NORMAL BASES OF PI-ALGEBRAS

ALEXEI KANEL-BELOV, LOUIS H. ROWEN, AND UZI VISHNE

Abstract. Normal bases of affine PI-algebras are studied through the following stages: essential height, monomial algebras, representability, and modular reduction.

1. Height

In this survey we review the algorithmic theory of PI-algebras, in terms of normal bases, and indicate directions for further research. In view of Kemer [17], one can study normal bases in terms of the codimension theory of PI-algebras, of which Regev is the pioneer. Thus we feel this paper is appropriate for a volume honoring Regev.

Let A be an associative affine algebra over an infinite field k, generated by the set $\Omega = \{a_1, \dots, a_\ell\}$. Ordering the letters $a_1 < \cdots < a_\ell$ induces the *lexicographic* order on the set Ω^* of words in the generators over the alphabet: $w < v$ if $|w| < |v|$, or if $|w| = |v|$ and w is lexicographically smaller than v. The *normal base* of the algebra A with respect to the ordered set Ω, is the set of all words in Ω^* that cannot be written as a linear combination of smaller words [3], [10], [27]. Obviously this is a base of A (as a vector space).

This paper investigates normal bases of PI-algebras, from an algorithmic point of view. We say that an algebra A *has PI-degree* d if some multilinear (noncommutative) polynomial of degree d, having at least one coefficient 1, vanishes identically on A. In particular, if A is any subring of a matrix algebra $M_n(F)$ over a field F, then A satisfies a PI of degree $d = 2n$, by the Amitsur-Levitzki Theorem. Such a PI-algebra is called *representable* (or *admissible* in [23]). Although there are only countably many affine representable algebras over $\mathbb{Q}$ up to isomorphism, Lewin showed there are uncountably many affine PI-algebras that are homomorphic images of subalgebras of $M_3(\mathbb{Q})$; thus there are uncountably many PI-algebras that are not representable.

Date: 20 June, 2005.

This work was carried out under the auspices of the Israel Academy of Science Center of Excellency at Bar-Ilan University.

One particularly direct example of a nonrepresentable PI-algebra of L. Small is given in [23, Example 4.4.22].

The first major breakthrough for normal bases of PI-algebras was obtained by A.I. Shirshov [25], [26], via his famous height theorem:

Definition 1. *An algebra A is said to have* height $\leq h$ *over a subset Y, if A is spanned as a vector space by*

$$Y^{[h]} = \{y_1^{m_1} \dots y_t^{m_t} : \; m_1, \dots, m_t \in \mathbb{N}, \; y_1, \dots, y_t \in Y, \; t \leq h\}.$$

Theorem 2 (Shirshov's height theorem [5, Chapter 2], [10], [26])**.** *Suppose $A = k\{a_1, \dots, a_\ell\}$ has PI-degree d. Let Y be the set of words of length $\leq d$ over the generators. Then A has some height over Y, bounded as a function of d and ℓ; furthermore, for a suitable $h \in \mathbb{N}$, $Y^{[h]}$ contains a normal base of A.*

In particular, every word in $\{a_1, \dots, a_\ell\}^*$ is a product of $\leq h$ periodic words, each of which has period $\leq d$.

Since the reader may not be familiar with Shirshov's theorem in this formulation, let us review the idea of the proof, following [5, Section 2.2]. We say a word w on ℓ letters is *d-decomposable* if it contains a subword $w_1 \cdots w_d$ such that $w_1 \cdots w_d > w_{\pi(1)} \cdots w_{\pi(d)}$ for any permutation π of $\{1, \dots, d\}$. It is easy to use a PI of degree d to rewrite any d-decomposable word as a sum of smaller words; thus the irreducible words are d-indecomposable. Shirshov proved *Shirshov's Lemma*, which asserts that, for any given $r > 0$, any long enough d-indecomposable word must contain a nonempty word u^r where $|u| \leq d$. Shirshov's height theorem then follows from an algorithmic argument given in [5, p. 50].

Accordingly, we say a subset $Y \subset A$ is a *Shirshov base* if A has finite height over Y. Shirshov's theorem also provides an immediate solution to Kurosch's problem for PI-algebras (solved earlier by Kaplansky):

Corollary 3. *If an affine PI algebra A is algebraic, then it is finite dimensional.*

Shirshov's Lemma being the key to Shirshov's theorems, we are led to a question of considerable interest:

Question 4. *How well can one bound "long enough" in Shirshov's Lemma as a function in d, r, and ℓ?*

The answer also provides a bound for the dimension of A, assuming it is algebraic. The best known bound, due to Belov, is described in detail in [10].

