J. OPERATOR THEORY
74:2(2015), 371–389
doi: 10.7900/jot.2014jul02.2060

© Copyright by THETA, 2015

NILPOTENT COMMUTATORS WITH A MASA

MITJA MASTNAK, MATJAŽ OMLADIČ and HEYDAR RADJAVI

Communicated by Kenneth R. Davidson

ABSTRACT. Let $\mathcal{H}$ be a complex Hilbert space, let $\mathcal{D} \subset \mathcal{B}(\mathcal{H})$ be a discrete masa (maximal abelian selfadjoint algebra) and let $\mathcal{A}$ be a linear subspace (or a singleton subset) of $\mathcal{B}(\mathcal{H})$ not necessarily having any nontrivial intersection with $\mathcal{D}$. Assume that the commutator $AD - DA$ is quasinilpotent for every $A \in \mathcal{A}$ and every $D \in \mathcal{D}$. We prove that $\mathcal{A}$ and $\mathcal{D}$ are simultaneously triangularizable. If $\mathcal{D}$ is a continuous masa, there exist compact operators satisfying this condition that fail to have a multiplicity-free triangularization together with $\mathcal{D}$. However, we prove an analogous result in the case where $\mathcal{A}$ is a finite-dimensional space of operators of finite rank.

KEYWORDS: *Reducibility, triangularizability, commutators, quasinilpotent operators, masa.*

MSC (2010): Primary 47A15; Secondary 28A99, 46J99, 47A66.

1. INTRODUCTION

Starting with finite dimensions, consider an upper triangular matrix A and the set $\mathcal{D}$ of all diagonal matrices. Clearly, the commutator $AD - DA$ is strictly upper triangular and thus nilpotent for every D in $\mathcal{D}$. This assertion has obvious infinite-dimensional analogues, where "nilpotent" can also be replaced with "quasinilpotent". What about the converse, that is, if for a given A, the commutators mentioned above are nilpotent, or more generally, quasinilpotent, then does it follow that A is upper triangular — after a permutation of the basis, of course? (A permutation of basis has no effect on the diagonal character of the set $\mathcal{D}$.) This is the question we are concerned with in this paper. We present an affirmative answer to this question in the discrete case and certain continuos cases. We further extend the result to simultaneous triangularizability of a space of operators whose members have quasinilpotent commutators with the members of $\mathcal{D}$.

Several authors have studied the effect of polynomial conditions on reducibility and (simultaneous) triangularizability of collections of operators in the following sense. Let $\mathcal{S}$ be a set of linear operators (usually with some structure, e.g., that of a group, semigroup or linear space). Let f be a noncommutative polynomial in two variables. One is interested in the effect of conditions such as

$$f(S,T)=0 \quad \text{for all } S \text{ and } T \text{ in } \mathcal{S}$$

or

$$\operatorname{tr} f(S,T)=0 \quad \text{for all } S \text{ and } T \text{ in } \mathcal{S}$$

on simultaneous reducibility or triangularizability of $\mathcal{S}$.

Starting with the finite-dimensional case (in which one can replace the field of complex numbers with any algebraically closed field) a well-known exercise in linear algebra is that if the polynomial $xy-yx$ vanishes on a set $\mathcal{S}$ of linear operators, then $\mathcal{S}$ is simultaneously triangularizable. Guralnick [1] proved that this hypothesis can be replaced with a much weaker one for a (multiplicative) semigroup $\mathcal{S}$, namely that $ST-TS$ is nilpotent for all S and T in $\mathcal{S}$. In particular, if $\mathcal{S}$ is a bounded group of operators on a complex vector space, then this condition implies that $\mathcal{S}$ is actually commutative (and therefore diagonalizable). For an extension of this result to infinite dimensions see [9]. For results on more general polynomials see [3] and [4].

In [5] polynomial conditions involving operators from outside the semigroup $\mathcal{S}$ were introduced. For example, if T is a fixed operator (not necessarily from the semigroup) and $f(S,T)=0$ for all S in $\mathcal{S}$, what can we conclude about the structure of invariant subspaces of $\mathcal{S}$? Even for simple polynomials, this question seems much harder than the previously mentioned kind of questions, perhaps not surprisingly.

In [5] the authors consider primarily reducibility of a semigroup satisfying a nilpotency condition with respect to an idempotent of rank one. The condition requires the order of nilpotency to be equal to 2. In general, higher orders of nilpotency do not yield reducibility, while higher ranks lead to much harder problems. Actually, in [6] the authors extend some of the results of the previous paper to the case when the rank of the idempotent P equals 2. In [7] the authors consider similar problems in a Hilbert space setting, mostly finite dimensional. They assume the "star conditions" on the semigroup, i.e. that $\mathcal{S}$ contains S^* for every $S\in\mathcal{S}$ (such semigroups are called *self-adjoint*), and also on the idempotent, i.e. $P^*=P$. In the particular case of a unitary group $\mathcal{U}$ they show that the condition $(UP-PU)^2=0$ for all $U\in\mathcal{U}$ implies that $UP-PU=0$ for all $U\in\mathcal{U}$. The main result of that paper is a structure theorem for closed self-adjoint semigroups.

In this paper we propose the study of a problem which has an opposite flavor in some sense. We start by a single operator and assume its commutativity (modulo nilpotent or at least quasinilpotent operators) with a large set of diagonal operators. We expect triangularizability of the operator under consideration in a basis in which the starting set remains diagonal. More precisely, we fix some maximal abelian self-adjoint subalgebra (masa) $\mathcal{D}$ of the algebra of bounded operators $\mathcal{B}(\mathcal{H})$ on a given Hilbert space $\mathcal{H}$. Moreover, we assume that A is a bounded operator on $\mathcal{H}$ such that $AD - DA$ is quasinilpotent for every D from the masa. We study the problem, whether it follows that A and $\mathcal{D}$ have a joint (multiplicity-free) triangularization. We are able to give a positive answer to this question in two special cases, the case of a discrete masa and the case of a continuous masa. Although the question itself is quite natural in view of the history of investigation explained above and in spite of the fact that the results do not come as a surprise, the methods we use to get them are quite involved and differ substantially from one case to the other. What does come as a surprise to us is the relation between our problem and the problem studied in [2] which concerns monotonicity of spectra of operator compressions. Although there is seemingly not much prior relations between the conditions of that paper and ours, the posterior relation in both the results and some methods used in the proofs are substantial.

In the first part of the paper (Sections 2 and 3) we focus on the case where $\mathcal{D}$ is discrete. We prove a somewhat stronger result than expected (Theorem 9): If $\mathcal{L}$ is a linear space of operators on $\mathcal{H}$ such that for every D in some fixed discrete masa we have that $DL - LD$ is quasinilpotent, then $\mathcal{L}$ and $\mathcal{D}$ have a joint triangularization. Our methods do not rely on the topology of $\mathcal{H}$ at all. A careful reading of the proofs establishes the following algebraic formulation of the theorem just mentioned:

Let $\mathcal{L}$ be a linear space of matrices of a fixed size (either finite or countably infinite) over some field k of large enough cardinality. If for every matrix $L \in \mathcal{L}$ and every diagonal matrix D with finitely many nonzero entries we have that the matrix $LD - DL$ is nilpotent, then there is a permutation of the standard basis with respect to which every matrix in $\mathcal{L}$ is upper triangular.

One of the key features of the proof is the study of (the absence of) cycles in $\mathcal{L}$. This idea comes naturally in the course of the proof. However, a similar idea also occurs in [2] and this is the first occasion where we notice similarities in the problems studied in the two papers.

