Exponential Growth for Codimensions of Some p.i. Algebras

Allan Berele\textsuperscript{1}

Department of Mathematics, DePaul University, Chicago, Illinois 60614
E-mail: aberele@condor.depaul.edu

and

Amitai Regev\textsuperscript{2}

Department of Mathematics, The Weizmann Institute of Science,
Rehovot 76100, Israel
E-mail: regev@wisdom.ac.il

Communicated by Susan Montgomery

Received January 3, 2000

By the Giambruno–Zaicev theorem for associative p.i. algebras, the exponential rate of growth of the codimensions of such a p.i. algebra is always a positive integer. Here we calculate that integer for various generic p.i. algebras which are given by a single identity. These include Capelli-type identities and the various powers of the standard polynomials. © 2001 Academic Press

1. INTRODUCTION

In this paper we study the codimension sequences of associative p.i. algebras in characteristic zero. If $A$ is a p.i. algebra and $\{c_n(A)\}$ its codimension sequence, then Regev showed that this sequence is exponentially bounded,
i.e., that

\[ c_n(A) \leq C a^n \]

for some \( C \) and \( a \) which depend on \( A \); see [L, R1] for the best known estimates. More recently, in [GZ1, GZ2] Giambruno and Zaicev improved this to show that

\[ g_1(n)a^n \leq c_n(A) \leq g_2(n)a^n \]

for rational functions \( g_1 \) and \( g_2 \), where \( 0 < g_1(n) \) for large \( n \). In addition they proved the striking result that \( a \) is always an integer. We will call \( a \) the exponential rate of growth of \( A \) and write \( a = \exp(A) \). The Giambruno Zaicev theorem raises the following general

**Problem.** Given a p.i. algebra \( A \), calculate the integer \( a = \exp(A) \).

We will be interested in the case in which \( A \) is a generic p.i. algebra, especially in the case in which \( A \) is the free algebra modulo a single identity. In this case, if the identity is \( f = f(x) \) we will write \( \exp(f) \) in place of \( \exp(A) \).

Codimensions are the degrees of the corresponding cocharacters. The asymptotic behaviour of cocharacter sequences has been studied for a number of algebras, especially verbally prime algebras and algebras related to them [BR3, BR4]. There has been much less success in describing the cocharacters of generic algebras whose identities were generated by a given set.

In this paper we will study \( \exp(f) \) for a number of important polynomials \( f \). Here are our main results in this vein:

**Theorem 3.1.** (1) Let \( n \geq 2 \). Then \( \exp(s_n(x)) = \lfloor \frac{n}{2} \rfloor^2 \). Here \( s_n(x) = s_n(x_1, \ldots, x_n) \) is the \( n \)th standard polynomial.

(2) Let \( f(x) \neq 0 \) be a polynomial of degree \( n \geq 4 \), then \( \exp(f(x)) \leq \exp(s_n(x)) = [\frac{n}{2}]^2 \). Thus, standard polynomials are the weakest identities among polynomials of the same degree—in the sense of having largest codimensions.

**Theorem 5.7.** If \( n = 3q + r \), \( 0 \leq r \leq 2 \), then \( \exp([x, y]^n) = 4q + r \), and for all \( d \geq 3 \), \( \exp([x_1, \ldots, x_d]^n) = 2n \). Here \([u, v] = uv - vu\) and \([x_1, \ldots, x_d] = [[x_1, \ldots, x_{d-1}], x_d]\).

**Theorem 5.8.** \( \exp([x, y, \ldots, y]^n) = 2n \).

By a theorem of Amitsur, every p.i. algebra satisfies \( s_n(x)^k \) for some \( n \) and \( k \). We prove

**Theorem 6.10.** For all \( n \geq 4 \), \( \exp(s_n(x)^k) = k \lfloor n/2 \rfloor^2 \).
THEOREM 7.3. If \( f(x_1, \ldots, x_n) \) is homogeneous in each variable, then \( \exp(f(x_1, \ldots, x_n)^k) \) is bounded above and below by a linear function of \( k \).

It is worthwhile to say a bit more about the material in Section 4. This section deals with the exponential growth of the Amitsur polynomials. These generalize the Capelli polynomials in the sense that the Capelli polynomials characterize which algebras have a cocharacter contained in a given strip and the Amitsur polynomials characterize which algebras have a cocharacter contained in a given hook. Denoting by \( E_{k, \ell}^* \) the Amitsur polynomial corresponding to the \((k, \ell)\) hook, we have

PROPOSITION 4.4. Let \( l \leq k \). Then \( k + l - 4 \leq E_{k, \ell} \leq \exp(E_{k, \ell}^*[x, y]) \leq k + l \).

Of course, one would like a more precise description of the exponential growth. In Section 4, we tell how it may be computed, and we compute it in various special cases. The corresponding problem for Capelli identities was solved by Mishchenko et al. [MRZ]. The solution there involved Lagrange’s four square theorem and the complete solution for the Amitsur polynomials would involve the solution of a similar number-theoretic problem. This and related questions lead to Waring type problems.

2. BACKGROUND

The starting point in the study of exponential rates of growths of associative p.i. algebras is this theorem of Kemer:

THEOREM (Kemer [K]). If \( A \) is any p.i. algebra in characteristic zero, then there exists a finite dimensional, \( \mathbb{Z}/2\mathbb{Z} \)-graded algebra \( B \) such that \( A \) is p.i. equivalent to \( G(B) = B_0 \oplus E_0 + B_1 \oplus E_1 \), the Grassmann envelope of \( B \).

Giambruno and Zaicev constructed a method to calculate the exponential behaviour of the codimension sequence of \( A \) using Kemer’s theorem. By Wedderburn’s principal decomposition theorem, we can write \( B = B' \oplus J \), where \( J \) is the Jacobson radical and \( B' \) is a semisimple subalgebra. This latter algebra \( B' \) can be further decomposed as a direct sum of graded simple algebras \( B' = B_1 \oplus \cdots \oplus B_k \). They then considered sequences of distinct \( B' \)'s \( B_{i_1}, \ldots, B_{i_s} \), with

\[
B_{i_1}B_{i_2}J \cdots J B_{i_s} \neq (0). \tag{2.1}
\]

It is worth reminding the reader that there are three types of simple \( \mathbb{Z}/2\mathbb{Z} \)-graded algebras: Matrices over the field, \( F_n \), concentrated in degree zero; the algebras \( M(k, \ell) \) which are \((k + \ell) \times (k + \ell)\) over \( F \) with degree zero part consisting of the \( k \times k \) and \( \ell \times \ell \) blocks on the diagonal and
degree one part consisting of the off-diagonal blocks; and $F_n + tF_n$, with $t$ a central, degree one element whose square is 1. Some authors prefer to consider this last algebra as simply $F_n$ with the degree one part equal to the degree zero part. At any rate, the dimensions are $n^2$, $(k + \ell)^2$, and $2n^2$, respectively. The Grassmann envelopes of the simple graded algebras are the verbally prime algebras, $F_n$, $M_{k, \ell}$, and $M_n(E)$. Of course, each of these is a subalgebra of matrices over $E$. Now let the maximum value of $\dim B_1 + \cdots + \dim B_s$ in products of the form (2.1) be $d$.

**Theorem (Giambruno and Zaicev [GZ1, GZ2]).** The limit of the $n$th root of the $n$th codimension of $A$, $\lim_{n \to \infty} \sqrt[n]{c_n(A)}$, exists and equals $d$.

We will define $\exp(A)$ to be $\lim_{n \to \infty} \sqrt[n]{c_n(A)}$. An important special case is the case in which $B$ is graded simple and so $A = G(B)$ is verbally prime. In this case the asymptotics of $c_n(A)$ were investigated in [R2, R3, BR3]. It follows from these papers, or as a consequence of the Giambruno–Zaicev theorem, that $\exp(A) = \dim(B)$.

Our main tool here is Theorem 2.4 which is a corollary of the Giambruno–Zaicev theorem.

Let $A_1, \ldots, A_n$ be verbally prime algebras, so $A_1 = G(B_1), \ldots, A_n = G(B_n)$ the $B$'s are all graded simple. We first define $B_1 \circ \cdots \circ B_n$ to be the $\mathbb{Z}_2$-graded matrix algebra

$$
\begin{pmatrix}
B_1 & * & \cdots & * \\
0 & B_2 & * & \\
& \vdots & \ddots & \vdots \\
0 & \cdots & 0 & B_n
\end{pmatrix}
$$

This may be graded in a manner consistent with the gradings on the $B_i$. The simplest way would be to let all of the $*$ entries to have degree both zero and one. Next we define $A_1 \circ \cdots \circ A_n$ to be the Grassmann envelope $G(B_1 \circ \cdots \circ B_n)$, namely, it will look like matrices of the form

$$
\begin{pmatrix}
G(B_1) & * & \cdots & * \\
0 & G(B_2) & * & \\
& \vdots & \ddots & \vdots \\
0 & \cdots & 0 & G(B_n)
\end{pmatrix},
$$

where all the entries come from $E$ and the $*$-entries are arbitrary in $E$. We will call an algebra of this form a prime product algebra. Let $A = G(B)$ be verbally prime, with $B$ graded simple. Then $\exp(A) = \dim(B)$ [R2, BR3]. By (the proof of) the Giambruno–Zaicev theorem it follows that

$$
\exp(A_1 \circ \cdots \circ A_n) = \exp(A_1) + \cdots + \exp(A_n) = \dim(B_1) + \cdots + \dim(B_n).
$$

Here are two useful properties of prime products.
Remark 2.1. If $f_i$ is an identity for $A_i$, $i = 1, \ldots, n$, then the product $f_1 \cdots f_n$ is an identity for $A_1 \circ \cdots \circ A_n$.

Remark 2.2. The Grassmann algebra $E$ is the exterior algebra of some vector space $V, E = E(V)$. Given a matrix with entries in $E$, the support is the smallest subspace of $V$ over which all the entries are defined. A set of such matrices has disjoint support if no two of these vector spaces intersect non-trivially. Then, if $0 \neq a_i \in A_i$, $i = 1, \ldots, n$, have disjoint support there exists $x_i \in A_1 \circ \cdots \circ A_n$ such that $a_i x_i \cdots x_{i-1} a_n \neq 0$.

It will sometimes be useful to speak of $A_1 \circ \cdots \circ A_n$ when the $A_i$'s are prime product algebras and not just verbally prime algebras. Even more generally, if $A \subseteq M_n(E)$ and $B \subseteq M_m(E)$, we may define $A \circ B \subseteq M_{n+m}(E)$ in the obvious way

$$A \circ B = \left\{ \begin{pmatrix} a & m \\ 0 & b \end{pmatrix} \mid a \in A, \ m \in M_{n,m}(E), b \in B \right\}.$$ 

Note that if the $B_i$'s are graded simple matrix algebras, then

$$B_1 \circ \cdots \circ B_n \cong B_1 \oplus \cdots \oplus B_n + J,$$

$J$ being the corresponding Jacobson radical.

