



General Preservers of Quasi-Commutativity

Gregor Dolinar and Bojan Kuzma

Abstract. Let M_n be the algebra of all $n \times n$ matrices over \mathbb{C} . We say that $A, B \in M_n$ quasi-commute if there exists a nonzero $\xi \in \mathbb{C}$ such that $AB = \xi BA$. In the paper we classify bijective not necessarily linear maps $\Phi: M_n \rightarrow M_n$ which preserve quasi-commutativity in both directions.

1 Introduction

Let M_n be the algebra of all $n \times n$ matrices over a complex field \mathbb{C} , with the standard basis E_{ij} . We say that $A, B \in M_n$ quasi-commute if there exists a nonzero $\xi \in \mathbb{C}$ such that $AB = \xi BA$. When A and B quasi-commute we write $A \overset{\#}{\leftrightarrow} B$. We also define the quasi-commutant of a set $\Omega \subset M_n$ by

$$\Omega^{\#} = \{X \in M_n : X \overset{\#}{\leftrightarrow} A \text{ for every } A \in \Omega\} = \bigcap_{A \in \Omega} \bigcup_{\xi \neq 0} \{X \in M_n : XA = \xi AX\}$$

and let $A^{\#} = \{A\}^{\#}$.

To avoid possible misunderstandings, we emphasize that there is a similar-looking, but in fact essentially different, relation from quasi-commutativity. It is called commutativity up to a (fixed) factor $\zeta \in \mathbb{C}$ and is defined as follows. We say that $A, B \in M_n$ commute up to a factor ζ if $AB = \zeta BA$. The difference from quasi-commutativity is that ζ is fixed here. In particular, the quasi-commutant of a given matrix is not necessarily a linear subspace in M_n (see Examples 2.1–2.3).

We also mention that the terms quasi-commutativity and commutativity up to a factor have also been used in a different sense, see for example McCoy [14], Potter [21], Holtz, Mehrmann, and Schneider [10], and Molnár [16]. Our notion of quasi-commutativity is the same as in Radjavi and Šemrl [22].

Commutativity up to a fixed factor and quasi-commutativity are interesting from different points of view. For example, the classical commutativity relation, that is, commutativity up to the factor $\xi = 1$, is important in quantum mechanics. It is linked to quantum observables, and one of the basic relations among observables is compatibility, where two observables are compatible if the operators representing them commute [18]. A related relation of quasi-commutativity has important applications in quantum mechanics as well, see [3]. Furthermore, transformations on quantum structures that preserve some relation or operation are usually called symmetries in physics and are of fundamental importance (see, for example, [4]).

Received by the editors February 28, 2008.

Published electronically May 20, 2010.

The authors were supported by a grant from the Ministry of Higher Education, Science and Technology, Slovenia.

AMS subject classification: 15A04, 15A27, 06A99.

Keywords: general preservers, matrix algebra, quasi-commutativity.

From a mathematical point of view, maps preserving a given algebraic property are called preservers and are extensively studied. We refer to [6, 19, 27] for linear (or additive) maps that preserve commutativity up to a factor ζ . Linear bijections that preserve quasi-commutativity in both directions were classified by Molnár [16]. Recently, Radjavi and Šemrl [22] improved his result by dropping the assumption of bijectivity and assuming that quasi-commutativity is preserved in one direction only. However, the maps they considered were still linear. It turned out that in many instances, the linearity assumption is superfluous. Without the linearity assumption, one can concentrate solely on the given structure. Among the first examples in this area, we mention Hua's result on adjacency preserving maps [13]. Recently the area of general preservers *i.e.*, not necessarily linear ones, has become very active, see, for example, [1, 5, 7, 15, 17, 20, 26].

In this paper, we follow this direction of research, and study (possibly nonlinear) maps that preserve quasi-commutativity. Without imposing linearity, however, the problem becomes much more difficult. We will therefore limit ourselves to bijective maps and will also assume that quasi-commutativity is preserved in both directions. Within these restrictions we obtain a nice structural result outside a certain rather small subset Ξ of M_n . This is best possible, since there is an example of a nonconstant bijection that preserves quasi-commutativity in both directions, fixes every matrix outside Ξ , but permutes elements of Ξ in a nonstandard way. It is given after the statement of the main theorem. We remark that, in a way, our result resembles other results on general preservers of commutativity up to a factor $\xi \in \{0, 1\}$ by Šemrl [23, 25] and by Chan, Li, and Sze [5], where similar obstructions were detected. In [23] it is also shown that, up to similarity, transposition, and field isomorphism applied entry-wise, nonlinear preservers of commutativity are regular locally polynomial maps outside a small pathological set. This means that for each A outside of this set, there are polynomials p_A and q_A such that $\Phi(A) = p_A(A)$ and $A = q_A(p_A(A))$. But then the classical commutant of A is equal to $A' = p_A(A)' = \Phi(A)'$. Our main theorem below is of the same kind, except that commutant is replaced with quasi-commutant.

Before giving our main result we introduce the necessary notation. First, define the relation $\overset{\#}{\sim}$ on M_n by $A \overset{\#}{\sim} B$ if and only if $A^\# = B^\#$. This is obviously an equivalence relation and as such it partitions M_n into equivalence classes $[\cdot]$. Next, given a matrix A and a field isomorphism $\sigma: \mathbb{C} \rightarrow \mathbb{C}$ (*i.e.*, an additive and multiplicative bijection on \mathbb{C}), let A^σ be the matrix obtained from A by applying σ entry wise. Also let A^{tr} be the transpose of A . Observe that $A \overset{\#}{\leftrightarrow} B$ is equivalent to $SAS^{-1} \overset{\#}{\leftrightarrow} SBS^{-1}$, to $A^{\text{tr}} \overset{\#}{\leftrightarrow} B^{\text{tr}}$, and to $A^\sigma \overset{\#}{\leftrightarrow} B^\sigma$. Moreover, it is easily seen (*cf.* Subsection 2.2) that any bijection $\Psi: M_n \rightarrow M_n$ with $\Psi(X) \in [X]$ also preserves quasi-commutativity in both directions.

Lastly, let $\Xi \subset M_n$ be the subset that consists of all matrices with at least two nonzero eigenvalues and whose minimal polynomial is of the form $m(\lambda) = \lambda^2 q(\lambda)$, $q(0) \neq 0$. Equivalently, their Jordan structure contains a nilpotent part of nil-index two and an invertible part with at least two invertible Jordan cells having different eigenvalues. Outside the set Ξ , our map will have a simple structure, but inside Ξ it is less tame. We are now ready to state our main theorem.

Theorem 1.1 Let $n \geq 3$. Suppose $\Phi: M_n \rightarrow M_n$ is a bijective map such that for any $A, B \in M_n$, we have

$$(1.1) \quad A \overset{\#}{\leftrightarrow} B \iff \Phi(A) \overset{\#}{\leftrightarrow} \Phi(B).$$

Then there exists a field isomorphism $\sigma: \mathbb{C} \rightarrow \mathbb{C}$ and a similarity S such that

- (i) $\Phi(A) \in [SA^\sigma S^{-1}]$, $A \in M_n \setminus \Xi$
- (ii) $\Phi(A) \in [S(A^\sigma)^{\text{tr}} S^{-1}]$, $A \in M_n \setminus \Xi$.

Remark 1.2 We can describe Φ on the whole M_n in the following way. Both maps $\Psi(X) = S^{-1}\Phi(X^{\sigma^{-1}})S$ and $\Psi(X) = S^{-1}\Phi(X^{\sigma^{-1}})^{\text{tr}}S$, where Φ is from (i) or (ii), respectively, satisfy $\Psi(A) \in \text{Poly}(A)$, see Lemma 2.36. Hence, Ψ is a locally polynomial map on the whole algebra M_n .

We conclude the section with an example that shows we cannot hope to get a nice structural result for every matrix in M_n .

Example 1.3 Let $\Phi: M_4 \rightarrow M_4$. For an arbitrary invertible $P \in M_2$, nonzero scalar $\alpha \in \mathbb{C}$, diagonal matrices $\text{diag}(\alpha, 2\alpha)$, $\text{diag}(\alpha, 3\alpha)$, and nonzero nilpotent $N \in M_2$, we define a map Φ with

$$\begin{aligned} \Phi(P \text{diag}(\alpha, 2\alpha)P^{-1} \oplus N) &= (P \text{diag}(\alpha, 3\alpha)P^{-1} \oplus N), \\ \Phi(P \text{diag}(\alpha, 3\alpha)P^{-1} \oplus N) &= (P \text{diag}(\alpha, 2\alpha)P^{-1} \oplus N), \\ \Phi(N \oplus P \text{diag}(\alpha, 2\alpha)P^{-1}) &= (N \oplus P \text{diag}(\alpha, 3\alpha)P^{-1}), \\ \Phi(N \oplus P \text{diag}(\alpha, 3\alpha)P^{-1}) &= (N \oplus P \text{diag}(\alpha, 2\alpha)P^{-1}), \end{aligned}$$

and $\Phi(A) = A$ for all other matrices. This is well defined since $P_1 \text{diag}(\alpha, 2\alpha)P_1^{-1} = P_2 \text{diag}(\alpha, 2\alpha)P_2^{-1}$ if and only if $P_2^{-1}P_1$ is diagonal, and this further gives

$$P_1 \text{diag}(\alpha, 3\alpha)P_1^{-1} = P_2 \text{diag}(\alpha, 3\alpha)P_2^{-1}.$$

It is easy to see that Φ is bijective and preserves quasi-commutativity in both directions. Moreover, $A = \Phi(A) \in [A]$ for every rank-one and every diagonalizable matrix A . However, a matrix $C = ((E_{11} + 2E_{22}) \oplus E_{12})$ is mapped into $\Phi(C) = ((E_{11} + 3E_{22}) \oplus E_{12}) \notin [C]$, because $(E_{21} + E_{33} + 2E_{44}) \in C^\# \setminus \Phi(C)^\#$.

2 Proof

2.1 Basic Properties

We start by listing some examples which will clarify the properties of the quasi-commutativity relation on certain matrices. These properties will be used in the sequel.

Example 2.1 Let $D = \text{diag}(d_1, \dots, d_n)$ be an invertible diagonal matrix. It induces the equivalence relation \sim_D between ordered pairs from the set $\mathbb{N}_n \times \mathbb{N}_n = \{(i, j) : 1 \leq i, j \leq n\}$, defined by

$$(i, j) \sim_D (u, v) \quad \text{if and only if} \quad \frac{d_i}{d_j} = \frac{d_u}{d_v}.$$

Then

$$D^\# = \bigcup_{\mathbf{e} \in (\mathbb{N}_n \times \mathbb{N}_n) / \sim_D} \text{Lin}\{E_{ij} : (i, j) \in \mathbf{e}\},$$

where $\text{Lin } \Omega$ is a linear span of a set Ω . We emphasize that $D^\#$ is not a linear subspace.

In the following example we expose a quasi-commutant of a very special invertible diagonal matrix: the one with only ± 1 on a diagonal. It will be shown later that its quasi-commutant is maximal in some sense.

Example 2.2 Let $D = \text{Id}_k \oplus (-\text{Id}_{n-k})$. Then,

$$D^\# = \left\{ \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}, \begin{pmatrix} 0 & * \\ * & 0 \end{pmatrix} \right\}.$$

Example 2.3 Let $D = \text{diag}(d_1, \dots, d_k, 0, \dots, 0)$ be a noninvertible diagonal matrix, with $d_1, \dots, d_k \neq 0$. Then $D^\# = D_k^\# \oplus M_{n-k}$, where $D_k = \text{diag}(d_1, \dots, d_k)$ is an invertible $k \times k$ matrix.

A special type of diagonalizable matrix is also an idempotent. In that case, however, the quasi-commutant is a linear subspace.

Example 2.4 If $A^2 = A$, then $A^\#$ equals A' , the classical commutant of A . The same is true also for scalar multiples of idempotents.

We continue by stating two well-known lemmas. The first one is an easy consequence of properties of elementary operators of length two. Namely, given complex matrices $A \in M_k$ and $B \in M_p$, the spectrum of the elementary operator $\mathbb{T} : M_{k \times p} \rightarrow M_{k \times p}$, defined on $k \times p$ complex matrices by $X \mapsto AX + \mu XB$, equals $\text{Sp } \mathbb{T} = \text{Sp } A + \mu \text{Sp } B = \{\alpha + \mu\beta : \alpha \in \text{Sp } A, \beta \in \text{Sp } B\}$. For a proof see, for example, [12, Theorem 4.4.5], or [2, page 2] for some further properties. We will use only the following weak consequence of this fact.

Lemma 2.5 *If $0 \notin \text{Sp } A + \mu \text{Sp } B$, then $AX + \mu XB = 0$ implies $X = 0$.*

The second lemma on the bicommutant of an arbitrary matrix A is also well known, see, for example, [28, page 106].

Lemma 2.6 *Let $A \in M_n$. Then $A'' = \text{Poly}(A) = \{p(A) : p \in \mathbb{C}[x]\}$.*

Using Lemma 2.6, some further properties of the quasi-commutant can be explained.

Lemma 2.7 *Let $A \in M_n$. Then $A^{\#\#} \subseteq A''$.*

Proof Pick any $Z \in A^{\#\#}$. Clearly $A' \subseteq A^{\#}$ so, in particular, Z must quasi-commute with every matrix $X \in A'$. We need to show that Z actually commutes with every $X \in A'$; once this is done, we have $Z \in A''$, as claimed. Now, assume erroneously that Z does not commute with some $X \in A'$. Then, for some nonzero $\mu \neq 1$,

$$(2.1) \quad ZX = \mu XZ.$$

However, $(\lambda \text{Id} + X) \in A' \subseteq A^{\#}$, so Z must quasi-commute with $(\lambda \text{Id} + X)$ for every $\lambda \in \mathbb{C}$. That is, $Z(\lambda \text{Id} + X) = \mu_\lambda(\lambda \text{Id} + X)Z$. If we rearrange this equation with the help of (2.1) we obtain

$$\lambda(1 - \mu_\lambda)Z = (\mu_\lambda - \mu)XZ$$

for every $\lambda \in \mathbb{C}$. If Z and XZ are linearly independent, this immediately gives $\mu = \mu_\lambda = 1$, a contradiction. If, on the other hand, $XZ = \alpha Z$, the above equation reduces to $\lambda(1 - \mu_\lambda)Z = \alpha(\mu_\lambda - \mu)Z$. Clearly $\alpha Z \neq 0$, otherwise $XZ = 0$ and equation (2.1) would force $ZX = 0$ hence Z would commute with X . So, with $\lambda = -\alpha \neq 0$, we get $-(1 - \mu_\lambda) = (\mu_\lambda - \mu)$ and once again $\mu = 1$, a contradiction. Indeed, Z must commute with $X \in A'$. ■

By Lemmas 2.6 and 2.7, we obtain the following corollary.

