Trace Identities and Polynomial Identities of $n \times n$ Matrices

URI LERON

Department of Mathematics, The Technion, Haifa, Israel

Communicated by I. N. Herstein

Received August 15, 1975

A simple criterion for trace identities of $n \times n$ matrices over a commutative ring is given, from which the Procesi–Razmyslov criterion in terms of Young symmetrizers is deduced. A new criterion for polynomial identities of $n \times n$ matrices is then obtained.

INTRODUCTION

This paper was inspired by Procesi’s recent work on trace identities [4] and owes much to it in spirit and ideas. Yet, it has turned out to be quite self-contained; in fact, in Section 1 we give a simple criterion for trace identities (Theorem 1), from which Procesi’s main theorem on trace identities [4, Theorem 4.3] follows. This criterion and its proof are quite elementary and do not depend on the classical theory of invariants or the theory of Young symmetrizers. In particular, the restriction to fields of characteristic 0 can be removed, and we work over any commutative ring with 1.

In Section 2, the main part of the paper, we interpret the criterion of Section 1 for ordinary polynomial identities, in the spirit of [4, Theorem 6.1]. The resulting criterion for polynomial identities of $n \times n$ matrices is then strengthened to yield the main theorem (Theorem 4).

I. TRACE IDENTITIES OF $n \times n$ MATRICES

Let $C$ be a commutative ring with 1, $C_n$ the ring of $n \times n$ matrices over $C$, $e_{ij}$ the usual matrix units in $C_n$. We denote by $S_h$ the symmetric group on $\{1, \ldots, h\}$ and, more generally, by $S_A$ the symmetric group on any set $A$ (whose elements are then called “letters”). Let $\sigma \in S_h$ be decomposed as a product

1 The referee has pointed out to me that Procesi’s basic theorems on trace identities were independently obtained by Razmyslov [5].
of disjoint cycles (including those of length 1), say \( \sigma = (i_1 \ldots i_r) (j_1 \ldots j_s) \ldots (l_1 \ldots l_t) \). Given matrices \( a_1, \ldots, a_k \in \mathbb{C}_n \) we define, following Procesi [4]:

\[
\Phi_\sigma(a_1, \ldots, a_k) = \text{tr}(a_{i_1} \ldots a_{i_r}) \text{tr}(a_{j_1} \ldots a_{j_s}) \ldots \text{tr}(a_{l_1} \ldots a_{l_t}),
\]

noting that this is well defined since the trace of a product is invariant under cyclic permutations of the factors. We shall also consider the formal expressions \( \Phi_\sigma(x_1, \ldots, x_k) \) (respectively, \( \sum_{\sigma \in \mathbb{S}_k} c_{\sigma} \Phi_\sigma(x_1, \ldots, x_k) \) with \( c_{\sigma} \in \mathbb{C} \)), where the \( x \)'s are noncommutative indeterminates, calling them multilinear trace monomials (respectively, polynomials) of degree \( k \). A multilinear trace polynomial \( f(x_1, \ldots, x_k) = \sum_{\sigma \in \mathbb{S}_k} c_{\sigma} \Phi_\sigma(x_1, \ldots, x_k) \) is a trace identity for \( \mathbb{C}_n \) (or \( n \)-trace identity, for short) if \( f(a_1, \ldots, a_k) = \sum_{\sigma \in \mathbb{S}_k} c_{\sigma} \Phi_\sigma(a_1, \ldots, a_k) = 0 \) for all substitutions \( x_v \rightarrow a_v \in \mathbb{C}_n \).

The object of this section is to give a criterion for a multilinear trace polynomial \( f = \sum_{\sigma \in \mathbb{S}_k} c_{\sigma} \Phi_\sigma \) to be a trace identity for \( \mathbb{C}_n \). We begin with a series of observations.

1. Since \( f(x_1, \ldots, x_k) \) is multilinear in the \( x_i \)'s, it is an \( n \)-trace identity if it vanishes for all \( n \)-simple substitutions \( x_v \rightarrow e_{v,i_v} \in \mathbb{C}_n \).

