

ALMOST PI ALGEBRAS ARE PI

MICHAEL LARSEN AND ANER SHALEV

ABSTRACT. We define the notion of an almost polynomial identity of an associative algebra R , and show that its existence implies the existence of an actual polynomial identity of R . A similar result is also obtained for Lie algebras and Jordan algebras. We also prove related quantitative results for simple and semisimple algebras.

By a well known theorem of Peter Neumann [N], for all $\epsilon > 0$ there exists $N > 0$ such that if G is a finite group with at least $\epsilon|G|^2$ pairs $(x, y) \in G^2$ satisfying $[x, y] = 1$, then $[x^N, y^N]^N = 1$ for all $x, y \in G$. We can express this by saying that if $[x, y]$ is an ϵ -probabilistic identity for G , then $[x^N, y^N]^N$ is an identity for G . See also Mann [M] for a similar result on finite groups in which x^2 is an ϵ -probabilistic identity. It is an open question whether every probabilistic identity in finite and residually finite groups implies an actual identity; see [LS2] and [Sh] for further discussion of this question.

In this paper, we consider the analogous problem for associative algebras, as well as Lie algebras and Jordan algebras. Here, the possible identities are polynomials in a non-commuting set of variables (or Lie polynomials or Jordan polynomials) rather than elements of a free group. We introduce the notion of an *almost identity* of an algebra as an analogue of a probabilistic identity and show that algebras with an almost identity satisfy an actual identity. While we focus here on infinite-dimensional algebras we also prove related quantitative results for finite dimensional simple algebras.

Throughout this paper let k denote an algebraically closed field of arbitrary characteristic. Let R be a (possibly infinite-dimensional) associative algebra over k . Let n be a positive integer and $V = R^n$, regarded as k -vector space. For each non-commutative polynomial $Q \in k\langle x_1, \dots, x_n \rangle$, evaluation of Q defines a map $e_Q: R^n \rightarrow R$. For each linear functional $\alpha \in R^*$, $\alpha \circ e_Q$ defines an element of the ring A of polynomial functions on V . Let $R_Q := e_Q^{-1}(0)$ denote the set of $\mathbf{r} = (r_1, \dots, r_n) \in V$ such that $e_Q(\mathbf{r}) = 0$.

We now introduce more notation and terminology. In particular we define the notions of codimension and almost identity, which play a key role in this paper. Let A be the ring of k -valued polynomial functions on a k -vector space V . If $\mathfrak{I}$ is an ideal in A , we denote by $V(\mathfrak{I}) \subset V$ the solution set of

2010 *Mathematics Subject Classification.* Primary 16R99.

ML was partially supported by NSF grant DMS-1702152. AS was partially supported by ISF grant 686/17 and the Vinik Chair of mathematics which he holds. Both authors were partially supported by BSF grant 2016072. AS is grateful to Efim Zelmanov for inspiring discussions.

the system of equations on V given by the elements of $\mathfrak{I}$. An *algebraic set* in V is a set of this form. Note that if $\dim V = \infty$, this may be a proper subset of the set of k -points of $\operatorname{Spec} A/\mathfrak{I}$. We say the algebraic set $V(\mathfrak{I})$ has *finite codimension* if it contains a translate of a vector subspace of V of finite codimension. In particular, this is the case whenever $\mathfrak{I}$ is finitely generated or (more generally) when $\mathfrak{I}$ is contained in a finitely generated (non-unit) ideal.

We say $V(\mathfrak{I})$ has codimension $\leq c$ if there exists a direct sum decomposition $V = V_1 \oplus V_2$, where V_2 is finite dimensional, and V contains $V_1 \times X_2$ for $X_2 \subset V_2$ an algebraic set of codimension c . This is, of course, the case whenever $V(\mathfrak{I})$ contains a translate of a subspace of codimension c . The codimension of an algebraic set of finite codimension is the smallest integer c for which the set has codimension $\leq c$.

Example 1. Let V denote the space of infinite sequences, A the algebra of polynomial k -valued functions on V , $x_i \in A$ the function sending a sequence to the value of its i th term, and

$$\mathfrak{I} = (x_1(x_1 - 1), x_1x_2(x_2 - 1), \dots, x_1 \cdots x_n(x_n - 1), \dots).$$

Then

$$V(\mathfrak{I}) = \{(1, 1, 1, \dots)\} \cup \bigcup_{i=0}^{\infty} \underbrace{(1, \dots, 1)}_i, 0, *, *, \dots).$$

In particular, $V(\mathfrak{I})$ has codimension 1, though it contains components of all positive integer codimensions.

We say that $Q \neq 0$ is a *c-almost identity* of R if $e_Q^{-1}(0)$ has codimension $\leq c$ in V . We say that R *satisfies an almost identity* if it satisfies a c -almost identity for some c ; of course, all finite dimensional algebras have this property. Our main result is the following:

Theorem 2. *For all positive integers c , d and n there exist an integer m and a non-zero non-commutative polynomial $P \in k\langle y_1, \dots, y_m \rangle$ with the following property. If $Q \in k\langle x_1, \dots, x_n \rangle$ is a non-commutative polynomial of degree d in n variables, and R is any associative k -algebra such that Q is a c -almost identity of R , then $P(r_1, \dots, r_m) = 0$ for all $r_1, \dots, r_m \in R$.*

In particular, every algebra satisfying an almost identity is PI.