Since the combinatoric results do not depend on A having a unit element (and in fact, can even be formulated for nonassociative algebras), Shirshov's theorem also implies that every nil affine PI-algebra is nilpotent. A well-known theorem of Wedderburn states that every nilpotent subring A of a matrix algebra $M_n(F)$ satisfies $A^n = 0$. (On the other hand, the ring of strictly upper triangular matrices satisfies $A^n = 0$ but $A^{n-1} \neq 0$.)

Putting these various facts together, if $A = k\langle\Omega\rangle$ is an affine subalgebra of $M_n(F)$ such that the words in Ω of length $\leq d = 2n$ are nilpotent, then $A^n = 0$. Amitsur and Shestakov conjectured that it is enough to require nilpotency of the words in Ω of length $\leq n$; this was proved independently by Ufnarovsky [27] and Chekanu [8]; a short proof of Belov [3] is given in [5, Corollary 2.82]. In fact, Belov improved this result to algebraicity:

Theorem 5. *If A is an affine PI algebra, and the matrix algebra $\mathrm{M}_{n+1}(F)$ does not satisfy all the identities of A, then the words of length $\leq n$ comprise a Shirshov base of A.*

The proof can be found in [3], [5, Exercise 9.18], [6] and (with a different approach) [10].

The height theorem leads to other questions for further investigation:

Problem 6. *Given a PI-algebra A, estimate its height over a given generating set. Upper bounds (in terms of the number of generators, the PI degree and the minimal degree of an identity not satisfied by $\mathrm{M}_n(F)$) were obtained in* [1] *and* [3].

Problem 7. *To describe those subsets Y over which A has some height.*

Let us formulate these concepts more precisely.

Definition 8. *An algebra A is said to have* essential height $\leq h$ *over a subset Y, if there is a finite set $S \subset A$ (which may depend on Y) such that A is spanned as a vector space by*

$$Y^{[h],S} = \{s_0 y_1^{m_1} s_1 \dots s_{t-1} y_t^{m_t} s_t \colon\ m_i \in \mathbb{N},\ y_i \in Y,\ s_i \in S,\ t \leq h\}.$$

In this case, Y is called an essential Shirshov base*, and S the* supplementary *set. Clearly we may always expand a supplementary set S; in particular we assume $1 \in S$.*

Note that if Y is an essential Shirshov base and generates A as an algebra, then Y is a Shirshov base. The minimal h in Definition 8 is called the 'essential height' of A (with respect to Y), and denoted by $\mathrm{H}_{\mathrm{ess}}(A, Y)$. By Shirshov's theorem, a PI-algebra A has finite essential height with respect to any finite set of generators.

From a different viewpoint, if Y is an essential Shirshov base of A, then any homomorphic image of A in which the elements of Y are algebraic, is finite dimensional.

The naive converse does not hold.

Example 9. *Let $A = k[x, 1/x]$ and $Y = \{x\}$. Then Y is not a Shirshov base of A, although every homomorphic image $\bar{A}$ of A in which x is algebraic is finite dimensional over k.*

A finite subset $W = \{w_1, \dots, w_t\}$ of A is a *Kurosch set*, if for some m,

$$A \otimes_k k[\Lambda] / \left\langle w_j^m - \sum_{i=0}^{m-1} \lambda_i^{(j)} w_j^i : 1 \leq j \leq t \right\rangle$$

is a finite module over $k[\Lambda]$ where $\Lambda = \{\lambda_i^{(j)}\}_{0 \leq i < m, 1 \leq j \leq t}$. In other words when we make each w_j integral of degree m over $k[\Lambda]$, the image of $A[\Lambda]$ becomes a finite module.

Theorem 10 ([5, Exer. 9.20])**.** *W is a Kurosch set iff W is an essential Shirshov base.*

2. Growth of affine PI-algebras vs. essential height

The usual way one nowadays studies growth of the affine algebra A generated by $\Omega = \{a_1, \dots, a_\ell\}$ is by means of the (Poincaré-)Hilbert series $H(A)$, defined as

$$H(A) = 1 + \sum_{n \geq 1} d_n \lambda^n,$$

where $d_n = \dim_k \left(\sum_{j=0}^n k\Omega^j \right)$. Of particular interest is the *Gelfand-Kirillov dimension*

$$\mathrm{GKdim}(A) = \overline{\lim_{n \to \infty}} \log_n d_n, \tag{1}$$

A good reference for Hilbert series and GK dimension is [20]. We say the Hilbert series is *rational* if it is a rational function in λ; otherwise it is called transcendental. Strictly speaking, the rationality of the series depends on the choice of generating set (even though the GK dimension is independent of the generating set). Nevertheless, the Hilbert series of a commutative affine algebra is always rational.

It is easy to see that if Y is an essential Shirshov base of A, then $\mathrm{GKdim}(A) \leq \mathrm{H}_{\mathrm{ess}}(A, Y)$.

