CrossMark

Polynomial functions on upper triangular matrix algebras

Sophie Frisch[1]

Abstract There are two kinds of polynomial functions on matrix algebras over commutative rings: those induced by polynomials with coefficients in the algebra itself and those induced by polynomials with scalar coefficients. In the case of algebras of upper triangular matrices over a commutative ring, we characterize the former in terms of the latter (which are easier to handle because of substitution homomorphism). We conclude that the set of integer-valued polynomials with matrix coefficients on an algebra of upper triangular matrices is a ring, and that the set of null-polynomials with matrix coefficients on an algebra of upper triangular matrices is an ideal.

Keywords Integer-valued polynomials · Null polynomials · Zero polynomials · Polynomial functions · Upper triangular matrices · Matrix algebras · Polynomials on non-commutative algebras · Matrices over commutative rings

Mathematics Subject Classification Primary 13B25; Secondary 13F20, 16A42, 15A24, 05A05, 11C08, 16P10

✉ Sophie Frisch
  frisch@math.tugraz.at

[1] Institut für Mathematik A, Technische Universität Graz, Kopernikusgasse 24, 8010 Graz, Austria

1 Introduction

Polynomial functions on non-commutative algebras over commutative rings come in two flavors: one, induced by polynomials with scalar coefficients, the other, by polynomials with coefficients in the non-commutative algebra itself [1,4–9,11,12].

For the algebra of upper triangular matrices over a commutative ring, we show how polynomial functions with matrix coefficients can be described in terms of polynomial functions with scalar coefficients. In particular, we express integer-valued polynomials with matrix coefficients in terms of integer-valued polynomials with scalar coefficients. The latter have been studied extensively by Evrard, Fares and Johnson [1] and Peruginelli [6] and have been characterized as polynomials that are integer-valued together with a certain kind of divided differences.

Also, our results have a bearing on several open questions in the subject of polynomial functions on non-commutative algebras. They allow us to answer in the affirmative, in the case of algebras of upper triangular matrices, two questions (see Sect. 5) posed by Werner [12]: whether the set of null-polynomials on a non-commutative algebra forms a two-sided ideal and whether the set of integer-valued polynomials on a non-commutative algebra forms a ring. In the absence of a substitution homomorphism for polynomial functions on non-commutative rings, neither of these properties is a given.

Also, our results on polynomials on upper triangular matrices show that we may be able to describe polynomial functions induced by polynomials with matrix coefficients by polynomial functions induced by polynomials with scalar coefficients, even when the relationship is not as simple as in the case of the full matrix algebra.

Let D be a domain with quotient field K, A a finitely generated torsion-free D-algebra and $B = A \otimes_D K$. To exclude pathological cases we stipulate that $A \cap K = D$. Then the set of right integer-valued polynomials on A is defined as

$$\text{Int}_B(A) = \{f \in B[x] \mid \forall a \in A \ f(a) \in A\},$$

where $f(a)$ is defined by substitution on the right of the coefficients, $f(a) = \sum_k b_k a^k$, for $a \in A$ and $f(x) = \sum_k b_k x^k \in B[x]$. Left integer-valued polynomials are defined analogously, using left-substitution:

$$\text{Int}_B^\ell(A) = \{f \in B[x] \mid \forall a \in A \ f(a)_\ell \in A\},$$

where $f(a)_\ell = \sum_k a^k b_k$. Also, we have integer-valued polynomials on A with scalar coefficients:

$$\text{Int}_K(A) = \{f \in K[x] \mid \forall a \in A \ f(a) \in A\}.$$

For $A = M_n(D)$ and $B = M_n(K)$ we could show [4] that

$$\text{Int}_B(A) = \text{Int}_K(A) \otimes_D A.$$

Peruginelli and Werner [7] have characterized the algebras for which this relationship holds. We now take this investigation one step further and show for algebras of upper triangular matrices, where $A = T_n(D)$ and $B = T_n(K)$, that $\text{Int}_B(A)$ can be described quite nicely in terms of $\text{Int}_K(A)$ (cf. Remark 1.6, Theorem 4.2), but that the connection is not as simple as merely tensoring with A.

For a ring R, and $n \geq 1$, let $M_n(R)$ denote the ring of $n \times n$ matrices with entries in R and $T_n(R)$ the subring of upper triangular matrices, i.e., the ring consisting of $n \times n$ matrices $C = (c_{ij})$ with $c_{ij} \in R$ and

$$c_{ij} \neq 0 \Longrightarrow j \geq i.$$

Remark 1.1 Let R be a commutative ring. We will make extensive use of the ring isomorphism between the ring of polynomials in one variable with coefficients in $M_n(R)$ and the ring of $n \times n$ matrices with entries in $R[x]$:

$$\varphi \colon (M_n(R))[x] \to M_n(R[x]), \quad \sum_k (a_{ij}^{(k)})_{1 \leq i, j \leq n} \, x^k \mapsto \left(\sum_k a_{ij}^{(k)} x^k \right)_{1 \leq i, j \leq n}$$

The restriction of φ to $(T_n(R))[x]$ gives a ring isomorphism between $(T_n(R))[x]$ and $T_n(R[x])$.

Notation 1.2 For $f \in (M_n(R))[x]$, F_k denotes the k-th coefficient of f, and $f_{ij}^{(k)} \in R$ the (i, j)-th entry in F_k. When we reinterpret f as an element of $M_n(R[x])$ via the ring isomorphism of Remark 1.1, we denote the (i, j)-th entry of f by f_{ij}.

In other words,

$$f = \sum_k F_k x^k, \quad \text{where} \quad F_k = (f_{ij}^{(k)})_{1 \leq i, j \leq n}$$

and

$$\varphi(f) = (f_{ij})_{1 \leq i, j \leq n}, \quad \text{where} \quad f_{ij} = \sum_k f_{ij}^{(k)} x^k$$

Also, we write $[M]_{i\,j}$ for the (i, j)-the entry of a matrix M. In particular, $[f(C)]_{i\,j}$ is the (i, j)-the entry of $f(C)$, the result of substituting the matrix C for the variable (to the right of the coefficients) in f, and $[f(C)_\ell]_{i\,j}$ the (i, j)-th entry of $f(C)_\ell$, the result of substituting C for the variable in f to the left of the coefficients.

