ЧЕБЫШЕВСКИЙ СБОРНИК
Том 21. Выпуск 3.

УДК 512.552 DOI 10.22405/2226-8383-2020-21-3-89-128

Теорема Размыслова — Кемера — Брауна для афинных PI-алгебр[1]

Канель-Белов А. Я., Роуэн Луис Халли

Александр Яковлевич Канель-Белов —доктор физико-математических наук, федеральный профессор математики MIPT, профессор Университета им. Бар-Илана (Израиль).
e-mail: kanelster@gmail.com
Роуэн Луис Халли — Университет им. Бар-Илана (Израиль).
e-mail: kanelster@gmail.com

Аннотация

Дается замкнутое в себе альтернативное комбинаторное изложение доказательство теоремы Размыслова-Кемера-Ббрауна о нильпотентности радикала афинной PI-алгебры над нетеровым ассоциативно-коммутативным кольцом. В свое время сообщество не верило в справедливость этого результата и вопреки общественному мнению соответствующий вопрос был поставлен В.Н.Латышевым в его докторской диссертации.

Начнем с определения:

1. Алгебра A является **аффинной** над коммутативным кольцом C, если A порождается как алгебра над C конечным числом элементов $a_1, \ldots, a_\ell$; в этом случае мы пишем $A = C\{a_1, \ldots, a_\ell\}$.
2. Мы говорим, что алгебра A является **конечной**, если A порождено как C-модуль конечным числом элементов.
3. Алгебры над полем называются **PI алгебрами**, если они удовлетворяют (нетривиальным) полиномиальным тождествам.
4. **Многочлен Капелли** Cap_k степени $2k$ определяется так:

$$\mathrm{Cap}_k(x_1, \ldots, x_k; y_1, \ldots, y_k) = \sum_{\pi \in S_k} \mathrm{sgn}(\pi) x_{\pi(1)} y_1 \cdots x_{\pi(k)} y_k$$

5. $\mathrm{Jac}(A)$ обозначает радикал Джекобсона алгебры A, который для PI-алгебр является пересечением максимальных идеалов A по теореме Капланского.

Ключевые слова: алгебры с полиномиальными тождествами, многообразия алгебр; представимые алгебры; относительно свободные алгебры; Ряды Гильберта; Проблема Шпехта

Библиография: 34 названий.

Для цитирования:
Канель-Белов А.Я., Роуэн Луис Халли Теорема Размыслова — Кемера — Брауна для афинных PI-алгебр // Чебышевский сборник, 2020, т. 21, вып. 3, с. 89–128.

[1]А.Я.Белов был поддержан РНФ (грант 17-11-01377), Л.Роуэн ISF 1623/16.

CHEBYSHEVSKII SBORNIK

Vol. 21. No. 3.

UDC 512.552 DOI 10.22405/2226-8383-2020-21-3-89-128

The Braun–Kemer–Razmyslov Theorem for affine PI-algebras

Alexei Kanel Belov, Louis Rowen

Alexei Kanel Belov — doctor of physical and mathematical sciences, federal professor MIPT, professor Bar Ilan University (Israel).
e-mail: kanelster@gmail.com
Louis Rowen — Bar-Ilan University (Israel).
e-mail: kanelster@gmail.com

Abstract

A self-contained, combinatoric exposition is given for the Braun–Kemer–Razmyslov Theorem over an arbitrary commutative Noetherian ring.At one time, the community did not believe in the validity of this result, and contrary to public opinion, the corresponding question was posed by V.N. Latyshev in his doctoral dissertation.

One of the major theorems in the theory of PI algebras is the Braun-Kemer-Razmyslov Theorem. We preface its statement with some basic definitions.

1. An algebra A is **affine** over a commutative ring C if A is generated as an algebra over C by a finite number of elements $a_1, \dots, a_\ell$; in this case we write $A = C\{a_1, \dots, a_\ell\}$.

We say the algebra A is **finite** if A is spanned as a C-module by finitely many elements.

2. Algebras over a field are called **PI algebras** if they satisfy (nontrivial) polynomial identities.

3. The **Capelli polynomial** Cap_k of degree $2k$ is defined as

$$\mathrm{Cap}_k(x_1, \dots, x_k; y_1, \dots, y_k) = \sum_{\pi \in S_k} \mathrm{sgn}(\pi) x_{\pi(1)} y_1 \cdots x_{\pi(k)} y_k$$

4. $\mathrm{Jac}(A)$ denotes the Jacobson radical of the algebra A which, for PI-algebras is the intersection of the maximal ideals of A, in view of Kaplansky's theorem.

The aim of this article is to present a readable combinatoric proof of the theorem: **The**

Braun-Kemer-Razmyslov Theorem *The Jacobson radical* $\mathrm{Jac}(A)$ *of any affine PI algebra A over a field is nilpotent.*

Keywords: algebras with polynomial identity; varieties of algebras; representable algebras; relatively free algebras; Hilbert series; Specht problem

Bibliography: 34 titles.

For citation:
A. Kanel Belov, L. Rowen, 2020, "The Braun–Kemer–Razmyslov Theorem for affine PI-algebras", *Chebyshevskii sbornik*, vol. 21, no. 3, pp. 89–128.

Светлой памяти Виктора Николаевича Латышева посвящается

1. The BKR Theorem

1.1. Introduction

One of the major theorems in the theory of PI algebras is the Braun-Kemer-Razmyslov Theorem (Theorem 1.1 below). We preface its statement with some basic definitions.

Definition 1. *1. An algebra A is* **affine** *over a commutative ring C if A is generated as an algebra over C by a finite number of elements $a_1, \dots, a_\ell$; in this case we write $A = C\{a_1, \dots, a_\ell\}$.*

2. We say the algebra A is **finite** *if A is spanned as a C-module by finitely many elements.*

3. Algebras over a field are called **PI algebras** *if they satisfy (nontrivial) polynomial identities.*

4. The **Capelli polynomial** Cap_k *of degree $2k$ is defined as*

$$\mathrm{Cap}_k(x_1, \dots, x_k; y_1, \dots, y_k) = \sum_{\pi \in S_k} \mathrm{sgn}(\pi) x_{\pi(1)} y_1 \cdots x_{\pi(k)} y_k$$

5. $\mathrm{Jac}(A)$ *denotes the Jacobson radical of the algebra A which, for PI-algebras is the intersection of the maximal ideals of A, in view of Kaplansky's theorem.*

Theorem 1.1 (The Braun-Kemer-Razmyslov Theorem). The Jacobson radical $\mathrm{Jac}(A)$ of any affine PI algebra A over a field is nilpotent.

The aim of this article is to present a readable combinatoric proof (essentially self-contained in characteristic 0).

Let us put the BKR Theorem into its broader context in PI theory. We say a ring is **Jacobson** if the Jacobson radical of every prime homomorphic image is 0. For PI-rings, this means every prime ideal is the intersection of maximal ideals. Obviously any field is Jacobson, since its only prime ideal 0 is maximal. Furthermore, any commutative affine algebra over a field is Noetherian by the Hilbert Basis Theorem and is Jacobson, in view of [28, Proposition 6.37], often called the "weak Nullstellensatz," implying the following two results:

- (cf. Proposition 1.11) If a commutative algebra A is affine over a field, then $\mathrm{Jac}(A)$ is nilpotent.
- (Special case of Theorem 2) If A is a finite algebra over an affine central subalgebra Z over a field, then $\mathrm{Jac}(A)$ is nilpotent. (Sketch of proof: Passing to homomorphic images modulo prime ideals, we may assume that A is prime PI, and Z is an affine domain over which A is torsion-free. The maximal ideals of Z lift up to maximal ideals of A, in view of Nakayama's lemma, implying $Z \cap \mathrm{Jac}(A) \subseteq \mathrm{Jac}(Z) = 0$. If $0 \neq a \in \mathrm{Jac}(A)$, then writing a as integral over Z, we have the nonzero constant term in $Z \cap \mathrm{Jac}(A) = 0$, a contradiction.)

Since either of these hypotheses implies that A is a PI-algebra, it is natural to try to find an umbrella result for affine PI-algebras, which is precisely the Braun-Kemer-Razmyslov Theorem. This theorem was proved in several stages. Amitsur [1, Theorem 5], generalizing the weak Nullstellensatz, proved that if A is affine over a commutative Jacobson ring, then $\mathrm{Jac}(A)$ is nil. In particular, A is a Jacobson ring. (Later, Amitsur and Procesi [3, Corollary 1.3] proved that $\mathrm{Jac}(A)$ is locally nilpotent.) Thus, it remained to prove that every nil ideal of A is nilpotent.

It was soon proved that this does hold for an affine algebra which can be embedded into a matrix algebra, see Theorem 1 below. However, examples of Small [33] showed the existence of affine PI algebras which can not be embedded into any matrix algebra. Thus, the following theorem by Razmyslov [22] was a major breakthrough in this area.

Theorem 1.2 (Razmyslov). If an affine algebra A over a field satisfies a Capelli identity, then its Jacobson radical $\mathrm{Jac}(A)$ is nilpotent.

Although Razmyslov's theorem was given originally in characteristic zero, he later found a proof that works in any characteristic. As we shall see, the same ideas yield the parallel result:

Theorem 1.3. Let A be an affine algebra over a commutative Noetherian ring C. If A satisfies a Capelli identity, then any nil ideal of A is nilpotent.

Following Razmyslov's theorem, Kemer [15] then proved

Theorem 1.4. [15] In characteristic zero, any affine PI algebra satisfies some Capelli identity (see Theorem 3.3).

Thus, Kemer completed the proof of the following theorem:

Theorem 1.5 (Kemer-Razmyslov). If A is an affine PI-algebra over a field F of characteristic zero, then its Jacobson radical $\mathrm{Jac}(A)$ is nilpotent.

Then, using different methods relying on the structure of Azumaya algebras, Braun proved the following result, which together with the Amitsur-Procesi Theorem immediately yields Theorem 1.1:

Theorem 1.6. Any nil ideal of an affine PI-algebra over an arbitrary commutative Noetherian ring is nilpotent.

Note that to prove directly that $\mathrm{Jac}(A)$ is nilpotent it is enough to prove Theorem 1.6 and show that $\mathrm{Jac}(A)$ is nil, which is the case case when A is Jacobson, and is called the "weak Nullstellensatz." But the weak Nullstellensatz requires some assumption on the base ring C. It can be proved without undue difficulty that A is Jacobson when C is Jacobson, cf. [26, Theorem 4.4.5]. Thus, in this case the proper general formulation of the nilpotence of $\mathrm{Jac}(A)$ is:

Theorem 1.7 (Braun). If A is an affine PI-algebra over a Jacobson Noetherian base ring, then $\mathrm{Jac}(A)$ is nilpotent.

Small has pointed out that Theorems 1.6 and 1.7 actually are equivalent, in view of a trick of [25]. Indeed, as just pointed out, Theorem 1.6 implies Theorem 1.7. Conversely, assuming Theorem 1.7, one needs to show that $\mathrm{Jac}(A)$ is nil. Modding out the nilradical, and then passing to prime images, one may assume that A is prime. Then one embeds A into the polynomial algebra $A[\lambda]$ over the Noetherian ring $C[\lambda]$, and localizes at the monic polynomials over $C[\lambda]$, yielding a Jacobson base ring by [25, Theorem 2.8].

Braun's qualitative proof was also presented in [27, Theorem 6.3.39], and a detailed exposition, by L'vov [19], is available in Russian. A sketch of Braun's proof is also given in [5, Theorem 3.1.1].

Meanwhile, Kemer [17] proved:

Theorem 1.8. [17] If A is a PI algebra (not necessarily affine) over a field F of characteristic $p > 0$, then A satisfies some Capelli identity.

Together with a characteristic-free proof of Razmyslov's theorem 1.2 due to Zubrilin [34], Kemer's Theorems 1.4 and 1.8 yield another proof of the Braun-Kemer-Razmyslov Theorem 1.1. The paper [34] is given in rather general circumstances, with some non-standard terminology. Zubrilin's method was given in [7], although [7, Remark 2.50] glosses over a key point (given here as Lemma 2.13), so a complete combinatoric proof had not yet appeared in print with all the details. Furthermore, full combinatoric details were provided in [7] only in characteristic 0 because the conclusion of the proof required Kemer's difficult Theorem 1.8. We need the special case, which we call "Kemer's Capelli Theorem," that every affine PI-algebra A over an arbitrary field satisfies some Capelli identity. This can be proved in two steps: First, that A satisfies a "sparse" identity, and then a formal argument that every sparse identity implies a Capelli identity. The version given here (Theorem 4.4) uses the representation theory of the symmetric group S_n, and provides a reasonable quartic bound ($(p-1)p\binom{u+1}{2}$, where $u=\frac{2pe(d-1)^2}{3}$) for the degree of the sparse identity of A in terms of the degree d of the given PI of A.

It should be noted that every proof that we have cited of the Braun-Kemer-Razmyslov Theorem ultimately utilizes an idea of Razmyslov defining a module structure on generalized polynomials with coefficients in the base ring, but we cannot locate full details of its implementation anywhere in the literature. One of the objectives of this paper is to provide these details, in §2.5 and §2.6.1. Although the proof is rather intricate for a general expository paper, we feel that the community deserves the opportunity to see the complete argument in print.

We emphasize the combinatoric approach here. Aside from the intrinsic interest in having such a proof available of this important theorem (and characteristic-free), these methods generalize easily to nonassociative algebras, and we plan to use this approach as a framework for the nonassociative PI-theory, as initiated by Zubrilin. (The proofs are nearly the same, but the statements are somewhat more complicated. See [6] for a clarification of Zubrilin's work in the nonassociative case.) To keep this exposition as readable as we can, we emphasize the case where the base ring C is a field and prove Theorem 1.1 directly by an induction argument without subdividing it into Theorem 1.6 and the weak Nullstellensatz, although we also treat these general cases.

§2 follows Zubrilin's short paper [34], and gives full details of Zubrilin's proof of Razmy-lov's theorem 1.2. This self–contained proof is characteristic free.

To complete the proof of the BKR Theorem, it remains to prove Kemer's Capelli Theorem. In §3 we provide the proof in characteristic 0, by means of Young diagrams, and §4 contains the characteristic p analog (for affine algebras). An alternative proof could be had by taking the second author's "pumping procedure" which he developed to answer Specht's question in characteristic p, and applying it to the "identity of algebraicity" [7, Proposition 1.59]. We chose the representation-theoretic approach since it might be more familiar to a wider audience. The proof of Theorem 1.6, over arbitrary commutative Noetherian rings, is given in §5.

Remark 1.9. An early version of Theorem 3.16 was written by Amitai Regev, to whom we are indebted for suggesting this project and providing helpful suggestions all along the way. Belov belongs lemma 3.8

1.2. Structure of the proof

We assume that A is an affine C-algebra and satisfies the $n+1$ Capelli identity Cap_{n+1} (but not necessarily the n Capelli identity Cap_n), and we induct on n: if such A satisfies Cap_n then we assume that $\mathrm{Jac}(A)$ is nilpotent, and we prove this for Cap_{n+1}. For the purposes of this sketch, in Steps 1 through 3 and Step 7 we assume that C is a field F.

The same argument shows that any nil ideal N of an affine algebra A over a Noetherian ring is nilpotent, yielding Theorem 1.3. For this result we would replace $\mathrm{Jac}(A)$ by N throughout our sketch.

We write $C\{x,y,t\}$ for the **free (associative) algebra** over the base ring C, with indeterminates x_i, y_j, t_k, z, containing one extra indeterminate z for further use. This is a free module over C, whose basis is the set of **words**, i.e., formal strings of the letters x_i, y_j, t_k, z. The x and y indeterminates play a special role and need to be treated separately. We write $C\{t\}$ for the free subalgebra generated by the t_k and z, omitting the x and y indeterminates.

