then $\text{val}_i(\partial f) = \text{val}_i(f) - 1$ and $\text{ann}_i(\partial f) = \partial(\text{ann}_i(f))$.*

Proof. Let $g = (f_{i+2}, f_{i+3}, \dots)$. Since $i + 1 \in R(f)$ we have $i + 2 \notin R(f)$, or equivalently $g \in \mathcal{A}$. Therefore

$$\partial f = ((\partial f)_1, \dots, (\partial f)_{i-1}, f_i + 1, f_{i+1} - 1, \partial g)$$

and

$$\begin{aligned} \text{val}_i(\partial f) &= i(f_i + 1) + (i + 1)(f_{i+1} - 1) + 2\ell(\partial g) \\ &= -1 + if_i + (i + 1)f_{i+1} + 2\ell(g) \\ &= \text{val}_i(f) - 1. \end{aligned}$$

Finally, $\text{ann}_i(\partial f) = ((\partial f)_1, \dots, (\partial f)_{i-1}, \partial g) = \partial(\text{ann}_i(f))$. \square

Proof of Proposition 36. Let i be right admissible in ∂f . If $i + 1 \in R(f)$ then the conclusion follows from Lemma 38, so we suppose that $i + 1 \notin R(f)$.

- We first deal with the case $i > 1$. Then we must have that both i and $i + 2$ belong to $R(f)$. Hence

$$\begin{aligned} \text{ann}_i(\partial f) &= \text{ann}_i(\dots, f_{i-1} + 1, f_i - 1, f_{i+1} + 1, f_{i+2} - 1, \dots) \\ &= (\dots, f_{i-1} + 1, f_{i+2} - 1, \dots) \\ &= \partial(\text{ann}_i(f)), \\ \text{val}_i(\partial f) &= (f_i - 1)i + (f_{i+1} + 1)(i + 1) + 2 \left((f_{i+2} - 1) + \sum_{k>i+2} f_k \right) \\ &= -1 + f_i i + f_{i+1}(i + 1) + 2 \sum_{k>i+1} f_k \\ &= \text{val}_i(f) - 1. \end{aligned}$$

- Now suppose that $i = 1$. We treat cases $(\partial f)_2 > 0$ and $(\partial f)_2 = 0$ separately.
 - Assume $(\partial f)_2 > 0$. Since $2 \notin R(f)$, this means that $3 \in R(f)$ and $(f_4, \dots) \in \mathcal{A}$. Hence

$$\begin{aligned} \text{ann}_1(\partial f) &= \text{ann}_1(f_1 - 1, f_2 + 1, f_3 - 1, \partial(f_4, \dots)) \\ &= (f_3 - 1, \partial(f_4, \dots)) \\ &= \partial(f_3, f_4, \dots) = \partial(\text{ann}_1 f), \\ \text{val}_1(\partial f) &= (f_1 - 1)(1) + (f_2 + 1)(2) + 2 \left((f_3 - 1) + \sum_{k>3} f_k \right) \\ &= -1 + f_1 + 2f_2 + 2 \sum_{k>2} f_k \\ &= \text{val}_1(\partial f) - 1. \end{aligned}$$

- Assume now that $(\partial f)_2 = 0$. Since $2 \notin R(f)$ this implies that $f_2 = 0$. Therefore we must have that $3 \notin R(f)$ and $(f_3, \dots) \in \mathcal{A}$, hence

$$\begin{aligned} \text{ann}_i(\partial f) &= \text{ann}_1(f_1 - 1, 0, \partial(f_3, \dots)) \\ &= \partial(f_3, \dots) = \partial(\text{ann}_i f), \\ \text{val}_i(\partial f) &= f_1 - 1 + 2 \sum_{k>2} f_k = \text{val}_i(f) - 1. \end{aligned}$$

Now let i be left admissible in f .