In the second part of the paper we focus on the case where $\mathcal{D}$ is a continuous masa. The dependence of the problem considered and the one studied in [2] becomes here even somewhat more apparent although still far from trivial. There, the authors consider finite- and infinite-dimensional versions of the following assertion. If A is a matrix with the property that whenever P and Q are diagonal idempotents with $P \leqslant Q$, the spectrum of PAP (considered as an operator on the range of P) is contained in that of QAQ (considered as an operator on the range of Q), then there is a permutation matrix U such that $U^{-1}AU$ is triangular. In other words, if the spectrum of compressions to standard idempotents (i.e. idempotents in $\mathcal{D}$) increases with increasing idempotents, then the operator has a standard triangularization (i.e. triangularization with standard idempotents).

We show that some compact operators satisfying our condition fail to have multiplicity-free triangularization jointly with the continuous masa $\mathcal{D}$. The counterexample is borrowed from [2]. So, in Section 4 we focus on an operator of finite rank that can be written as $A = f_1 \otimes g_1 + f_2 \otimes g_2 + \cdots + f_n \otimes g_n$. For every idempotent $P \in \mathcal{D}$ we then introduce the P-matrix of A as

$$M(P) = M_A(P) = \left\{ \int_P f_i \overline{g}_j \,\mathrm{d}\mu \right\}_{i,j=1}^{n}.$$

We will usually omit A in this notation since it will be a fixed operator throughout. Here we abuse the notation by conflating the notions of the idempotent P and its support. The key observation of our proof is (Lemma 4.1): Under certain technical conditions and if N is large enough, it holds for disjoint idempotents $P_1, P_2, \dots, P_N$ that $\sum_{\pi \in S_N} M(P_{\pi 1})M(P_{\pi 2}) \cdots M(P_{\pi N}) = 0$. A careful estimate involving some matrix norms shows that the idempotents in this equality are allowed to repeat.

In Section 5 we first show using some non-trivial combinatorics that any linear combination of matrices $M(P_1), M(P_2), \dots, M(P_k)$, under the technical conditions mentioned above, is a nilpotent matrix; and then we show using some more combinatorics that the mentioned conditions may always be assumed with no loss. This brings us to Theorem 5.4 saying that the span $\mathcal{M}$ of all matrices of the form $M(P)$ for $P \in \mathcal{D}$ is made up of nilpotent matrices. As a simple corollary of this result we get that the set of all operators $PAP|_{\operatorname{ran} P}$ for $P \in \mathcal{D}$ is made up of nilpotents only, so that the spectrum of compressions to standard idempotents increases with increasing idempotents. The desired result then follows by Theorem 3.17 of [2]. Although the two conditions "spectrum of the compressions increases with increasing standard idempotents" and "the commutators with standard idempotents are nilpotent" are seemingly unrelated, they are consequently equivalent at least in the case of a finite rank operator both in the discrete and in the continuous case. A simple extension of our result for finite-dimensional linear space of such operators is given as Theorem 5.9.

2. STANDARD TRIANGULARIZATION AND NON-TRIVIAL CYCLES

In Sections 2 and 3 we assume that $\mathcal{H}$ is a separable Hilbert space, either complex or real, and that $\mathcal{D} \subset \mathcal{B}(\mathcal{H})$ is a discrete masa, i.e. a maximal abelian selfadjoint subalgebra of $\mathcal{B}(\mathcal{H})$. We write any $A \in \mathcal{B}(\mathcal{H})$ as a (possibly infinite) matrix with respect to the discrete basis in which $\mathcal{D}$ is diagonal. This basis will be called the *default basis*. Unless specified otherwise, the default basis will be assumed to have a default indexation by positive integers $\mathbb{N}$. The idempotent of rank one in $\mathcal{D}$ whose range equals the span of the i-th basis vector will be denoted by P_i. Also, a_{ij} will mean the (i,j) entry of A, $A_{i\cdot}$ will mean the i-th row of A and $A_{\cdot j}$ will mean the j-th column of A. So, loosely speaking, $A_{i\cdot}$ equals $P_i A$ and $A_{\cdot j}$ equals AP_j. Observe that a column may be viewed as a vector in $\mathcal{H}$ and a row as a functional on $\mathcal{H}$, where the interchange between rows and columns is seen as an isomorphism $\mathcal{H}^* \equiv \mathcal{H}$ as above. So, $A_{i\cdot}\, B_{\cdot j}$ equals the scalar product of the functional $P_i A$ with the vector BP_j which in turn equals the (i,j) entry of AB.

Let $\mathcal{C}$ be a collection of operators in $\mathcal{B}(\mathcal{H})$. We say that $\mathcal{C}$ admits a non-degenerate cycle (relative to an orthonormal basis $(e_n)_n$) if there exists an integer $m \geqslant 2$, operators $A^{(1)}, \dots, A^{(m)} \in \mathcal{C}$, and indices $i_1, i_2, \dots, i_m$, not all equal, not necessarily pairwise distinct, such that

$$a^{(1)}_{i_1,i_2} a^{(2)}_{i_2,i_3} \cdots a^{(m)}_{i_m,i_1} \neq 0.$$

Note that if $\mathcal{C}$ admits a non-degenerate cycle, then it admits one where the indices $i_1, \dots, i_m$ are pairwise distinct. If indices $i_1, \dots, i_m$ are pairwise distinct, then we say that the cycle is of length m. We say that an operator A admits a non-degenerate cycle if the collection $\{A\}$ admits a non-degenerate cycle.

The following is a straightforward generalization of Theorem 2.3 from [2]. We include the proof for the sake of completeness.

LEMMA 2.1. *Let $\mathcal{C}$ be a collection of operators in $\mathcal{B}(\mathcal{H})$. If $\mathcal{C}$ does not admit any non-degenerate cycles, then there is a permutation of the basis $(e_n)_n$ for which every matrix corresponding to operators from $\mathcal{C}$ is upper triangular.*

Proof. The problem is equivalent to the existence of total order $\preceq$ on $\mathbb{N}$ such that for every $A \in \mathcal{C}$ we have $a_{i,j} = 0$ whenever $i \prec j$. We shall establish such an order below. To this end, define $i \prec j$, whenever $i \neq j$ and there exist an integer $p \geqslant 2$, operators $A^{(1)}, \dots, A^{(p-1)} \in \mathcal{C}$, and indices $i_1 = j, i_2, \dots, i_{p-1}, i_p = i$, such that

$$a^{(1)}_{i_1,i_2} a^{(2)}_{i_2,i_3} \cdots a^{(p-1)}_{i_{p-1},i_p} \neq 0.$$

The relation $\preceq$ is transitive: Assume $i \preceq j$ and $j \preceq k$. With no loss of generality we assume that i, j, k are pairwise distinct. Then there exist operators $A^{(1)}, \dots, A^{(p-1)}, B^{(1)}, \dots, B^{(q-1)} \in \mathcal{C}$ and indices $k = i_1, i_2, \dots, i_p = j, j = j_1, j_2, \dots, j_q = i$ such that

$$a^{(1)}_{i_1,i_2} a^{(2)}_{i_2,i_3} \cdots a^{(p-1)}_{i_{p-1},i_p} \neq 0, \quad b^{(1)}_{j_1,j_2} b^{(2)}_{j_2,j_3} \cdots b^{(q-1)}_{j_{q-1},j_q} \neq 0.$$

Now set $m = p + q$, $A^{(p)} = B^{(1)}, \dots, A^{(p+q-2)} = B^{(q-1)}$, and $i_p = j, i_{p+1} = j_2, \dots, i_{p+q} = j_q$ and note that we get $a^{(1)}_{i_1,i_2} \cdots a^{(m-1)}_{i_{m-1},i_m} \neq 0$ proving that $i \preceq k$.