We shall need

Lemma 2.3. Let $B = B' + J, J = J(B)$ the Jacobson radical, and $B' = B_1 \oplus \cdots \oplus B_n$ where the $B_i$ are matrix algebras. Assume that for some $x_i, \ldots, x_{j-1} \in J, B_1 x_i B_2 x_2 \cdots x_{j-1} B_j \neq 0$.

Then $B$ contains the sub-algebra

$$D = \sum_{1 \leq i \leq j \leq n} \oplus B_i x_i B_{i+1} x_{i+1} \cdots x_{j-1} B_j,$$

and $D$ is isomorphic with $B_1 \circ \cdots \circ B_n$.

Proof. That the sum in $D$ is direct easily follows from the orthogonality of the $B_i$.

Let $B_i = F_{a_i}$. First, denote by $e'_{a_i, \beta}$ the matrix units in $B_i$. For each $i$ there are matrix units $e_{a_i, \beta}, e_{i+1}$ such that $e_{a_i, \beta} x_i e_{i+1} \neq 0$. By replacing $x_i$ by $e'_{a_i, \beta} e_{i+1} e_{\gamma, \delta} e_{\delta, 1}$ we may assume without loss of generality that $e'_{i+1} x_i e'_{i+1} \neq 0$ and $e'_{i+1} x_i e'_{i+1} = 0$ if $\alpha \neq 1$ or $\beta \neq 1$.

We can index the matrix units in $B_1 \circ \cdots \circ B_s$ as $e(i, j, \alpha, \beta)$, where $1 \leq i \leq j \leq s, 1 \leq \alpha \leq \dim B_i$, and $1 \leq \beta \leq \dim B_j$. Thus, in the matrix algebra $B_1 \circ \cdots \circ B_s, e(i, j, \alpha, \beta)$ is the matrix-unit $e_{u,v} = e(i, j, \alpha, \beta)$, where $u = a_1 + \cdots + a_{i-1} + \alpha$ and $v = a_1 + \cdots + a_{j-1} + \beta$. It is easy to verify that multiplication is given by $e(i, j, \alpha, \beta) e(k, \ell, \gamma, \delta) = 0$ unless $j = k$ and
\( \beta = \gamma \), in which case it equals \( e(i, \ell, \alpha, \delta) \). Since the \( B_i \)'s are orthogonal, a straightforward computation shows that the correspondence

\[
e(i, j, \alpha, \beta) \leftrightarrow e^i_{a_1,1}x_i e^{i+1}_{1,1}x_{i+1}e^{i+2}_{1,1} \cdots x_{j-2}e^{j-1}_{1,1}x_{j-1}e_i^j
\]

is the required isomorphism.

We can now prove

**Theorem 2.4.** Let \( A \) be any p.i. algebra. Then \( \exp(A) \) is the maximum value of \( \exp(\bar{A}) \), where \( \bar{A} \) runs over all prime product algebras which satisfy all the identities of \( A \).

**Proof.** Consider finite dimensional algebras of the form \( A_1 = G(B_1 \oplus \cdots \oplus B_n + J) \), satisfying all the identities of \( A \). Here \( B = B_1 \oplus \cdots \oplus B_n + J, J = J(B) \) the Jacobson radical, and the \( B_i \) are \( \mathbb{Z}_2 \)-graded simple matrix algebras. By (the proof of) the Giambruno–Zaicev Theorem, \( \exp(A) \) is the maximum value of

\[
\dim(B_{i_1}) + \cdots + \dim(B_{i_s})
\]

taken over all such algebras \( A_1 \), and such that \( B_{i_1}J B_{i_2} \cdots J B_{i_s} \neq 0, i_1, \ldots, i_s \) distinct. The proof now clearly follows from Lemma 2.3.

We now turn to the proof that the ideal of identities of \( A_1 \circ \cdots \circ A_n \) is the product of the ideals of identities of the individual \( A_i \).

**Lemma 2.5.** Let \( a_1, \ldots, a_k \in M_n(E) \), not all zero, and \( b_1, \ldots, b_k \in M_m(E) \), not all zero, be such that each \( a_i \) and \( b_j \) have disjoint support and \( \sum_{i=1}^k a_i x b_i = 0 \) for all \( x \in M_{n,m}(E) \). Then, \( a_1, \ldots, a_k \) are F-linearly dependent and \( b_1, \ldots, b_k \) are F-linearly dependent.

**Proof.** Write \( a_i = (a_{\alpha, \beta}^{(i)}) \) and \( b_i = (b_{\alpha, \beta}^{(i)}) \), and assume W.L.O.G. that \( b_{1,1}^{(1)} \neq 0 \). For any \( \alpha, \beta \),

\[
\sum_{i=1}^k e_{1,\alpha} a_i e_{\beta,1} b_{i,1,1} = 0.
\]

This implies that \( \sum_{i} a_{\alpha, \beta}^{(i)} b_{1,1}^{(i)} = 0 \), which in turn implies that \( \sum a_i b_i^{(i)} = 0 \). Hence, \( a_1, \ldots, a_k \) are dependent over \( E \). To get dependence over \( F \) we use the disjoint supports. For each \( i \) let \( b_{1,1}^{(i)} = \sum_w f_i(w) w \), where \( w \) runs over monomials in \( E \) and \( f_i(w) \in F \) is the coefficient of \( w \). If some \( f_i(w_0) \neq 0 \), then \( \sum_i a_i f_i(w_0) = 0 \) gives a non-trivial \( F \)-relationship among \( a_1, \ldots, a_k \). The case of \( b_1, \ldots, b_k \) is similar.
DEFINITION. Given a prime product algebra $A$ and an integer $k$, we can construct $k$ generic matrices $X_1, \ldots, X_k$ as follows. By definition of prime product algebras, $A \subseteq M_n(E)$ for some $n$, and for each $1 \leq \alpha, \beta \leq n$ the $(\alpha, \beta)$ entry of each element of $A$ is constrained to either be $0$, an element of $F$, or $E_0$, an element of $E_1$, or it can be any element of $E$. Now the $X_i$ will be elements of the algebra of $n \times n$ matrices over the free supercommutative algebra $F[t^{(i)}_{\alpha, \beta}, v^{(i)}_{\alpha, \beta}|i = 1, \ldots, k, \alpha, \beta = 1, \ldots, n]$. For each $i$, the matrix $X_i$ will have $(\alpha, \beta)$-entry equal to 0, or $t^{(i)}_{\alpha, \beta}$, or $v^{(i)}_{\alpha, \beta}$, or $t^{(i)}_{\alpha, \beta} + v^{(i)}_{\alpha, \beta}$, depending on the restrictions on the $(\alpha, \beta)$-entries of $A$. Moreover, $U_k(A)$ will be defined to be $F[X_1, \ldots, X_k]$, the $F$-algebra generated by $X_1, \ldots, X_k$.

This algebra has two important properties: First, it is generic in the sense that given any $a_1, \ldots, a_k \in A$, there is a homomorphism $U_k(A) \to A$ that takes each $X_i$ to $a_i$. This implies that if $f(x_1, \ldots, x_k)$ is a non-commutative polynomial and if $f(X_1, \ldots, X_k) = 0$ in $U_k(A)$, then $f$ is a polynomial identity for $A$. Each of the verbally prime algebras, and hence each of the prime product algebras, is defined using a Grassmann algebra whose definition depends on a vector space over the field. By a Vandermonde argument, the polynomial identities are not sensitive to which characteristic zero field we use nor which infinite dimensional Grassmann algebra we use. So, since $U_k(A)$ is contained in the algebra obtained from $A$ by extending the field and the underlying vector space, $U_k(A)$ must satisfy all of the identities of $A$.

**Lemma 2.6.** Given prime product algebras $A \subseteq M_n(E)$ and $B \subseteq M_m(E)$, and given non-zero polynomials $f_i(x_1, \ldots, x_k), g_i(x_1, \ldots, x_k)$ in the free algebra such that

$$\sum_{i=1}^{d} f_i(a_1, \ldots, a_k)mg_i(b_1, \ldots, b_k) = 0$$

for all $a_1, \ldots, a_k \in A$, $m \in M_{n,m}(E)$, and $b_1, \ldots, b_k \in B$, then either some linear combination of the $f_i$ is an identity for $A$ or some linear combination of the $g_i$ is an identity for $B$.

**Proof.** Let $X_1, \ldots, X_k$ be generic for $A$ and $Y_1, \ldots, Y_k$ be generic for $B$ with disjoint supports. Then $\sum f_i(X_1, \ldots, X_k)mg_i(Y_1, \ldots, Y_k) = 0$ for all $m \in M_{n,m}(E)$. By Lemma 2.5 either some linear combination of the $f_i(X_1, \ldots, X_k)$ is zero, in which case the corresponding linear combination of the $f_i(x_1, \ldots, x_k)$ would be an identity for $A$; or every $g_i(Y_1, \ldots, Y_k)$ would be zero, in which case every $g_i(x_1, \ldots, x_k)$ would be an identity for $B$. 


LEMMA 2.7. Given prime product algebras $A \subseteq M_n(E)$ and $B \subseteq M_m(E)$, given $A \sim A'$, $B \sim B'$ (p.i. equivalent) and an $A'-B'$ bimodule $M$, and given polynomials $f_i(x_1, \ldots, x_k), g_i(x_1, \ldots, x_k)$ in the free algebra such that
\[
\sum_{i=1}^d f_i(a_1, \ldots, a_k)xg_i(b_1, \ldots, b_k) = 0
\]
for all $a_1, \ldots, a_k \in A$, $x \in M_{n,m}(E), b_1, \ldots, b_k \in B$, then
\[
\sum_{i=1}^d f_i(a'_1, \ldots, a'_k)x'g_i(b'_1, \ldots, b'_k) = 0
\]
for all $a'_1, \ldots, a'_k \in A'$, $x' \in M$, and $b'_1, \ldots, b'_k \in B'$.

Proof. If the theorem were false, we could choose a counterexample with $d$ as small as possible. By Lemma 2.6, some linear combination of the $f_i$'s would be an identity for $A$, and hence for $A'$. Say $f_1 - c_2f_2 - \cdots - c_df_d$ is such an identity. Then
\[
\sum_{i=2}^d f_i(a_1, \ldots, a_k)x(g_i(b_1, \ldots, b_k) + c_ig_1(b_1, \ldots, b_k)) = 0
\]
for all $a_1, \ldots, a_k \in A$, $x \in M_{n,m}(E), b_1, \ldots, b_k \in B$, but for some $a'_1, \ldots, a'_k \in A', x' \in M$, and $b'_1, \ldots, b'_k \in B'$,
\[
0 \neq \sum_{i=1}^d f_i(a'_1, \ldots, a'_k)x'g_i(b'_1, \ldots, b'_k)
\]
\[
= \sum_{i=2}^d f_i(a'_1, \ldots, a'_k)x'(g_i(b'_1, \ldots, b'_k) + c_ig_1(b'_1, \ldots, b'_k)).
\]
This gives a counterexample in which the number of summands is decreased by 1, and so contradicts the minimality of $d$.