In the case of matrix algebras $R = M_s(k)$, if P is an identity of $M_s(k)$ then $s \leq \deg P/2$, and this is tight by the Amitsur-Levitski theorem [AL]. Thus, Theorem 2 implies

Corollary 3. *If $Q \in k\langle x_1, \dots, x_n \rangle$ is non-zero, then the codimension of*

$$M_{s,Q} := \{\mathbf{r} \in M_s^n \mid Q(\mathbf{r}) = 0\}$$

in M_s^n (regarded as an affine space of dimension ns^2) grows without bound as $s \rightarrow \infty$.

In fact, something much stronger is true. The second theorem of this paper is the following:

Theorem 4. *For each positive integer d there exist real numbers $a > 0$ and b such that if $Q \in k\langle x_1, \dots, x_n \rangle$ is a non-commutative polynomial of degree $d \geq 0$, then the codimension of $M_{s,Q}$ in M_s^n is at least $as^2 + b$ for all positive integers s .*

Let Q, d, n be as above and let R be a finite dimensional k -algebra. We define the (normalized) Hausdorff dimension of the algebraic subset $R_Q \subseteq R^n$ by $\dim R_Q / \dim R^n$. Theorem 4 shows that, when R is simple of sufficiently large dimension given d , the Hausdorff dimension of R_Q is at most $1 - \epsilon$ for some $\epsilon > 0$ depending only on d .

Before proving Theorems 2 and 4, we derive some consequences and related results.

Recall that the Jacobson radical $J(R)$ of an associative ring R is the intersection of all primitive ideals of R . We say that a ring R is J -semisimple if $J(R) = 0$.

Combining Theorems 2 and 4 with other tools we obtain the following:

Corollary 5. *For each $d > 0$ there exist real numbers $f = f(d), g = g(d)$ such that the following holds. Let R be an associative k -algebra having a c -almost identity of degree d . Suppose R is J -semisimple. Then*

(i) *R is a subdirect sum of central simple algebras $M_{s_i}(D_i)$ ($i \in I$), where D_i is a division algebra of degree d_i^2 over $k_i := Z(D_i)$ and $s_i d_i \leq fc^{1/2} + g$ for all $i \in I$.*

(ii) *R can be embedded in a matrix ring $M_s(C)$ for some commutative k -algebra C , where $s \leq fc^{1/2} + g$.*

Proof. To deduce this, note that R is PI by Theorem 2. By Theorem 1 of Amitsur [A], R is a subdirect sum of central simple algebras C_i ($i \in I$) over suitable fields $k_i \supseteq k$ (where $k_i = k$ or $[k_i : k] = \infty$ since k is algebraically closed). We have $C_i = M_{s_i}(D_i)$ where D_i is a division algebra over k_i of degree d_i^2 . Let Q be a c -almost identity of R of degree d . Each C_i is a quotient of R , therefore Q is a c -almost identity of $M_{s_i}(D_i)$ and hence of $M_{s_i d_i}(\bar{k}_i)$ for all $i \in I$. Let a, b be as in Theorem 4 above. Then this theorem yields $a(s_i d_i)^2 + b \leq c$ for all i , so $s_i d_i \leq \sqrt{(c-b)/a}$ for all $i \in I$. This implies conclusion (i), which in turn implies conclusion (ii) (with C a direct sum of $\bar{k}_i$). $\square$

Another consequence of Theorem 4 deals with simple Lie algebras.

Corollary 6. *For each $d > 0$ there exist real numbers $a > 0$ and b such that if Q is a non-zero Lie polynomial of degree d in n variables, L is a finite dimensional simple Lie algebra over k of characteristic zero, then the codimension of L_Q in L^n is at least $a \cdot \dim L + b$.*

Proof. To deduce this, we may ignore the exceptional algebras (whose dimension is bounded) and focus on the classical ones $L = A_r, B_r, C_r, D_r$.

Any such algebra L contains the full matrix algebra $M_r(k)$ and satisfies $\dim L \leq 2r^2 + r$. The Lie polynomial Q can be viewed as an associative polynomial of degree d . Its codimension in $M_r(k)^n$ is a lower bound on the codimension of Q in L^n . Applying Theorem 4 now completes the proof. $\square$

Theorem 4 and Corollary 6 have several applications. First, they immediately imply that *if R is an associative (resp. Lie) k -algebra defined by n generators and by relators whose minimal degree is d , and $S = M_s(k)$ (resp. a classical simple Lie k -algebra) of sufficiently large dimension (given d), then the representation variety $\text{Hom}(R, S)$ has dimension at most $(n - \epsilon) \dim S$, where $\epsilon > 0$ depends only on d .*

Secondly, they imply the following.