Corollary 2.8 *Let $A \in M_n$. Then $A^{\#\#} \subseteq \text{Poly}(A)$.*

Corollary 2.9 *Let $A, B \in M_n$. Then $A^{\#} \subseteq B^{\#}$ implies $B \in A^{\#\#}$. More precisely, $B = p(A)$ for some polynomial p .*

Proof $A^{\#} \subseteq B^{\#}$ implies $B^{\#\#} \subseteq A^{\#\#}$. We also have $B \in B^{\#\#}$, hence $B \in A^{\#\#} \subseteq \text{Poly}(A)$. ■

Lemma 2.10 *Let $A \in M_n$. Then $A^{\#} = M_n$ if and only if A is a scalar matrix.*

Proof Suppose $A^{\#} = M_n$, that is, A quasi-commutes with every matrix. Since $A \overset{\#}{\leftrightarrow} E_{ii}$ for every $i \in \mathbb{N}_n$, it follows that A is diagonal. And since $A \overset{\#}{\leftrightarrow} (E_{ii} + E_{ij})$ for every $i, j \in \mathbb{N}_n$, we obtain that A is a scalar matrix. The opposite implication is trivial. ■

Corollary 2.11 *Let Φ be as in Theorem 1.1. Then $\Phi(\mathbb{C}\text{Id}) = \mathbb{C}\text{Id}$.*

Proof The map Φ is surjective, therefore condition (1.1) is equivalent to $\Phi(A^{\#}) = \Phi(A)^{\#}$, and the result follows immediately from Lemma 2.10. ■

We finish this subsection with the quasi-commutant of elementary Jordan nilpotent.

Lemma 2.12 *If $N_k = \sum_{i=1}^{k-1} E_{i,i+1}$ is an elementary Jordan nilpotent with nil-index k and rank $k - 1$, $k \geq 2$, then*

$$N_k^{\#\#} = \{\mathbb{C}\text{Id}, \mathbb{C}N_k, \mathbb{C}N_k^2, \dots, \mathbb{C}N_k^{k-1}\}.$$

Proof There exists an invertible matrix X with $N_k X = 2XN_k$; for example, we might take $X = \text{diag}(1, 2, 2^2, \dots, 2^k) \oplus (2\text{Id}_{n-k})$. Then $N_k^i X = 2^i XN_k$. Now, by Corollary 2.8, any $B \in N_k^{\#}$ is a polynomial in N_k . So for some α_i , we have

$$B = \alpha_0 \text{Id} + \alpha_1 N_k + \dots + \alpha_{k-1} N_k^{k-1}.$$

Clearly $X \in N_k^{\#}$ so $B \overset{\#}{\leftrightarrow} X$. That is, for some $\mu \neq 0$ we have

$$\sum_{i=0}^{k-1} \alpha_i N_k^i X = BX = \mu XB = \mu \sum_{i=0}^{k-1} \alpha_i XN_k^i = \mu \sum_{i=0}^{k-1} 2^{-i} \alpha_i N_k^i X.$$

Compare the first and the last equation. We may cancel out the invertible X to get

$$\sum_{i=0}^{k-1} \alpha_i N_k^i = \mu \sum_{i=0}^{k-1} 2^{-i} \alpha_i N_k^i.$$

But N_k^0, \dots, N_k^{k-1} are linearly independent matrices, so $\alpha_i = \mu 2^{-i} \alpha_i$ for each $i = 0, \dots, k-1$. Now, if $\alpha_i \neq 0$ for some i , we have $\mu = 2^i$ in which case $\alpha_j = 0$ for indices $j \neq i$. Hence, $B = \alpha_i N_k^i$ as claimed. ■

2.2 Poset Structure

Recall the definition of equivalence relation $A \overset{\#}{\sim} B$ by $A^{\#} = B^{\#}$. The corresponding equivalence classes are $[A] = \{X : X^{\#} = A^{\#}\}$. On these equivalence classes we define the partial order with $[A] \leq [B]$ if and only if $A^{\#} \subseteq B^{\#}$, and we say that $[A]$ connects to $[B]$.

We have already observed that a bijective map $\Phi: M_n \rightarrow M_n$ preserves quasi-commutativity in both directions if and only if $\Phi(A^{\#}) = \Phi(A)^{\#}$ for each $A \in M_n$. Hence, $[A] = [B]$ if and only if $[\Phi(A)] = [\Phi(B)]$. Therefore such Φ induces a well defined bijection $\hat{\Phi}: M_n / \overset{\#}{\sim} \rightarrow M_n / \overset{\#}{\sim}$ via $\hat{\Phi}: [A] \mapsto [\Phi(A)]$, which preserves the partial order in both directions.

Observe also that a map satisfying condition (1.1) is determined only up to an equivalence class. Namely, assume Ψ is any bijection with $\Psi(X) \in [\Phi(X)]$ for every X . If $A \overset{\#}{\leftrightarrow} B$ then $\Phi(A) \overset{\#}{\leftrightarrow} \Phi(B)$, so $\Phi(A) \in \Phi(B)^{\#} = \Psi(B)^{\#}$, and therefore $\Phi(A) \overset{\#}{\leftrightarrow} \Psi(B)$. Hence, $\Psi(B) \in \Phi(A)^{\#} = \Psi(A)^{\#}$, so $\Psi(B) \overset{\#}{\leftrightarrow} \Psi(A)$. Therefore, $A \overset{\#}{\leftrightarrow} B$ implies $\Psi(A) \overset{\#}{\leftrightarrow} \Psi(B)$. The reversed implication is similar. This proves that Ψ also satisfies condition (1.1).

We say that a nonscalar matrix $A \in M_n$ is maximal if its equivalence class $[A]$ is maximal, that is, for any $B \in M_n$, the relation $[A] \leq [B]$ implies $[A] = [B]$ or $[B] = [\text{Id}]$. Notice that $[B] = [\text{Id}]$ is equivalent to $B^{\#} = M_n$. Similarly, we say that $A \in M_n$ is minimal if its equivalence class $[A]$ is minimal, that is $[B] \leq [A]$ implies $[A] = [B]$ for any $B \in M_n$. Moreover, B is an immediate predecessor of A if the following conditions are met:

- (i) $[B] \lessdot [A]$,
- (ii) there exists no matrix C with $[B] \lessdot [C] \lessdot [A]$.

Likewise we define an immediate successor.

We next consider the poset structure of equivalence classes of diagonal matrices.

Observe first that each nonscalar diagonal matrix with ± 1 on the diagonal is always maximal. Indeed, let $D = \text{diag}(\pm 1, \dots, \pm 1)$ be a nonscalar diagonal matrix. Using similarity, we may assume that $D = \text{Id}_k \oplus (-\text{Id}_{n-k})$. Suppose now that $D^\# \subseteq B^\#$ for some nonscalar $B \in M_n$. Then by Corollary 2.9, $B = p(D)$ for some polynomial p , that is $B = \alpha \text{Id}_k \oplus \beta \text{Id}_{n-k}$. If either $\alpha = 0$ or $\beta = 0$, then $B^\# = B' \subseteq D^\#$ by Examples 2.4 and 2.2. If α and β are both nonzero and $\alpha \neq -\beta$, then again $B^\# \subseteq D^\#$ by Example 2.1 and 2.2. Lastly, if $\alpha = -\beta \neq 0$, then $D^\# = B^\#$.

However, there are more maximal diagonal matrices. For example, it can be shown that $D = \text{diag}(1, \omega, \omega^2)$, $\omega = e^{2\pi i/3}$ is also maximal.

Lemma 2.13 *There are only finitely many maximal equivalence classes among diagonal matrices, and each of them contains only invertible diagonal matrices. No singular diagonal matrix is maximal.*

Proof Clearly, a singular diagonal matrix cannot be maximal. For if a diagonal D is singular, we can find a permutation matrix S such that $D = S^{-1}(\tilde{D} \oplus 0_{n-k})S$, with $\tilde{D} \in M_k$ invertible diagonal. Then, $D^\# = S^{-1}(\tilde{D}^\# \oplus M_{n-k})S$. This is strictly contained in a quasi-commutant of a diagonal $A = S^{-1}(\text{Id}_k \oplus (-\text{Id}_{n-k}))S$, see Example 2.2.

Hence, among diagonal matrices, maximal equivalence classes consist only of invertible matrices. But in view of Example 2.1, the quasi-commutant of each invertible diagonal matrix is in bijective correspondence with equivalence relations on $\mathbb{N}_n \times \mathbb{N}_n$ subject to some constraints. Clearly, on a set with n^2 elements, we may choose at most finitely many such relations. ■

Lemma 2.14 *Let $\lambda_1, \dots, \lambda_k$ be nonzero scalars and let $n_1, \dots, n_k \geq 1$ be integers. Then $\lambda_1 \text{Id}_{n_1} \oplus \dots \oplus \lambda_k \text{Id}_{n_k}$ is a maximal $n \times n$ matrix ($n = \sum n_i$) if and only if $\text{diag}(\lambda_1, \dots, \lambda_k)$ is a maximal $k \times k$ matrix.*

Proof First, let us prove that given invertible $A = \text{diag}(\lambda_1, \dots, \lambda_k)$ and $B = \text{diag}(\mu_1, \dots, \mu_k)$, we have $A^\# \subseteq B^\#$ if and only if $\lambda_i/\lambda_j = \lambda_u/\lambda_v$ implies $\mu_i/\mu_j = \mu_u/\mu_v$ for every $i, j, u, v \in \mathbb{N}_k$.

Indeed, suppose $A^\# \subseteq B^\#$. Then $\lambda_i/\lambda_j = \lambda_u/\lambda_v$ implies $X = E_{ij} + E_{uv} \in A^\# \subseteq B^\#$. So there exists a nonzero ξ with $XB = \xi BX$ which yields $(\mu_j E_{ij} + \mu_v E_{uv}) = \xi(\mu_i E_{ij} + \mu_u E_{uv})$, that is $\mu_i/\mu_j = \xi^{-1} = \mu_u/\mu_v$. Inversely, given the condition on the quotients, $X \in A^\#$ implies

$$X \in \text{Lin}\{E_{ij} : \lambda_i/\lambda_j = \lambda_{u_0}/\lambda_{v_0}\} \subseteq \text{Lin}\{E_{ij} : \mu_i/\mu_j = \mu_{u_0}/\mu_{v_0}\} \subseteq B^\#$$

(see Example 2.1).

From there we easily deduce that $A^\# \subseteq B^\#$ if and only if $\hat{A}^\# \subseteq \hat{B}^\#$, where $\hat{A} = \lambda_1 \text{Id}_{n_1} \oplus \dots \oplus \lambda_k \text{Id}_{n_k}$ and $\hat{B} = \mu_1 \text{Id}_{n_1} \oplus \dots \oplus \mu_k \text{Id}_{n_k}$.

Now, we argue as follows: let $A = \text{diag}(\lambda_1, \dots, \lambda_k)$ be maximal. Consider $\hat{A} = \lambda_1 \text{Id}_{n_1} \oplus \dots \oplus \lambda_k \text{Id}_{n_k}$. Suppose $\hat{A}^\# \subseteq \hat{B}^\#$. Then \hat{B} is a polynomial in \hat{A} , so $\hat{B} = \mu_1 \text{Id}_{n_1} \oplus \dots \oplus \mu_k \text{Id}_{n_k}$. If all $\mu_i = 0$, then \hat{B} is scalar. Further, if at least one diagonal element of \hat{B} is nonzero and at least one is zero, say $\mu_1 \neq 0$ and $\mu_k = 0$, then $E_{1n} \in \hat{A}^\#$.

But $E_{1n} \notin \hat{B}^\#$, a contradiction. Thus \hat{B} is invertible in this case. By the previous paragraph, $B = \text{diag}(\mu_1, \dots, \mu_k)$ also satisfies $A^\# \subseteq B^\#$. Regarding maximality, B is scalar or $B^\# = A^\#$, wherefrom $\hat{A}^\# = \hat{B}^\#$. Thus, A is maximal. Similar arguments can be used for the reverse implication. ■

We now turn our attention to minimal matrices. Let us start by showing that $D_0 = \text{diag}(2, 3, \dots, p_n)$, where p_n is the n -th consecutive prime, is minimal. The matrix D_0 will play an important role in the proof of Theorem 1.1.

Lemma 2.15 *The matrix D_0 is minimal. Moreover, given any invertible diagonal B , we have $[D_0] \leq [B]$.*

Proof Since $p_i/p_j = p_u/p_v$ precisely when either $(i, j) = (u, v)$ or $i = j$ and $u = v$, Example 2.1 gives that D_0 is below every invertible diagonal matrix B . To prove its minimality, suppose $[A] \leq [D_0]$, i.e., $A^\# \subseteq D_0^\#$ for some $A \in M_n$. Then, $D_0 = p(A)$ for some polynomial p by Corollary 2.9. Observe that D_0 has n distinct eigenvalues. So, by the Spectral Mapping Theorem, A also has n distinct eigenvalues and is therefore diagonalizable. Moreover A and D_0 are simultaneously diagonalizable since $D_0 = p(A)$. So we may assume that $A = \text{diag}(a_1, \dots, a_n)$. If, say, $a_n = 0$ then $E_{12} + E_{nn} \in A^\# \setminus D_0^\#$, a contradiction. Hence, $a_n \neq 0$, and likewise for any other a_i . Therefore, A is an invertible diagonal matrix, and as we showed at the beginning of the proof, $A^\# \supseteq D_0^\#$. Hence, D_0 is minimal. ■

As a consequence of Lemma 2.15, there is only one minimal equivalence class among invertible diagonal matrices, and it equals $[D_0]$.

Lemma 2.16 *Let $D \in [D_0]$. Then*

$$D^{\#\#} = \{\text{diag}(a_1, \dots, a_n) : a_i \in \mathbb{C} \setminus \{0\}, i = 1, \dots, n\} \cup \{0\}.$$

Proof Since $D^{\#\#} \subseteq \text{Poly} D$ by Corollary 2.8, any $X \in D^{\#\#}$ is diagonal. Note that $E_{ij} \in D^\#$ for all $i, j \in \mathbb{N}_n$ and $X = \text{diag}(x_1, \dots, x_n) \overset{\#}{\leftrightarrow} E_{ij}$ if and only if x_i and x_j are either both nonzero or both zero. Hence X is invertible diagonal or zero. ■

Lemma 2.17 *A nonzero singular diagonal matrix $D = \text{diag}(d_1, \dots, d_n)$ is minimal if and only if $d_i = 0$ for exactly one index $i = i_0$ and the diagonal matrix $\tilde{D} = \text{diag}(d_1, \dots, d_{i_0-1}, d_{i_0+1}, \dots, d_n)$ is a minimal invertible $(n-1) \times (n-1)$ matrix.*

Proof Without loss of generality we may assume that $D = D_k \oplus 0_{n-k}$ where D_k is an invertible diagonal matrix. Then $D^\# = D_k^\# \oplus M_{n-k}$ by Example 2.3. If D_k is not minimal then, by Lemma 2.15, there exists a minimal invertible matrix \tilde{D} , $\tilde{D}^\# \subsetneq D_k^\#$ and $(\tilde{D} \oplus 0_{n-k})^\# \subsetneq D^\#$. If $k \geq 2$, then $(D_k \oplus J_{n-k}(0))^\# \subsetneq D^\#$, where $J_s(\alpha) = \alpha \text{Id}_s + \sum_{i=1}^{s-1} E_{i,i+1}$.