1. The induction starts with $n=1$. Then $n+1=2$, and any algebra satisfying Cap_2 is

commutative. We therefore need to show that if A is commutative affine over a field F, then $\mathrm{Jac}(A)$ is nilpotent. This classical case is reviewed in §1.3.1.

2. Next is the *finite* case: If A is affine over a field F and a finite module over an affine central subalgebra, then $\mathrm{Jac}(A)$ is nilpotent. This case was known well before Razmyslov's Theorem, and is reviewed in §1.3.2. Theorem 1.3 follows whenever A is a subring of a finite dimensional algebra over a field.

3. Let $\mathcal{CAP}_n = T(\mathrm{Cap}_n)$ be the T-ideal generated by Cap_n, and let $\mathcal{CAP}_n(A) \subseteq A$ be the ideal generated in A by the evaluations of Cap_n on A, so $A/\mathcal{CAP}_n(A)$ satisfies Cap_n. Therefore, by induction on n, $\mathrm{Jac}(A/\mathcal{CAP}_n(A))$ is nilpotent. Hence there exists q such that

$$\mathrm{Jac}(A)^q \subseteq \mathcal{CAP}_n(A), \quad \text{so} \quad \mathrm{Jac}(A)^{2q} \subseteq (\mathcal{CAP}_n(A))^2.$$

4. In §2.2.4 we work over an arbitrary base ring C (which need not even be Noetherian), and for any algebra A introduce the ideal $I_{n,A} \subset A[\xi_{n,A}]$, for commuting indeterminates $\xi_{n,A}$, which provides "generic" integrality relations for elements of A. Let $\widehat{C\{x,y,t\}} := C\{x,y,t\}/\mathcal{CAP}_{n+1}$, the relatively free algebra of Cap_{n+1}. Taking the "doubly alternating" polynomial

$$f = t_1 \,\mathrm{Cap}_n(x_1,\dots,x_n) t_2 \,\mathrm{Cap}_n(y_1,\dots,y_n) t_3,$$

we construct, in Section 2.2.1, the key $\widehat{C\{t\}}$-module $\widehat{\mathcal{M}} \subset \widehat{C\{x,y,t\}}$, which contains the polynomial $\hat{f}$. A combinatoric argument given in Proposition 2.17 applied to $\widehat{C\{x,y,t\}}$ (together with substitutions) shows that $I_{n,\widehat{C\{x,y,t\}}} \cdot \widehat{\mathcal{M}} = 0$.

5. We introduce the **obstruction to integrality** $\mathrm{Obst}_n(A) = A \cap I_{n,A} \subset A$ and show that $A/\mathrm{Obst}_n(A)$ can be embedded into a finite algebra over an affine central F-subalgebra; hence $\mathrm{Jac}(A/\mathrm{Obst}_n(A))$ is nilpotent. This implies that there exists m such that

$$\mathrm{Jac}(A)^m \subseteq \mathrm{Obst}_n(A).$$

The proof of this step applies Shirshov's Height Theorem [32], [7, Theorem 2.3].

6. We prove that $\mathrm{Obst}_n(A) \cdot (\mathcal{CAP}_n(A))^2 = 0$ over an arbitrary ring C. This is obtained from Step 4 via a sophisticated specialization argument involving **free products**.

7. We put the pieces together. When C is a field, Step 3 shows that $\mathrm{Jac}(A)^q \subseteq \mathcal{CAP}_n(A)$ for some q, and Step 5 shows that $\mathrm{Jac}(A)^m \subseteq \mathrm{Obst}_n(A)$ for some m. Hence

$$\mathrm{Jac}(A)^{2q+m} \subseteq \mathrm{Obst}_n(A) \cdot (\mathcal{CAP}_n(A))^2 = 0,$$

which completes the proof of Theorem 1.2. When C is Noetherian, any nil ideal N of C is nilpotent, so the analogous argument shows that $N^m \subseteq \mathrm{Obst}_n(A)$ for some m. Hence

$$N^{2q+m} \subseteq \mathrm{Obst}_n(A) \cdot (\mathcal{CAP}_n(A))^2 = 0,$$

proving Theorem 1.3.

1.3. Special cases

We need some classical special cases.

1.3.1. The commutative case

Our main objective is to prove that the Jacobson radical $\mathrm{Jac}(A)$ of an affine PI-algebra A (over a field) is nilpotent. We start with the classical case for which A is commutative.

Remark 1.10. Any commutative affine algebra A over a Noetherian base ring C is Noetherian, by Hilbert's Basis Theorem, and hence the intersection of its prime ideals is nilpotent, cf. [29, Theorem 16.24].

But for any ideal $I \triangleleft A$, the algebra A/I is also Noetherian, so the intersection of the prime ideals of A containing I is nilpotent modulo I.

Proposition 1.11. If H is a commutative affine algebra over a field, then $\mathrm{Jac}(H)$ is nilpotent.

Proof. The "weak Nullstellensatz" [28, Proposition 6.37] says that H is Jacobson, and thus the Jacobson radical $\mathrm{Jac}(H)$ is contained in the intersection of the prime ideals of H. But any commutative affine algebra is Noetherian, so we conclude with Remark 1.10. □

1.3.2. The finite case

To extend this to noncommutative algebras, we start with some other classical results:

1. [29, Theorem 15.23] (Wedderburn) Any nil subring of an $n \times n$ matrix algebra over a field is nilpotent, of nilpotence index $\leqslant n$ (in view of [29, Lemma 15.22]).
2. [29, Theorem 15.18] (Jacobson) The Jacobson radical of an n-dimensional algebra over a field is nilpotent, and thus has nilpotence index $\leqslant n$, by (1).
3. Any algebra finite over a Noetherian central subring C, is Noetherian (This follows at once from induction applied to [28, Proposition 7.5].

Theorem 1. *Suppose $A = C\{a_1, \ldots, a_\ell\}$ is an affine algebra over a commutative Noetherian ring C, with $A \subseteq M_n(K)$ for a suitable commutative C-algebra K. Then*

1. *Any nil subalgebra N of A is nilpotent, of bounded nilpotence index $\leqslant mn$, where m is given in the proof. When K is reduced, i.e., without nonzero nilpotent elements, then $m = 1$, so $N^n = 0$.*
2. *If C is a field, then $\mathrm{Jac}(A)$ is nilpotent.*

Proof. For each $1 \leqslant k \leqslant \ell$, write each a_k as an $n \times n$ matrix $(a_{ij}^{(k)})$, for $a_{ij}^{(k)} \in K$, and let H be the commutative C-subalgebra of K generated by these finitely many $a_{ij}^{(k)}$; then H is C-affine. We can view each a_k in $M_n(H)$, so $A \subseteq M_n(H)$.

(1) Let $N \subseteq A$ be a nil subalgebra. Now $A \subseteq M_n(H)$, so $N \subseteq M_n(H)$ and is nil. Let $P \subseteq H$ be prime. The homomorphism $H \to H/P$ extends to

$$M_n(H) \to M_n(H/P) \ \ (\cong M_n(H)/M_n(P)).$$

Let $\bar{N}$ be the image of N, so $\bar{N} = (N + M_n(P))/M_n(P)$ so $\bar{N} \subseteq M_n(H)/M_n(P) \cong M_n(H/P) \subseteq M_n(L)$ where L is the field of fractions of the domain H/P. By Wedderburn's theorem $\bar{N}^n = 0$ which implies that $N^n \subseteq M_n(P)$ (since $P = 0$ in H/P and in L). Hence, letting U denote the prime radical of H, we have $N^n \subseteq M_n(U)$. But, in view of Remark 1.10, we have $U^m = 0$ for some m. (If K is reduced then $U = 0$, implying $m = 1$.) We conclude that

$$N^{mn} = (N^n)^m \subseteq (M_n(U))^m = M_n(U^m) = 0.$$

(2) We need here the well-known fact [29, Exercise 15.28] that when $J \triangleleft A$, with J nilpotent, then $\mathrm{Jac}(A/J) = \mathrm{Jac}(A)/J$. It follows at once that if $\mathrm{Jac}(A/J)$ is nilpotent, then $\mathrm{Jac}(A)$ is nilpotent.

By hypothesis H is affine over the field C, so $\mathrm{Jac}(H)$ is nilpotent, and thus $M_n(\mathrm{Jac}(H)) = \mathrm{Jac}(M_n(H))$ is nilpotent. Let $\tilde{A} = A/(A \cap M_n(\mathrm{Jac}(H)))$ and $\tilde{H} = H/\mathrm{Jac}(H)$. Then

$$\tilde{A} \subseteq M_n(H/\mathrm{Jac}(H)) = M_n(\tilde{H}),$$

and $\mathrm{Jac}(\tilde{H}) = 0$. Thus we may assume that $\mathrm{Jac}(H) = 0$, and we shall prove that $J^n = 0$, where $J = \mathrm{Jac}(A)$.

For any maximal ideal P of H, we see that H/P is an affine field extension of C, and thus is finite dimensional over C, by [28, Theorem 5.11]. But then the image of A in $M_n(H/P)$ is finite dimensional over C, so the image $\bar{J}$ of J is nilpotent, implying $\bar{J}^n = 0$. Hence J^n is contained in $\cap M_n(P) = M_n(\cap P) = 0$, where P runs over the maximal ideals of H. □

Theorem 2. *Suppose A is an algebra that is finite over C, itself an affine algebra over a field. Then* $\mathrm{Jac}(A)$ *is nilpotent.*

Proof. Since A is Noetherian, its nilradical N is nilpotent by [29, Remark 16.30(ii)], so modding out N we may assume that A is semiprime, and thus the subdirect product of prime algebras $\{A_i = A/P_i : i \in I\}$ finite over their centers. If $\mathrm{Jac}(A)^n \subset P_i$ for each $i \in I$, then $\mathrm{Jac}(A)^n \subset \cap P_i = 0$.

So we may assume that A is prime. But localizing over the center, we may assume that C is a field. Let $n = \dim_C A$. Then A is embedded via the regular representation into $n \times n$ matrices over a field, and we are done by Theorem 1. □

Since not every affine PI-algebra might satisfy the hypotheses of Theorem 1, cf. [33] and [18], we must proceed further.

2. Proof of Razmyslov's Theorem

In this section we give full details for Zubrilin's proof of Theorem 1.2.

2.1. Zubrilin's approach

2.1.1. The operator $\delta_{k,z}^{(x,n)}$

Let us fix notation for the next few sections. C is an arbitrary commutative ring. We start with a polynomial $f := f(x_1, \dots, x_n) \in C\{x, y, t\}$ in $\vec{x} = \{x_1, \dots, x_n\}$ (which we always notate), as well as possibly $\vec{y} = \{y_1, \dots, y_n\}$ (which we sometimes notate), and $\vec{t} = \{t_1, \dots\}$, all of which are noncommutative indeterminates.

Definition 2.1. Let $f(\vec{x}, \vec{y}, \vec{t})$ be multilinear in the x_i (and perhaps involving additional indeterminates $\vec{y}$ and $\vec{t}$). Take $0 \leqslant k \leqslant n$, and expand

$$f^* = f((z+1)x_1, \dots, (z+1)x_n, \vec{y}, \vec{t}),$$

where z is a new noncommutative indeterminate. Then we write

$$\delta_{k,z}^{(x,n)}(f) := \delta_{k,z}^{(x,n)}(f)(x_1, \dots, x_n, z)$$

for the homogeneous component of f^* of degree k in the noncommutative indeterminate z. (We have suppressed $\vec{y}$, $\vec{t}$ in the notation, as indicated above.)

For example let $n = 2$ and $f = x_1x_2$. Then

$$(z+1)x_1(z+1)x_2 = zx_1zx_2 + zx_1x_2 + x_1zx_2 + x_1x_2.$$

Hence $\delta^{(x,2)}_{0,z}(x_1x_2) = x_1x_2$, $\delta^{(x,2)}_{1,z}(x_1x_2) = zx_1x_2 + x_1zx_2$, and $\delta^{(x,2)}_{2,z}(x_1x_2) = zx_1zx_2$.

More generally, for any $h \in C\{t\}$ we write $\delta^{(x,n)}_{k,h}(f) := \delta^{(x,n)}_{k,h}(f)(x_1, \dots, x_n, h)$, i.e., the specialization of $\delta^{(x,n)}_{k,z}(f)$ under $z \mapsto h$.

Remark 1.

1. In calculating $\delta^{(x,n)}_{k,z}(f)$, the substitution $x_i \to (z+1)x_i$ is applied to the first n positions in f but not to the other positions. For example, the last (i.e. $n+1$ st) variable in $f(x_1, \dots, x_{n-1}, x_{n+1}, x_n)$ is x_n, not x_{n+1}. Hence, to calculate $\delta^{(x,n)}_{k,z}(f(x_1, \dots, x_{n-1}, x_{n+1}, x_n))$ we apply $x_i \to (z+1)x_i$ to all x_i's except x_n.

2. We can also write

$$\delta^{(x,n)}_{k,z}(f(x_1, \dots, x_n, \vec{t})) = \sum_{1 \leqslant i_1 < \dots < i_k \leqslant n} f(x_1, \dots, x_n, \vec{t}) \,|_{x_{i_j} \to zx_{i_j}} =$$

$$= \sum_{1 \leqslant i_1 < \dots < i_k \leqslant n} f(x_1, \dots, zx_{i_1}, \dots zx_{i_k}, \dots, x_n, \vec{t}).$$

3. In case $f = f(x_1, \dots, x_n, y_1, \dots, y_n)$ also involves indeterminates $y_1, \dots, y_n$, we still have

$$\delta^{(x,n)}_{k,z}(f) = \sum_{1 \leqslant i_1 < \dots < i_k \leqslant n} f \,|_{x_{i_j} \to zx_{i_j}},$$

indicating that the other indeterminates $y_1, \dots, y_n$ remain fixed. Analogously,

$$\delta^{(y,n)}_{k,z}(f) = \sum_{1 \leqslant i_1 < \dots < i_k \leqslant n} f \,|_{y_{i_j} \to zy_{i_j}},$$

and the indeterminates $x_1, \dots, x_n$ are fixed.

Definition 2.2. A polynomial $f(x_1, \dots, x_n, \vec{t}\,)$ is **alternating** in $x_1, \dots, x_n$ if f is multilinear in the x_i and

$$f(x_1, \dots, x_i, \dots, x_j, \dots, x_n, \vec{t}\,) + f(x_1, \dots, x_j, \dots, x_i, \dots, x_n, \vec{t}\,) = 0 \text{ for all } i < j. \tag{1}$$

A stronger definition, which would suffice for our purposes, is to require that

$$f(x_1, \dots, x_i, \dots, x_i, \dots, x_n, \vec{t}\,) = 0; \tag{2}$$

i.e., we get 0 when specializing x_j to x_i for any $1 \leqslant i < j \leqslant n$. We get (2.2) by linearizing (2), and can recover (2) from (2.2) in characteristic $\neq 2$.)

Lemma 2.3. Let $f(x_1, \dots, x_n, \vec{t}\ )$ be multilinear and alternating in $x_1, \dots, x_n$. Then for each $0 \leqslant k \leqslant n$, $\delta^{(x,n)}_{k,z}(f(x_1, \dots, x_n, \vec{t}\,))$ is also alternating in $x_1, \dots, x_n$.

Proof. Let $v = 1 + \varepsilon z$ where ε is a central indeterminate. Obviously $f(vx_1, \ldots, vx_n, \vec{t}\,)$ is also alternating in $x_1, \ldots, x_n$. Since

$$f(vx_1, \ldots, vx_n, \vec{t}\,) = \sum_{k=0}^{n} \left(\delta_{k,z}^{(x,n)}(f(x_1, \ldots, x_n, \vec{t}\,))\right) \cdot \varepsilon^k$$

is alternating in $x_1, \ldots, x_n$, it follows that each $\delta_{k,z}^{(x,n)}(f(x_1, \ldots, x_n, \vec{t})$ is alternating in $x_1, \ldots, x_n$. □

Remark 2.4.