2. \( \text{tr}(e_{i_1,j_1} e_{i_2,j_2} \ldots e_{i_k,j_k}) = \delta(i_1,i_2) \delta(j_2,j_3) \ldots \delta(j_{k-1},i_k) \delta(j_k,i_1) \); hence \( \Phi_\sigma(e_{i_1,j_1} e_{i_2,j_2} \ldots e_{i_k,j_k}) = \delta(i_1,o(1)) \ldots \delta(j_k,o(k)) \). Abbreviating \( I = (i_1, \ldots, i_k) \), \( J = (j_1, \ldots, j_k) \), \( I^o = (i_{o(1)}, \ldots, i_{o(k)}) \), and \( e(I, J) = (e_{i_1,j_1}, \ldots, e_{i_k,j_k}) \), we have

\[
\Phi_\sigma(e(I, J)) = 1 \quad \text{if} \quad I^o = J,
\]

\[
= 0 \quad \text{if} \quad I^o \neq J.
\]

3. Substituting the sequence \( e(I, J) \) in the trace polynomial \( f(x_1, \ldots, x_k) = \sum_{\sigma \in \mathbb{S}_k} c_{\sigma} \Phi_\sigma(x_1, \ldots, x_k) \), we have from (2)

\[
f(e(I, J)) = \sum_{\sigma \in T} c_{\sigma}
\]

where \( T = \{ \sigma \in \mathbb{S}_k | I^o = J \} \).

(The expression \( \sum_{\sigma \in T} c_{\sigma} \) is defined to be 0 if \( T \) is empty.) Assume now \( T \neq \emptyset \) and fix a \( \tau \) in \( T \). Then \( T = \{ \sigma \in \mathbb{S}_k | I^o = I^\tau \} = \{ \sigma \in T | I^{\sigma^{-1}} = I \} \), so letting \( G_{I^\tau} = \{ \sigma \in \mathbb{S}_k | I^o = I \} \) (a subgroup of \( \mathbb{S}_k \)), we have \( T = \{ \sigma \in \mathbb{S}_k | \sigma^{-1} \in G_{I^\tau} \} = G_{I^\tau} \). Thus

\[
f(e(I, I^\tau)) = \sum_{\sigma \in G_{I^\tau}} c_{\sigma}.
\]

**Lemma 1.** A multilinear trace polynomial \( f(x_1, \ldots, x_k) = \sum_{\sigma \in \mathbb{S}_k} c_{\sigma} \Phi_\sigma(x_1, \ldots, x_k) \) is a trace identity for \( \mathbb{C}_n \) if and only if it satisfies the following condition:

\[
\sum_{\sigma \in G_{I^\tau}} c_{\sigma} = 0 \quad \text{for all sequences } I = (i_1, \ldots, i_k) \text{ with } 1 \leq i_v \leq n, \text{ and for all } \tau \in \mathbb{S}_k, \text{ where } G_{I^\tau} = \{ \sigma \in \mathbb{S}_k | I^o = I \}.
\]

A different version of Lemma 1, for ordinary polynomial identities, was previously obtained by I. Beck, to whom I am also indebted for many stimulating discussions. Remark 2 above was used by Formanek in [2].
Proof. The preceding remarks establish the necessity of the condition. Conversely, assume the condition holds and let \( x_v \to e_{ij} \) be an \( n \)-simple substitution. We use the notation \( I, J, e(I, J) \) as in (2). If \( J = I^\tau \) for some \( \tau \in S_k \) then \( f(e(I, J)) = f(e(I, I^\tau)) = \sum_{\sigma \in G^\tau} c_\sigma = 0 \) by assumption. If on the other hand \( \{ \sigma \mid I^\sigma = J \} \) is empty, then \( \Phi_\sigma(e(I, J)) = 0 \) for all \( \sigma \in S_k \), hence \( f(e(I, J)) = 0 \), completing the proof of the lemma.