To prove the opposite direction, assume $D = D_{n-1} \oplus 0$ with D_{n-1} a minimal invertible $(n-1) \times (n-1)$ diagonal matrix, and suppose $[A] \leq [D]$ for some $A \in M_n$. Observe that D has n distinct eigenvalues since D_{n-1} has $n-1$ distinct nonzero eigenvalues (this follows easily from Lemma 2.15 and Example 2.1). By Corollary 2.9, $D = p(A)$ for some polynomial p , so by the Spectral Mapping Theorem, A also has n distinct eigenvalues and is therefore diagonalizable. Moreover, A and D

are simultaneously diagonalizable since $D = p(A)$. So we may assume that $A = \text{diag}(a_1, \dots, a_n)$ and $D = D_{n-1} \oplus 0$. If $a_n \neq 0$ then $a_i \neq 0$ for some $i \in \mathbb{N}_{n-1}$, and so $E_{ni} \in A^\# \setminus D^\#$, a contradiction. Hence $a_n = 0$ and $a_i \neq 0$ for every $i = 1, \dots, n - 1$. It follows that $A^\# = \text{diag}(a_1, \dots, a_{n-1})^\# \oplus \mathbb{C}$, so $\text{diag}(a_1, \dots, a_{n-1})^\# \subseteq D_{n-1}^\#$. Since D_{n-1} is minimal we obtain $\text{diag}(a_1, \dots, a_{n-1})^\# = D_{n-1}^\#$, hence $[A] = [D]$. ■

We continue by showing how to distinguish between minimal invertible diagonal matrices and singular diagonal ones, using only the quasi-commutativity relation.

Corollary 2.18 *Let D be a minimal diagonal matrix. Then the following two facts are equivalent:*

- (i) D is singular;
- (ii) $D^{\#\#}$ contains precisely one nonscalar maximal matrix, up to equivalence.

Proof (i) \implies (ii) Suppose D is singular. We can find a permutation matrix S such that $SDS^{-1} = D_1 \oplus 0_{n-k}$ for some invertible diagonal $k \times k$ matrix D_1 . Clearly $k \geq 1$, for otherwise $D = 0$ which is not minimal. Now, Corollary 2.8 gives $D^{\#\#} \subseteq \text{Poly}(D)$. Hence, any $X \in D^{\#\#}$ takes the form $X = S(\text{diag}(x_1, \dots, x_k) \oplus \alpha \text{Id}_{n-k})S^{-1}$.

If $\alpha = 0$, such X can never be maximal (see Lemma 2.13). On the other hand, if $\alpha \neq 0$, then X must quasi-commute with every matrix from $D^\# = S(D_1^\# \oplus M_{n-k})S^{-1}$. In particular, with $Z = S(E_{ij}^{k \times k} \oplus \text{Id}_{n-k})S^{-1}$, where $E_{ij}^{s \times t}$ is an $s \times t$ rectangular matrix unit. Hence, for some nonzero μ , we have $XZ = \mu ZX$. Therefore, $x_i E_{ij} = \mu x_j E_{ij}$ and $\alpha \text{Id}_{n-k} = \alpha \mu \text{Id}_{n-k}$, which gives $\mu = 1$ and, consequently, $x_i = x_j$ for $i, j \in \mathbb{N}_k$. So, $X = S(x \text{Id}_k \oplus \alpha \text{Id}_{n-k})S^{-1}$. Clearly, this is maximal precisely when $x = -\alpha \neq 0$. However, all such matrices are equivalent.

(i) \impliedby (ii) Suppose D is not singular. By Lemma 2.16, $D^{\#\#}$ contains nonequivalent matrices $\text{diag}(-1, 1, -1, \dots, -1, -1)$ and $\text{diag}(-1, \dots, -1, 1)$ (cf. Example 2.2), which we have already shown to be maximal. ■

2.3 Classifying Minimal Diagonal Matrices

In this subsection, we will show how to distinguish minimal diagonalizable matrices from minimal nondiagonalizable ones. This will be done by counting the number of maximal equivalence classes that an equivalence class of a matrix connects to.

Let $\mathfrak{d}(n)$ denote the number of pairwise, nonequivalent, maximal diagonal $n \times n$ matrices. In view of Lemma 2.13, there are only finitely many such, and each one contains only invertible diagonal matrices. So $\mathfrak{d}(n) < \infty$. On the other hand, we can show that $\mathfrak{d}(n) \geq 2^{(n-1)} - 1$. Namely, recall that each nonscalar diagonal matrix with ± 1 on its diagonal is maximal. Moreover, two such, say B_1, B_2 , are equivalent precisely when $B_1 = -B_2$. We have 2^n possible choices for ± 1 . Grouping together the equivalent pairs, while avoiding the scalars, yields that there are at least $2^{(n-1)} - 1$ pairwise nonequivalent maximal diagonal matrices.

Lemma 2.19 *Given a positive integer k , we have $\mathfrak{d}(k + 1) \geq \mathfrak{d}(k) + 1$. Moreover, if $k \geq 2$, then $\mathfrak{d}(k + 1) \geq \mathfrak{d}(k) + 2$.*

Proof By Lemma 2.14, if $\text{diag}(\mu_1, \dots, \mu_k) \in M_k$ is maximal, then so also is $\text{diag}(\mu_1, \dots, \mu_k, \mu_k) \in M_{k+1}$. Hence we easily deduce $\mathfrak{d}(k) \leq \mathfrak{d}(k+1)$. In addition, we

can choose precisely $2^{k-1} - 1$ maximal diagonal $k \times k$ matrices with ± 1 entries on the diagonal, while we can choose $2^{(k+1)-1} - 1 > 2^{k-1} - 1$ such $(k + 1) \times (k + 1)$ matrices. This gives $\mathfrak{d}(k) + 1 \leq \mathfrak{d}(k + 1)$. Actually, $k \geq 2$ implies $2^{(k+1)-1} - 1 - (2^{k-1} - 1) \geq 2$, and so $\mathfrak{d}(k + 1) \geq 2 + \mathfrak{d}(k)$. ■

Lemma 2.20 *Let an invertible diagonal matrix A be minimal. Then, $[A]$ connects to precisely $\mathfrak{d}(n)$ maximal equivalence classes.*

Proof By Lemma 2.15, $[A] = [D_0]$, and it connects to every maximal diagonal matrix. So, it connects to at least $\mathfrak{d}(n)$ maximal classes. Next, let $[B]$ be any maximal equivalence class. If $[A] \leq [B]$ then, by Corollary 2.9, $B \in \text{Poly}(A)$. Consequently B is diagonal and invertible (see Lemma 2.13). Hence, $[A]$ connects to precisely $\mathfrak{d}(n)$ maximal classes. ■

We continue with analyzing minimal nondiagonalizable matrices. It should be remarked that we will not prove that such a matrix exists. We will only show that, if there is one, it never connects to precisely $\mathfrak{d}(n)$ maximal equivalence classes. To demonstrate this we require some auxiliary results.

Lemma 2.21 *Any square-zero nonscalar matrix is maximal.*

Proof Let $N^2 = 0$. If $[N] \leq [B]$ then B is a polynomial in N . The only possibilities are $B = \lambda \text{Id} + \mu N$, $\lambda, \mu \in \mathbb{C}$. If $\mu = 0$ then B is scalar. Assume $\mu \neq 0$. On the one hand, if $\lambda = 0$, then, trivially, $[B] = [N]$. On the other hand, if $\lambda \neq 0$, then $\text{Sp } B = \{\lambda\} \neq \{0\}$ implies $B^\# = B'$ (see Lemma 2.5). However, $B' = N' \subseteq N^\#$, so $[B] = [N]$, as anticipated. ■

Lemma 2.22 *We have $X^\# \subseteq (X^2)^\#$. Consequently, $[X] \leq [X^2]$.*

Proof Suppose $Z \in X^\#$. Then there exists nonzero $\mu \in \mathbb{C}$ such that $XZ = \mu ZX$. Then $X^2Z = \mu^2 ZX^2$ and $Z \in (X^2)^\#$. ■

Lemma 2.23 *Suppose $A = \hat{A} \oplus N(0)$ is a block-diagonal matrix, with \hat{A} its invertible part and $N(0)$ nilpotent. If a corresponding block-matrix $Z = (Z_{ij})_{1 \leq i, j \leq 2}$ is in $A^\#$, then $Z_{12} = 0 = Z_{21}$.*

Proof If $Z \in A^\#$, then $AZ = \mu ZA$, $\mu \neq 0$. So the block (12) of this equation is $\hat{A}Z_{12} = \mu Z_{12}N(0)$. Equivalently, $\hat{A}Z_{12} - Z_{12}(\mu N(0)) = 0$. However, $\text{Sp } \hat{A} - \text{Sp}(\mu N(0)) = \text{Sp } \hat{A} - \{0\} = \text{Sp } \hat{A}$ does not contain zero so, by Lemma 2.5, this equation has only a zero solution. In the same way we prove that $Z_{21} = 0$. ■

Recall that a matrix is nonderogatory provided each eigenvalue has geometric multiplicity one, i.e., no two Jordan cells share the same eigenvalue.

Lemma 2.24 *Let $A = \hat{A} \oplus 0$ with \hat{A} invertible. There exists $t \in \mathbb{N}$ and a nonderogatory matrix \hat{B} such that $(\hat{B} \oplus 0)^t = A$ and such that the quotients of its eigenvalues are pairwise distinct, i.e., for any $\lambda, \mu, \alpha, \beta \in \text{Sp } \hat{B}$ with $\lambda \neq \mu$, we have $\lambda/\mu = \alpha/\beta$ precisely when $(\lambda, \mu) = (\alpha, \beta)$.*

Proof Assume with no loss of generality that $\hat{A} = \bigoplus_{i=1}^k J_{n_i}(\lambda_i)$ is already in its Jordan form. Now, given any p -th root $\mu_i = \sqrt[p]{\lambda_i}$, it is easy to construct a matrix \hat{B} , similar to $\bigoplus_{i=1}^k J_{n_i}(\mu_i)$, such that $\hat{B}^p = \hat{A}$. Therefore, it only remains to find $p = 2^t$ and p -th roots μ_i of $\lambda_i = |\lambda_i|e^{2\pi\sqrt{-1}\phi_i}$ ($0 \leq \phi_i \leq 1$) that have pairwise distinct quotients.

To this end, we rearrange ϕ_i in nondecreasing order and recursively construct a strictly increasing sequence of positive integers r_i in such a way that $|(r_i + \phi_i) - (r_1 + \phi_1)| < |(r_{i+1} + \phi_{i+1}) - (r_i + \phi_i)|$. For $x_i = (r_i + \phi_i)$, we obtain that $|x_i - x_j| = |x_u - x_v|$ precisely when $(i, j) = (u, v)$ or $i = j$ and $u = v$.

Pick any $p = 2^t > 4(r_k + \phi_k)$ and let $\mu_i = \sqrt[p]{|\lambda_i|}e^{2\pi\sqrt{-1}(r_i + \phi_i)/p}$. Then, μ_i are the p -th roots of λ_i and for these $\mu_i/\mu_j = \mu_u/\mu_v$ is possible only when either $(i, j) = (u, v)$ or $i = j$ and $u = v$. ■

We now count the maximal equivalence classes that a nondiagonalizable minimal matrix A connects to. Clearly, this number depends solely on the Jordan cell structure of A . It turns out that in the presence of invertible Jordan cells this number is different than in the presence of singular Jordan cells. We will cover all possibilities by examining the following cases separately.

- (i) At least one cell of dimension ≥ 4 is invertible.
- (ii) The above does not apply, but at least one invertible cell and at least one singular cell have dimensions ≥ 2 .
- (iii) Neither of the above applies, but all invertible cells of A (if any) have dimensions 1, and there exists a singular cell of dimension ≥ 2 .
- (iv) All remaining possibilities: the maximal dimension of invertible Jordan cells of A is $d \in \{2, 3\}$, however, all singular cells (if any) are 1×1 .

We start with (i).

Lemma 2.25 *If a Jordan structure of A contains an invertible cell of dimension at least four, then $[A]$ connects to infinitely many maximal classes.*

Proof Assume with no loss of generality that $A = \hat{A} \oplus N(0) = J_{n_1}(\lambda_1) \oplus \dots \oplus J_{n_k}(\lambda_k) \oplus N(0)$ is already in its Jordan form with dimensions n_i of invertible Jordan cells arranged in decreasing order and with $N(0)$ containing all nilpotent Jordan cells (we omit $N(0)$ if A itself is invertible). For an arbitrary $\alpha \in \mathbb{C}$, consider the matrix

$$A_\alpha = \hat{A}_\alpha \oplus 0 = \bigoplus_{i=1}^k \left(\alpha J_{n_i}(0)^{\max-2} + \frac{\lambda_i}{\lambda_i} J_{n_i}(0)^{\max-1} \right) \oplus 0, \quad \max = n_1.$$

By the assumptions, $n_1 \geq 4$, so A_α is a nonzero but square-zero matrix. Moreover, $[A_\alpha] = [A_\beta]$ implies $A_\alpha \in A_\alpha^\# = A_\beta^\# \subseteq \text{Poly}(A_\beta)$. So A_α is a polynomial in A_β which is possible only when $\alpha = \beta$. We conclude that, as $\alpha \in \mathbb{C}$ varies, $[A_\alpha]$ consists of infinitely many pairwise distinct maximal equivalence classes. It remains to show that $[A] \leq [A_\alpha]$, i.e., $A^\# \subseteq A_\alpha^\#$.

To this end, let $Z \in A^\#$. There exists some nonzero μ such that $AZ = \mu ZA$. Now, decompose $Z = \begin{pmatrix} (X_{ij})_{ij} & U \\ V & Y \end{pmatrix}$ according to the block structure of $A = \hat{A} \oplus N(0)$ and consider the above equation. It follows from invertibility of \hat{A} that $U = 0 = V$ (see

Lemma 2.23). That is, Z is block-diagonal. Consider now the X_{ij} block of the above equation. We get

$$(2.2) \quad J_{n_i}(\lambda_i)X_{ij} = \mu X_{ij}J_{n_j}(\lambda_j).$$

So by Lemma 2.5, either $X_{ij} = 0$ or else $\lambda_i = \mu\lambda_j$, in which case (2.2) further reduces into $(\lambda_i\text{Id}_{n_i} + J_{n_i}(0))X_{ij} = X_{ij}(\mu\lambda_j\text{Id}_{n_j} + \mu J_{n_j}(0))$, or equivalently into $J_{n_i}(0)X_{ij} = \mu X_{ij}J_{n_j}(0)$. However, if we consecutively pre-multiply this by $J_{n_i}(0)$, we deduce, by induction, that $J_{n_i}(0)^r X_{ij} = \mu^r X_{ij}J_{n_j}(0)^r = (\frac{\lambda_i}{\lambda_j})^r X_{ij}J_{n_j}(0)^r$ holds for any positive integer r . Therefore,

$$(\alpha J_{n_i}(0)^{\max-2} + \frac{\lambda_1}{\lambda_i} J_{n_i}(0)^{\max-1}) X_{ij} = \mu^{\max-2} X_{ij} (\alpha J_{n_j}(0)^{\max-2} + \frac{\lambda_1}{\lambda_j} J_{n_j}(0)^{\max-1}).$$

This remains valid even if $X_{ij} = 0$ (when possibly $\lambda_i \neq \mu\lambda_j$). Since the indices (i, j) were completely arbitrary, we deduce that $(X_{ij})_{ij}$ satisfies $\hat{A}_\alpha(X_{ij})_{ij} = \mu^{\max-2}(X_{ij})_{ij}\hat{A}_\alpha$. On the other hand, $Z = (X_{ij})_{ij} \oplus Y$ is block-diagonal, and it is easy to deduce that then

$$A_\alpha Z - \mu^{\max-2} Z A_\alpha = (\hat{A}_\alpha \oplus 0) ((X_{ij})_{ij} \oplus Y) - \mu^{\max-2} ((X_{ij})_{ij} \oplus Y) (\hat{A}_\alpha \oplus 0) = 0,$$

so, indeed, $Z \in A_\alpha^\#$. ■

We now proceed with (ii) of the above plan.

