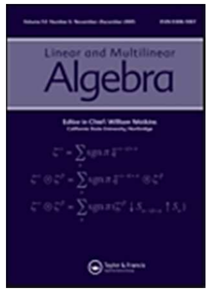

A Spatial Version of Wedderburn's Principal Theorem

Journal:	*Linear and Multilinear Algebra*
Manuscript ID:	GLMA-a-2013-0247.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	18-Apr-2014
Complete List of Authors:	Livshits, Leo; Colby College, Mathematics and Statistics MacDonald, Gordon; UPEI, Mathematics and Statistics Marcoux, Laurent; University of Waterloo, Pure Mathematics Radjavi, Heydar; University of Waterloo, Pure Mathematics
Keywords:	Wedderburn's Principal Theorem, Wedderburn-Artin Theorem, block-upper-triangular matrix algebras, irreducible matrix algebras, semi-simple matrix algebras

This is an Accepted Manuscript of an article published by Taylor & Francis in 'Linear and Multilinear Algebra' on 7/2014, available online: http://www.tandfonline.com/10.1080/03081087.2014.925452.

To appear in *Linear and Multilinear Algebra*
Vol. 00, No. 00, Month 20XX, 1–26

Research Article

A Spatial Version of Wedderburn's Principal Theorem

L. Livshits*, G. W. MacDonald, L. W. Marcoux, H. Radjavi
(*Received 00 Month 20XX; final version received 00 Month 20XX*)

In this article we verify that "Wedderburn's Principal Theorem" has a particularly pleasant spatial implementation in the case of cleft subalgebras of the algebra of all linear transformations on a finite-dimensional vector space.

Once such a subalgebra $\mathcal{A}$ is represented by block-upper-triangular matrices with respect to a maximal chain of its invariant subspaces, after an application of a block-upper-triangular similarity, the resulting algebra is a linear direct sum of an algebra of block-diagonal matrices and an algebra of strictly block-upper-triangular matrices (i.e. the radical), while the block-diagonal matrices involved have a very nice structure.

We apply this result to demonstrate that, when the underlying field is algebraically closed, and $\left(Rad(\mathcal{A})\right)^{\mu(\mathcal{A})-1} \neq \{0\}$, the algebra is unicellular, i.e. the lattice of all invariant subspaces of $\mathcal{A}$ is totally ordered by inclusion. The quantity $\mu(\mathcal{A})$ stands for the length of (every) maximal chain of non-zero invariant subspaces of $\mathcal{A}$.

Keywords: Wedderburn's Principal Theorem, Wedderburn-Artin Theorem, block-upper-triangular matrix algebras, irreducible matrix algebras, semi-simple matrix algebras

AMS Subject Classification: 15A21, 15A30, 15A33, 16D60, 16D70, 16K20, 16N40, 16N60, 16P10, 16S50

1. Introduction

Parts of this article are somewhat expository, since some of the theorems we include in this paper can be considered to be a part of the folklore. The results were essential to our work presented in two papers [1] and [2], and since we did not find these theorems in print we were compelled to verify their validity by supplying the proofs. As far as we know these results have not been published in a form easily accessible to mathematicians working in the fields of linear algebra and matrix theory, even though the theorems provide remarkably useful tools that can be used in both subjects. It is likely that the theorems, expressed in a different language, are known to specialists in representation theory and ring theory.

The goal is to verify that several fundamental theorems of Wedderburn have spatial realizations in the context of algebras of linear transformations over general fields. Subsequently an application of the spatial version of Wedderburn's Principal Theorem is presented. As we have mentioned already, papers [1] and [2] involve further applications of the result.

The first author was supported by the Colby College Natural Science Division Grant. The second, third and fourth authors acknowledge the support of NSERC Canada.
*Corresponding author. Email: llivshi@colby.edu

One of the fundamental structure theorems for algebras over a field, proved by J. H. M. Wedderburn [3, 4] in the case of a field of characteristic zero, extended by L. E. Dickson [5, 6], and often referred to as "Wedderburn's Principal Theorem", states that under very general conditions (for example, if the underlying field is perfect) a finite-dimensional algebra can be decomposed as a linear direct sum of its nil radical and a semi-simple algebra (usually called a "Wedderburn factor"). A theorem of Malcev [7] shows that any two Wedderburn factors of an algebra are "conjugate" in the sense that one can be mapped to the other via (an almost inner) similarity.

A subalgebra $\mathcal{A}$ of the algebra of all linear transformations on a finite-dimensional vector space $\mathcal{V}$ can be represented as an algebra of block-upper-triangular matrices with respect to a maximal chain of invariant subspaces of $\mathcal{A}$. In this form the radical of $\mathcal{A}$ is (represented by) the set of all strictly block-upper-triangular matrices in $\mathcal{A}$. The subalgebra of $\mathcal{A}$ represented by the block-diagonal matrices is semi-simple, but may not be a Wedderburn factor of $\mathcal{A}$. For example, this is the case for

$$\left\{ \begin{bmatrix} a & 0 & c \\ 0 & a & a-b \\ 0 & 0 & b \end{bmatrix} \;\middle|\; a, b, c \in \mathrm{C} \right\},$$

where the radical is the span of the matrix unit E_{13}, and the only diagonal matrices in the algebra are the scalar multiples of the identity matrix.

Still, a well-known theorem of Watters [8] indicates that after an application of a block-diagonal similarity, the block-diagonals of the matrices involved can be assumed to have a particularly nice form. It is worth mentioning that if the hypothesis of Wedderburn's Principal Theorem applies to $\mathcal{A}$, the similarities described by Malcev's theorem are also block-upper-triangular.

In this article we are able to show that, when $\mathcal{A}$ satisfies the hypothesis of Wedderburn's Principal Theorem, it is always possible to apply a block-upper-triangular similarity to $\mathcal{A}$ in such a way that the block-diagonal matrices in the resulting algebra have the structure described in Watter's theorem and comprise a Wedderburn factor (with the radical still being the set of all strictly-block-upper-triangular matrices). In particular, the resulting algebra decomposes as a linear direct sum of an algebra of block-diagonal matrices and an algebra of strictly block-upper-triangular matrices.

Furthermore, we demonstrate that for any Wedderburn factor $\mathcal{S}$ of $\mathcal{A}$ there is a maximal chain of invariant subspaces of $\mathcal{A}$, such that $\mathcal{S}$ is represented by the algebra of the corresponding block-diagonal matrices.

These results are then used to show, for example, that if the underlying field is algebraically closed and $\left(Rad(\mathcal{A})\right)^{\mu(\mathcal{A})-1} \neq \{0\}$, then $\mathcal{A}$ is unicellular, i.e. the lattice of all invariant subspaces of $\mathcal{A}$ is totally ordered by inclusion. The quantity $\mu(\mathcal{A})$ stands for the length of (any) maximal chain of non-zero invariant subspaces of $\mathcal{A}$.

The main impetus behind the research that produced these results was the inquiry into a natural weakening of the concept of transitivity for algebras of linear transformations acting on a finite-dimensional vector space. The classical notion of transitivity is defined for an arbitrary collection $\mathcal{C}$ of linear transformations from a vector space $\mathcal{V}$ to another vector space $\mathcal{W}$ over the same field. In general, $\mathcal{C}$ is said to be transitive if for any non-zero vector x in $\mathcal{V}$ and any $y \in \mathcal{W}$ there is a member T of $\mathcal{C}$ with $Tx = y$. When the underlying field is algebraically closed, a well-known theorem of Burnside dictates that the only transitive algebra of linear transformations on $\mathcal{V}$ is the algebra $\mathcal{L}(\mathcal{V})$ of all linear transformations on $\mathcal{V}$.

The aforementioned weakening of the concept of transitivity for algebras, referred to as paratransitivity, is defined as follows: given positive integers k and m, an algebra $\mathcal{A}$ is said to be (k, m)-transitive if for every pair of subspaces $\mathcal{W}_1$ and $\mathcal{W}_2$ of $\mathcal{V}$, of respective dimensions k and m, the orbit $\mathcal{A}\mathcal{W}_1 := \{Ax : A \in \mathcal{A}, x \in \mathcal{W}_1\}$ meets $\mathcal{W}_2$ non-trivially.

The spatial version of Wedderburn's Principal Theorem (see Corollary 30) and its consequences, that we develop in the present article, are crucial to our method of proof of the results on paratransitivity. We refer the reader to the papers [1] and [2] for further information. We believe that this form of the well-known result will be useful in connection with other problems dealing with spatial structure of sets of operators.

Convention 1. Unless specified otherwise all vector spaces in this article are assumed to be non-zero and finite-dimensional.

As is common, we shall be using the same notation $\mathcal{I}$ to denote both the identity transformation (independent of the underlying vector space) and an identity matrix (or a block-matrix). This will not cause ambiguity since the context shall always indicate which is which.

The same goes for a zero transformation.

Terminology 2. Let $\mathcal{V}$ be a (necessarily finite-dimensional) vector space over a field F, and let $\mathcal{L}_{\mathrm{F}}(\mathcal{V})$ stand for the algebra of all linear transformations on $\mathcal{V}$; (we shall omit the subscript "F" when possible).

A subalgebra of $\mathcal{L}(\mathcal{V})$ is said to be irreducible if it has no non-trivial invariant subspaces. When $\dim(\mathcal{V}) > 1$, subalgebra of $\mathcal{L}(\mathcal{V})$ is irreducible if and only if it is transitive; i.e. if and only if

$$\mathcal{A}x \overset{\text{def}}{=} \{\ Ax \mid\ A \in \mathcal{A}\ \} = \mathcal{V}$$

for every non-zero $x \in \mathcal{V}$. (Transitive algebras are irreducible even if $\dim(\mathcal{V}) = 1$.)

A unital algebra $\mathcal{A}$ is said to be central if its center is the set of all scalar multiples of its multiplicative identity element.

Remark 3. Non-zero ideals of irreducible algebras are irreducible.

The following classical result is presented here for the convenience of the reader. For a stronger version of this result see Theorem 10 below.

Proposition 4. *Every non-zero irreducible subalgebra of $\mathcal{L}_{\mathrm{F}}(\mathcal{V})$ is simple and contains the identity transformation.*

Proof. If $\mathcal{A}$ is a non-zero irreducible (and so, transitive) subalgebra of $\mathcal{L}(\mathcal{V})$ then every non-zero ideal $\mathcal{J}$ of $\mathcal{A}$ is itself an irreducible algebra, and hence has a trivial common kernel.

In particular the (nil) radical $Rad(\mathcal{A})$ is either zero or irreducible. Since $Rad(\mathcal{A})$ is a nilpotent ideal, if it is not zero, it has a common kernel, and thus is not irreducible. Hence it must be that $\mathcal{A}$ is semi-simple.

Since every finite-dimensional semi-simple algebra is unital, there is an idempotent $E \in \mathcal{A}$ such that $E\mathcal{A} = \mathcal{A}$. Yet by transitivity, for a non-zero x,

$$\mathcal{V} = \mathcal{A}x = E\mathcal{A}x \subset range(E) \subset \mathcal{V},$$

so that E is surjective, and therefore $E = \mathcal{I}$.

We have shown that every non-zero irreducible subalgebra of $\mathcal{L}(\mathcal{V})$ contains the identity transformation. Therefore any non-zero ideal of a non-zero irreducible algebra $\mathcal{A}$ contains the identity transformation, and thus equals the whole algebra $\mathcal{A}$. Since $\mathcal{A}$ is unital, this shows that $\mathcal{A}$ is simple and completes the proof. ■

Definition 5. A linear transformation ϕ between algebras $\mathcal{A}$ and $\mathcal{C}$ is said to be a derivation if

$$\phi(AB) = A\phi(B) + \phi(A)B$$

for all $A, B \in \mathcal{A}$.

When $\mathcal{A} = \mathcal{C}$ such a derivation is inner if

$$\phi(A) = AC - CA$$

for some $C \in \mathcal{A}$.

Theorem 6 (Jacobson [9]). *If $\mathcal{B}$ is a semi-simple unital subalgebra of a finite-dimensional central simple algebra $\mathcal{A}$, then any derivation $\delta : \mathcal{B} \longrightarrow \mathcal{A}$ extends to an inner derivation on $\mathcal{A}$.*

Combining theorems 4 and 6 with the fact that $\mathcal{L}_{\mathrm{F}}(\mathcal{V})$ is a central simple algebra, we arrive at the following.