Proposition 7. *For every $d \in \mathbb{N}$ there are positive real numbers $\epsilon = \epsilon(d), N = N(d)$ depending only on d such that the following holds. Let F be a finite field, $M = M_s(F)$ a matrix algebra and L a classical simple Lie algebra over F .*

(i) *If Q is a non-zero associative polynomial of degree d with n variables over F and $\dim M \geq N$ then*

$$|M_Q| \leq c|M|^{n-\epsilon},$$

where c is a constant depending on Q . Moreover, all fibers of the evaluation map $e_Q : M^n \rightarrow M$ have size at most $c|M|^{n-\epsilon}$.

(ii) *If Q is a non-zero Lie polynomial of degree d with n variables over F and $\dim L \geq N$ then*

$$|L_Q| \leq c|L|^{n-\epsilon},$$

where c is a constant depending on Q . Moreover, all fibers of the evaluation map $e_Q : L^n \rightarrow L$ have size at most $c|L|^{n-\epsilon}$.

Proof. Extending scalars to $k = \overline{F}$ does not change the dimension of fibers of e_Q , so all such fibers are bounded above by $(n - \epsilon) \dim M_s$ for part (i) and by $(n - \epsilon) \dim L$ for part (ii). The dimensions of the polynomial equations defining any fiber are bounded above, depending only on the polynomial Q , so the estimate of fiber cardinality follows from the Lefschetz trace formula and the upper bounds on Betti numbers for affine varieties defined by polynomial equations of bounded degree [K]. $\square$

This result may be viewed as a ring-theoretic analogue of [LS1, Theorem 1.2].

In particular it follows that if the finite simple algebras M, L above satisfy a probabilistic identity Q (namely, the probability that e_Q vanishes on a random n -tuple of elements of M or L is at least some fixed $\delta > 0$), then $\dim M, \dim L$ are bounded above (in terms of Q and δ).

Next we turn to almost nil algebras and algebras with an almost identity of the form x^d , as well as their Lie analogues, namely almost Engel Lie algebras. Recall that Zelmanov's work on Engel Lie algebras and related

objects has remarkable group theoretic applications, such as the solution to the Restricted Burnside Problem [Z1, Z2], as well as stronger results [Z3].

We start with finite dimensional algebras. We denote by $Rad(L)$ the solvable radical of a finite dimensional Lie algebra L .

Proposition 8. (i) *Let R be a finite dimensional associative k -algebra, and let N be the subvariety of nilpotent elements of R . If $\text{codim } N = c$ then $\dim R/J(R) \leq c^2$, so R has a nilpotent ideal of codimension at most c^2 .*

(ii) *Let L be a finite dimensional Lie k -algebra, where k has characteristic zero, and let N be the subvariety of ad-nilpotent elements of L . If $\text{codim } N = c$ then $\dim L/Rad(L) \leq 4c^2 - c$. Moreover, L has a nilpotent subalgebra K (consisting of ad-nilpotent elements) such that $\dim L/K \leq 4c^2$.*

Proof. To prove part (i), write $R/J(R) = \prod_{i=1}^m M_{s_i}(k)$ and let $\pi : R \rightarrow R/J(R)$ be the canonical epimorphism. For each $i = 1, \dots, m$ let N_i denote the set of nilpotent matrices in $M_{s_i}(k)$. Since the elements of $J(R)$ are nilpotent we have $N = \pi^{-1}(N_1 \times \dots \times N_m)$. It is well known (see for instance [H2, §1.3]) that each N_i is irreducible of dimension $s_i^2 - s_i$. Hence

$$\sum_{i=1}^m s_i = \text{codim } N = c.$$

This yields

$$\dim R/J(R) = \sum_{i=1}^m s_i^2 \leq c^2.$$

The proof of the first assertion in part (ii) is similar, using the fact that the codimension of the variety of ad-nilpotent elements of a finite dimensional simple Lie algebra of rank r is r . To prove the second assertion, set $R = Rad(L)$, $Z = Z(L)$. Then $R/Z \leq L/Z \cong ad(L) \leq gl(L)$. Applying Lie's Theorem (see [H1, 4.1]) to the solvable Lie algebra R/Z we conclude that there is a basis for L with respect to which R/Z is represented by upper-triangular matrices. Let K/Z denote the nilpotent elements of R/Z in its action on L ; these are the elements represented by upper triangular matrices with zero diagonal. Hence K/Z is a Lie subalgebra of R/Z , and K is a nilpotent Lie subalgebra of L (consisting of ad-nilpotent elements). By our assumption on N the codimension of K in R is at most c . Therefore

$$\dim L/K = \dim L/R + \dim R/K \leq 4c^2.$$

□

The example of $R = M_c(k)$ shows that the bound in Proposition 8 (i) is sharp; the above argument shows that it is attained if and only if $R/J(R) = M_c(k)$. Similarly, the bound in part (ii) of the above result is dictated by E_8 .