We will work in a setting that allows us to consider integer-valued polynomials and null-polynomials on upper triangular matrices at the same time.

From now on, R is a commutative ring, S a subring of R, and I an ideal of S.

Notation 1.3 Let R be commutative ring, S a subring of R and I an ideal of S.

$$\text{Int}_R(\text{T}_n(S), \text{T}_n(I)) = \{f \in R[x] \mid \forall C \in T_n(S) : f(C) \in T_n(I)\}$$
$$\text{Int}_{\text{T}_n(R)}(\text{T}_n(S), \text{T}_n(I)) = \{f \in (T_n(R))[x] \mid \forall C \in T_n(S) : f(C) \in T_n(I)\}$$
$$\text{Int}^\ell{}_{\text{T}_n(R)}(\text{T}_n(S), \text{T}_n(I)) = \{f \in (T_n(R))[x] \mid \forall C \in T_n(S) : f(C)_\ell \in T_n(I)\},$$

where

$$f(C) = \sum_k F_k C^k \quad \text{and} \quad f(C)_\ell = \sum_k C^k F_k \quad \text{for} \quad f = \sum_k F_k x^k.$$

Example 1.4 When D is a domain with quotient field K and we set $R = K$ and $S = I = D$, then $\text{Int}_R(\text{T}_n(S), \text{T}_n(I)) = \text{Int}_K(\text{T}_n(D))$ is the ring of integer-valued polynomials on $T_n(D)$ with coefficients in K and $\text{Int}_{\text{T}_n(R)}(\text{T}_n(S), \text{T}_n(I)) = \text{Int}_{\text{T}_n(K)}(\text{T}_n(D))$, the set of right integer-valued polynomials on $T_n(D)$ with coefficients in $(T_n(K))$. We will show that the latter set is closed under multiplication and, therefore, a ring, cf. Theorem 5.4.

Example 1.5 When R is a commutative ring and we set $S = R$ and $I = (0)$, then $\text{Int}_R(\text{T}_n(S), \text{T}_n(I)) = \text{N}_R(\text{T}_n(R))$ is the ideal of those polynomials in $R[x]$ that map every matrix in $T_n(R)$ to zero, and $\text{Int}_{\text{T}_n(R)}(\text{T}_n(S), \text{T}_n(I)) = \text{N}_{\text{T}_n(R)}(\text{T}_n(R))$ is the set of polynomials in $(T_n(R))[x]$ that map every matrix in $T_n(R)$ to zero under right substitution. We will show that the latter set is actually an ideal of $(T_n(R))[x]$, cf. Theorem 5.2.

We illustrate here in matrix form our main result on the connection between the two kinds of polynomial functions on upper triangular matrices, those induced by polynomials with matrix coefficients on one hand, and those induced by polynomials with scalar coefficients on the other hand. (For details, see Theorem 4.2.)

Remark 1.6 If we identify $\text{Int}_{\text{T}_n(R)}(\text{T}_n(S), \text{T}_n(I))$ and $\text{Int}^\ell{}_{\text{T}_n(R)}(\text{T}_n(S), \text{T}_n(I))$—a priori subsets of $(T_n(R))[x]$—with their respective images in $T_n(R[x])$ under the ring isomorphism

$$\varphi: (T_n(R))[x] \to T_n(R[x]), \quad \sum_k (f_{ij}^{(k)})_{1 \le i,j \le n}\, x^k \mapsto \left(\sum_k f_{ij}^{(k)} x^k\right)_{1 \le i,j \le n}$$

then

(1) $\text{Int}_{\text{T}_n(R)}(\text{T}_n(S), \text{T}_n(I))$

$$= \begin{pmatrix} \text{Int}_R(\text{T}_n(S), \text{T}_n(I)) & \text{Int}_R(\text{T}_{n-1}(S), \text{T}_{n-1}(I)) & \ldots & \text{Int}_R(\text{T}_2(S), \text{T}_2(I)) & \text{Int}_R(S, I) \\ 0 & \text{Int}_R(\text{T}_{n-1}(S), \text{T}_{n-1}(I)) & \ldots & \text{Int}_R(\text{T}_2(S), \text{T}_2(I)) & \text{Int}_R(S, I) \\ & & \ddots & & \\ 0 & 0 & \ldots & \text{Int}_R(\text{T}_2(S), \text{T}_2(I)) & \text{Int}_R(S, I) \\ 0 & 0 & \ldots & 0 & \text{Int}_R(S, I) \end{pmatrix}$$

(2) $\operatorname{Int}^{\ell}{}_{\mathrm{T}_n(R)}(\mathrm{T}_n(S), \mathrm{T}_n(I))$

$$
= \begin{pmatrix}
\operatorname{Int}_{\mathrm{R}}(S, I) & \operatorname{Int}_{\mathrm{R}}(S, I) & \dots & \operatorname{Int}_{\mathrm{R}}(S, I) & \operatorname{Int}_{\mathrm{R}}(S, I) \\
0 & \operatorname{Int}_{\mathrm{R}}(\mathrm{T}_2(S), \mathrm{T}_2(I)) & \dots & \operatorname{Int}_{\mathrm{R}}(\mathrm{T}_2(S), \mathrm{T}_2(I)) & \operatorname{Int}_{\mathrm{R}}(\mathrm{T}_2(S), \mathrm{T}_2(I)) \\
 & & \ddots & & \\
0 & 0 & \dots & \operatorname{Int}_{\mathrm{R}}(\mathrm{T}_{n-1}(S), \mathrm{T}_{n-1}(I)) & \operatorname{Int}_{\mathrm{R}}(\mathrm{T}_{n-1}(S), \mathrm{T}_{n-1}(I)) \\
0 & 0 & \dots & 0 & \operatorname{Int}_{\mathrm{R}}(\mathrm{T}_n(S), \mathrm{T}_n(I))
\end{pmatrix}
$$

2 Path polynomials and polynomials with scalar coefficients

We will use the combinatorial interpretation of the (i, j)-th entry in the k-th power of a matrix as a weighted sum of paths from i to j.