Lemma 2.26 *If a Jordan structure of A contains invertible and singular cells of dimension ≥ 2 simultaneously, then $[A]$ connects to infinitely many maximal classes.*

Proof Decompose A as in the proof of Lemma 2.25 and retain the notation. Note that by Lemma 2.25 we only need to consider the case where \max , the maximal dimension of invertible Jordan cell, is at most three. Let $\max_0 \geq 2$ be the nil-index of $N(0)$ (so, $N(0)^{\max_0-1} \neq 0$, but $N(0)^{\max_0} = 0$), and consider

$$A_\beta = \bigoplus_{i=1}^k ((\frac{\lambda_1}{\lambda_i})^m J_{n_i}(0)^{\max-1}) \oplus (\beta N(0)^{\max_0-1})$$

for $m = \max - \max_0$. Since $\max, \max_0 \geq 2$, the matrices A_β are nonzero but square-zero, and no polynomial transforms A_β into A_γ if $\beta \neq \gamma$. Hence, as $\beta \in \mathbb{C}$ varies, $[A_\beta]$ is comprised of infinitely many maximal equivalence classes.

Similar arguments as in the proof of Lemma 2.25 also validate $[A] \leq [A_\beta]$. ■

Case (iii), however, is fundamentally different.

Lemma 2.27 *Suppose all invertible Jordan cells of nondiagonalizable A (if present) have dimension 1. Then, $[A]$ connects to less than $\mathfrak{d}(n)$ maximal classes.*

Proof As before, we may assume $A = J_{n_1}(\lambda_1) \oplus \dots \oplus J_{n_k}(\lambda_k)$ is already in its Jordan block form with dimensions n_i of Jordan cells arranged in a nonincreasing order. Unlike the previous two cases, however, we do not treat nilpotent Jordan cells

separately, so λ_i can now be zero as well. If for a maximal B , $[A] \leq [B]$, then B must be a polynomial in A . Assume first that B is diagonalizable. Then it must be diagonal (use that B is a polynomial in an upper-triangular A), and of the form $B = (\mu_1 \text{Id}_{n_1}) \oplus \dots \oplus (\mu_k \text{Id}_{n_k})$. Moreover, by Lemma 2.13, it is invertible. In view of Lemma 2.14, B is maximal precisely when $\text{diag}(\mu_1, \dots, \mu_k)$ is maximal $k \times k$. We may, consequently, choose a total of $\mathfrak{d}(k)$ distinct maximal classes of diagonal matrices.

Now assume a maximal B is nondiagonalizable. It is easy to see that each Jordan cell of B must have zero diagonal, for otherwise the matrix B_0 with zero diagonal and the same off diagonal entries as B satisfies $B^\# \subseteq B_0^\#$. So B is a nilpotent – actually, a nilpotent of nil-index 2 by the fact that $[B] \leq [B^2]$ (see Lemma 2.22).

Recall once more that B is a polynomial in A , and simultaneously $B^2 = 0$. Recall also that at least one singular cell of A has dimension greater than one and all invertible cells (if present) are 1×1 . Therefore, the only possibilities for a maximal nondiagonalizable $[B] \geq [A]$ are that

$$B = \bigoplus_{n_i > 1} p(J_{n_i}(0)) \oplus 0$$

for some polynomial $p(x) = \zeta_r x^r + \zeta_{r+1} x^{r+1} + \dots$, where $r \geq 1/2 \max_0$, and $\max_0 \geq 2$ is the maximal dimension of singular cells. Note that $A^\# \subseteq B^\#$ implies

$$B \in B^{\#\#} \subseteq A^{\#\#} \subseteq \bigoplus_{n_i > 1} (J_{n_i}(0))^{\#\#} \oplus \left(\bigoplus_{\lambda_i \neq 0} \mathbb{C} \right).$$

However, $J_{n_i}(0)^{\#\#} = \mathbb{C}\{\text{Id}, J_{n_i}(0), J_{n_i}(0)^2, \dots, J_{n_i}(0)^{n_i-1}\}$ in view of Lemma 2.12. This rules out the possibility of a polynomial, which would not act in monomial fashion on the aggregation of singular Jordan cells. Hence,

$$B = \xi_r \left(\left(\bigoplus_{n_i > 1} J_{n_i}(0) \right)^r \oplus 0 \right), \quad r \geq \frac{1}{2} \max_0.$$

Denote the integer part of x by $\lfloor x \rfloor$. We can choose $\lfloor 1/2 \max_0 \rfloor$ such r , which gives at most $\lfloor 1/2 \max_0 \rfloor$ additional maximal nondiagonalizable equivalence classes that A connects to.

All together A connects to at most $\mathfrak{d}(k) + \lfloor 1/2 \max_0 \rfloor = \mathfrak{d}(k) + \lfloor 1/2 n_1 \rfloor$ maximal nonscalar equivalence classes. On the other hand, note that $k + \lfloor n_1/2 \rfloor \leq n$ so, after an easy exercise, Lemma 2.19 gives $\mathfrak{d}(k) + \lfloor n_1/2 \rfloor \leq \mathfrak{d}(n)$. ■

Up till now we did not use the fact that A is minimal. In the final case (iv), however, we will use a weak consequence of this.

Lemma 2.28 *Let $A = \hat{A} \oplus 0$, with \hat{A} invertible, be minimal and nondiagonalizable. Then it does not connect to precisely $\mathfrak{d}(n)$ maximal classes.*

Proof Assume with no loss of generality that \hat{A} is in its Jordan form, so $A = (\bigoplus_1^k J_{n_i}(\lambda_i)) \oplus 0$, with $n_1 \geq n_2 \geq \dots$. Since it is minimal, we may further assume that λ_i are pairwise distinct with distinct quotients. Otherwise we would use Lemma 2.24 to find B with that property, and with $B^{2^t} = A$. Lemma 2.22 would then give $[B] \leq [A]$.

We only need to examine the possibility when $n_1 \leq 3$; if it is greater, we already know A connects to infinitely many maximal classes. Now, as in the proof of Lemma 2.27, we may choose a total of $\mathfrak{d}(k)$ distinct maximal classes of diagonalizable matrices. In addition, every nondiagonalizable maximal B that A connects to is equivalent to a nilpotent of nil-index 2, which is at the same time a polynomial in A (see the proof of Lemma 2.27). We may therefore assume that

$$B = \left(\bigoplus_1^k \alpha_i J_{n_i}(0)^{n_i-1} \right) \oplus 0$$

for some α_i . We recall that $3 \geq n_1 \geq n_2 \dots$. If $n_1 = 3$, then all α_u , corresponding to 2×2 cells vanish. To see this, take $X = (X_{ij})_{ij}$, with all blocks zero, except for

$$X_{1u} = \begin{pmatrix} \lambda_u & 0 \\ 0 & \lambda_1 \\ 0 & 0 \end{pmatrix}.$$

It is elementary that $AX = (\lambda_1/\lambda_u)XA$, so $X \in A^\#$. However, $XB = \mu BX$ is equivalent to $\alpha_u \lambda_u \mu = 0$, forcing $\alpha_u = 0$.

With this information, we may rewrite B as

$$(2.3) \quad B = \left(\bigoplus_1^k \alpha_i J_{n_i}(0)^{\max-1} \right) \oplus 0; \quad \max = n_1 \in \{2, 3\}.$$

We are now facing two options.

Option 1 Only the first cell of A is of maximal dimension. Then A connects to exactly one nondiagonalizable maximal equivalence class: $[J_{n_1}(0)^{n_1-1} \oplus 0]$. Counting also diagonalizable ones, the totality of maximal classes that such an A connects to is $\mathfrak{d}(k) + 1$. But this is strictly less than $\mathfrak{d}(n)$ because $k \leq n - 1$, and hence $\mathfrak{d}(n) \geq \mathfrak{d}(n - 1) + 2 \geq \mathfrak{d}(k) + 2$ (see Lemma 2.19).

Option 2 Two or more invertible cells of A are of maximal dimension. We will show that A connects to infinitely many pairwise, nonequivalent, maximal B from equation (2.3) where in addition each $\alpha_i \neq 0$. Now, $Z \in A^\#$ implies $AZ = \mu ZA$. Since \hat{A} is invertible, we deduce $Z = (X_{ij})_{ij} \oplus Y$. Moreover, by Lemma 2.5, $X_{ij} \neq 0$ precisely when $\lambda_i = \mu \lambda_j$. Eigenvalues $\lambda_1, \dots, \lambda_k$ have pairwise distinct quotients, therefore either $\mu = 1$ and $(X_{ij})_{ij}$ is block-diagonal or all its blocks but at most one are zero.

Consequently, if $\mu = 1$, then Z commutes with A , and therefore also with any polynomial in A , that is also with B . Otherwise, at most one block, say $X_{i_0 j_0}$, is nonzero and we can show that $BZ = \frac{\alpha_{i_0}}{\alpha_{j_0}} \mu^{\max-1} ZB$ similarly as in the proof of Lemmas 2.25 and 2.26.

In either case, Z also quasi-commutes with B , whatever the nonzero scalars α_i in (2.3) are. Such A , therefore, connects to infinitely many maximal equivalence classes. ■

Corollary 2.29 *If X is a minimal invertible diagonal matrix, then $\Phi(X)$ is a minimal diagonalizable invertible matrix.*

Proof Clearly, $\Phi(X)$ must be minimal. By Lemmas 2.20, and 2.25–2.28, we know that $\Phi(X)$ must be diagonalizable. However, by Corollary 2.18, $\Phi(X)$ cannot be singular. The only possibility left is that $\Phi(X)$ is indeed a minimal invertible and diagonalizable. ■

2.4 Proof of Theorem 1.1

Recall from Lemma 2.15 that $D_0 = \text{diag}(2, 3, 5, \dots, p_n)$ is, up to equivalence, the only minimal, diagonal, invertible matrix. By Corollary 2.29, $\Phi(D_0)$ is also a minimal, diagonalizable, invertible matrix. Hence, with a suitably chosen similarity, $S\Phi(D_0)S^{-1} \in [D_0]$.

By Lemma 2.16, the set $D_0^\#$ consists of precisely all invertible diagonal matrices plus the zero matrix. Consequently, the map $S\Phi(\cdot)S^{-1}$ which remains bijective and preserves quasi-commutativity in both directions, maps the set of invertible diagonal matrices bijectively onto itself.

Next, if $\mu \neq 0$ and $i \neq j$, then $\mu E_{ij} \in D_0^\#$. In view of Example 2.1, $S\Phi(\mu E_{ij})S^{-1} \in (S\Phi(D_0)S^{-1})^\# = D_0^\#$ equals either some λE_{uv} , $u \neq v$, or it must be mapped into a diagonal matrix which would have to be maximal because μE_{ij} is maximal by Lemma 2.21. However, the latter contradicts bijectivity of $S\Phi(\cdot)S^{-1}$. Namely, we already know that invertible diagonal matrices are permuted among themselves and that singular diagonal matrices are not maximal by Lemma 2.13. Hence, for each nonzero μ and for each $i \neq j$, we can find some $u \neq v$ and nonzero scalar λ such that $S\Phi(\mu E_{ij})S^{-1} = \lambda E_{uv}$.

Next, $D_0^\#$ consists precisely of all diagonal matrices and all $\mathbb{C}E_{ij}$. We have just shown that $S\Phi(\cdot)S^{-1}$ permutes the elements of $\bigcup_{i \neq j} (\mathbb{C} \setminus \{0\})E_{ij}$ and since the same holds true for invertible diagonal matrices, we see that $S\Phi(\cdot)S^{-1}$ also permutes singular diagonal matrices among themselves. Therefore, it also permutes minimal singular diagonal ones.

Further, let us show that Φ preserves rank one nilpotents. Indeed, let $N = TE_{12}T^{-1}$. We define a bijective map $\Psi_T(X) = \Phi(TXT^{-1})$. Then $\Phi(N) = \Psi_T(E_{12})$. However, the map Ψ_T preserves quasi-commutativity in both directions so, as we did for Φ , we can find invertible S_T with $S_T\Psi_T(E_{12})S_T^{-1} = \lambda E_{uv}$ for some $\lambda \neq 0$ and $u \neq v$. Consequently, $\Phi(N) = \Psi_T(E_{12})$ is a nilpotent of rank-one. Applying the same procedure to Φ^{-1} , we see that the set of rank-one nilpotents is bijectively mapped onto itself by Φ .

We proceed by showing that N_1, N_2 and $(N_1 + N_2)$ are all rank-one nilpotents precisely when the same holds for their Φ -images. Indeed, it is known, and easy to see with the help of tensors, that (N_1, N_2) are simultaneously similar to either $(E_{12}, \mu E_{12})$, or (E_{12}, E_{13}) , or (E_{12}^t, E_{13}^t) . Assume T is this similarity. We then introduce a bijection $\Psi_T(X) = \Phi(TXT^{-1})$ which possesses all the properties we have proved so far for Φ . So, there exists an invertible matrix W_T , such that $W_T\Psi_T(D_0)W_T^{-1} \in [D_0]$. Moreover, $\Phi(N_1)$, which equals either $\Psi_T(E_{12})$ or $\Psi_T(E_{21})$, is mapped into some $\lambda_1 W_T^{-1} E_{st} W_T$, $s \neq t$, $\lambda_1 \neq 0$. Likewise, $\Phi(N_2) = \lambda_2 W_T^{-1} E_{uv} W_T$, $u \neq v$,

$\lambda_2 \neq 0$. Notice that neither of $E_{12}, \mu E_{12}$, nor E_{12}, E_{13} , nor $E_{12}^{\text{tr}}, E_{13}^{\text{tr}}$ quasi-commutes with a minimal singular diagonal matrix $\tilde{D} = \text{diag}(0, 2, 3, \dots, p_{n-1})$, p_i being the i -th prime. Also, notice that $W_T \Psi_T(\cdot) W_T^{-1}$ maps singular diagonal matrices into themselves. Hence, $W_T \Psi_T(\tilde{D}_0) W_T^{-1}$ is a singular minimal diagonal matrix $\text{diag}(d_1, \dots, d_n)$. In particular, it must have precisely one zero diagonal entry, say $d_{i_0} = 0$ (see Lemma 2.17). It follows that neither $\lambda_1 E_{st}$ nor $\lambda_2 E_{uv}$ quasi-commutes with $\text{diag}(d_1, \dots, d_n)$, hence either $s = i_0$ or $t = i_0$, and either $u = i_0$ or $v = i_0$. If $s = v$ or $u = t$, then E_{st} and E_{uv} do not quasi-commute, but $E_{12} \overset{\#}{\leftrightarrow} \mu E_{12}$, $E_{12} \overset{\#}{\leftrightarrow} E_{13}$, and $E_{12}^{\text{tr}} \overset{\#}{\leftrightarrow} E_{13}^{\text{tr}}$, a contradiction. In the remaining two cases $s = u$ or $t = v$, we obtain $\Phi(N_1) + \Phi(N_2)$ is a rank-one nilpotent. We apply the map Φ^{-1} to show also the other direction.