- We first deal with the case $i \geq 1$. If $i + 1 \in R(f)$, then the conclusion again comes from Lemma 38, so we suppose $i + 1 \notin R(f)$. Since i is left admissible in f , this means $f_i > 0$, and both $i, i + 2 \in R(f)$. Hence

$$\begin{aligned} \text{ann}_i(\partial f) &= \text{ann}_i(\dots, f_{i-1} + 1, f_i - 1, f_{i+1} + 1, f_{i+2} - 1, \dots) \\ &= (\dots, f_{i-1} + 1, f_{i+2} - 1, \dots) = \partial(\text{ann}_i(f)), \text{ and} \\ \text{val}_i(\partial f) &= (f_i - 1)i + (f_{i+1} + 1)(i + 1) + 2 \left((f_{i+2} - 1) + \sum_{k>i+2} f_k \right) \\ &= -1 + f_i i + f_{i+1}(i + 1) + 2 \sum_{k>i+1} f_k = \text{val}_i(f) - 1. \end{aligned}$$

- We now analyse the case $i = 0$. In this case we must have that $f_1 > 0$ and $f_2 = 0$. We analyse the cases $3 \in R(f)$ (i.e., $(f_3, \dots) \in \mathcal{B}$) and $3 \notin R(f)$ separately.
 - Assume that $3 \in R(f)$. Then

$$\begin{aligned} \text{ann}_1(\partial f) &= (f_1 - 1, 1, f_3 - 1, \partial(f_4, \dots)) \\ &= (f_3 - 1, \partial(f_4, \dots)) = \partial(f_3, \dots) = \partial(\text{ann}_1 f), \\ \text{val}_1(\partial f) &= (f_1 - 1) + 2 + 2 \left((f_3 - 1) + \sum_{k>3} f_k \right) \\ &= \text{val}_1(f) - 1. \end{aligned}$$

- Assume that $3 \notin R(f)$. Then $(f_3, \dots) \in \mathcal{A}$ and hence

$$\begin{aligned} \text{ann}_1(\partial f) &= (f_1 - 1, 0, \partial(f_3, \dots)) \\ &= \partial(f_3, \dots) = \partial(\text{ann}_1 f), \\ \text{val}_1(\partial f) &= (f_1 - 1) + 2 \sum_{k>2} f_k = \text{val}_1(f) - 1. \end{aligned}$$

□

We now wish to extend Proposition 36 to address when maximality of the index i in f ensures maximality of i in ∂f . This requires a somewhat roundabout approach.

Definition 39. Fix $f \in \mathcal{F}$. We say indices $i, j \in \mathbb{N}$ are equivalent with respect to f if $\text{ann}_i(f) = \text{ann}_j(f)$ and $\text{val}_i(f) = \text{val}_j(f)$.

For example, if $f = (3, 0, 2, 0, 0, 1, 1, 1, 0, 1)$, then the index set \mathbb{N} is partitioned into equivalence classes $\{0, 1\}$, $\{2, 3\}$, $\{4\}$, $\{5\}$, $\{6, 7\}$, $\{8, 9, 10\}$, and $\{11, 12, \dots\}$. We remark that the condition $\text{val}_i(f) = \text{val}_j(f)$ in Definition 39 is actually redundant as Lemma 29 shows it to follow from $\text{ann}_i(f) = \text{ann}_j(f)$. It is included only to emphasize the equality of both annihilations and evaluations at equivalent indices.

Lemma 40. *Let $f \in \mathcal{F}$ with $f \neq \varepsilon$. If i is maximal for f then it is equivalent to some right admissible index j and also equivalent to some left admissible index k .*

Proof. Since $\text{val}_0 = \text{val}_1$ we assume, with no loss of generality, that $i \geq 1$.

We first consider the case $f_i = 0$. Due to maximality of $\text{val}_i(f)$ we cannot have $f_r = 0$ for all $r > i$. Let j be the smallest index such that $j > i$ and $f_j > 0$. From definition of val_i it is clear that $\text{val}_i(f) \leq \text{val}_j(f)$ with equality if and only if $j = i + 1$ and $f_{i+2} = 0$. Maximality of $\text{val}_i(f)$ forces equality, in which case i is left admissible and equivalent to the right admissible index $j = i + 1$.