THEOREM 2.8. For any prime product algebra $A = A_1 \circ \cdots \circ A_k$, the ideal of identities $\text{Id}(A_1 \circ \cdots \circ A_k)$ is the product $\text{Id}(A_1) \cdots \text{Id}(A_k)$.

Proof. By induction it suffices to prove that if $A$ and $B$ are arbitrary prime product algebras, then $\text{Id}(A \circ B) = \text{Id}(A) \text{Id}(B)$. The proof will be based on Lewin's theorem [Lc]. (Indeed, this theorem was the inspiration for our definition of the circle product.) Lewin proved that there is an algebra $L$ consisting of $2 \times 2$ matrices of the form
\[
\begin{pmatrix}
a & m \\ 0 & b
\end{pmatrix},
\]
where $a \in U(A)$, the generic p.i. algebra for $A$, $b \in U(B)$, the generic p.i. algebra for $B$, and $m$ lies in a certain $U(A) - U(B)$ bimodule $M$ which
he constructs, with the property that \( \text{Id}(L) = \text{Id}(A) \text{Id}(B) \). We will prove that every identity for \( A \circ B \) is also an identity for \( L \). This implies that \( \text{Id}(A \circ B) \subseteq \text{Id}(L) \). The opposite inclusion is trivial.

Assume by way of contradiction that \( f = \sum_{\sigma} a_\sigma x_{\sigma_1} \cdots x_{\sigma_k} \) is a multilinear identity for \( A \circ B \) but not for \( L \). Since \( f \) is multilinear we can find a substitution with each \( x_i \mapsto U(A) \cup M \cup U(B) \subseteq L \) which does not make \( f \) equal to zero. If all of the \( x_i \) are substituted by elements of \( U(A) \), then \( f \) will be zero, because \( A \subseteq A \circ B \), so \( f \) is an identity for \( A \) and so for \( U(A) \). Likewise, \( f \) will be zero if all of the \( x_i \) are substituted by elements of \( U(B) \). Moreover, if two of the \( x_i \) are in \( M \), then any product \( x_{\sigma_1} \cdots x_{\sigma_k} \) will be zero. So, without loss of generality, we may assume that

\[
f(a'_1, \ldots, a'_{i-1}, m', b'_1, \ldots, b'_{n-i}) \neq 0,
\]

for some \( a'_1, \ldots, a'_{i-1} \in U(A) \), \( m' \in M \), and \( b'_1, \ldots, b'_{n-i} \in U(B) \). The only non-zero product of these elements must have all of the \( a' \) to the left of the \( m' \), and all of the \( b' \) to the right. So, we may write

\[
0 \neq f(a'_1, \ldots, a'_{i-1}, m', b'_1, \ldots, b'_{n-i}) = \sum_{\sigma \in S_{i-1} \cap S} \beta_{\sigma, \tau} a'_{\sigma_1} \cdots a'_{\sigma_{i-1}} m b'_{\tau_1} \cdots b'_{\tau_{n-i}}.
\]

On the other hand, if we make any substitution \( x_1, \ldots, x_{i-1} \mapsto a_1, \ldots, a_{i-1} \in A \), \( x_i \mapsto m \in M_{n,m}(E) \), \( x_1, \ldots, x_n \mapsto b_1, \ldots, b_{n-i} \in B \) and use the fact that \( f \) is an identity for \( A \circ B \), we get

\[
0 = \sum_{\sigma \in S_{i-1} \cap S} \beta_{\sigma, \tau} a_{\sigma_1} \cdots a_{\sigma_{i-1}} m b_{\tau_1} \cdots b_{\tau_{n-i}}.
\]

This contradicts Lemma 2.7 and so completes the proof.

3. \( \exp(s_n(x)) \) IS THE LARGEST

The main result of this section says that, in a sense, standard polynomials are the weakest possible identities. This is part (2) of the following

**Theorem 3.1.** (1) Let \( n \geq 2 \). Then \( \exp(s_n(x)) = \left\lfloor \frac{n}{2} \right\rfloor^2 \), where \( s_n(x) = s_n(x_1, \ldots, x_n) \) is the \( n \)th standard polynomial.

(2) Let \( f(x) = f(x_1, \ldots, x_n) \) be any non-zero polynomial of degree \( n \geq 4 \). Then

\[
\exp(f(x)) \leq \exp(s_n(x)) = \left\lfloor \frac{n}{2} \right\rfloor^2.
\]

The case \( n = 3 \) is a true exception, since it is well known that for the infinite dimensional Grassmann algebra \( E \), \( \exp(E) = \exp([x, y], z) = 2 \).
Proof of (1). Let \( A = A_1 \circ \cdots \circ A_i \) where all the \( A_i \) are verbally prime, and \( A \) satisfies \( s_n(x) \). Since the algebras \( M_k(E) \) and \( M_{k, \ell} \) do not satisfy any \( s_m(x) \), all \( A_i \) must be matrix algebras: \( A_i = F_{a_i} \). By (2-2), \( \exp(A) = a_1^2 + \cdots + a_i^2 \), and \( \exp(s_n(x)) \) is the maximal such sum \( a_1^2 + \cdots + a_i^2 \).

The constraint comes from the fact that \( A = F_{a_1} \circ \cdots \circ F_{a_i} \) does not satisfy any (proper) identity of degree \( \leq 2a - 1 \) where \( a = a_1 + \cdots + a_i \).

This follows from a classical argument of Amitsur and Levitski: for \( 1 \leq i \leq j \leq a, e_{i, j} \in A \), and we can form the non-zero product
\[ e_{1,1} e_{1,2} e_{2,2} \cdots e_{a-1,a} e_{a,a} \]
of length \( 2a - 1 \); however, the product in any other order is zero.

Now, since \( A \) satisfies \( s_n(x) \), it follows that \( 2a \leq n \), i.e., \( a_1 + \cdots + a_i \leq \left\lfloor \frac{n}{2} \right\rfloor \). Thus,
\[ \exp(s_n(x)) = \max \left\{ a_1^2 + \cdots + a_i^2 \mid a_1 + \cdots + a_i \leq \left\lfloor \frac{n}{2} \right\rfloor \right\} = \left\lfloor \frac{n}{2} \right\rfloor^2. \]

Proof of (2). This is based on the following lemmas and remarks. Recall that \( \pideg(R) \) is the minimal degree of an identity satisfied by \( R \).

Lemma 3.2. Let \( R \) be a verbally prime algebra with \( \pideg(R) = n \). Then \( \pideg(M_k(R)) \geq kn. \) In particular, \( \pideg(M_k(E)) \geq 3k. \)

Proof. This easily follows since
\[ R^k = R \circ \cdots \circ R \subseteq (M_k(R)) \]
and \( \pideg(A_1 \circ \cdots \circ A_i) = \pideg(A_1) + \cdots + \pideg(A_i) \) by Theorem 2.8.

Proposition 3.3 [P2]. \( \pideg(M_3(E)) \geq 7. \)

Note that since \( M_2(E) \circ E \subseteq M_3(E) \) and \( M_2(E) \circ M_2(E) \circ E \subseteq M_3(E) \), hence \( \pideg(M_3(E)) \geq 10 \) and \( \pideg(M_5(E)) \geq 17. \)

Lemma 3.4. \( \pideg(M_{k, \ell}) \geq 2(k + \ell), \) i.e., \( M_{k, \ell} \) satisfies no identity of degree \( \leq 2(k + \ell) - 1. \)

Proof. Essentially, one can repeat here the previous “Amitsur–Levitski” argument, now in \( M_{k, \ell} \), using \( g_{i,j} e_{i,j}, 1 \leq i \leq j \leq k + \ell, \) and with appropriate \( g_{i,j} \in E. \)

The proof of part (2) of Theorem 3.1 obviously follows from

Lemma 3.5. Let
\[ A = F_{a_1} \circ \cdots \circ F_{a_i} \circ M_{p_1, q_1} \circ \cdots \circ M_{p_i, q_i} \circ M_{b_1} \circ \cdots \circ M_{b_i}(E) \]
satisfy an identity of degree \( n, n \geq 4. \) Then \( \exp(A) \leq \left\lfloor \frac{n}{2} \right\rfloor^2. \)
Proof. Assume first that \( r = s = 0, t = 1 \), hence \( A = M_s(E) \) for some \( u \), and further assume that \( u = 2v + 1 \) is odd.

If \( v = 0 \) then \( A = E \), and \( \exp(E) = 2 < 4 \leq \left\lfloor \frac{u}{2} \right\rfloor^2 \).

If \( v = 1 \) then \( A = M_3(E) \) and by the remark following Proposition 3.3 and by Lemma 3.4, \( 10 \leq \pideg(A) \leq n \); therefore

\[
\exp(A) = 18 < 5^2 = \left\lfloor \frac{10}{2} \right\rfloor^2.
\]

Similarly, if \( v = 2, A = M_5(E), 17 \leq \pideg(A) \leq n \), and

\[
\exp(A) = 50 < 8^2 = \left\lfloor \frac{17}{2} \right\rfloor^2 \left( \leq \left\lfloor \frac{n}{2} \right\rfloor^2 \right).
\]

For any \( v \geq 0, 3(2v + 1) \leq \pideg(A) \leq n \), hence \( 3v + 1 \leq \left\lfloor \frac{u}{2} \right\rfloor \), and if \( 3 \leq v \) then

\[
\exp(A) = 2(2v + 1)^2 \leq (3v + 1)^2 \leq \left\lfloor \frac{n}{2} \right\rfloor^2.
\]

Turn now to the general case. Note that

\[
\exp(A) = a_1^2 + \cdots + a_t^2 + (p_1 + q_1)^2 + \cdots + (p_s + q_s)^2 + 2(b_1 + \cdots + b_t),
\]

Denote \( h = a_1 + \cdots + a_t + p_1 + q_1 + \cdots + p_s + q_s \) and \( u = b_1 + \cdots + b_t \).

If \( t \geq 2 \) then \( 2(b_1^2 + \cdots + b_t^2) \leq 2(b_1 + \cdots + b_t)^2 - 2 = 2u^2 - 2 \), so \( \exp(A) \leq h^2 + 2u^2 - 2 \). Now using Lemma 3.2, \( \pideg(A) \leq 2h + 3u \leq n \), hence it suffices to show that

\[
h^2 + 2u^2 - 2 \leq \left\lfloor \frac{2h + 3u}{2} \right\rfloor^2 = \left( h + \left\lfloor \frac{3u}{2} \right\rfloor \right)^2,
\]

and this is easily verified in both cases when \( u = 2v \) and \( u = 2v + 1 \).