Proposition 8 (i) can be extended as follows. For each monic polynomial $P(x) \in k[x]$ of degree s , let $M = M_s(k)$ and let M_P be the variety of matrices in M whose characteristic polynomial is P . Then $\dim M_P = s^2 - s$. Using

arguments as above, it follows that if P is a c -almost identity of a finite dimensional associative k -algebra R , then $\dim R/J(R) \leq c^2$.

The well known Nagata-Higman Theorem states that, if x^d is an identity of an associative (non-unital) k -algebra R , where the characteristic of k is zero or greater than d , then R is nilpotent. See [DF, Chapter 6] for this and for explicit bounds on the degree of nilpotency of R in terms of d . Our next result deals with rings in which x^d is an almost identity.

Proposition 9. *Let R be an associative k -algebra, and suppose x^d is a c -almost identity of R . Then*

- (i) *$R/J(R)$ is finite dimensional; in fact $\dim R/J(R) \leq c^2$.*
- (ii) *If R is a finitely generated k -algebra then R is virtually nilpotent.*

Here an algebra R is said to be virtually nilpotent if it has a (two-sided) nilpotent ideal of finite codimension.

Proof. To prove Proposition 9, use Corollary 5 and its notation to deduce that $R/J(R)$ is a subdirect sum of central simple algebras $C_i = M_{s_i}(D_i)$ over k_i ($i \in I$). We claim that $D_i = k$ for all i . Otherwise $[D_i : k] = \infty$ for some i , so the set N of nilpotent elements of C_i has infinite codimension. This contradicts the fact that x^d is a c -almost identity of C_i (being a quotient of R).

Since C_i are all finite-dimensional over k , $R/J(R)$ is residually finite dimensional, so to prove part (i) it suffices to show that all its finite dimensional quotients have dimension at most c^2 . Let S be such a quotient. Then $J(S) = 0$ and the codimension of the solutions of $x^d = 0$ in S is at most c . In particular, the codimension of the nilpotent elements of S is at most c . Applying Proposition 8 (i) to S we obtain $\dim S = \dim S/J(S) \leq c^2$, proving part (i).

To prove part (ii) we apply the Kemer-Braun Theorem on the nilpotency of $J(R)$ for finitely generated PI algebras R [B2]. Since R is PI by Theorem 2 above, part (ii) now follows from part (i). $\square$

We now turn to the proofs of the main results, namely Theorems 2 and 4.

Lemma 10. *If $\mathfrak{J} \subset A$ is a non-zero ideal and $\pi_1, \pi_2: V \rightarrow W$ k -linear maps onto a finite dimensional vector space, then $W = \pi_1(V \setminus V(\mathfrak{J})) + \pi_2(V \setminus V(\mathfrak{J}))$.*

Proof. As $\mathfrak{J}$ is non-zero, it contains (f) for some non-zero polynomial $f \in A$, so it suffices to prove the theorem when $\mathfrak{J} = (f)$. We choose a surjective linear map $\phi: V \rightarrow W_0$, for some finite-dimensional W_0 , such that f factors through ϕ . Replacing V by its quotient by $\ker \pi_1 + \ker \pi_2 + \ker \phi$, π_i and f are still well defined, so we may assume V is finite dimensional. For finite dimensional vector spaces, the image of a non-empty open subset under any k -linear surjective map contains a non-empty open subset. By the irreducibility of affine spaces, non-empty open subsets U_i of a finite dimensional vector space W satisfy $U_1 + U_2 = W$. $\square$

We fix positive integers m and d . Let M^d denote the set of ordered d -tuples of elements (or, equivalently, d -term sequences) in $\{1, 2, \dots, m\}$. We identify $I = (i_1, \dots, i_d) \in M^d$ with $x_{i_1} \cdots x_{i_d} \in k\langle x_1, \dots, x_m \rangle$. For $\mathbf{r} = (r_1, \dots, r_m) \in R^m$, we define

$$\mathbf{r}_I := r_{i_1} \cdots r_{i_d}.$$

Let

$$M^{\leq d} = \bigcup_{i=0}^d M^i,$$

and let $k\langle x_1, \dots, x_m \rangle^{\leq d}$ denote its linear span.

Proposition 11. *Given m and d as above, there exists a non-zero homogeneous non-commutative polynomial $P \in k\langle x_1, \dots, x_m \rangle$ such that if R is an associative k -algebra and $\mathbf{r} \in R^m$ satisfies $P(\mathbf{r}) \neq 0$, then $\{\mathbf{r}_I \mid I \in M^{\leq d}\}$ is linearly independent.*

To prove this we need some preparations. For $I \in M^d$ and any positive integer N , we let $S_{N,I}$ denote the set of distinct N -term sequences in $\{1, \dots, m\}$ for which I is not a consecutive subsequence. We identify $S_{N,I}$ with a subset of the monomials in $k\langle x_1, \dots, x_m \rangle$.