1. Since $\mathcal{CAP}_n$ is generated as a T-ideal by polynomials alternating in $x_1, \ldots, x_n$, we have

$$\delta_{k,z}^{(x,n)}(\mathcal{CAP}_n) \subseteq \mathcal{CAP}_n \qquad \text{and} \qquad \delta_{k,z}^{(x,n)}(\mathcal{CAP}_{n+1}) \subseteq \mathcal{CAP}_{n+1}.$$

2. The results proved for the indeterminate z specialize to an arbitrary polynomial h, and thus can be formulated for h.

Lemma 2.5. The $\delta_{k,z}^{(x,n)}(f)$-operator is functorial, in the sense that if $\vec{a} = (a_1, \ldots, a_m) \in A$ and $h(\vec{a}) = h'(\vec{a})$, then $\delta_{k,h}^{(x,n)}(f)(\vec{a}) = \delta_{k,h'}^{(x,n)}(f)(\vec{a})$.

Proof. We get the same result in Definition 2.1 by specializing z to h and then to $\vec{a}$, as we get by specializing z to h', and then to $\vec{a}$. □

This observation is needed in our later specialization arguments.

The following observation, which is rather well known, motivates Proposition 2.10 below. Let $V = Cx_1 \oplus \cdots \oplus Cx_n$ and let $z : V \to V$ be a linear transformation from V to V. Let

$$\det(\lambda I - z) = \sum_{k=0}^{n} c_k(z)\lambda^k$$

be the characteristic ("Cayley-Hamilton") polynomial of z. Then we have the following formula from [26, Theorem 1.4.12]:

$$\delta_{k,z}^{(x,n)}(\mathrm{Cap}_n(x_1, \ldots, x_n; \vec{y})) = \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} \mathrm{Cap}_n(x_1, \ldots, zx_{i_1}, \ldots, zx_{i_k}, \ldots, x_n; \vec{y}) =$$
$$= c_k(z) \cdot \mathrm{Cap}_n(x_1, \ldots, x_n; \vec{y}),$$

and the coefficients $c_k(z)$ are independent of the particular indeterminates $x_1, \ldots, x_n$. Proposition 2.10 below displays a similar phenomenon.

2.2. Zubrilin's Proposition

Our goal in this section is Proposition 2.17. Let us define the terms used there.

Let $\widehat{C\{x, y, t\}}$ denote the relatively free algebra $C\{x, y, t\}/\mathcal{CAP}_{n+1}$. We denote the image of $f \in C\{x, y, t\}$ in $\widehat{C\{x, y, t\}}$ by $\hat{f}$.

Remark 2.6. If A satisfies Cap_{n+1}, then any algebra homomorphism $\varphi : C\{x, y, t\} \to A$ naturally induces an algebra homomorphism $\widehat{\varphi} : \widehat{C\{x, y, t\}} \to A$ given by

$$\widehat{\varphi}(\hat{f}) = \varphi(f),$$

since $\mathcal{CAP}_{n+1} \subseteq \ker \varphi$.

Remark 2. *Let $f(x_1,\dots,x_{n+1})$ be multilinear in $x_1,\dots,x_{n+1}$ and alternating in $x_1,\dots,x_n$. Construct*

$$\tilde f=\tilde f(x_1,\dots,x_{n+1})=\sum_{k=1}^{n+1}(-1)^{k-1}f(x_1,\dots,x_{k-1},x_{k+1},\dots,x_{n+1},x_k). \tag{3}$$

(All other variables occurring in f are left untouched.)

Then $\tilde f$ is $(n+1)$-alternating in $x_1,\dots,x_{n+1}$.

Proposition 2.7. Let $f(x_1,\dots,x_n,x_{n+1})$ be multilinear in $x_1,\dots,x_n,x_{n+1}$ and alternating in $x_1,\dots,x_n$ (so $\tilde f$ of Equation (3) is $(n+1)$-alternating). Then

$$\sum_{j=0}^{n}(-1)^j\delta_{j,z}^{(x,n)}(f(x_1,\dots,x_n,z^{n-j}x_{n+1}))\equiv 0 \quad \text{modulo } \mathcal{CAP}_{n+1}.$$

Proof. Throughout we work modulo $\mathcal{CAP}_{n+1}$. Since $\tilde f$ is $(n+1)$-alternating, we have

$$0\equiv\tilde f=f(x_2,x_3\dots,x_{n+1},x_1)-f(x_1,x_3\dots,x_{n+1},x_2)+\dots+(-1)^nf(x_1,x_2,\dots,x_n,x_{n+1}).$$

Thus, modulo $\mathcal{CAP}_{n+1}$ the last summand $(-1)^nf(x_1,x_2,\dots,x_n,x_{n+1})$ can be replaced by minus the sum of the other summands:

$$(-1)^nf(x_1,x_2,\dots,x_n,x_{n+1})\equiv\sum_{k=1}^{n}(-1)^kf(x_1,\dots,x_{k-1},x_{k+1},\dots,x_n,x_{n+1},x_k),$$

Given $0\leqslant j\leqslant n$, substitute $x_{n+1}\mapsto z^{n-j}x_{n+1}$, so

$$(-1)^nf(x_1,x_2,\dots,x_n,z^{n-j}x_{n+1})\equiv\sum_{k=1}^{n}(-1)^kf(x_1,\dots,x_{k-1},x_{k+1},\dots .x_n,z^{n-j}x_{n+1},x_k).$$

Applying $\delta_{j,z}^{(x,n)}$ and summing with sign, we get

$$(-1)^n\sum_{j=0}^{n}(-1)^j\delta_{j,z}^{(x,n)}(f(x_1,\dots,x_n,z^{n-j}x_{n+1})\equiv$$

$$\equiv\sum_{j=0}^{n}(-1)^j\sum_{k=1}^{n}(-1)^k\delta_{j,z}^{(x,n)}(f(x_1,\dots,x_{k-1},x_{k+1},\dots,x_n,z^{n-j}x_{n+1},x_k))=$$

$$=\sum_{k=1}^{n}(-1)^k\sum_{j=0}^{n}(-1)^j\delta_{j,z}^{(x,n)}(f(x_1,\dots,x_{k-1},x_{k+1},\dots,x_n,z^{n-j}x_{n+1},x_k)).$$

Denote $g_{j,k}=f(x_1,\dots,x_{k-1},x_{k+1},\dots,x_n,z^{n-j}x_{n+1},x_k)$, and

$$Q_k=\sum_{j=0}^{n}(-1)^j\delta_{j,z}^{(x,n)}(g_{j,k}). \tag{4}$$

It suffices to show that $Q_k\equiv 0$ for each k. Note that in calculating $\delta_{j,z}^{(x,n)}(g_{j,k})=\delta_{j,z}^{(x,n)}(f(x_1,\dots,x_{k-1},x_{k+1},\dots,x_n,z^{n-j}x_{n+1},x_k))$, x_k is unchanged (since it is the last indeterminate), while for all other x_i's (in particular – for x_{n+1}) we substitute $x_i\mapsto(z+1)x_i$, cf. Remark 1.1. Therefore

$$\delta_{j,z}^{(x,n)}(g_{j,k})=\delta_{j,z,[k']}^{(x,n)}(g_{j,k})+\delta_{j,z,[k'']}^{(x,n)}(g_{j,k})$$

where

$\delta^{(x,n)}_{j,z,[k']}(g_{j,k})$ is the sum of the monomials of $\delta^{(x,n)}_{j,z}(g_{j,k})$ having z-degree j, where x_{n+1} was replaced by zx_{n+1};

and

$\delta^{(x,n)}_{j,z,[k'']}(g_{j,k})$ is the sum of the monomials of $\delta^{(x,n)}_{j,z}(g_{j,k})$ having z-degree j, where x_{n+1} was unchanged.

It is not difficult to see that for $j > 0$,

$$\delta^{(x,n)}_{j,z,[k']} f(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n, z^{n-j}x_{n+1}, x_k) =$$

$$\delta^{(x,n)}_{j-1,z,[k'']} f(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n, z^{n-j+1}x_{n+1}, x_k),$$

namely

$$\delta^{(x,n)}_{j,z,[k']}(g_{j,k}) = \delta^{(x,n)}_{j-1,z,[k'']}(g_{j-1,k}).$$

It also follows from the definitions that $\delta^{(x,n)}_{0,z,[k']}(g_{0,k}) = \delta^{(x,n)}_{n,z,[k'']}(g_{n,k}) = 0$. Hence

$$\sum_{j=0}^{n}(-1)^j \delta^{(x,n)}_{j,z,[k']}(g_{j,k}) = \sum_{j=1}^{n}(-1)^j \delta^{(x,n)}_{j,z,[k']}(g_{j,k}) =$$

$$= \sum_{j=1}^{n}(-1)^j \delta^{(x,n)}_{j-1,z,[k'']}(g_{j-1,k}) = -\sum_{j=0}^{n-1}(-1)^j \delta^{(x,n)}_{j,z,[k'']}(g_{j,k}),$$

and

$$\sum_{j=0}^{n}(-1)^j \delta^{(x,n)}_{j,z,[k'']}(g_{j,k}) = \sum_{j=0}^{n-1}(-1)^j \delta^{(x,n)}_{j,z,[k'']}(g_{j,k}).$$

Summing in (4) we get

$$Q_k = \sum_{j=0}^{n}(-1)^j \delta^{(x,n)}_{j,z}(g_{j,k}) \equiv \sum_{j=0}^{n}(-1)^j \left(\delta^{(x,n)}_{j,z,[k']} g_{j,k} + \delta^{(x,n)}_{j,z,[k'']} g_{j,k}\right) \equiv 0.$$

□

2.2.1. The module $\widehat{\mathcal{M}}$ over the relatively free algebra of Cap_{n+1}

We need a special sort of alternating polynomials.

Definition 2.8. A polynomial $f(x_1, \ldots, x_n; y_1, \ldots, y_n; \vec{t})$, where $\vec{t}$ denotes other possible indeterminates, is **doubly alternating** if f is linear and alternating in $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$.

Our main example is the **double Capelli** polynomial

$$\mathrm{DCap}_n = t_1 \,\mathrm{Cap}_n(x_1, \ldots, x_n; \vec{t}) t_2 \,\mathrm{Cap}_n(y_1, \ldots, y_n; \vec{t'}) t_3. \tag{5}$$

Here $\vec{t}$ and $\vec{t'}$ are arbitrary sets of extra indeterminates. We suppress the indeterminates $\vec{t}$, $\vec{t'}$, and t_1, t_2, t_3 from the notation, since we do not alter them.

Definition 2.9. Let $\mathcal{M}$ denote the C-submodule of $C\{x, y, t\}$ consisting of all doubly alternating polynomials (in $x_1 \ldots, x_n$, and in $y_1, \ldots, y_n$).

$\widehat{\mathcal{M}}$ denotes the image of $\mathcal{M}$ in $C\widehat{\{x, y, t\}}$, i.e., the C-submodule of $C\widehat{\{x, y, t\}}$ consisting of the images of all doubly alternating polynomials (in $x_1 \ldots, x_n$, and in $y_1, \ldots, y_n$).

Remark 3. *$\widehat{\mathcal{M}}$ is a $\widehat{C\{t\}}$-submodule of $C\widehat{\{x, y, t\}}$, namely $\widehat{C\{t\}}\widehat{\mathcal{M}} \subseteq \widehat{\mathcal{M}}$. Indeed, let $h \in C\{t\}$ and $f \in \mathcal{M}$. If either h or f is in $\mathcal{CAP}_{n+1}$ then $hf \in \mathcal{CAP}_{n+1}$; hence the product $\hat{h}\hat{f} = \widehat{hf}$ is well defined. Moreover, if $f = f(x_1, \ldots, x_n, y_1, \ldots, y_n, \vec{t}\,)$ is doubly alternating in the x's and in the y's, and $h \in C\{\vec{t}\,\}$, then hf is doubly alternating in the x's and in the y's.*

2.2.2. The Zubrilin action

The theory hinges on the following amazing result, which we prove in Section 2.7 below. (This is also proved in [7, Theorem 4.82], but more details are given here.)

Proposition 2.10. Let $f(x_1, \ldots, x_n; y_1, \ldots, y_n)$ be doubly alternating in $x_1, \ldots, x_n$ and in $y_1, \ldots, y_n$ (perhaps involving additional indeterminates). Then for any polynomial h,

$$\delta_{k,h}^{(x,n)}(f) \equiv \delta_{k,h}^{(y,n)}(f) \quad \textit{modulo } \mathcal{CAP}_{n+1}; \tag{6}$$

namely,

$$\sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} f\,|_{x_{i_j} \to h x_{i_j}} \equiv \sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} f\,|_{y_{i_j} \to h y_{i_j}} \quad \textit{modulo } \mathcal{CAP}_{n+1}.$$

Before proving Proposition 2.10 we deduce some of its consequences.

Remark 2.11. It follows from Proposition 2.10 that $\delta_{k,h}^{(x,n)}(f) - \delta_{k,h}^{(y,n)}(f) \in \mathcal{CAP}_{n+1}$ whenever $f \in \widehat{\mathcal{M}}$, so working modulo $\mathcal{CAP}_{n+1}$ we can suppress x in the notation, writing $\hat{\delta}_{k,h}^{(n)}(\hat{f})$ for $\hat{\delta}_{k,h}^{(x,n)}(\hat{f})$.

2.2.3. Commutativity of the operators $\delta_{k,h_j}^{(n)}$ modulo $\mathcal{CAP}_{n+1}$

We use $\widehat{\mathcal{M}}$ instead of $\mathcal{M}$ because of the following lemma.

Lemma 2.12. (i) $\delta_{k,h}^{(n)}$ induces a well-defined map $\hat{\delta}_{k,h}^{(n)} : \widehat{\mathcal{M}} \to \widehat{\mathcal{M}}$ given by $\hat{\delta}_{k,h}^{(n)}(\hat{f}) = \widehat{\delta_{k,h}^{(x,n)}(f)}$.

(ii) $\hat{\delta}_{k,h}^{(n)}$ produces the same result using the indeterminates x or y.

Proof. (i) If $f(x_1, \ldots, x_n, y_1, \ldots, y_n)$ and $g(x_1, \ldots, x_n, y_1, \ldots, y_n)$ are doubly alternating polynomials, with $\hat{f} = \hat{g}$, then $f - g \in \mathcal{CAP}_{n+1}$, so by Remark 2.4(1), $\delta_{k,h}^{(n)}(f - g) \in \mathcal{CAP}_{n+1}$ and hence $\widehat{\delta_{k,h}^{(n)}(f - g)} = 0$. Therefore we have

$$0 = \widehat{\delta_{k,h}^{(n)}(f - g)} = \widehat{\delta_{k,h}^{(n)}(f)} - \widehat{\delta_{k,h}^{(n)}(g)} = \hat{\delta}_{k,h}^{(n)}(\hat{f}) - \hat{\delta}_{k,h}^{(n)}(\hat{g}),$$

proving that $\hat{\delta}_{k,h}^{(n)}$ is well-defined.