Consider a set V together with a subset E of $V \times V$. We may regard the pair (V, E) as a set with a binary relation or as a directed graph. Choosing the interpretation as a graph, we may associate monomials to paths and polynomials to finite sets of paths.

For our purposes, a path of length $k \geq 1$ from a to b (where $a, b \in V$) in a directed graph (V, E) is a sequence $e_1 e_2 \dots e_k$ of edges $e_i = (a_i, b_i) \in E$ such that $a_1 = a$, $b_k = b$ and $b_j = a_{j+1}$ for $1 \leq j < k$. Also, for each $a \in V$, there is a path of length 0 from a to a, which we denote by ε_a. (For $a \neq b$ there is no path of length 0 from a to b.)

We introduce a set of independent variables $X = \{x_{v\,w} \mid (v, w) \in V \times V\}$ and consider the polynomial ring

$$
R[X] = R[\{x_{v\,w} \mid (v, w) \in V \times V\}]
$$

with coefficients in the commutative ring R.

To each edge $e = (a, b)$ in E, we associate the variable x_{ab} and to each path $e_1 e_2 \dots e_k$ of length k, with $e_i = (a_i, b_i)$, the monomial of degree k which is the product of the variables associated to the edges of the path: $x_{a_1 b_1} x_{a_2 b_2} \dots x_{a_k b_k}$. To each path of length 0 we associate the monomial 1.

If E is finite, or, more generally, if for any pair of vertices $a, b \in V$ and fixed $k \geq 0$ there are only finitely many paths in (V, E) from a to b, we define the k-th path polynomial from a to b, denoted by $p_{ab}^{(k)}$, as the sum in $R[X]$ of the monomials corresponding to paths of length k from a to b. If there is no path of length k from a to b in (V, E), we set $p_{ab}^{(k)} = 0$.

From now on, we fix the relation $(\mathbb{N}, \leq)$ and all path polynomials will refer to the graph of $(\mathbb{N}, \leq)$, where $\mathbb{N} = \{1, 2, 3, \dots\}$, or one of its finite subgraphs given by intervals. The finite subgraph given by the interval

$$
[i, j] = \{k \in \mathbb{N} \mid i \leq k \leq j\}
$$

is the graph with set of vertices $[i, j]$ and set of edges $\{(a, b) \mid i \leq a \leq b \leq j\}$.

Because of the transitivity of the relation "$\leq$", a path in $(\mathbb{N}, \leq)$ from a to b involves only vertices in the interval $[a, b]$. The path polynomial $p_{ab}^{(k)}$, therefore, is the same whether we consider a, b as vertices in the graph $(\mathbb{N}, \leq)$, or any subgraph given by

an interval $[i, j]$ with $a, b \in [i, j]$. So we may suppress all references to intervals and subgraphs and define:

Definition 2.1 Let R be a commutative ring. The k-th path polynomial from i to j (corresponding to the relation $(\mathbb{N}, \leq)$) in $R[x]$ is defined by

(1) For $1 \leq i \leq j$ and $k > 0$ by

$$p_{ij}^{(k)} = \sum_{i=i_1 \leq i_2 \leq \ldots \leq i_{k+1}=j} x_{i_1 i_2} x_{i_2 i_3} \ldots x_{i_{k-1} i_k} x_{i_k i_{k+1}},$$

(2) for $1 \leq i \leq j$ and $k = 0$ by $p_{ij}^{(0)} = \delta_{ij}$,

(3) for $i > j$ and all k: $p_{ij}^{(k)} = 0$.

For $a, b \in \mathbb{N}$, we define the sequence of path polynomials from a to b as

$$p_{ab} = (p_{ab}^{(k)})_{k \geq 0}.$$

Remark 2.2 Note that $p_{ij}^{(k)}$ is the (i, j)-th entry of the k-th power of a generic upper triangular $n \times n$ matrix (with $n \geq i, j$) whose (i, j)-th entry is x_{ij} when $i \leq j$ and zero otherwise.

Example 2.3 The sequence of path polynomials from 2 to 4 is

$$p_{24} = (0, \quad x_{24}, \quad x_{22}x_{24} + x_{23}x_{34} + x_{24}x_{44}, \quad \ldots)$$

and

$$p_{24}^{(3)} = x_{22}x_{22}x_{24} + x_{22}x_{23}x_{34} + x_{22}x_{24}x_{44} + x_{23}x_{33}x_{34} + x_{23}x_{34}x_{44} + x_{24}x_{44}x_{44}$$

Again, note that p_{24} is the sequence of entries in position $(2, 4)$ in the powers $G^0, G, G^2, G^3, \ldots$ of a generic $n \times n$ (with $n \geq 4$) upper triangular matrix $G = (g_{ij})$ with $g_{ij} = x_{ij}$ for $i \leq j$ and $g_{ij} = 0$ otherwise.

In addition to right and left substitution of a matrix for the variable in a polynomial in $R[x]$ or $(T_n(R))[x]$, we are going to use another way of plugging matrices into polynomials, namely, into polynomials in $R[X] = R[\{x_{ij} \mid i, j \in \mathbb{N}\}]$. For this purpose, the matrix $C = (c_{ij}) \in M_n(R)$ is regarded as a vector of elements of R indexed by $\mathbb{N} \times \mathbb{N}$, with $c_{ij} = 0$ for $i > n$ or $j > n$:

Definition 2.4 For a polynomial $p \in R[X] = R[\{x_{ij} \mid i, j \in \mathbb{N}\}]$ and a matrix $C = (c_{ij}) \in M_n(R)$ we define $p(C)$ as the result of substituting c_{ij} for those x_{ij} in p with $i, j \leq n$ and substituting 0 for all x_{kh} with $k > n$ or $h > n$.

To be able to describe the (i, j)-th entry in $f(C)$, where $f \in R[x]$, we need one more construction: for sequences of polynomials $p = (p_i)_{i \geq 0}, q = (q_i)_{i \geq 0}$ in $R[X]$, at least one of which is finite, we define a scalar product $\langle p, q \rangle = \sum_i p_i q_i$. Actually,

we only need one special instance of this construction, that where one of the sequences is the sequence of coefficients of a polynomial in $R[x]$ and the other a sequence of path polynomials from a to b.