Corollary 7. *If $\mathcal{B}$ is an irreducible subalgebra of $\mathcal{L}_{\mathrm{F}}(\mathcal{V})$, then every derivation $\delta : \mathcal{B} \longrightarrow \mathcal{L}_{\mathrm{F}}(\mathcal{V})$ is inner.*

Convention 8. When $\mathcal{D}$ is a subalgebra of $\mathrm{M}_k(\mathrm{F})$, we can identify $\mathrm{M}_m(\mathcal{D})$ with a subalgebra of $\mathrm{M}_{mk}(\mathrm{F})$ by partitioning matrices in $\mathrm{M}_{mk}(\mathrm{F})$ in a natural fashion.

Suppose that $\mathcal{V}$ is an n-dimensional vector space over a field F. Given a direct sum decomposition $\mathcal{V} = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \mathcal{V}_3 \dotplus \ldots \dotplus \mathcal{V}_k$, every element T of $\mathcal{L}(\mathcal{V})$ can be represented by a matrix $[T_{ij}]$ of linear transformations, where $T_{ij} : \mathcal{V}_j \longrightarrow \mathcal{V}_i$. Once a basis β_i for each $\mathcal{V}_i$ has been chosen, each T_{ij} can be represented by a matrix $\langle T_{ij}\rangle_{\beta_i \leftarrow \beta_j}$ in $\mathrm{M}_{n_i \times n_j}(\mathrm{F})$, where $n_t = dim(\mathcal{V}_t)$. It is common to write $\langle T_{ii}\rangle_{\beta_i}$ for $\langle T_{ii}\rangle_{\beta_i \leftarrow \beta_i}$.

Definition 9. A subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$ is said to have a reduced block-upper-triangular form with respect to a direct sum decomposition $\mathcal{V} = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \mathcal{V}_3 \dotplus \ldots \dotplus \mathcal{V}_k$, if every transformation $A \in \mathcal{A}$ has a matrix form

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} & \ldots & A_{1k} \\ 0 & A_{22} & A_{23} & \ldots & A_{2k} \\ 0 & 0 & A_{33} & \ldots & A_{3k} \\ \ldots & \ldots & \ldots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & A_{kk} \end{bmatrix}$$

with respect to this decomposition, and for each i the subalgebra $\{ \ A_{ii} \mid A \in \mathcal{A} \ \}$ of $\mathcal{L}(\mathcal{V}_i)$ is irreducible. We shall denote this subalgebra by $\mathcal{A}_{ii}$, making the notation $\mathcal{A}_{ij}$ self-explanatory, whenever the underlying decomposition of the space is unambiguous.

A subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$ is said to have a standard block-upper-triangular form with respect to $\mathcal{V} = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \mathcal{V}_3 \dotplus \ldots \dotplus \mathcal{V}_k$, if it has a reduced block-upper-triangular form with respect to this decomposition, and the set $\{1, 2, \dots, k\}$ can be partitioned into non-empty subsets $\Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_m$ such that the following hold.

(1) There are bases $\beta_1, \beta_2, \beta_3, \dots, \beta_k$ of $\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3, \dots, \mathcal{V}_k$ respectively, such that if i and j are in the same Γ_s (in which case we refer to them as "linked"), then $\langle A_{ii} \rangle_{\beta_i} = \langle A_{jj} \rangle_{\beta_j}$ for every $A \in \mathcal{A}$;

(2) If $\mathcal{A}_{ii} \neq \{0\}$ then there exists $G^{<i>} \in \mathcal{A}$ such that $G^{<i>}_{ii} = \mathcal{I}_{\mathcal{V}_i}$ and $G^{<i>}_{jj} = 0$ for all j not linked to i.

Note that it follows from part (1) that $dim(\mathcal{V}_i) = dim(\mathcal{V}_j)$ whenever i is linked to j.

It also follows that

$$\{ \ (A_{ii}, A_{jj}) \mid A \in \mathcal{A} \ \} = \{ \ A_{ii} \mid A \in \mathcal{A} \ \} \times \{ \ A_{jj} \mid A \in \mathcal{A} \ \},$$

when i is not linked to j.

Let the basis β of $\mathcal{V}$ be the concatenation of the bases $\beta_1, \beta_2, \beta_3, \dots, \beta_k$ described above. The subalgebra $\hat{\mathcal{A}} = \left\{ \ \langle A \rangle_\beta \ \middle| \ A \in \mathcal{A} \ \right\}$ of $\mathrm{M}_{dim(\mathcal{V})}(\mathrm{F})$ is said to be a standard matricial form of $\mathcal{A}$.

We can interpret $\hat{\mathcal{A}}$ as an algebra of block-upper-triangular matrices with the (i,j)-th block-entry of the block-matrix for A being the matrix $\langle A_{ij} \rangle_{\beta_i \leftarrow \beta_j}$. In this case we say that $\hat{\mathcal{A}}$ is a standard block-upper-triangular matricial form of $\mathcal{A}$.

If the underlying field is algebraically closed, then, for each i, either

$$\{ \ A_{ii} \mid A \in \mathcal{A} \ \} = \{0\} \quad \text{and} \quad \dim \mathcal{V}_i = 1,$$

or

$$\{ \ A_{ii} \mid A \in \mathcal{A} \ \} = \mathcal{L}(\mathcal{V}_i),$$

by Burnside's Theorem (see, for example, [10]). While Burnside's Theorem does not hold for general fields (without some additional hypotheses), a spatial characterization of irreducible matrix algebras does exist.

Of course we have already seen (Proposition 4) that every irreducible subalgebra of $\mathcal{L}_{\mathrm{F}}(\mathcal{V})$ is simple. The fact that every finite-dimensional simple F-algebra is, up to an algebra isomorphism, a full matrix algebra over a division F-algebra, is a classical theorem of J.M.H. Wedderburn. The advantage of the theorem below is that it characterizes irreducible matrix algebras spatially up to similarity. This very useful theorem is certainly part of the "folklore", but since we were not able to find a proof of this result in the literature, we offer one here for the sake of completeness and as a service to those entering the discipline.

Theorem 10. *A subalgebra $\mathcal{A}$ of $\mathrm{M}_n(\mathrm{F})$ is irreducible if and only if the minimum non-zero value r of the rank function on $\mathcal{A}$ divides n and there is an irreducible division subalgebra D of $\mathrm{M}_r(\mathrm{F})$ such that $\mathcal{A}$ is (simultaneously) similar to $\mathrm{M}_{\frac{n}{r}}(\mathcal{D})$.*

Proof. If $\mathcal{A}$ is a division algebra, the proof is trivial (with $n = r$ and $\mathcal{D} = \mathcal{A}$). Let us therefore assume that $\mathcal{A}$ is not a division algebra. For obvious reasons we lose no generality if we replace $\mathcal{A}$ with an algebra simultaneously similar to $\mathcal{A}$; we do so several times in the proof that follows.

Let r be the minimum non-zero value of the rank function on $\mathcal{A}$. Then $r < n$ and the set $\mathcal{J}$ of the elements of $\mathcal{A}$ of rank at most r is a semigroup ideal in $\mathcal{A}$; i.e. $\mathcal{J}\mathcal{A} \cup \mathcal{A}\mathcal{J} \subset \mathcal{J}$.

It is well-known (and is easy to check by considering the common kernel of $\mathcal{J}$, and $\mathcal{J}\mathcal{M}$, where $\mathcal{M}$ is an invariant subspace of $\mathcal{J}$) that a non-zero semigroup ideal of an irreducible matrix semigroup is also irreducible. By a celebrated theorem of Levitzki (see, for example, Theorem 35 of part II in [10]), $\mathcal{J}$ cannot be nil.

Claim 1: $\mathcal{J}$ contains a non-zero idempotent.

Let A be a non-nilpotent element of $\mathcal{J}$. Since $range(p(A)) \subset range(A)$, when $p(x) \in \mathrm{F}[x]$ has no constant term and $p(A) \neq 0$, we conclude, using the definition of $\mathcal{J}$, that $range(p(A)) = range(A)$.

Let $\mu(x)$ be the minimal polynomial of A in $\mathrm{F}[x]$. If μ has a (non-zero) constant term then A is invertible, and consequently $r = n$, so that $\mathcal{A}$ is a division algebra, which we have assumed it is not. Therefore it must be that μ has no constant term.

We claim that in this case $\mu(x) = xq(x)$, where $q(x)$ is a polynomial with a non-zero constant term. Indeed, if q has no constant term then $\mu(x) = x\,q(x) = x^2\,h(x)$ for some $0 \neq h(x) \in \mathrm{F}[x]$, and therefore $0 = \mu(A) = h(A)\,A^2$. Since $A^2 \neq 0$, we can conclude that $range(A^2) = range(A)$ and thus $0 = q(A) = h(A)\,A$, contradicting the minimality of μ.

Thus we can write $q(x) = x\,\rho(x) - \alpha$ for some $\rho(x) \in \mathrm{F}[x]$ and a non-zero $\alpha \in \mathrm{F}$. Consequently

$$0 = \alpha^{-1}\mu(A) = \big(\alpha^{-1}A\,\rho(A) - I\big)\,A,$$

or equivalently:

$$\big(\alpha^{-1}\rho(A)A\big)\,A = A.$$

Let us write $\nu(x) = \alpha^{-1}\rho(x)\,x$, so that $0 \neq \nu(x) \in \mathrm{F}[x]$, $\nu(0) = 0$ and

$$0 \neq A = \nu(A)A \in \mathcal{A}\mathcal{J} \subset \mathcal{J}.$$

Since we have that $range(\nu(A)) = range(A)$, $\nu(A)$ acts as an identity transformation on its own range, and is therefore a non-zero idempotent matrix in $\mathcal{J}$, which demonstrates the validity of Claim 1.

Now, let E be a non-zero idempotent in $\mathcal{J}$, so that $\mathrm{F}^n = range(E) \dotplus kernel(E)$. By changing to a basis that is a concatenation of a basis of $range(A)$ and a basis of $kernel(E)$, we can assume without loss of generality that the elements of $\mathcal{A}$ are now represented as 2×2 block-matrices with respect to the given decomposition of the underlying space F^n. Clearly

$$E = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Since $I_n \in \mathcal{A}$ by Proposition 4, we see that

$$I_n - E = \begin{bmatrix} 0 & 0 \\ 0 & I_{n-r} \end{bmatrix} \in \mathcal{A}.$$

In view of Definition 9, the notation $\mathcal{A}_{ij}$ is self-explanatory, and, for example, we can write

$$E\mathcal{A}E = \begin{bmatrix} \mathcal{A}_{11} & 0 \\ 0 & 0 \end{bmatrix}.$$

Since every non-zero element of the algebra $E\mathcal{A}E$ has rank r, we conclude that $\mathcal{A}_{11}$ is a division subalgebra of $\mathrm{M}_r(\mathrm{F})$. Let us write

$$\mathcal{D} \stackrel{\text{def}}{=} \mathcal{A}_{11}.$$

Since the rank of $I_n - E$ is not zero, it must be at least r, and we can conclude that $n - r \geq r$. Furthermore,

$$(I_n - E)\mathcal{A}(I_n - E) = \begin{bmatrix} 0 & 0 \\ 0 & \mathcal{A}_{22} \end{bmatrix},$$

and it follows that $\mathcal{A}_{22}$ is a subalgebra of $\mathrm{M}_{n-r}(\mathrm{F})$.

Let us note that $\mathcal{A}_{21}\mathcal{A}_{12} \subset \mathcal{A}_{22}$. We claim that $\mathcal{A}_{21}\mathcal{A}_{12} \neq \{0\}$, from which it follows that $\mathcal{A}_{22}$ contains a non-zero element of rank at most r, and thus of rank exactly r. Now, suppose for the sake of contradiction that $\mathcal{A}_{21}\mathcal{A}_{12} = \{0\}$. We know that $\mathcal{A}_{12} \neq \{0\}$, since $kernel(E)$, being non-zero and proper, cannot be an invariant subspace of the irreducible algebra $\mathcal{A}$. Thus the common kernel $\mathcal{K}$ of $\mathcal{A}_{21}$ is non-trivial. Since the common kernel of $\mathcal{A}$ is trivial, we must have:

$$\{0\} \neq \mathcal{A}\mathcal{K} \subset range(E) \subsetneq \mathrm{F}^n,$$

which implies that $\mathcal{A}\mathcal{K}$ is a proper non-trivial invariant subspace of $\mathcal{A}$, contradicting the hypothesis of the irreducibility of $\mathcal{A}$.