The subgroups \( G_I \) of \( S_k \) can be more usefully described as Young subgroups which we now define. Sets \( A_1, \ldots, A_n \) form a partition of \( \{1, \ldots, k\} \) if they are pairwise disjoint and their union is \( \{1, \ldots, k\} \). We call each \( A_i \) a row, and its cardinality the length of the row. The subgroup \( G = S_{A_1} \times S_{A_2} \times \cdots \times S_{A_n} \) of \( S_k \) is called [3, p. 171] the Young subgroup on the rows \( A_1, \ldots, A_n \). Any Young subgroup \( G \) on \( n \) rows (some of which may be empty) is called an \( n \)-Young subgroup, and its (right) cosets \( G\tau, \tau \in S_k \), are called \( n \)-Young cosets (on \( k \) letters). We remark that if \( G \) is the Young subgroup on rows \( A_1, \ldots, A_n \) and \( \tau \in S_k \), then \( \tau G \tau^{-1} \) is the Young subgroup on \( \tau A_1, \ldots, \tau A_n \); thus any \( n \)-Young left coset is also an \( n \)-Young right coset and vice versa.

**Theorem 1.** A multilinear trace polynomial \( f(x_1, \ldots, x_k) = \sum_{\sigma \in S_k} c_\sigma \Phi_\sigma(x_1, \ldots, x_k) \) of degree \( k \) is a trace identity for \( C_n \) if and only if \( \sum_{\sigma \in G^\tau} c_\sigma = 0 \) for every \( n \)-Young coset \( G\tau \) of \( S_k \).

Proof. The relation \( i_v = j \iff v \in A_j \) establishes a bijective correspondence between Young subgroups on the rows \( A_1, \ldots, A_n \) and subgroups \( G_I \) with \( I = (i_1, \ldots, i_k), 1 \leq i_v \leq n \), under which corresponding subgroups are equal. Thus Theorem 1 follows from Lemma 1.

**Remark 1.** Given an \( n \)-Young coset \( G\tau \) we may, of course, replace \( \tau \) by any other representative \( \pi \tau \) with \( \pi \in G \). In particular we may always choose \( \tau \) to be “\( G \)-reduced” in the following sense. Each cycle of \( \tau \) contains at most one letter from each row of \( G \). Indeed, if a cycle \( \gamma \) of \( \tau \) contains two letters \( i, j \) from the same row of \( G \) then \( (ij) \in G \), and in \( (ij)\tau \) the cycle breaks off into two cycles (possibly of length 1), each containing one of the letters \( i, j \). (Compare Remark 4 in the proof of [4, Theorem 4.5].)

**Remark 2.** In Theorem 1, it is enough to let \( G \) range over \( n \)-Young subgroups having no empty rows. (The proof is similar to that of the corresponding case in Theorem 3 below.)

**Example.** As a first application, we deduce Procesi’s fundamental trace identity [4; 2, Theorem 1]. Let \( p(x_1, \ldots, x_{n+1}) = \sum_{\sigma \in S_{n+1}} (sgn) \Phi_\sigma(x_1, \ldots, x_{n+1}) \), and let \( G\tau \) be an \( n \)-Young coset on \( n + 1 \) letters. By the pigeon-hole principle, at least one row of \( G \) contains two letters \( i \neq j \), so \( G \) contains the transposition \( (ij) \). Thus \( G\tau \) contains an equal number of odd and even permutations,
\[ \sum_{\sigma \in \mathbb{S}_n} (sg\sigma) = 0, \] and Theorem 1 implies that \( p(x_1, \ldots, x_{n+1}) \) is a trace identity for \( C_n \).

Let \( CS_k \) be the group ring of \( S_k \) over \( C \). Theorem 1 can be stated in a different form by using the "inner product" in \( CS_k \), defined for \( a = \sum_{\sigma \in \mathbb{S}_k} c_\sigma \) and \( b = \sum_{\sigma \in \mathbb{S}_k} d_\sigma \) by \( \langle a, b \rangle = \sum_{\sigma \in \mathbb{S}_k} c_\sigma d_{\sigma^{-1}} \). This is a symmetric bilinear form on \( CS_k \) satisfying \( \langle ab, c \rangle = \langle a, bc \rangle \) for all \( a, b, c \in CS_k \). For a subset \( T \subseteq S_k \) put \( \sum T = \sum_{\sigma \in T} \sigma \in CS_k \), and let \( J_n \) be the \( C \)-submodule of \( CS_k \) spanned by the elements \( \sum G_T, G_T \) ranging over all \( n \)-Young cosets of \( S_k \).