Definition 2.5 Given $f = f_1 + f_1 x + \cdots f_m x^m \in R[x]$ (which we identify with the sequence of its coefficients), a, $b \in \mathbb{N}$, and $p_{ab} = (p_{ab}^{(k)})_{k=0}^{\infty}$ the sequence of path polynomials from a to b as in Definition 2.1, we define

$$\langle f, p_{ab} \rangle = \sum_{k \geq 0} f_k \, p_{ab}^{(k)}.$$

Definition 2.6 For a polynomial $p \in R[X] = R[\{x_{ij} \mid i, j \in \mathbb{N}\}]$ and $S \subseteq R$, we define the image $p(S^*) \subseteq R$ as the set of values of p as the variables occurring in p range through S independently. (The star in S^* serves to remind us that the arguments of p are not elements of S, but k-tuples of elements of S for unspecified k.)

We define $\mathrm{Int}(S^*, I)$ as the set of those polynomials in $R[X] = R[\{x_{ij} \mid i, j \in \mathbb{N}\}]$ that take values in I whenever elements of S are substituted for the variables.

The notation $\mathrm{Int}(S^*, I)$ is suggested by the convention that $\mathrm{Int}(S, I)$ consists of polynomials in one indeterminate mapping elements of S to elements of I and, for $k \in \mathbb{N}$, $\mathrm{Int}(S^k, I)$ consists of polynomials in k indeterminates mapping k-tuples of elements of S to I.

We summarize here the connection between path polynomials and the related constructions of Definitions 2.1, 2.4, 2.5 and 2.6 with entries of powers of matrices and entries of the image of a matrix under a polynomial function.

Remark 2.7 Let R be a commutative ring, $C \in \mathrm{T}_n(R)$, $k \geq 0$, $1 \leq i, j \leq n$, and $p_{ij}^{(k)}$ the k-the path polynomial from i to j in $R[x]$ as in Definition 2.1.

(1) $\left[C^k \right]_{ij} = p_{ij}^{(k)}(C)$
(2) For $f \in R[x]$, $[f(C)]_{ij} = \langle f, p_{ij} \rangle(C)$.
(3) If the i-th row or the j-th column of C is zero then $p_{ij}^{(k)}(C) = 0$, and for all $f \in R[x]$, $\langle f, p_{ij} \rangle(C) = 0$.
(4) $p_{ij}^{(k)}(S^*) = \{ p_{ij}^{(k)}(C) \mid C \in \mathrm{T}_n(S) \}$.

Proof (1) and (2) follow immediately from Definitions 2.1, 2.4 and 2.5. Compare Remark 2.2. (3) follows from (2) and Definition 2.4, since every monomial occurring in $p_{ij}^{(k)}$ involves a variable x_{im} for some m and a variable x_{hj} for some h. Also, (4) follows from Definitions 2.1 and 2.4.

Lemma 2.8 *Let $f \in R[x]$. The image of $\langle f, p_{ij} \rangle$ under substitution of elements of S for the variables depends only on f and $j - i$, that is, for all $i \leq j$ and all $m \in \mathbb{N}$*

$$\langle f, p_{ij} \rangle(S^*) = \langle f, p_{i+m \, j+m} \rangle(S^*)$$

Proof The R-algebra isomorphism

$$\psi \colon R[\{x_{hk} \mid i \leq h \leq k \leq j\}] \to R[\{x_{hk} \mid i + m \leq h \leq k \leq j + m\}]$$

with $\psi(x_{hk}) = \psi(x_{h+m\,k+m})$ and $\psi(r) = r$ for all $r \in R$ maps $\langle f, p_{ij} \rangle$ to $\langle f, p_{i+m\,j+m} \rangle$.

Applying ψ amounts to a renaming of variables; it doesn't affect the image of the polynomial function resulting from substituting elements of S for the variables.   □

Proposition 2.9 *Let $f \in R[x]$. The following are equivalent*

(1) $f \in \mathrm{Int}_R(T_n(S), T_n(I))$
(2) $\forall\, 1 \le i \le j \le n$ $\langle f, p_{ij} \rangle \in \mathrm{Int}(S^*, I)$
(3) $\forall\, 0 \le k \le n - 1$ $\exists i \in \mathbb{N}$ $\langle f, p_{i\,i+k} \rangle \in \mathrm{Int}(S^*, I)$

Proof The (i, j)-th entry of $f(C)$, for $C \in T_n(R)$, is $\langle f, p_{ij} \rangle(C)$, by Remark 2.7 (2). If C varies through $T_n(S)$, then all variables occurring in $\langle f, p_{ij} \rangle$ vary through S independently. This shows the equivalence of (1) and (2).

By Lemma 2.8, the image of $\langle f, p_{ij} \rangle$ as the variables range through S depends only on f and $j - i$. This shows the equivalence of (2) and (3).   □

3 Lemmata for polynomials with matrix coefficients

Notation 3.1 Given $C \in M_n(R)$ and $1 \le h \le j \le n$, let $C^{[h,j]}$ be the matrix obtained from C by replacing all entries with row- or column-index outside the interval $[h, j] = \{m \in \mathbb{N} \mid h \le m \le j\}$ by zeros; and for $S \subseteq R$, let

$$T_n^{[h,j]}(S) = \{C^{[h,j]} \mid C \in T_n(S)\}.$$

Remark 3.2 Note that, for $f \in R[x]$ and $C \in T_n(R)$,

$$[f(C)]_{ij} = \langle f, p_{ij} \rangle(C) = \langle f, p_{ij} \rangle(C^{[i,j]}) = [f(C^{[i,j]})]_{ij}.$$

This is so because no variables other than x_{st} with $i \le s \le t \le j$ occur in $p_{ij}^{(k)}$, for any k.

We derive some technical, but useful, formulas for the (i, j)-th entry in $f(C)$ and $f(C)_\ell$, respectively, where $f \in (T_n(R))[x]$ and $C \in T_n(R)$.