The relation $\preceq$ is anti-symmetric: Assume $i \preceq j$ and $j \preceq i$. Suppose, if possible, that $i \neq j$. Let $p, q, A^{(1)}, \dots, A^{(p-1)}, B^{(1)}, \dots, B^{(q-1)}, i_1 = i, i_2, \dots, i_p = j, j = j_1, j_2, \dots, j_q = i$ be such that

$$a^{(1)}_{i_1,i_2} a^{(2)}_{i_2,i_3} \cdots a^{(p-1)}_{i_{p-1},i_p} \neq 0, \quad b^{(1)}_{j_1,j_2} b^{(2)}_{j_2,j_3} \cdots b^{(q-1)}_{j_{q-1},j_q} \neq 0.$$

Then setting $m = p + q$, $A^{(p)} = B^{(1)}, \dots, A^{(p+q-1)} = B^{(q-1)}$, and $i_{p+1} = j = j_1, i_{p+1} = j_2, \dots, i_{p+q} = j_q$ we get a non-degenerate cycle $a^{(1)}_{i,i_2} \cdots a^{(m-1)}_{i_{m-1},i} \neq 0$ contradicting the assumption of the lemma.

Hence $\preceq$ defines a partial order on $\mathbb{N}$. Next apply Zorn's lemma to complete it to a total order.

Now note that if for $i \neq j$ we have $a_{i,j} \neq 0$, then $j \prec i$ (set $m = 2$, $A^{(1)} = A$, $i_1 = i$, $i_2 = j$). ∎

LEMMA 2.2. *Let $\mathcal{L}$ be a linear space of operators. Then $\mathcal{L}$ admits a non-degenerate cycle if and only if there is an operator $A \in \mathcal{L}$ that admits a non-degenerate cycle.*

Proof. The implication $(\Leftarrow)$ is obvious. We now prove $(\Rightarrow)$. Let $m \geqslant 2$, $A^{(1)}, \dots, A^{(m)} \in \mathcal{L}$ and let $i_1, \dots, i_m$ be indices such that

$$a^{(1)}_{i_1,i_2} a^{(2)}_{i_2,i_3} \cdots a^{(m)}_{i_m,i_1} \neq 0.$$

For various $t = (t_1, \dots, t_m)$ consider linear combinations $A^{(t)} = t_1 A^{(1)} + \cdots + t_m A^{(m)}$. Note that for a suitably chosen t we must have that

$$a^{(t)}_{i_1,i_2} \cdots a^{(t)}_{i_m,i_1} \neq 0.$$

One of the many ways to see this is described below.

We define t_i's as follows. Let $t_1 = 1$. After $t_1, \dots, t_n$ have been chosen so that $a^{(t_1,\dots,t_n,0,\dots,0)}_{i_j,i_{j+1}} \neq 0$ for $j = 1, \dots, n$, (if $n = m$, then we abbreviate $i_{m+1} = i_1$) then define t_{n+1} to be any value such that for $j = 1, \dots, n+1$ we have $t_{n+1} a^{(n+1)}_{i_j,i_{j+1}} \neq -a^{(t_1,\dots,t_n,0,\dots,0)}_{i_j,i_{j+1}}$. This will always be possible as our ground field has infinitely many (and hence it has at least $m+1$) elements. ∎

3. NULLITY OF THE CYCLES FOR OPERATORS COMMUTING WITH A MASA MODULO QUASI-NILPOTENT OPERATORS

We recall the notation introduced in the first paragraph of Section 2. Furthermore, we fix an operator $A \in \mathcal{B}(\mathcal{H})$ satisfying

$$\rho(AD - DA) = 0 \tag{3.1}$$

for every $D \in \mathcal{D}$ of finite rank, where $\mathcal{D} \subset \mathcal{B}(\mathcal{H})$ is the discrete masa under consideration. Here, ρ denotes the spectral radius of an operator.

Let us start by a couple of side results to be needed in the sequel, possibly of some independent interest. Let $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ be a decomposition of the Hilbert space $\mathcal{H}$ into subspaces $\mathcal{H}_1$ and $\mathcal{H}_2$ for which we assume that at least the first one is finite-dimensional and denote its dimension by m. Consider operators on $\mathcal{H}$ in the block partition with respect to this decomposition:

$$\begin{pmatrix} A & B \\ C & * \end{pmatrix}. \tag{3.2}$$

This means, in particular, that A may be viewed as an $m \times m$ matrix in a default basis, B as a column of m members of $\mathcal{H}_2^* \equiv \mathcal{H}_2$, C as a row of m members of $\mathcal{H}_2$, and $*$ as an operator on $\mathcal{H}_2$.

LEMMA 3.1. *Let $\mathcal{S}$ be a (multiplicative) semigroup of $m \times m$ matrices, let B be a column of m members of $\mathcal{H}_2^* \equiv \mathcal{H}_2$, and C a row of m members of $\mathcal{H}_2$ such that $BC = 0$. Then, written with respect to the decomposition* (3.2), *operators*

$$\begin{pmatrix} S & TB \\ CR & CUB \end{pmatrix},$$

where $R, S, T, U \in \mathcal{S}$, form a (multiplicative) semigroup of operators on $\mathcal{H}$.

The proof is a simple computation.

LEMMA 3.2. *For every idempotent $P \in \mathcal{D}$ of finite rank we have that*

$$PA(I-P)AP$$

is nilpotent.

Proof. Denote $B = PA(I-P)$ and $C = (I-P)AP$. By (3.1) we have that $PA - AP = B - C$ is nilpotent. However, $B^2 = C^2 = 0$ so that $(PA - AP)^2$ equals $-BC - CB$ and we get that

$$(PA - AP)^{2n} = (-BC - CB)^n = (-BC)^n + (-CB)^n$$

as $(BC)(CB) = 0 = (CB)(BC)$. So, $\mathrm{tr}(-BC)^n = -\mathrm{tr}(-CB)^n$ and by cyclic permutability of the trace also $\mathrm{tr}(-BC)^n = \mathrm{tr}(-CB)^n$. Consequently, $\mathrm{tr}(-BC)^n = 0$ for all positive integers n, hence BC is nilpotent and the lemma follows. ∎

For any operator A of finite rank denote by $\Delta_m(A)$ the sum of principal minors of A of a given order m. Note that this sum is always finite and independent of the choice of basis in which it is computed. So, we may and will compute it in our default basis with no loss. Observe that $\Delta_m(A) = 0$ for all m strictly greater than the rank of A. Note that $\Delta_1(A) = \mathrm{tr}(A)$, the trace of A. The nilpotency of a finite-rank operator N may be characterized by the condition that

$$\Delta_m(N) = 0 \quad \text{for all } m \in \mathbb{N}. \tag{3.3}$$

This is because $\Delta_m(N)$ are equal to the non-leading coefficients of the characteristic polynomial of N up to a sign.

Observe that condition (3.1), to be studied from now on in this section, does not change if we set all the diagonal entries of A equal to zero. This will be our standing hypothesis from now on with no loss of generality until we withdraw this assumption. It is the main goal of the section to prove that for any distinct indices $i_1, i_2, \ldots, i_k$, where k is no smaller than 2 and no greater than the dimension of $\mathcal{H}$, it holds that

$$a_{i_1,i_2} a_{i_2,i_3} \cdots a_{i_{k-1},i_k} a_{i_k,i_1} = 0. \tag{3.4}$$

We will show this by induction on k.

Actually, we will assume inductively at this point for a given m no smaller than 2 and no greater than the dimension of $\mathcal{H}$ that relation (3.4) is true for all indices k strictly smaller than m. Observe that in the case $m = 2$ the inductive assumption is fulfilled by the standing hypothesis.