We also claim that $\mathcal{A}_{22}$ is an irreducible subalgebra of $\mathrm{M}_{n-r}(\mathrm{F})$. If this were not the case, and $\{0\} \neq \mathcal{M} \neq \mathrm{F}^{n-r}$ is an invariant subspace of $\mathcal{A}_{22}$, writing $\mathcal{M}'$ for the corresponding subspace of $kernel(E)$, we would have that either $\mathcal{A}\mathcal{M}' = \{0\}$, or

$$\{0\} \neq \mathcal{A}\mathcal{M}' \subset range(E) \dotplus \mathcal{M}' \neq \mathrm{F}^n.$$

In the former case $\mathcal{M}$' is a non-zero proper invariant subspace of $\mathcal{A}$, while in the latter it is $\mathcal{A}\mathcal{M}$' that plays that role. Either case contradicts the hypothesis of the irreducibility of $\mathcal{A}$.

Now, $\mathcal{A}_{22}$ is an irreducible subalgebra of $\mathrm{M}_{n-r}(\mathrm{F})$, and as we have shown above, the minimum non-zero value of the rank function on $\mathcal{A}_{22}$ is r. If $\mathcal{A}_{22}$ is not a division algebra, i.e. if $n - r > r$, the whole argument above can be repeated in the case of $\mathcal{A}_{22}$, to show that there is a direct sum decomposition

$$\mathrm{F}^n = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \mathcal{V}_3,$$

such that

$$dim\mathcal{V}_1 = dim\mathcal{V}_2 = r \leq n - 2r = dim\mathcal{V}_3,$$

and with respect to a congenial basis the elements of $\mathcal{A}$ can be represented by 3×3 block-matrices, with

$$\begin{bmatrix} I_r & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & I_r & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & I_{n-2r} \end{bmatrix} \in \mathcal{A}.$$

The process of the repetition of the argument (applied to A_{33}, A_{44}, etc., as needed) must terminate, and it does so exactly when we arrive at the case "$n-(k-1)r = r$" with $\mathcal{A}_{k\,k}$ being a division subalgebra of $\mathrm{M}_r(\mathrm{F})$.

Hence we conclude that $n = kr$ for some $k \in \mathrm{N}$, and there is a direct sum decomposition

$$\mathrm{F}^n = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \ldots \dotplus \mathcal{V}_k,$$

with $dim\mathcal{V}_i = r$, and that with respect to a congenial basis the elements of $\mathcal{A}$ are represented by $k \times k$ block-matrices, such that

$$E_{11} \stackrel{\text{def}}{=} \begin{bmatrix} I_r & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}, \quad E_{22} \stackrel{\text{def}}{=} \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ 0 & I_r & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}, \ldots, \quad E_{kk} \stackrel{\text{def}}{=} \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & I_r \end{bmatrix} \quad \in \mathcal{A}.$$

We continue to use the notation $\mathcal{A}_{ij}$ to indicate the linear space of all matrices in $\mathrm{M}_r(\mathrm{F})$ that appear as the (i,j)-th block entry of a matrix in $\mathcal{A}$, and equivalently, of a matrix in $E_{ii}\mathcal{A}E_{jj}$. Each non-zero matrix in $\mathcal{A}_{ij}$ has rank at least r, and so exactly r; in other words, each non-zero matrix in $\mathcal{A}_{ij}$ is invertible. Obviously $\mathcal{A}_{ij}\mathcal{A}_{jk} \subset \mathcal{A}_{ik}$. Since $\mathcal{A}$ is irreducible, so is each $\mathcal{A}_{ij}$, and each is non-zero; (one justifies this via the arguments similar to those presented above). In particular each $\mathcal{A}_{ii}$ is a division algebra.

Since each of $\mathcal{A}_{12}, \mathcal{A}_{13}, \ldots, \mathcal{A}_{1k}$ contains an invertible element, applying a block-diagonal similarity to $\mathcal{A}$ if necessary, we can assume that $I_r \in \mathcal{A}_{1j}$, for all j. Let us write $C \otimes E_{ij}$ for the block matrix that has C as its (i,j)-th block entry, and zero blocks in all other positions. We will simply write E_{ij} for $I_r \otimes E_{ij}$, and we already know that E_{jj} and E_{1j} are elements of $\mathcal{A}$ for all j.

The next step is to show that $E_{i1} \in \mathcal{A}$ for all i. To this end, note that $\mathcal{A}_{i1}$ contains an invertible element C, and therefore

$$C \otimes E_{i1} \in E_{ii}\mathcal{A}E_{11} \subset \mathcal{A}.$$

Thus $C \otimes E_{ii} = (C \otimes E_{i1})E_{1i}$, so that C is an invertible element of the division algebra A_{ii}, and consequently $C^{-1} \otimes E_{ii} \in \mathcal{A}$. Hence:

$$E_{i1} = (C^{-1} \otimes E_{ii})(C \otimes E_{i1}) \in \mathcal{A}.$$

To complete the proof of the theorem we demonstrate that $\mathcal{A}_{ij} = \mathcal{A}_{11} = \mathcal{D}$ for all i, j. If $C \in \mathcal{A}_{ij}$ then $C \otimes E_{ij} \in \mathcal{A}$, and

$$E_{1i}(C \otimes E_{ij})E_{j1} = C \otimes E_{11},$$

so that $C \in \mathcal{A}_{11}$. Conversely, if $B \in \mathcal{A}_{11}$ then $B \otimes E_{11} \in \mathcal{A}$, and

$$E_{i1}(B \otimes E_{11})E_{1j} = B \otimes E_{ij}$$

so that $B \in \mathcal{A}_{ij}$.

Thus $\mathcal{A}$ contains "$\mathcal{D} \otimes E_{ij}$" for every i, j and consequently $\mathcal{A} = \mathrm{M}_k(\mathcal{D})$, which completes the proof. ■

We will need the following standard result when we treat the uniqueness aspect of the block-diagonal part of an algebra in a standard block-upper-triangular form.

Theorem 11. *If $\mathcal{A}$ and $\mathcal{B}$ are irreducible subalgebras of $\mathrm{M}_m(\mathrm{F})$ and $\mathrm{M}_n(\mathrm{F})$ respectively, and $\phi : \mathcal{A} \longrightarrow \mathcal{B}$ is an algebra isomorphism, then $m = n$ and ϕ is spatial; i.e. there is an invertible $T \in \mathrm{M}_n(\mathrm{F})$ such that*

$$\phi(A) = T^{-1}AT \quad \text{for all} \quad A \in \mathcal{A}.$$

Proof. By Theorem 10, after applying similarities, we can assume that $\mathcal{A} = \mathrm{M}_{\frac{n}{r}}(\mathcal{D})$, where D is an irreducible division subalgebra of $\mathrm{M}_r(\mathrm{F})$, and $\mathcal{B} = \mathrm{M}_{\frac{m}{p}}(\mathcal{G})$, where G is an irreducible division subalgebra of $\mathrm{M}_p(\mathrm{F})$. By Proposition 4, $\mathcal{A}$ and $\mathcal{B}$ are simple, and therefore the uniqueness part of the classical Wedderburn structure theorem for simple algebras (see, for example, Theorem 4.23 in [11]) dictates that $\frac{n}{r} = \frac{m}{p}$ and that $\mathcal{D}$ is isomorphic to $\mathcal{G}$.

Since $\mathcal{D}$ is irreducible, it contains the identity matrix (Proposition 4). Since $\mathcal{D}$ is also a division algebra, every non-zero element of $\mathcal{D}$ is an invertible matrix. Thus, if two non-zero elements of $\mathcal{D}$ agree on a non-zero vector, they must be equal, since otherwise their difference is a non-zero element of $\mathcal{D}$ with a non-trivial kernel. The same applies to $\mathcal{G}$.

Thus the first basis vector e_1 in F^r is a separating vector for $\mathcal{D}$, in the sense that

$$De_1 = D'e_1 \Longrightarrow D = D',$$

and a similar statement holds true for $\mathcal{G}$.

Since $\mathcal{D}$ is irreducible, the "evaluation at the first standard basis vector" function $\alpha : \mathcal{D} \longrightarrow \mathrm{F}^r$ is a linear bijection. The same is true for the corresponding function $\beta : \mathcal{G} \longrightarrow \mathrm{F}^p$. Being isomorphic, $\mathcal{D}$ and $\mathcal{G}$ have the same dimension, and so it follows that $r = p$, and consequently $m = n$.

To see that ϕ must be spatial, we can now apply the Noether-Skolem Theorem within $\mathrm{M}_n(\mathrm{F})$ (see, for example, Theorem 4.3.1 of [12]), via Proposition 4. Note that $\phi(\mathcal{I})$ must be the identity transformation, being the multiplicative identity in $\mathcal{B}$, and thus an idempotent whose range and kernel are invariant under $\mathcal{B}$. ■

Definition 12. Each invertible $S \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ induces an algebra isomorphism

$$(\ ) \longrightarrow S^{-1} \circ (\ ) \circ S$$

between $\mathcal{L}(\mathcal{W})$ and $\mathcal{L}(\mathcal{V})$. These maps are called "similarity transformations". The corresponding concept for $\mathrm{M}_n(\mathrm{F})$ is self-explanatory.

Theorem 13 (Watters [8])**.** *If a subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$ has a reduced block-upper-triangular form with respect to a decomposition $\mathcal{V} = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \mathcal{V}_3 \dotplus \ldots \dotplus \mathcal{V}_k$, then after an application of a block-diagonal similarity $\mathcal{A}$ has a standard block-upper-triangular form with respect to this decomposition.*

Corollary 14. *If a subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$ has a reduced block-upper-triangular form with respect to a decomposition $\mathcal{V} = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \mathcal{V}_3 \dotplus \ldots \dotplus \mathcal{V}_k$, then the set $\{1, 2, \ldots, k\}$ can be partitioned into non-empty subsets $\Gamma_1, \Gamma_2, \Gamma_3, \ldots, \Gamma_m$ such that*

(1) If $\mathcal{A}_{ii} \neq \{0\}$ then there exists $G^{<i>} \in \mathcal{A}$ such that $G^{<i>}_{jj} = \mathcal{I}_{\mathcal{V}_j}$ for all j linked to i, and $G^{<i>}_{jj} = 0$ for all j not linked to i.

(2) When i is linked to j,

$$dim(\mathcal{V}_i) = dim(\mathcal{V}_j),$$

and there is an invertible $S_{ij} \in \mathcal{L}(\mathcal{V}_i, \mathcal{V}_j)$ *such that*

$$A_{ii} = S_{ij}^{-1} A_{jj} S_{ij}, \quad \text{for all } A \in \mathcal{A}.$$

(3) When i *is not linked to* j*,*

$$\{ \ (A_{ii}, A_{jj}) \mid A \in \mathcal{A} \ \} = \{ \ A_{ii} \mid A \in \mathcal{A} \ \} \times \{ \ A_{jj} \mid A \in \mathcal{A} \ \}.$$

As before, indices i *and* j *are "linked" if they are in the same* Γ_s*.*

Observation 15. If a subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$ has a reduced block-upper-triangular form with respect to a given decomposition of $\mathcal{V}$, and an invertible $S \in \mathcal{L}(\mathcal{V})$ is block-upper-triangular with respect to the same decomposition, then $S^{-1}\mathcal{A}S$ has a reduced block-upper-triangular form with respect to the decomposition, and index i is linked to an index j for $S^{-1}\mathcal{A}S$ if and only if i is linked to j for $\mathcal{A}$.

Terminology 16. Given a direct sum decomposition $\mathcal{V} = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \mathcal{V}_3 \dotplus \ldots \dotplus \mathcal{V}_k$, if some of the spaces $\mathcal{V}_i$ are in turn decomposed as direct sums of their subspaces, the resulting direct sum decomposition of $\mathcal{V}$ is said to be a refinement of the original decomposition.

Linear transformations which are block-diagonal with respect to a refinement are automatically block-diagonal with respect to the original decomposition.

It is a standard fact that if an algebra is block-upper-triangular with respect to a given direct sum decomposition, then it has a reduced block-upper-triangular form with respect to some refinement of this decomposition.

Hence we can use Theorem 13 to draw the following conclusion.

Corollary 17. *If a subalgebra* $\mathcal{A}$ *of* $\mathcal{L}(\mathcal{V})$ *has a block-upper-triangular form with respect to a decomposition* $\mathcal{V} = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \mathcal{V}_3 \dotplus \ldots \dotplus \mathcal{V}_k$*, then after an application of a block-diagonal similarity* $\mathcal{A}$ *has a standard block-upper-triangular form with respect to a refinement of this decomposition.*

Notation 18. For a matrix $A \in \mathrm{M}_n(\mathrm{F})$ we denote by $Diag(A)$ the diagonal matrix in $\mathrm{M}_n(\mathrm{F})$ that has the same diagonal as A.