Lemma 3.3 *Let $f \in (T_n(R))[x]$ and $C \in T_n(R)$ with notation as in 1.2 and in Definitions 2.1, 2.4, and 2.5. Then, for all $1 \le i \le j \le n$, we have*

(R)

$$[f(C)]_{i\,j} = \sum_{h \in [i,j]} [f_{ih}(C)]_{h\,j} = \sum_{h \in [i,j]} [f_{ih}(C^{[h,j]})]_{h\,j}$$

and also

$$[f(C)]_{i\,j} = \sum_{h \in [i,j]} \langle f_{ih}, p_{hj} \rangle(C) = \sum_{h \in [i,j]} \langle f_{ih}, p_{hj} \rangle(C^{[h,j]}).$$

(L)

$$[f(C)_\ell]_{i\,j} = \sum_{h\in[i,j]} [f_{hj}(C)]_{i\,h} = \sum_{h\in[i,j]} [f_{hj}(C^{[i,h]})]_{i\,h}$$

and also

$$[f(C)_\ell]_{i\,j} = \sum_{h\in[i,j]} \langle f_{hj}, p_{ih}\rangle(C) = \sum_{h\in[i,j]} \langle f_{hj}, p_{ih}\rangle(C^{[i,h]}).$$

Proof Let $d = \deg f$.

$$[f(C)]_{i\,j} = \sum_{k=0}^{d}[F_k C^k]_{i\,j} = \sum_{k=0}^{d}\sum_{h\in[1,n]} f_{ih}^{(k)}[C^k]_{h\,j} = \sum_{k=0}^{d}\sum_{h\in[1,n]} f_{ih}^{(k)} p_{hj}^{(k)}(C).$$

Changing summation, we get $\sum_{h\in[1,n]}\sum_{k=0}^{d} f_{ih}^{(k)} p_{hj}^{(k)}(C)$.

Considering that $f_{ih}^{(k)} \neq 0$ only if $i \le h$, and that $p_{hj}^{(k)} \neq 0$ only if $h \le j$, we can restrict summation to $h \in [i, j]$. Consequently, $[f(C)]_{i\,j}$ equals

$$\sum_{h\in[i,j]}\sum_{k=0}^{d} f_{ih}^{(k)} p_{hj}^{(k)}(C) = \sum_{h\in[i,j]} \langle f_{ih}, p_{hj}\rangle(C) = \sum_{h\in[i,j]} [f_{ih}(C)]_{h\,j}.$$

By Remark 3.2, we can replace C by $C^{[h,j]}$, the matrix obtained from C by replacing all entries with row or column index outside the interval $[h, j]$ by zeros. Therefore,

$$[f(C)]_{i\,j} = \sum_{h\in[i,j]} \langle f_{ih}, p_{hj}\rangle(C^{[h,j]}) = \sum_{h\in[i,j]} \left[f_{ih}(C^{[h,j]})\right]_{h\,j}.$$

This shows the formulas for right substitution.

Now, if we substitute C for the variable of f to the left of the coefficients,

$$[f(C)_\ell]_{i\,j} = \sum_{k=0}^{d}[C^k F_k]_{i\,j} = \sum_{k=0}^{d}\sum_{h\in[1,n]} [C^k]_{i\,h} f_{hj}^{(k)} = \sum_{k=0}^{d}\sum_{h\in[1,n]} p_{ih}^{(k)}(C) f_{hj}^{(k)}.$$

Changing summation, we get

$$\sum_{h\in[1,n]}\sum_{k=0}^{d} p_{ih}^{(k)}(C) f_{hj}^{(k)} = \sum_{h\in[1,n]}\sum_{k=0}^{d} f_{hj}^{(k)} p_{ih}^{(k)}(C).$$

Considering that $f_{hj}^{(k)} \neq 0$ only if $h \leq j$, and that $p_{ih}^{(k)} \neq 0$ only if $i \leq h$, we can restrict summation to $h \in [i, j]$. Consequently, $[f(C)_\ell]_{i\,j}$ equals

$$\sum_{h \in [i,j]} \sum_{k=0}^{d} f_{hj}^{(k)} p_{ih}^{(k)}(C) = \sum_{h \in [i,j]} \langle f_{hj}, p_{ih} \rangle(C) = \sum_{h \in [i,j]} [f_{hj}(C)]_{i\,h}.$$

By Remark 3.2, we can replace C by $C^{[i,h]}$, the matrix obtained from C by replacing all entries in rows and columns with index outside $[i, h]$ by zeros. Therefore,

$$[f(C)_\ell]_{i\,j} = \sum_{h \in [i,j]} \langle f_{hj}, p_{ih} \rangle(C^{[i,h]}) = \sum_{h \in [i,j]} [f_{hj}(C^{[i,h]})]_{i\,h}.$$

□

Lemma 3.4 *Let $f \in (T_n(R))[x]$ with notation as in 1.2 and in Definitions 2.1, 2.4, and 2.5. Let $1 \leq i \leq j \leq n$.*

[right:] *The following are equivalent*
    (1) $[f(C)]_{i\,j} \in I$ *for all $C \in T_n(S)$*
    (2) $[f_{ih}(C)]_{h\,j} \in I$ *for all $C \in T_n(S)$, for all $h \in [i, j]$.*
    (3) $\langle f_{ih}, p_{hj} \rangle \in \text{Int}_R(S^*, I)$ *for all $h \in [i, j]$.*
[left:] *The following are equivalent*
    (1) $[f(C)_\ell]_{i\,j} \in I$ *for all $C \in T_n(S)$*
    (2) $[f_{hj}(C)]_{i\,h} \in I$ *for all $C \in T_n(S)$, for all $h \in [i, j]$.*
    (3) $\langle f_{hj}, p_{ih} \rangle \in \text{Int}_R(S^*, I)$ *for all $h \in [i, j]$.*

Proof For **right** substitution: $(2 \Rightarrow 1)$ follows directly from

$$[f(C)]_{i\,j} = \sum_{h \in [i,j]} [f_{ih}(C)]_{h\,j},$$

which is Lemma 3.3.