LEMMA 3.3. *Under the inductive hypothesis for m let k be such that $2 \leqslant k < m$ and let $a_{12}\, a_{23} \cdots a_{k-1,k} \neq 0$. Then the northwest $k \times k$ corner of A is strictly upper triangular.*

Proof. The diagonal entries are zero by the standing hypothesis. We get zeros on the first subdiagonal from the overall inductive hypothesis for 2-cycles: e.g. $a_{12}\, a_{21} = 0$ and $a_{12} \neq 0$ imply that $a_{21} = 0$; similar arguments yield $a_{32} = \cdots = a_{k,k-1} = 0$. Zeros on the second subdiagonal are obtained from the standing inductive hypothesis for 3-cycles: $a_{12}\, a_{23}\, a_{31} = 0$ and $a_{12}\, a_{23} \neq 0$ imply that $a_{31} = 0$. Proceed inductively. ∎

In the following key proposition we need a technical lemma. Denote $\mathbf{x}^{(k)} := (x_1, x_2, \ldots, x_k)$ for $k \geqslant 2$ and variables $x_1, x_2, \ldots, x_k$.

LEMMA 3.4. *Define the polynomial*

$$\Pi_k(\mathbf{x}^{(k)}) := (x_1 - x_2)\,(x_2 - x_3)\,\cdots\,(x_{k-1} - x_k)\, x_1 x_k$$

in variables $\mathbf{x}^{(k)} := (x_1, x_2, \ldots, x_k)$ for $k \geqslant 2$. Then the sum of its monomials divisible by the product $x_1 x_2 \cdots x_k$ equals $\Lambda_k(\mathbf{x}^{(k)}) x_1 x_2 \cdots x_k$, where

$$\Lambda_k(\mathbf{x}^{(k)}) := \Big(\sum_{j=1}^{k} (-1)^{j-1} x_j \Big).$$

Proof. Actually, the claim amounts to saying that the mixed partial derivative of $\Pi_k(\mathbf{x}^{(k)})$ with respect to all the variables $x_1, x_2, \ldots, x_k$ equals $2\Lambda_k(\mathbf{x}^{(k)})$. The lemma is clear for $k = 2$. For $k > 2$ introduce

$$\widehat{\Pi}_k(\mathbf{x}^{(k)}) := (x_1 - x_2)\,(x_2 - x_3)\,\cdots\,(x_{k-1} - x_k) \quad \text{and}$$
$$\widehat{\Lambda}_k(\mathbf{x}^{(k)}) := x_1 + 2\Big(\sum_{j=2}^{k-1} (-1)^{j-1} x_j \Big) + (-1)^{k-1} x_k.$$

It suffices to show that

$$\Big(\prod_{j=2}^{k-1} \frac{\partial}{\partial x_j} \Big) \widehat{\Pi}_k(\mathbf{x}^{(k)}) = \widehat{\Lambda}_k(\mathbf{x}^{(k)}). \tag{3.5}$$

The proof will go by induction on k. In order to start the induction we have to show that

$$\frac{\partial}{\partial b}[(a - b)(b - c)] = a - 2b + c \tag{3.6}$$

which can be easily verified using calculus techniques. Assume now inductively that equation (3.5) holds for indices up to a certain index $k > 2$ and proceed towards the next index. First observe that, using inductive hypothesis, we have

$$\Big(\prod_{j=2}^{k-1}\frac{\partial}{\partial x_j}\Big)\widehat{\Pi}_{k+1}(\mathbf{x}^{(k+1)}) = \widehat{\Lambda}_k(\mathbf{x}^{(k)})(x_k - x_{k+1})$$

and apply the partial derivative with respect to x_k to this equation. Use equation (3.6), with $b = x_k$, $c = x_{k+1}$ and

$$(-1)^k a = x_1 + 2\Big(\sum_{j=2}^{k-1}(-1)^{j-1}x_j\Big),$$

and multiply the result by $(-1)^k$ to get the desired conclusion. ∎

PROPOSITION 3.5. *Under the inductive hypothesis for m let k be such that $2 \leqslant k < m$ and assume for some distinct indices $i_1, i_2, \ldots, i_k$ it holds that*

$$a_{i_1,i_2}a_{i_2,i_3}\cdots a_{i_{k-1},i_k} \neq 0.$$

Then it holds that $A_{i_k\cdot}\ A_{\cdot i_1} = 0$.

Proof. Recall the notation P_j for rank one idempotent in $\mathcal{D}$ whose range equals the span of the j-th basis vector. To simplify the notation assume (after a possible permutation of the basis vectors) with no loss that $i_1 = 1, i_2 = 2, \ldots, i_k = k$. By Lemma 3.3 the inductive hypothesis implies the northwest $k \times k$ corner of A is strictly upper triangular. We let $D = \sum_{j=1}^{k} x_j P_j$ for some choice of $x_1, x_2, \ldots, x_k$ and compute $\Delta_{k+1}(\mathbf{x}^{(k)}) := \Delta_{k+1}(DA - AD)$. Note that the assumption that $DA - AD$ (of finite rank) is nilpotent implies that $\Delta_{k+1}(\mathbf{x}^{(k)}) = 0$. We separate the sum of determinants $\Delta_{k+1}(DA - AD)$ corresponding to minors containing the northwest $k \times k$ corner, to be denoted by S from the rest of the sum to be denoted by R.

First, observe that a summand S_j of S is determined by an index $j > k$ such that the j-th column and j-th row (actually the according parts of them) are added to the northwest $k \times k$ corner of the matrix under consideration, i.e. $DA - AD$. To compute this determinant we observe that the only not necessarily zero entry of the first column of it is the $(k+1)$-st one and it equals $-a_{j,1}x_1$. Similarly, the only not necessarily zero entry of the k-th row of this determinant is the $(k+1)$-st one and it equals $a_{k,j}x_k$. So, the determinant can be computed easily by first expanding it in terms of the first column and then in terms of the last remaining row. What we are left with is an upper triangular determinant so that we simply multiply its diagonal entries. The result is

$$S_j = (-1)^{k-1}(x_1 - x_2)\,(x_2 - x_3)\,\cdots\,(x_{k-1} - x_k)\,x_1\,x_k a_{12}\,a_{23}\,\cdots\,a_{k-1,k}\,a_{k,j}\,a_{j,1}$$

which simplifies into

$$S_j = (-1)^{k-1}\Pi_k(\mathbf{x}^{(k)})\,C_k\,a_{k,j}\,a_{j,1},$$

where we have denoted $\mathbf{x}^{(k)}$ and for $k > 2$ $\Pi_k(\mathbf{x}^{(k)})$ as above, and

$$C_k := a_{12}\, a_{23} \,\cdots\, a_{k-1,k}.$$

Note that in the sum $S = \sum_{j>k} S_j$ the product $\Pi_k(\mathbf{x}^{(k)})\, C_k$ can be factored out so that

$$S = \Pi_k(\mathbf{x}^{(k)})\, C_k \sum_{j>k} a_{k,j}\, a_{j,1} = \Pi_k(\mathbf{x}^{(k)})\, C_k\, A_{k\cdot}\, A_{\cdot 1}. \tag{3.7}$$

It is clear that the series in this expression is convergent (even if not finite) since it represents the inner product of two members of $\mathcal{H}$.