Assume therefore that \( t \leq 1 \) and show that \( \exp(A) \leq \left\lfloor \frac{u}{2} \right\rfloor^2 \). If \( h = 0 \) then \( A = E \) and we only need to check the case \( u = 2v \) is even. In this case, as well as in the cases \( h \geq 1 \), it suffices to show that \( h^2 + 2u^2 \leq \left\lfloor \frac{2h + 3u}{2} \right\rfloor^2 \) (\( \leq \left\lfloor \frac{u}{2} \right\rfloor^2 \)), or equivalently, that \( 2u^2 \leq 2h \left\lfloor \frac{3u}{2} \right\rfloor + \left\lfloor \frac{3u}{2} \right\rfloor^2 \).

Again, these are easily verified in both cases \( u = 2v \) and \( u = 2v + 1 \).

Q.E.D.

Remark. Let \( \pideg(R) = n \). The previous best known bound for \( \exp(R) \) was \( \exp(R) \leq (n - 1)^2 \) [L, R1], while here, Theorem 3.1.2 gives \( \exp(R) \leq \left\lfloor \frac{u}{2} \right\rfloor^2 \), which is, roughly, an improvement by about a factor of 1/4.
4. AMITSUR’S CAPELLI-TYPE POLYNOMIALS

Let $\lambda$ be a partition of $n$ (i.e., $\lambda \vdash n$) with $\chi_{\lambda}$ the corresponding irreducible $S_n$ character. In [AR] Amitsur introduced the Capelli-type polynomials

$$E_{\lambda}^*[x; y] = \sum_{\sigma \in S_n} \chi_{\lambda}(\sigma)x_{\sigma(1)}y_1x_{\sigma(2)}y_2 \cdots y_{n-1}x_{\sigma(n)}.$$ 

When $\lambda = ((k + 1)^{l+1})$ is the $(k + 1) \times (l + 1)$ rectangle, we denote

$$E_{((k+1)^{l+1})}^*[x; y] = E_{k, l}^*[x; y]$$

and $\exp(E_{k, l}^*[x; y]) = E_{k, l}$. The polynomial $E_{k, l}^*[x; y]$ characterizes when cocharacters are contained by the $(k, l)$ hook $H(k, l)$. This is the following obvious corollary of Theorem B in [AR].

**Corollary 4.1 [AR].** Let $\lambda = ((k + 1)^{l+1})$ be the $(k + 1) \times (l + 1)$ rectangle, and let $A$ be a p.i. algebra. Then the cocharacters $\chi_n(A)$ are contained by the $(k, l)$ hook $H(k, l)$ if and only if $A$ satisfies the identity $E_{k, l}^*[x; y] = 0$.

We use this corollary to calculate $\exp(E_{k, l}^*[x; y]) = E_{k, l}$.

Since the cocharacters of $E_{k, l}^*[x; y]$ are contained in $H(k, l)$, it follows [BR1, BR2] that

$$\exp(E_{k, l}^*[x; y]) \leq k + l.$$ 

We write $E_{k, l} = k + l - g$ and below we investigate the gap $g \geq 0$.

**Remark 4.2.** Let

$$A = F_{a_1} \circ \cdots \circ F_{a_r} \circ M_{p_1, q_1} \circ \cdots \circ M_{p_s, q_s} \circ M_{b_1}(E) \circ \cdots \circ M_{b_t}(E),$$

let $w = r + s + t - 1$, let

$$k_1 = a_1^2 + \cdots + a_r^2 + p_1^2 + q_1^2 + \cdots + p_s^2 + q_s^2 + b_1^2 + \cdots + b_t^2,$$

and

$$l_1 = 2(p_1q_1 + \cdots + p_sq_s) + b_1^2 + \cdots + b_t^2.$$ 

It follows from [BR4] that

$$\chi_n(A) \subseteq H(k_1, l_1) \hat{\otimes} (\chi_{1}^{\otimes w}),$$

where $\hat{\otimes}$ represents the “Littlewood–Richardson” outer product. Given $u, v, \in \mathbb{N}$, if $w < (u + 1)(v + 1)$ it is easy to check that $\chi_1^{\otimes w} \subseteq H(u, v)$ and similarly that $H(k_1, l_1) \hat{\otimes} (\chi_{1}^{\otimes w}) \subseteq H(k_1 + u, l_1 + v)$. It follows that the above algebra $A$ (with corresponding $k_1, l_1$, and $w$) satisfies $E_{k, l}^*[x; y]$ if and only if there exist $u, v$ such that $k_1 + u \leq k$, $l_1 + v \leq l$, and
\[ w < (u + 1)(v + 1). \] By Theorem 2.4 we get

**Proposition 4.3.** \( E_{k, l} = \exp(E_{k, l}^* [x; y]) \) is the maximum value of

\[
a_1^2 + \cdots + a_r^2 + (p_1 + q_1)^2 + \cdots + (p_s + q_s)^2 + 2(b_1^2 + \cdots + b_t^2)
\]  

(4.3.1)

with various \( r, s, t, a_i's, b_j's, p_i's \) and \( q_j's \), subject to the restrictions that there exist \( u, v \in \mathbb{N} \) satisfying

\[
r + s + t \leq (u + 1)(v + 1),
\]

\[
a_1^2 + \cdots + a_r^2 + p_1^2 + q_1^2 + \cdots + p_s^2 + q_s^2 + b_1^2 + \cdots + b_t^2 + u \leq k,
\]

and

\[ 2(p_1q_1 + \cdots + p_sq_s) + b_1^2 + \cdots + b_t^2 + v \leq l. \]  

(4.3.2)

As shown in [MRZ], in general, the gap \( g = k + l - \exp(E_{k, l}^* [x; y]) \) can be arbitrarily large. However, if we further assume that \( l \leq k \), we show below (Proposition 4.5) that as in the Capelli case, this gap is bounded by 3.

The **verbally prime hooks** are \( H(p^2 + q^2, 2pq) \) (corresponding to the algebra \( A = M_{p, q} \), here \( M_{p, 0} = M_p(F) \)) and \( H(b^2, b^2) \) (where \( A = M_b(E) \)). In the “strip” (i.e., “Capelli”) case, these are \( H(p^2, 0) \), corresponding to the squares \( p^2 \in \mathbb{N} \).

By corresponding \( H(k, l) \leftrightarrow (k, l) \), the verbally prime hooks give rise to the following “super” (or \( \mathbb{Z}_2 \)-graded) squares, which we call “generalized” squares:

\[ \mathcal{P} = \{(r^2, r^2), (r^2 + s^2, 2rs) \mid r, s \in \mathbb{N} \}. \]

The cocharacters of \( E_{k, l}^* \) are supported on \( H(k, l) \leftrightarrow (k, l) \), and via cocharacters, \( \{E_{k, l}^* \mid k \geq l \geq 0\} \) corresponds to \( \mathcal{R} = \{(k, l) \mid k \geq l \geq 0\} \).

In general, if the hook of \( A_i \) is \( H(k_i, l_i) \), \( i = 1, 2 \) (i.e., the cocharacters of \( A_i \) are supported on \( H(k_i, l_i) \)) then by the Littlewood–Richardson rule and by [BR4], the cocharacters of \( A_1 \circ A_2 \) are contained (i.e., supported) in \( H(k_1 + k_2, l_1 + l_2) \mathop{\otimes} \chi_1 \).

Define \( (a, b) + (c, d) = (a + c, b + d) \). Clearly, \( \mathcal{P} \subseteq \mathcal{R} \) and \( \mathcal{R} \) is closed under summations. We say that \( (k, l) \in \mathcal{R} \) is of class \( r \) if there exist \( z_1, \ldots, z_r \in \mathcal{P} \) such that \( z_1 + \cdots + z_r = (k, l) \), with \( r \) minimal. Define

\[ \mathcal{C}_r = \{(k, l) \in \mathcal{R} \mid (k, l) \text{ is of class } r \}. \]

Based on computer evidence, in an earlier version of this paper we conjectured Theorem 4.4 below. That theorem is proved in [CR].

**Theorem 4.4 [CR].** In \( \mathcal{R} \), every element is of class \( \leq 6 \).
Applying Theorem 4.4 [CR] we can prove

**Theorem 4.5.** Let \( k \geq l \geq 0 \). Then

\[
k + l - 3 \leq E_{k,l} \leq k + l.
\]

**Proof.** Case 1. \( l = 1 \). If \( k = 1 \) (or \( k = 2 \)) there is nothing to prove. If \( k \geq 2 \), represent \((k - 2, l - 1) = (k - 2, 0)\) as a sum of (at most) six generalized squares (i.e., elements of \( \mathcal{P} \)), then proceed as in Case 2 below.

Case 2. \((k \geq l \geq 2)\). Write \((k - 1, l - 2)\) as a sum of (at most) six generalized squares, \((k - 1, l - 2) = (a^2_1, 0) + \cdots + (a^2_r, 0) + (p^2_1 + q^2_1, 2p_1q_1) + \cdots + (b^2_i, b^2_i)\) with \( r + s + t \leq 6 \). Let \( u = 1, v = 2 \). Then

\[
r + s + t \leq (u + 1)(v + 1) = 6,
\]

\[
a^2_1 + \cdots + a^2_r + p^2_1 + q^2_1 + \cdots + p^2_s + q^2_s + b^2_1 + \cdots + b^2_t + u
\]

\[
= k - 1 + 1 = k,
\]

and similarly

\[
2(p_1q_1 + \cdots + p_sq_s) + b^2_1 + \cdots + b^2_t + v = l - 2 + 2 = l.
\]

Trivially,

\[
a^2_1 + \cdots + a^2_r + (p_1 + q_1)^2 + \cdots + (p_s + q_s)^2 + 2(b^2_1 + \cdots + b^2_t)
\]

\[
= (k - 1) + (l - 2) = k + l - 3.
\]

By Proposition 4.3 it follows that \( k + l - 3 \leq E_{k,l} \). Q.E.D.

The possible gaps. Let \((k, l) \in \mathcal{H}\) and denote

\[
g = g(k, l) = k + l - \exp(E_{k,l}) = k + l - E_{k,l}.
\]

- \( g = 0 \) if and only if \((k, l) \in \mathcal{P}\), i.e., if and only if \((k, l)\) is a generalized square.
- \( g = 1 \) if and only if \( g \neq 0 \) but either \((k - 1, l)\) or \((k, l - 1)\) is of class \( \leq 2 \) in \( \mathcal{H} \), i.e., is a sum of at most two elements of \( \mathcal{P} \).
- \( g = 2 \) if and only if \( g \neq 0, 1 \), and either \((k - 2, l)\) or \((k, l - 2)\) is of class \( \leq 3 \) in \( \mathcal{H} \) or \((k - 1, l - 1)\) is of class \( \leq 4 \) in \( \mathcal{H} \).
- \( g = 3 \) if and only if \( g \neq 0, 1, 2 \).