When interpreting elements of $\mathcal{L}(\mathcal{V})$ as block-matrices with respect to a given decomposition of $\mathcal{V}$, we denote by $BlockDiag(B)$ the block-diagonal matrix obtained by replacing the block-"off-diagonal" entries of B with zeros. We refer to the map $B \longrightarrow BlockDiag(B)$ as "the compression to the block-diagonal".

The following result is certainly not new. We give a short proof for the sake of completeness.

Proposition 19. *If a subalgebra* $\mathcal{A}$ *of* $\mathcal{L}(\mathcal{V})$ *has a reduced block-upper-triangular form with respect to a direct sum decomposition of* $\mathcal{V}$*, then* $Rad(\mathcal{A})$ *is exactly the set of all strictly block-upper-triangular elements of* $\mathcal{A}$*.*

Proof. It is obvious that the set of strictly block-upper-triangular elements of $\mathcal{A}$ is a subset of the radical of $\mathcal{A}$. If $\mathcal{A}$ is strictly block-upper-triangular the proof is complete. Henceforth assume that $\mathcal{A}$ has elements with a non-zero block-diagonal.

If $\mathcal{V} = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \mathcal{V}_3 \dotplus \ldots \dotplus \mathcal{V}_k$ is the underlying decomposition, and i is such that $\mathcal{A}_{ii} \neq \{0\}$, then the set

$$\mathcal{J} = \{ \ B \in \mathcal{L}(\mathcal{V}_i) \mid B = A_{ii} \text{ for some } A \in Rad(\mathcal{A}) \ \},$$

does not contain the identity transformation (by spectral considerations), and so is a proper ideal in the irreducible algebra $\mathcal{A}_{ii}$. By Proposition 4, $\mathcal{J} = \{0\}$. ■

Observation 20. If a subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$ has a standard block-diagonal form with respect to a decomposition $\mathcal{V} = \mathcal{W}_1 \dotplus \mathcal{W}_2 \dotplus \mathcal{W}_3 \dotplus \ldots \dotplus \mathcal{W}_k$, then $\mathcal{A}$ is semi-simple by Proposition 19, and the idempotents $G^{<i>} \in \mathcal{A}$ described in Definition 9 are exactly the minimal central idempotents of $\mathcal{A}$. Obviously $G^{<i>}G^{<j>} = 0$, whenever $G^{<i>} \neq G^{<j>}$, i.e. whenever i is not linked to j. If i is linked to j, then $G^{<i>} = G^{<j>}$.

Furthermore, if $G^{<i_1>}, \ldots, G^{<i_m>}$ is the complete list of the distinct minimal central idempotents of $\mathcal{A}$ [1], then the sub-algebras $G^{<i_t>}\mathcal{A}G^{<i_t>}$ are exactly the simple components of $\mathcal{A}$, in the sense of the classical Wedderburn-Artin structure theorem for semi-simple algebras (see, for example, section 4 of chapter III in [13]). It is also obvious that each $G^{<i_t>}\mathcal{A}G^{<i_t>}$ is algebra-isomorphic to $\mathcal{A}_{i_t i_t}$, since a similarity transforms $\mathcal{A}_{ii}$ into $\mathcal{A}_{jj}$ whenever i is linked to j.

2. Main results

A standard module theory result yields that, up to similarity, every semi-simple subalgebra of $\mathrm{M}_n(\mathrm{F})$ is block-diagonal with irreducible blocks. Our first main theorem shows that for a semi-simple algebra of matrices in a reduced block-upper-triangular form there is *a block-upper-triangular similarity* that implements the compression to the block-diagonal.

This result will yield a spatial version of Wedderburn's Principal Theorem.

Theorem 21. *Suppose that a semi-simple subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$ has a reduced block-upper-triangular form with respect to the decomposition*

$$\mathcal{V} = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \mathcal{V}_3 \dotplus \ldots \dotplus \mathcal{V}_k. \tag{1}$$

Then there exists an invertible $T \in \mathcal{L}(\mathcal{V})$, which is block-upper-triangular with respect to the decomposition (1), such that

$$T^{-1}AT = BlockDiag(A) \tag{2}$$

for every $A \in \mathcal{A}$.

We present a proof of this theorem in section 4, but consider some of its consequences at present.

Corollary 22. *If a semi-simple subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$ is block-upper-triangular with respect to a given direct sum decomposition of $\mathcal{V}$, then there exists an invertible block-upper-triangular T such that $T^{-1}AT$ is block-diagonal for every $A \in \mathcal{A}$.*

Proof. By Corollary 17, after a block-diagonal similarity $\mathcal{A}$ has a standard block-upper-triangular form with respect to a refinement of the given decomposition, and any matrix that is block-upper-triangular (or block-diagonal) with respect to a refinement is still block-upper-triangular (resp. block-diagonal) with respect to the original decomposition. Hence the result follows by Theorem 21. ■

[1] i.e. $i_1, i_2, i_3, \ldots, i_m$ is a complete set of representatives of the partition $\Gamma_1, \Gamma_2, \Gamma_3, \ldots, \Gamma_m$.

The following Jordan-Hölder-type result addresses the uniqueness of the structure of the block-diagonal for a semi-simple algebra in a standard block-upper-triangular form.

Theorem 23. *Suppose that a subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$ has a standard block-diagonal form with respect to a decomposition $\mathcal{V} = \mathcal{W}_1 \dot{+} \mathcal{W}_2 \dot{+} \mathcal{W}_3 \dot{+} \ldots \dot{+} \mathcal{W}_k$, as well as with respect to a decomposition $\mathcal{V} = \mathcal{Z}_1 \dot{+} \mathcal{Z}_2 \dot{+} \mathcal{Z}_3 \dot{+} \ldots \dot{+} \mathcal{Z}_p$.*

Then $k = p$ and there is a permutation π on $\{1, 2, \ldots, k\}$ such that:

(1) i is linked to j (in the $\mathcal{W}$-decomposition) if and only if $\pi(i)$ is linked to $\pi(j)$ (in the $\mathcal{Z}$-decomposition);
(2) for each i there exists an invertible $T_i \in \mathcal{L}(\mathcal{W}_i, \mathcal{Z}_{\pi(i)})$ such that

$$A_{|_{\mathcal{W}_i}} = T_i^{-1} A_{|_{\mathcal{Z}_{\pi(i)}}} T_i \,, \quad \textit{for every} \;\; A \in \mathcal{A}.$$

The use of the terminology "linked" applied to the indices in this context is based on the result of Corollary 14.

Proof. We shall use subscripts and superscripts $(\mathcal{W})$ and $(\mathcal{Z})$ to distinguish the two decompositions. By Observation 20 we see that the set of the idempotents $G_{\mathcal{W}}^{<t>}$ equals the set of the idempotents $G_{\mathcal{Z}}^{<s>}$, being just the set of the minimal central idempotents of $\mathcal{A}$.

Furthermore, if $G_{\mathcal{W}}^{<i>} = G_{\mathcal{Z}}^{<s_i>}$, then the irreducible algebra $\mathcal{A}_{ii}^{(\mathcal{W})}$ is algebra-isomorphic to the irreducible algebra $\mathcal{A}_{s_i s_i}^{(\mathcal{Z})}$, since both are algebra-isomorphic to $G_{\mathcal{W}}^{<i>} \mathcal{A} G_{\mathcal{W}}^{<i>}$. By Theorem 11, it follows that $dim(\mathcal{W}_i) = dim(\mathcal{Z}_{s_i})$ and that there exists an invertible $T_i \in \mathcal{L}(\mathcal{W}_i, \mathcal{Z}_{s_i})$ such that

$$A_{|_{\mathcal{W}_i}} = T_i^{-1} A_{|_{\mathcal{Z}_{s_i}}} T_i \,, \quad \text{for every} \;\; A \in \mathcal{A}. \tag{3}$$

Note that the sets $\left\{ \; i \mid \mathcal{A}_{ii}^{(\mathcal{W})} = \{0\} \; \right\}$ and $\left\{ \; j \mid \mathcal{A}_{jj}^{(\mathcal{Z})} = \{0\} \; \right\}$ have the same cardinality, that being the dimension of the common kernel of $\mathcal{A}$. Since

$$G^{<i>} = G^{<j>} \iff \; i \;\text{ is linked to }\; j,$$

and there is a similarity mapping $\mathcal{A}_{ii}$ onto $\mathcal{A}_{jj}$, whenever i is linked to j, for either decomposition, the conclusion of the theorem follows by (3) and a dimensionality argument.

■

One immediate consequence of Theorems 10, 13 and 21 is the spatial version of the classical Wedderburn-Artin theorem for semi-simple matrix algebras. When $\mathcal{A}$ is an algebra of matrices we use the notation $\mathcal{A}^{(p)}$ for the matrix algebra

$$\left\{ \begin{bmatrix} A & 0 & \ldots & 0 \\ 0 & A & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & A \end{bmatrix} \;\middle|\; A \in \mathcal{A} \right\} \subset \mathrm{M}_p(\mathcal{A}).$$

Corollary 24 (A spatial version of Wedderburn-Artin theorem)**.** *If $\mathcal{A}$ is a semi-simple subalgebra of $\mathrm{M}_n(\mathrm{F})$, then there exist irreducible division algebras $\mathcal{D}_i$ of* *matrices over* F*, such that* $\mathcal{A}$ *is simultaneously similar to an internal direct sum*

$$\bigoplus (\mathrm{M}_{k_i}(\mathcal{D}_i))^{(p_i)}.$$

Definition 25. A subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$ is said to be a cleft algebra if there exists a semi-simple subalgebra $\mathcal{S}$ of $\mathcal{A}$ (referred to as a Wedderburn factor of $\mathcal{A}$) such that

$$\mathcal{A} = \mathcal{S} \dotplus Rad(\mathcal{A}) \quad \text{(a vector space direct sum).}$$

Such an $\mathcal{A}$ has an unhinged block-upper-triangular form with respect to a given decomposition of $\mathcal{V}$ if $\mathcal{A} = BlockDiag(\mathcal{A}) \dotplus Rad(\mathcal{A})$, where $BlockDiag(\mathcal{A})$ is semi-simple and $Rad(\mathcal{A})$ coincides with the set of all strictly block-upper-triangular elements of $\mathcal{A}$.

Of course terms such as "standard" and "reduced" may apply to an unhinged block-upper-triangular form of $\mathcal{A}$ in a way described in Definition 9.

Terminology 26. A finite-dimensional semi-simple F-algebra $\mathcal{A}$ is said to be separable, if for every field extension E of F, $\mathcal{A}$ is semi-simple as an algebra over E. This is equivalent to the statement that the center of each of the "simple components" of $\mathcal{A}$ is a separable field extension of F (see, for example, Theorem 35 of Chapter 5 in [14]).

In particular, if F is a perfect field then every finite-dimensional semi-simple algebra over F is separable. Finite fields, fields of characteristic zero and algebraically closed fields are perfect (Exercise 13 in section 6 of chapter V in [15]).

A subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$ is said to be separable, if the semi-simple algebra $\mathcal{A}/Rad(\mathcal{A})$ is separable.

Theorem 27 (Wedderburn's Principal Theorem)**.** *Every separable subalgebra of* $\mathcal{L}(\mathcal{V})$ *is a cleft algebra.*

Corollary 28. *If* $\mathcal{V}$ *is a vector space over a perfect* F*, then every subalgebra of* $\mathcal{L}(\mathcal{V})$ *is a cleft algebra.*

Theorem 29 (Malcev [7])**.** *If* $\mathcal{A}$ *is a cleft subalgebra of* $\mathcal{L}(\mathcal{V})$*, then for any two Wedderburn factors* $\mathcal{S}$ *and* $\mathcal{T}$ *of* $\mathcal{A}$ *there is an element* N *of* $Rad(\mathcal{A})$ *such that*

$$\mathcal{T} = (\mathcal{I} - N)^{-1}\mathcal{S}(\mathcal{I} - N);$$

(where the use of $\mathcal{I}$ *is formal if* $\mathcal{I} \notin \mathcal{A}$*, with* $(\mathcal{I} - N)^{-1} = \mathcal{I} + N + N^2 + \ldots + N^p$*, where* $N^{p+1} = 0$*).*

With this in mind, one can see how our Theorem 21 yields the following.