Corollary 11. *The Gelfand-Kirillov of an affine PI algebra is finite.*

This raises the question of when is the $\mathrm{GKdim}(A)$ equal to $\mathrm{H}_{\mathrm{ess}}(A, Y)$. Clearly, the growth of A is maximal when A is *relatively free*, i.e., satisfies no relations other than those required by its polynomial identities. See [5, Chapter 3] for a more precise definition. On the other hand, our estimates of essential height were all made in terms of d, k, and ℓ, which remain the same when we pass to the relatively free affine algebra. Thus it is reasonable to start with relatively free algebras.

Proposition 12. *Relatively free PI-algebras are representable.*

This result follows without difficulty from Kemer's theorem that any affine PI-algebra over a field k satisfies the same PI's as a suitable finite dimensional k-algebra A, say with base $b_1, \dots, b_n$. Indeed, there is a construction for the relatively free algebra of a finite dimensional algebra, using "generic elements" $\tilde{a}_i = \sum_{j=1}^{n} \lambda_{ij} b_j$, where the λ_{ij} are commuting indeterminates, and this algebra is clearly contained in $A \otimes k(\Lambda) \subset M_n(F)$ where $F = k(\{\lambda_{ij}\})$; thus it is representable.

Details are given in characteristic 0 in [5, Corollary 4.67]. Kemer [16] handles the characteristic $p > 0$ case.

Representable PI algebras do exhibit good behavior with respect to the Gelfand-Kirillov dimension:

Theorem 13. *If A is a representable affine algebra with an essential Shirshov base Y, then* $\mathrm{GKdim}(A) = \mathrm{H}_{\mathrm{ess}}(A, Y)$.

Corollary 14. *If A is representable, then* $\mathrm{GKdim}(A)$ *is an integer, and* $\mathrm{H}_{\mathrm{ess}}(A, Y)$ *is independent of respect to the essential Shirshov base* Y.

In particular, the Gelfand-Kirillov dimension of a relatively free affine PI-algebra is an integer. Also, any relatively free PI-algebra has a rational Hilbert series, cf. [5, Theorem 9.44 and Corollary 9.45], although [5, Example 9.39] presents a representable algebra with transcendental Hilbert series (but clearly with integral Gelfand-Kirillov dimension). We summarize the various interrelations in the following diagram.

$$\begin{tikzcd}
\text{relatively free PI} \arrow[r, Rightarrow] \arrow[d, Rightarrow] & \text{representable} \arrow[rr, Rightarrow] \arrow[d, Rightarrow] \arrow[dl, "|" marking] & & \text{PI} \arrow[d, Rightarrow] \\
H_A \text{ rational} \arrow[r, Rightarrow] & \text{integral GKdim} & & \text{has Shirshov base}
\end{tikzcd}$$

FIGURE 1

3. Monomial algebras

An algebra is **monomial** if it can be described in terms of relations that are monomials in the generators. Besides being basic to computer science, monomial algebras play an important role in the theory of growth, since given a presentation of an affine algebra A, it is an easy matter to define the *associated monomial algebra* having the same Hilbert series; namely one factors the free algebra by the set of reducible words in the generators of A, *cf.* [5, Proposition 9.8]. Note that the associated monomial algebra of A also has the same Shirshov basis. This procedure provides a way to study an arbitrary affine algebra. However, this construction does not respect polynomial identities (or other key properties, such as finite presentation).

If a monomial algebra is representable, then it is PI and so has finite height over some finite set of words in the generators. The converse does not hold (for example, an algebra with a nonintegral Gelfand-Kirillov dimension cannot be representable, by Corollary 14). In this section we formulate and prove a criterion for the representability of a monomial algebra.

Let A be an affine PI monomial algebra. By the height theorem, A has bounded essential height over a (finite) Shirshov base Y, which we may assume to be a set of words in the generators. Let S be a supplementary set as in the notation of Definition 8; moreover assume $Y \subseteq S$. Choose a subset of $Y^{[h],S}$ which is a basis of A. Given

$$w = s_0 y_1^{m_1} s_1 \dots s_{t-1} y_t^{m_t} s_t \tag{2}$$

(with $y_i \in Y$ and $s_i \in S$, and t bounded by the height), we rewrite it in the same manner with $s_0 \in S$ of maximal possible length, then with $y_1^{m_1}$ of maximal possible length, and so on. The assumption that $Y \subseteq S$ guarantees that no s_i equals 1. Then we call $(s_0, y_1, s_1, \dots, s_{t-1}, y_t, s_t)$ the *type* of w. We may assume that $m_i > 0$ for all i, by adjoining $s_i s_{i+1}$ to S if necessary. Furthermore, we may assume that the exponents m_i in words of any given type are unbounded; otherwise, by enlarging S the type could be replaced by a shorter one.