For example, if \( l = 1 \) then \( g \leq 2 \). Indeed, in that case \((k - 1, l - 1) = (k - 1, 0)\) and by the four-squares-theorem, \( k - 1 \) is a sum of four squares in \( \mathbb{N} \), hence is of class \( \leq 4 \) in \( \mathcal{H} \).
The General $\exp(E_1^+[x; y])$

The cell $c = (k + 1, l + 1)$ determines the $(k + 1) \times (l + 1)$ rectangle $R_c$, the infinite hook $H(k, l)$ (of the partitions avoiding $c$), and the exponent $\exp(E_{k,l}^+[x; y])$.

It is possible to prove

**Theorem 4.6.** Let $\lambda = (\lambda_2, \lambda_1, \ldots)$ be a partition. Call $c \in \lambda$ extreme if $c = (i, \lambda_i)$ and $\lambda_i > \lambda_{i+1}$. Then

$$\exp(E_{k}^+[x; y]) = \max_{c \in \lambda} \{ \exp(E_{k,c}^+[x; y]) \}.$$ 

**Some Special Cases**

It is interesting to calculate $E_{k,l} = \exp(E_{k,l}^+)$—with the corresponding relative densities in $\mathbb{N}$—in some special cases.

**Example 1:** $l = 1$. Trivially, $E_{1, 1} = E_{2, 1} = \exp(E) = 2$ and $E_{3, 1} = \exp(E \circ F) = 3$. If $k \geq 2$ then $H(k, 1)$ is not vph (i.e., $(k, 1) \not\in \mathcal{A}$), hence $E_{k, 1} \leq (k + 1) - 1 = k$.

**Claim.** Let $k \geq 2$. Then $k - 1 \leq E_{k, 1} \leq k$. Denote $W_i = \{ k \mid E_{k, i} = k - i \}, i = 0, 1$, with $d(W_i)$ the corresponding density in $\mathbb{N}$. Then both $W_0$, $W_1$ are infinite, with $d(W_i) = i, i = 0, 1$.

**Proof.** This can easily be derived from (4.3.1) and (4.3.2). Let $k - 1 = a_1^2 + \cdots + a_s^2$. Thus $r \leq 4, s \geq 0$ in (4.3.1). Let $u = v = 1$. Then $w \leq 3 < (u + 1)(v + 1)$, as required in (4.3.2). We get $k - 1 = a_1^2 + \cdots + a_s^2 \leq E_{k, 1} \leq k$.

Now $E_{k, 1} = k$ if and only if either $k - 2 = a^2(A = E \circ F_a)$ or $k = a_1^2 + a_2^2$.

$(A = F_{a_1} \circ F_{a_2})$ Clearly, $W_0$ is infinite, and by a classical theorem of E. Landau (see [MRZ, Addendum]), $d(W_0) = 0$. It follows that $d(W_i) = 1$.

**Example 2:** $2 \leq l \leq k$. Here $E_{2, 2} = \exp(M_{1, 1}) = 4$. If $k \geq 3$, $H(k, 2)$ is not vph, hence $E_{k, 2} \leq k + 1$.

**Claim.** Let $k \geq 3$. Then $k \leq E_{k, 2} \leq k + 1$. Denote $W_i = \{ k \mid E_{k, 2} = k + 1 - i \}, i = 0, 1$, with $d(W_i)$ the corresponding density in $\{ (k, 1) \mid k \in \mathbb{N} \}$.

Then $d(W_i) = i, i = 0, 1$.

**Proof.** The proof is based on the remark that for any $n \in \mathbb{N}, n$ or $n + 2$ is a sum of—at most—three squares. (If $n$ is not then $n = 4'(8s + 7)$. If also $n + 2$ is not then $n + 2 = 4''(8b + 7)$, hence $2 = 4''(8b + 7) - 4'(8s + 7)$, so $r = 0$ or $a = 0$. But 2 is even, hence $r = a = 0$, therefore $2 = 8(b - s)$, a contradiction.)
There is now an easy argument based on (4.3.1) and (4.3.2), but we prefer to prove the lower bound by directly constructing appropriate verbally prime-product algebras. If \( k = a_1^2 + a_2^3 + a_3^3 \), let \( A = F_{a_1} + F_{a_2} + F_{a_3} \). Then \( \exp(A) = k \) and the hook of \( A \) is contained by \( H(k, 2) \), hence \( k \leq E_{k,2} \), and similarly if \( k + 2 = b_1^2 + b_2^3 + b_3^3 \), with \( A = E \circ F_{b_1} \circ F_{b_2} \circ F_{b_3} \).

To calculate the (relative) densities, assume \( E_{k,2} = k + 1 \) and find all possible \( A \)'s as in Remark 4.2. By Proposition 4.3, \( r + s + t \leq 2 \). The only possibilities here are either \( k - 1 = a^2 (A = E \circ F_a) \) or \( k - 3 = a^2 (A = M_{1,1} \circ F_a) \).

This shows that \( W_0 \) is infinite and that \( d(W_0) = 0 \). Hence \( d(w_1) = 1 \).

**Example 3:** \( 3 \leq l \leq k \). Here we

**Claim.** Let \( 3 \leq k \). Then \( k \leq E_{k,3} \leq k + 2 \). Let \( W_i = \{ k \mid E_{k,3} = k + 2 - i \}, i = 0, 1, 2 \). Then all three sets are infinite and with the following densities in \( \mathbb{N} \): \( d(W_0) = 0 \), \( d(W_1) = 5/6 \), and \( d(W_2) = 1/6 \).

**Proof.** \( H(k, 3) \) is never vph, hence \( E_{k,3} \leq k + 3 - 1 = k + 2 \). For example, \( E_{3,3} = 5(A = M_{1,1} \circ F) \). Assume \( k \geq 4 \). Write \( k - 4 = a_1^2 + \cdots + a_4^2 \), \( A = M_{1,1} \circ F_{a_1} \circ \cdots \circ F_{a_4} \) and deduce that \( k \leq E_{k,3} \).

Let \( E_{k,3} = k + 2 \). Then, in Remark 4.2, \( r + s + t \leq 2 \). The only possibilities are that \( (A = M_{1,1} \circ F_a) \), hence \( k - 2 = a^2 \). This shows that \( d(W_0) = 0 \) and that \( W_0 \) is infinite.

Similarly, \( E_{k,3} = k + 1 \) if and only if the following holds: \( k - 2 \) is not a square, but either \( k - 1 = b_1^2 + b_2^2 \), which yields a subset with density \( 0 \), or \( k - 3 = a_1^2 + a_2^2 + a_3^2 \), which yields a subset with density \( 5/6 \) [MRZ]. Thus, \( d(W_1) = 5/6 \). Obviously, this implies for the remaining set \( d(W_2) = 1/6 \).

**Example 4:** \( E_{k,k} \). Here it can be shown that \( 2k - 2 \leq E_{k,k} \leq 2k \). Let \( W_i = \{ k \mid E_{k,k} = 2k - i \}, i = 0, 1, 2 \). Then all \( W_i \) are infinite, with densities \( d(W_0) = d(W_1) = 0 \) and \( d(W_2) = 1 \). We leave the details to the reader.

### 5. POWERS OF HIGHER COMMUTATORS

A polynomial of the form

\[
x_1 \cdots x_n - \sum_{\sigma(1) \neq 1} \alpha_{\sigma} x_{\sigma(1)} \cdots x_{\sigma(n)}
\]

is called a \( J \)-polynomial. The most important example of \( J \)-polynomials is (higher) commutators. It is known that \( 2 \times 2 \) matrices do not satisfy any \( J \)-identities. (Proof. Let \( x_1 = e_{21} \) and \( x_i = e_{11} \) for \( i \geq 2 \). By a theorem of Amitsur, they also would not satisfy any powers of \( J \)-polynomials. Also \( M_{1,1} \) does not satisfy and \( J \)-identity (essentially the same argument as for \( M_2(F) \)).
However, $M_{1,1}$ might satisfy a power of a $J$-polynomial, for example, $[x, y]^3$ [P1]. Clearly, $M_r(F) \subseteq M_r(E)$, $M_{r,s}$. Hence, if $2 \leq r$, $M_r(F)$, $M_r(E)$, and $M_{r,s}$ satisfy no power of a $J$-polynomial: only $F$, $E$, and $M_{1,1}$ might. Thus, up to p.i. equivalence, the only verbally prime algebras which could satisfy $J$-identities are $F$ and $E$, while those that might satisfy powers of $J$-polynomials are $F$, $E$, and $M_{1,1}$.

Henceforth, $f(x_1, \ldots, x_n)$ will be a fixed higher commutator (more generally, $J$-polynomial) and for a p.i. algebra $A$ we will let $d_f(A) = d(A)$ be the minimum $a \leq \infty$ such that $f(x_1, \ldots, x_n)^a$ is an identity for $A$. Note that $d(F)$ is always 1 or $\infty$. The latter case is uninteresting, so we will always take $f$ to be an identity for $F$.

**Lemma 5.1.** Let $A$ and $B$ be prime product algebras (see Section 2) with $d(A) = \alpha$ and $d(B) = \beta$. Then $d(A \circ B) = \alpha + \beta$.

**Proof.** It follows from Remark 2.1 that $d(A \circ B) \leq \alpha + \beta$. To complete the proof we need to show that $A \circ B$ does not satisfy $f(x_1, \ldots, x_n)^{\alpha+\beta-1}$. Let

$$
x_1 = \begin{pmatrix} a_1 & m \\ 0 & b_1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} a_2 & 0 \\ 0 & b_2 \end{pmatrix}, \ldots, \quad x_n = \begin{pmatrix} a_n & 0 \\ 0 & b_n \end{pmatrix}
$$

be in $A \circ B$. Here $m$ is a matrix over $E$ of appropriate dimensions. Then

$$f(x_1, \ldots, x_n) = \left( \begin{array}{c} f(a_1, \ldots, a_n) \\ 0 \\ f(b_1, \ldots, b_n) \end{array} \right) = \left( \begin{array}{c} f(a) \\ 0 \\ f(b) \end{array} \right),$$

where $\tilde{m} = mb_2 \cdots b_n$ terms with lower degree in the $b_i$’s. Now,

$$f(x_1, \ldots, x_n)^{\alpha+\beta-1} = \left( \begin{array}{cccc} 0 & \sum_{i=0}^{\alpha+\beta-2} f(a)^i \tilde{m}^i f(b)^{\alpha+\beta-2-i} \\ 0 & \tilde{m}^{\alpha+\beta-1} b_2 \cdots b_n f(b)^{\beta-1} \end{array} \right).$$

In the $(1, 2)$-component the unique non-zero summand corresponds to $i = \alpha - 1$; its unique term with degree $(\beta - 1)n + (n - 1)$ in the $b_i$’s is

$$f(a)^{\alpha-1} m b_2 \cdots b_n f(b)^{\beta-1}.$$
Indeed, write \( f(x_1 + 1, x_2, \ldots, x_n) = \sum_{i=0}^{k} f_i \), where \( f_i \) is a polynomial of degree \( i \) in \( x_1 \). Then, since \( B \) satisfies \((x_1 + 1)f(x_1 + 1, x_2, \ldots, x_n)\) it will satisfy the part of degree \( i \) in \( x_1 \). This equals

\[
x_1 f_{i-1} + f_i,
\]

for all \( i = 0, \ldots, k \), with the understanding that \( f_{-1} = 0 \). We now prove by induction that each \( f_i \) is an identity for \( A \). This will complete the proof because \( f = f_k \). To start the induction, set \( i = 0 \) in the above equation to get that \( f_0 \) is an identity for \( B \). To do the induction step, note that if \( f_{i-1} \) is an identity for \( B \) then \( f_i \) must be an identity for \( B \).