Corollary 30 (Spatial Wedderburn's Principal Theorem)**.** *If* $\mathcal{V}$ *is a vector space, then every cleft subalgebra* $\mathcal{A}$ *of* $\mathcal{L}(\mathcal{V})$ *has an unhinged standard block-upper-triangular form with respect to a direct sum decomposition of* $\mathcal{V}$*.*

In fact more is true: if $\mathcal{A}$ *has a reduced block-upper-triangular form with respect to a given direct sum decomposition of* $\mathcal{V}$*, then after a block-upper-triangular similarity* $\mathcal{A}$ *has an unhinged standard block-upper-triangular form with respect to the same decomposition.*

Consequently, if $\mathcal{A}$ *is block-upper-triangular with respect to a given direct sum decomposition of* $\mathcal{V}$*, then after a block-upper-triangular similarity* $\mathcal{A}$ *has an unhinged* *standard block-upper-triangular form with respect to a refinement of the original decomposition.*

Corollary 31. *If a subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$ is a cleft algebra and $\mathcal{S}$ is any Wedderburn factor of $\mathcal{A}$ then there is a direct sum decomposition of $\mathcal{V}$ with respect to which $\mathcal{A}$ has an unhinged standard block-upper-triangular form, with $\mathcal{S}$ being the set of the block-diagonal elements of $\mathcal{A}$.*

Proof. Pick a decomposition $\mathcal{V} = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \mathcal{V}_3 \dotplus \ldots \dotplus \mathcal{V}_k$ with respect to which $\mathcal{A}$ has an unhinged standard block-upper-triangular form (see Corollary 30). The set of elements in $\mathcal{A}$ that are block-diagonal with respect to this decomposition is a Wedderburn factor $\mathcal{S}_0$ of $\mathcal{A}$.

By Theorem 29 there is an invertible T which is block-upper-triangular with respect to the decomposition $\mathcal{V} = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \mathcal{V}_3 \dotplus \ldots \dotplus \mathcal{V}_k$ and is such that $T^{-1}ST = S_0$.

Then $\mathcal{A}$ has an unhinged standard block-upper-triangular form with respect to the decomposition $\mathcal{V} = T\mathcal{V}_1 \dotplus T\mathcal{V}_2 \dotplus T\mathcal{V}_3 \dotplus \ldots \dotplus T\mathcal{V}_k$. Indeed $\mathcal{A}$ has a reduced block-upper-triangular form with respect to this decomposition, since $T^{-1}\mathcal{A}T$ has a reduced block-upper-triangular form with respect to $\mathcal{V} = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \mathcal{V}_3 \dotplus \ldots \dotplus \mathcal{V}_k$.

Furthermore, since every $\mathcal{V}_i$ is invariant under $\mathcal{S}_0$, every $T\mathcal{V}_i$ is invariant under $\mathcal{S}$, so that elements in $\mathcal{S}$ are block-diagonal with respect to $\mathcal{V} = T\mathcal{V}_1 \dotplus T\mathcal{V}_2 \dotplus T\mathcal{V}_3 \dotplus \ldots \dotplus T\mathcal{V}_k$.

Since $\mathcal{A} = \mathcal{S} \dotplus Rad(\mathcal{A})$, it follows that $\mathcal{A}$ has an unhinged standard block-upper-triangular form with respect to $\mathcal{V} = T\mathcal{V}_1 \dotplus T\mathcal{V}_2 \dotplus T\mathcal{V}_3 \dotplus \ldots \dotplus T\mathcal{V}_k$ with $\mathcal{S}$ being the set of block-diagonal elements of $\mathcal{A}$.

■

Corollary 32. *If a cleft subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$ has a reduced block-upper-triangular form with respect to a given direct sum decomposition of $\mathcal{V}$, then it contains idempotent elements $G^{<i>}$ with the properties described in Corollary 14.*

Proof. [1] By Corollary 30 there is an invertible S, block-upper-triangular with respect to the given decomposition, such that $S^{-1}\mathcal{A}S$ has an unhinged standard block-upper-triangular form with respect to the same decomposition. The Wedderburn factor $\mathcal{S}$ of $S^{-1}\mathcal{A}S$ has minimal central idempotents $\hat{G}^{<i>}$, as described in the Observation 20, which correspond to the structure of the "index linking" of the diagonal blocks of $\mathcal{S}$. Since we have already noted in the Observation 15 that the structure of the "index linking" of the diagonal blocks is not affected by a block-upper-triangular similarity, we can obtain the required idempotents $G^{<i>}$ in $\mathcal{A}$ by setting

$$G^{<i>} \stackrel{\text{def}}{=} S\hat{G}^{<i>}S^{-1}.$$

■

3. Some applications

We use Corollary 30 extensively to explore the structure of paratransitive matrix algebras in articles [1] and [2]. Let us demonstrate another application of the theorem presently.

[1] A proof of this result can be also extracted from our proof of Theorem 21 in the last section of the paper.

First let us remind the reader of the following standard (finite-dimensional) result.

Proposition 33. *The only non-zero $\mathcal{L}(\mathcal{W})$-$\mathcal{L}(\mathcal{V})$-bimodule of $\mathcal{L}(\mathcal{V},\mathcal{W})$ is $\mathcal{L}(\mathcal{V},\mathcal{W})$ itself.*

Observation 34. Suppose that the underlying field is algebraically closed and a subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$ has an unhinged standard block-upper-triangular form with respect to a decomposition

$$\mathcal{V} = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \mathcal{V}_3 \dotplus \ldots \dotplus \mathcal{V}_k.$$

We write $\mathcal{S}$ for the block-diagonal Wedderburn factor (i.e. the block-diagonal) of $\mathcal{A}$.

The linear subspace $\mathcal{A}_{ij}$, is a left $\mathcal{A}_{ii}$-submodule and a right $\mathcal{A}_{jj}$-submodule of $\mathcal{L}(\mathcal{V}_j,\mathcal{V}_i)$. This is because $\left(\mathcal{S}\mathcal{A}\right)_{ij} = \mathcal{A}_{ii}\mathcal{A}_{ij}$, and $\left(\mathcal{A}\mathcal{S}\right)_{ij} = \mathcal{A}_{ij}\mathcal{A}_{jj}$.

Since we have restricted our attention to the case of an algebraically closed field, Burnside's theorem implies that for each i, $\mathcal{A}_{ii}$ is either equal to $\mathcal{L}(\mathcal{V}_i)$ or is $\{0\}$ (in which case $dim(\mathcal{V}_i) = 1$).

Thus by Proposition 33 we can conclude that $\mathcal{A}_{ij}$ is either $\mathcal{L}(\mathcal{V}_j,\mathcal{V}_i)$ or $\{0\}$, for all i,j.

Observation 35. Combining the results of Corollary 30, Theorem 23 and Corollary 14, we can see that for a subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$, besides the number of subspaces in any decomposition of $\mathcal{V}$ with respect to which $\mathcal{A}$ has a reduced block-upper-triangular form being intrinsic, the linking pattern between the subspaces is intrinsic as well. In other words, the type of the partition $\Gamma_1,\Gamma_2,\ldots,\Gamma_m$ of the indices of the decomposition, as described in Corollary 14, including the dimensions of the subspaces corresponding to the indices in each part is intrinsic.

Notation 36. Given a subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$, let us denote by $\mu(\mathcal{A})$ the number of subspaces in any decomposition of $\mathcal{V}$ with respect to which $\mathcal{A}$ has a reduced block-upper-triangular form.

Theorem 37. *When the underlying field is algebraically closed, the following are equivalent for a subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$:*

(1) $\left(Rad(\mathcal{A})\right)^{\mu(\mathcal{A})-1} \neq \{0\}$*;*
(2) There is an element $R \in Rad(\mathcal{A})$ such that $R^{\mu(\mathcal{A})-1} \neq 0$.

Proof. First of all, it is obvious that an algebra generated by $\mathcal{A}$ and the identity transformation is a cleft algebra with exactly the same radical as $\mathcal{A}$; (see, for example, Proposition 19). Thus, using Corollary 30 we may assume without loss of generality that $\mathcal{A}$ is already unital, and has an unhinged standard block-upper-triangular form with respect to a decomposition $\mathcal{V} = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \mathcal{V}_3 \dotplus \ldots \dotplus \mathcal{V}_{\mu(\mathcal{A})}$, which shall be fixed for the remainder of the proof. Let us write $\mathcal{S}$ for the block-diagonal compression (i.e. the corresponding Wedderburn factor) of $\mathcal{A}$.

The implication (2) $\Rightarrow$ (1) is trivial. To demonstrate the validity of the converse, let us suppose that $\left(Rad(\mathcal{A})\right)^{\mu(\mathcal{A})-1} \neq \{0\}$. Proposition 19 and a standard property of strictly block-upper-triangular $\mu(\mathcal{A}) \times \mu(\mathcal{A})$ block-matrices dictate that for $R \in$

$Rad(\mathcal{A})$,

$$R^{\mu(\mathcal{A})-1} = 0 \iff \prod_{i=1}^{\mu(\mathcal{A})-1} R_{i\,i+1} = 0 \tag{4}$$

Thus the hypothesis implies that $\mathcal{A}_{i\,i+1} \neq \{0\}$ for every i, and therefore

$$\mathcal{A}_{i\,i+1} = \mathcal{L}(\mathcal{V}_{i+1}, \mathcal{V}_i), \tag{5}$$

by Observation 34.

If $\mathcal{A}$ contains an element A such that $A_{i\,i+1} = 0$ and $A_{j\,j+1} \neq 0$, we shall say that the index j is "independent" of the index i. Let us note that the relation thus defined is symmetric. Indeed, if A is as described, then

$$\begin{aligned}\{\ (B_{i\,i+1}, B_{j\,j+1}) \mid B \in \mathcal{S}A\mathcal{S}\ \} &= \{0\} \times \mathcal{A}_{jj} A_{j\,j+1} \mathcal{A}_{j+1\,j+1} \\ &= \{0\} \times \mathcal{L}(\mathcal{V}_j) A_{j\,j+1} \mathcal{L}(\mathcal{V}_{j+1}) \\ &= \{0\} \times \mathcal{L}(\mathcal{V}_{j+1}, \mathcal{V}_j).\end{aligned}$$

The last equality holds because $A_{j\,j+1} \neq 0$, and so $\mathcal{L}(\mathcal{V}_j) A_{j\,j+1} \mathcal{L}(\mathcal{V}_{j+1})$ is a non-zero $\mathcal{L}(\mathcal{V}_j) - \mathcal{L}(\mathcal{V}_{j+1})$-bimodule of $\mathcal{L}(\mathcal{V}_{j+1}, \mathcal{V}_j)$; (see Proposition 33).

Therefore

$$\begin{aligned}\{\ (B_{i\,i+1}, B_{j\,j+1}) \mid B \in \mathcal{A}\ \} &= \Big(\mathcal{A}_{i\,i+1} \times \{0\}\Big) + \Big(\{0\} \times \mathcal{L}(\mathcal{V}_{j+1}, \mathcal{V}_j)\Big) \\ &= \Big(\mathcal{L}(\mathcal{V}_{i+1}, \mathcal{V}_i) \times \{0\}\Big) + \Big(\{0\} \times \mathcal{L}(\mathcal{V}_{j+1}, \mathcal{V}_j)\Big) \\ &= \mathcal{L}(\mathcal{V}_{i+1}, \mathcal{V}_i) \times \mathcal{L}(\mathcal{V}_{j+1}, \mathcal{V}_j),\end{aligned}$$

and in particular i is "independent" of j, so that the symmetry of the relation has been established. In fact we have demonstrated that

$$i \text{ is independent of } j \iff \{\ (B_{i\,i+1}, B_{j\,j+1}) \mid B \in \mathcal{A}\ \} = \mathcal{L}(\mathcal{V}_{i+1}, \mathcal{V}_i) \times \mathcal{L}(\mathcal{V}_{j+1}, \mathcal{V}_j). \tag{6}$$

Suppose that either i is unlinked from j, or $i+1$ is unlinked from $j+1$ (in the sense of Definition 9). We claim that in such a case i and j are independent. First let us suppose that i is unlinked from j, but i and j are not independent. Then

$$\{\ (B_{ii}, B_{jj}) \mid B \in \mathcal{S}\ \} = \{\ (B_{ii}, B_{jj}) \mid B \in \mathcal{A}\ \} = \mathcal{L}(\mathcal{V}_i) \times \mathcal{L}(\mathcal{V}_j), \tag{7}$$

and, by a standard argument (recalling that (5) holds), we conclude that there is a linear function $\varphi : \mathcal{L}(\mathcal{V}_{i+1}, \mathcal{V}_i) \longrightarrow \mathcal{L}(\mathcal{V}_{j+1}, \mathcal{V}_j)$ such that

$$\varphi(A_{i\,i+1}) = A_{j\,j+1}, \tag{8}$$

for all $A \in \mathcal{A}$. Since

$$B_{jj} A_{j\,j+1} = (BA)_{j\,j+1} = \varphi((BA)_{i\,i+1}) = \varphi(B_{ii} A_{i\,i+1}) \tag{9}$$

for every $A \in \mathcal{A}$ and $B \in \mathcal{S}$, it follows from (7) that

$$T A_{j\,j+1} = \varphi(S A_{i\,i+1}),$$

for every $A \in \mathcal{A}$ and $T \in \mathcal{L}(\mathcal{V}_j)$, $S \in \mathcal{L}(\mathcal{V}_i)$. This is impossible in our setting and thus is a contradiction.