Next, consider a term R' of R. There exists at least one index i such that $1 \leqslant i \leqslant k$ and that determinant of R' contains no i-th column nor i-th row, so that it is independent of x_i. It follows that the k-th mixed partial derivative of R' to all the variables $x_1, x_2, \ldots, x_k$ is zero. This implies that monomials of R' divisible by the product $x_1 x_2 \cdots x_k$ are all zero and the same is consequently true for R. Following Lemma 3.4, R is not divisible by $\Pi_k(\mathbf{x}^{(k)})$ if nonzero. This is contradicting the fact that $\Delta_k = R + S$ equals zero unless R and S are both zero. Now, since $C_k \neq 0$ by the assumption of the proposition, equation (3.7) implies that $A_{k\cdot}\, A_{\cdot 1} = 0$. ∎

PROPOSITION 3.6. *Under the inductive hypothesis for m let $i_1, i_2, \ldots, i_m$ be any distinct indices. Then*

$$a_{i_1,i_2}\, a_{i_2,i_3} \,\cdots\, a_{i_{m-1},i_m}\, a_{i_m,i_1} = 0.$$

Proof. Assume the contrary. As in the proof of Proposition 3.5 we assume with no loss that $i_1 = 1, i_2 = 2, \ldots, i_m = m$. Using Lemma 3.3 we will show that all entries of the northwest $m \times m$ corner of A equal zero except for the entries a_{12}, $a_{23}, \ldots, a_{m-1,m}$, and $a_{m,1}$. First, by Lemma 3.3 the principal minor $\{1, \ldots, m-1\}$ (in this proof principal minor means the corresponding submatrix of A) is strictly upper triangular and hence for $1 \leqslant j < i \leqslant m-1$ we have that $a_{ij} = 0$. The fact that the principal minor $\{2, \ldots, m\}$ is strictly upper triangular then implies that a_{m1} is the only nonzero entry below the diagonal in the northwest $m \times m$ corner. Now, suppose towards contradiction that some entry in the corner a_{ij} with $i + 2 \leqslant j$ is nonzero. Then, by Lemma 3.3 the principal minor $\{1, \ldots, i, j, \ldots, m\}$ is strictly upper triangular contradicting the fact that $a_{1,m} \neq 0$.

By the starting assumption of this proof (opposing the conclusion of the proposition), all the elements a_{12}, $a_{23}, \ldots, a_{m-1,m}$, and $a_{m,1}$ are nonzero and after multiplying the matrix A, if necessary, by a scalar we may assume with no loss that their product equals 1. Moreover, by going to a diagonal similarity, if necessary, we may further assume with no loss that $a_{12} = a_{23} = \cdots = a_{m-1,m} = a_{m,1} = 1$.

Denote by A_0 the northwest $m \times m$ corner of A, and by B, respectively C, the corresponding northeast, respectively the southwest corner of A. Note that

A_0 is a cyclic matrix of the form

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

so that $A_0^m = I$. Let $\mathcal{S}$ be the semigroup, generated by A_0 and all the diagonal $m \times m$ matrices with all the diagonal entries nonzero. Observe that $\mathcal{S}$ is actually a group so that in particular all the powers of its members are nonzero. Now, choose a $D \in \mathcal{D}$ such that the first m of its diagonal entries are nonzero and distinct, while all the remaining ones are equal to zero. Consider the 2×2 block decomposition of our matrices where the first block is made of the first m rows, respectively columns. It follows that commutators $DA - AD$ are of the form

$$DA - AD = \begin{pmatrix} S & TB \\ CR & 0 \end{pmatrix}$$

with respect to this decomposition, where $S, T, R \in \mathcal{S}$. We want to show that $BC = 0$; this will imply that all the products of commutators $DA - AD$ are of the form given in Lemma 3.1 (with $U = 0$). This will tell us, in particular, that all the powers of any $DA - AD$ of the kind are nonzero in contradiction with the fact that $DA - AD$ is nilpotent and the proposition will follow.

Therefore, it remains to show that $BC = 0$. With this goal in mind we first want to show that for every $j = 1, \ldots, k$ it holds that $A_{1\cdot}\, A_{\cdot j} = 0$. Choose $D = P_1$, compute $DA - AD$, and observe that the trace of the square of $DA - AD$ equals $2 A_{1\cdot}\, A_{\cdot 1}$ giving the desired conclusion in case $j = 1$. Next, let $j > 1$ and observe that $a_{12} \cdots a_{j-1,j} \neq 0$ so that $A_{1\cdot}\, A_{\cdot j} = 0$ by Proposition 3.5. These facts together imply that the product of the first row of B with C equals zero. Let us now consider a block-diagonal similarity that equals the direct sum of A_0^j, for some power j, and identity. When we apply the kind of similarity, it does not change the northwest corner of A and leaves the masa invariant. This similarity permutes the rows of B (and less importantly the columns of C) and it does that transitively. Apply the above consideration on the matrix obtained from A by this similarity to see that the product of any row of B with C equals zero. ∎

THEOREM 3.7. *Let $\mathcal{D} \subset \mathcal{B}(\mathcal{H})$ be a discrete* masa *on a separable Hilbert space $\mathcal{H}$ and let $A \in \mathcal{B}(\mathcal{H})$ be a bounded operator satisfying condition* (3.1), *i.e.*

$$\rho(AD - DA) = 0, \quad \text{for all } D \in \mathcal{D}.$$

Then there exists an indexation of the basis in which $\mathcal{D}$ consists of diagonal matrices and A is an upper triangular matrix.

For the proof combine Proposition 3.6 and Lemma 2.1.

At this point we may remove the standing hypothesis that the diagonal entries of A are zero. Namely if this were not so, we could first set them equal to zero, then prove the theorem and finally put the original diagonal entries back in.

Let us now prove the main result for linear spaces.

THEOREM 3.8. *Let $\mathcal{D} \subset \mathcal{B}(\mathcal{H})$ be a discrete* masa *on a separable Hilbert space $\mathcal{H}$ and let $\mathcal{L}$ be a linear space of bounded operators on $\mathcal{H}$, each $A \in \mathcal{L}$ satisfying condition* (3.1), *i.e.*

$$\rho(AD - DA) = 0, \quad \text{for all } D \in \mathcal{D}.$$

Then there exists an indexation of the basis in which $\mathcal{D}$ consists of diagonal matrices and all elements of $\mathcal{L}$ are simultaneously upper triangular matrices.

For the proof combine Proposition 3.6, Lemma 2.1, and Lemma 2.2.

4. PRELIMINARIES FOR THE CONTINUOUS CASE

Since we will rely on a result of [2], we assume that the underlying topological space S is a Hausdorff, Lindelöf, and locally compact topological space and that μ is a σ-finite, regular Borel measure. Moreover, we will assume μ to be a (purely) non-atomic measure. Note that all these assumptions are satisfied for the usual Lebesgue measure on subsets of $\mathbb{R}^n$. We assume that $\mathcal{H} = L^2(\mu)$ is a Hilbert space of (equivalence classes of) (complex or real) measurable functions on S whose absolute squares are integrable with respect to μ. It is well known that the set $\mathcal{D}$ of multiplications M_f by bounded measurable functions f on S forms a continuous masa of operators in $\mathcal{B}(\mathcal{H})$. We will assume from now on that $A \in \mathcal{B}(\mathcal{H})$ is an operator of finite rank satisfying

$$\rho(AD - DA) = 0 \tag{4.1}$$

for every $D \in \mathcal{D}$. Then the commutator in (4.1) is also of finite rank and consequently it is nilpotent. Let us write A in the form

$$A = f_1 \otimes g_1 + f_2 \otimes g_2 + \cdots + f_n \otimes g_n, \tag{4.2}$$

for some $f_j, g_j \in \mathcal{H}$ for $j = 1, 2, \ldots, n$. We may and will assume throughout that n is actually equal to the rank of A, which is equivalent to saying that both sets of vectors $\{f_j\}_{j=1}^n$ and $\{g_k\}_{k=1}^n$ are linearly independent. For any idempotent $P \in \mathcal{D}$ let us introduce the *P-matrix of A* as

$$M(P) = M_A(P) = \Big\{ \int_P f_j \overline{g_k} \, \mathrm{d}\mu \Big\}_{j,k=1}^n. \tag{4.3}$$

Since A is a fixed operator throughout, we will be omitting it. Here we abuse the notation by conflating the notions of the idempotent P and its support. Note that all these matrices are square matrices of order n.