We now examine the case $f_i > 0$. If we also have $f_{i+1} > 0$, then i itself is both left and right admissible. Now suppose that $f_{i+1} = 0$. If $f_{i-1} > 0$ then $\text{val}_{i-1}(f) > \text{val}_i(f)$, contradicting the maximality of $\text{val}_i(f)$. Hence we have $f_{i-1} = f_{i+1} = 0$. So i is right admissible and equivalent to the left admissible index $k = i - 1$. \square

Proposition 41. *Let $f \in \mathcal{F}$ with $f \neq (1)$. If index i is left admissible and maximal for f , then i is maximal for ∂f and $\text{val}_i(\partial f) = \text{val}_i(f) - 1$.*

Proof. Let $m_f = \text{val}_i(f)$ be the maximum evaluation for f , and let $m_{\partial f}$ be the maximum evaluation value for ∂f . Left admissibility of i implies $f \neq \varepsilon$, and since $f \neq (1)$ we have $\partial f \neq \varepsilon$. Lemma 40 ensures that $\text{val}_j(\partial f) = m_{\partial f}$ for some j that is right admissible in ∂f . Proposition 36 then gives both

$$1 + m_{\partial f} = 1 + \text{val}_j(\partial f) = \text{val}_j(f) \leq m_f$$

and

$$1 + m_{\partial f} \geq 1 + \text{val}_i(\partial f) = \text{val}_i(f) = m_f,$$

so the inequalities to be tight and $m_{\partial f} = \text{val}_i(\partial f) = \text{val}_i(f) - 1$. \square

4.3. Oblak Chains and the proof of Basili's Theorem. The ingredients are now in place to assemble the proof of Theorem 34. We begin by defining an *Oblak chain* for $f \in \mathcal{F}$ to be any sequence (f^0, \dots, f^k) that could be generated through the Oblak process applied to f . (Compare with Definition 31.)

Definition 42. *Let $f \in \mathcal{F}$. An Oblak chain for f is a sequence $C = (f^0, f^1, \dots, f^k)$ such that $f^0 = f$, $f^k = \varepsilon$ and for every $r = 1, \dots, k$ we have $f^r = \text{ann}_{i_r}(f^{r-1})$ for some index i_r that is maximal for f^{r-1} . (Note that i_r is unique up to equivalence in f^{r-1} .) Such a sequence of indices (i_1, \dots, i_k) is called an index sequence for C . The multiset*

$$\text{val}(C) := [\text{val}_{i_1}(f^0), \dots, \text{val}_{i_k}(f^{k-1})] = [|f^0| - |f^1|, \dots, |f^{k-1}| - |f^k|],$$

is called the valuation of C .

Evidently, if $C = (f^0, \dots, f^k)$ is an Oblak chain for f then $\text{val}(C)$ is a partition of size $|f|$ and no two of the f^r are equal. Note that $C = (\varepsilon)$ is the unique Oblak chain for ε , with valuation $\text{val}(C) = \varepsilon$.

We extend ∂ to act on Oblak chains as follows.

Definition 43. *Let $C = (f^0, \dots, f^k)$ be an Oblak chain. We define*

$$\partial C := \begin{cases} (\partial f^0, \dots, \partial f^k), & f^{k-1} \neq (1) \\ (\partial f^0, \dots, \partial f^{k-1}), & f^{k-1} = (1) \end{cases}.$$

Theorem 44. *Let $f \in \mathcal{F}$ and let C be an Oblak chain for f . Then ∂C is an Oblak chain for ∂f and $\text{val}(\partial C) = \text{val}(C) - 1$.*

Proof. Let $C = (f^0, \dots, f^k)$ and suppose (i_1, \dots, i_k) is an index sequence for C . By Lemma 40, each i_r is equivalent in f^{r-1} to a left admissible index j_r . Proposition 36 gives

$$\begin{aligned}\partial f^r &= \partial(\text{ann}_{j_r}(f^{r-1})) = \text{ann}_{j_r}(\partial f^{r-1}), \\ \text{val}_{j_r}(\partial f^{r-1}) &= \text{val}_{j_r}(f^{r-1}) - 1,\end{aligned}$$

and, if $f^{r-1} \neq (1)$ then Proposition 41 ensures j_r is maximal for $\partial(f^{r-1})$.

The condition $f^{r-1} = (1)$ implies $f^r = \varepsilon$ and is therefore only possible at $r = k$. Moreover, it is clear that $\text{val}_{j_r}(f^{r-1}) = 1$ if and only if $f^{r-1} = (1)$. So if indeed $f^{k-1} = (1)$, then $\partial C = (\partial f^0, \dots, \partial f^{k-1})$ is an Oblak chain for ∂f with index sequence (j_1, \dots, j_{k-1}) and valuation

$$\begin{aligned}\text{val}(\partial C) &= [\text{val}_{j_1}(f^0) - 1, \dots, \text{val}_{j_{k-1}}(f^{k-2}) - 1] \\ &= [\text{val}_{j_1}(f^0), \dots, \text{val}_{j_{k-1}}(f^{k-2}), 1] - 1 \\ &= \text{val}(C) - 1.\end{aligned}$$