By the remark preceding Theorem 1, \( J_n \) is a (two-sided) ideal of \( CS_k \). Let \( I_n = \{ a \in CS_k \mid \langle a, b \rangle = 0 \text{ for all } b \in J_n \} \) be the orthogonal complement of \( J_n \).

**Theorem 1'** The trace polynomial \( f(x_1, \ldots, x_k) = \sum_{\sigma \in \mathbb{S}_k} c_\sigma \Phi_\sigma(x_1, \ldots, x_k) \) is a trace identity for \( C_n \) if and only if the element \( \sum_{\sigma \in \mathbb{S}_k} c_\sigma \sigma \) of \( CS_k \) belongs to \( I_n \).

**Proof.** By Theorem 1, \( f = \sum_{\sigma \in \mathbb{S}_k} c_\sigma \Phi_\sigma \) is an \( n \)-trace identity if and only if \( 0 = \sum_{\sigma \in G_T} c_\sigma = \sum_{\sigma^{-1}G} c_\sigma = \langle \sum_{\sigma \in \mathbb{S}_k} c_\sigma \sigma, \sum \tau^{-1}G \rangle \) for all \( n \)-Young cosets \( G_T \). Since the elements \( \sum_\tau \tau^{-1}G \) span \( J_n \), this is equivalent to \( \sum_{\sigma \in \mathbb{S}_k} c_\sigma \sigma \in I_n \).

Procesi's main theorem on trace identities [4, Theorem 4.3], now follows by the definition and elementary properties of the Young symmetrizers [1, pp. 190 ff]. We include the simple proof for completeness.

**Corollary (Procesi-Razmyslov).** Let \( C \) be a field of characteristic 0 and let \( I'_n \) be the ideal of \( CS_k \) generated by the Young symmetrizers on \( \geq n + 1 \) rows. Then a trace polynomial \( f(x_1, \ldots, x_k) = \sum_{\sigma \in \mathbb{S}_k} c_\sigma \Phi_\sigma(x_1, \ldots, x_k) \) is a trace identity for \( C_n \) if and only if \( \sum_{\sigma \in \mathbb{S}_k} c_\sigma \sigma \in I'_n \).

**Proof.** We have to show that \( I'_n \) is the orthogonal complement of \( J_n \), i.e., \( I'_n = I_n \). Let \( e \) be a Young symmetrizer on a diagram \( D \) with \( \geq n + 1 \) rows and let \( b = \sum G \), where \( G \) is an \( n \)-Young subgroup. The first column of \( D \) contains at least \( n + 1 \) letters, two of which must therefore fall in the same row of \( G \); call them \( i, j \). Then \( (ij) \) belongs both to \( G \) and to the vertical permutations of \( D \). Thus for every \( \sigma, \tau \in S_k \) we have \( \langle ae, b\tau \rangle = \langle ae, (ij)b\tau \rangle = \langle ae(ij), b\tau \rangle = -\langle ae, b\tau \rangle \), hence \( \langle ae, b\tau \rangle = 0 \). Since the elements of the form \( ae \) and \( b\tau \) span \( I'_n \) and \( J_n \), respectively, \( I'_n \) is orthogonal to \( J_n \), placing \( I'_n \subseteq I_n \). On the other hand, \( CS_k \) is generated (as, say, a right ideal) by the Young symmetrizers, and all the Young symmetrizers on \( < n \) rows are in \( J_n \), so \( CS_k = I'_n + J_n \). This, together with \( I'_n \cap J_n = \{0\} \), implies \( I'_n = I_n \).

2. POLYNOMIAL IDENTITIES OF \( n \times n \) MATRICES

In this section we use the following simple fact to render the theory of trace identities applicable to problems on polynomial identities.
For $a \in C_n$, $a = 0$ if and only if $\text{tr}(ab) = 0$ for all $b \in C_n$.