(ii) The assertion follows from Remark 2.11, which shows that $\widehat{\delta_{k,h}^{(x,n)}(f)} = \widehat{\delta_{k,h}^{(y,n)}(f)}$. $\square$

Lemma 2.13. Let $f = f(x_1, \ldots, x_n; y_1, \ldots, y_n)$ be doubly alternating in $x_1, \ldots, x_n$ and in $y_1, \ldots, y_n$ (and perhaps involving other indeterminates). Let $1 \leqslant k, \ell \leqslant n$. Then for any $h_1, h_2 \in C\{t\}$,

$$\hat{\delta}_{k,h_1}^{(n)} \hat{\delta}_{\ell,h_2}^{(n)}(\hat{f}) = \hat{\delta}_{\ell,h_2}^{(n)} \hat{\delta}_{k,h_1}^{(n)}(\hat{f}). \tag{7}$$

Proof. Equation (7) claims that modulo $\mathcal{CAP}_{n+1}$,

(i)
$$\delta^{(x,n)}_{k,h_1}\delta^{(x,n)}_{\ell,h_2}(f) \equiv \delta^{(x,n)}_{\ell,h_2}\delta^{(x,n)}_{k,h_1}(f) \quad \text{and}$$

(ii)
$$\delta^{(x,n)}_{k,h_1}\delta^{(y,n)}_{\ell,h_2}(f) \equiv \delta^{(y,n)}_{\ell,h_2}\delta^{(x,n)}_{k,h_1}(f) \quad \text{and}$$

(iii)
$$\delta^{(y,n)}_{k,h_1}\delta^{(y,n)}_{\ell,h_2}(f) \equiv \delta^{(y,n)}_{\ell,h_2}\delta^{(y,n)}_{k,h_1}(f).$$

The middle equivalence (ii) is an obvious equality. The first and third equivalences are similar, and we prove the first. By Proposition 2.10, by (ii), and again by Proposition 2.10, modulo $\mathcal{CAP}_{n+1}$ we can write
$$\delta^{(x,n)}_{k,h_1}\delta^{(x,n)}_{\ell,h_2}(f) \equiv \delta^{(x,n)}_{k,h_1}\delta^{(y,n)}_{\ell,h_2}(f) \equiv \delta^{(y,n)}_{\ell,h_2}\delta^{(x,n)}_{k,h_1}(f) \equiv \delta^{(x,n)}_{\ell,h_2}\delta^{(x,n)}_{k,h_1}(f).$$

Note that in the last step, Lemma 2.3 was applied (to $\delta^{(x,n)}_{k,h_1}(f)$). $\square$

2.2.4. The ideal $I_{n,A} \subset A[\xi_{n,A}]$ and the annihilator of $\hat{M}$

Definition 2.14. For each $a \in A$ let $\xi_{1,a}, \ldots, \xi_{n,a}$ be n corresponding new commuting variables, and construct $A[\xi_{n,A}] = A[\xi_{1,a}, \ldots, \xi_{n,a} \mid a \in A]$. Let $I_{n,A} \subseteq A[\xi_{n,A}]$ be the ideal generated in $A[\xi_{n,A}]$ by the elements
$$a^n + \xi_{1,a}a^{n-1} + \cdots + \xi_{n,a}, \quad a \in A,$$

namely
$$I_{n,A} = \langle a^n + \xi_{1,a}a^{n-1} + \cdots + \xi_{n,a} \mid a \in A\rangle.$$

Remark 2.15. In view of Proposition 2.10, the map $\hat{\delta}^{(n)}_{k,h} : \widehat{\mathcal{M}} \to \widehat{\mathcal{M}}$ of Lemma 2.12 yields an action of the C-algebra $\widehat{C\{t\}}[\xi_{n,\widehat{C\{t\}}}]$ on $\widehat{\mathcal{M}}$, given by $\xi_{k,h}f = \delta^{(n)}_{k,h}(f)$.

Working with the relatively free algebra, our next goal is to prove that $I_{n,\widehat{C\{t\}}} \cdot \widehat{\mathcal{M}} = 0$. For that we shall need the next result.

Proposition 2.16. (Zubrilin) Assume that a multilinear polynomial $g(x_1, \ldots, x_n)$ is alternating in $x_1, \ldots, x_n$. Then, modulo $\mathcal{CAP}_{n+1}$,
$$\sum_{k=0}^{n}(-1)^k h^{n-k}\delta^{(n)}_{k,h}(g) \equiv 0$$

for any $h \in C\{t\}$. In particular, if g is doubly alternating, then (again modulo $\mathcal{CAP}_{n+1}$)
$$\sum_{k=0}^{n}(-1)^k h^{n-k}\delta^{(n)}_{k,h}(g) \equiv 0.$$

Proof. First we take h to be an indeterminate z. Let $f(x_1, \ldots, x_{n+1}) = x_{n+1}g(x_1, \ldots, x_n)$. By Proposition 2.7, $\widehat{C\{x, y, t\}}$ satisfies the identity
$$\sum_{j=0}^{n}(-1)^j\delta^{(n)}_{j,z}(f(x_1, \ldots, x_n, z^{n-j}x_{n+1})) \equiv 0.$$

Note that in computing $\delta_{j,z}^{(n)}(f(x_1,\dots,x_n,z^{n-j}x_{n+1}))$, the last indeterminate is x_{n+1} and is unchanged, cf. Remark 1.1, so

$$\delta_{j,z}^{(n)}(f(x_1,\dots,x_n,z^{n-j}x_{n+1})) = z^{n-j}x_{n+1}\delta_{j,z}^{(n)}g(x_1,\dots,x_n).$$

Using Proposition 2.7, we have that

$$\sum_{j=0}^{n}(-1)^j z^{n-j}x_{n+1}\delta_{j,z}^{(n)}(g(x_1,\dots,x_n)) \in \mathcal{CAP}_{n+1}.$$

The proof now follows by substituting $x_{n+1}\mapsto 1$ and $z\to h\in C\{t\}$.

□

As a consequence we can now prove the key result:

PROPOSITION 2.17. Let $\widehat{\mathcal{M}}$ be the module given by Definition 2.9. Then, $I_{n,\widehat{C\{t\}}}\cdot\widehat{\mathcal{M}} = 0$.

PROOF. We prove that $I_{n,\widehat{C\{t\}}}\cdot\widehat{\mathcal{M}} = 0$, by showing for any doubly alternating polynomial $f(x_1,\dots,x_n,y_1,\dots,y_n)\in\widehat{\mathcal{M}}$ and $h\in\widehat{C\{t\}}$, that

$$(h^n+\xi_{1,h}h^{n-1}+\dots+\xi_{n,h})f \equiv 0 \pmod{\mathcal{CAP}_{n+1}}.$$

It follows from the action $\xi_{k,h}f = \delta_{k,h}^{(n)}(f)$ and from Proposition 2.16 that modulo $\mathcal{CAP}_{n+1}$,

$$(h^n+\xi_{1,h}h^{n-1}+\dots+\xi_{n,h})f = \sum_{k=0}^{n}(-1)^k h^{n-k}\delta_{k,h}^{(n)}(f) \equiv 0.$$

□

2.3. The ideal $\mathrm{Obst}_n(A)\subseteq A$

In order to utilize these results about integrality, we need another concept. We define $\mathrm{Obst}_n(A) = A\cap I_{n,A}$, viewing $A\subset A[\xi_{n,A}]$.

REMARK 4.

1. *Let*

$$\bar{A} = A[\xi_{n,A}]/I_{n,A}, \tag{8}$$

with $f: A[\xi_{n,A}]\to\bar{A}$ the natural homomorphism, and $f_r: A\to\bar{A}$ be the restriction of f to A. Then

$$\ker(f_r) = A\cap I_{n,A} = \mathrm{Obst}_n(A).$$

2. *Note that for every $a\in A$, $f(a)$ is n-integral (i.e., integral of degree n) over $C[\bar{\xi}_{i,A}]$, and thus over the center of $\bar{A}$. Indeed, apply the homomorphism f to the element*

$$a^n+\xi_{1,a}a^{n-1}+\dots+\xi_{n,a} \quad (\in I_{n,A})$$

to get

$$\bar{a}^n+\bar{\xi}_{1,a}\bar{a}^{n-1}+\dots+\bar{\xi}_{n,a} = (a^n+\xi_{1,a}a^{n-1}+\dots+\xi_{n,a})+I_{n,A} = 0.$$

LEMMA 2.18. $\ker(f_r)$ also is the intersection of all kernels $\ker(g)$ of the following maps g:

$g: A\to B$, where B is a C-algebra, and $g: A\to B$ is a homomorphism such that for any $a\in A$, $g(a)$ is n-integral over the center of B.

Proof. Denote the above intersection $\cap_g \ker(g)$ as $\mathrm{Obst}'_n(A)$. Then $\mathrm{Obst}'_n(A) \subseteq \mathrm{Obst}_n(A)$ since $ker(f_r)$ is among these $\ker(g)$. To show the opposite inclusion we prove

Claim: For such $g : A \to B$, $\ker(g) \supseteq A \cap I_{n,A} = \mathrm{Obst}_n(A)$.

Extend g to $g^* : A[\xi_{n,A}] \to B$ as follows: $g^*(a) = a$ if $a \in A$, while $g^*(\xi_{i,a}) = \beta_{i,a}$. We claim that $g^*(I_{n,A}) = 0$. Indeed, let

$$r = a^n + \xi_{1,a}a^{n-1} + \cdots + \xi_{n,a}$$

be one of the generators of $I_{n,A}$.

By assumption there exist $\beta_{1,a}, \ldots, \beta_{n,a}$ in the center of B satisfying

$$g(a)^n + \beta_{1,a}g(a)^{n-1} + \cdots + \beta_{n,a} = 0. \tag{9}$$

Hence,

$$g^*(r) = g(a)^n + \beta_{1,a}g(a)^{n-1} + \cdots + \beta_{n,a} = 0.$$

This shows that as claimed, $g^*(I_{n,A}) = 0$.

Finally, if $a \in A \cap I_{n,A}$ then $g(a) = g^*(a) = 0$. Hence $a \in \ker(g)$, so $\ker(g) \supseteq A \cap I_{n,A} = \mathrm{Obst}_n(A)$. □

Corollary 2.19. If every $a \in A$ is n-integral (over the base field), then $\mathrm{Obst}_n(A) = 0$.

Proof. The assumption implies that in the above, the identity map $id = g : A \to A$ satisfies the condition of Lemma 2.18. Hence $0 = \ker(g) \supseteq \mathrm{Obst}_n(A)$, and the proof follows. □

This corollary explains the notation $\mathrm{Obst}_n(A)$: it is the **obstruction** for each $a \in A$ to be n-integral. The next result technically is not needed, but helps to show how Obst behaves.

Lemma 2.20. $\mathrm{Obst}_{n-1}(A) \supseteq \mathrm{Obst}_n(A)$.

Proof. Represent

$$\mathrm{Obst}_{n-1}(A) = \cap_h \ker(h) \qquad \text{and} \qquad \mathrm{Obst}_n(A) = \cap_g \ker(g),$$

with the respective conditions of $n-1$ integrality and of n integrality. Take $a \in A$ and $h : A \to B$ with $h(a)$ being $n-1$ integral over the center of B. Then $h(a)$ is also n integral over the center of B. Hence every $\ker(h)$ in $\mathrm{Obst}_{n-1}(A)$ also appears in the intersection $\mathrm{Obst}_n(A) = \cap_g \ker(g)$, and the assertion follows. □

2.4. Reduction to finite modules

The reduction to finite modules is done using Shirshov's theorem.

Proposition 2.21. Let $A = C\{a_1, \ldots, a_\ell\}$ have PI degree d over the base ring C. Then the affine algebra $A/\mathrm{Obst}_n(A)$ can be embedded in an algebra which is finite over a central affine subalgebra.

Proof.

Let $B \subseteq A$ be the subset of the words in the alphabet $a_1, \ldots, a_\ell$ of length $\leqslant d$. By Shirshov's Height Theorem there exists an integer h such that the set

$$W = \{b_1^{k_1} \cdots b_h^{k_h} \mid b_i \in B, \quad \text{any } k_i \geqslant 0\}$$

spans A over the base ring C.

Similarly to $A[\xi_{n,A}]$, construct $A[\xi_{n,B}] \subseteq A[\xi_{n,A}]$:

$$A[\xi_{n,B}] = A[\xi_{1,b}, \ldots, \xi_{n,b} \mid b \in B],$$

and let $I_{n,B}$ be the ideal

$$I_{n,B} = \langle b^n + \xi_{1,b}b^{n-1} + \cdots + \xi_{n,b} \mid b \in B\rangle \triangleleft A[\xi_{n,B}]$$

Denote

$$A' = A[\xi_{n,B}]/I_{n,B}. \tag{10}$$

We show that A' is finite over an affine central subalgebra and thus is Noetherian.

Given $a \in A$, denote $a' = a + I_{n,B} \in A'$, and similarly $\xi'_{i,b} = \xi_{i,b} + I_{n,B}$. Then for every $b \in B$, b' is n-integral over $C[\xi'_{n,B}]$, where

$$C[\xi'_{n,B}] = C[\xi'_{1,b}, \ldots, \xi'_{n,b} \mid b \in B] \subseteq \operatorname{center}(A').$$

Hence the finite subset

$$W' = \{b'^{k_1}_1 \cdots b'^{k_h}_h \mid b_i \in B, \quad k_i \leqslant n-1\} \quad (\subseteq A') \tag{11}$$

spans A' over $C[\xi'_{n,B}]$. Thus A' is finite over the affine central subalgebra $C[\xi'_{n,B}] \subseteq \operatorname{center}(A')$ and thus is Noetherian.

Restricting the natural map $g : A[\xi_{n,B}] \to A' = A[\xi_{n,B}]/I_{n,B}$ to A, we have

$$g_r : A \to A' \quad (a \mapsto a' = a + I_{n,B}) \tag{12}$$

which satisfies

$$\ker(g_r) = A \cap I_{n,B} \subseteq A \cap I_{n,A} = \operatorname{Obst}_n(A). \tag{13}$$

Let

$$\tilde{A} = A/\operatorname{Obst}_n(A), \tag{14}$$

and for $a \in A$ denote $\tilde{a} = a + \operatorname{Obst}_n(A) \in \tilde{A}$. We then have the corresponding subset $\tilde{B} = \{\tilde{b} \mid b \in B\} \subseteq \tilde{A}$, as well as the set of commutative variables $\xi_{n,\tilde{B}}$ and the ideal $\tilde{I}_{n,\tilde{B}}$. Let $A^* = \tilde{A}[\xi_{n,\tilde{B}}]/\tilde{I}_{n,\tilde{B}}$.

Replacing A by $\tilde{A}$ and $\xi_{i,B}$ by $\xi_{i,\tilde{B}}$, we clearly have the natural homomorphism

$$\tilde{g} : \tilde{A}[\xi_{n,\tilde{B}}] \to \tilde{A}[\xi_{n,\tilde{B}}]/\tilde{I}_{n,\tilde{B}} := A^*,$$

with restriction

$$\tilde{g}|_{\tilde{A}} = \tilde{g}_r : \tilde{A} \to A^*.$$

Note that each $\tilde{a} \in \tilde{A}$ is n-integral over the center of $\tilde{A}$, implying, by Corollary 2.19, that $\operatorname{Obst}_n(\tilde{A}) = 0$. Then, as in (13), we deduce that

$$\ker(\tilde{g}_r) \subseteq \operatorname{Obst}_n(\tilde{A}) \ \ (= 0).$$

Hence $\tilde{g}_r$ embeds $\tilde{A} = A/\operatorname{Obst}_n(A)$ into A^*. Note that A^* is a finite algebra over the affine central subalgebra $Q \subseteq A^*$ generated by the finitely many central elements $\xi_{k,\tilde{B}} + \tilde{I}_{n,\tilde{B}}$.

Denote $b^* = \tilde{b} + I_{n,\tilde{B}}$. Then, as in (11), the finite subset

$$W^* = \{b^{*k_1}_1 \cdots b^{*k_h}_h \mid b_i \in B, \ \ k_i \leqslant n-1\} \quad (\subseteq A^*)$$

spans A^* over Q. $\square$

2.5. Proving that $\mathrm{Obst}_n(A) \cdot (\mathcal{CAP}_n(A))^2 = 0$

In this section we show how Proposition 2.17 implies that $\mathrm{Obst}_n(A) \cdot (\mathcal{CAP}_n(A))^2 = 0$, thereby completing the proof of Razmyslov's Theorem. For this, we need to specialize down to given algebra A, requiring a new construction, the **relatively free product**, which enables us to handle A together with polynomials. Since, to our knowledge, this crucial step, which is needed one way or another in every published proof of the BKR theorem, has not yet appeared in print in full detail, we present two proofs, one faster but more ad hoc (since we intersect with A and bypass certain difficulties), and the second more structural.