The type of a subword of a w of type θ is called a *subtype* of θ. The type of a word equal to zero is the empty type, which we disregard. If $\theta = (s_0, y_1, s_1, \dots, y_t, s_t)$ does not occur as the type of a (nonzero) word, we say that θ is empty.

The exponents $(k_1, \dots, k_s)$ related to an type form a subset of $\mathbb{N}^s$, which we will denote by Λ_θ. Summarizing:

Proposition 15. *Every word in the generators of A has a unique type, and there are finitely many types.*

Let k denote our base field. By an *exponential polynomial* in the variables $m_1, \dots, m_t$ we mean a polynomial in the m_i, as well as in expressions of the form α^{m_i} where α is algebraic over k — more precisely, an expression of the form

$$\sum f_j(m_1, \dots, m_t)\alpha_{1j}^{m_1} \cdots \alpha_{tj}^{m_t}$$

where f_j are polynomials over a finite algebraic extension K of k, and $\alpha_{ij} \in K$.

We can now formulate the representability criterion.

Theorem 16 ([3, Thm. 6.26]). *A monomial algebra A over k is representable iff:*

1. *A has essential height over a finite set Y (with a supplementary set S), such that Proposition 15 holds.*
2. *For each type $\theta = (s_0, y_1, s_1, y_2, \dots, y_t, s_t)$, there is a finite system $P_{\theta,j}$ of exponential polynomials over an algebraic extension of k in the variables $m_1, \dots, m_t$, such that the following condition holds:*
$$s_0 y_1^{m_1} s_1 \cdots s_{t-1} y_t^{m_t} s_t \neq 0$$
if and only if
$$\exists j \quad P_{\theta,j}(m_1, \dots, m_t) \neq 0,$$
and these are the only nonzero words of A.
3. *Any solution for the system of equations for a subtype is also a solution for the system of equations for the type.*

As the system of equations associated to an empty type, we may take the zero polynomial. The 'only if' part of the proof follows easily from Jordan decomposition:

Proposition 17. *Let C be a square $r \times r$ matrix over a field k. Then there are: a finite field extension K/k, matrices $C_1, \dots, C_r \in \mathrm{M}_r(K[\lambda])$ of polynomials over K (of degree $\leq r$), and elements $\alpha_1, \dots, \alpha_r \in K$, such that for every $m \in \mathbb{N}$, the mth power of C is*

$$C^m = \sum_{i=1}^{r} C_i(m) \cdot \alpha_i^m,$$

where $C_i(m) \in \mathrm{M}_r(K)$ is the matrix obtained by substituting m for λ in C_i.

Proof. If $C = \alpha I_r + \sum_{i=1}^{r-1} e_{r,r+1}$ is a Jordan block, then the (i,j)th entry in C^m is $\alpha^{-(j-i)}\binom{m}{j-i}\alpha^m$ for $j \geq i$, and 0 otherwise. The assertion then follows from the fact that over a suitable algebraic extension of the base field, C is similar to its Jordan decomposition. □

Necessity of the conditions in the theorem follows easily, since by the proposition, equality to zero of a word of the given type means the vanishing of the components of the corresponding matrix.

Conversely, suppose A is a monomial algebra with a Shirshov basis Y and supplementary set S, with finitely many types, each endowed with a system of exponential polynomials $P_{\theta,j}$ as in the theorem. We need to show that A is representable. Alternatively, since Y and S with the system of equations specifies a presentation of A, it is enough to construct a representable algebra with the given presentation.

Reduction 1. We may assume that A has only one (nonempty) type. Indeed, suppose A has types $\theta_1, \dots, \theta_k$, and let $A_1, \dots, A_k$ be monomial algebras generated by copies of Y and S, where the only nonempty type of A_i is θ_i. Then the algebra $A \subseteq A_1 \times \cdots \times A_k$ generated by the diagonal elements $(y, y, \dots, y)$ $(y \in Y)$ and $(s, s, \dots, s)$ $(s \in S)$ has precisely the given presentation, as seen by comparing components. Moreover (as we shall see) the A_i are representable, say each acting on a vector space V_i, and thus so is $A_1 \times \cdots \times A_k$, by its action on the direct sum $V_1 \oplus \cdots \oplus V_k$.

Reduction 2. Recall that the elements of Y and S are words on the original generators Ω. We claim that one may assume that the generators composing the s_i and y_i are all distinct (and, by construction, no s_i or y_i of a type equals 1).

For simplicity of notation, we (temporarily) renumber the components of the type as $(s_0, s_1, s_2, \dots, s_u)$, and agree that $s_{2i} \in S$ and $s_{2i+1} \in Y$. Write each s_i as a product $s_i = \omega_{i1} \dots \omega_{it_i}$, where the ω_{ij} are in Ω, not necessarily distinct. Let $\hat{\Omega}$ be a 'generic' set of generators, composed of new generators $\hat{\omega}_{ij}$ which are by definition distinct. Now let $\hat{s_i} = \hat{\omega}_{i1} \dots \hat{\omega}_{it_i}$. Let A' be the algebra having the single type $(\hat{s_0}, \dots, \hat{s_u})$. Then A embeds into A' by sending each ω_{ij} to the sum of all $\hat{\omega}_{uv}$ such that $\omega_{uv} = \omega_{ij}$. But A' satisfies the assertion of the reduction, and representability of A follows from that of A'.