Corresponding to every multilinear polynomial $F(x_1, \ldots, x_{k-1}) = \sum_{\sigma \in S_{k-1}} c_{\sigma} x_{\sigma(1)} \cdots x_{\sigma(k-1)}$ in the noncommutative indeterminates $x_1, \ldots, x_{k-1}$ over $C$, we consider the following trace polynomial of degree $k$.

\[ f(x_1, \ldots, x_k) = \text{tr}(F(x_1, \ldots, x_{k-1})x_k) = \sum_{\sigma \in S_{k-1}} c_{\sigma} \Phi_\sigma(x_1, \ldots, x_k), \]

where $\tilde{\sigma} = (\sigma(1) \sigma(2) \cdots \sigma(k - 1)k) \in S_k$. Note that, putting $\sigma$ in $S_k$ by letting $\sigma(k) = k$, we have $\tilde{\sigma} = \sigma \gamma \sigma^{-1}$, where $\gamma = (1 \ 2 \ \cdots \ k)$. By the opening remark, $F(x_1, \ldots, x_{k-1})$ is a polynomial identity for $C_n$ if and only if $f(x_1, \ldots, x_k)$ is a trace identity for $C_n$. Note that as $\sigma$ ranges over $S_{k-1}$, $\tilde{\sigma}$ ranges over the set $\Gamma$ of all cycles of length $k$ in $S_k$ (we call them full cycles). We are thus led to study trace polynomials having the special form $f(x_1, \ldots, x_k) = \sum_{\sigma \in \Gamma} c_{\sigma} \Phi_\sigma(x_1, \ldots, x_k)$.

From Theorems 1 and 1' we read off immediately:

**Theorem 2.** The trace polynomial $f(x_1, \ldots, x_k) = \sum_{\sigma \in \Gamma} c_{\sigma} \Phi_\sigma(x_1, \ldots, x_k)$ is a trace identity for $C_n$ if and only if $\sum_{\sigma \in \Gamma_n} c_{\sigma} = 0$ for all $n$-Young cosets $G_T$ of $S_k$.

(As before, a sum on the empty set is defined to be 0.)

**Theorem 2'.** Let $T$ be the $C$-submodule of $CS_k$ spanned by the elements $\sum G_T \cap \Gamma$, $G_T$ ranging over all the $n$-Young cosets of $S_k$. Then $f(x_1, \ldots, x_k) = \sum_{\sigma \in \Gamma} c_{\sigma} \Phi_\sigma(x_1, \ldots, x_k)$ is a trace identity for $C_n$ if and only if the element $\sum_{\sigma \in \Gamma} c_{\sigma}$ of $CS_k$ is orthogonal to $T$.

**Remark 1.** In applying Theorems 2 and 2' one can always choose the representative $\tau$ of $G_T \cap \Gamma$ to be a full cycle, since the conditions are met trivially when $G_T \cap \Gamma$ is empty.

**Remark 2.** By working out some examples, it becomes clear that these criteria are not very efficient, since many different cosets $G_T$ give rise to the same set $G_T \cap \Gamma$. It is therefore desirable to increase the efficiency of the criteria by finding smaller sets of generators for the $C$-module $T$. This can apparently be achieved by putting restrictions on the cosets $G_T$ in two distinct ways: Restricting the "shape" of $G$ (i.e., the lengths of its rows) or restricting the form of $\tau$ by insisting that it has a special relation with $G$ (apart from the freedom to select representatives as in Remark 1 following the proof of Theorem 1, or Remark 1 here). The first method is illustrated by Theorem 3 below.

We now state our main theorem.
DEFINITION. An n-Young subgroup $G$ of $S_k$ (as well as any of its cosets) is called full if all its rows have lengths $\geq 2$.

**Theorem 3.** Let $k \geq 2n$. The trace polynomial $f(x_1, ..., x_k) = \sum_{\sigma \in \Gamma} x_{\sigma}(x_1, ..., x_k)$ is a trace identity for $C_n$ if and only if $\sum_{\sigma \in G\cap \Gamma} c_{\sigma} = 0$ for all full n-Young cosets $G \tau$ of $S_k$.

For the proof, which will proceed in several steps, let us first introduce some terminology and notation.

For any n-Young coset $G \tau$ we write $\sum_G G \tau = \sum G \tau \cap \Gamma = \sum_{\sigma \in G\cap \Gamma} \sigma \in CS_k$.