Both approaches are taken in the context of varieties in universal algebra, by taking the free product of A with the free associative algebra, and then modding out the identities defining its variety.

2.5.1. The relatively free product

Definition 2.22. The **free product** $A *_C B$ of C-algebras A and B is their coproduct in the category of algebras.

(For C-algebras with 1, there are canonical C-module maps

$$A \to A \otimes 1 \subset A \otimes_C B, \qquad B \to 1 \otimes B \subset A \otimes_C B,$$

viewed naturally as C-modules, so $A *_C B$ can be identified with the tensor algebra of $A \otimes_C B$, as reviewed in [29, Example 18.38].)

Although the results through Theorem 2.32 hold over any commutative base ring C, it is easier to visualize the situation for algebras over a field F, in which case we have an explicit description of $A[\xi_{n,A}] * F\{x;y;t\}$:

Fix a base $\mathcal{B}_A = \{1\} \cup \mathcal{B}_0$ of A over F, and let $\mathcal{B}$ be the monomials in the $\{\xi_{n,a} : a \in A\}$ with coefficients in $\mathcal{B}_A$. (For algebras without 1, we take $\mathcal{B} = \mathcal{B}_0$.) Thus $\mathcal{B}$ is an F-base of $A[\xi_{n,A}]$, and $A[\xi_{n,A}] * F\{x;y;t\}$ is the vector space having base comprised of all elements of the form $b_0 h_1 b_1 h_2 b_2 \cdots h_m b_m$ where $m \geqslant 0$, $b_0, b_m \in \mathcal{B}$, $b_1, \dots, b_{m-1} \in \mathcal{B} \setminus \{1\}$, and the h_i are nontrivial words in the indeterminates x_i, y_j, t_k. The free product $A[\xi_{n,A}] * F\{x;y;t\}$ becomes an algebra via juxtaposition of terms. In other words, given

$$g_j = b_{j,0} h_{j,1} b_{j,1} h_{j,2} b_{j,2} \cdots h_{j,m_j} b_{j,m_j}$$

for $j = 1, 2$, we write $b_{1,m_1} b_{2,0} = \alpha_1 + \sum_k \alpha_k b_k$ for $\alpha_k \in F$ and b_k ranging over $\mathcal{B} \setminus \{1\}$, and define

$$\begin{aligned} g_1 g_2 = &\alpha_1 b_{1,0} h_{1,1} b_{1,1} h_{1,2} b_{1,2} \cdots (h_{1,m_1} h_{2,1}) b_{2,1} h_{2,2} b_{2,2} \cdots h_{2,m_2} \\ &+ \sum_k \alpha_k b_{1,0} h_{1,1} b_{1,1} h_{1,2} b_{1,2} \cdots h_{1,m_1} b_k h_{2,1} b_{2,1} h_{2,2} b_{2,2} \cdots h_{2,m_2}. \end{aligned} \tag{15}$$

For example, if $b_1 b_2 = 1 + b_3 + b_4$, then

$$(b_2 h_{1,1} b_1)(b_2 h_{2,1} b_2) = b_2 (h_{1,1} h_{2,1}) b_2 + b_2 h_{1,1} b_3 h_{2,1} b_2 + b_2 h_{1,1} b_4 h_{2,1} b_2.$$

2.5.2. The relatively free product of A and $C\{x;y;t\}$ modulo a T-ideal

Even for algebras over an arbitrary base ring C, we can describe the free product of a C-algebra with $C\{x;y;t\}$ by going over the same construction and mimicking the tensor product. Namely we form the free C-module M having base comprised of all elements of the form $a_0h_1a_1h_2a_2\cdots h_ma_m$, $h_1a_1h_2a_2\cdots h_ma_m$, $a_0h_1a_1h_2a_2\cdots h_m$, and $h_1a_1h_2a_2\cdots h_m$ where $m \geqslant 0$, $a_0,\dots,a_m \in A$, and the h_i are nontrivial words in the indeterminates x_i, y_j, t_k.

The free product $A * C\{x;y;t\}$ is M/N, where N is the submodule generated by all

$$a_0h_1a_1h_2\cdots a_i\cdots h_ma_m + a_0h_1a_1h_2\cdots a_i'\cdots h_ma_m - a_0h_1a_1h_2\cdots (a_i+a_i')\cdots h_ma_m,$$

$$(ch_1) - ch_1,$$

$$ca_0h_1a_1h_2\cdots a_i - a_0h_1a_1h_2\cdots (ca_i)\cdots h_ma_m, \qquad a_i \in A,\ c \in C;$$

$A * C\{x;y;t\}$ becomes an algebra via juxtaposition of terms, i.e., given

$$g_j = a_{j,0}h_{j,1}a_{j,1}h_{j,2}a_{j,2}\cdots h_{j,m_j}a_{j,m_j}$$

for $j = 1, 2$, we define

$$g_1g_2 = ca_{1,0}h_{1,1}a_{1,1}h_{1,2}a_{1,2}\cdots (h_{1,m_1}h_{2,1})a_{2,1}h_{2,2}a_{2,2}\cdots h_{2,m_2} \tag{16}$$

when $a_{1,m_1}a_{2,0} = c \in C$, or

$$g_1g_2 = \alpha_1 a_{1,0}h_{1,1}a_{1,1}h_{1,2}a_{1,2}\cdots h_{1,m_1}(a_{1,m_1}a_{2,0})h_{2,1}a_{2,1}h_{2,2}a_{2,2}\cdots h_{2,m_2} \tag{17}$$

when $a_{1,m_1}a_{2,0} \notin C$.

We write $A\langle x;y;t\rangle$ for the free product $A * C\{x;y;t\}$.

We have the natural embedding $C\{x;y;t\} \to A\langle x;y;t\rangle$. For $g \in C\{x;y;t\}$, we write $\bar{g}$ for its natural image in $A * C\{x;y;t\}$.

Definition 2.23. Suppose $\mathcal{I}$ is a T-ideal of $C\{x;y;t\}$, for which $\mathcal{I} \subseteq \mathrm{id}(A)$. The **relatively free product** $A\langle x;y;t\rangle_{\mathcal{I}}$ of A and $C\{x;y;t\}$ **modulo** $\mathcal{I}$ is defined as $(A *_C C\{x;y;t\})/\hat{\mathcal{I}}$, where $\hat{\mathcal{I}}$ is the two-sided ideal $\mathcal{I}(A *_C C\{x;y;t\})$ consisting of all evaluations on $A * C\{x;y;t\}$ of polynomials from $\mathcal{I}$.

We can consider $A\langle x;y;t\rangle_{\mathcal{I}}$ as the ring of (noncommutative) polynomials but with coefficients from A interspersed throughout, taken modulo the relations in $\mathcal{I}$.

This construction is universal in the following sense: Any homomorphic image of $A\langle x;y;t\rangle$ satisfying these identities (from $\mathcal{I}$) is naturally a homomorphic image of $A\langle x;y;t\rangle_{\mathcal{I}}$. Thus, we have:

Lemma 2.24. (i) For any $g_1,\dots,g_k,h_1,\dots,h_k$ in $A\langle x;y;t\rangle$, there is a natural endomorphism $A\langle x;y;t\rangle \to A\langle x;y;t\rangle$ which fixes A and all t_i and sends $x_i \mapsto g_i$, $y_i \mapsto h_i$.

(ii) For any $g_1,\dots,g_k,h_1,\dots,h_k$ in $A\langle x;y;t\rangle_{\mathcal{I}}$, there is a natural endomorphism

$$A\langle x;y;t\rangle_{\mathcal{I}} \to A\langle x;y;t\rangle_{\mathcal{I}},$$

which fixes A and all t_i and sends $x_i \mapsto g_i$, $y_i \mapsto h_i$.

Although difficult to describe explicitly, the relatively free product is needed implicitly in all known proofs of the Braun-Kemer-Razmyslov Theorem in the literature. *From now on, we assume that $\mathcal{I}$ contains $\mathcal{CAP}_{n+1}$, so that we can work with $\widehat{M}$.*

Let $\widehat{\mathcal{M}}_A$ denote the image of $\mathcal{M}$ under substitutions to A, i.e., the C-submodule of $C\widehat{\{x,y,t\}}$ consisting of the images of all doubly alternating polynomials (in $x_1\dots,x_n$, and in $y_1,\dots,y_n$). In view of Lemma 2.13, the natural action of $\mathrm{Obst}_n(A)$ on $\widehat{\mathcal{M}}_A$ respects multiplication by the $\delta_{k,h}^{(n)}$-operators.

Proposition 2.25. $\mathrm{Obst}_n(A)\widehat{\mathcal{M}}_A = 0$.

Proof. If $a \in \mathrm{Obst}_n(A)$, then $a\mathcal{M} \in \mathcal{I}$, in view of Lemmas 2.5 and 2.12 and Proposition 2.17, so is 0 modulo $\mathcal{I}$. $\square$

Corollary 2.26. If $b \in A$ belongs to the T-ideal generated by doubly alternating polynomials, then $\mathrm{Obst}_n(A)b = 0$.

Proof. The element b belongs to the linear combinations images of $\widehat{\mathcal{M}}_A$ under specializations $x_i \mapsto a_i$. $\square$

By Step 7 of Section 1.2, this will complete the proof of the nilpotence of $\mathrm{Jac}(A)$ when C is a field, or more generally of any nil ideal when C is Noetherian, once we complete the proof of Proposition 2.10.

2.6. A more formal approach to Zubrilin's argument

Rather than push immediately into A, one can perform these computations first at the level of polynomials and then specialize. This requires a bit more machinery, since it requires adjoining the commuting indeterminates $\xi_{n,A}$ to the free product, but might be clearer conceptually.

Note that $\widehat{C\{t\}}[\xi_{n,C\{t\}}] = R \otimes_C \widehat{C\{t\}}$.

Lemma 2.27. $\widehat{\mathcal{M}}$ becomes an $\widehat{C\{t\}}[\xi_{n,C\{t\}}]$-module via the action given as follows:

Order the $\xi_{k,h}$ as $\xi_j = \xi_{k_j,h_j}$ for $1 \leqslant j < \infty$.

For a letter $\xi_j = \xi_{k_j,h_j}$, define

$$\xi_j \hat{f} = \hat{\delta}^{(n)}_{k_j,h_j}(\hat{f}),$$

and, inductively,

$$\xi_j^d \hat{f} = \hat{\delta}^{(n)}_{k_j,h_j}(\xi_j^{d-1}\hat{f}).$$

For a monomial $h = \xi_j^{d_j} \dots \xi_1^{d_1}$ of degree $d = d_1 + \dots + d_j$, define

$$hf = \xi_j^{d_j}(\xi_{j-1}^{d_{j-1}} \dots \xi_1^{d_1}\hat{f})$$

inductively on j.

Finally, define

$$\sum (c_i h_i)\hat{f} = \sum \alpha_i(h_i \hat{f})$$

where $c_i \in C$ and h_i are distinct monomials.

Proof.

The action is clearly well-defined, so we need to verify the associativity and commutativity of the action. It is enough to show that $(h_i h_{i'})\hat{f} = h_i(h_{i'})\hat{f}$ for any two monomials h_i and h_i'. But this follows inductively from induction on their length, plus the fact that $\xi_j(\xi_{j'}\hat{f}) = \xi_{j'}(\xi_j\hat{f})$ for any ξ_j and $\xi_{j'}$. $\square$

Let us continue to take $\mathcal{I} = \mathcal{CAP}_{n+1}$.

Remark 2.28. Clearly $A[\xi_{n,A}] *_C C\{t\} \subset A[\xi_{n,A}] *_C C\{x;y;t\}$ in the natural way, and then

$$\mathcal{I}(A[\xi_{n,A}] *_C C\{t\}) = (A[\xi_{n,A}] *_C C\{t\}) \cap \mathcal{I}(A[\xi_{n,A}] *_C C\{x;y;t\})$$

since we are just restricting the indeterminates $\vec{x}, \vec{y}, \vec{t}$ to the indeterminates $\vec{t}$.

It follows from Noether's Isomorphism Theorem that

$$\mathcal{F} := (A[\xi_{n,A}] *_C C\{t\})/\mathcal{I}(A[\xi_{n,A}] *_C C\{t\}),$$

can be viewed naturally in $(A[\xi_{n,A}] *_C C\{x;y;t\})/\mathcal{I}(A[\xi_{n,A}] *_C C\{x;y;t\})$.

Viewing $\mathcal{M} \subset C\{x;y;t\} \subset A[\xi_{n,A}] *_C C\{x;y;t\}$, we define

$$\tilde{\mathcal{M}}' = (A[\xi_{n,A}] *_C C\{t\})\mathcal{M} \subset A[\xi_{n,A}] *_C C\{x;y;t\}, \tag{18}$$

and its image in $(A[\xi_{n,A}] *_C C\{x;y;t\})/\mathcal{I}(A[\xi_{n,A}] *_C C\{x;y;t\})$, which we call $\tilde{\mathcal{M}}$ (intuitively consisting of terms ending with images of doubly alternating polynomials), which acts naturally by right multiplication on $\mathcal{F}$. To understand how $\tilde{\mathcal{M}}$ works, we look at the Capelli polynomial acting on $A *_C C\{x;y;t\}$ for an arbitrary algebra A satisfying Cap_{n+1}.

There is a more subtle action that we need. $\tilde{\mathcal{M}}$ can be viewed as an R-module where $R = C[\xi_{n,\widehat{C\{x;y;t\}}}]$, via the crucial Lemma 2.12. But as above, $\mathcal{M}$ is an $A * C\{t\}$-module where the algebra multiplication is induced from (15) (viewing $\mathcal{M} \subset C\{x;y;t\}$), implying $\widehat{\mathcal{M}}$ is an $A * C\{t\}$-module annihilated by CAP_{n+1}, and $\tilde{\mathcal{M}}$ thereby becomes an $A[\xi_{n,\widehat{C\{x;y;t\}}}] * \widehat{C\{t\}}$-module, where we define

$$\xi_{k,h}\hat{f} = \hat{\delta}^{(n)}_{k,h}\hat{f}$$

for $h \in \widehat{C\{x;y;t\}}$ and $\hat{f} \in \widehat{\mathcal{M}}$, by means of the action given in Lemma 2.27, also cf. Remark 2.15. Our main task is to identify these two actions when they are specialized to A.

2.6.1. The specialization argument

Having in hand the module $\tilde{\mathcal{M}}$ on which $A[\xi_{n,\widehat{C\{x;y;t\}}}]$ acts, we can specialize the assertion of Proposition 2.17 down to A once we succeed in matching the actions of $A[\xi_{n,\widehat{C\{x;y;t\}}}]$ and $A[\xi_{n,A}]$ when specializing to A.

Remark 2.29. $\mathcal{CAP}_k(A[\xi_{n,A}]) = \mathcal{CAP}_k(A)[\xi_{n,A}]$, since Cap_k is multilinear.

We write $\mathcal{DCAP}_n$ for the $C\{t\}$-submodule of $C\{x;y;t\}$ generated by DCap_n, cf. (5), and $\widehat{\mathcal{DCAP}}_n$ for its image in $\widehat{C\{x;y;t\}}$. This is a set of doubly alternating polynomials in $x_1,\dots,x_n$ and $y_1,\dots,y_n$, with variables t_i interspersed arbitrarily.