Otherwise, $\partial C = (\partial f^0, \dots, \partial f^k)$ is an Oblak chain for ∂f with index sequence (j_1, \dots, j_k) and valuation $\text{val}(\partial C) = [\text{val}_{j_1}(f^0) - 1, \dots, \text{val}_{j_k}(f^{k-1}) - 1] = \text{val}(C) - 1$. \square

Example 45. Let $P = [1, 2^3, 5, 10, 14]$ and $f = f(P)$. Figure 3 illustrates the interaction between the Oblak chains (displayed horizontally) and Burge chains (displayed vertically) obtained by repeated application of Theorem 44, starting with a particular Oblak chain for f . The labels above the horizontal arrows list the equivalent maximal indices at each step, with the left admissible ones in bold. The labels below the horizontal arrows are the corresponding evaluations, which are seen to decrease by one row-by-row down each column.

The following corollary of Theorem 44 says that the output of the Oblak process is not dependent on the choices of maximal indices made along the way. That is, we have recovered Proposition 32.

Corollary 46 (Proposition 32). *Let $f \in \mathcal{F}$. Any two Oblak chains for f have the same valuation.*

Proof. We proceed by induction on the size of f . The statement is obvious if $f = \varepsilon$. Suppose $f \neq \varepsilon$ and let C_1, C_2 be Oblak chains for f . Then Theorem 44 implies ∂C_1 and ∂C_2 are Oblak chains for ∂f , with $\text{val}(\partial C_i) = \text{val}(C_i) - 1$. The inductive hypothesis gives $\text{val}(\partial C_1) = \text{val}(\partial C_2)$, hence $\text{val}(C_1) - 1 = \text{val}(C_2) - 1$. But this implies $\text{val}(C_1) = \text{val}(C_2)$, since $\text{val}(C_1)$ and $\text{val}(C_2)$ are of common size $|f|$. (In particular, if $\text{val}(C_i) - 1 = [1^{m_1}, 2^{m_2}, \dots]$ then $\text{val}(C_i) = [1^k, 2^{m_1}, 3^{m_2}, \dots]$, where $k = |f| - \sum_{i \geq 2} (i+1)m_i$.) \square

With Corollary 46 in hand we can unambiguously regard val as a function from \mathcal{F} to \mathcal{P} by setting $\text{val}(f) := \text{val}(C)$, where C is any Oblak chain for f . Of course this identifies $\text{val}(f)$ with $\text{Obk}(f)$, and we arrive at Basili's Theorem as promised.

Corollary 47 (Theorem 34). $\mathfrak{D}(P) = \text{val}(f(P))$ for any $P \in \mathcal{P}$.

Proof. Let $\text{val}(P) := \text{val}(f(P))$. We know $|\text{val}(P)| = |P|$ and according to Theorem 44 we have $\text{val}(\partial P) = \text{val}(P) - 1$. Hence Lemma 15 gives $\text{val}(P) = \text{Des}(P)$, and the result follows by Theorem 1. \square