A similar argument shows that the supposition that i is unlinked from j, but $i+1$ and $j+1$ are not independent, leads to a contradiction as well. Thus we have demonstrated that i is linked to j and $i+1$ is linked to $j+1$, whenever i and j are not independent. In that case the linear function $\varphi_{ij} : \mathcal{L}(\mathcal{V}_{i+1}, \mathcal{V}_i) \longrightarrow \mathcal{L}(\mathcal{V}_{j+1}, \mathcal{V}_j)$ satisfying (8) must satisfy (9) and

$$A_{j\,j+1}C_{j+1\,j+1} = (AC)_{j\,j+1} = \varphi_{ij}((AC)_{i\,i+1}) = \varphi_{ij}(A_{i\,i+1}C_{i+1\,i+1}),$$

for every $A \in \mathcal{A}$ and $B, C \in \mathcal{S}$.

Selecting bases $\beta_1, \beta_2, \beta_3, \ldots, \beta_{\mu(\mathcal{A})}$ of $\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3, \ldots, \mathcal{V}_{\mu(\mathcal{A})}$ as described in Definition 9, and passing to matrices (with the notation of Definition 9), we see that the corresponding linear function $\hat{\varphi}_{ij} : \mathrm{M}_{n_i \times n_{i+1}} \longrightarrow \mathrm{M}_{n_i \times n_{i+1}}$, (here $n_i = n_j$ and $n_{i+1} = n_{j+1}$), must satisfy

$$\begin{aligned}
\langle B_{jj}\rangle_{\beta_j} \langle A_{j\,j+1}\rangle_{\beta_j \leftarrow \beta_{j+1}} \langle C_{j+1\,j+1}\rangle_{\beta_{j+1}} &= \langle (BAC)_{j\,j+1}\rangle_{\beta_j \leftarrow \beta_{j+1}} \\
&= \hat{\varphi}_{ij}\Big(\langle (BAC)_{i\,i+1}\rangle_{\beta_j \leftarrow \beta_{j+1}}\Big) \\
&= \hat{\varphi}_{ij}\Big(\langle B_{ii}A_{i\,i+1}C_{i+1\,i+1}\rangle_{\beta_i \leftarrow \beta_{i+1}}\Big) \\
&= \hat{\varphi}_{ij}\Big(\langle B_{ii}\rangle_{\beta_i} \langle A_{i\,i+1}\rangle_{\beta_i \leftarrow \beta_{i+1}} \langle C_{i+1\,i+1}\rangle_{\beta_{i+1}}\Big)
\end{aligned}$$

for every $A \in \mathcal{A}$ and $B, C \in \mathcal{S}$. Recalling that $\langle B_{jj}\rangle_{\beta_j} = \langle B_{ii}\rangle_{\beta_i}$ and $\langle C_{j+1\,j+1}\rangle_{\beta_{j+1}} = \langle C_{i+1\,i+1}\rangle_{\beta_{i+1}}$, because i is linked to j and $i+1$ is linked to $j+1$, by (5) we get that

$$\hat{\varphi}_{ij}(BAC) = B\hat{\varphi}_{ij}(A)\,C$$

for every $B \in \mathrm{M}_{n_i}$, $A \in \mathrm{M}_{n_i \times n_{i+1}}$ and $C \in \mathrm{M}_{n_{i+1}}$. [1] It is common knowledge (and an easy exercise to show) that the only such linear functions are the scalar multiples of the identity function on $\mathrm{M}_{n_i \times n_{i+1}}$. In particular, we now see that i is not independent of j if and only if for any bases $\beta_1, \beta_2, \beta_3, \ldots, \beta_{\mu(\mathcal{A})}$ of $\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3, \ldots, \mathcal{V}_{\mu(\mathcal{A})}$ described in Definition 9, there is a non-zero scalar α_{ij} such that

$$\alpha_{ij}\ \langle A_{i\,i+1}\rangle_{\beta_i \leftarrow \beta_{i+1}} = \langle A_{j\,j+1}\rangle_{\beta_j \leftarrow \beta_{j+1}} \tag{10}$$

for all $A \in \mathcal{A}$. We can therefore conclude that the relation "not independent of" defined (in our setting) in the obvious way on the indices $1, 2, 3, \ldots, \mu(\mathcal{A}) - 1$ is an equivalence relation.

It is now not hard to see that there are bases $\beta_1, \beta_2, \beta_3, \ldots, \beta_{\mu(\mathcal{A})}$ of $\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3, \ldots, \mathcal{V}_{\mu(\mathcal{A})}$, described in Definition 9, which also satisfy the condition

$$\langle A_{i\,i+1}\rangle_{\beta_i \leftarrow \beta_{i+1}} = \langle A_{j\,j+1}\rangle_{\beta_j \leftarrow \beta_{j+1}},$$

for every $A \in \mathcal{A}$, whenever i and j are not independent. Indeed, start with any choice of bases $\beta_1, \beta_2, \beta_3, \ldots, \beta_{\mu(\mathcal{A})}$ described in Definition 9. For each i that is the minimal index in its "not independent of" equivalence class, let $\gamma_{i\,i+1} = 1$, and for

[1] In other words, $\hat{\varphi}$ is an $(\mathrm{M}_{n_i}, \mathrm{M}_{n_{i+1}})$-bimodule map on $\mathrm{M}_{n_i \times n_{i+1}}$.

each other j in that same equivalence class define $\gamma_{_{j\,j+1}} \overset{\text{def}}{=} \alpha_{ij}$. Then it is clear that there exist scalars $\delta_1, \delta_2, \delta_3, \ldots, \delta_{_{\mu(\mathcal{A})}}$ such that

$$\gamma_{_{l\,l+1}} = \frac{\delta_{l+1}}{\delta_l},$$

for all $l \in \{1, 2, 3, \ldots, \mu(\mathcal{A}) - 1\}$. Multiplying each element of β_l by δ_l produces the required new basis $\hat{\beta}_l$.

Let us summarize: we can now assume that

$$\langle A_{i\,i+1} \rangle_{_{\beta_i \leftarrow \beta_{i+1}}} = \langle A_{j\,j+1} \rangle_{_{\beta_j \leftarrow \beta_{j+1}}},$$

for every $A \in \mathcal{A}$, whenever i and j are not independent. Furthermore

$$\left\{ \langle A_{i\,i+1} \rangle_{_{\beta_i \leftarrow \beta_{i+1}}} \;\middle|\; A \in \mathcal{A} \right\} = \mathrm{M}_{n_i \times n_{i+1}}$$

for every i, and

$$\left\{ \left(\langle A_{i\,i+1} \rangle_{_{\beta_i \leftarrow \beta_{i+1}}}, \langle A_{j\,j+1} \rangle_{_{\beta_j \leftarrow \beta_{j+1}}} \right) \;\middle|\; A \in \mathcal{A} \right\} = \mathrm{M}_{n_i \times n_{i+1}} \times \mathrm{M}_{n_j \times n_{j+1}},$$

whenever i and j are independent.

Thus

$$\left\{ \prod_{i=1}^{\mu(\mathcal{A})-1} R_{i\,i+1} \;\middle|\; R \in Rad(\mathcal{A}) \right\}$$

$$= \left\{ \prod_{i=1}^{\mu(\mathcal{A})-1} P_i \;\middle|\; P_i \in \mathrm{M}_{n_i \times n_{i+1}}, \text{ and } \; P_i = P_j \; \text{ whenever } i \text{ and } j \text{ are not independent} \right\}$$

Taking each P_i to be either $\begin{bmatrix} I \\ 0 \end{bmatrix}$ or $\begin{bmatrix} I \; 0 \end{bmatrix}$, we see that the latter set is not $\{0\}$, and so the proof is complete by (4). ■

It is certainly known to the algebraists that the use of $\mu(\mathcal{A})$ is essential in Theorem 37. For example, consider the nilpotent algebra

$$\mathcal{A} = \left\{ \begin{bmatrix} 0 & a & b & c \\ 0 & 0 & 0 & -b \\ 0 & 0 & 0 & a \\ 0 & 0 & 0 & 0 \end{bmatrix} \;\middle|\; a, b, c \in \mathrm{C} \right\}.$$

It is clear that $\mathcal{A}^3 = \{0\}$, but $\mathcal{A}^2 \neq \{0\}$ since

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \circ \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \neq 0.$$

Yet it is easy to check that $T^2 = 0$ for every $T \in \mathcal{A}$.

The proof of Theorem 37 shows the following as well.

Corollary 38. *If the underlying field is algebraically closed, and a subalgebra $\mathcal{A}$ of $\mathcal{L}(\mathcal{V})$ has an unhinged standard block-upper-triangular form with respect to the decomposition $\mathcal{V} = \mathcal{V}_1 \dot{+} \mathcal{V}_2 \dot{+} \mathcal{V}_3 \dot{+} \ldots \dot{+} \mathcal{V}_{\mu(\mathcal{A})}$, then the following are equivalent:*

(1) $\Big(Rad(\mathcal{A})\Big)^{\mu(\mathcal{A})-1} \neq \{0\}$;

(2) There is an element $R \in Rad(\mathcal{A})$ *such that* $R^{\mu(\mathcal{A})-1} \neq 0$.

(3) There exist bases $\beta_1, \beta_2, \beta_3, \ldots, \beta_{\mu(\mathcal{A})}$ *of* $\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3, \ldots, \mathcal{V}_{\mu(\mathcal{A})}$*, as described in Definition 9, which also satisfy:*

(a) $\Big\{ \langle A_{i\,i+1}\rangle_{\beta_i \leftarrow \beta_{i+1}} \;\Big|\; A \in \mathcal{A} \Big\} = \mathrm{M}_{n_i \times n_{i+1}}$ *for every* i;

(b) There exists an equivalence relation $\sim$ *on* $\{1, 2, 3, \ldots, \mu(\mathcal{A}) - 1\}$ *such that*

(i) If $i \sim j$ *then*

$$\langle A_{i\,i+1}\rangle_{\beta_i \leftarrow \beta_{i+1}} = \langle A_{j\,j+1}\rangle_{\beta_j \leftarrow \beta_{j+1}},$$

for every $A \in \mathcal{A}$*, and if* $i \neq j$ *then* i *is linked to* j *and* $i+1$ *is linked to* $j+1$ *(in the sense of Definition 9);*

(ii) If $i \not\sim j$ *then*

$$\Big\{ \Big(\langle A_{i\,i+1}\rangle_{\beta_i \leftarrow \beta_{i+1}}, \langle A_{j\,j+1}\rangle_{\beta_j \leftarrow \beta_{j+1}}\Big) \;\Big|\; A \in \mathcal{A} \Big\} = \mathrm{M}_{n_i \times n_{i+1}} \times \mathrm{M}_{n_j \times n_{j+1}}.$$

Theorem 39. *If the underlying field is algebraically closed, and* $\Big(Rad(\mathcal{A})\Big)^{\mu(\mathcal{A})-1} \neq \{0\}$ *for a subalgebra* $\mathcal{A}$ *of* $\mathcal{L}(\mathcal{V})$*, then* $\mathcal{A}$ *is unicellular, i.e. the lattice of the invariant subspaces of* $\mathcal{A}$ *is totally ordered by inclusion.*

Proof. Without loss of generality we may assume that $\mathcal{A}$ contains the identity transformation. By Corollary 30, $\mathcal{A}$ has an unhinged standard block-upper-triangular form with respect to some decomposition $\mathcal{V} = \mathcal{V}_1 \dot{+} \mathcal{V}_2 \dot{+} \mathcal{V}_3 \dot{+} \ldots \dot{+} \mathcal{V}_{\mu(\mathcal{A})}$. As usual we denote by $\mathcal{S}$ the block diagonal compression of $\mathcal{A}$.

Let us write $\mathcal{L}_i$ for $\mathcal{V}_1 \dot{+} \mathcal{V}_2 \dot{+} \mathcal{V}_3 \dot{+} \ldots \dot{+} \mathcal{V}_i$, with $i = 1, 2, 3, \ldots, \mu(\mathcal{A})$, and we shall let $\mathcal{L}_0 = \{0\}$. Then each $\mathcal{L}_i$ is an invariant subspace for $\mathcal{A}$. We claim that

$$(\{0\} =)\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3, \ldots, \mathcal{L}_{\mu(\mathcal{A})}(= \mathcal{V})$$

is the complete list of the invariant subspaces for $\mathcal{A}$; equivalently, that for each j the only invariant subspaces for $\mathcal{A}$ inside $\mathcal{L}_j$ are the spaces $\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3, \ldots, \mathcal{L}_j$. (Once this is verified the proof is complete.)