Lemma 2.30. Any specialization $\varphi : C\{x;y;t\} \to A$ (together with its accompanying specialization $\widehat{\varphi} : \widehat{C\{x;y;t\}} \to A$) gives rise naturally to a map

$$\Phi : A[\xi_{n,A}]\widehat{\mathcal{DCAP}}_n \to \tilde{\mathcal{M}}$$

given by

$$\sum_i a_i \xi_{k,\varphi(h_{j_i})}\widehat{f}_i \mapsto \sum_i a_i \widehat{\varphi}(\widehat{\xi_{k,h_{j_i}} f_i}) = \sum_i a_i \widehat{\varphi}(\widehat{\delta^{(x,n)}_{k,h_{j_i}} f_i})$$

where $\widehat{f}_i \in \widehat{\mathcal{DCAP}}_n$.

Proof. We need to show that this is well-defined, which follows from the functoriality property given in Lemma 2.5. Namely, if $\varphi(h_{j_i}) = \varphi(h'_{j_i})$, then $\widehat{\varphi}(\widehat{h_{j_i}}) = \widehat{\varphi}(\widehat{h'_{j_i}})$ and

$$\sum_i a_i \varphi(\widehat{\delta^{(x,n)}_{k,h_{j_i}}}\widehat{f}_i) = \sum_i a_i \varphi(\delta^{(x,n)}_{k,h_{j_i}} f_i) = \sum_i a_i \varphi(\delta^{(x,n)}_{k,h'_{j_i}} f_i) = \sum_i a_i \varphi(\widehat{\delta^{(x,n)}_{k,h'_{j_i}} f_i}).$$

□

The objective of this lemma was to enable us to replace $A[\xi_{n,A}]$ by A in our considerations. $\ker \Phi$ contains all $\widehat{\delta^{(n)}_{k,h} f} - \xi_{k,h}\hat{f}$ (cf. Remark 2.11) as well as $(\hat{h}^n - \sum_{k=0}^{n-1} \hat{h}^k \xi_{k,h})\widehat{f}$, where h ranges over all words and $\hat{f} \in \widehat{\mathcal{DCAP}}_n$, so we see that the Zubrilin integrality relations are passed on.

Lemma 2.31. If t is an infinite set of noncommuting indeterminates whose cardinality $\aleph$ is at least that of A, then for any given evaluation w in $\mathcal{DCAP}_n(A *_C C\{x;y;t\})$, there is a map

$$\varphi_w : C\{x;y;t\} \to A *_C C\{x;y;t\},$$

sending $\mathcal{DCAP}_n$ to $\mathcal{DCAP}_n(A *_C C\{x;y;t\})$, such that w is in the image of φ_w.

Proof. Note that $A *_C C\{x;y;t\}$ has cardinality $\aleph$. Setting aside indeterminates

$$\{t_g : g \in A *_C C\{x;y;t\}\},$$

we still have $\aleph$ indeterminates left over, to map onto our original set t of $\aleph$ indeterminates. But any evaluation w of $\mathcal{DCAP}_n$ on $A *_C C\{x;y;t\}$ can be written as

$$w = g\,\mathrm{Cap}_n(x_1,\dots,x_n;g_1,\dots,g_n)g'\,\mathrm{Cap}_n(y_1,\dots,y_n;h_1,\dots,h_n)g'', \tag{19}$$

for suitable $g,g',g'',g_i,h_j \in A *_C C\{x;y;t\}$. Defining φ_w by sending $x_i \mapsto x_i$, $y_j \mapsto y_j$, and sending the appropriate $t_g \mapsto g$, $t_{g'} \mapsto g'$, $t_{g''} \mapsto g''$, $t_{g_i} \mapsto g_i$, and $t_{h_j} \mapsto h_j$, we have an element in $\varphi_w^{-1}(w)$. □

Clearly $\varphi_w(\mathcal{CAP}_{n+1}) \subseteq \mathcal{CAP}_{n+1}(A *_C C\{x;y;t\})$, so, when $\mathrm{Cap}_{n+1} \in \mathcal{I}$, φ_w induces a map

$$\hat{\varphi}_w : \widehat{C\{x;y;t\}} \to (A *_C C\{x;y;t\})_{\mathcal{I}},$$

which sends $\widehat{\mathcal{M}} \to \tilde{\mathcal{M}}$.

Although we do not see that $\mathcal{CAP}_{n+1}$ need be mapped onto $\mathcal{CAP}_{n+1}(A *_C C\{x;y;t\})$, Lemma 2.31 says that it is “pointwise” onto, according to any chosen point, and this is enough for our purposes.

Theorem 2.32. $\mathrm{Obst}_n(A)\cdot(\mathcal{CAP}_n(A))^2 = 0$, for any PI-algebra $A = C\{a_1,\dots,a_\ell\}$ satisfying the Capelli identity Cap_{n+1}.

Proof.

We form the free algebra $C\{x;y;t\}$ by taking a separate indeterminate t_j for each element of $A[\xi_{n,A}]\widehat{\mathcal{DCAP}}_n$. We work with $A[\xi_{n,A}]\widehat{\mathcal{DCAP}}_n$, viewed in the relatively free product $\tilde{A} := (A[\xi_{n,A}] *_C C\{x;y;t\})_{\mathcal{I}}$, where $\mathcal{I} = \mathcal{CAP}_{n+1}(A[\xi_{n,A}] *_C C\{x;y;t\})$. In view of Lemma 2.30, the relation

$$I_{n,\widehat{C\{x;y;t\}}}\cdot\widehat{\mathcal{M}} \equiv 0 \pmod{\mathcal{CAP}_{n+1}(A[\xi_{n,A}] *_C C\{x;y;t\})}$$

restricts to the relation $I_{n,\widehat{C\{x;y;t\}}}\widehat{\mathcal{DCAP}}_n \equiv 0 \pmod{\mathcal{CAP}_{n+1}(A[\xi_{n,A}] *_C C\{x;y;t\})}$. But the various specializations of Lemma 2.31 cover all of $\mathcal{DCAP}_n(A)$. Hence Lemma 2.5 applied to Proposition 2.17 and Lemma 2.30 implies $I_{n,A}\mathcal{DCAP}_n(A) = 0$, and thus

$$\mathrm{Obst}_n(A)\cdot(\mathcal{CAP}_n(A))^2 \subseteq I_{n,A}\mathcal{DCAP}_n(A) = 0.$$

□

2.7. The proof of Proposition 2.10

Now we present the proof of the crucial Proposition 2.10, stating that for a doubly alternating polynomial $f = f(x_1,\dots,x_n,y_1,\dots,y_n,\vec{t})$,

$$\delta_{k,h}^{(x,n)}(f) \equiv \delta_{k,h}^{(y,n)}(f) \quad \textit{modulo } \mathcal{CAP}_{n+1}.$$

2.7.1. The connection to the group algebra of S_n

We begin with the basic correspondence between multilinear identities and elements of the group algebra over S_n.

$V_n = V_n(x_1, \dots, x_n)$ denotes the C-module of multilinear polynomials in $x_1, \dots, x_n$, i.e.,

$$V_n = \operatorname{span}_C\{x_{\sigma(1)}x_{\sigma(2)} \cdots x_{\sigma(n)} \mid \sigma \in S_n\}.$$

Definition 2.33. We identify V_n with the group algebra $C[S_n]$, by identifying a permutation $\sigma \in S_n$ with its corresponding monomial (in $x_1, x_2, \dots, x_n$):

$$\sigma \leftrightarrow M_\sigma(x_1, \dots, x_n) = x_{\sigma(1)} \cdots x_{\sigma(n)}.$$

Any polynomial $\sum \alpha_\sigma x_{\sigma(1)} \cdots x_{\sigma(n)}$ corresponds to an element $\sum \alpha_\sigma \sigma \in C[S_n]$, and conversely, $\sum \alpha_\sigma \sigma$ corresponds to the polynomial

$$\left(\sum \alpha_\sigma \sigma\right) x_1 \cdots x_n = \sum \alpha_\sigma x_{\sigma(1)} \cdots x_{\sigma(n)}.$$

Here is a combinatorial identity of interest of its own.

Consider two disjoint sets $X \cap Y = \emptyset$, each of cardinality n, and the symmetric group $S_{2n} = S_{X \cup Y}$ acting on $X \cup Y$. For each subset $Z \subseteq X$ we define an element $P(Z) \in C[S_{2n}]$ as follows:

$$P(Z) = \sum_{\sigma(Z) \subseteq Y} \operatorname{sgn}(\sigma) \cdot \sigma.$$

In particular

$$P(\emptyset) = \sum_{\sigma \in S_{2n}} \operatorname{sgn}(\sigma) \cdot \sigma.$$

Proposition 2.34.

$$\sum_{Z \subseteq X} (-1)^{|Z|} P(Z) = \sum_{\sigma(X) = X} \operatorname{sgn}(\sigma) \cdot \sigma. \tag{20}$$

Proof. Let $\sigma \in S_{2n}$ and let a_σ (resp. b_σ) be the coefficient of σ on the l.h.s. (resp. r.h.s.) of (20). We show that $a_\sigma = b_\sigma$.

Let $Z(\sigma) = \sigma^{-1}(Y)$ be the largest subset $Z \subseteq X$ such that $\sigma(Z) \subseteq Y$. Note that $\sigma(X) = X$ if and only if $Z(\sigma) = \emptyset$. Therefore

$$b_\sigma = \operatorname{sgn}(\sigma) \text{ if } Z(\sigma) = \emptyset \quad \text{and} \quad b_\sigma = 0 \text{ if } Z(\sigma) \neq \emptyset, \tag{21}$$

since $P(\emptyset) = \sum \operatorname{sgn}(\sigma) \cdot \sigma$. We claim that

$$a_\sigma = \operatorname{sgn}(\sigma) \cdot \sum_{Z \subset Z(\sigma)} (-1)^{|Z|}.$$

To show this, recall that

$$l.h.s = \sum_{Z \subseteq X} (-1)^{|Z|} \sum_{\sigma(Z) \subseteq Y} \operatorname{sgn}(\sigma) \cdot \sigma.$$

In $P(Z)$ the coefficient of σ is $\operatorname{sgn}(\sigma)$ if $Z \subseteq Z(\sigma)$ (since then $\sigma(Z) \subseteq Y$), and is zero if $Z \not\subseteq Z(\sigma)$ (since if $\sigma(Z) \subseteq Y$ then $\sigma(Z \cup Z(\sigma)) \subseteq Y$, contradicting the maximality of $Z(\sigma)$). It follows that as claimed,

$$a_\sigma = \operatorname{sgn}(\sigma) \cdot \sum_{Z \subseteq Z(\sigma)} (-1)^{|Z|}.$$

It is well known that $\sum_{Z\subseteq Z(\sigma)}(-1)^{|Z|} = 1$ when $Z(\sigma) = \emptyset$ and $= 0$ otherwise. Therefore

$$a_\sigma = \operatorname{sgn}(\sigma) \quad \text{if} \quad Z(\sigma) = \emptyset \quad \text{and} \quad a_\sigma = 0 \quad \text{if} \quad Z(\sigma) \neq \emptyset. \tag{22}$$

The proof now follows by comparing (21) with (22).
□

Lemma 2.35. Let $f(x_1, \ldots, x_n, y_1, \ldots, y_n, \vec{t}\,)$ be doubly alternating. Then

$$f(x_1, \ldots, x_n, y_1, \ldots, y_n, \vec{t}\,) \equiv f(y_1, \ldots, y_n, x_1, \ldots, x_n, \vec{t}\,) \quad \text{modulo} \ \mathcal{CAP}_{n+1}.$$

Proof. Let $X = \{x_1, \ldots, x_n\}$ and $Y = \{y_1, \ldots, y_n\}$. Then $|X| = |Y| = n$ and $X \cap Y = \emptyset$, and we identify $S_{2n} = S_{X\cup Y}$. Let $M = \{x_{i_1}, \ldots, x_{i_k}\} \subseteq X$, with $1 \leqslant i_1 < \cdots < i_k \leqslant n$, and $N = \{y_{j_1}, \ldots, y_{j_k}\} \subseteq Y$, with $1 \leqslant j_1 < \cdots < j_k \leqslant n$. Thus, $|M| = |N| = k \leqslant n$. M will play the role of Z in Proposition 2.34. We consider permutations $\sigma \in S_{2n}$ with $\sigma(M) = N$. Define the permutation

$$\tau_{MN} = (x_{i_1}, y_{j_1}) \cdots (x_{i_k}, y_{j_k}).$$

Since $M \cap N = \emptyset$, τ_{MN} has order 2 in S_{2n}, and satisfies $\operatorname{sgn}(\tau_{MN}) = (-1)^k$. If $M = X$ then $N = Y$ and $\operatorname{sgn}(\tau_{MN}) = \operatorname{sgn}(\tau_{XY}) = (-1)^n$. Moreover $\tau_{MN}(M) = N$ and $\tau_{MN}(N) = M$.

Next, we define

$$T_{MN} = \sum_{\pi(M)=N} \operatorname{sgn}(\pi)\cdot\pi \in C[S_n].$$

Let $\rho = \tau_{MN}\cdot\pi$, so that $\rho(M) = M$. Then $\pi = \tau_{MN}\cdot\rho$ and

$$T_{MN} = \operatorname{sgn}(\tau_{MN})\cdot\tau_{MN}\cdot\left(\sum_{\rho(M)=M}\operatorname{sgn}(\rho)\cdot\rho\right).$$

But by Proposition 2.34,

$$\sum_{M\subseteq X}(-1)^{|M|}P(M) = \sum_{\sigma(X)=X}\operatorname{sgn}\sigma\cdot\sigma. \tag{23}$$

If $M \subsetneq X$, then $P(M)$ is alternating on $2n - |M| \geqslant n+1$ indeterminates, and hence is 0 modulo $\mathcal{CAP}_{n+1}$. Thus, modulo $\mathcal{CAP}_{n+1}$, the left hand side of (23) equals the unique summand with $M = X$, which is

$$(-1)^n\sum_{\sigma(X)=Y}\operatorname{sgn}(\sigma)\cdot\sigma = (-1)^n\operatorname{sgn}(\tau_{XY})\cdot\tau_{XY}\cdot\left(\sum_{\sigma(Y)=Y}\operatorname{sgn}(\sigma)\cdot\sigma\right) =$$

$$= \tau_{XY}\cdot\left(\sum_{\sigma(Y)=Y}\operatorname{sgn}(\sigma)\cdot\sigma\right).$$

Since $\sigma(X) = X$ if and only if $\sigma(Y) = Y$, it follows that

$$\sum_{\sigma(Y)=Y}\operatorname{sgn}(\sigma)\cdot\sigma = \sum_{\sigma(X)=X}\operatorname{sgn}(\sigma)\cdot\sigma \equiv \tau_{XY}\cdot\left(\sum_{\sigma(Y)=Y}\operatorname{sgn}(\sigma)\cdot\sigma\right), \quad \text{modulo } \mathcal{CAP}_{n+1}.$$

Now we identify elements in $C[S_{2n}]$ with polynomials multilinear in $x_1, \dots, x_n, y_1, \dots, y_n$. Taking a monomial $h(x_1, \dots, x_n, y_1, \dots, y_n; \vec{t})$ multilinear in $x_1, \dots, x_n, y_1, \dots, y_n$, we define

$$f(x_1, \dots, x_n, y_1, \dots, y_n; \vec{t}) := \left(\sum_{\sigma(Y)=Y} \operatorname{sgn}(\sigma) \cdot \sigma \right) h.$$

Then

$$\tau_{XY} \cdot \left(\sum_{\sigma(Y)=Y} \operatorname{sgn}(\sigma) \cdot \sigma \right) h = f(y_1, \dots, y_n, x_1, \dots, x_n; \vec{t}).$$

Again, since $\sigma(X) = X$ if and only if $\sigma(Y) = Y$, it follows that $f(x_1, \dots, x_n, y_1, \dots, y_n; \vec{t})$ is doubly alternating, and we have proved that

$$f(x_1, \dots, x_n, y_1, \dots, y_n; \vec{t}) \equiv f(y_1, \dots, y_n, x_1, \dots, x_n; \vec{t}) \quad \text{modulo} \;\; \mathcal{CAP}_{n+1},$$

as desired.