Next, let

$$P_0, P_1, \ldots, P_N \in \mathcal{D}, \tag{4.4}$$

be disjoint idempotents whose sum equals the identity. We assume that P_0 *preserves independence* in the sense that both sets of vectors $\{P_0 f_j\}_{j=1}^n$ and $\{P_0 g_k\}_{k=1}^n$ as members of the L^2 of restrictions to the support of P_0 are still linearly independent.

LEMMA 4.1. *Let $N = 2n+1$, where $n = \operatorname{rank} A$, and let the idempotents given by* (4.4) *be such that P_0 preserves independence. Furthermore, let the P_i-matrices $M(P_i)$, $i = 1, \dots, N$, be defined as in* (4.3)*. Then, it holds that*

$$\sum_{\pi \in S_N} M(P_{\pi 1}) M(P_{\pi 2}) \cdots M(P_{\pi N}) = 0. \tag{4.5}$$

Proof. Introduce $f = x_0 P_0 + x_1 P_1 + \cdots + x_N P_N$, where $x_0, x_1, \dots, x_N$ are arbitrary unknowns, and observe that $D = M_f \in \mathcal{D}$. Clearly, $AD - DA$ is a nilpotent of rank no more than $2n$, so that its N-th power is equal to zero. Consider the coefficient of the N-th power of $AD - DA$ at the product $x_0 x_1 x_2 \cdots x_{N-1} x_N$ of the $(0,0)$ block entry with respect to the block partition given by the idempotents $P_0, \dots, P_N$. Since coefficient of $x_0 x_1 \cdots x_N$ in $(x_0 - x_1)(x_1 - x_2) \cdots (x_{N-1} - x_N)(x_N - x_0)$ is $1 + (-1)^{N+1} = 2 \neq 0$, it follows that

$$\sum_{j=1}^n \sum_{k=1}^n P_0 f_j \otimes P_0 g_k \Big[\sum_{\pi \in S_N} M(P_{\pi 1}) M(P_{\pi 2}) \cdots M(P_{\pi N}) \Big]_{jk} = 0,$$

and (4.5) follows by the fact that P_0 preserves independence. ∎

A careful reader may have noticed that in the proof above it was crucial to have N odd. However, it is easy to see that as soon as we prove the conclusion for certain N the same is true for all larger numbers. We do not explain this in detail since we do not need this fact.

We now introduce for any idempotent $P \in \mathcal{D}$ a number

$$\Lambda(P) = \sum_{j=1}^n \sum_{k=1}^n \int_P |f_j|\,|g_k| \,\mathrm{d}\mu$$

which represents a sort of upper bound of some norm of the matrix of the idempotent P. Namely, $\sum_{j=1}^n \sum_{k=1}^n |a_{jk}| \leqslant \Lambda(P)$ for $\{a_{jk}\}_{j,k=1}^n = M(P)$, and consequently there is a constant C, possibly depending on n, such that $\|M(P)\| \leqslant C\Lambda(P)$ (here $\|\cdot\|$ denotes the operator norm on matrices). The latter conclusion follows from the fact that all the matrix norms on square matrices of the same order are equivalent. Due to the fact that the measure μ has no atoms, it follows that there is for every $K \in \mathbb{N}$ a partition of P, $P = Q_1 + Q_2 + \cdots + Q_K$ such that $\Lambda(Q_i) = \frac{1}{K}\Lambda(P)$ for $i = 1, 2, \dots, K$. This implies that

$$\|M(Q_i)\| \leqslant \frac{C}{K}\Lambda(P), \quad \text{for } i = 1, 2, \dots, K. \tag{4.6}$$

LEMMA 4.2. *Let $N = 2n + 1$, where $n = \operatorname{rank} A$. Let $\{P_1, P_2, \dots, P_N\}$ be a set of standard idempotents such that either $P_i = P_j$ or $P_iP_j = 0$ for every pair of indices $1 \leqslant i, j \leqslant N$. As before denote by P_0 the complement of the span of $\{P_i\}_{i=1}^N$. Then,* (4.5) *is still true.*

Proof. We will prove this by induction on a lexicografically ordered number of repetitions. Note that in the case that there is no repetitions the lemma is true by Lemma 4.1. We would like to show that the lemma holds for all words with $r \geqslant 1$ repeating idempotents, where the first one is repeated k_1 times, the second one k_2 times and so on to the r-th one which is repeated k_r times. Here we assume with no loss of generality that $k_r \geqslant k_{r-1} \geqslant \cdots \geqslant k_1 \geqslant 2$ and that the lemma is valid inductively for all smaller indices k_j satisfying these relations and for all smaller indices r. More precisely, if we have another word with $1 \leqslant s \leqslant r$ repetitions, where the first idempotent is repeated l_1 times, the second one l_2 times and so on to the s-th one which is repeated l_s times (here again we assume with no loss of generality that $l_s \geqslant l_{s-1} \geqslant \cdots \geqslant l_1 \geqslant 2$), then the conclusion of the lemma holds for that word provided that either $s < r$ or $s = r$ and there exists $1 \leqslant i \leqslant s$ such that $l_i < k_i$ and $l_j = k_j$ for all $j > i$.

Take a word of the kind and denote by $X = M(P)$ the letter that is repeating k_r times. To simplify the notation let $k = k_r$. Furthermore, choose a permutation $\pi \in S_N$ and let w_0^π be the part of the word $w^\pi = M(P_{\pi 1})M(P_{\pi 2}) \cdots M(P_{\pi N})$ that stands to the left of the first appearance of X, by w_1^π the part of the word that stands in between the first and the second appearance of X and so on, up to w_k^π which stands to the right of the last appearance of X. If any of these words w_j^π are empty, we write $w_j^\pi = I$, the identity matrix. So, the word under consideration can be written as $w^\pi = M(P_{\pi 1})M(P_{\pi 2}) \cdots M(P_{\pi N}) = w_0^\pi X w_1^\pi X \cdots X w_k^\pi$, so that we get, when going through all the permutations, the sum of the words

$$Z = \sum_{\pi \in S_N} w_0^\pi X w_1^\pi X \cdots X w_k^\pi \tag{4.7}$$

which we want to show to be equal to zero.