Since the bijective map Φ has the property that N_1, N_2 and $(N_1 + N_2)$ are all rank-one nilpotents precisely when the same holds for their Φ -images, we can apply the result by Du, Hou, and Bai [8, Lemma 2.2] which states that there exists an invertible matrix T such that

$$\begin{aligned} \Phi(N) &= \lambda_N T N^\sigma T^{-1} \quad \text{for every rank-one nilpotent } N, \text{ or} \\ \Phi(N) &= \lambda_N T (N^\sigma)^{\text{tr}} T^{-1} \quad \text{for every rank-one nilpotent } N, \end{aligned}$$

where λ_N is a nonzero scalar depending on N and σ is a field automorphism.

If we compose Φ with either similarity, transposition, or the map $X \mapsto X^{\sigma^{-1}}$, the resulting map will still preserve quasi-commutativity in both directions. So we can assume in the sequel that Φ fixes all rank-one nilpotents, modulo a scalar factor.

We proceed by showing that Φ fixes rank-one idempotents and minimal invertible diagonalizable matrices, up to equivalence.

Lemma 2.30 *Φ fixes the equivalence classes of every minimal invertible diagonalizable matrix, and of every rank-one matrix.*

Proof By Lemma 2.15, every minimal, invertible, diagonalizable matrix is of the form $A = SDS^{-1}$, $D \in [D_0]$. We temporarily replace Φ with a bijection $\Psi: X \mapsto S^{-1}\Phi(SXS^{-1})S$, which still preserves quasi-commutativity in both directions and fixes all rank-one nilpotents, modulo a scalar factor.

It is easy to see that the j -th row of any matrix from $E_{ij}^{\#}$ can have a nonzero entry only on the diagonal. Therefore, the intersection $\bigcap_{i \neq j} E_{ij}^{\#}$ is included in the set of all diagonal matrices. Note that $D \overset{\#}{\leftrightarrow} E_{ij}$, so $\Psi(D) \in \bigcap_{i \neq j} \Psi(E_{ij})^{\#} = \bigcap_{i \neq j} E_{ij}^{\#}$. Moreover, the matrix $\Psi(D)$ is also minimal, diagonal, and invertible, and since there is only one such up to equivalence, it follows that $\Psi(D) \in [D_0]$. Indeed,

$$\Phi(A) = S\Psi(D)S^{-1} \in S[D_0]S^{-1} = S[D]S^{-1} = [A].$$

Since Φ fixes equivalence classes of scalar multiples of rank-one nilpotents, it remains to consider scalar multiples of rank-one idempotents. Suppose $P = S(\lambda E_{11})S^{-1}$ is such a matrix. Again, we temporarily replace Φ by the map $\Psi: X \mapsto S^{-1}\Phi(SXS^{-1})S$. By the first part of the proof, $\Phi(SD_0S^{-1}) \in [SD_0S^{-1}]$, so $\Psi(D_0) \in [D_0]$. As we have shown at the beginning of Subsection 2.4, Ψ then preserves the

set of singular diagonal matrices, and so $\Psi(\lambda E_{11}) = \tilde{D}$ with $\tilde{D} = \text{diag}(d_1, \dots, d_n)$ noninvertible.

Next, λE_{11} does not quasi-commute with rank-one nilpotent matrices E_{1i} for $i = 2, \dots, n$. The same then holds for $\Psi(\lambda E_{11}) = \tilde{D}$ and $\Psi(E_{1i}) = \lambda_{E_{1i}} E_{1i}$. Consequently, no nonzero μ can satisfy $d_1 E_{1i} = \mu d_i E_{1i}$, so either $d_1 = 0$ and $d_2, \dots, d_n \neq 0$ or else $d_1 \neq 0$ and $d_2 = 0 = \dots = d_n$.

In the last case we have $\Phi(P) = S\Psi(\lambda E_{11})S^{-1} = S(d_1 \oplus 0_{n-1})S^{-1} \in [P]$, as claimed. In the first case, note that λE_{11} does quasi-commute with the rank-one nilpotent $R = (E_{2n} + \dots + E_{(n-1)n})$. The same holds for their Ψ -images, i.e., $d_2 E_{2n} + \dots + d_{n-1} E_{(n-1)n} = \mu(d_n E_{2n} + \dots + d_n E_{(n-1)n})$. Therefore, $d_2 = \dots = d_{n-1} = \mu d_n = d$. To see that $\mu = 1$, we repeat the arguments on a rank-one nilpotent $R_1 = \sum_{i=2}^n (E_{in} - E_{i2})$ in place of R . Hence, in the first case, $\Phi(P) = S\Psi(\lambda E_{11})S^{-1} = S(0 \oplus d \text{Id}_{n-1})S^{-1}$. But this is a scalar multiple of an idempotent, so its quasi-commutant equals the usual commutant, which equals $P' = P^\#$. So, also in the first case, $\Phi(P) = S(0 \oplus d \text{Id}_{n-1})S^{-1} \in [P]$. ■

Remark 2.31 Let $Y \in [X]$, that is, $X^\# = Y^\#$. If B is any matrix, then $Y \overset{\#}{\leftrightarrow} B$ precisely when $X \overset{\#}{\leftrightarrow} B$. To see this, note that $Z \overset{\#}{\leftrightarrow} B$ if and only if $B \in Z^\#$. This fact will be used to show that Φ fixes the equivalence class of some matrix A as follows. Suppose $\Phi(X) \in [X]$. Then, $A \overset{\#}{\leftrightarrow} X$ precisely when $\Phi(A) \overset{\#}{\leftrightarrow} X$. When there are enough such matrices X , this will give $\Phi(A) \in [A]$.

We use the above idea to show that Φ acts locally polynomially on diagonalizable matrices. First, assume A is diagonal. Then, $B = \Phi(A)$ is also. Given column vectors $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, we form the rank-one $R = \mathbf{x}\mathbf{y}^{\text{tr}}$. Then, R quasi-commutes with A if and only if

$$A\mathbf{x}\mathbf{y}^{\text{tr}} = \mu\mathbf{x}\mathbf{y}^{\text{tr}}A = \mu\mathbf{x}(A^{\text{tr}}\mathbf{y})^{\text{tr}},$$

which is possible if and only if \mathbf{x} is an eigenvector of A , and \mathbf{y} is an eigenvector of A^{tr} . Note that A quasi-commutes with R precisely when $\Phi(A) = B$ quasi-commutes with $\Phi(R) \in [R]$. This, in turn, is equivalent to B quasi-commuting with R . Consequently, both A and B have exactly the same eigenvectors (though not necessarily corresponding to same eigenvalues). That is, if $A = \sum \lambda_i \text{Id}_{n_i}$, with λ_i pairwise distinct, then $\Phi(A) = \sum \tilde{\lambda}_i \text{Id}_{n_i}$, with $\tilde{\lambda}_i$ pairwise distinct. Therefore, $\Phi(A)$ is a polynomial in A .

The above holds also for diagonalizable $\tilde{A} = SDS^{-1}$; just apply the arguments to $\Psi: X \mapsto S^{-1}\Phi(SXS^{-1})S$ while keeping in mind that, in view of Lemma 2.30, Ψ maps D_0 into $[D_0]$, so Ψ preserves the set of diagonal matrices and maps R into $[R]$ for any rank-one R . Therefore, Φ acts locally polynomially on diagonalizable matrices as well as on rank-one nilpotents.

With some extra work we will show that Φ actually fixes equivalence classes of those matrices. Recall from Lemma 2.30 that $\Phi(R) \in [R]$ for every rank-one R .

Lemma 2.32 Φ fixes the equivalence classes of every diagonalizable matrix with at most two nonzero eigenvalues.

Proof Pick any such A . If A is scalar, we have nothing to do (see Corollary 2.11).

Assume next that A has precisely one nonzero eigenvalue λ . Then, there is a similarity S such that $A = S(\lambda \text{Id}_k \oplus 0_{n-k})S^{-1}$ for some $1 \leq k \leq n - 1$. We already know

that $\Phi(A)$ is a polynomial in A . So, $\Phi(A) = p(A) = S(\xi \text{Id}_k \oplus \zeta \text{Id}_{n-k})S^{-1}$. Now, A does not quasi-commute with $R = SE_{1n}S^{-1}$, so $\Phi(A) = S(\xi \text{Id}_k \oplus \zeta \text{Id}_{n-k})S^{-1}$ does not quasi-commute with $\Phi(R) \in [R]$, and hence also with R . This is possible only when $\xi = 0, \zeta \neq 0$ or when $\zeta = 0, \xi \neq 0$. In both cases, however, $A^\# = \Phi(A)^\#$, so indeed $[\Phi(A)] = [A]$.

Assume lastly that $A = S(\lambda_1 \text{Id}_{k_1} \oplus \lambda_2 \text{Id}_{k_2} \oplus 0_{n-k_1-k_2})S^{-1}$ has two nonzero eigenvalues (possibly $k_1 + k_2 = n$; then, the last term is omitted). Again $\Phi(A) = p(A) = S(\mu_1 \text{Id}_{k_1} \oplus \mu_2 \text{Id}_{k_2} \oplus \alpha \text{Id}_{n-k_1-k_2})S^{-1}$, where μ_1, μ_2, α are distinct (see discussion following Remark 2.31). Note that, when $n - k_1 - k_2 \neq 0$, A does not quasi-commute with $SE_{1n}S^{-1}$ nor with $SE_{(k_1+1)n}S^{-1}$. The same holds for $\Phi(A)$. Due to the fact that μ_1, μ_2, α are distinct, this gives $\alpha = 0$. If, however, $n - k_1 - k_2 = 0$, then $A \stackrel{\#}{\leftrightarrow} E_{1n}$, which yet again yields $\mu_1 \mu_2 \neq 0$. So,

$$\Phi(A) = S(\mu_1 \text{Id}_{k_1} \oplus \mu_2 \text{Id}_{k_2} \oplus 0_{n-k_1-k_2})S^{-1}.$$

Consequently, Φ preserves the set

$$\{S(\gamma_1 \text{Id}_{k_1} \oplus \gamma_2 \text{Id}_{k_2} \oplus 0_{n-k_1-k_2})S^{-1} : \gamma_1 \neq \gamma_2, \gamma_1 \gamma_2 \neq 0\},$$

which splits into precisely two equivalence classes:

$$[S(\text{Id}_{k_1} \oplus 2\text{Id}_{k_2} \oplus 0_{n-k_1-k_2})S^{-1}] \preceq [S(-\text{Id}_{k_1} \oplus \text{Id}_{k_2} \oplus 0_{n-k_1-k_2})S^{-1}].$$

However, Φ induces a bijection on equivalence classes, so this relation is also preserved by Φ . Indeed, both classes are fixed by Φ . ■

Lemma 2.33 *Given any matrix units E_{ij}, E_{uv} , we have $\Phi(E_{ij} + E_{uv}) \in [E_{ij} + \alpha E_{uv}] = [E_{uv} + (1/\alpha)E_{ij}]$ for some nonzero α .*

Proof If $\text{rk}(E_{ij} + E_{uv}) = 1$, we can take $\alpha = 1$, by Lemma 2.30. Assume $\text{rk}(E_{ij} + E_{uv}) = 2$. We can find a permutation matrix S such that the unordered pair $\{E_{ij}, E_{uv}\}$ is simultaneously similar to precisely one of the following: (i) $\{E_{12}, E_{34}\}$, (ii) $\{E_{12}, E_{23}\}$, (iii) $\{E_{11}, E_{23}\}$, (iv) $\{E_{11}, E_{22}\}$, or (v) $\{E_{12}, E_{21}\}$. Actually, we may assume that $\{E_{ij}, E_{uv}\}$ already equals one of the above, for otherwise we might temporarily replace Φ with a map $X \mapsto S\Phi(S^{-1}XS)S^{-1}$, which still fixes the equivalence classes of all rank-one matrices, and D_0 , and preserves quasi-commutativity in both directions.

Now, the last two possibilities fit the assumptions of Lemma 2.32, so their equivalence class is fixed, and we might take $\alpha = 1$.

Next, consider case (i). $A = (E_{12} + E_{34})$ quasi-commutes with every idempotent $X_i = E_{ii}$ for $i \geq 5$ as well as with the idempotent $X = E_{11} + E_{22}$. By Lemma 2.32 and Example 2.4, $B = \Phi(A) \in \Phi(X_i^\#) = X_i^\# = X_i'$ and $B \in X^\# = X' = M_2 \oplus M_{n-2}$. But then the matrix B commutes with idempotents X_i , so its i -th row/column is zero, except possibly at the diagonal (ii) entry. This holds for every $i \geq 5$. Moreover, B also commutes with $X = E_{11} + E_{22}$. Combining these we easily derive that

$$B = B_1 \oplus \text{diag}(\lambda_5, \dots, \lambda_n); \quad B_1 \in M_2 \oplus M_2.$$

In addition, $A = (E_{12} + E_{34})$ also quasi-commutes with diagonal

$$Y = (E_{11} + 2E_{22} + E_{33} + 2E_{44}) \oplus \text{Id}_{n-4},$$

which fits the assumptions of Lemma 2.32. Therefore, $B \overset{\#}{\leftrightarrow} \Phi(Y) \in [Y]$. We derive that either B is diagonal, or $B = (uE_{12} + vE_{34})^{\text{tr}}$, or $B = xE_{12} + yE_{34}$. The first case is contradictory, for the diagonal B would quasi-commute with the idempotent $E_{11} + E_{33}$, but $\Phi^{-1}(B) = A$ does not quasi-commute with $\Phi^{-1}(E_{11} + E_{33}) \in [E_{11} + E_{33}]$. The second case is also contradictory. Namely, $A \overset{\#}{\leftrightarrow} E_{12}$ and $A \overset{\#}{\leftrightarrow} E_{34}$, and so $B \overset{\#}{\leftrightarrow} \Phi(E_{12}) \in [E_{12}]$ as well as $B \overset{\#}{\leftrightarrow} \Phi(E_{34}) \in [E_{34}]$. This is possible only when $u = 0 = v$, so B would be a scalar matrix, and hence so $A = \Phi^{-1}(B)$, a contradiction. Hence, $B = xE_{12} + yE_{34}$.

Finally, $[E_{12}] \neq [E_{12} + E_{34}] \neq [E_{34}]$ implies $[E_{12}] \neq [\Phi(E_{12} + E_{34})] \neq [E_{34}]$ and therefore, $xy \neq 0$ so that $B \in [E_{12} + y/xE_{34}]$, as anticipated.

We proceed with (ii) and show that $E_{12} + E_{23}$ is also mapped into $[E_{12} + \alpha E_{23}]$. To this end, note that $E_{12} + E_{23} \in M_3 \oplus 0_{n-3}$. With the help of idempotents $X_i = E_{ii}$ we may, as above, deduce that

$$B = \Phi(E_{12} + E_{23}) = B_1 \oplus \text{diag}(\lambda_4, \dots, \lambda_n); \quad B_1 \in M_3.$$

Note that $E_{12} + E_{23}$ is elementary Jordan nilpotent. It is easy to see that, of all rank-one matrices inside $M_3 \oplus 0$, it quasi-commutes with precisely one equivalence class, namely with $[E_{13}] = [\mathbf{e}_1 \mathbf{e}_3^{\text{tr}}]$. The same must hold for B . This gives rather easily that \mathbf{e}_1 is the only eigenvector of B_1 , and \mathbf{e}_3 the only eigenvector of B_1^{tr} . Hence, B_1 also has only one eigenvalue, say λ . Now, from $B_1 \mathbf{e}_1 = \lambda \mathbf{e}_1$ and $B_1^{\text{tr}} \mathbf{e}_3 = \lambda \mathbf{e}_3$, we easily derive

$$B_1 = \lambda \text{Id}_3 + \begin{pmatrix} 0 & x_{12} & x_{13} \\ 0 & x_{22} & x_{23} \\ 0 & 0 & 0 \end{pmatrix}.$$

Also, $x_{22} = 0$, otherwise we would have an additional eigenvalue.