To this end, as part of a proof by contradiction, suppose that j_0 is the smallest index for which the latter claim is false. Clearly $j_0 > 1$ since $\mathcal{A}_{11}$ is an irreducible algebra. Suppose that $\mathcal{W}$ is an invariant subspace for $\mathcal{A}$ which is contained in $\mathcal{L}_{j_0}$ but which is not one of $\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3, \ldots, \mathcal{L}_{j_0}$.

Then $\mathcal{W}$ is not a subset of $\mathcal{L}_{j_0-1}$, and so $\mathcal{W}$ contains an element of the form $x + v$, where $x \in \mathcal{L}_{j_0-1}$ and $0 \neq v \in \mathcal{V}_{j_0}$. Let R be an element of $Rad(\mathcal{A})$ such that $R^{\mu(\mathcal{A})-1} \neq 0$, the existence of which is guaranteed by Theorem 37. By (4) it follows that $R_{j_0-1\;j_0} \neq 0$. Since $\mathcal{A}_{j_0 j_0} = \mathcal{L}(\mathcal{V}_{j_0})$, there is an $A \in \mathcal{S}$ such that

$$A_{j_0 j_0} v \notin kernel(R_{j_0-1\;j_0}).$$

Then

$$RA(x+v) \in \mathcal{L}_{j_0-1} \cap \mathcal{W}, \quad \text{and} \quad RA(x+v) \notin \mathcal{L}_{j_0-2},$$

so that $\mathcal{A}RA(x+v)$ is a non-zero subspace of $\mathcal{L}_{j_0-1} \cap \mathcal{W}$, and is invariant under $\mathcal{A}$, but is not contained in $\mathcal{L}_{j_0-2}$.

By the definition of j_0, it must be that $\mathcal{A}RA(x+v) = \mathcal{L}_{j_0-1}$, but since $\mathcal{A}RA(x+v)$ is a subspace of $\mathcal{L}_{j_0-1} \cap \mathcal{W}$, we must have $\mathcal{L}_{j_0-1} \subset \mathcal{W}$.

In particular then $v \in \mathcal{W}$, and thus $\mathcal{V}_{j_0} = \mathcal{S}v \subset \mathcal{W}$.

Consequently

$$\mathcal{L}_{j_0} = \mathcal{L}_{j_0-1} + \mathcal{V}_{j_0} \subset \mathcal{W} \subsetneq \mathcal{L}_{j_0},$$

which is a contradiction, and the proof is complete. ∎

4. The proof of the main result

Let us finally proceed to the proof of Theorem 21. We start by presenting some auxiliary results needed for the proof.

Lemma 40. *If* $\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$ *is diagonalizable, then so are* A *and* C. *(The converse is false.)*

Proof. This is a standard undergraduate exercise. A matrix is diagonalizable if and only if its minimal polynomial is a product of distinct linear factors; (see, for example, Theorem 6 in section 6.4 of [16]). Obviously, the minimal polynomials of A and C divide that of $\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$. ∎

Lemma 41. *Suppose that* $D \in \mathrm{M}_n(\mathrm{F})$ *is a diagonal matrix,* $\alpha \in \mathrm{F}$, *and* $v \in \mathrm{F}^n$. *Then* $\begin{bmatrix} D & v \\ 0 & \alpha \end{bmatrix}$ *is diagonalizable as a matrix in* $\mathrm{M}_{n+1}(\mathrm{F})$ *if and only if*

$$v_j = 0 \quad \textit{whenever} \quad D_{jj} = \alpha.$$

Proof. Let T be the matrix in question. After applying a similarity generated by a matrix of the form

$$\begin{bmatrix} P & 0 \\ 0 & 1 \end{bmatrix},$$

where P is a permutation, we can assume that D has a block form

$$\begin{bmatrix} D_1 & 0 \\ 0 & \alpha\mathcal{I} \end{bmatrix},$$

where α does not appear on the diagonal of D_1, and we allow a possibility that either D_1 or $\alpha\mathcal{I}$ is absent altogether. After writing

$$T = \begin{bmatrix} D_1 & 0 & v_1 \\ 0 & \alpha\mathcal{I} & v_2 \\ 0 & 0 & \alpha \end{bmatrix},$$

we invoke Lemma 40 to conclude that $\begin{bmatrix} \alpha\mathcal{I} & v_2 \\ 0 & \alpha \end{bmatrix}$ must be diagonalizable, but since α is the only characteristic value of this matrix, this can only happen if the matrix is already diagonal. Hence v_2 is either absent or zero, and the proof is complete. ∎

Theorem 42. *If $\mathcal{C}$ is an abelian collection of diagonalizable upper-triangular matrices in $\mathrm{M}_n(\mathrm{F})$, then there is an upper-triangular invertible $S \in \mathrm{M}_n(\mathrm{F})$, such that $S^{-1}AS = Diag(A)$, for all $A \in \mathcal{C}$.*

Proof. The proof is by induction on n. The result is trivially true when $n = 1$. Assume that the theorem holds for some n. We shall demonstrate its validity for $n+1$. Suppose $\mathcal{C} \subset \mathrm{M}_{n+1}(\mathrm{F})$ satisfies the hypothesis of the theorem.

The span $\mathcal{W}$ of the first n standard basis vectors $e_1, e_2, e_3, \ldots, e_n$ is a common invariant subspace for $\mathcal{C}$, and the restriction $\mathcal{C}_{11}$ of $\mathcal{C}$ to $\mathcal{W}$ is a collection satisfying the hypothesis of the theorem within $\mathrm{M}_n(\mathrm{F})$ (by Lemma 40). Therefore by the inductive assumption there exists an upper-triangular invertible $S \in \mathrm{M}_n(\mathrm{F})$ such that $S^{-1}A_{11}S = Diag(A_{11})$ for every $A_{11} \in \mathcal{C}_{11}$.

Let us express matrices in $\mathrm{M}_{n+1}(\mathrm{F})$ as block-matrices with respect to the decomposition $\mathcal{W} \dotplus span\,(e_{n+1})$ of the underlying space. Then, for every $A \in \mathcal{C}$,

$$\begin{bmatrix} S & 0 \\ 0 & 1 \end{bmatrix}^{-1} A \begin{bmatrix} S & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} D & v \\ 0 & \alpha \end{bmatrix},$$

where

$$\begin{bmatrix} D & 0 \\ 0 & \alpha \end{bmatrix} = Diag(A).$$

Therefore we can assume without loss of generality that every element A of $\mathcal{C}$ already has the form

$$\begin{bmatrix} D_A & v_A \\ 0 & \alpha_A \end{bmatrix},$$

where D_A is an invertible diagonal matrix in $\mathrm{M}_n(\mathrm{F})$ and α_A is a non-zero scalar.

Note that

$$\begin{bmatrix} \mathcal{I} & -x \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} D_A & v_A \\ 0 & \alpha_A \end{bmatrix} \begin{bmatrix} \mathcal{I} & -x \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathcal{I} & x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} D_A & v_A \\ 0 & \alpha_A \end{bmatrix} \begin{bmatrix} \mathcal{I} & -x \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} D_A & v_A - (D_A - \alpha_A\mathcal{I})\,x \\ 0 & \alpha_A \end{bmatrix}.$$

In particular, to complete the proof it is sufficient to demonstrate the existence of an x such that

$$(D_A - \alpha_A\mathcal{I})\,x = v_A \tag{11}$$

for all $A \in \mathcal{C}$. Since $D_{_A} - \alpha_{_A}\mathcal{I}$ is a diagonal matrix, this amounts to showing that, for each j, there is a scalar x_j such that (with a slight abuse of notation)

$$\frac{(v_{_A})_{_j}}{(D_{_A} - \alpha_{_A}\mathcal{I})_{_{jj}}} \in \left\{x_j\,,\ \frac{0}{0}\right\} \quad \text{for every } A \in \mathcal{C}.$$

Equivalently, we can verify that

$$(v_{_C})_{_j} = 0 \ \text{ whenever } \ (D_{_C})_{_{jj}} = \alpha_{_C}, \tag{12}$$

and that

$$(D_{_B} - \alpha_{_B}\mathcal{I})\, v_{_A} = (D_{_A} - \alpha_{_A}\mathcal{I})\, v_{_B} \quad \text{for all } A, B \in \mathcal{C}. \tag{13}$$

It is easy to see that (13) is simply the statement that $\mathcal{C}$ is abelian, and hence true. The validity of (12) is a direct consequence of Lemma 41.

■

Remark 43. It is well known (see for example Theorem 1.3.19 in [17]) that for collections of diagonalizable matrices, simultaneous diagonalizability is equivalent to the commutativity of the collection. What Theorem 42 demonstrates is that in the case of diagonalizable upper-triangular matrices the simultaneous similarity can be implemented by an upper-triangular matrix.

Since idempotents are always diagonalizable, and it is clear that matrices E_i satisfying the hypothesis of the following lemma must be idempotent, the lemma is a direct consequence of Theorem 42.

Lemma 44. *If $E_1, E_2, E_3, \ldots, E_k$ are upper-triangular matrices in $\mathrm{M}_n(\mathrm{F})$, such that*

$$\begin{cases} \sum E_i = \mathcal{I}_n \\ E_i E_j = 0 \ \ \textit{for} \ \ i \neq j, \end{cases}$$

then there is an upper-triangular invertible $S \in \mathrm{M}_n(\mathrm{F})$, such that $S^{-1}E_iS = Diag(E_i)$ for every i.

Remark 45. A standard undergraduate exercise demonstrates that the result still holds true if the hypothesis $E_iE_j = 0$ for $i \neq j$ in Lemma 44 is replaced by the hypothesis that F has characteristic zero, since for such fields the prior hypothesis can be actually recovered from the fact that the idempotents add up to the identity; (see for example Exercise 10 in section 6.6 of [16]).

We are now ready to present the proof we have been working towards.

Proof of Theorem 21. First of all, the algebra generated by $\mathcal{A}$ and $\mathcal{I}$ is still semi-simple, and by passing to that algebra we may assume without loss of generality that $\mathcal{A}$ contains $\mathcal{I}$.

Next we argue that if we can prove the theorem under a stronger hypothesis that $\mathcal{A}$ has a standard block-upper-triangular form with respect to a given decomposition of $\mathcal{V}$, then the desired theorem can be deduced with little work.

Indeed, by Theorem 13 there is a block-diagonal invertible S such that $S^{-1}\mathcal{A}S$ has a standard block-upper-triangular form with respect to the given decomposition. From the "stronger hypothesis result" we would deduce the existence of an invertible block-upper-triangular T such that

$$T^{-1}(S^{-1}\mathcal{A}S)T = BlockDiag(S^{-1}\mathcal{A}S) = S^{-1}\big(BlockDiag(\mathcal{A})\big)S$$

for every $A \in \mathcal{A}$, which shows that $\hat{T}^{-1}A\hat{T} = BlockDiag(A)$, where

$$\hat{T} = STS^{-1} = \text{block-upper-triangular},$$

as required.

So, for the rest of the proof we assume that $\mathcal{A}$ has a standard block-upper-triangular form with respect to the given decomposition of $\mathcal{V}$.

Since $\mathcal{A}$ is block-upper-triangular and semi-simple,

$$\left.\begin{array}{r} A \in \mathcal{A} \\ BlockDiag(A) = 0 \end{array}\right\} \Longrightarrow A \in Rad(\mathcal{A}) \Longrightarrow A = 0. \tag{14}$$

Since $\mathcal{A}$ has a standard block-upper-triangular form, we can consider elements $G^{<i>}$ of $\mathcal{A}$ described in Definition 9 $(i = 1, 2, \ldots, m)$. The fact that $\mathcal{I} \in \mathcal{A}$ implies the existence of $G^{<i>}$ for every i. Let $\{t_1, t_2, t_3, \ldots, t_m\}$ be a complete set of representatives of the partition $\Gamma_1, \Gamma_2, \Gamma_3, \ldots, \Gamma_m$.

Note that for $i \neq j$, $BlockDiag\left(G^{<t_i>}G^{<t_j>}\right) = 0$ and so $G^{<t_i>}G^{<t_j>} = 0$. Similarly $BlockDiag(\sum_i G^{<t_i>}) = \mathcal{I}$, so that $BlockDiag(\sum_i G^{<t_i>} - \mathcal{I}) = 0$ and thus $\sum_i G^{<t_i>} = \mathcal{I}$.