Let us now find a partition of P, $P = Q_1 + Q_2 + \cdots + Q_K$, for a given $K \in \mathbb{N}$ as in the paragraph just preceding the lemma. Observe that $X = M(P) = \sum_{i=1}^{K} M(Q_i)$ and insert this into the sum Z given by (4.7). For each permutation π the corresponding summand of (4.7) develops into a sum of K^r words of the form $w_0^\pi Y_1 w_1^\pi Y_2 \cdots Y_k w_k^\pi$, where each of the Y_j's equals a certain $X_i = M(Q_i)$. Consider a given variation of K of the X_i's that are not all equal and sum up the corresponding words over all $\pi \in S_N$. By the inductive hypothesis the sum is zero, so that (4.7) reduces to

$$Z = \sum_{\pi \in S_N} \sum_{i=1}^{K} w_0^\pi X_i w_1^\pi X_i \cdots X_i w_k^\pi.$$

Next apply the estimate (4.6) to this sum to get

$$\|Z\| \leqslant \sum_{\pi \in S_N} \sum_{i=1}^{K} \|w_0^\pi\| \, \|w_1^\pi\| \cdots \|w_k^\pi\| \, \|X_i\|^k,$$

so that

$$\|Z\| \leqslant \sum_{\pi \in S_N} \|w_0^\pi\| \, \|w_1^\pi\| \cdots \|w_k^\pi\| K\Big(\frac{C}{K}\Lambda(P)\Big)^k.$$

Finally, send K to infinity to get the desired result. ∎

5. OPERATORS COMMUTING WITH A CONTINUOUS MASA MODULO NILPOTENT OPERATORS

PROPOSITION 5.1. *Let $X_i = M(P_i)$ for $i = 1, 2, \ldots, k$ be P_i-matrices, where $P_i \in \mathcal{D}$ are disjoint idempotents such that the complement P_0 of their span preserves independence. Then, any linear combination of X_i's is nilpotent.*

Proof. Any linear combination of X_i's can be written as $Y = x_1 X_1 + x_2 X_2 + \cdots + x_k X_k$, where $x_1, x_2, \ldots, x_k$ are arbitrary unknowns. In order to see that it is nilpotent let us compute the N-th power of Y, where $N \geqslant 2n + 1$ and $n = \operatorname{rank} A$. For a given variation of indices $1 \leqslant i_1 \leqslant i_2 \leqslant \cdots \leqslant i_N \leqslant k$ the coefficient of the N-th power of Y at $x_{i_1} x_{i_2} \cdots x_{i_N}$ can be written in terms of permutations $\pi \in S_n$ as

$$\frac{1}{m_1!}\frac{1}{m_2!}\cdots\frac{1}{m_k!}\sum_{\pi \in S_n} X_{i_{\pi 1}} X_{i_{\pi 2}} \cdots X_{i_{\pi N}},$$

where m_i is the number of repetitions of the index $i \in \{1, 2, \ldots, k\}$ in the given variation. However, this sum is equal to zero by Lemma 4.2. ∎

LEMMA 5.2. *For any idempotent $P \in \mathcal{D}$ that preserves independence there exists a partition $P = P_1 + P_2$ such that P_1 and P_2 each preserves independence.*

Proof. Let $\{h_1, h_2, \ldots, h_m\}$ be a maximal linearly independent subset of the union of the sets $\{f_1, f_2, \ldots, f_n\}$ and $\{g_1, g_2, \ldots, g_n\}$. It is a simple exercise in linear algebra to show that P preserves independence if the set $\{Ph_1, Ph_2, \ldots, Ph_m\}$ is linearly independent. We will prove the lemma using this observation inductively on m. The start of induction is clear since for a function h_1 which is not equal to zero a.e. on the support of P we can always find two subsets of nonzero measure such that it is not equal to zero a.e. on either of them. Assume that the set of functions $\mathcal{H} = \{h_1, h_2, \ldots, h_k, h_{k+1}\}$ is linearly independent and that the lemma is true for all sizes of $\mathcal{H}$ that are strictly smaller. Inductively find a partition $P = P_1 + P_2$ such that the set $\mathcal{H}' = \{h_1, h_2, \ldots, h_k\}$ is linearly independent on both the support of P_1 and P_2. Now if after adjoining h_{k+1} they are still linearly independent on both the supports, we are done. If not, then at least on one of them (and we may and will assume with no loss that this is the support of P_1) they are linearly dependent, so that there are scalars $\alpha_1, \alpha_2, \dots, \alpha_k$ such that $h = h_{k+1} - \sum_{i=1}^{k} \alpha_i h_i = 0$ a.e. on the support of P_1. The fact that $\mathcal{H}$ is linearly independent now yields that h is not equal to zero a.e. on the support of P_2. Apply the inductive hypothesis for the set $\mathcal{H}'$ on P_1 to find a partition $P_1 = Q_1 + Q_2$ such that the set $\mathcal{H}'$ is linearly independent on both the supports of Q_1 and Q_2. Next apply the inductive hypothesis for the set $\{h\}$ on P_2 to find a partition $P_2 = Q_3 + Q_4$ such that $\{h\}$ is not equal to zero a.e. on both the supports of Q_3 and Q_4. It follows that idempotents $R_1 = Q_1 + Q_3$ and $R_2 = Q_2 + Q_4$ form a partition $P = R_1 + R_2$ that solves the problem. ∎

Denote by $\mathcal{M}$ the span of all $M(P)$ when P runs through the idempotents of $\mathcal{D}$.

PROPOSITION 5.3. *There exists a $k \in \mathbb{N}$ and a pairwise disjoint set of idempotents $P_1, P_2, \dots, P_k \in \mathcal{D}$ such that the complement of their span preserves independence and such that the matrices $M(P_1), M(P_2), \dots, M(P_k)$ span $\mathcal{M}$.*

Proof. It is first clear that we may choose a finite number of P_i's, $P_1, P_2, \dots, P_k$ say, whose matrices would span $\mathcal{M}$ since the dimension of all square matrices of order n cannot exceed n^2. Next we show that we may assume them pairwise disjoint with no loss using the law of inclusion-exclusion. Namely, for any partition α, β of $\{1, 2, \dots, n\}$ we let

$$P_{\alpha,\beta} = \Big(\bigcap_{i\in\alpha} P_i\Big) \setminus \Big(\bigcup_{l\in\beta} P_l\Big)$$

to get a set of pairwise disjoint idempotents that span $\mathcal{M}$.

Finally, we want to show that the idempotents can be chosen so that the complement of their span preserves independence. To this end we first apply Lemma 5.2 inductively to find a partition $I = Q_1 + Q_2 + \cdots + Q_{n^2+1}$ such that each of Q_l's preserve independence. It follows that matrices of idempotents $P_{il} = P_i \cap Q_l$ for $i = 1, 2, \dots, k$ and $l = 1, 2, \dots, n^2 + 1$ span $\mathcal{M}$. Choose a maximal linearly independent subset $R_1, R_2, \dots, R_m$ among them and observe that it must still span $\mathcal{M}$ and that it can be no greater than n^2. This implies that each of the chosen idempotents R_j can belong to no more than one of Q_l's showing that there exists one of them, say Q_l, disjoint with all R_i's, so that Q_l is in the complement of the union of R_j's. ∎

THEOREM 5.4. *The linear space $\mathcal{M}$ contains only nilpotent matrices.*

Proof. By Proposition 5.3 the space is spanned by matrices that belong to pairwise disjoint idempotents whose complement preserves independence. However, by Proposition 5.1 a linear combination of matrices of the kind is always nilpotent. ∎

COROLLARY 5.5. *For any idempotent $P \in \mathcal{D}$ the compression PAP is nilpotent.*

Proof. Observe that after the sets of vectors $\{f_j\}_{j=1}^n$ and $\{g_k\}_{k=1}^n$ are chosen, then $M_{(PAP)^k}(P) = M_A(P)^k$. So, a high enough power of PAP is zero due to the fact that $M(P)$ is nilpotent, which is true by Theorem 5.4. ∎

THEOREM 5.6. *The operator A admits standard multiplicity-free triangularization.*

The proof follows from the corollary above and Theorem 3.17 of [2].

Observe that the condition of the rank to be finite is necessary in this theorem.

EXAMPLE 5.7. There exist a compact operator W on a Hilbert space such that for every $D \in \mathcal{D}$, where $\mathcal{D}$ is a continuous masa, it holds that $WD - DW$ is quasinilpotent but that W does not admit a standard multiplicity-free triangularization.