$(1 \Rightarrow 2)$ Induction on h from j down to i. Given $m \in [i, j]$, we show (2) for $h = m$, assuming that the statement holds for all values $h \in [i, j]$ with $h > m$. In the above formula from Lemma 3.3, we let C vary through $T_n^{[m,j]}(S)$ (as in Notation 3.1). For such a C, the summands $[f_{ih}(C)]_{h\,j}$ with $h < m$ are zero, by Remark 2.7 (3). The summands with $h > m$ are in I, by induction hypothesis. Therefore $[f(C)]_{i\,j} \in I$ for all $C \in T_n(S)$ implies $[f_{im}(C)]_{m\,j} \in I$ for all $C \in T_n^{[m,j]}(S)$. Since $[f_{im}(C)]_{m\,j} = [f_{im}(C^{[m,j]})]_{m\,j}$, by Remark 3.2, the statement follows for all $C \in T_n(S)$.

Finally, $(3 \Leftrightarrow 2)$ because, firstly, $[f_{ih}(C)]_{h\,j} = \langle f_{ih}, p_{hj} \rangle(C)$, for all $C \in T_n(R)$, and secondly, as far as the image of $\langle f_{ih}, p_{hj} \rangle$ is concerned, all possible values under substitution of elements from S for the variables are obtained as C varies through $T_n(S)$, because no variables other that x_{ij} with $1 \leq i \leq j \leq n$ occur in this polynomial.

For **left** substitution: $(2 \Rightarrow 1)$ follows directly from

$$[f(C)_\ell]_{i\,j} = \sum_{h \in [i,j]} [f_{hj}(C)]_{i\,h},$$

which is Lemma 3.3.

$(1 \Rightarrow 2)$ Induction on h, from $h = i$ to $h = j$. Given $m \in [i, j]$, we show (2) for $h = m$ under the hypothesis that it holds for all $h \in [i, j]$ with $h < m$. In the above formula from Lemma 3.3, the summands corresponding to $h < m$ are in I by induction hypothesis. If we let C vary through $T_n^{[i,m]}(S)$ (as in 3.1), then, for such C, the summands $[f_{hj}(C)]_{i\,h} = \langle f_{hj}, p_{ih} \rangle (C)$ corresponding to $h > m$ are zero, by Remark 2.7 (3). Therefore $[f_{mj}(C^{[i,m]})]_{i\,m} \in I$ for all $C \in T_n(S)$. But $[f_{mj}(C^{[i,m]})]_{i\,m} = [f_{mj}(C)]_{i\,m}$ for all $C \in T_n(S)$, by Remark 3.2. Thus, (2) follows by induction.

Finally, $(3 \Leftrightarrow 2)$ because, firstly, $[f_{hj}(C)]_{i\,h} = \langle f_{hj}, p_{ih} \rangle (C)$, for all $C \in T_n(R)$, and secondly, as far as the image of $\langle f_{hj}, p_{ih} \rangle$ is concerned, all possible values under substitution of elements of S for the variables are realized as C varies through $T_n(S)$, since no variables other than x_{ij} with $1 \leq i \leq j \leq n$ occur in this polynomial. □

4 Results for polynomials with matrix coefficients

Considering how matrix multiplication works, it is no surprise that, for a fixed row-index i and $f \in (T_n(R))[x]$, whether the entries of the i-th row of $f(C)$ are in I for every $C \in T_n(S)$, depends only on the i-th rows of the coefficients of f. Indeed, if $f = \sum_k F_k x^k$, and $[B]_i$ denotes the i-th row of a matrix B, then

$$[f(C)]_i = \sum_k [F_k]_i C^k.$$

Likewise, for a fixed column index j, whether the entries of the j-th column of $f(C)_\ell$ are in I for every $C \in T_n(S)$ depends only on the j-th columns of the coefficients of $f = \sum_k F_k x^k$: if $[B]^j$ denotes the j-th column of a matrix B, then

$$[f(C)_\ell]^j = \sum_k C^k [F_k]^j.$$

Now we can formulate the precise criterion which the i-th rows, or j-th columns, of the coefficients of $f \in (T_n(R))[x]$ have to satisfy to guarantee that the entries of the i-th row of $f(C)$, or the j-th column of $f(C)_\ell$, respectively, are in I for every $C \in T_n(S)$.

Lemma 4.1 *Let $f \in (T_n(R))[x]$. We use the notation of 1.2 and Definitions 2.1. 2.4, and 2.5.*

[right:] *Let $1 \leq i \leq n$. The following are equivalent*
  (1) For every $C \in T_n(S)$, all entries of the i-th row of $f(C)$ are in I.

(2) *For all $C \in T_n(S)$, for all h, j with $i \leq h \leq j \leq n$, $[f_{ih}(C)]_{hj} \in I$.*
(3) *For all h, j with $i \leq h \leq j \leq n$, $\langle f_{ih}, p_{hj} \rangle \in \text{Int}_R(S^*, I)$.*
(4) *$f_{ih} \in \text{Int}_R(T_{n-h+1}(S), T_{n-h+1}(I))$ for $h = i, \ldots, n$.*

[left:] *Let $1 \leq j \leq n$. The following are equivalent*
(1) *For every $C \in T_n(S)$, all entries of the j-th column of $f(C)_\ell$ are in I.*
(2) *For all $C \in T_n(S)$, for all i, h with $1 \leq i \leq h \leq j$, $[f_{hj}(C)]_{ih} \in I$.*
(3) *For all i, h with $1 \leq i \leq h \leq j$, $\langle f_{hj}, p_{ih} \rangle \in \text{Int}_R(S^*, I)$.*
(4) *$f_{hj} \in \text{Int}_R(T_h(S), T_h(I))$ for $h = 1, \ldots, j$.*

Proof For **right** substitution: by Lemma 3.4, (1–3) are equivalent conditions for the (i, j)-th entry of $f(C)$ to be in I for all j with $i \leq j \leq n$, for every $C \in T_n(S)$.

(3⇒4) For each h with $i \leq h \leq n$, letting j vary from h to n shows that criterion (3) of Proposition 2.9 for $f_{ih} \in \text{Int}_R(T_{n-h+1}(S), T_{n-h+1}(I))$ is satisfied.