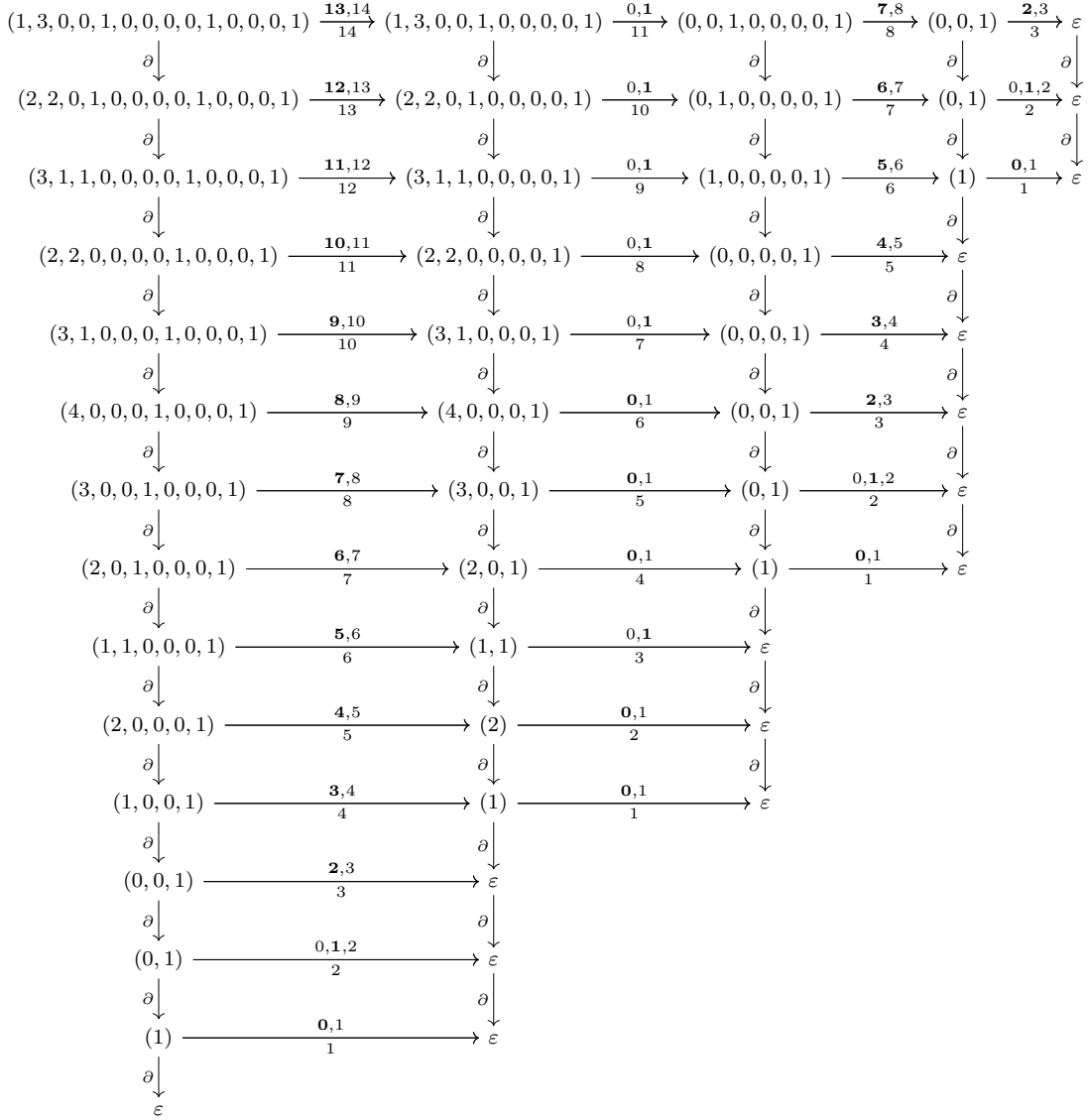


FIGURE 3. An illustration of the iterated application of Theorem 44. (See Example 45.)

ACKNOWLEDGMENTS

T. Košir was supported in part by the Research and Innovation Agency of Slovenia (research core funding P1-0222, research projects N1-0154 and J1-3004 until December 31, 2023, and research core funding P1-0448 and research project N1-0217 after January 1, 2024). M. Mastnak was supported in part by the NSERC Discovery Grant 371994-2019.

Most of research presented in the paper was done during an extended visit of T. Košir at Saint Mary's University in Halifax. He is truly grateful for the hospitality he received during his stay.