Lemma 12. *With notation as above,*

$$\limsup_N |S_{N,I}|^{1/N} < m.$$

Proof. Write $N = qd + r$ where $q, r \in \mathbb{N}$ and $0 \leq r < d$. Avoiding the subsequence I in locations $td + 1, \dots, (t+1)d \leq N$ ($t \in \mathbb{N}$) yields $|S_{N,I}| \leq (m^d - 1)^q m^r \leq (m^d - 1)^{N/d} m^d$. Therefore

$$|S_{N,I}|^{1/N} \leq (m^d - 1)^{1/d} m^{d/N},$$

so $\limsup_N |S_{N,I}|^{1/N} \leq (m^d - 1)^{1/d} < m$. □

Given an m -tuple of functions $\mathbf{F} := (F_1(x), \dots, F_m(x))$, $F_i: \mathbb{Z}^{>0} \rightarrow \mathbb{Z}^{>0}$, and a sequence $\mathbf{a} := a_1, a_2, \dots, a_N \in \{1, \dots, m\}$, we define the *score* of $\mathbf{a}$ with respect to $\mathbf{F}$ to be

$$\sigma(\mathbf{a}) := \sum_{j=1}^N F_{a_j}(j).$$

Given d and m , we construct $\mathbf{F}$ as follows. We fix a prime $p > 2^{dm}$. For $1 \leq i \leq m$ and $j > 0$, we let r denote the remainder when $jm + i$ is divided by dm and define $F_i(j)$ to be the unique integer congruent to $2^r \pmod{p}$ for which

$$jp \leq F_i(j) < (j+1)p.$$

Thus, for any sequence $\mathbf{a}$ of length N ,

$$(1) \quad \frac{N(N+1)}{2}p \leq \sigma(\mathbf{a}) < \frac{(N+1)(N+2) - 2}{2}p.$$

This implies that if the sequence $\mathbf{a}$ is longer than the sequence $\mathbf{b}$, then $\sigma(\mathbf{a}) > \sigma(\mathbf{b})$.

Lemma 13. *For all integers t and N with $1 \leq t \leq N - d + 1$, if $\mathbf{a}$ and $\mathbf{b}$ are distinct N -term sequences as above, and $a_j = b_j$ except if $t \leq j < t + d$, then $\sigma(\mathbf{a}) \neq \sigma(\mathbf{b})$.*

Proof. It suffices to prove that

$$\sum_{j=t}^{t+d-1} F_{a_j}(j) \neq \sum_{j=t}^{t+d-1} F_{b_j}(j).$$

We show that the two sides are not congruent to one another (mod p). Each is congruent (mod p) to a sum of d elements of the set $\{1, 2, 4, 8, \dots, 2^{dm-1}\}$. As $p > 2^{dm}$, two such sums are congruent (mod p) if and only if the sums of powers of 2 are the same, i.e., if and only if the set of remainders obtained when $jm + a_j$ is divided by dm is the same as the set of remainders obtained when $jm + b_j$ is divided by dm as j ranges over $t, \dots, t + d - 1$. As $1 \leq a_j, b_j \leq m$, this is equivalent to the condition that $a_j = b_j$ for all j in this range, contrary to hypothesis. $\square$

We now prove Proposition 11.

Proof. By induction on d , we may assume that $\mathbf{r}$ satisfies some non-commutative polynomial equation of degree exactly d , i.e., that for some $I \in M^d$, $\mathbf{r}_I$ can be expressed

$$(2) \quad \mathbf{r}_I = \sum_{J \in M^{\leq d} \setminus \{I\}} c_J \mathbf{r}_J.$$

By Lemma 13, without loss of generality, we may assume that $c_J \neq 0$ implies $\sigma(J) < \sigma(I)$.

We introduce formal (commuting) variables z_J indexed by $J \in M^{\leq d} \setminus \{I\}$. Let B denote the (commutative) polynomial algebra over k generated by the variables z_J and consider the free associative B -algebra $B\langle x_1, \dots, x_m \rangle$. Let $\mathfrak{J}$ denote the two-sided ideal in this algebra generated by

$$x_I - \sum_{\{J | c_J \neq 0\}} z_J x_J.$$

Let $B\langle x_1, \dots, x_m \rangle_{\leq N}$ denote the k -subspace of $B\langle x_1, \dots, x_m \rangle$ consisting of monomials in $S_{K,I}$ for $K \leq N$ multiplied by polynomials in the z_J of degree at most the maximal score of a sequence of length $\leq N$.

We claim that every element α of $k\langle x_1, \dots, x_m \rangle_{\leq N}$ is congruent (mod $\mathfrak{J}$) to an element β of $B\langle x_1, \dots, x_m \rangle_{\leq N}$. We iteratively construct a sequence of elements of $B\langle x_1, \dots, x_m \rangle$ which lie in $\alpha + \mathfrak{J}$. Each monomial $x_{j_1} \cdots x_{j_K}$ which is not in $S_{K,I}$ can be replaced by a linear combination of monomials associated to sequences of lower score than $j_1, \dots, j_K$, with coefficients of the form z_J . The number of steps before this process terminates is at most the maximum score of a monomial of degree N , which, by (1), is less than

$p(N^2 + 3N)/2$. The resulting element, β , lies in $B\langle x_1, \dots, x_m \rangle_{\leq N}$. The space of polynomials in $|M^{\leq d}| - 1$ variables of degree less than $p(N^2 + 3N)/2$ has dimension at most $CN^{2(|M^{\leq d}| - 2)}$ for some C not depending on N .