**Corollary 5.2.** If \( B_1, \ldots, B_n \) are verbally prime, then \( d(B_1 \circ \cdots \circ B_n) = d(B_1) + \cdots + d(B_n) \). Thus, if \( B_1 \circ \cdots \circ B_n \) satisfies \( f(x)^k \) then \( d(B_1) + \cdots + d(B_n) \leq k \).

Here now is our main technical devise:

**Lemma 5.3.** Given a higher commutator \( f(x_1, \ldots, x_n) \), or more generally a \( J \)-polynomial and let \( f(x_1, \ldots, x_n)^\beta \) be an identity for \( E \) and \( f(x_1, \ldots, x_n)^\gamma \) be an identity for \( M_{1,1} \), with \( \beta \) and \( \gamma \) minimal. Then

\[
\exp(f(x_1, \ldots, x_n)^k) = \max\{a + 2b + 4c \mid a + \beta b + \gamma c \leq k\}.
\]

**Proof.** A prime product \( B_1 \circ \cdots \circ B_t \) satisfies \( f(x)^k \) only if each factor is \( F, E \), or \( M_{1,1} \). Moreover, if \( a, b, \) and \( c \) are the number of factors of each of \( F, E \), and \( M_{1,1} \), respectively, then the previous corollary implies that such a product satisfies \( f(x)^k \) precisely when \( a + \beta b + \gamma c \leq k \). The proof now follows from Eq. (2.2) using \( \exp(F) = 1 \), \( \exp(E) = 2 \), \( \exp(M_{1,1}) = 4 \).

**Corollary 5.4.** Let \( f(x_1, \ldots, x_n) \) be a \( J \)-polynomial. Then

\[
\exp(f(x)) = \begin{cases} 1 & \text{if } f(x) \text{ is not an identity of } E \\ 2 & \text{if } f(x) \text{ is an identity of } E. \end{cases}
\]

**Corollary 5.5.** Let \( f(x_1, \ldots, x_n) \) be a \( J \)-polynomial. Then

\[
\exp(f(x)^k) \leq 2k.
\]

**Proof.** This easily follows from Corollary 5.2 and from the following fact: let \( 1 \leq \beta, 2 \leq \gamma \), and let \( a + \beta b + \gamma c \leq k \). Then \( a + 2b + 4c \leq 2k \).

**Lemma 5.6.** If \( f = [x, y] \), then, keeping the notations of Lemma 5.3, \( \beta = 2 \) and \( \gamma = 3 \); and if \( f = [x_1, \ldots, x_n] \) with \( n \geq 3 \), then \( \beta = 1 \) and \( 2 \leq \gamma \leq 3 \).
Proof. The identities of $E$ are well known and the present instances are easy to verify. The fact that commutators cubed vanish in $M_{1,1}$ is due to Popov, see [P1]. It remains to check that higher commutators don’t vanish in $M_{1,1}$.

Kemer showed that $M_{1,1}$ is p.i. equivalent to $E \otimes E$, hence we will work in $E \otimes E$ instead. The algebra $E$ has a $\mathbb{Z}/2\mathbb{Z}$ grading, $E = E_0 + E_1$ and so $E \otimes E$ has a $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$-grading, $E \otimes E = E_0 \otimes E_0 + E_0 \otimes E_1 + E_1 \otimes E_0 + E_1 \otimes E_1$. In the higher commutator, if we take $x_1 \in E_0 \otimes E_1$ and the remaining $x_i \in E_1 \otimes E_1$ then we don’t get zero in general.

Theorem 5.7. If $k = 3q + r$, $0 \leq r \leq 2$, then $\exp([x, y]^k) = 4q + r$, and for all $d \geq 3$, $\exp([x_1, \ldots, x_d]^k) = 2k$.

Proof. In the case of the commutator we need to maximize $a + 2b + 4c$ subject to $a + 2b + 3c \leq k$. Clearly we do this by taking $c$ as large as possible, so $c = q$ and $a = r$. For higher commutators we have the restraint $a + b + 3c \leq k$ or possibly $a + b + 2c \leq k$. In either case we can maximize $a + 2b + 4c$ by taking $b = k$.

The calculations we did apply equally well to the Engel identity $[x, y, \ldots, y] = 0$. Although this is not a $J$-polynomial because it is not linear, our proof did not use linearity in any essential way. Now, if $f(x, y) = [x, y, \ldots, y]$ has at least two $y$’s, then $f(x, y)$ is an identity for $F$ and for $E$, but not for $M_{1,1}$; and $f(x, y)^2$ is an identity for $M_{1,1}$. This implies:

Theorem 5.8. $\exp([x, y, \ldots, y]^k) = 2k$.

6. POWERS OF STANDARD IDENTITIES

In this section we investigate the exponential behaviour of $s_n(x_1, \ldots, x_n)^k$. In light of Theorem 2.3, our main job is to investigate which prime product algebras satisfy which powers of standard polynomials. Here is a theorem of Amitsur that we will need.

Theorem (Amitsur). The matrix algebra $F_m$ satisfies $f(x_1, \ldots, x_n)^k$ if and only if it satisfies $f(x_1, \ldots, x_n)$.

The next two classes of algebras we consider are $E$ and $M_{1,1}$.

Lemma 6.1. The algebra $E$ satisfies $s_n(x_1, \ldots, x_n)^k$ for all $n, k \geq 2$, and $M_{1,1}$ satisfies $s_n(x_1, \ldots, x_n)^k$ for all $n \geq 2$ and $k \geq 3$. 
Proof. It is fairly well known and easy to verify that $E$ satisfies $[x, y][x, z]$. If we replace $z$ by $uz$ and use the Jacobi identity, it follows that $E$ satisfies $[x, y]u[x, z]$. But the standard identity $s_n(x_1, \ldots, x_n)$ can be written as a linear combination of terms each involving $[x_1, y]$, for some $y$. Hence, the square will be zero.

As for $M_{1,1}$, Popov proved that it satisfies $[x, y][x, z][x, u]$. As in the previous case, this implies that $[x, y]a[x, z]b[x, u]$ which in turn implies that the cube of any standard identity is zero.

We define the function $f(k, \ell)$ as follows. The algebra $E$ has a natural $Z/2Z$-grading in which the degree one elements $E_0$ form the center and the degree one elements $E_1$ anticommute. This defines a $Z/2Z \times Z/2Z$ grading on the tensor product $E \otimes E$. Our main interest will be in $E_0 \otimes E_1$ and in $E_1 \otimes E_0$. Elements of these two subspaces commute with each other and anticommute among themselves. Let $e_1, \ldots, e_k \in E_0 \otimes E_1$ and $g_1, \ldots, g_\ell \in E_1 \otimes E_0$ have non-zero product. Then we may define the function $f(k, \ell)$ via

$$s_{k+\ell}(e_1, \ldots, e_k, g_1, \ldots, g_\ell) = f(k, \ell)e_1 \cdots e_k g_1 \cdots g_\ell.$$  

Note that

$$s_{k+\ell}(g_1, \ldots, g_k, e_1, \ldots, e_\ell) = f(k, \ell)g_1 \cdots g_k e_1 \cdots e_\ell.$$  

Lemma 6.2. The function $f(k, \ell)$ defined above equals

$$\frac{1 + (-1)^{k+\ell}}{2} k! \ell! \left( \left[ \frac{k+\ell}{2} \right] \right).$$

In particular, $f(k, \ell) \geq 0$ for all $k$ and $\ell$ and it is zero precisely when both are odd.

Proof. Use the expansion

$$\sum_{i=1}^n x_i s_{n-1}(x_1, \ldots, \hat{x}_i, \ldots, x_n)$$

and induction on $n$ to get the recurrence relation $f(k, \ell) = kf(k-1, \ell) + (-1)^{k+\ell}f(k, \ell-1)$. The proof of the lemma follows from verifying that the right hand side satisfies the same recurrence.

Lemma 6.3. The polynomial $s_a(x_1, \ldots, x_a) s_{n+b}(x_1, \ldots, x_a, y_1, \ldots, y_b) \times s_b(y_1, \ldots, y_b)$ is not an identity for $E \otimes E$ for any $a, b \geq 0$.

Proof. Without loss of generality, we may assume that $a$ and $b$ are both even by substituting $x_a \mapsto x_a + 1$ and $y_b \mapsto y_b + 1$ and using the fact that

$$s_n(x_1, \ldots, x_{n-1}, 1) = \begin{cases} 0, & \text{if } n \text{ is even} \\ s_{n-1}(x_1, \ldots, x_{n-1}), & \text{if } n \text{ is odd}. \end{cases}$$
Now, let
\[ e_1, \ldots, e_a, \epsilon_1, \ldots, \epsilon_b \in E_0 \otimes E_1 \]
\[ g_1, \ldots, g_a, \gamma_1, \ldots, \gamma_b \in E_1 \otimes E_0 \]
and let each \( x_i = e_i + g_i \) and \( y_i = \epsilon_i + \gamma_i \). We will show that under this substitution \( s_a(x)s_{a,b}(x, y)s_b(\gamma) \) is not zero. Using multilinearity and the fact that any term with degree two or more in some \( e, g, \epsilon, \) or \( \gamma \) is zero, we get
\[
s_a(e_1 + g_1, \ldots, e_a + g_a)(\epsilon_1 + g_1, \ldots, \epsilon_b + \gamma_1, \ldots, \epsilon_b + \gamma_1, \ldots) = \sum s_a(u_1, \ldots, u_a)s_{a+b}(u_1', \ldots, u_a', w_1', \ldots, w_b')s_b(w_1, \ldots, w_b),
\]
where for each \( i \), \( \{u_i, u_i'\} = \{e_i, g_i\} \) and \( \{w_i, w_i'\} = \{\epsilon_i, \gamma_i\} \). For each summand
\[
s_a(u_1, \ldots, u_a)s_{a+b}(u_1', \ldots, u_a', w_1', \ldots, w_b')s_b(w_1, \ldots, w_b) \quad (6.1)
\]
there exist \( 1 \leq k \leq a \), \( 1 \leq l \leq b \), and permutations \( \sigma \in S_a \) and \( \tau \in S_b \) such that
\[
(u_{\sigma(1)}, \ldots, u_{\sigma(a)}) = (e_{\sigma(1)}, \ldots, e_{\sigma(k)}, g_{\sigma(k+1)}, \ldots, g_{\sigma(a)})
\]
\[
(w_{\tau(1)}, \ldots, w_{\tau(b)}) = (\epsilon_{\tau(1)}, \ldots, \epsilon_{\tau(l)}, \gamma_{\tau(l+1)}, \ldots, \gamma_{\tau(b)}),
\]
where \( \sigma(1) < \cdots < \sigma(k), \sigma(k+1) < \cdots < \sigma(a), \tau(1) < \cdots < \tau(\ell), \) and \( \tau(\ell+1) < \cdots < \tau(b) \). Note that
\[
(u_1', \ldots, u_a', w_1', \ldots, w_b') = (g_{\sigma(1)}, \ldots, g_{\sigma(k)}, e_{\sigma(k+1)}, \ldots, e_{\sigma(a)}, \gamma_{\tau(1)}, \ldots, \gamma_{\tau(\ell+1)}, \ldots, \gamma_{\tau(b)}).
\]
By Lemma 6.2, (6.1) is zero unless \( k \) and \( \ell \) are even. We will show that in this case it equals \( \alpha e_1 \cdots e_a g_1 \cdots g_a e_1 \cdots e_b \gamma_1 \cdots \gamma_b \) for some \( \alpha > 0 \).