□

2.7.2. Proof of Proposition 2.10

We may assume that h is a new indeterminate z. Recall that

$$\delta_{k,z}^{(x,n)}(f(x_1, \dots, x_n, y_1, \dots, y_n, \vec{t}\,)) = \\ = \sum_{1 \leqslant i_1 < \dots < i_k \leqslant n} f(x_1, \dots, x_n, y_1, \dots, y_n, \vec{t}\,)|_{x_{i_u} \mapsto z x_{i_u}}; \quad u = 1, \dots, k,$$

and

$$\delta_{k,z}^{(y,n)}(f(x_1, \dots, x_n, y_1, \dots, y_n, \vec{t}\,)) = \\ = \sum_{1 \leqslant i_1 < \dots < i_k \leqslant n} f(x_1, \dots, x_n, y_1, \dots, y_n, \vec{t}\,)|_{y_{i_u} \mapsto z y_{i_u}}; \quad u = 1, \dots, k.$$

Let $z' = 1 + \varepsilon z$, ε being a central indeterminant. Then clearly

$$f(z'x_1, \dots, z'x_n, y_1, \dots, y_n, \vec{t}\,) = \sum_{k=0}^{n} \varepsilon^k \cdot \delta_{k,z}^{(x,n)}(f(x_1, \dots, x_n, y_1, \dots, y_n, \vec{t}\,)) \tag{24}$$

and

$$f(x_1, \dots, x_n, z'y_1, \dots, z'y_n, \vec{t}\,) = \sum_{k=0}^{n} \varepsilon^k \cdot \delta_{k,z}^{(y,n)}(f(x_1, \dots, x_n, y_1, \dots, y_n, \vec{t}\,)). \tag{25}$$

By Equations (24) and (25) it is enough to show that

$$f(z'x_1, \dots, z'x_n, y_1, \dots, y_n, \vec{t}\,) \equiv f(x_1, \dots, x_n, z'y_1, \dots, z'y_n, \vec{t}\,) \quad \text{modulo} \;\; \mathcal{CAP}_{n+1}.$$

Let

$$g_1(x_1, \dots, x_n, y_1, \dots, y_n, \vec{t}\,) = f(z'x_1, \dots, z'x_n, y_1, \dots, y_n, \vec{t}\,)$$

and

$$g_2(x_1, \dots, x_n, y_1, \dots, y_n, \vec{t}\,) = f(x_1, \dots, x_n, z'y_1, \dots, z'y_n, \vec{t}\,).$$

We have to show that

$$g_1 \equiv g_2 \quad \text{modulo } \mathcal{CAP}_{n+1}.$$

Denote $x_i' = z'x_i,\ y_i' = z'y_i;\ i = 1, \ldots, n$. Then

$$\begin{aligned} g_1(x_1, \ldots, x_n, y_1, \ldots, y_n, \vec{t}\,) = f(x_1', \ldots, x_n', y_1, \ldots, y_n, \vec{t}\,) &\equiv \\ \equiv f(y_1, \ldots, y_n, x_1', \ldots, x_n', \vec{t}\,) &= \\ = g_2(y_1, \ldots, y_n, x_1, \ldots, x_n, \vec{t}\,) &\equiv \\ \equiv g_2(x_1, \ldots, x_n, y_1, \ldots, y_n, \vec{t}\,) & \quad \text{modulo } \mathcal{CAP}_{n+1}. \end{aligned}$$

The congruences follow from Lemma 2.35 since both f and g_2 are doubly alternating.

3. Proof of Kemer's "Capelli Theorem,"

To complete the proof of Theorem 1.1, it remains to present an exposition of Kemer's "Capelli Theorem," that any affine PI algebra over a field F satisfies a Capelli identity Cap_n for large enough n. This is done by abstracting a key property of Cap_n, called **spareseness**.

Definition 3.1. A multilinear polynomial $g = \sum \alpha_\sigma x_{\sigma(1)} \ldots x_{\sigma(d)}$ is a **sparse identity** of A if, for any monomial $f(x_1, \ldots, x_d; \vec{t})$ we have

$$\sum \alpha_\sigma f(x_{\sigma(1)}, \ldots, x_{\sigma(d)}; \vec{t}) \in \mathrm{id}(A).$$

See [7, §2.5.2] for more detail. The major example of a sparse identity is the Capelli identity. One proves rather quickly that any sparse identity implies a Capelli identity, so it remains to show that any affine PI algebra over a field satisfies a sparse identity. There are two possible approaches, both using the classical representation theory of S_n. One proof relies on "the branching theorem," which requires characteristic 0, and the other relies more on the structure of the group algebra $F[S_n]$, also with the technique of "pumping" polynomial identities, and works in arbitrary characteristic.

3.1. Affine algebras satisfying a sparse identity

Sparse identities work well with the left lexicographic order $<$. If $b_1 < \cdots < b_m$ and $1 \neq \sigma \in S_m$, then $(b_1, \ldots, b_m) < (b_{\sigma(1)}, \ldots, b_{\sigma(m)})$. Any sparse identity over a field yields a powerful **sparse reduction procedure**. Namely, we may assume $\alpha_{(1)} = 1$; given $a_1, \ldots, a_d$ in A, we can replace any term $f(a_1, \ldots, a_d)$ by

$$-\sum_{\sigma \neq 1} \alpha_\sigma f(x_{\sigma(1)} \ldots x_{\sigma(d)}, x_{d+1}, \ldots, x_n).$$

(The analogous assertion also holds for c_d.)

Lemma 3.2. Let $A = C\{a_1, a_2 \ldots\}$ be a PI algebra, satisfying a sparse multilinear identity $p = \sum_{\sigma \in S_d} \beta_\sigma x_{\sigma(1)} \cdots x_{\sigma(d)}$ of degree d, with $d \leqslant n$, and let $M(x_1, \ldots, x_n; \vec{y}\,)$ be a monomial multilinear in $x_1, \ldots, x_n$ and perhaps involving extra indeterminates $\vec{y}$. We consider $\Delta = M(v_1, \ldots, v_n; \overline{\vec{y}})$, where $v_1, \ldots, v_n$ are words in the generators $a_1, a_2, \ldots$ and $\overline{\vec{y}}$ is an arbitrary specialization of $\vec{y}$ in A. Assume that k of the v_i satisfy $|v_i| \geqslant d$ (length as words in $a_1, a_2, \ldots$). If $\ell \geqslant d$, then Δ is a linear combination of monomials $\Delta' = M(v_1', \ldots, v_n'; \overline{\vec{y}})$ where at most $\ell - 1$ of the words v_i' have length $\geqslant d$.

This clearly implies that A is spanned by monomials $\Delta' = M(v'_1, v'_2, \ldots)$, with at most $d-1$ of the v'_i having length $\geqslant d$.
Proof.

Claim: If $|v_{i_1}|, \ldots, |v_{i_d}| \geqslant d$, then $\Delta = M(v_1, \ldots, v_n; \overline{\vec{y}})$ is a linear combination of terms $\Delta' = M(v'_1, \ldots, v'_n; \overline{\vec{y}})$ satisfying

$$(|v'_1|, \ldots, |v'_n|) < (|v_1|, \ldots, |v_n|).$$

The above Claim implies the existence of descending sequences of monomials, under the left lexicographic order. Such a descending sequence must stop. When it stops we have a corresponding monomial having strictly fewer words v'_i for which $|v'_i| \geqslant d$. Therefore proving the above Claim will prove the lemma. We now prove the Claim.

We rewrite $\Delta = M(v_{i_1}, \ldots, v_{i_d}; \overline{\vec{y}})$, where $i_1 < i_2, \cdots < i_d$; then we may assume that $i_1 = 1, \ldots, i_d = d$. We write $v_i = w_i u_i$ where $|u_i| = d - i$, $1 \leqslant i \leqslant d$. The sparse identity p implies that Δ is a linear combination of terms $\Delta_\sigma = M(w_1 u_{\sigma(1)}, \ldots, w_h u_{\sigma(d)}; \overline{\vec{y}}) = M(v'_1, \ldots, v'_d; \overline{\vec{y}})$ where $1 \neq \sigma \in S_d$. (Δ itself corresponds to $\sigma = 1$.) To see this, we rewrite $\Delta = M(w_1 u_1, \ldots, w_d u_d; \overline{\vec{y}})$ as $N(u_1, \ldots, u_d; \overline{W})$. The sparse identity p implies that $N(u_1, \ldots, u_d; \overline{W})$ is a linear combination of elements of the form

$$N(u_{\sigma(1)}, \ldots, u_{\sigma(d)}; \overline{W}) = M(w_1 u_{\sigma(1)}, \ldots, w_d u_{\sigma(d)}; \overline{\vec{y}}), \quad 1 \neq \sigma \in S_d.$$

Denote $w_i u_{\sigma(i)} = v'_i$, $1 \leqslant i \leqslant d$. But then $(|v'_1|, \ldots, |v'_d|) < (|v_1|, \ldots, |v_d|)$ for such $\sigma \neq 1$. This proves the Claim, and completes the proof of the lemma. □

Although we did not apply Shirshov's Height Theorem, the main argument here is similar. Note also that Lemma 3.2 applies to any PI algebra, not necessarily affine. In the next theorem, due to Kemer, we do assume that A is affine.

Theorem 3.3. Let $A = C\{a_1, \ldots, a_r\}$ be an affine PI algebra over a commutative ring C, satisfying a sparse identity p of degree d, and let $n \geqslant r^d + d$. Then A satisfies the Capelli identity $\mathrm{Cap}_n[x; y]$.

Proof. We may assume that $r \geqslant 2$, since otherwise A is commutative. Consider

$$\mathrm{Cap}_n(v_1, \ldots, v_n; w_1, \ldots, w_n)$$

where $v_i, w_i \in A$. By Lemma 3.2 we may assume that at most $d-1$ of the v_i have length $\geqslant d$ (as words in the generators $a_1 \ldots, a_r$). Hence at least $n - (d-1)$ of the v_i have length $\leqslant d-1$. The number of distinct words of length q is $\leqslant r^q$. Hence the number of words of length $\leqslant d-1$ is

$$\leqslant 1 + r + r^2 + \cdots + r^{d-1} = \frac{r^d - 1}{r - 1} < r^d \quad (\text{since } r \geqslant 2).$$

But we have at least $n - (d-1)$ such words appearing in $v_1, \ldots, v_n$, and $n - (d-1) > r^d$ (since by assumption $n \geqslant r^d + d$). It follows that there must be repetitions among $v_1, \ldots, v_n$, so $\mathrm{Cap}_n(v_1, \ldots, v_n; w_1, \ldots, w_n) = 0$. □

3.2. Actions of the group algebra

It remains to prove the existence of sparse identities for affine PI-algebras. For this, we turn to the representation theory of S_n. After a brief review of actions of S_n on Young diagrams, we treat the characteristic 0 case, cf. Kemer [15]. The characteristic $p>0$ proof, which requires some results about modular representations but bypassing branching, is done in §3.2 and §4.

Given $\sigma,\pi\in S_n$, by convention we take $\sigma\pi(i)=\pi(\sigma(i))$. The product $\sigma\pi$ corresponding (by Definition 2.33) to the monomial

$$M_{\sigma\pi}=x_{\sigma\pi(1)}\cdots x_{\sigma\pi(n)}$$

can be interpreted in two ways, according to **left** and **right** actions of S_n on V_n, described respectively as follows:

Let $\sigma,\pi\in S_n$. Let $y_i=x_{\sigma(i)}$. Then

(i) $\sigma M_\pi(x_1\dots,x_n):=M_{\sigma\pi}=M_\pi(x_{\sigma(1)},\dots,x_{\sigma(n)})$ and
(ii) $M_\sigma(x_1\dots,x_n)\pi:=(y_1\cdots y_n)\pi=M_{\sigma\pi}=y_{\pi(1)}\cdots y_{\pi(n)}$.
Thus, the effect of the right action of π on a monomial is to permute the places of the indeterminates according to π.

Extending by linearity, we obtain for any $f=f(x_1,\dots,x_n)\in V_n$ that

(i) $\sigma p(x_1,\dots,x_n)=p(x_{\sigma(1)},\dots,x_{\sigma(n)})$;

(ii) $p(x_1,\dots,x_n)\pi=q(y_1,\dots,y_n)$, where $q(y_1,\dots,y_n)$ is obtained from $p(x_1,\dots,x_n)$ by place-permuting all the monomials of p according to the permutation π.

For any finite group G and field F, there is a well-known correspondence between the $F[G]$-modules and the representations of G. The simple modules correspond to the irreducible representations.

Remark 3.4. If $p\in\mathrm{Id}(A)$, then $\sigma p\in\mathrm{Id}(A)$ since the left action is just a change of variables.

Hence, for any PI-algebra A, the spaces

$$\mathrm{Id}(A)\cap V_n\subseteq V_n$$

are in fact left ideals of $F[S_n]$ (thereby affording certain S_n representations), but need not be two-sided ideals. However, we prove below the existence of a nonzero two-sided ideal in $\mathrm{Id}(A)\cap V_n$, a fact which is of crucial importance in what follows.

Remark 3.5. Let λ be a partition. As explained in [7, p. 147], any tableau T of λ gives rise to an element

$$a_T=\sum_{q\in\mathcal{C}_{T_\lambda},\ p\in\mathcal{R}_{T_\lambda}}\mathrm{sgn}(q)qp\in C[S_n],$$

where $\mathcal{C}_{T_\lambda}$ (resp. $\mathcal{R}_{T_\lambda}$) denotes the set of column (resp. row) permutations of the tableau T_λ. $a_T^2=\alpha_T a_t$ for some α_T in the base field F. When $\alpha_T\neq 0$, which by [29, Lemma 19.59(i)] is always the case when $\mathrm{char}(F)$ does not divide n, in particular, when $\mathrm{char}(F)=0$, we will call the idempotent $e_T:=\alpha_T^{-1}a_T$ the **Young symmetrizer** of the tableau T.

Furthermore, by [29, Lemma 19.59(i)], if $a_T\neq 0$ and then $F[S_n]a_T=Fa_T$, implying $F[S_n]a_T$ (if nonzero) is a minimal left ideal, which we call J_λ. Thus, if J_λ contains an element corresponding to a nontrivial PI of A, a_T itself must correspond to a PI of A.

$s^\lambda:=\dim J_\lambda$ is given by the “hook"formula, see for example [30] or [14], where we recall that each “hook"number h_x for a box x is the number of boxes in “hook"formed by taking all boxes to the right of x and beneath x. (In the literature, one writes f^λ instead of s^λ, but here we have used f throughout for polynomials.)

Lemma 3.6. Suppose L is a minimal left ideal of a ring R. Then the minimal two-sided ideal of R containing L is a sum of minimal left ideals of R isomorphic to L as modules.

We let I_λ denote the minimal two-sided ideal of $F[S_n]$ containing J_λ.

We define the **codimension** $c_n(A) = \dim\left(\frac{V_n}{\mathrm{Id}(A)\cap V_n}\right)$. The characteristic 0 version of the next result is in [24].

LEMMA 3.7. Let A be an F-algebra, and let λ be a partition of n. If $\dim J_\lambda > c_n(A)$, then $I_\lambda \subseteq \mathrm{Id}(A)\cap V_n$.