Note that $(E_{12} + E_{23})$ quasi-commutes also with the matrix

$$D = (E_{11} - E_{22} + E_{33}) \oplus \text{Id}_{n-3},$$

which fits the assumptions of Lemma 2.32. Thus, $B \in D^{\#}$, wherefrom, after an easy calculation, $B = x_{12}E_{12} + x_{23}E_{23}$, or $B = (x_{13}E_{13} + \lambda \sum_{i=1}^3 E_{ii}) + \sum_{i=4}^n \lambda_i E_{ii}$ or $B = 0$. The latter two cases would wrongly imply that $B_1 = (x_{13}E_{13} + \lambda \sum_{i=1}^3 E_{ii})$ will quasi-commute with at least two nonequivalent rank-one matrices: E_{22} as well as E_{13} . Hence, $B = x_{12}E_{12} + x_{23}E_{23}$. Since $[E_{12}] \neq [E_{12} + E_{23}] \neq [E_{23}]$ implies $[E_{12}] \neq [\Phi(E_{12} + E_{23})] \neq [E_{23}]$, we have $x_{12}x_{23} \neq 0$, as anticipated.

Consider lastly (iii). As in case (ii), we deduce

$$B = \Phi(E_{11} + E_{23}) = B_1 \oplus \text{diag}(\lambda_4, \dots, \lambda_n); \quad B_1 \in M_3.$$

Moreover, $E_{11} + E_{23}$ quasi-commutes with the idempotent $X = E_{22} + E_{33}$. Hence, $B \overset{\#}{\leftrightarrow} \Phi(X) \in [X]$, which gives $B_1 \in \mathbb{C} \oplus M_2$. We next follow the arguments from

case (ii). Of all rank-one matrices inside $0 \oplus M_2 \oplus 0$, $E_{11} + E_{23}$ quasi-commutes with precisely one equivalence class, i.e., $[E_{23}] = [\mathbf{e}_2 \mathbf{e}_3^t]$. The same holds for B , which gives $B = xE_{11} + (\lambda E_{22} + yE_{23} + \lambda E_{33}) + \sum_{i=4}^n \lambda_i E_{ii}$. Further, $E_{11} + E_{23}$ quasi-commutes with $D = -E_{22} + \sum_{i=3}^n E_{ii}$, which fits the assumptions of Lemma 2.32. This gives that either B is diagonal, or $B = xE_{11} + yE_{23}$. The former case is contradictory, for then B would quasi-commute with at least two rank-one matrices $E_{22}, E_{23} \in 0 \oplus M_2 \oplus 0$. Only the last case is possible, and we must also have $xy \neq 0$ due to $[E_{11}] \neq [E_{11} + E_{23}] \neq [E_{23}]$, for example, $E_{32} \in E_{11}^\# \setminus (E_{11} + E_{23})^\#$, and $E_{21} \in E_{23}^\# \setminus (E_{11} + E_{23})^\#$. ■

Lemma 2.34 Φ fixes equivalence classes of every diagonalizable matrix.

Proof It suffices to prove this for diagonal matrices. Indeed, once this is established, we can choose any $A = SDS^{-1}$, where D is diagonal, and temporarily replace Φ with a map $\Psi: X \mapsto S^{-1}\Phi(SXS^{-1})S$, which still preserves quasi-commutativity, fixes every rank-one matrix, and acts bijectively on equivalence classes. So all the above claims are available for Ψ and, consequently, Ψ must fix equivalence classes of diagonal matrix D . Equivalently, Φ must fix $[A]$.

By the same reasoning, it suffices to assume $D = \text{diag}(d_1, \dots, d_k) \oplus 0_{n-k}$ with d_1, \dots, d_k nonzero, and $1 \leq k \leq n$ (when $k = n$, D is invertible, and the last term is omitted). Now, if D has at most two nonzero eigenvalues, the result follows from Lemma 2.32. So we only need to consider the case when D has at least three nonzero eigenvalues. Furthermore, as in the proof of Lemma 2.32, we can then see that

$$B = \Phi(D) = \text{diag}(\beta_1, \dots, \beta_k) \oplus 0_{n-k},$$

with β_i nonzero.

Now, assume $d_i/d_j = d_u/d_v$. Then, by Examples 2.1–2.3, $E_{ij} + E_{uv}$ is in $D^\#$, hence $\Phi(E_{ij} + E_{uv}) \in \Phi(D^\#) = B^\#$. But $\Phi(E_{ij} + E_{uv}) \in [E_{ij} + \alpha E_{uv}]$. Therefore, $E_{ij} + \alpha E_{uv} \in B^\#$ and, since α is nonzero, we derive $\beta_i/\beta_j = \beta_u/\beta_v$. We may reverse implications, since quasi-commutativity is preserved in both directions. If $\beta_i/\beta_j = \beta_u/\beta_v$, then $X \in B^\#$ for every member $X \in [E_{ij} + \alpha E_{uv}]$ and each nonzero α . We know $\Phi(E_{ij} + E_{uv})$ is one such a member for certain α , so $E_{ij} + E_{uv} \in D^\#$, and so $d_i/d_j = d_u/d_v$. That is, $d_i/d_j = d_u/d_v$ precisely when $\beta_i/\beta_j = \beta_u/\beta_v$.

In view of Examples 2.1–2.3, this gives $D^\# = B^\#$, so $B = \Phi(D) \in [D]$. ■

We next show that every nilpotent is fixed, modulo equivalence.

Lemma 2.35 Let $\Omega \subseteq M_n$ be the subset of all matrices, equivalent to rank-one or to diagonalizable ones, and let $N \in M_n$ be a nonzero nilpotent. Then, the following claims are equivalent for $B \in M_n$:

- (i) $N^\# \cap \Omega = B^\# \cap \Omega$;
- (ii) $B = \lambda N$ for some nonzero λ .

Proof Let us prove only the nontrivial implication, (i) implies (ii). Without loss of generality we may assume $N = J_{n_1}(0) \oplus \dots \oplus J_{n_k}(0)$ is already in its Jordan form, with decreasing size of Jordan cells $J_{n_i}(0)$, where $J_{n_i}(0) = \sum_{j=1}^{n_i-1} E_{j(j+1)}$ and $J_1(0)$ is a zero matrix of dimension 1×1 .

Now, a diagonal matrix $D = \text{diag}(2, 2^2, \dots, 2^n)$ quasi-commutes with N (namely, $D^\#$ contains every super-diagonal matrix). So, it also quasi-commutes with B . This gives that $B \in D^\#$, so B is either diagonal or equals some matrix with nonzero entries only on one of sub-diagonals or on one of the super-diagonals. That is, $B = \sum_{i=1}^{n-r_0} \alpha_i E_{(i+r_0)i}$ or $B = \sum_{i=1}^{n-r_0} \alpha_i E_{i(i+r_0)}$ for some $0 \leq r_0 \leq n - 1$ and some scalars α_i . Note that B cannot be diagonal (and in particular, $B \neq 0$), for otherwise, E_{11} would quasi-commute with B , which is not the case with E_{11} and N .

Actually, B lies on the first super-diagonal. To see this, notice first that the idempotents $P_i = 0_{n_1} \oplus \dots \oplus 0_{n_{i-1}} \oplus \text{Id}_{n_i} \oplus 0_{n_{i+1}} \oplus \dots \oplus 0_{n_k}$ quasi-commute with N , so also with B . Hence,

$$(2.4) \quad B \in \bigcap_i P_i^\# = \bigcap_i P_i' = M_{n_1} \oplus \dots \oplus M_{n_k}.$$

In particular, since $n_1 \geq n_2 \geq \dots$, this implies that a nonzero B cannot occupy the r 'th sub/super diagonal when $r \geq n_1$. Hence, $0 \leq r_0 \leq (n_1 - 1)$. Also, $r_0 = 0$ is ruled out since B is not diagonal. Assume erroneously that $r_0 \geq 2$. Then, we would consider the idempotent

$$Q_{r_0} = \sum_{i=0}^{\lfloor (n_1-1)/r_0 \rfloor} E_{(1+ir_0)(1+ir_0)} = E_{11} + E_{(1+r_0)(1+r_0)} + \dots + E_{(1+sr_0)(1+sr_0)}$$

with $(n_1 - r_0) < 1 + sr_0 \leq n_1$, which does not quasi-commute with N , but Q_{r_0} does so with B , a contradiction. Therefore, B lies either on the first sub-diagonal or on the first super-diagonal. Now, since E_{n_1} and N do not quasi-commute, the same is true for B and E_{n_1} . However, any sub-diagonal B would quasi-commute with E_{n_1} , a contradiction. So, $B = \sum_{i=1}^{n-1} \alpha_i E_{i(i+1)}$ is a super-diagonal matrix. From equation (2.4) we further deduce that $\alpha_{n_1} = 0 = \alpha_{n_1+n_2} = \dots = \alpha_{n_1+\dots+n_k}$ because $B = B_1 \oplus \dots \oplus B_k$ is block-diagonal with blocks of the same sizes as in $N = J_{n_1}(0) \oplus \dots \oplus J_{n_k}(0)$. In particular, if $n_j = 1$ for some j then $B_j = 0$.

Consider now a block-diagonal matrix A , with all blocks zero except the j -th one. Suppose $n_j > 1$ and define the j -th block with $A_j = \sum_{i=1}^{n_j} 2^{i-1} E_{ii} + \sum_{i=1}^{n_j-1} 2^i E_{i(i+1)}$. Clearly, A is diagonalizable and quasi-commutes with N , so it also quasi-commutes with B . That is, A_j quasi-commutes with B_j . But we already know B_j can be nonzero only on the first super-diagonal, so this implies that $B_j = \lambda_j J_{n_j}(0)$. Since A_j does not quasi-commute with E_{11} , the same holds for B_j , so $\lambda_j \neq 0$. Therefore,

$$B = \lambda_1 J_{n_1}(0) \oplus \dots \oplus \lambda_k J_{n_k}(0).$$

It only remains to show that all λ_j are equal. Clearly, we need to consider only the blocks of dimensions $n_j \geq 2$. For simplicity, we will consider only the first two blocks and assume, therefore, that $n_1 \geq n_2 \geq 2$. Now, the matrix $X = E_{(n_1+1)(n_1-1)} + 2E_{(n_1+2)n_1} + \sum_{i=1}^n 2^{i-1} E_{ii}$ quasi-commutes with N . Since it is diagonalizable, it must quasi-commute also with B . This easily gives the desired $\lambda_1 = \lambda_2$. ■

Recall that every $X \in \Omega$ is mapped into $\Phi(X) \in [X] \subset \Omega$. We can thus apply Lemma 2.35 on nilpotent N and matrix $B = \Phi(N)$. As a consequence, $B = \Phi(N) = (\lambda N) \in [N]$, i.e., Φ fixes equivalence class of every nilpotent matrix.

Assume now that $A = S(\bigoplus_{i=1}^k J_{n_i}(\alpha_i))S^{-1}$ is an arbitrary matrix, where $J_{n_i}(\alpha_i)$, as usual, denotes the Jordan cell of dimension n_i with α_i on its diagonal. If A is nilpotent or diagonalizable, we already know $[A]$ is fixed. Otherwise, we can further permute its blocks to achieve that the first $r \geq 1$ blocks are invertible, and the remaining $k - r$, if any, are nilpotent. Moreover, we can assume that $S = \text{Id}$, that is $A = A_1 \oplus \dots \oplus A_k$, $A_i = J_{n_i}(\alpha_i)$, is block diagonal, otherwise we would regard a $\Psi(\cdot) = S^{-1}\Phi(S \cdot S^{-1})S$. We also collect the first r invertible blocks into \hat{A} and the last $n - r$ into a nilpotent $N(0)$, thus $A = \hat{A} \oplus N(0)$. If $n - r = 0$, we have $A = \hat{A}$ and omit $N(0)$.

Let $B = \Phi(A)$. Notice first that the idempotents $P_i = 0_{n_1} \oplus \dots \oplus 0_{n_{i-1}} \oplus \text{Id}_{n_i} \oplus 0_{n_{i+1}} \oplus \dots \oplus 0_{n_k}$ quasi-commute with A . So, $\Phi(P_i) \in [P_i]$ also quasi-commutes with $\Phi(A) = B$ for every i . As in equation (2.4), we deduce that $B = B_1 \oplus \dots \oplus B_k$ is block-diagonal, with $B_i \in M_{n_i}$. Notice that $A \overset{\#}{\leftrightarrow} (0_{n_1} \oplus \dots \oplus X_i \oplus \dots \oplus 0_{n_k})$ precisely when $A_i \overset{\#}{\leftrightarrow} X_i$. With this in mind we can temporarily regard only one block or one group of blocks. We proceed in several steps.

Step 1 Assume $A = \hat{A} \oplus N(0)$ has a nilpotent part; that is, $N(0)$ is a matrix of dimension $m \geq 1$. When $N(0) \neq 0$, the same arguments as in the proof of Lemma 2.35 give $B_{r+1} \oplus \dots \oplus B_k = \lambda N(0)$ for some nonzero λ . When $N(0) = 0$, then $B_{r+1} \oplus \dots \oplus B_k$ is a scalar, for otherwise A quasi-commutes with the rank-one $X = 0_{n-m} \oplus (E_{11}^{m \times m} + \dots + E_{1m}^{m \times m})$, $m = n_{r+1} + \dots + n_k$, but $B = \Phi(A)$ would not. Actually, this scalar is 0. Assume otherwise. Since A is not diagonalizable, the same holds for B by Lemma 2.34. So, at least one block, say B_1 , is nonzero and of dimension $n_1 \geq 2$. By Lemma 2.35, B_1 is not nilpotent, since the corresponding A_1 is invertible. Find a similarity $S = S_1 \oplus \text{Id}_{n-n_1}$ such that the matrix $S_1^{-1}B_1S$ is in Jordan form with $(1, 1)$ -element nonzero. It follows that a rank-one matrix $SE_{1n}S^{-1}$ quasi-commutes with B , but it does not quasi-commute with $\Phi^{-1}(B) = A$. In this case $B_{r+1} \oplus \dots \oplus B_k = \lambda N(0)$, and we can take any nonzero λ .

Step 2 Consider any invertible cell A_i of A . Then we claim that the corresponding cell B_i of $B = \Phi(A)$ satisfies $B_i \in [A_i]$. In particular, $B_i \in \text{Poly}(A_i)$ is invertible upper-Toeplitz.