We say that the coset $G \tau$ is $\Gamma$-generated by a collection $A \ell$ of subgroups of $S_k$ if $\sum_G G \tau$ is generated additively by the set $\{\sum G' \sigma \mid G' \in A \ell, \sigma \in S_k\}$. $G$ itself is said to be $\Gamma$-generated by $A \ell$ if every coset $G \tau$ is $\Gamma$-generated by $\ell$. (If $\ell$ consists of a single subgroup $G'$, we replace $\ell \ell = \{G'\}$ by $G'$ in the above terminology.) Finally, let $G$ be a Young subgroup on rows $A_1, ..., A_n$, and let $1 \leq \mu, \nu \leq n$ and $j \in A_\mu$ be given. Let $G'$ be the Young subgroup on the rows $A_1', ..., A_n'$, where $A_\nu' = A_\nu \cup \{j\}$, $A_\mu' = A_\mu - \{j\}$ and $A_\lambda' = A_\lambda$ for $\lambda \neq \mu, \nu$. We then say that $G'$ is obtained from $G$ by adjoining $j$ to the row $A_\nu$.

**Lemma 4.** Let $G \tau$ be a Young coset on rows $A_1, ..., A_n$ and let $1 \leq \nu \leq n$ be given. Let $G'$ be the n-Young subgroup obtained from $G$ by adjoining $1 \leq j \leq k$ to $A_\nu$, where either (i) $A_\nu$ is empty and $j$ is arbitrary, or (ii) $A_\nu = \{i\}$ and $j = \tau^{-1}(i)$. Then $G \tau$ is $\Gamma$-generated by $G'$.

**Proof.** Case (i). The $\nu$th row of $G'$ is $A_\nu' = \{j\}$, so $S_{A_\nu} = \{1\}$ and $G' \subset G$. We write $G$ as a disjoint union of cosets of $G'$, $G = \bigcup_{\lambda=1}^t G' \sigma_\lambda$, $\sigma_\lambda \in G$, thereby obtaining $\sum_{G \tau} G \tau = \sum_{\lambda=1}^t \sum_{G' \sigma_\lambda} \tau$, i.e., $G \tau$ is $\Gamma$-generated by $G'$.

Case (ii). Let $H = G \cap G'$, or, more explicitly, $H = \{\sigma \in G \mid \sigma(j) = j\} = \{\sigma \in G' \mid \sigma(j) = j\}$. Let $A_\mu = \{j = j_1, j_2, ..., j_t\}$ be the row of $G$ containing $j$. Then $G$ and $G'$ decompose into disjoint cosets of $H$ as follows. $G = \bigcup_{\lambda=1}^t H(j, j_\lambda)$ (where $(jj)$ means the identity) and $G' = H \cup H(jj)$. Let $\sigma \in S_k$ satisfy $\sigma(j) = i$. We claim that $\sum_{G' \sigma} G' \tau = \sum_{H \sigma}$ $H \sigma$. Indeed, $\sum_{G' \sigma} G' \tau = \sum_{H \sigma} + \sum_{H(jj)} \sigma$ so we have only to show that $H(jj) \sigma \cap \Gamma$ is empty. But for any $\pi \in H$, $\pi(j) = j$ so $[\pi(jj) \sigma(j) = \pi(jj)(i) = \pi(j) = j$, excluding $\pi(jj) \sigma$ from $\Gamma$. Thus our claim is established. Since each $\sigma(j) = i$ (recall $j = \tau^{-1}(i)$), we have

$$\sum_{G \tau} G \tau = \sum_{\lambda=1}^t \sum_{H \sigma} H(jj_\lambda) \tau = \sum_{\lambda=1}^t \sum_{G' \sigma_\lambda} G'(jj_\lambda) \tau,$$

completing the proof of the lemma.
For the next lemma we fix two letters $i, j$ out of $\{1, \ldots, k\}$ and let $B = \{1, \ldots, k\} - \{i, j\}$. $S_B$ denotes, as before, the symmetric group on the letters of $B$, and $\Gamma_B \subseteq S_B$ is the set of full cycles (of length $k - 2$) on the letters of $B$. We embed $S_n \subseteq S_k$ in the usual way.