Reduction 3. We may assume the single nonempty type of A has a single defining exponential polynomial. Indeed, as in Reduction 1, if $A_1, \dots, A_k$ are the algebras defined for the given type with distinct equations $P_1, \dots, P_k$, and each A_i acts on a vector space V_i, then the diagonal embedding (as before) can be presented with the system $\{P_1, \dots, P_k\}$.

Thus we are left with a single type $(s_0, y_1, \dots, y_t, s_t)$ and an exponential equation $Q(m_1, \dots, m_t)$. We need to find a monomial representable algebra generated by y_i and s_i (over an extension of k), such that $s_0 y_1^{m_1} \dots y_t^{m_t} s_t \neq 0$ iff $Q(m_1, \dots, m_t) \neq 0$.

The first step is to note that any power m_i^u is a linear combination of binomial expressions of the form $\binom{m_i}{u'}$ for $u' \leq u$. Hence, Q can be written in the form

$$(3)\qquad Q(m_1,\dots,m_t) = \\ = \sum_{u_1,\dots,u_t,j} c_{\vec{u}} \cdot \alpha_{\vec{u},1}^{m_1}\binom{m_1}{u_1-1}\alpha_{\vec{u},2}^{m_2}\binom{m_2}{u_2-1}\dots\alpha_{\vec{u},t}^{m_t}\binom{m_t}{u_t-1},$$

where $\vec{u} = (u_1,\dots,u_t,j)$ (j is added to allow more than one product with the same $u_1,\dots,u_t$). Let K denote the (finite dimensional) extension of k generated by all the $\alpha_{\vec{u},i}$, and take $\bar{K} = K(\lambda_1,\dots,\lambda_t)$. We will construct the representable algebra directly as $\bar{K}$-maps of a suitable $\bar{K}$-vector space.

For each monomial $\vec{u} = (u_1,\dots,u_t)$ in the expression for Q, let $V_{\vec{u}} = V_{\vec{u},1} \oplus \cdots \oplus V_{\vec{u},t}$, where $V_{\vec{u},i}$ is a u_i-dimensional vector space over K (with an ordered basis). Take $V = \oplus V_{\vec{u}}$. Define operators $T_i : V \to V$ by acting on each $V_{\vec{u},i'}$ as follows: For $i' = i$, T_i acts as a Jordan matrix of size u_i with the eigenvalue $\alpha_{\vec{u},i}$; and if $i' \neq i$, then T_i is the zero operator.

For $0 < i < t$, we define s_i on $V_{\vec{u},i'}$ as follows: for $i' = i$, s_i maps the last basis element of $V_{\vec{u},i}$ to the first basis vector of $V_{\vec{u},i+1}$ (and sends the others to zero). For $i' \neq i$, s_i is the zero operator. With this choice of the operators, the coefficient of the matrix unit $e_{1,u_1+\cdots+u_t}$ in $T_1^{m_1}s_1\cdots s_{t-1}T_t^{m_t}$ (operating from left to right on $V_{\vec{u}}$) is the monomial corresponding to $\vec{u}$ in (3). (This follows from the calculation of Proposition 17). Let $V_{\vec{0}}$ be a designated one-dimensional component of V. We define $s_0 : V \to V$ by letting $s_0 : V_{\vec{0}} \to V_{\vec{u}}$ be the injection into the first entry (and zero on every other pair of components); dually we let $s_t : V \to V$ be defined on $V_{\vec{u}} \to V_{\vec{0}}$ by multiplying the $u_1 + \cdots + u_t$ entry by $c_{\vec{u}}$. Now $s_0 T_1^{m_1} s_1 \cdots s_{t-1} T_t^{m_t} s_t$ equals $Q(m_1,\dots,m_t)$ times the matrix unit $e_{\vec{0},\vec{0}}$. Finally take $y_i = \lambda_i T_i$. Then

$$s_0 y_1^{m_1} s_1 \cdots s_{t-1} y_t^{m_t} s_t = \lambda_1^{m_1} \dots \lambda_t^{m_t} Q(m_1,\dots,m_t) e_{\vec{0},\vec{0}}$$

which are linearly independent over k, so there are no additional relations, and we have constructed the desired monomial algebra.