Given bases $\beta_1, \beta_2, \beta_3, \ldots, \beta_k$ of $V_1, V_2, V_3, \ldots, V_k$ respectively, each $G^{<t_i>}$ is represented by an upper-triangular matrix in $\mathrm{M}_n(\mathrm{F})$ with respect to the corresponding basis of $\mathcal{V}$. By Lemma 44 there is an upper-triangular similarity that turns to zero the off-diagonal entries of the $G^{<t_i>}$'s. Applying this similarity to $\mathcal{A}$, we may assume without loss of generality that each $G^{<t_i>}$ is block-diagonal with each diagonal block being either $\mathcal{I}$ or 0. Using the notation of Definition 9 we further observe that

$$G^{<t_i>}_{jj} = \mathcal{I} \Longleftrightarrow j \in \Gamma_i \Longleftrightarrow i \text{ and } j \text{ are linked.}$$

For the rest of the proof $G^{<p>}$ stands for $G^{<t_i>}$ such that $p \in \Gamma_i$.

Next we induct on k. The result is trivial when $k = 1$. Assuming that the theorem holds true for $k-1$ $(k \geq 2)$, let us verify its validity for k as well. The subspace $\mathcal{V} \stackrel{\text{def}}{=} \mathcal{V}_1 \dotplus V_2 \dotplus \mathcal{V}_3 \dotplus \ldots \dotplus \mathcal{V}_{k-1}$ is invariant under $\mathcal{A}$, and the restriction $\mathcal{A}_{\mathcal{V}}$ of $\mathcal{A}$ to $\mathcal{V}$ satisfies the assumption in the inductive hypothesis, so that we can conclude that there exists an invertible $T \in \mathcal{L}(\mathcal{V})$ which is block-upper-triangular with respect to the decomposition $\mathcal{V} = \mathcal{V}_1 \dotplus V_2 \dotplus \mathcal{V}_3 \dotplus \ldots \dotplus \mathcal{V}_{k-1}$ and such that

$$T^{-1}AT = BlockDiag(A)$$

for every $A \in \mathcal{A}_{\mathcal{V}}$.

The appropriate direct sum $S \in \mathcal{L}(\mathcal{V})$ of T and $\mathcal{I}_{\mathcal{V}_k}$ is block-upper-triangular with respect to the decomposition $\mathcal{V} = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \mathcal{V}_3 \dotplus \ldots \dotplus \mathcal{V}_k$, and

$$S^{-1}AS = BlockDiag(A)$$

for all $A \in \mathcal{A}$, so that after passing to $S^{-1}\mathcal{A}S$ we may assume without loss of generality that every element of $\mathcal{A}$ has a block-matrix form

$$\begin{bmatrix} A_{11} & 0 & 0 & \dots & A_{1k} \\ 0 & A_{22} & 0 & \dots & A_{2k} \\ 0 & 0 & A_{33} & \dots & A_{3k} \\ \dots & \dots & \dots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & A_{kk} \end{bmatrix} \tag{15}$$

There are two alternatives: either k is linked to some $i \in \{1, 2, \dots, k-1\}$, or it is not.

Consider the latter case first. In this case $G^{<k>}$ has the block-form

$$\begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \mathcal{I} \end{bmatrix}$$

and therefore for every $A \in \mathcal{A}$, $BlockDiag\left(AG^{<k>} - G^{<k>}A\right) = 0$, so that by (14) $AG^{<k>} = G^{<k>}A$, which demonstrates that

$$A_{1k} = A_{2k} = A_{3k} = \dots = A_{k-1\,k} = 0.$$

In other words, in this case every matrix in $\mathcal{A}$ is block-diagonal already and there is nothing left to prove.

Let us now deal with the case that k is linked to some $i \in \{1, 2, \dots, k-1\}$. To simplify notation, let us assume that an appropriate choice of bases $\beta_1, \beta_2, \beta_3, \dots, \beta_k$ of $\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3, \dots, \mathcal{V}_k$ respectively has been made, so that $\langle A_{ii}\rangle_{\beta_i} = \langle A_{kk}\rangle_{\beta_k}$ for all $A \in \mathcal{A}$ if i is linked to k. We shall now treat $\mathcal{A}$ as an algebra of block-matrices with matrix blocks, where $A_{ij} \in \mathrm{M}_{n_i \times n_j}(\mathrm{F})$ for each $A \in \mathcal{A}$. Since for i linked to k, $dim(\mathcal{V}_i) = dim(\mathcal{V}_k)$, we have $n_i = n_k$ for all such i.

Consider the subalgebra $G^{<k>}\mathcal{A}G^{<k>}$ of $\mathcal{A}$. For each $B \in G^{<k>}\mathcal{A}G^{<k>}$

$$B_{ii} = \begin{cases} B_{kk}, & \text{if } i \text{ is linked to } k \\ 0, & \text{otherwise} \end{cases}.$$

In particular, if $B_{kk} = C_{kk}$ for some $B, C \in G^{<k>}\mathcal{A}G^{<k>}$ then $BlockDiag(B-C) = 0$ and so $B = C$ by (14). It follows that for each i linked to k, there exists a linear transformation $\phi_i : \mathcal{A}_{kk} \longrightarrow \mathcal{A}_{ik}$ such that $A_{ik} = \phi_i(A_{kk})$ for every $A \in \mathcal{A}$.

Hence if $i < k$ and i is linked to k, then for all $A, B \in \mathcal{A}$:

$$\begin{aligned} \phi_i(A_{kk}B_{kk}) = \phi_i\left((AB)_{kk}\right) = (AB)_{ik} = A_{ii}B_{ik} + A_{ik}B_{kk} &= A_{kk}B_{ik} + A_{ik}B_{kk} \\ &= A_{kk}\phi_i(A_{kk}) + \phi_i(A_{kk})B_{kk}, \end{aligned}$$

which shows that for all such i, the transformation ϕ_i is a derivation.

Being an irreducible algebra, $\mathcal{A}_{kk}$ is a unital semi-simple subalgebra (see Proposition 4) of the central simple algebra $\mathrm{M}_{n_k}(\mathrm{F})$, and therefore each ϕ_i is an inner derivation on $\mathrm{M}_{n_k}(\mathrm{F})$ by Theorem 6. Hence for each i $(< k)$ linked to k there exists $S_i \in \mathrm{M}_{n_k}(\mathrm{F})$ such that $A_{ik} = \phi_i(A_{kk}) = A_{kk}S_i - S_iA_{kk}$ for all $A \in G^{<k>}\mathcal{A}G^{<k>}$.

Consider

$$T = \begin{bmatrix} \mathcal{I} & 0 & 0 & \dots & T_{1k} \\ 0 & \mathcal{I} & 0 & \dots & T_{2k} \\ 0 & 0 & \mathcal{I} & \dots & T_{3k} \\ \dots & \dots & \dots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \mathcal{I} \end{bmatrix}$$

where

$$T_{ik} = \begin{cases} S_i, & \text{if } \ i < k \ \text{ and } \ i \ \text{ is linked to } \ k \\ 0, & \text{otherwise} \end{cases}$$

Such T is invertible with

$$T^{-1} = \begin{bmatrix} \mathcal{I} & 0 & 0 & \dots & -T_{1k} \\ 0 & \mathcal{I} & 0 & \dots & -T_{2k} \\ 0 & 0 & \mathcal{I} & \dots & -T_{3k} \\ \dots & \dots & \dots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \mathcal{I} \end{bmatrix},$$

and it is easy to check that for all $A \in G^{<k>}\mathcal{A}G^{<k>}$: $TAT^{-1} = BlockDiag(A)$.

In particular $TG^{<k>}T^{-1} = G^{<k>}$ and $TG^{<k>}\mathcal{A}G^{<k>}T^{-1} = G^{<k>}T\mathcal{A}T^{-1}G^{<k>}$.

After passing to $T\mathcal{A}T^{-1}$ we may assume without loss of generality that every element of $G^{<k>}\mathcal{A}G^{<k>}$ is block-diagonal, i.e. $A_{ik} = 0$ for every $A \in \mathcal{A}$ and every i $(< k)$ linked to k.

Since we have assumed that k is linked to at least one other index, let us write i_0 for the smallest such index. The elements of $\mathcal{A}$ have the form described in (15), and therefore the subspace $\mathcal{U}$ of $\mathcal{V}$ with the direct sum decomposition

$$\mathcal{U} = \sum_{\substack{i=1 \\ i \neq i_0}}^{k} \mathcal{V}_i \tag{16}$$

is invariant under $\mathcal{A}$, and the restriction $\mathcal{A}_{\mathcal{U}}$ of $\mathcal{A}$ to $\mathcal{U}$ satisfies the assumption in the inductive hypothesis, so that we can conclude that there exists an invertible $T \in \mathcal{L}(\mathcal{U})$ which is block-upper-triangular with respect to the decomposition (16), and such that

$$T^{-1}AT = BlockDiag(A)$$

for every $A \in \mathcal{A}_{\mathcal{U}}$.

The appropriate direct sum $S \in \mathcal{L}(\mathcal{V})$ of T and $\mathcal{I}_{\mathcal{V}_{i_0}}$ is block-upper-triangular with respect to the decomposition $\mathcal{V} = \mathcal{V}_1 \dotplus \mathcal{V}_2 \dotplus \mathcal{V}_3 \dotplus \dots \dotplus \mathcal{V}_k$, and all elements of $S^{-1}\mathcal{A}S$ are block-diagonal. This completes the proof. ∎

References

[1] Livshits L, MacDonald GW, Marcoux LW, Radjavi H. Paratransitive algebras of linear operators. (To appear in) Linear Algebra and its Applications. 2013;.

[2] Livshits L, MacDonald GW, Marcoux LW, Radjavi H. Paratransitive algebras of linear operators II. (To appear in) Linear Algebra and its Applications. 2013;.

[3] Wedderburn JHM. On Hypercomplex Numbers. Proc London Math Soc. 1908;S2-6(1):77. Available from: http://dx.doi.org/10.1112/plms/s2-6.1.77.

[4] Wedderburn JHM. Lectures on matrices. American Mathematical Society Colloquium Publications. 1934;Volume XVII.

[5] Dickson LE. Algebras and their arithmetics. Bull Amer Math Soc. 1924;30(5-6):247–257. Available from: http://dx.doi.org/10.1090/S0002-9904-1924-03895-6.

[6] Jacobson N. Simple Lie algebras over a field of characteristic zero. Duke Math J. 1938;4(3):534–551. Available from: http://dx.doi.org/10.1215/S0012-7094-38-00444-2.

[7] Malcev A. On the representation of an algebra as a direct sum of the radical and a semi-simple subalgebra. C R (Doklady) Acad Sci URSS (NS). 1942;36:42–45.

[8] Watters JF. Block triangularization of algebras of matrices. Linear Algebra Appl. 1980;32:3–7. Available from: http://dx.doi.org/10.1016/0024-3795(80)90003-8.

[9] Jacobson N. Abstract derivation and Lie algebras. Trans Amer Math Soc. 1937; 42(2):206–224. Available from: http://dx.doi.org/10.2307/1989656.

[10] Kaplansky I. Fields and rings. Chicago Lectures in Mathematics. Chicago, IL: University of Chicago Press; 1995. reprint of the second (1972) edition.

[11] Farenick DR. Algebras of linear transformations. Universitext. New York: Springer-Verlag; 2001.

[12] Herstein IN. Noncommutative rings. Vol. 15 of Carus Mathematical Monographs. Washington, DC: Mathematical Association of America; 1994. reprint of the 1968 original, With an afterword by Lance W. Small. Available from: http://links.jstor.org/sici?sici=0002-9890(199308/09)100:7<698:AFCINR>2.0.CO;2-I&origin=MSN.

[13] Jacobson N. Structure of rings. American Mathematical Society, Colloquium Publications, vol. 37. 190 Hope Street, Prov., R. I.: American Mathematical Society; 1956.

[14] Jacobson N. The Theory of Rings. American Mathematical Society Mathematical Surveys, vol. I. New York: American Mathematical Society; 1943.

[15] Hungerford TW. Algebra. Vol. 73 of Graduate Texts in Mathematics. New York: Springer-Verlag; 1980. reprint of the 1974 original.

[16] Hoffman K, Kunze R. Linear algebra. Second edition. Englewood Cliffs, N.J.: Prentice-Hall Inc.; 1971.

[17] Horn RA, Johnson CR. Matrix analysis. Cambridge: Cambridge University Press; 1990. corrected reprint of the 1985 original.