REFERENCES

- [1] N. Altafi, A. Iarrobino, P. Marcias Marques. *Jordan Type of an Artinian algebra, a survey*. arXiv:2307.00957
- [2] G. E. Andrews, S. Bhattacharjee, M. Dastidar. *Sequences in partitions*. *Integers*. **22** (2022), A#32.
- [3] G. E. Andrews, D. M. Bressoud. *On the Burge correspondence between partitions and binary words*. *Rocky Mount. J. Math.* **15** (1985), no. 2, 225-233.
- [4] G. E. Andrews, S. Chern, Z. Li. *On the k -measure of partitions and distinct partitions*. *Algebr. Combin.* **5** (2022), no. 6, 1353-1361.
- [5] F. V. Atkinson. *Multiparameter Eigenvalue Problems*. Academic Press, 1972.
- [6] V. Baranovsky, *The variety of pairs of commuting nilpotent matrices is irreducible*, *Transform. Groups* **6** (2001), no. 1, 3-8.
- [7] R. Basili. *On the irreducibility of varieties of commuting matrices*. *J. Pure Appl. Algebra* **149** (2000), 107-120.
- [8] R. Basili. *On the irreducibility of commuting varieties of nilpotent matrices*. *J. Algebra* **268** (2003), 58-80.
- [9] R. Basili. *On the maximal nilpotent orbit that intersects the centralizer of a matrix*. *Transform. Groups* **27** (2022), no. 1, 1-30.
- [10] R. Basili, A. Iarrobino. *Pairs of commuting nilpotent matrices and Hilbert function*. *J. Algebra* **320** (2008), 1235-1254.
- [11] R. Basili, A. Iarrobino, L. Khatami. *Commuting nilpotent matrices and Artinian algebras*. *J. Commut. Alg.* **2** (2010), no. 3, 295-325.
- [12] D. Bogdanić, A. Djurić, S. Koljančić, P. Oblak, K. Šivic. *Jordan structures of nilpotent matrices in the centralizer of a nilpotent matrix with two Jordan blocks of the same size*. *Lin. Alg. Appl.* **690** (2024), 59-90.
- [13] D. M. Bressoud. *Lattice paths and the Rogers-Ramanujan identities*. In *Lecture Notes in Math.* **1395**, Springer-Verlag, 1989, 140-172.
- [14] J. R. Britnell, M. Wildon. *On types and classes of commuting matrices over finite fields*. *J. London Math. Soc.* **383** (2011), no. 2, 470-492.
- [15] W. H. Burge. *A correspondence between partitions related to generalizations of the Rogers-Ramanujan identities*. *Discrete Math.* **34** (1981), 9-15.
- [16] W. H. Burge. *A three-way correspondence between partitions*. *Europ. J. Combinatorics* **3** (1982), 195-213.
- [17] D. H. Collingwood, W. M. McGovern. *Nilpotent Orbits in Semisimple Lie Algebras*. Van Nostrand Reinhold Publ., 1993.
- [18] I. M. Gelfand, V. A. Ponomarev. *Remarks on the classification of a pair of commuting linear transformations in a finite-dimensional space*. *Funkcional. Anal. Prilozh.* **3** (1969), no. 4, 81-82.
- [19] M. Gerstenhaber. *On dominance and varieties of commuting matrices*. *Ann. Math.* **73** (1961), no. 2, 324-348.
- [20] I. Gohberg, P. Lancaster, L. Rodman. *Invariant Subspaces of Matrices with Applications*. Wiley Interscience Publ., 1986.
- [21] L. Grunenfelder, T. Košir. *Coalgebras and spectral theory in one and several parameters*. In *Recent developments in operator theory and its applications* (Winnipeg, MB, 1994), 177-192. *Oper. Theory Adv. Appl.*, **87**, Birkhäuser Verlag, Basel, 1996.
- [22] R. M. Guralnick. *A note on commuting pairs of matrices*. *Lin. Multilin. Alg.* **31** (1992), 71-75.
- [23] R. M. Guralnick, B. A. Sethuraman. *Commuting pairs and triples of matrices and related varieties*. *Lin. Alg. Appl.* **310** (2000), nos. 1-3, 139-148.
- [24] W. J. Haboush, D. Hyeon. *Conjugacy classes of commuting nilpotents*. *Trans. Amer. Math. Soc.* **372** (2019), no. 6, 4293-4311.
- [25] Y. Han. *Commuting triples of matrices*. *Electron. J. Lin. Alg.* **13** (2005), 274-343.
- [26] J. Holbrook, M. Omladić. *Approximating commuting operators*. *Lin. Alg. Appl.* **327** (2001), nos. 1-3, 131-149.
- [27] A. Iarrobino, L. Khatami. *Bound on the Jordan type of a generic nilpotent matrix commuting with a given matrix*. *J. Algebr. Combin.* **38** (2013), no. 4, 947-972.
- [28] A. Iarrobino, L. Khatami, B. Van Steirtenghem, R. Zhao. *Nilpotent matrices having a given Jordan type as maximum commuting nilpotent orbit*. *Lin. Alg. Appl.* **546** (2018), 210-260.
- [29] A. Iarrobino, P. Macias Marques, C. McDaniel. *Artinian algebras and Jordan type*. *J. Commut. Alg.* **14** (2022), no. 3, 365-414.