Example 18. *Let us construct a representable monomial algebra with the type* $(s_0, y_1, s_1, y_2, s_2)$ *and the equation*

$$Q(m_1,m_2) = 4^{m_1} m_1 5^{m_2} m_2 - 3 \cdot 2^{m_1} 5^{m_2} m_2.$$

There are two products, corresponding to $\vec{u} = (2,2)$ *and* $\vec{u'} = (1,2)$. *To these we add* $\vec{0} = (1,1)$, *and so we act on* $V = V_{\vec{u}} \oplus V_{\vec{u'}} \oplus V_{\vec{0}}$, *a*

8-dimensional space. We view $V_{\vec{u}}$ as occupying the first to fourth entry, and so on.

The construction above suggests

$$T_1 = \left(\begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix} \oplus \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \oplus \left((2) \oplus \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \oplus (1),$$

$$T_2 = \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \oplus \begin{pmatrix} 5 & 1 \\ 0 & 5 \end{pmatrix}\right) \oplus \left((1) \oplus \begin{pmatrix} 5 & 1 \\ 0 & 5 \end{pmatrix}\right) \oplus (1),$$

$s_0 = e_{1,8} + e_{5,8}$, $s_1 = e_{3,2} + e_{6,5}$, *and* $s_2 = e_{8,4} - 3e_{8,7}$. *One can check that indeed* $s_0 T_1^{m_1} s_1 T_2^{m_2} s_2 = Q(m_1, m_2) \cdot e_{8,8}$, *and every product not of this form is zero.*

4. Power vectors

Suppose A has height $\leq h$ over a set Y. Having studied the relation between A and Y, it remains for us to consider the 'power vectors' of A with respect to Y, defined as all vectors $(m_1, \dots, m_h) \in \mathbb{N}^h$ such that

$$w = y_1^{m_1} \dots y_h^{m_h}$$

is irreducible for some choice of $y_1, \dots, y_h \in Y$.

Our goal is to describe the power vectors of A. This can be fairly complicated even in the monomial case, cf. Theorem 16. The construction of monomial algebras is thus equivalent to the solution of arbitrary exponential polynomials. But this is algorithmically unsolvable by the celebrated theorem of Davis-Putnam-Robinson [9]. Thus we conclude in characteristic zero:

Proposition 19 ([3]). *The isomorphism problem for two subalgebras of the algebra of matrices over the ring of polynomials, given by their generators, is algorithmically unsolvable.*

However, the situation is different in positive characteristic. Let $p > 0$ be prime, and $(m_1, \dots, m_\ell) \in \mathbb{N}^\ell$. Write

$$m_i = \sum_{j=0}^{N} a_{ij} p^j,$$

for $a_{ij} \in \{0, \dots, p-1\}$ and some $N \in \mathbb{N}$. The 'p-adic presentation' of $(m_1, \dots, m_\ell)$ is defined as the series of vectors $(a_{i0})_i, (a_{i1})_i, \dots, (a_{iN})_i$.

Let X be a finite set. Recall that a set of (finite) words in X^* is called a 'language', and that a language W is 'regular' if there is a finite graph with two designated vertices e_0, e_1 and edges labelled by letters from X, such that W is the set of words obtained by concatenating the labels over a path, ranging over all paths from e_0 to e_1. We say that the graph 'presents' the language. The main result of [4] is as follows:

Theorem 20 ([4]). *Suppose $P_u(m_1, \dots, m_\ell) = 0$ are finitely many exponential Diophantine equations in the parameters $m_1, \dots, m_\ell$, over a field of characteristic $p > 0$.*

Let W be the language composed of all the p-adic presentations of vectors $(m_1, \dots, m_\ell)$ satisfying the equations P_u. Then W is a regular language. Moreover there is an algorithm to construct a graph presenting W.

Corollary 21. *Suppose we are given finitely many exponential polynomials $P_u(m_1, \dots, m_t)$ in characteristic $p > 0$. There is an algorithm to decide whether or not there is a solution to the system of equations $\forall u\colon\ P_u(m_1, \dots, m_t) = 0$.*

Proof. Since the details of Theorem 20 and Corollary 21 are only available in Russian (cf. [4]), let us give the main idea of the proof of Corollary 21. We assume there is a single indeterminate m, and that the base field $k = \mathbb{F}_p(x)$ is the field of rational functions in one variable over the prime field. Furthermore we assume each P_u can be written in the form

$$P_u(x, m) = \sum_{j=1}^{t} r_u^{(j)}(x)\alpha_j(x)^m \tag{4}$$

where $r_u^{(j)}(x), \alpha_j(x) \in \mathbb{F}_p[x]$. If we did not make the assumption that the α_j are in k, but rather permitted them to be in a finite extension field, we would need to view the coefficients as matrices over k via the regular representation; the proof would be along the same lines, but much more intricate.

Let $\mathcal{F}$ denote the original system of equations $\{\exists m \forall u : P_u(x, m) = 0\}$.