(4⇒3) For each fixed h, f_{ih} satisfies, by criterion (2) of Proposition 2.9, in particular, $\langle f_{ih}, p_{1k} \rangle \in \text{Int}(S^*, I)$ for all $1 \leq k \leq n-h+1$. By Lemma 2.8, $\langle f_{ih}, p_{1k} \rangle(S^*)$ equals $\langle f_{ih}, p_{h\,h+k-1} \rangle(S^*)$. Therefore $\langle f_{ih}, p_{hj} \rangle \in \text{Int}(S^*, I)$ for all j with $h \leq j \leq n$.

For **left** substitution: by Lemma 3.4, (1-3) are equivalent conditions for the (i, j)-th entry of $f(C)_\ell$ to be in I for all i with $1 \leq i \leq j$, for every $C \in T_n(S)$.

(3⇒4) For each h with $1 \leq h \leq j$, letting i range from 1 to h shows that criterion (3) of Proposition 2.9 for $f_{hj} \in \text{Int}_R(T_h(S), T_h(I))$ is satisfied.

(4⇒3) For each h with $1 \leq h \leq j$, applying criterion (2) of Proposition 2.9 to f_{hj} shows, in particular, $\langle f_{hj}, p_{ih} \rangle \in \text{Int}(S^*, I)$ for all $1 \leq i \leq h$. □

We are now ready to prove the promised characterization of polynomials with matrix coefficients in terms of polynomials with scalar coefficients:

Theorem 4.2 *Let $f \in (T_n(R))[x]$. Let $F_k = (f_{ij}^{(k)})_{1 \leq i, j \leq n}$ denote the coefficient of x^k in f and set*

$$f_{ij} = \sum_k f_{ij}^{(k)} x^k.$$

Then

$$f \in \text{Int}_{T_n(R)}(T_n(S), T_n(I)) \quad \Leftrightarrow \quad \forall i, j \quad f_{ij} \in \text{Int}_R(T_{n-j+1}(S), T_{n-j+1}(I))$$

and

$$f \in \text{Int}^\ell_{T_n(R)}(T_n(S), T_n(I)) \quad \Leftrightarrow \quad \forall i, j \quad f_{ij} \in \text{Int}_R(T_i(S), T_i(I)).$$

Proof The criterion for $f \in \text{Int}_{T_n(R)}(T_n(S), T_n(I))$ is just Lemma 4.1 applied to each row-index i of the coefficients of $f \in (T_n(R))[x]$, and the criterion for $f \in \text{Int}^\ell_{T_n(R)}(T_n(S), T_n(I))$ is Lemma 4.1 applied to each column index j of the coefficients of f.

With the above proof we have also shown Remark 1.6 from the introduction, which is the representation in matrix form of Theorem 4.2.

5 Applications to null-polynomials and integer-valued polynomials

Applying our findings to null-polynomials on upper triangular matrices, we can describe polynomials with coefficients in $T_n(R)$ that induce the zero-function on $T_n(R)$ in terms of polynomials with coefficients in R that induce the zero function on $T_n(R)$. As before, we denote substitution for the variable in a polynomial $f = \sum_k b_k x^k$ in $T_n(R)[x]$, to the right or to the left of the coefficients, as

$$f(C) = \sum_k b_k C^k \quad \text{and} \quad f(C)_\ell = \sum_k C^k b_k$$

and define

$$N_{T_n(R)}(T_n(R)) = \{f \in (T_n(R))[x] \mid \forall C \in T_n(R) \ f(C) = 0\}$$
$$N^\ell_{T_n(R)}(T_n(R)) = \{f \in T_n(R)[x] \mid \forall C \in T_n(R) \ f(C)_\ell = 0\}$$
$$N_R(T_k(R)) = \{f \in R[x] \mid \forall C \in T_n(R) \ f(C) = 0\}$$

Null-polynomials on matrix algebras occur naturally in two circumstances: in connection with null-ideals of matrices [2,10], and in connection with integer-valued polynomials. For instance, if D is a domain with quotient field K and $f \in (T_n(K))[x]$, we may represent f as g/d with $d \in D$ and $g \in (T_n(D))[x]$. Then f is (right) integer-valued on $T_n(D)$ if and only if the residue class of g in $T_n(D/dD)[x]$ is a (right) null-polynomial on $T_n(D/dD)$ [3].

From Theorem 4.2 we derive the following corollary (see also Remark 1.6):

Corollary 5.1 *If we identify polynomials in $(T_n(R))[x]$ with their images in $T_n(R[x])$ under the isomorphism of Remark 1.1, then*
$N_{T_n(R)}(T_n(R))$

$$= \begin{pmatrix} N_R(T_n(R)) & N_R(T_{n-1}(R)) & \dots & N_R(T_2(R)) & N_R(T_1(R)) \\ 0 & N_R(T_{n-1}(R)) & \dots & N_R(T_2(R)) & N_R(T_1(R)) \\ & & \ddots & & \\ 0 & 0 & \dots & N_R(T_2(R)) & N_R(T_1(R)) \\ 0 & 0 & \dots & 0 & N_R(T_1(R)) \end{pmatrix}$$

and $N^\ell_{T_n(R)}(T_n(R))$

$$= \begin{pmatrix} N_R(T_1(R)) & N_R(T_1(R)) & \dots & N_R(T_1(R)) & N_R(T_1(R)) \\ 0 & N_R(T_2(R)) & \dots & N_R(T_2(R)) & N_R(T_2(R)) \\ & & \ddots & & \\ 0 & 0 & \dots & N_R(T_{n-1}(R)) & N_R(T_{n-1}(R)) \\ 0 & 0 & \dots & 0 & N_R(T_n(R)) \end{pmatrix}$$

This allows us to conclude:

Theorem 5.2 *Let R be a commutative ring. The set $N_{T_n(R)}(T_n(R))$ of right null-polynomials on $T_n(R)$ with coefficients in $(T_n(R))[x]$, and the set $N^\ell_{T_n(R)}(T_n(R))$ of left null-polynomials on $T_n(R)$ with coefficients in $T_n(R)$, are ideals of $(T_n(R))[x]$.*