**Lemma 5.** Let $G$ be an $n$-Young subgroup of $S_k$ having a row $A_{\mu} = \{i, j\}$, $H$ the subgroup consisting of permutations of $G$ fixing $i$ and $j$, and let $\tau \in S_k$ satisfy $\tau(j) = i, \tau(i) \neq j$. Then $\tau(lij) \in S_B$ and $(G_{\tau} \cap \Gamma)(lij) = H_{\tau}(lij) \cap \Gamma_B$, where $l = \tau^{-1}(j)$.

**Proof.** We first repeat (briefly) the argument in Lemma 4, Case (ii), to show that $G_{\tau} \cap \Gamma = H_{\tau} \cap \Gamma$: Thus $G_{\tau} = H_{\tau} \cup H(lij)_{\tau}$, but for $\pi \in H [\pi(ij)\tau](j) = j$, so $H(lij)_{\tau} \cap \Gamma$ is empty, and the claim follows. The proof therefore reduces to showing that $(H_{\tau} \cap \Gamma)(lij) = H_{\tau}(lij) \cap \Gamma_B$. For this, one has merely to note that if $\sigma \in S_k$ is any permutation containing the string $lij$ in one of its cycles, then right multiplication of $\sigma$ with $(lij)$ simply amounts to deleting $i$ and $j$ from that cycle. Since $\tau$, and together with it any permutation in $H_{\tau}$, contains the string $lij$ in one of its cycles (recall $\tau(l) = j, \tau(j) = i$, and $l \neq i$), the desired equality follows, and the proof of the lemma is complete.

**Proof of Theorem 3.** Define the deficiency $d(G)$ of an $n$-Young subgroup $G$ to be the minimal number of letters which have to be added to the various rows of $G$ to make $G$ full. If $G$ is an $n$-Young subgroup of deficiency $d$, let us for the proof call the ordered pair $(n, d)$ the type of $G$. We order these pairs in the lexicographic order, i.e., $(n, d) < (n', d')$ if $n < n'$ or $n = n'$ and $d < d'$. We now prove the following proposition by induction on the type of $G$.

**Proposition.** Every Young subgroup of type $\langle n, d \rangle$ on $k \geq 2n$ letters is $\Gamma$-generated by full $n$-Young subgroups on $k$ letters.

The proposition is clearly true if $G$ has type $\langle n, 0 \rangle$, $n$ arbitrary, or type $\langle 1, d \rangle$, $d$ arbitrary (in fact, $n = 1$ and $k \geq 2n$ implies $d = 0$). Assume now that $G$ has type $\langle n, d \rangle$ with $n > 1$ and $d > 0$ and note that, since $k \geq 2n$, this implies that at least one row of $G$ has length $\geq 3$. Consider any coset $G_{\tau}$.

If $G$ has an empty row $A_{\mu}$, let $j$ be any letter belonging to a row of length $\geq 3$; if $G$ has no empty rows but has a row $A_{\nu} = \{i\}$ such that $\tau^{-1}(i)$ belongs to a row of length $\geq 3$, let $j = \tau^{-1}(i)$. In both cases, let $G'$ be the $n$-Young subgroup on $k$ letters obtained from $G$ by adjoining $j$ to $A_{\nu}$. Then $G'$ has type $\langle n, d - 1 \rangle$ and it $\Gamma$-generates $G$ by Lemma 4. Since, by the induction hypothesis, $G'$ is $\Gamma$-generated by full $n$-Young subgroups on $k$ letters, the same is true for $G_{\tau}$. Since $\tau$ was arbitrary, the proposition is proved for this case.
In the remaining cases, \( G \) has a row \( A_v = \{ i \} \) such that \( j = \tau^{-1}(i) \) belongs to a row of length 2, say \( A_u = \{ j, m \} \). By Lemma 4, \( G_T \) is \( \Gamma \)-generated by \( G' \), the subgroup obtained by adjoining \( j \) to \( A_v \); more explicitly, \( \sum \Gamma G_T = \sum_r G' \tau + \sum_r G'(jm)\tau \). Note that both \( G' \tau \) and \( G'(jm)\tau \) satisfy the hypotheses of Lemma 5 (the only hypothesis requiring comment is \( \tau(i) \neq j \); but if \( \tau(i) = j \) then the set \( \{ i, j \} \) is invariant under \( G' \tau \) and \( G' \tau \cap \Gamma \) is empty; similarly for \( (jm)\tau \)).