At this point it is helpful to introduce some shorthand. We let
\[
I = \{\sigma(1) < \cdots < \sigma(k)\},
\]
\[
I' = \{\sigma(k+1) < \cdots < \sigma(a)\},
\]
\[
J = \{\tau(1) < \cdots < \tau(\ell)\},
\]
and
\[
J' = \{\tau(\ell+1) < \cdots < \tau(b)\}.
\]
Then we define \( e_I \) to be \( (e_{\sigma(1)}, \ldots, e_{\sigma(k)}) \). Likewise \( e_{I'}, g_I, \) etc. Later we shall use the same notation to denote the order product of the elements, e.g., \( e_I = e_{\sigma(1)} \cdots e_{\sigma(k)} \).

Now
\[
s_a(u_1, \ldots, u_a) = (-1)^\sigma s_a(e_I, g_I)
\]
\[
s_b(w_1, \ldots, w_b) = (-1)^\gamma s_b(\epsilon_I, \gamma_I)
\]
\[
s_{a+b}(u_1', \ldots, u_a', w_1', \ldots, w_b') = (-1)^{\sigma + \gamma} s_{a+b}(g_I, e_{I'}, \gamma_I, \epsilon_{I'}).
Hence, (6.1) equals
\[ s_n(e_{Ic}, g_{Jc})s_n(b, e_{Ic}, e_{Jc}, e_{Jc})s_n(e_{Ic}, e_{Jc}). \]
Since \( I^c \) and \( J \) have even cardinality we may switch \( e_{Ic} \) and \( e_{Jc} \) without changing the sign. Then we may apply Lemma 6.2 to get a positive constant times
\[ e_I g_{Jc} g_I e_J e_I g_{Jc} e_J. \]
Again, since all of the index sets are even, all of the products are central so we may rearrange terms to get
\[ e_I g_{Jc} g_I e_J e_I g_{Jc} e_J. \]
But, \( e_I e_J e_I = (-1)^\varepsilon e_1 \cdots e_{\alpha} g_{Jc} g_{Jc} = (-1)^\varepsilon g_1 \cdots g_{\alpha}, e_J e_{Jc} = (-1)^\tau e_1 \cdots e_{\beta}, \)
and \( e_{Jc} e_{Jc} = (-1)^\tau \gamma_1 \cdots \gamma_b \) and this completes the proof.

**Lemma 6.4.** If the prime product \( A \circ B \circ C \) satisfies \( s_n(x_1, \ldots, x_n)^k \) and if for some \( 1 \leq t \leq k \) \( B \) does not satisfy
\[ s_i(x_1, \ldots, x_i) s_n(x_1, \ldots, x_n)^{t-1} s_n(x_{i+1}, \ldots, x_{n-1}) = 0 \]
for any \( i = 0, \ldots, n-1 \), then \( A \circ C \) satisfies \( s_n(x)^{k-t} \).

**Proof.** Let
\[ X_1 = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & b_1 & 0 \\ 0 & 0 & c_1 \end{pmatrix}, \ldots, X_{n-1} = \begin{pmatrix} a_{n-1} & 0 & 0 \\ 0 & b_{n-1} & 0 \\ 0 & 0 & c_{n-1} \end{pmatrix}, \]
\[ X_n = \begin{pmatrix} a_n & x & 0 \\ 0 & b_n & y \\ 0 & 0 & c_n \end{pmatrix}, \]
be in \( A \circ B \circ C \). Then we may write
\[ s_n(X_1, \ldots, X_n) = \begin{pmatrix} s_n(a_1, \ldots, a_n) & \tilde{x} & 0 \\ 0 & s_n(b_1, \ldots, b_n) & \tilde{y} \\ 0 & 0 & s_n(c_1, \ldots, c_n) \end{pmatrix}. \]
To write \( \tilde{x} \) and \( \tilde{y} \) explicitly, we will use the notation \( s_I(a) \) to denote \( s_n(a_{i_1}, \ldots, a_{i_a}) \) where \( I = \{ i_1 < \cdots < i_a \} \subseteq \{ 1, 2, \ldots, n-1 \} \). Likewise \( s_I(b), s_I(c), s_I(x) \). It will also be helpful to write \( s_n(a) \) for \( s_n(a_{i_1}, \ldots, a_{i_a}) \), etc. Then
\[ \tilde{x} = \sum_{I \subseteq \{ 1, 2, \ldots, n-1 \}} s_I(a) x s_I(b) \quad \text{and} \quad \tilde{y} = \sum_{I \subseteq \{ 1, 2, \ldots, n-1 \}} s_I(b) y s_I(c). \]
where $\overline{T}$ denotes the complement of $I$ in $\{1, \ldots, n-1\}$. Let’s first consider the $(1, 2)$ entry of $s_n(x_1, \ldots, x_n)^k$. It is
\[
\sum_{\alpha + \beta = k-1} s_n(a_1, \ldots, a_n)^\alpha \bar{x} s_n(b_1, \ldots, b_n)^\beta = \sum_{\alpha + \beta = k-1} \sum_{I} s_n(a)^\alpha s_I(a)x s_T(b)s_n(b)^\beta.
\]
This must be zero and, by a Vandermonde argument, the term with $I = \emptyset$ and $\alpha = k-1$ must be zero. Since $s_T(b)s_n(b)^{k-1}$ does not vanish identically $A$ must satisfy $s_n(x)^{k-1}$. By a similar argument, $B$ also satisfies $s_n(x)^{k-1}$.

Now consider the $(1, 3)$ entry of $s_n(x_1, \ldots, x_n)^k$. It is
\[
\sum_{\alpha + \beta + \gamma = k-2} s_n(a_1, \ldots, a_n)^\alpha \bar{x} s_n(b_1, \ldots, b_n)^\beta y s_n(c_1, \ldots, c_n)^\gamma = \sum_{\alpha + \beta + \gamma = k-2} \sum_{I} s_n(a)^\alpha s_I(a)x s_T(b)s_n(b)^\beta s_J(b) y s_T(c)s_n(c)^\gamma.
\]
It follows from a Vandermonde argument that each individual summand must be zero. If we take $\beta = t-1$ and $I = J$ then, by hypothesis, $s_T(b)s_n(b)^{t-1}$ is not zero. Hence, either $s_n(a)^\alpha s_I(a) = 0$ or $s_T(c)s_n(c)^\gamma = 0$ for any $\alpha + \gamma = k-1-t$. A straightforward calculation now shows that $s_n(x)^{k-t}$ is an identity for $A \circ C$.

**Lemma 6.5.** $E$ satisfies the hypothesis of the above lemma with $t = 1$ and $M_{1,1}$ satisfies the hypothesis of the above lemma with $t = 2$.

**Proof.** If $e_1, \ldots, e_{n-1}$ are degree one Grassmann elements then
\[
s_t(e_1, \ldots, e_t) s_{n-t-1}(e_{t+1}, \ldots, e_{n-1}) = t!(n-i-1)! e_1 \cdots e_{n-1}
\]
which is not zero in general. The statement about $M_{1,1}$ follows from Lemma 6.3.

These lemmas have some useful corollaries.

**Corollary 6.6.** If the prime product $A \circ M_a(E) \circ B$ satisfies $s_n(x)^k$, then $A \circ B$ satisfies $s_n(x)^k-d$. In particular, $M_a(E)$ does not satisfy $s_n(x)^d$ for any $n$.

**Proof.** $M_a(E)$ contains $E \circ E \circ \cdots \circ E$ (a factors).

**Corollary 6.7.** If $a \geq b \geq 1$, and if $A \circ M_{a,b} \circ B$ satisfies $s_n(x)^k$, then $A \circ B$ satisfies $s_n(x)^{k-2b}$. In particular, $M_{a,b}$ does not satisfy $s_n(x)^{2b}$ for any $n$.

**Proof.** There is an embedding of $M_{1,1} \circ M_{1,1} \circ \cdots \circ M_{1,1}$ (b factors) into $M_{b,b}$ and hence into $M_{a,b}$. To see this, write $M_{1,1}$ as the Grassmann envelope $G(\text{End}(V))$, where $V$ is a $Z/2Z$-graded vector space where the dimensions of the degree zero and degree one part are each one. So, the
product of $b$ factors of $M_{1,1}$ will be the Grassmann envelope of the product $\text{End}(V_1) \circ \cdots \circ \text{End}(V_6)$. This latter embeds in the endomorphism of $V_1 \oplus \cdots \oplus V_6$, which is a graded vector space in which the degree zero part and the degree one part are each $\ell$ dimensional. So, the Grassmann envelope of the endomorphism ring will be isomorphic to $M_{b,b}$. This completes the proof.

**Lemma 6.8.** The algebra $M_{2,1}$ does not satisfy an identity of the form

$$s_i(x_1, \ldots, x_i)s_4(x_1, \ldots, x_4)^2s_{3-i}(x_{i+1}, \ldots, x_3),$$

where $i = 0, \ldots, 3$.

**Proof.** The hard part of the proof is done by a computer computation. Let

$$x_i = \begin{pmatrix} 0 & 0 & a_i \\ 0 & 0 & b_i \\ c_i & d_i & 0 \end{pmatrix}, \quad i = 1, \ldots, 4,$$

where the matrix entries are anticommuting variables. So the $x_i$ are degree one elements in $M_{2,1}$. To further simplify the computation let $d_1 = c_2 = b_3 = a_4 = 0$. Then a computer computation using Macsyma shows that the $s_4(x)^3$ has $(3, 3)$ entry equal to $960a_1b_1c_2d_3 \neq 0$. Now, in our lemma, if $i = 0$ we have $s_4(x_1, \ldots, x_4)^2s_3(x_1, x_2, x_3)$. This implies each of

$$s_3(x_1, \ldots, x_4)^2s_3(x_1, \ldots, 2, \ldots, x_4),$$

and hence $s_4(x_1, \ldots, x_4)^3$. Next, if $i = 1$ the polynomial is $x_1s_4(x_1, \ldots, x_4)^2s_3(x_2, x_3)$. By replacing $x_1$ by $x_1 + 1$, we can eliminate the factor of $x_1$ on the left. But this also has $s_4(x)^3$ as a consequence: One may right multiply by $s_2(x_1, x_4)$, take various permutations of the variables, and add. Hence, neither the $i = 0$ nor $i = 1$ cases are identities. The arguments in the other two cases are essentially the same.