Let $C = \max_{u,j}\{\deg(r_u^{(j)}), \deg(\alpha_j)\}$. Writing $m = m_0 + m_1 p$ (where $0 \le m_0 < p$), we have that $\alpha_j(x)^m = \alpha_j(x)^{m_0}\alpha_j(x^p)^{m_1}$ and so

$$P_u(x, m) = \sum_{j=1}^{t} r_u^{(j)}(x)\alpha_j(x)^{m_0}\alpha_j(x^p)^{m_1}.$$

For every u and j and every m_0, write $r_u^{(j)}(x)\alpha_j(x)^{m_0} = \sum_{i=0}^{p-1} x^i R_{m_0,u,i}^{(j)}(x^p)$ for suitable polynomials $R_{m_0,u,i}^{(j)}$, and note that $\deg(R_{m_0,u,i}^{(j)}) \le \frac{1}{p}(\deg(r_u^{(j)}) + (p-1)\deg(\alpha_j)) \le C$. Every equation of the form $P_u(x, m) = 0$ can now be written as

$$P_u(x, m_0 + pm_1) = \sum_{i=0}^{p-1} x^i \sum_{j=1}^{t} R_{m_0,u,i}^{(j)}(x^p)\alpha_j(x^p)^{m_1} = 0,$$

or equivalently

$$\sum_{j=1}^{t} R^{(j)}_{m_0,u,i}(y)\alpha_j(y)^{m_1} = 0 \tag{5}$$

for every $i = 0,\dots,p-1$, replacing x^p by y. Obviously the system of equations (4) (ranging over u) has a solution iff the system (5) (ranging over i and u, with m_0 fixed) has one.

The degree of the polynomial coefficients $R^{(j)}_{m_0,u,i}$ remains bounded by C, and so there are finitely many possible vectors of coefficients. It follows that only finitely many systems of equations are obtained in this process; let us denote this collection of systems by $\mathcal{L}$. Any solution m to a system can be reduced to a solution $m_1 = [m/p]$ of another system, and so the original system $\mathcal{F}$ has a solution if and only if one of the systems in $\mathcal{L}$ has a solution with $m \leq p$. This reduces the solution of $\mathcal{F}$ to a finite number of steps. □

Theorem 22. *The isomorphism problem for monomial subalgebras of the matrix algebra in polynomials over a field of char $p > 0$ (defined in terms of their generators) is algorithmically solvable.*

Proof. By Theorem 16, a representable monomial algebra is determined by finitely many exponential polynomials (the proof is constructive), and in characteristic p we have an algorithm to find their solution. □

It would be nice if Theorem 22 held for arbitrary representable algebras (not necessarily monomial) in characteristic > 0. Towards this end, we pose a conjecture:

Conjecture 23. *If A is a representable algebra over a field of positive characteristic, then the set of power vectors determines a regular language.*

Note that this conjecture holds for monomial algebras, by Theorems 16 and 20.

Given this discrepancy between characteristic 0 and characteristic $p > 0$, we would like to study affine algebras, at least in the relatively free case, by passing modulo p. Unfortunately this cannot be done naively, due to counterexamples of Schelter [24] and Asparouhov-Drensky-Koev-Tsiganchev [2]. But the idea does work for large enough p.

Theorem 24. *Suppose A is a relatively free affine algebra over $\mathbb{Z}$, with the standard set of generators. Then a normal base of $A\otimes\mathbb{Q}$ is mapped onto a normal base of $A\otimes\mathbb{Z}/p$ for all sufficiently large p. In particular $H_{A\otimes\mathbb{Q}} = H_{A\otimes\mathbb{Z}/p}$.*

Proof. See [5, Exercise 9.32]. This exercise follows readily from the extensive hint given in [5, Exercise 9.31]. □

Thus, taking p as in the theorem, one could solve the isomorphism problem for two relatively free PI-algebras. Unfortunately, we do not yet have a way of determining p. We can make these conditions more precise using the following program.

Proposition 25 ([18])**.** *Let $S \subseteq \mathbb{N}$. Suppose for primes $p \neq p'$, that the languages of p-adic and p'-adic presentations of S are both regular. Then S is a union of a finite set and a finite union of arithmetic progressions.*

Conjecture 26 (Generalization to the multivariate case)**.** *Suppose $S \subseteq \mathbb{N}^k$ has p-adic and p'-adic regular presentations. Then S is a finite union of shifts of finitely generated semigroups of $\mathbb{N}^k$.*

Conjectures 23 and 26 together with Theorem 24 would imply:

Conjecture 27. *Let A be a relatively free affine algebra over a field of characteristic zero. Then the set of power vectors of a Shirshov base of A is regular, i.e., can be described as in Conjecture 26.*

References

[1] Amitsur, S. A. and Small, L. W. *Affine algebras with polynomial identities*, Recent developments in the theory of algebras with polynomial identities (Palermo, 1992). Rend. Circ. Mat. Palermo (2) Suppl. No. 31, 9–43, (1993).