- [30] P. Jacob, P. Mathieu. *Multiple partitions, lattice paths and a Burge-Bressoud-type correspondence*. Discrete Math. **309** (2009), 878-886.
- [31] J. Jelisiejew, K. Šivic. *Components and singularities of Quot schemes and varieties of commuting matrices*. J. Reine Angew. Math. **788** (2022), 129-187.
- [32] L. Khatami. *The poset of the nilpotent commutator of a nilpotent matrix*. Linear Algebra Appl. **439** (2013), no. 12, 3763-3776.
- [33] L. Khatami. *The smallest part of the generic partition of the nilpotent commutator of a nilpotent matrix*. J. Pure Appl. Alg. **218** (2014), 1496-1516.
- [34] L. Khatami. *Commuting Jordan types: a survey*. arXiv:2304.08550
- [35] T. Košir. *The Cayley-Hamilton theorem and inverse problems for multiparameter systems*. Lin. Alg. Appl. **367**, (2003), 155-163.
- [36] T. Košir, P. Oblak. *On pairs of commuting nilpotent matrices*. Transform. Groups **14** (2009), no. 1, 175-182.
- [37] P. D. Levy, N. V. Ngo, K. Šivic. *Commuting varieties and cohomological complexity theory*. J. Lond. Math. Soc. (2) **106** (2022), no. 4, 2847-2883.
- [38] N. V. Ngo, K. Šivic. *On varieties of commuting nilpotent matrices*. Lin. Alg. Appl. **452** (2014), 237-262.
- [39] P. Oblak. *The upper bound for the index of nilpotency for a matrix commuting with a given nilpotent matrix*. Lin. Multilin. Algebra **56** (2008), no. 6, 701-711.
- [40] P. Oblak. *Jordan forms for mutually annihilating nilpotent pairs*. Lin. Alg. Appl. **428** (2008), no.7, 1476-1491.
- [41] P. Oblak. *On the nilpotent commutator of a nilpotent matrix*. Lin. Multilin. Algebra **60** (2012), no. 5, 599-612.
- [42] D. I. Panyushev. *Two results on centraliser of nilpotent elements*. J. Pure Appl. Algebra **212** (2008), no. 4, 774-779.
- [43] A. Premet. *Nilpotent commuting varieties of reductive Lie algebras*. Invent. Math. **154** (2003), no. 3, 653-683.
- [44] B. Sagan, C. Savage. *Mahonian pairs*. J. Combin. Theory Ser. A. **119** (2012), no. 3, 526-545.
- [45] M. A. Shayman. *On the variety of invariant subspaces of a finite-dimensional linear operator*. Trans. Amer. Math. Soc. **274** (1982), no. 2, 721-747.
- [46] M. A. Shayman. *Parametrization of the flags fixed by a unipotent matrix*. Contemp. Math. **47** (1985), 355-368.
- [47] K. Šivic. *On varieties of commuting triples*. Lin. Alg. Appl. **428** (2008), 2006-2029.
- [48] K. Šivic. *On varieties of commuting triples II*. Lin. Alg. Appl. **437** (2012), 461-489.
- [49] K. Šivic. *On varieties of commuting triples III*. Lin. Alg. Appl. **437** (2012), 393-460.
- [50] D. A. Suprunenko, R. I. Tyshkevich. *Commutative Matrices*. Academic Press, 1968.
- [51] H. Volkmer. *Multiparameter Eigenvalue Problems and Expansion Theorems*. Lecture Notes in Math. **1356**, Springer-Verlag, 1988.

DEPARTMENT OF MATHEMATICS AND COMPUTING SCIENCE, SAINT MARY'S UNIVERSITY, 923 ROBIE ST,
 HALIFAX, NOVA SCOTIA, CANADA B3N 1Z9
Email address: john.irving@smu.ca

FACULTY OF MATHEMATICS AND PHYSICS, UNIVERSITY OF LJUBLJANA, JADRANSKA 19, SI-1000 LJUB-
 LJANA, SLOVENIA
Email address: tomaz.kosir@fmf.uni-lj.si

DEPARTMENT OF MATHEMATICS AND COMPUTING SCIENCE, SAINT MARY'S UNIVERSITY, 923 ROBIE ST,
 HALIFAX, NOVA SCOTIA, CANADA B3N 1Z9
Email address: mitja.mastnak@smu.ca