Our computations lead us to hazard a conjecture.

**Conjecture.** If $x_1, \ldots, x_{2a}$ are generic odd elements in $M_{a,1}$ then $s_{2a}(x_1, \ldots, x_{2a})^{2a} \neq 0$. (It is easy to see that the $(2a+1)$st power will be zero.)

**Lemma 6.9.** If a product $\overline{A} = F_{n_1} \circ \cdots \circ F_{n_k}$ of matrix algebras (over the field $F$) satisfies $s_n(x)^k$, then $\overline{A}$ embeds into $F_{\lceil n/2 \rceil} \circ \cdots \circ F_{\lceil n/2 \rceil}$ ($k$ factors).

**Proof.** The proof will be by induction on $k$. For $k = 1$ it follows from a staircase argument that if $\overline{A}$ satisfies $s_n(x)$, then $n_1 + \cdots + n_i \leq \lceil n/2 \rceil$ and so the lemma holds.

For the general case, note that each $F_{n_i}$ must satisfy $s_n(x)^k$ and so by Amitsur’s theorem, it must satisfy $s_n(x)$. Now, let $\alpha$ be as small as possible
such that $F_{n_1} \circ \cdots \circ F_{n_k}$ does not satisfy $s_n(x)$. (If no such $\alpha$ exists, then we are done by the $k = 1$ case.) Take $A = F_{n_1} \circ \cdots \circ F_{n_{k-1}}$, $B = F_{n_k}$ and $C = F_{n_{k+1}} \circ \cdots \circ F_{n_1}$. We will prove that $B \circ C$ must satisfy $s_n(x)^k$ and this will provide the induction step.

In Eq. 6.2 both $A$ and $B$ satisfy $s_n(x)$ and so we must consider only terms with $\alpha = \beta = 0$:

$$s_f(a)m_1 s_T(b)s_j(b)m_2 s_T(c)s_n(c)^{k-2}.$$  

Since $A \circ B$ does not satisfy $s_n(x)$ there is an $I$ such that $s_f(a)m_1 s_T$ is not zero. It follows that $s_f(b)m_2 s_T(c)s_n(c)^{k-2}$ must be zero and so $B \circ C$ satisfies $s_n(x)^{k-1}$.

We now have all the ingredients we need to prove our main theorem of this section.

**Theorem 6.10.** For all $n \geq 4$, $\exp(s_n(x)^k) = k \lceil n/2 \rceil^2$.

**Proof.** We first prove that if $A$ is a prime product algebra satisfying $s_n(x)^k$ then $\exp(A) \leq k \lceil n/2 \rceil^2$.

The theorem will follow since the algebra $F_{\lceil n/2 \rceil} \circ \cdots \circ F_{\lceil n/2 \rceil}$ has exponential rate of growth $k \lceil n/2 \rceil^2$ and satisfies $s_n(x)^k$.

The proof will be by induction on $k$. The case of $k = 1$ is true by the previous lemma, since algebras which satisfy standard polynomials do not contain $E$.

Let $A = A_1 \circ \cdots \circ A_k$. We now consider cases. If some $A_i = M_{a_i}(E)$, then $A_1 \circ \cdots \circ A_i \circ \cdots \circ A_k$ will satisfy $s_n(x)^{k-a}$ and so, by induction,

$$\exp(A) \leq (k - a) \lceil n/2 \rceil^2 + 2a^2 = k \lceil n/2 \rceil^2 + a(- \lceil n/2 \rceil^2 + 2a).$$

But since $M_{a_i}(E)$ contains $F_a$, $a \leq \lceil n/2 \rceil$. Hence, if $n \geq 4$, $(- \lceil n/2 \rceil^2 + 2a) \leq 0$, and we are done.

Next, if some $A_i = M_{a,b}$, then $A_1 \circ \cdots \circ A_i \circ \cdots \circ A_k$ will satisfy $s_n(x)^{k-2b}$ and so, by induction,

$$\exp(A) \leq (k - 2b) \lceil n/2 \rceil^2 + (a + b)^2 = k \lceil n/2 \rceil^2 + ((a + b)^2 - 2b \lceil n/2 \rceil^2).$$

We need to prove that the second summand is zero. If $n \geq 6$ then, using the fact that $a \leq \lceil n/2 \rceil$ we get

$$(a + b)^2 - 2b \lceil n/2 \rceil^2 \leq (\lceil n/2 \rceil + b)^2 - 2b \lceil n/2 \rceil^2 = b^2 + (2 \lceil n/2 \rceil - \lceil n/2 \rceil^2)b + \lceil n/2 \rceil^2,$$

where $1 \leq b \leq \lceil n/2 \rceil$. Considered as a function of $b$ the minimum value of this function on this interval is at $b = 1$ and it is $2 \lceil n/2 \rceil - \lceil n/2 \rceil^2 + 1$. Since $n \geq 6$ this will be less than or equal to zero and we are done in this case.
Next, if \( n = 5 \) then \( A \) satisfies \( s_5(x)^k \). But since prime product algebras have unit, \( A \) must also satisfy \( s_5(x)^k \) and so we are in the \( n = 4 \) case. In this case \( A_i = M_{n_i} \), equals either \( M_{2_1} \) or \( M_{2_2} \). We consider these cases separately. If \( A_i = M_{2_1} \) then \( A_1 \circ \cdots \circ A_i \circ \cdots \circ A \) satisfies \( s_n(x)^{k-3} \) and so by induction

\[
\exp(A) \leq (k-3)[n/2]^2 + (2+1)^2 = k[n/2]^2 - 3 < k[n/2]^2.
\]

Finally, if \( A_i = M_{2_2} \) then \( A_1 \circ \cdots \circ A_i \circ \cdots \circ A \) satisfies \( s_n(x)^{k-4} \) so

\[
\exp(A) \leq (k-4)[n/2]^2 + (2+2)^2 = k[n/2]^2.
\]

This completes the proof.

7. GROWTH OF POWERS OF POLYNOMIALS

The main result of this section generalizes the results of the previous two sections. Let \( f(x_1, \ldots, x_n) \) be any polynomial which is homogeneous in each variable. We will show that \( \exp(f(x_1, \ldots, x_n)^k) \) is bounded above and below by linear functions in \( k \).

**Lemma 7.1.** Given any polynomial \( f(x_1, \ldots, x_n) \) there exists a finite set of verbally prime algebras \( S \) such that if \( A \) is a verbally prime algebra not p.i. equivalent to an element of \( S \) then \( A \) does not satisfy any power of \( f \).

**Proof.** If \( m \) is large enough then \( f(x_1, \ldots, x_n) \) will not be an identity for \( F_m \). By Amitsur's theorem, no power of \( f(x_1, \ldots, x_n) \) will be an identity for \( F_m \). However, up to p.i. equivalence, there are only finitely many p.i. algebras which don't contain \( F_m \): \( F_a, M_{a,b}, \) and \( E_a \) for \( a < m \).

**Lemma 7.2.** If \( f(x_1, \ldots, x_n)^k \) is an identity for \( A \circ B \) and \( f(x_1, \ldots, x_n) \) is not an identity for \( A \), then \( f(x_1, \ldots, x_n)^{k-1} \) is an identity for \( B \).

**Proof.** Let \( f(x_1, \ldots, x_n) = \sum f_i(x_1, \ldots, x_n)x_i \) and assume without loss of generality that \( f_1(x_1, \ldots, x_n) \) is not an identity for \( A \). As in Lemma 5.1, we let

\[
x_1 = \begin{pmatrix} a_1 & m \\ 0 & b_1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} a_2 & 0 \\ 0 & b_2 \end{pmatrix}, \ldots, \quad x_n = \begin{pmatrix} a_n & 0 \\ 0 & b_n \end{pmatrix}
\]

be in \( A \circ B \). Under this substitution,

\[
f(x_1, \ldots, x_n) = \begin{pmatrix} f(a_1, \ldots, a_n) & \tilde{m} \\ 0 & f(b_1, \ldots, b_n) \end{pmatrix}.
\]

Hence, the \((1, 2)\) entry of \( f(x_1, \ldots, x_n)^k \) will be

\[
\sum_{\alpha + \beta = k-1} f(a_1, \ldots, a_n)\alpha \tilde{m} f(b_1, \ldots, b_n)\beta.
\]
Let $N$ be the degree of $f$. We would like to identify the term with total degree $N(k-1)$ in the $b$'s. By a Vandermonde argument, this term will be an identity for $A \circ B$. Note that $\bar{m}$ is a sum of terms having degree at most $N-1$ in the $b$'s. Hence, the desired identity comes from the part of the sum with $\alpha = 0$, $\beta = k-1$, and the terms in $\bar{m}$ with degree zero in the $b$ (and so degree $N-1$ in the $a$). The latter is $f_1(a_1, \ldots, a_n)\bar{m}$, hence

$$f_1(a_1, \ldots, a_n)m\bar{m}(b_1, \ldots, b_n)^{k-1} = 0$$

for all $a_1, \ldots, a_n \in A$, $m \in M$, and $b_1, \ldots, b_n \in B$. Since the first factor is not zero, $f(x_1, \ldots, x_n)^{k-1}$ is an identity for $B$ by Remark 2.2.

**Theorem 7.3.** If $f(x_1, \ldots, x_n)$ is homogeneous in each variable, then $\exp(f(x_1, \ldots, x_n)^k)$ is bounded above and below by a linear function of $k$.

**Proof.** Let $I$ be the $T$-ideal generated by $f(x_1, \ldots, x_n)$. Then $f(x_1, \ldots, x_n)^k \in I^k$, so $\exp(I^k) \leq \exp(f(x_1, \ldots, x_n)^k)$. But $\exp(I^k) = k \cdot \exp(I)$ by [BR4].

For the upper bound, let $S$ be as in Lemma 7.1. For $t$ large enough, no prime product of more than $t$ (not necessarily distinct) elements of $S$ satisfies $f(x_1, \ldots, x_n)$. An induction argument using Lemma 7.2 now shows that if $A_1 \circ \cdots \circ A_N$ satisfies $f(x_1, \ldots, x_n)^k$ then $N \leq kt$. Let

$$C = \max\{|\exp(A_1 \circ \cdots \circ A_t)\}, A_1, \ldots, A_t \in S\}.$$ 

Then $\exp(f(x_1, \ldots, x_n)^k) \leq Ck$.

**REFERENCES**


EXEMPLARY GROWTH OF p.i. ALGEBRAS