We are therefore reduced to proving the proposition for a coset \( G_T \) satisfying the hypotheses of Lemma 5. Thus, using the notation of Lemma 5, we have \( (\sum \Gamma G_T)(lij) = \sum \Gamma_B H_{\tau_B} \), where \( \tau_B = \tau(lij) \in S_B \). Since \( H \) is an \((n - 1)\)-Young subgroup of \( S_B \), \( \tau_B \in S_B \) and \( B \) has \( k - 2 \geq 2(n - 1) \) letters, the induction hypothesis applies to \( \sum \Gamma H_{\tau_B} = \sum \Gamma_B H_{\tau_B} = \sum \Gamma_B H_{\sigma_1} \), where \( H_{\sigma_1} \) are full \((n - 1)\)-Young cosets of \( S_B \). Right-multiply both sides of this equality by \( (lij)^{-1} = (lji) \) to get \( \sum \Gamma G_T = (\sum \Gamma_B H_{\tau_B})(lji) = \sum \Gamma_B H_{\sigma_1}(lji) \). For each \( 1 \leq \lambda \leq t \), let \( G_{\lambda} \) be the \( n \)-Young subgroup of \( S_k \) obtained by adding the row \( \{ i, j \} \) to the rows of \( H_{\lambda} \), and let \( \sigma_\lambda = \sigma_0(lji) \). Then \( G_{\lambda} \sigma_\lambda \) satisfies the hypotheses of Lemma 5, so we have \( (\sum \Gamma G_{\lambda} \sigma_\lambda)(lij) = \sum \Gamma_B H_{\sigma_\lambda}(lij) = \sum \Gamma_B H_{\sigma_0}(lij) = \sum \Gamma_B G_{\lambda} \tau_\lambda \). Thus, finally, \( \sum \Gamma G_T = \sum \lambda \sum \Gamma_B H_{\sigma_0}(lij) = \sum \lambda \sum \Gamma G_{\lambda} \tau_\lambda \), and since \( G_{\lambda} \) are full, this completes the proof of the proposition.

To conclude the proof of Theorem 3, assume \( f = \sum_{\sigma \in \Gamma} c_\sigma \Phi_\sigma \) satisfies \( \sum_{\sigma \in G_T \cap \Gamma} c_\sigma = 0 \) for all full \( n \)-Young cosets of \( S_k \). Then \( \sum_{\sigma \in \Gamma} c_\sigma \Phi_\sigma \) is orthogonal to all the elements \( \sum \Gamma G_T \) with \( G_T \) a full \( n \)-Young coset. But these elements span, by the proposition, the \( C \)-module \( T \) of Theorem 2', hence \( \sum_{\sigma \in \Gamma} c_\sigma \Phi_\sigma \) is orthogonal to \( T \). The conclusion of Theorem 3 now follows from Theorem 2'.

Recalling the opening remarks of this section, we derive from Theorem 3 a criterion for polynomial identities of \( C_n \).

**Theorem 4.** Let \( k \geq 2n - 1 \) and let \( F(x_1, \ldots, x_k) = \sum_{\sigma \in S_k} c_{\sigma} x_{\sigma(a)} \cdots x_{\sigma(k)} \) be a multilinear polynomial of degree \( k \) over \( C \). For each \( \sigma \in S_k \), write \( \tilde{\sigma} = (\sigma(1) \sigma(2) \cdots \sigma(k) k + 1) \in S_{k+1} \). Then \( F(x_1, \ldots, x_k) \) is a polynomial identity for \( C_n \) if and only if \( \sum_{\sigma \in S_k} c_\sigma = 0 \) for all full \( n \)-Young cosets \( G_{\tau} \) of \( S_{k+1} \), the sum being taken over all \( \sigma \in S_k \) with \( \tilde{\sigma} \in G_{\tau} \).

We remark in conclusion that a proof of the Amitsur–Levitzki theorem can be obtained using Theorem 4. This proof is omitted here in view of the recent simple proofs in [2, 6].

**References**