Say for simplicity that $A_i = A_1 = J_{n_1}(\alpha_1)$. Notice that $J_{n_1}(0) \overset{\#}{\leftrightarrow} A_1$, so also $J_{n_1}(0) \overset{\#}{\leftrightarrow} B_1$. It is elementary, then, that

$$(2.5) \quad B_1 = \sum_{i=1}^{n_1} \beta_0 \mu^{i-1} E_{ii} + \sum_{i=1}^{n_1-1} \beta_1 \mu^i E_{i(i+1)} + \dots + \beta_{n_1-1} \mu^{n_1-1} E_{1n_1},$$

for some β_i and some nonzero μ . We claim B_1 is invertible, that is, $\beta_0 \neq 0$. To see this, assume first that the dimension of A_1 , i.e., $n_1 \geq 2$. We then argue with contradiction. If B_1 would be nilpotent, it would quasi-commute with a nonscalar $n_1 \times n_1$ diagonalizable matrix (use the Jordan form of B_1). However, this is not the case with $A_1 = \alpha_1 \text{Id}_{n_1} + J_{n_1}(0)$ because $\alpha_1 \neq 0$ forces $A_1^\# = A_1'$ (see Lemma 2.5) and $A_1' = \text{Poly}(A_1)$. On the other hand, assume A_1 is a 1×1 block. Then, B_1 is also of the same dimension. We next consider two cases separately. First, if $A = \hat{A} \oplus N(0)$ has a nilpotent part, then $E_{1n} \notin A^\#$. Recall from Step 1 that $B = (B_1 \oplus \dots \oplus B_r) \oplus (\lambda N(0))$.

But then, $E_{1n} \notin B^\#$ implies $B_1 \neq 0$. Second, if A is invertible and diagonal, we have nothing left to do (see Lemma 2.34). Otherwise, at least one block, say A_2 , is of dimension greater than one. We have already proven that the corresponding block B_2 is also upper-triangular and invertible. Notice that $E_{1(1+n_2)} \in A^\#$, and since nilpotents are fixed, also $E_{1(1+n_2)} \in B^\#$. But this is possible only if $B_1 \neq 0$, as anticipated.

Next we show B_1 is upper-Toeplitz with $\beta_1 \neq 0$. If $n_1 \geq 3$, then $A_1 \overset{\#}{\leftrightarrow} (J_{n_1}(0) + J_{n_1}(0)^2)$, therefore also $B_1 \overset{\#}{\leftrightarrow} (J_{n_1}(0) + J_{n_1}(0)^2)$, that is $B_1(J_{n_1}(0) + J_{n_1}(0)^2) = \xi(J_{n_1}(0) + J_{n_1}(0)^2)B_1$ for some nonzero scalar ξ . Comparing the first row of the products on both sides, we obtain equations $\beta_0 = \xi\beta_0\mu$ and $\beta_0 + \beta_1\mu = \xi(\beta_0\mu^2 + \beta_1\mu^2)$. Since $\beta_0 \neq 0$, it follows from the first equation that $\xi = \frac{1}{\mu} \neq 0$, and then from the second equation, that $\mu = 1$. If $n_1 = 2$, then also $\mu = 1$, for otherwise A_1 would quasi-commute with rank-one matrix $E_{11} + \frac{\beta_1}{\mu\beta_0 - \mu^2\beta_0}E_{12}$, but B_1 would not. The case $n_1 = 1$ is trivial. It now follows from equation (2.5) that B_1 is upper-Toeplitz, hence a polynomial in A_1 . Moreover, $\beta_1 \neq 0$, for otherwise B_1 would quasi-commute with nilpotent $E_{1,n_1-1} + 2E_{2,n_1}$, but A_1 does not. Since Φ fixes nilpotents, up to a scalar multiple, we have a contradiction. So $\beta_1 \neq 0$.

It remains to show that B_1 and A_1 are equivalent. If their dimension equals 1, there is nothing to do. Otherwise, note that the spectrum, $\text{Sp} A_1 = \{\alpha_1\} \notin \{0\}$, implies $A_1^\# = A_1' = \text{Poly}(A_1)$ (see Lemma 2.5). With the same argument, we also deduce that $B_1^\# = B_1' = \text{Poly}(B_1)$. Obviously, $\text{Poly}(A_1) = \text{Poly}(J_{n_1}(0))$ is an algebra of dimension n_1 . Since its generator $J_{n_1}(0)$ commutes with B_1 , we actually have

$$\text{Poly}(A_1) = \text{Poly}(J_{n_1}(0)) \subseteq B_1' = (B_1 - \beta_0\text{Id}_{n_1})'$$

But $\beta_1 \neq 0$ implies $(B_1 - \beta_0\text{Id}_{n_1})$ is a nilpotent of maximal nil-index, so $\dim B_1' = n_1$. This forces equality in $A' = \text{Poly}(A) \subseteq B'$. Therefore, $B_1 \in [A_1]$.

Similar arguments show that each invertible cell A_i is equivalent to the corresponding cell B_i .

Step 3 We now consider two invertible cells simultaneously, $A_i = J_{n_i}(\alpha_i)$ and $A_j = J_{n_j}(\alpha_j)$. We have just shown that the corresponding cells in B , B_i, B_j are invertible, upper-Toeplitz. Hence, $B_i = \sum_{j=0}^{n_i-1} \beta_{i,j}J_{n_i}(0)^j$, and likewise for B_j . We claim now that $\alpha_i = \alpha_j$ implies $\beta_{i,0} = \beta_{j,0}$.

Assume first that $A_1 = J_{n_1}(\alpha_1)$ and $A_2 = J_{n_2}(\alpha_2)$, where $\alpha_1\alpha_2 \neq 0$. Then $N_1 = E_{1n_1} + E_{1(n_1+n_2)}$ is of rank-one, and A quasi-commutes with N_1 if and only if $\alpha_1 = \alpha_2$. Since rank-one matrices are fixed by Φ , we obtain the same conditions for B . Hence if $\alpha_1 = \alpha_2$, then $\beta_{1,0} = \beta_{2,0}$. The same arguments and the same results are valid for arbitrary invertible blocks in place of A_1, A_2 .

Step 4 We now describe the quasi-commutant of $A = \hat{A} \oplus N(0)$. The following notation will prove handy. Given a nonzero μ , let $A^{\#\mu} = \{X : AX = \mu XA\}$. Also, given a subset $\Gamma \subset \mathbb{C} \setminus \{0\}$, let $\frac{\Gamma}{\Gamma} = \{\gamma_1/\gamma_2 : \gamma_1/\gamma_2 \in \Gamma\}$. We claim that

$$(2.6) \quad A^\# = (0 \oplus N(0)^\#) \cup \bigcup_{\mu \in \frac{\text{Sp } \hat{A}}{\text{Sp } \hat{A}}} \hat{A}^{\#\mu} \oplus N(0)^{\#\mu},$$

and that a block matrix $(X_{ij})_{i,j}$ is an element of $\hat{A}^{\# \mu}$, with $\mu = \alpha_{i_0}/\alpha_{j_0}$, precisely when the following condition is met: $X_{ij} = 0$ for $\alpha_i/\alpha_j \neq \mu$, and $J_{n_i}(\alpha_i)X_{ij} = \alpha_i/\alpha_j X_{ij} J_{n_j}(\alpha_j)$, when $\alpha_i/\alpha_j = \mu$.

To show equation (2.6), pick any $X = \begin{pmatrix} X_1 & U \\ V & X_2 \end{pmatrix} \in A^\#$. Then, there is a nonzero μ with $AX = \mu XA$. Comparing the blocks gives four equations:

$$\begin{aligned} \hat{A}X_1 - \mu X_1 \hat{A} &= 0, & \hat{A}U - \mu U N(0) &= 0, \\ N(0)V - \mu V \hat{A} &= 0, & N(0)X_2 - \mu X_2 N(0) &= 0. \end{aligned}$$

Note that in the second equation, $\text{Sp} \hat{A} - \mu \text{Sp} N(0) = \text{Sp} \hat{A} \not\cong 0$. Hence, by Lemma 2.5, its only solution is $U = 0$. Likewise we see $V = 0$. Furthermore, unless $0 \in \text{Sp} \hat{A} - \mu \text{Sp} \hat{A}$, we have $X_1 = 0$, and therefore $X = 0 \oplus X_2$ with $X_2 \in N(0)^\#$, which gives $X \in 0 \oplus N(0)^\#$. However, if $0 \in \text{Sp} \hat{A} - \mu \text{Sp} \hat{A}$ or equivalently $\mu \in \text{Sp} \hat{A}/\text{Sp} \hat{A}$, then $X_1 \in \hat{A}^{\# \mu}$, while $X_2 \in N(0)^{\# \mu}$, giving $X \in (\hat{A}^{\# \mu} \oplus N(0)^{\# \mu})$.

Next, let us determine $\hat{A}^\#$. It will turn out that the result is similar to Example 2.1. Start with any block-matrix $X = (X_{ij})_{i,j} \in \hat{A}^\# = (\bigoplus_{i=1}^r J_{n_i}(\alpha_i))^\#$. So, there is some nonzero μ such that $\hat{A}X = \mu X \hat{A}$. Considering only the (ij) -th block gives $J_{n_i}(\alpha_i)X_{ij} = \mu X_{ij} J_{n_j}(\alpha_j)$. If $\alpha_i \neq \mu \alpha_j$, Lemma 2.5 implies that $X_{ij} = 0$. However, when $\alpha_i = \mu \alpha_j$, we easily solve this equation and see that the $n_i \times n_j$ matrix X_{ij} has a similar structure to the square matrix in (2.5): it is upper-triangular, equal to

$$(2.7) \quad X_{ij} = \left(x_1 \sum_{s=1}^{m_0} \mu^{s-1} E_{s(n_j-m_0+s)}^{n_i \times n_j} \right) + \dots + x_{m_0} \mu^{m_0-1} E_{1n_j}^{n_i \times n_j},$$

where $m_0 = \min\{n_i, n_j\}$. We define $\Lambda^{(ij)} \subseteq \hat{A}^\#$ to be the set of all block matrices $(X_{st})_{st}$ with all the blocks equal to zero, except for the (ij) -th block, which is of the same form as in equation (2.7) for $\mu = \alpha_i/\alpha_j$. In particular, $E^{(ij)} \in \Lambda^{(ij)}$, where $E^{(ij)}$ has only one nonzero entry, which equals 1, and it lies at the upper-right corner of the (ij) -th block. Notice that $\text{rk} E^{(ij)} = 1$. Notice also that $\Lambda^{(ii)} = 0 \oplus \dots \oplus A_i^\# \oplus 0 \oplus \dots \oplus 0$, and corresponds to $\mu = 1$.

As usual, given two subsets $\Omega_1, \Omega_2 \in M_n$, we let

$$\Omega_1 + \Omega_2 = \{X + Y : X \in \Omega_1, Y \in \Omega_2\}.$$

It is now easy to deduce the following for $\hat{A}^\#$:

$$(2.8) \quad \hat{A}^\# = \bigcup_{\mu \in \frac{\text{Sp} \hat{A}}{\text{Sp} \hat{A}}} \left(\sum_{(i,j) \in \{(i,j): \frac{\alpha_i}{\alpha_j} = \mu\}} \Lambda^{(ij)} \right).$$

Step 5 This can also be rewritten as follows. First, according to the block-structure of \hat{A} , which is the same as that for \hat{B} , we introduce $\mathcal{B}^{(ij)}$ to denote the set of all those block-matrices of same dimension as \hat{A} or \hat{B} that have all blocks but the (ij) -th one equal to zero. It now follows from equation (2.8) that $\Lambda^{(ij)} = \hat{A}^\# \cap \mathcal{B}^{(ij)}$. Moreover,

we introduce the equivalence relation on ordered pairs from $\mathbb{N}_r \times \mathbb{N}_r$ by $(i, j) \sim (u, v)$ if and only if $\alpha_i/\alpha_j = \alpha_u/\alpha_v$. Then, equation (2.8) can be rewritten as

$$(2.9) \quad \hat{A}^\# = \bigcup_{\mathfrak{p} \in (\mathbb{N}_r \times \mathbb{N}_r)/\sim} \left(\sum_{(i,j) \in \mathfrak{p}} (\hat{A}^\# \cap \mathcal{B}^{(ij)}) \right).$$

Step 6 Recall from Steps 1 and 2 that B also decomposes into $B = \hat{B} \oplus (\lambda N(0))$, where each cell B_i in \hat{B} is invertible, upper-Toeplitz, with $\beta_{i,0}$ on the main diagonal, and equivalent to the corresponding cell in \hat{A} . We claim here that $\alpha_i/\alpha_j = \alpha_u/\alpha_v$ precisely when $\beta_{i,0}/\beta_{j,0} = \beta_{u,0}/\beta_{v,0}$.

Namely, it follows from equation (2.8) that $\alpha_i/\alpha_j = \alpha_u/\alpha_v$ precisely when $E^{(ij)} + E^{(uv)} \in \hat{A}^\#$. In view of equation (2.6), this is further equivalent to $(E^{(ij)} \oplus 0) + (E^{(uv)} \oplus 0) \in A^\#$. Notice that $(E^{(ij)} \oplus 0), (E^{(uv)} \oplus 0)$ are both matrix units. By Lemma 2.33, $\Phi((E^{(ij)} + E^{(uv)}) \oplus 0) \in [(E^{(ij)} + \alpha E^{(uv)}) \oplus 0]$ for some $\alpha \neq 0$. Therefore, also $(E^{(ij)} + \alpha E^{(uv)}) \oplus 0 \in B^\#$. But then, $E^{(ij)} + \alpha E^{(uv)} \in \hat{B}^\#$. A simple calculation shows that $\beta_{i,0}/\beta_{j,0} = \beta_{u,0}/\beta_{v,0}$.

The reversed implication follows since Lemma 2.33 is valid also for Φ^{-1} because it also fixes equivalence classes of diagonalizable and nilpotent matrices.