For any $\epsilon > 0$, we have

$$(3) \quad \dim k\langle x_1, \dots, x_m \rangle^{\leq N} > m^N > CN^{2(|M^{\leq d}| - 2)}(m - \epsilon)^N.$$

By Lemma 12, when N is sufficiently large,

$$\dim k\langle x_1, \dots, x_m \rangle^{\leq N} > \dim B\langle x_1, \dots, x_m \rangle_{\leq N},$$

and it follows that there exists a non-zero α which is equivalent to $\beta = 0 \pmod{\mathfrak{I}}$.

Substituting c_J for each z_J and r_i for each x_i in α , by (2), we get 0. Defining P_I to be α and defining P to be the product of P_I over all $I \in M^d$, we obtain a polynomial which vanishes on any $\mathbf{r}$ for which $Q(\mathbf{r}) = 0$ for any non-commutative polynomial Q of degree d . $\square$

Lemma 14. *Let $r_1, \dots, r_n$ denote elements of an associative algebra R such that the monomials of degree $\leq d$ in the r_i are linearly independent. If $s_1, \dots, s_m$ are linearly independent in $\text{Span}_k(r_1, \dots, r_n)$, then the monomials of degree $\leq d$ in the s_j are linearly independent.*

Proof. Let V denote the span of the r_i and W the span of the s_j . The linear independence of the degree d monomials in the r_i (resp. s_j) is equivalent to the injectivity of the product map $\bigoplus_{i=0}^d V^{\otimes i} \rightarrow R$ (resp. $\bigoplus_{i=0}^d W^{\otimes i} \rightarrow R$), so the lemma follows from the fact that the latter map factors through the former and the map

$$\bigoplus_{i=0}^d W^{\otimes i} \rightarrow \bigoplus_{i=0}^d V^{\otimes i}$$

is injective. $\square$

Lemma 15. *If $n < m$ are positive integers, the algebraic set $N_{n,m}$ consisting of $n \times m$ matrices of rank strictly less than n has codimension at least $1 + m - n$ in $M_{n \times m}(k) = k^{mn}$.*

Proof. If $Z_{n,m}$ consists of ordered pairs (X, v) consisting of an $n \times m$ matrix and a column vector of size n such that $Xv = 0$, then $Z_{n,m}$ projects onto the n -dimensional space of column vectors v , and the dimension of every fiber except the 0-fiber is $mn - m$, so $\dim Z_{n,m} = mn + n - m$. Projecting onto the first factor, the non-empty fibers have dimension at least 1, so the dimension of the image is at most $mn + n - m - 1$. Its codimension is therefore at least $1 + m - n$. $\square$

In fact, the codimension is exactly $1 + m - n$; this follows immediately from [E, Exercise 10.10].

We can now prove Theorem 2.

Proof. Assume that Q is a c -almost identity of R of degree d in n variables. We fix a k -linear direct sum decomposition $R^n = V = V_1 \oplus V_2$ where V_2 is finite dimensional and contains an algebraic set X_2 of codimension c such that Q vanishes on $V_1 \times X_2$. We fix $m := n + \max(n, c)$. By Proposition 11, there exists a non-zero element $P \in k\langle y_1, \dots, y_m \rangle$ such that $P(r_1, \dots, r_m) \neq 0$ implies that the monomials in r_i of degree $\leq d$ are linearly independent.

We assume that P is not identically zero on R . Applying Lemma 10 to $\mathfrak{J} = (P)$ with $\pi_1(r_1, \dots, r_m)$ (resp. $\pi_2(r_1, \dots, r_m)$) equal to the projection of $(r_1, r_3, \dots, r_{2n-1})$ (resp. $(r_2, r_4, \dots, r_{2n})$) to V_2 , we may choose $r_1, \dots, r_m$ such that $P(r_1, \dots, r_m) \neq 0$ and $(r_1 + r_2, r_3 + r_4, \dots, r_{2n-1} + r_{2n}) \in V_1 \times X_2$.

Right-multiplication by the column vector $(r_1 r_2 \cdots r_m)^t$ embeds $M_{n \times m}(k)$ linearly in V , and we have already seen that the image contains at least one point of $V_1 \times X_2$. By Lemma 14, every matrix of rank n in $M_{n \times m}(k)$ maps to an element of R^n which does not satisfy any non-commutative polynomial equation of degree $\leq d$ and in particular does not lie in $V_1 \times X_2$. By Lemma 15, the intersection of $M_{n \times m}(k)$ with $V_1 \times X_2$ has codimension $> c$. However, in a finite dimensional vector space, a non-empty intersection of a subvariety of codimension c with any irreducible subvariety is of codimension $\leq c$ in the latter.

Thus P is a polynomial identity of R . $\square$

We remark that the analogue of Theorem 2 holds for non-unital associative algebras as well. The proof is the same except that polynomials cannot have a constant term.