[2] Asparouhov, T., Drensky, V., Koev, P. and Tsiganchev D. *Generic* 2×2 *matrices in positive characteristic*, J. Algebra **225**(1), 451–486, (2000).

[3] Belov, A.Ya., Borisenko, V.V., and Latyshev, V.N., "Monomial algebras". Algebra 4, J. Math. Sci. (New York) **87**(3), 3463-3575, (1997).

[4] Belov, A.Ya. and Chilikov, A.A. Exponential Diophantine equations in rings of positive characteristic (Russian) Fundam. Prikl. Mat. **6**(3), 649–668, (2000).

[5] Belov, A.Ya. and Rowen, L.H. "Computational aspects of polynomial identities". Research Notes in Mathematics, 9. AK Peters, Ltd., Wellesley, MA, 2005.

[6] Belov, A.Ya., *On Shirshov bases in relatively free algebras of complexity n*, Mat. Sb. **135** (1988), no. 3, 373–384.

[7] Bernšteĭn, I. N., Gelfand, I. M. and Ponomarev, V. A. *Coxeter functors, and Gabriel's theorem* (Russian) Uspehi Mat. Nauk **28**(2(170)), 19–33, (1973).

[8] Chekanu, G.P., *Independency and quasiregularity in algebras*, Dokl. Akad. Nauk **337**(3), 316-319, (1994); translation: Russian Acad. Sci. Dokl. Math. **50**(1), 84–89, (1995).

[9] Davis M., Putnam, H., and Robsinson, J. *The decision problem for exponential differential equations*, Annals of Math. **74**, 425-436, (1961).

[10] Drensky, V., "Free Algebras and PI-algebras: Graduate Course in Algebra", Springer-Verlag, Singapore, (2000).

[11] Formanek, E., *Invariants and the ring of generic matrices*, J. Algebra **89**(1), 178–223, (1984).

[12] Kaplansky, I., *Topoloigcal representation of algebras. II*, Trans. Amer. Math. Soc. **66** (1949), 464–491.

[13] Kemer, A.R., *The representability of reduced-free algebras*, Algebra i Logika **27**(3), 274–294, (1988).

[14] Kemer, A.R., "Identities of Associative Algebras", Transl. Math. Monogr., **87**, Amer. Math. Soc. (1991).

[15] Kemer, A.R., "Ideals of Identities of Associative Algebras", Amer. Math. Soc. Translations of monographs **87** (1991).

[16] Kemer, A.R., *Identities of finitely generated algebras over an infinite field*, Math-USSR Izv.**37**, 69-96, (1991).

[17] Kemer, A.R., *Multilinear Identities of the algebras over a field of characteristic p*, Inter. J. Algebra Comp. **5**(2), 189–197, (1995).

[18] Kudryavtsev, V. B., Aleshin, S. V. and Podkolzin, A. S. "Введение в теорию автоматов", 320 pp. (Russian) [Introduction to automata theory] "Nauka", Moscow, 1985.

[19] Lewin, J., *A matrix representation for associative algebras. I and II*, Trans. Amer. Math. Soc. **188**(2), 293–317, (1974).

[20] Krause G.R. and Lenagan T.H., "Growth of Algebras and Gelfand-Kirillov Dimension", rev. edition, GSM **22**, AMS press, 2000.

[21] Razmyslov, Yu.P., and Zubrilin, K.A., *Capelli identities and representations of finite type*, Comm. Algebra **22**(14), 5733–5744, (1994).

[22] Regev, A., *Codimensions and trace codimensions of matrices are asymptotically equal*, Israel J. Math. **47** (1984), 246–250.

[23] Rowen, L.H., "Polynomial Identities in Ring Theory", Acad. Press Pure and Applied Math, **84**, New York, 1980.

[24] Schelter, W.F. , *On a question concerning generic matrices over the integers*, J. Algebra **96**(1) (1985), 48–53.

[25] Shirshov, A.I., *On some nonassociative nil-rings and algebraic algebras*, Mat. Sb. **41**(3), 381–394, (1957).

[26] Shirshov, A.I., *On rings with identity relations*, Mat. Sb. **43**(2), 277–283, (1957).

[27] Ufnarovskiĭ, V.A., *A theorem on independence and its corollaries* (Russian) Mat. Sb. (N.S.) **128**(170(1)), 124–132, 144, (1985).

Einstein Institute of Mathematics, Givat Ram, The Hebrew University, Jerusalem, 91904, Israel

E-mail address: kanel@mccme.ru

Deptartment of Mathematics, Bar Ilan University, Ramat Gan 52900, Israel

E-mail address: {rowen,vishne}@math.biu.ac.il