Proof. Define on the direct sum $\mathcal{H}^{(2)}$ of $\mathcal{H} = L^2[0,1]$ with itself the compact operator, defined by an operator matrix

$$W = \begin{pmatrix} V & V \\ -V & -V \end{pmatrix},$$

where V is the Volterra operator on $L^2[0,1]$. It was shown in Example 3.2 of [2] that W admits no standard multiplicity-free triangularization. Here, the standard masa $\mathcal{D}$ is generated by idempotents $D = P \oplus Q$, where P and Q run through multiplications by characteristic functions of measurable subsets of the interval $[0,1]$. We want to show that $[W,D] = WD - DW$ is always quasinilpotent. Compute

$$[W,D] = \begin{pmatrix} VP - PV & VQ - PV \\ -VP + QV & -VQ + QV \end{pmatrix},$$

and write the following estimate in the notation that treats the Hilbert space $L^2[0,1]$ as a Riesz space using the fact that V, P, $I-P$, Q, and $I-Q$ are all positive in the Riesz sense. We thus have the following relation, where $|X|$ denotes the modulus of an operator X:

$$|\,[W,D]\,| \leqslant \begin{pmatrix} VP + PV & VQ + PV \\ VP + QV & VQ + QV \end{pmatrix} \leqslant 2|W|.$$

Now, since $|W|$ is clearly quasinilpotent, so is $[W,D]$. ∎

LEMMA 5.8. *Let $\mathcal{D}$ be a continuous* masa *and let $\mathcal{C} \subseteq \mathcal{D}$ be an uncountable collection of nontrivial projections. If n is a positive integer, then the collection of products*

$$\{P_1 P_2 \cdots P_n \,:\, P_i \in \mathcal{C} \text{ for all } i\}$$

also contains an uncountable collection of nontrivial projections.

Proof. If $n = 2$ this follows from the fact that the measure on the underlying topological space is σ-finite proving that no uncountable collection of projections is disjoint. This argument extends to an inductive argument on k, where $n = 2^k$. ∎

THEOREM 5.9. *Let $\mathcal{A}$ be a finite-dimensional linear space of operators satisfying the standing hypothesis of Theorem* 5.6 *(i.e. they are of finite rank and satisfy condition* (4.1)*). Then they simultaneously admit a standard multiplicity-free triangularization.*

Proof. We first prove that every member $A \in \mathcal{A}$ has a nontrivial standard idempotent in its kernel. We will prove this by induction on the rank of A. To this end we write A in the block-form

$$A = \begin{pmatrix} A' & * \\ 0 & A'' \end{pmatrix}$$

with respect to some standard idempotent P. By the fact that A is triangularizable with respect to $\mathcal{D}$ such a P different from 0 and I always exists. Let the rank of A be 1. If we show the existence of such a P with $A' = 0$ we are done. If not, then $A'' = 0$. Choose a descending chain of P's of this kind whose intersection is trivial. If we have always $A' \neq 0$ this yields $A = 0$ in contradiction with the fact that it is of rank 1. For the inductive step it is enough to find a standard idempotent P such that the rank of A' is strictly smaller than the rank of A. If this cannot be done we have $A'' = 0$ which brings us to contradiction in a similar way.

Now we pick a basis $A_1, \ldots, A_n$ (actually we only need it to be a spanning set) for the linear space $\mathcal{A}$. We will establish simultaneous standard multiplicity free triangularization by induction on the sum s of the ranks of A_i's. If $s = 1$ the conclusion follows from Theorem 5.6. Now assume $s > 1$. We will first show that there exists a projection $P \in \mathcal{D}$ in the common kernel of all of them and consequently in the common kernel of all operators of $\mathcal{A}$. In order to see this we first take an uncountable set $\mathcal{S}$ of n-tuples of scalars every n of which form a basis for $\mathbb{C}^n$. Next consider all linear combinations of the A_i's with coefficients coming from $\mathcal{S}$. For each such combination pick a nontrivial P in $\mathcal{D}$ that is in its kernel. Now use Lemma 5.8 to get a nontrivial projection P (actually, an uncountable lot of them!) in the common kernel of some set of n linear combinations of $\mathcal{S}$, and since these combinations still span $\mathcal{A}$, P is in the common kernel of all of the A_i's.

Next, denote by P the supremum of the projections in the common kernel of all members of $\mathcal{A}$. Introduce $\widehat{\mathcal{A}} = \{\widehat{A} : A \in \mathcal{A}\}$, where $\widehat{A}$ is the compression of $A \in \mathcal{A}$ to $I - P$. Note that the linear space of all such compressions satisfies the same condition as the starting linear space did. Observe also that $\widehat{A}_1, \ldots, \widehat{A}_n$ obtained from the spanning set picked in the previous paragraph, span $\widehat{\mathcal{A}}$. Using the considerations of the previous paragraph we see that $\widehat{\mathcal{A}}$ have a standard idempotent in their joint kernel. Now, since P was maximal, we observe that in at least one A_i, the corresponding compression $\widehat{A}_i$ has strictly diminished the rank. Hence the sum of the ranks of $\widehat{A}_i$'s is strictly smaller than s and therefore by inductive hypothesis the theorem follows. ∎

Acknowledgements. Research of the first and the third author supported in part by grants from NSERC (Canada). Research of the second author supported in part by a grant from ARRS (Slovenia).

REFERENCES

[1] R.M. GURALNICK, Triangularization of sets of matrices, *Linear Multilinear Algebra* **9**(1980), 133–140.

[2] L.W. MARCOUX, M. MASTNAK, H. RADJAVI, Triangularizability of operators with increasing spectrum, *J. Funct. Anal.* **257**(2009), 3517–3540.

[3] H. RADJAVI, Polynomial conditions on operator semigroups, *J. Operator Theory* **53**(2005), 197–220.

[4] H. RADJAVI, Nilpolynomials and reducibility of operator semigroups, *J. Operator Theory* **64**(2010), 245–264.

[5] H. RADJAVI, M. OMLADIČ, Nilpotent commutators and reducibility of semigroups, *Linear Multilinear Algebra* **57**(2009), 307–317.

[6] H. RADJAVI, M. OMLADIČ, Reducibility of semigroups and nilpotent commutators with idempotents of rank two, *Ars Math. Contemp.* **3**(2010), 99–108.

[7] H. RADJAVI, M. OMLADIČ, Self-adjoint semigroups with nilpotent commutators, *Linear Algebra Appl.* **436**(2012), 2597–2603.

[8] H. RADJAVI, P. ROSENTHAL, *Simultaneous Triangularization*, Springer-Verlag, New York 2000.

[9] H. RADJAVI, P. ROSENTHAL, V. SHULMAN, Operator semigroups with quasinilpotent commutators, *Proc. Amer. Math. Soc.* **128**(2000), 2413–2420.

MITJA MASTNAK, DEPARTMENT OF MATHEMATICS, SAINT MARY'S UNIVERSITY, HALIFAX, B3H 3C3, CANADA
E-mail address: mmastnak@cs.smu.ca

MATJAŽ OMLADIČ, DEPARTMENT OF MATHEMATICS, INSTITUTE OF MATHEMATICS, PHYSICS AND MECHANICS, LJUBLJANA, SI-1000, SLOVENIA
E-mail address: matjaz@omladic.net

HEYDAR RADJAVI, DEPARTMENT OF PURE MATHEMATICS, UNIVERSITY OF WATERLOO, WATERLOO, N2L 3G1, CANADA
E-mail address: hradjavi@math.uwaterloo.ca

Received July 2, 2014; revised April 29, 2015.