There is also an analogue for Lie algebras. Namely, given positive integers c, d, n , there exists a Lie polynomial P such that for every Lie polynomial Q of degree d in n variables and every Lie algebra L , either P is an identity on L , or the variety defined by Q on L^n has codimension $> c$. The proof is the same except that $k\langle x_1, \dots, x_n \rangle^{\leq N}$ must be replaced by the subspace of Lie polynomials in $x_1, \dots, x_n$ of degree $\leq N$, and the left hand side of (3) must be replaced by the dimension of this subspace. By [B, II, §3, Théorème 2], it is

$$\sum_{M=1}^N \frac{1}{M} \sum_{d|M} \mu(d) n^{M/d} \geq \frac{1}{N} \sum_{d|N} \mu(d) n^{N/d} > \frac{n^N - \sum_{i \leq N/2} n^i}{N} > \frac{n^N - 2n^{N/2}}{N}$$

It follows that for fixed n and N sufficiently large,

$$\frac{1}{N} \sum_{d|N} \mu(d) n^{N/d} > \binom{n+N-1}{N} (n-1)^N.$$

The analogue of Theorem 2 holds also for Jordan algebras. It suffices to show that for some $\epsilon < 1$ the dimension of the space of Jordan polynomials in $x_1, \dots, x_n$ of degree $\leq N$ is greater than $(n - \epsilon)^N$ for some N . By [R, Theorem 9], if $c_{n,i}$ is the dimension of the part of the free Jordan algebra

on n generators homogeneous of degree i , then

$$\sum_{i=0}^{\infty} c_{n,i} t^i = \exp \sum_{i=1}^{\infty} \frac{n^{\theta(i)} t^i}{i},$$

where $\theta(i)$ denotes the largest odd divisor of i . As $\sum_{i=1}^{\infty} \frac{n^{\theta(i)} t^i}{i}$ grows without bound as t approaches $1/n$ from below, it follows that the radius of convergence of $\sum_{i=0}^{\infty} c_{n,i} t^i$ is at most $1/n$, from which we deduce that

$$\limsup \frac{\log c_{n,i}}{i} \geq n,$$

which implies the needed analogue of (3).

The proof of Theorem 4 follows the method used to bound the size of fibers of word maps for groups of finite simple groups of Lie type in [LS1], except that since we are working over an algebraically closed field instead of a finite field, we use dimension theory as a substitute for the counting arguments in that paper.

Proof. Let N denote the cardinality of $M^{\leq d}$. We order $M^{\leq d}$ by increasing degree and within each degree, lexicographically, denoting its elements, in increasing order, $I_1, I_2, \dots, I_N$.

For $\mathbf{r} = (r_1, \dots, r_n) \in M_s(k)^n$, and $\mathbf{v} = (v_1, \dots, v_t) \in (k^s)^t$, we consider the sequence

$$(4) \quad \mathbf{r}_{I_1} v_1, \dots, \mathbf{r}_{I_N} v_1, \mathbf{r}_{I_1} v_2, \dots, \mathbf{r}_{I_N} v_2, \dots, \mathbf{r}_{I_N} v_t.$$

If $e_Q(\mathbf{r}) = 0$ for some non-zero $Q \in k\langle x_1, \dots, x_m \rangle^{\leq d}$, then $e_Q(\mathbf{r})v_i = 0$ for $1 \leq i \leq t$, so for each i in this range, there exists $j_i \leq N$ such that $r_{I_{j_i}} v_i$ is a linear combination of previous terms in the sequence (4). Let $\Sigma_{n,d,t}$ denote the algebraic set consisting of pairs $(\mathbf{r}, \mathbf{v}) \in M_s(k)^n \times (k^s)^t$ such that this condition holds. We claim

$$(5) \quad \dim \Sigma_{n,d,t} \leq ns^2 + Nt^2.$$

Since $M_{s,Q} \times (k^s)^t \subset \Sigma_{n,d,t}$, this implies

$$\dim M_{s,Q} \leq ns^2 + Nt^2 - st.$$

Setting $t = \lfloor s/2N \rfloor$, we obtain

$$\dim M_{s,Q} \leq (n - 1/4N)s^2 + s,$$

which implies the theorem.

To prove (5), we consider a fixed $(\mathbf{r}, \mathbf{v}) \in \Sigma_{n,d,t}$ and define j_i , for $1 \leq i \leq t$, to have the smallest value for which $r_{I_{j_i}} v_i$ is a linear combination of previous terms in the sequence (4). It suffices to prove that for each i , if we fix all the terms preceding $r_{I_{j_i}} v_i$, the condition on $(\mathbf{r}, \mathbf{v})$ that $r_{I_{j_i}} v_i$ lies in the span of these fixed vectors is a linear condition of codimension greater than $s - Nt$, and moreover, the conditions imposed by successive values of i are independent of one another. If $j_i = 1$, since I_1 is the tuple of length 0, the linear dependence condition requires v_i to lie in the span of the $(i-1)N$