Proof Note that $N_R(T_m(R)) \subseteq N_R(T_k(R))$ for $m \geq k$. Also, $N_R(T_m(R))R[x] \subseteq N_R(T_m(R))$ and $R[x]N_R(T_m(R)) \subseteq N_R(T_m(R))$ by substitution homomorphism for polynomials with coefficients in the commutative ring R. This observation together with matrix multiplication shows that the image of $N_{T_n(R)}(T_n(R))$ in $T_n(R[x])$ under the ring isomorphism from $(T_n(R))[x]$ to $T_n(R[x])$ is an ideal of $T_n(R[x])$, and like-wise the image of $N^\ell_{T_n(R)}(T_n(R))$ in $T_n(R[x])$. □

Now let D be a domain with quotient field K. Applying Theorem 4.2 to integer-valued polynomials (see also Remark 1.6)

$$\text{Int}_{T_n(K)}(T_n(D)) = \{f \in (T_n(K))[x] \mid \forall C \in T_n(D)\ f(C) \in T_n(D)\}$$
$$\text{Int}^\ell_{T_n(K)}(T_n(D)) = \{f \in (T_n(K))[x] \mid \forall C \in T_n(D)\ f(C)_\ell \in T_n(D)\}$$
$$\text{Int}_K(T_n(D)) = \{f \in K[x] \mid \forall C \in T_n(D)\ f(C) \in T_n(D)\}$$

yields the following corollary:

Corollary 5.3 *If we identify polynomials in $(T_n(K))[x]$ with their images in $T_n(K[x])$ under the isomorphism of Remark 1.1, then*
$\text{Int}_{T_n(K)}(T_n(D))$

$$= \begin{pmatrix} \text{Int}_K(T_n(D)) & \text{Int}_K(T_{n-1}(D)) & \ldots & \text{Int}_K(T_2(D)) & \text{Int}_K(T_1(D)) \\ 0 & \text{Int}_K(T_{n-1}(D)) & \ldots & \text{Int}_K(T_2(D)) & \text{Int}_K(T_1(D)) \\ & & \ddots & & \\ 0 & 0 & \ldots & \text{Int}_K(T_2(D)) & \text{Int}_K(T_1(D)) \\ 0 & 0 & \ldots & 0 & \text{Int}_K(T_1(D)) \end{pmatrix}$$

and $\text{Int}^\ell_{T_n(K)}(T_n(D))$

$$= \begin{pmatrix} \text{Int}_K(T_1(D)) & \text{Int}_K(T_1(D)) & \ldots & \text{Int}_K(T_1(D)) & \text{Int}_K(T_1(D)) \\ 0 & \text{Int}_K(T_2(D)) & \ldots & \text{Int}_K(T_2(D)) & \text{Int}_K(T_2(D)) \\ & & \ddots & & \\ 0 & 0 & \ldots & \text{Int}_K(T_{n-1}(D)) & \text{Int}_K(T_{n-1}(D)) \\ 0 & 0 & \ldots & 0 & \text{Int}_K(T_n(D)) \end{pmatrix}$$

We can see the following:

Theorem 5.4 *Let D be a domain with quotient field K. The set $\text{Int}_{T_n(K)}(T_n(D))$ of right integer-valued polynomials with coefficients in $(T_n(K))[x]$ and the set $\text{Int}^\ell_{T_n(K)}(T_n(D))$ of left integer-valued polynomials with coefficients in $(T_n(K))[x]$ are subrings of $(T_n(K))[x]$.*

Proof Note that $\mathrm{Int}_K(\mathrm{T}_m(D)) \subseteq \mathrm{Int}_K(\mathrm{T}_k(D))$ for $m \geq k$. Together with substitution homomorphism for polynomials with coefficients in K, this implies

$$\mathrm{Int}_K(\mathrm{T}_i(D)) \cdot \mathrm{Int}_K(\mathrm{T}_j(D)) \subseteq \mathrm{Int}_K(\mathrm{T}_{\min(i,j)}(D)).$$

This observation together with matrix multiplication shows that the image of $\mathrm{Int}_{\mathrm{T}_n(K)}(\mathrm{T}_n(D))$ in $T_n(K[x])$ under the ring isomorphism of $(\mathrm{T}_n(K))[x]$ and $T_n(K[x])$ is a subring of $T_n(K[x])$, and, likewise, the image of $\mathrm{Int}^{\ell}{}_{\mathrm{T}_n(K)}(\mathrm{T}_n(D))$ is a subring of $T_n(K[x])$. □

References

1. Evrard, S., Fares, Y., Johnson, K.: Integer valued polynomials on lower triangular integer matrices. Monatsh. Math. **170**, 147–160 (2013)
2. Frisch, S.: Integrally closed domains, minimal polynomials, and null ideals of matrices. Commun. Algebra **32**, 2015–2017 (2004)
3. Frisch, S.: Polynomial separation of points in algebras. In: Arithmetical Properties of Commutative Rings and Monoids. Lecture Notes Pure Appl. Math., vol 241. Chapman & Hall/CRC, Boca Raton, pp 253–259 (2005)
4. Frisch, S.: Integer-valued polynomials on algebras. J. Algebra **373**, 414–425 (2013). (see also Corrigendum 412, p 282 (2014))
5. Loper, K.A., Werner, N.J.: Generalized rings of integer-valued polynomials. J. Number Theory **132**, 2481–2490 (2012)
6. Peruginelli, G.: Integer-valued polynomials over matrices and divided differences. Monatsh. Math. **173**, 559–571 (2014)
7. Peruginelli, G., Werner, N.: Decomposition of integer-valued polynomial algebras (2016). **(preprint)**. arXiv:1604.08337
8. Peruginelli, G., Werner, N.: Non-triviality conditions for integer-valued polynomial rings on algebras. Monatsh. Math. (2016). **(to appear, preprint)**. arXiv:1604.08337
9. Peruginelli, G., Werner, N.J.: Properly integral polynomials over the ring of integer-valued polynomials on a matrix ring. J. Algebra **460**, 320–339 (2016)
10. Rissner, R.: Null ideals of matrices over residue class rings of principal ideal domains. Linear Algebra Appl. **494**, 44–69 (2016)
11. Werner, N.J.: Integer-valued polynomials over matrix rings. Commun. Algebra **40**, 4717–4726 (2012)
12. Werner, N.J.: Polynomials that kill each element of a finite ring. J. Algebra Appl. **13**, 1350111 (2014)