Step 7 We now consider all invertible cells combined and claim that $\hat{A}^\# = \hat{B}^\#$. The same arguments we used to prove equation (2.6) are valid also for $B = \hat{B} \oplus (\lambda N(0))$, so

$$(2.10) \quad B^\# = (0 \oplus (\lambda N(0))^\#) \cup \bigcup_{\mu \in \frac{\text{Sp } \hat{B}}{\text{Sp } B}} \hat{B}^{\#\mu} \oplus (\lambda N(0))^\#.$$

By Lemma 2.5,

$$\hat{B}^\# \oplus 0 = \bigcup_{\mu \in (\text{Sp } \hat{B}/\text{Sp } B)} \hat{B}^{\#\mu} \oplus 0,$$

hence $\hat{B}^\# \oplus 0 \subseteq B^\#$; similarly, $\hat{A}^\# \oplus 0 \subseteq A^\#$. Next, let $\mathcal{B}^{(ij)}$ be as in Step 5. When $N(0)$ is present, we can augment matrices from $\mathcal{B}^{(ij)}$ with a zero tail, so that $\mathcal{B}^{(ij)} \oplus 0 \subseteq M_n$. Now, if $i \neq j$, then $\mathcal{B}^{(ij)} \oplus 0$ consists of nilpotents only, and their equivalence classes are held fixed by Φ and by Φ^{-1} . Therefore, given $i \neq j$, we have $A^\# \cap (\mathcal{B}^{(ij)} \oplus 0) = B^\# \cap (\mathcal{B}^{(ij)} \oplus 0)$. By equations (2.6) and (2.10), this is equivalent to

$$\Lambda^{(ij)} = \hat{A}^\# \cap \mathcal{B}^{(ij)} = \hat{B}^\# \cap \mathcal{B}^{(ij)}$$

(recall that $\Lambda^{(ij)}$ was defined in equation (2.8)). If, however, $i = j$, then $\Lambda^{(ii)} = 0 \oplus \dots \oplus 0 \oplus A_i^\# \oplus 0 \oplus \dots \oplus 0$. But $A_i^\# = B_i^\#$ by Step 2, so yet again, $\Lambda^{(ii)} = \hat{A}^\# \cap \mathcal{B}^{(ii)} = \hat{B}^\# \cap \mathcal{B}^{(ii)}$. A formula similar to (2.9) is valid also for $\hat{B}^\#$, namely

$$\hat{B}^\# = \bigcup_{\mathfrak{p} \in (\mathbb{N}_r \times \mathbb{N}_r)/\approx} \left(\sum_{(i,j) \in \mathfrak{p}} (\hat{B}^\# \cap \mathcal{B}^{(ij)}) \right),$$

with $(i, j) \approx (u, v)$ if and only if $\frac{\beta_{i,0}}{\beta_{j,0}} = \frac{\beta_{u,0}}{\beta_{v,0}}$, an equivalence relation similar to \sim for \hat{A} . But we have already shown that $(\hat{B}^\# \cap \mathcal{B}^{(ii)}) = \Lambda^{(ij)}$, and we know that $\alpha_i/\alpha_j = \alpha_u/\alpha_v$ precisely when $\beta_{i,0}/\beta_{j,0} = \beta_{u,0}/\beta_{v,0}$ by Step 6. Therefore, $(i, j) \sim (u, v)$ precisely when $(i, j) \approx (u, v)$. So,

$$\hat{B}^\# = \bigcup_{\mathfrak{p} \in (\mathbb{N}_r \times \mathbb{N}_r)/\sim} \left(\sum_{(i,j) \in \mathfrak{p}} \Lambda^{(ij)} \right) = \hat{A}^\#.$$

Step 8 It remains to compare $A^\#$ and $B^\#$. They are clearly equal when A is invertible (i.e., when there is no nilpotent part) by Step 7. Assume next the nilpotent part contains a cell of nil-index ≥ 3 . With no loss of generality,

$$A = \hat{A} \oplus J_{n_{r+1}}(0) \oplus \cdots \oplus J_{n_k}(0) \text{ and } n_{r+1} \geq 3.$$

We claim that the quotients of eigenvalues in \hat{A} match those in \hat{B} , that is, $\alpha_i/\alpha_j = \beta_{i,0}/\beta_{j,0}$.

Indeed, if $\alpha_i/\alpha_j = 1$, then $\alpha_i/\alpha_j = \alpha_1/\alpha_1$, and we already know from Step 6 that the same holds true for $\beta_{i,0}/\beta_{j,0} = \beta_{1,0}/\beta_{1,0}$, i.e., $\beta_{i,0}/\beta_{j,0} = 1$. If, however, $\mu = \alpha_i/\alpha_j \neq 1$, then in view of equation (2.6), we know that

$$X_\mu = E^{(ij)} \oplus \left(\sum_s (\alpha_i/\alpha_j)^{s-1} E_{s(s+1)}^{n_{r+1} \times n_{r+1}} \right) \oplus 0 \oplus \cdots \oplus 0 \subseteq A^\#.$$

It now follows from $i \neq j$ that X_μ is nilpotent, hence its equivalence class is held fixed by Φ . So, also $X_\mu \in B^\#$. Then, however, we see from the middle block, i.e., from $\sum_s \mu^{s-1} E_{s(s+1)}^{n_{r+1} \times n_{r+1}}$, that $BX_\mu = \xi X_\mu B$, $\xi = \beta_{i,0}/\beta_{j,0}$, and so $\beta_i/\beta_j = \mu = \alpha_i/\alpha_j$. In particular, $\text{Sp } \hat{B}/\text{Sp } \hat{B} = \text{Sp } \hat{A}/\text{Sp } \hat{A}$.

It is now easy to see that $\hat{A}^{\#\mu} = \hat{B}^{\#\mu}$, so that

$$\begin{aligned} A^\# &= (0 \oplus N(0)^\#) \cup \bigcup_{\mu \in \frac{\text{Sp } \hat{A}}{\text{Sp } \hat{A}}} \hat{A}^{\#\mu} \oplus N(0)^{\#\mu} \\ &= (0 \oplus (\lambda N(0))^\#) \cup \bigcup_{\mu \in \frac{\text{Sp } \hat{A}}{\text{Sp } \hat{A}}} \hat{B}^{\#\mu} \oplus (\lambda N(0))^{\#\mu} \\ &= (0 \oplus (\lambda N(0))^\#) \cup \bigcup_{\mu \in \frac{\text{Sp } \hat{B}}{\text{Sp } \hat{B}}} \hat{B}^{\#\mu} \oplus (\lambda N(0))^{\#\mu} \\ &= B^\#. \end{aligned}$$

This completes the proof of Theorem 1.1. ■

We finish this section by proving Remark 1.2.

Lemma 2.36 *Let Φ satisfy (i)–(ii) in Theorem 1.1. Then for $\Psi(X) = S^{-1}\Phi(X^{\sigma^{-1}})S$, respectively, for $\Psi(X) = S^{-1}\Phi(X^{\sigma^{-1}})^{\text{tr}}S$ we have that $\Psi(A) \in \text{Poly}(A)$.*

Proof We use the same notation as in the proof of Theorem 1.1. Let B_i be an arbitrary invertible cell. By Step 2 we have $B_i \in \text{Poly}(A_i)$. Now, assume $\text{Sp } A_{i_1} = \text{Sp } A_{i_2} = \dots = \text{Sp } A_{i_u}$. Consider any pair among this tuple, say i_1 and i_2 . For simplicity we can take $i_1 = 1, i_2 = 2$; then $\text{Sp } A_1 = \text{Sp } A_2$. Then also $\text{Sp } B_1 = \text{Sp } B_2$, by Step 3. In this case, we take a nilpotent $N = (N_{ij})_{ij}$ with all blocks, but $(1, 2)$ -th, zero and with $N_{12} = \sum_{s=1}^{\min\{n_1, n_2\}} E_{s(\max\{n_1, n_2\} + s)} \in A^\#$. Being nilpotent, it is fixed by Φ , so $N \in B^\#$. Therefore, $BN = \mu NB$ hence $B_1 N_{12} = \mu N_{12} B_2$, for some $\mu \neq 0$. However, $N_{12} \neq 0$ while $\text{Sp } B_1 = \text{Sp } B_2$. Consequently, Lemma 2.5 gives $\mu = 1$. Recall that B_1, B_2 are upper Toeplitz which easily shows that B_1, B_2 have the same elements on the corresponding super-diagonals. This implies that $(B_1 \oplus B_2) \in \text{Poly}(A_1 \oplus A_2)$. But our choice of pair was arbitrary, and it follows that actually $(B_{i_1} \oplus \dots \oplus B_{i_u}) \in \text{Poly}(A_{i_1} \oplus \dots \oplus A_{i_u})$.

Recall from Step 1 that for noninvertible cells we also have

$$(B_{r+1} \oplus \dots \oplus B_k) = \lambda(A_{r+1} \oplus \dots \oplus A_k) \in \text{Poly}(A_{r+1} \oplus \dots \oplus A_k).$$

Now, recall that, given any two matrices X, Y with distinct minimal polynomial, we have $\text{Poly}(X \oplus Y) = \text{Poly}(X) \oplus \text{Poly}(Y)$. This follows since the projections $\text{Id}_k \oplus 0_{n-k}$ and $0_k \oplus \text{Id}_{n-k}$ are polynomials in $X \oplus Y$ [11, p. 221]; see also [9, Theorem 2 and Remark on p. 127] for more on this theme. This gives $B = (B_1 \oplus \dots \oplus B_k) \in \text{Poly}(A_1 \oplus \dots \oplus A_k) = \text{Poly}(A)$. ■

3 Concluding Remark

The map in Theorem 1.1 is described with the help of equivalence classes, defined by $[A] = \{X : X^\# = A^\#\}$. Here we describe the elements inside $[A]$ in more detail.

Remark 3.1 When A is nilpotent, then $[A] = (\mathbb{C} \setminus \{0\})A$, see Lemma 2.35.

When A is invertible, then $B \in [A]$ implies

- (i) $B \in \text{Poly}(A)$,
- (ii) B is invertible, with Jordan cells of the same size as in A , and
- (iii) if the spectrum, $\text{Sp } A = (\alpha_1, \dots, \alpha_k)$, respectively, $\text{Sp } B = (\beta_1, \dots, \beta_k)$ is a k -tuple, arranged according to the Jordan cell structure of A , and of B , respectively, then $\alpha_i/\alpha_j = \alpha_u/\alpha_v$ if and only if $\beta_i/\beta_j = \beta_u/\beta_v$.

See Step 6 of the proof.

When $A = S(\hat{A} \oplus N)S^{-1}$ is similar to a block-diagonal with invertible part \hat{A} and nilpotent N , then $B \in [A]$ implies $B = S(\hat{B} \oplus (\lambda N))S^{-1}$ for some nonzero λ , and for some $\hat{B} \in [\hat{A}]$ with $\alpha_i/\alpha_j = \alpha_u/\alpha_v$ if and only if $\beta_i/\beta_j = \beta_u/\beta_v$. When the nil-index of N is greater than two, we have even more restrictive condition on B : $\alpha_i/\alpha_j = \beta_i/\beta_j$. See Step 8 of the proof.

Acknowledgement The authors are indebted to Peter Šemrl for proposing this kind of research and for his advice which improved the first draft.

References

- [1] L. Baribeau and T. Ransford, *Non-linear spectrum-preserving maps*. Bull. London Math. Soc. **32**(2000), no. 1, 8–14. doi:10.1112/S0024609399006426
- [2] R. Bhatia and P. Rosenthal, *How and why to solve the operator equation $AX - XB = Y$* . Bull. London Math. Soc. **29**(1997), no. 1, 1–21. doi:10.1112/S0024609396001828
- [3] J. A. Brooke, P. Busch, and D. B. Pearson, *Commutativity up to a factor of bounded operators in complex Hilbert space*. R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. **458**(2002), no. 2017, 109–118. doi:10.1098/rspa.2001.0858
- [4] G. Cassinelli, E. De Vito, P. J. Lahti, and A. Levrero, *The theory of symmetry actions in quantum mechanics. With an application to the Galilei group*. Lecture Notes in Physics, 654, Springer-Verlag, Berlin, 2004.
- [5] J.-T. Chan, C.-K. Li, and N.-S. Sze, *Mappings on matrices: invariance of functional values of matrix products*. J. Aust. Math. Soc. **81**(2006), no. 2, 165–184. doi:10.1017/S1446788700015809
- [6] J. Cui, J. Hou, and C. Park, *Additive maps preserving commutativity up to a factor*. Chinese Ann. Math. Ser. A **29**(2008), no. 5, 583–590.
- [7] G. Dolinar, S. Du, J. Hou, and P. Legiša, *General preservers of invariant subspace lattices*. Linear Algebra Appl. **429**(2008), no. 1, 100–109. doi:10.1016/j.laa.2008.02.007
- [8] S. Du, J. Hou, and Z. Bai, *Nonlinear maps preserving similarity on $B(H)$* . Linear Algebra Appl. **422**(2007), no. 2–3, 506–516. doi:10.1016/j.laa.2006.11.008
- [9] F. A. Fillmore, D. A. Herrero, and W. E. Longstaff, *The hyperinvariant subspace lattice of a linear transformation*. Linear Algebra Appl. **17**(1977), no. 2, 125–132. doi:10.1016/0024-3795(77)90032-5
- [10] O. Holtz, V. Mehrmann, and H. Schneider, *Potter, Wielandt, and Drazin on the matrix equation $AB = \omega BA$: new answers to old questions*. Amer. Math. Monthly **111**(2004), no. 8, 655–667. doi:10.2307/4145039
- [11] K. Hoffman and R. Kunze, *Linear algebra*. Second ed., Prentice-Hall, Englewood Cliffs, NJ, 1971.
- [12] R. A. Horn and C. R. Johnson, *Topics in matrix analysis*. Cambridge University Press, Cambridge, 1991.
- [13] L.-K. Hua, *A theorem on matrices over a sfield and its applications*. J. Chinese Math. Soc. (N.S.) **1**(1951), 110–163.
- [14] N. McCoy, *On quasi-commutative matrices*. Trans. Amer. Math. Soc. **36**(1934), no. 2, 327–340. doi:10.2307/1989841
- [15] L. Molnár, *Orthogonality preserving transformations on indefinite inner product spaces: generalization of Uhlhorn's version of Wigner's theorem*. J. Funct. Anal. **194**(2002), no. 2, 248–262. doi:10.1006/jfan.2002.3970
- [16] ———, *Linear maps on matrices preserving commutativity up to a factor*. Linear and Multilinear Algebra **57**(2009), no. 1, 13–18. doi:10.1080/03081080701210211
- [17] L. Molnár and P. Šemrl, *Nonlinear commutativity preserving maps on self-adjoint operators*. Q. J. Math. **56**(2005), no. 4, 589–595. doi:10.1093/qmath/hah058
- [18] J. von Neumann, *Mathematical foundations of quantum mechanics*. Princeton University Press, Princeton, NJ, 1955.
- [19] M. Omladič, H. Radjavi, and P. Šemrl, *Preserving commutativity*. J. Pure Appl. Algebra **156**(2001), no. 2–3, 309–328. doi:10.1016/S0022-4049(99)00154-1
- [20] P. G. Ovchinnikov, *Automorphisms of the poset of skew projections*. J. Funct. Anal. **115**(1993), no. 1, 184–189. doi:10.1006/jfan.1993.1086
- [21] H. S. A. Potter, *On the latent roots of quasi-commutative matrices*. Amer. Math. Monthly **57**(1950), 321–322. doi:10.2307/2306202
- [22] H. Radjavi and P. Šemrl, *Linear maps preserving quasi-commutativity*. Studia Math. **184**(2008), no. 2, 191–204. doi:10.4064/sm184-2-7
- [23] P. Šemrl, *Non-linear commutativity preserving maps*. Acta Sci. Math. (Szeged) **71**(2005), no. 3–4, 781–819.
- [24] ———, *Maps on idempotent matrices over division rings*. J. Algebra **298**(2006), no. 1, 142–187. doi:10.1016/j.jalgebra.2005.08.010
- [25] ———, *Maps on matrix spaces*. Linear Algebra Appl. **413**(2006), no. 2–3, 364–393. doi:10.1016/j.laa.2005.03.011

- [26] ———, *Commutativity preserving maps*. *Linear Algebra Appl.* **429**(2008), no. 5–6, 1051–1070.
doi:10.1016/j.laa.2007.05.006
- [27] W. Watkins, *Linear maps that preserve commuting pairs of matrices*. *Linear Algebra and Appl.* **14**(1976), no. 1, 29–35. doi:10.1016/0024-3795(76)90060-4
- [28] J. H. M. Wedderburn, *Lectures on matrices*. Dover Publications, New York, 1964.

Faculty of Electrical Engineering, University of Ljubljana, Ljubljana, Slovenia
e-mail: gregor.dolinar@fe.uni-lj.si

University of Primorska, Koper, Slovenia

and

Institute of Mathematics, Physics, and Mechanics, Ljubljana, Slovenia
e-mail: bojan.kuzma@pef.upr.si