preceding terms of the sequence (4). This is a condition of codimension greater than $s - Nt$, and as it is the only condition we consider which applies to v_i , it is linearly independent of our other conditions. If $j_i > 1$, then $x_{I_{j_i}}$ can be written $x_{p_i}x_{J_i}$ for some $1 \leq p_i \leq n$ and some $J_i < I_{j_i}$. The linear dependence condition can now be viewed as a condition on r_{p_i} , and it asserts that $r_{p_i}(\mathbf{r}_{J_i}v_i)$ lies in a specified vector space of dimension less than Nt . Note that $\mathbf{r}_{J_i}v_i$ belongs to the part of the sequence (4) that we are assuming fixed, so this is a linear condition on r_{p_i} of codimension $> s - Nt$. Note also that by definition of j_i , $\mathbf{r}_{J_i}v_i$ is linearly independent from all previous terms in the sequence (4), so the condition on r_{p_i} is linearly independent of any previous conditions on r_{p_i} .

This concludes the proof of (5) and therefore of Theorem 4. □

REFERENCES

- [A] Amitsur, Shimshon: An embedding of PI-rings, *Proc. Amer. Math. Soc.* **3** (1952), 3–9.
- [AL] Amitsur, Shimshon; Levitzki, Jacob: Minimal identities for algebras, *Proc. Amer. Math. Soc.* **1** (1950), 449–463.
- [B] Bourbaki, N.: Éléments de mathématique. Fasc. XXXVII. Groupes et algèbres de Lie. Chapitre II: Algèbres de Lie libres. Chapitre III: Groupes de Lie. Actualités Scientifiques et Industrielles, No. 1349. Hermann, Paris, 1972.
- [B2] Braun, Amiram: The nilpotency of the radical in a finitely generated PI ring, *J. Algebra* **89** (1984), 375–396.
- [CRS] Cvetković, Dragoš; Rowlinson, Peter; Simić, Slobodan: An introduction to the theory of graph spectra. London Mathematical Society Student Texts, 75. Cambridge University Press, Cambridge, 2010.
- [DF] Drensky, Vesselin; Formanek, Edward: Polynomial identity rings. Advanced Courses in Mathematics, CRM Barcelona, Birkhäuser Verlag, Basel, 2004.
- [E] Eisenbud, David: Commutative algebra. With a view toward algebraic geometry. Graduate Texts in Mathematics, 150. Springer-Verlag, New York, 1995.
- [H1] Humphreys, James E.: Introduction to Lie algebras and their representations. Graduate texts in Mathematics, 9. Springer-Verlag, New York, 1972.
- [H2] Humphreys, James E.: Conjugacy classes in semisimple algebraic groups. Mathematical Surveys and Monographs, 43. American Mathematical Society, Providence, RI, 1995.
- [K] Katz, Nicholas M.: Sums of Betti numbers in arbitrary characteristic, *Finite Fields Appl.* **7** (2001), 29–44.
- [LW] Lang, S.; Weil A.: Number of points of varieties over finite fields, *Amer. J. Math.* **76** (1954), 819–827.
- [LS1] Larsen, Michael; Shalev, Aner: Fibers of word maps and some applications, *J. Algebra* **354** (2012), 36–48.
- [LS2] Larsen, Michael; Shalev, Aner: A probabilistic Tits alternative and probabilistic identities, *Algebra and Number Theory* **10** (2016), 1359–1371.
- [M] Mann, Avinoam: Finite groups containing many involutions, *Proc. Amer. Math. Soc.* **122** (1994), 383–385.
- [N] Neumann, Peter M.: Two combinatorial problems in group theory, *Bull. London Math. Soc.* **21** (1989), 456–458.
- [R] Robbins, David P.: Jordan elements in a free associative algebra, I, *J. of Algebra* **19** (1971), 354–378.

- [Sh] Shalev, Aner: Probabilistically nilpotent groups, *Proc. Amer. Math. Soc.* **146** (2018), 1529–1536.
- [Z1] Zelmanov, Efim: Solution of the restricted Burnside problem for groups of odd exponent, (Russian) *Izv. Akad. SSSR Ser. Mat.* **54** (1990), 4259, 1990; translation in *Math USSR-Izv.* **36** (1990), 41-60.
- [Z2] Zelmanov, Efim: Solution of the restricted Burnside problem for 2-groups, (Russian) *Mat. Sb.* **182** (1991), 568–592; translation in *Math USSR-Sb* **72** (1992), 543–565.
- [Z3] Zelmanov, Efim: Lie algebras and torsion groups with identity, *J. Comb. Algebra* **1** (2017), 289–340.

E-mail address: mjlarsen@indiana.edu

DEPARTMENT OF MATHEMATICS, INDIANA UNIVERSITY, BLOOMINGTON, IN 47405,
U.S.A.

E-mail address: shalev@math.huji.ac.il

EINSTEIN INSTITUTE OF MATHEMATICS, HEBREW UNIVERSITY, GIVAT RAM, JERUSALEM
91904, ISRAEL