PROOF. By Lemma 3.6, J_λ is a sum of minimal left ideals, with each such minimal left ideal J isomorphic to J_λ. Thus, $\dim J = \dim J_\lambda > c_n(A)$. Since J is minimal, either $J \subseteq \mathrm{Id}(A)\cap V_n$ or $J\cap\big(\mathrm{Id}(A)\cap V_n\big) = 0$. If $J\cap\big(\mathrm{Id}(A)\cap V_n\big) = 0$ then it follows that

$$c_n(A) = \dim V_n/\big(\mathrm{Id}(A)\cap V_n\big) \geqslant \dim J > c_n(A),$$

a contradiction. Therefore each $J \subseteq \mathrm{Id}(A)\cap V_n$. $I_\lambda \subseteq \mathrm{Id}(A)\cap V_n$ since I_λ equals the sum of these minimal left ideals. □

3.3. The characteristic 0 case [15]

The characteristic 0 case is treated separately here, since it can be handled via the classical representation theory of the symmetric group. By Maschke's Theorem, the group algebra FS_n now is a finite direct product of matrix algebras over F. We have the decomposition $FS_n = \bigoplus_{\lambda\vdash n} I_\lambda$.

Thus, Lemma 3.7 yields at once:

LEMMA 3.8. [24] Let $\mathrm{char}(F) = 0$, let A be an F algebra, and let λ be a partition of n. If $s^\lambda > c_n(A)$, then $I_\lambda \subseteq \mathrm{Id}(A)\cap V_n$.

(Here I_λ is the sum of those $F[S_n]e_T$ for which T is a standard tableau with partition λ. These I_λ are minimal two sided ideals, each a sum of s^λ minimal left ideals isomorphic to J_λ.)

EXAMPLE 3. *Consider the "rectangle" of u rows and v columns. By [20, page 11], the hook numbers of the partition $\mu = (u^v)$ satisfy*

$$\sum_{x\in\mu} h_x = uv(u+v)/2 = n\frac{u+v}{2}.$$

Let us review the proof, for further reference. For any box x in the $(1,j)$ position, the hook has length $u+v-j$, so the sum of all hook numbers in the first row is

$$\sum_{j=1}^{v}(u+v-j) = uv + \frac{v(v-1)}{2} = v\left(u+\frac{v-1}{2}\right).$$

Summing this over all rows yields

$$v\frac{u(u+1)}{2} + uv\frac{v-1}{2} = uv\left(\frac{u+1}{2}+\frac{v-1}{2}\right) = uv\frac{u+v}{2},$$

as desired.

3.3.1. Strong identities

DEFINITION 3.9. Let A be a PI algebra. The multilinear polynomial $g \in V_n$ is a *strong* identity of A if for every $m \geqslant n$ we have $FS_m \cdot g \cdot FS_m \subseteq \mathrm{Id}(A)$.

Note that every strong identity is sparse. To obtain strong identities, we utilize the following construction, due to Amitsur.

The natural embedding $S_n \subset S_{n+1}$ (via $\sigma(n+1) = n+1$ for $\sigma \in S_n$) induces the embedding $V_n \subset V_{n+1}$: $f(x_1, \ldots, x_n) \equiv f(x_1, \ldots, x_n) \cdot x_{n+1}$. More generally, for any $n < m$ we have the inclusion $V_n \subset V_m$ via $f(x_1, \ldots, x_n) \equiv f(x_1, \ldots, x_n) \cdot x_{n+1} \cdots x_m$.

For $f(x) = f(x_1, \ldots, x_n) = \sum_{\sigma \in S_n} \alpha_\sigma x_{\sigma(1)} \cdots x_{\sigma(n)} \in V_n$, we define

$$f^*(x_1, \ldots, x_n; x_{n+1}, \ldots, x_{2n-1}) = \sum_{\sigma \in S_n} \alpha_\sigma x_{\sigma(1)} x_{n+1} x_{\sigma(2)} x_{n+2} \cdots x_{\sigma(n-1)} x_{2n-1} x_{\sigma(n)} \tag{26}$$

$$= (f(x_1, \ldots, x_n) x_{n+1} \cdots x_{2n-1})\eta,$$

where $\eta \in S_{2n-1}$ is the permutation

$$\eta = \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & 2n-1 \\ 1 & n+1 & 2 & n+2 & \cdots & n \end{pmatrix}. \tag{27}$$

Let $L \subseteq \{x_{n+1}, \ldots, x_{2n-1}\}$ and denote by f_L^* the polynomial obtained from f^* by substituting $x_j \to 1$ for all $x_j \in L$. Rename the indeterminates in $\{x_{n+1}, \ldots, x_{2n-1}\} \setminus L$ as $\{x_{n+1}, \ldots, x_{n+q}\}$ (where $q = n - 1 - |L|$) and denote the resulting polynomial as f_L^*. Then similarly to (26), there exists a permutation $\rho \in S_{n+q}$ such that $f_L^* = (f x_{n+1} \cdots x_{n+q})\rho$.

Note that if $1 \in A$ and $f^* \in \mathrm{Id}(A)$, then also $f_L^* \in \mathrm{Id}(A)$ for any such L, and in particular $f \in \mathrm{Id}(A)$. The converse is not true: it is possible that $f \in \mathrm{Id}(A)$ but $f^* \notin \mathrm{Id}(A)$.

Lemma 3.10. Let A be a PI algebra, let $I \subseteq V_n$ be a two–sided ideal in V_n, and assume for any $f \in I$ that $f^* \in \mathrm{Id}(A)$ (and thus $f \in \mathrm{Id}(A)$). Then for any $m \geqslant n$,

$$(FS_m) I (FS_m) \subseteq \mathrm{Id}(A).$$

Proof. Since $(FS_m)I \subseteq \mathrm{Id}(A)$, it suffices to prove:

Claim: If $f \in I$ and $\pi \in S_m$, then $f_L^* \pi = (f(x_1, \ldots, x_n) x_{n+1} \cdots x_m)\pi \in \mathrm{Id}(A)$.

If $f = \sum_{\sigma \in S_n} a_\sigma \sigma(x_1 \cdots x_n \cdots x_m)$, then $f_L^* \pi = \sum_{\sigma \in S_n} a_\sigma \sigma(\pi(x_1 \cdots x_n \cdots x_m))$.

Consider the positions of $x_1, \ldots, x_n$ in the monomial $\pi(x_1 \cdots x_m)$: There exists $\tau \in S_n$ such that

$$\pi(x_1 \cdots x_n \cdots x_m) = g_0 x_{\tau(1)} g_1 x_{\tau(2)} g_2 \cdots g_{n-1} x_{\tau(n)} g_n = \tau(g_0 x_1 g_1 x_2 g_2 \cdots g_{n-1} x_n g_n),$$

where each g_j is $= 1$ or is a monomial in some of the indeterminates $x_{n+1}, \ldots, x_m$. It follows that $f_L^* \pi = (f\tau)(g_0 x_1 g_1 x_2 g_2 \cdots g_{n-1} x_n g_n)$. Since $f \in V_n$ and $\tau \in S_n$, $f\tau$ only permutes the indeterminates $x_1, \ldots, x_n$, and hence (see (26))

$$f_L^* \pi = (f\tau)(g_0 x_1 g_1 x_2 g_2 \cdots g_{n-1} x_n g_n) = g_0((f\tau)^*[x_1, \ldots, x_n; g_1, \ldots, g_{n-1}])g_n.$$

Since I is two-sided, $f\tau \in I$, hence by assumption $(f\tau)^* \in \mathrm{Id}(A)$, which by the last equality implies that $f\pi \in \mathrm{Id}(A)$. □

3.3.2. Existence of nonzero two-sided ideals $I_\lambda \subseteq FS_n$ of identities

Let $c_n(A) \leqslant \alpha^n$ for all n. The next lemma yields rectangles $\mu = (u^v) \vdash n$ such that $\alpha^n < s^\mu$.

Lemma 3.11. Let $0 < u, v$ be integers and let μ be the $u \times v$ rectangle $\mu = (u^v) \vdash u \cdot v$. Let $n = uv$. Then

$$\left(\frac{n}{u+v}\right)^n \cdot \left(\frac{2}{e}\right)^n < s^\mu \qquad (\text{where} \quad e = 2.718281828\ldots).$$

In particular, if $\alpha \leqslant \frac{n}{u+v} \cdot \frac{2}{e}$ then $\alpha^n \leqslant s^\mu$.

Proof.

Since the geometric mean is bounded by the arithmetic mean,

$$\left(\prod_{x\in\mu} h_x\right)^{1/n} \leqslant \frac{1}{n}\sum_{x\in\mu} h_x = \frac{u+v}{2},$$

in view of Example 3, and hence

$$\left(\frac{2}{u+v}\right)^n \leqslant \frac{1}{\prod_{x\in\mu} h_x}.$$

Together with the classical inequality $(n/e)^n < n!$, this implies that

$$\left(\frac{uv}{u+v}\right)^n \cdot \left(\frac{2}{e}\right)^n = \left(\frac{n}{e}\right)^n \cdot \left(\frac{2}{u+v}\right)^n < \frac{n!}{\prod_{x\in\mu} h_x} = s^\mu.$$

□

Remark 3.12. To apply this, we need Regev's estimate [23] of codimensions,

$$c_m(A) \leqslant (d-1)^{2m},$$

as explained in [7, Theorem 5.38].

Proposition 3.13. [4] Let A be a PI algebra satisfying an identity of degree d. Choose natural numbers u and v such that

$$\frac{uv}{u+v} \cdot \frac{2}{e} \geqslant (d-1)^4. \qquad \text{For example, choose} \quad u = v \geqslant e \cdot (d-1)^4.$$

Let $n = uv$ and let $\mu = (u^v)$ be the $u \times v$ rectangle. Let $n \leqslant m \leqslant 2n$ and let $\lambda \vdash m$ be any partition of m which contains μ: $(u^v) \subseteq \lambda$. Then the elements of the corresponding two–sided ideal $I_\lambda \subseteq FS_m$ are identities of A: $I_\lambda \subseteq \mathrm{Id}(A) \cap V_m$.

Proof. Since $m \leqslant 2n$, $(d-1)^{2m} \leqslant (d-1)^{4n}$, and by assumption $(d-1)^4 \leqslant \frac{n}{u+v} \cdot \frac{2}{e}$. By Lemma 3.11, $\left(\frac{n}{u+v} \cdot \frac{2}{e}\right)^n < s^\mu$ and since $\mu \subseteq \lambda$, we know that $s^\mu \leqslant s^\lambda$. Thus, by Remark 3.12,

$$c_m(A) \leqslant (d-1)^{2m} \leqslant (d-1)^{4n} \leqslant \left(\frac{uv}{u+v} \cdot \frac{2}{e}\right)^n < s^\mu \leqslant s^\lambda,$$

and the assertion now follows from Lemma 3.8. □

Corollary 3.14. Hypotheses as in Proposition 3.13, for $n \leqslant m \leqslant 2n$,

$$\bigoplus_{\substack{\lambda\vdash m\\ \mu\subseteq\lambda}} I_\lambda \subseteq \mathrm{Id}(A).$$

Consequently, if $f \in I_\mu$ then $f^* \in \mathrm{Id}(A) \cap V_{2n-1}$ (see (26)). Also, for any subset $L \subseteq \{n+1, \ldots, 2n-1\}$, $f^*_L \in \mathrm{Id}(A)$, and in particular $f \in \mathrm{Id}(A)$.

Proof. By "branching," the two–sided ideal generated in V_m by I_μ is

$$V_m I_\mu V_m = (FS_m) I_\mu (FS_m) = \bigoplus_{\substack{\lambda \vdash m \\ \mu \subseteq \lambda}} I_\lambda.$$

Hence, $(FS_m)I_\mu(FS_m) \subseteq \mathrm{Id}(A)$ for any $n \leqslant m \leqslant 2n-1$, and in particular, if $f \in I_\mu$ and $\rho \in S_m$ then $f\rho \in \mathrm{Id}(A)$. (26) concludes the proof. □

By Proposition 3.13 and Lemma 3.10 we have just proved

Proposition 3.15. Every PI algebra in characteristic 0 satisfies non-trivial strong identities. Explicitly, let $\mathrm{char}(F) = 0$ and let A satisfy an identity of degree d. Let u, v be natural numbers such that $\frac{uv}{u+v} \cdot \frac{2}{e} \geqslant (d-1)^4$, and let $\mu = (u^v)$ be the $u \times v$ rectangle. Then every $g \in I_\mu$ is a strong identity of A. The degree of such a strong identity g is uv. We can choose for example $u = v = \lceil e \cdot (d-1)^4 \rceil$, so $\deg(g) = \lceil e \cdot (d-1)^4 \rceil^2 = e^2(d-1)^8$.

We summarize:

Theorem 3.16. Every affine PI algebra over a field of characteristic 0 satisfies some Capelli identity. Explicitly, we have the following:

(a) Suppose the F-algebra A satisfies an identity of degree d. Then A satisfies a strong identity of degree

$$d' = \lceil e(d-1)^4 \rceil^2 = e^2(d-1)^8.$$

(b) Suppose $A = F\{a_1, \ldots, a_r\}$, and A satisfies an identity of degree d and take d' as in (a). Let $n = r^{d'} + d' \approx r^{e^2(d-1)^8}$. Then A satisfies the Capelli identity Cap_n.

Proof. (a) is by Proposition 3.15, and then (b) follows from Theorem 3.3, since every strong identity is sparse. □

3.4. Actions of the group algebra on sparse identities

Although the method of §3.2 is the one customarily used in the literature, it does rely on branching and thus only is effective in characteristic 0. A slight modification enables us to avoid branching. The main idea is that any sparse identity follows from an identity of the form

$$f = \sum_{\sigma \in S_n} \alpha_\sigma x_{\sigma(1)} x_{n+1} \cdots x_{\sigma(n)} x_{2n},$$

since we could then specialize $x_{n+1}, \ldots, x_{2n}$ to whatever we want. Thus, letting V_n' denote the subspace of V_{2n} generated by the words $x_{\sigma(1)} x_{n+1} \cdots x_{\sigma(n)} x_{2n}$, we can identify the sparse identities with $F[S_n]$-subbimodules of V_n' inside V_{2n}. But there is an as $F[S_n]$-bimodule isomorphism $\varphi : V_n \to V_n'$, given by $x_{\sigma(1)} \cdots x_{\sigma(n)} \to x_{\sigma(1)} x_{n+1} \cdots x_{\sigma(n)} x_{2n}$. In particular V_n' has the same simple $F[S_n]$-subbimodules structure as V_n and can be studied with the same Young representation theory, although now we only utilize the left action of permutations.

Thus, for any PI-algebra A, the spaces

$$\mathrm{Id}(A) \cap V_n' \subseteq V_n'$$

are $F[S_n]$-subbimodules of V_n'.

REMARK 3.17. Again, any tableau T of $2n$ boxes gives rise to an element

$$a_T = \varphi\left(\sum_{q \in \mathcal{C}_{T_\lambda},\ p \in \mathcal{R}_{T_\lambda}} \operatorname{sgn}(q) qp\right) \in F[S_{2n}],$$

where $\mathcal{C}_{T_\lambda}$ (resp. $\mathcal{R}_{T_\lambda}$) denotes the set of column (resp. row) permutations of the tableau T_λ.

Thus, $F[S_n]a_T$ (if nonzero) is an $F[S_n]$-submodule, which we call J_λ. If J_λ contains an element corresponding to a nontrivial PI of A, a_T itself must correspond to a PI of A.

We let I_λ denote the minimal $F[S_n]$-bisubmodule of $F[S_{2n}]$ containing J_λ.

LEMMA 3.18. Let A be an F-algebra, and let λ be a partition of n. If dim $J_\lambda > c_{2n}(A)$ and J_λ is a simple $F[S_n]$-module, then $I_\lambda \subseteq \operatorname{Id}(A) \cap V_n'$.

PROOF. Same as Lemma 3.7, noting that I_λ is a sum of $F[S_n]$-submodules $J_\lambda a$ each isomorphic to J_λ. Thus, taking such J, one has

$$c_n(A) = \dim\left(\frac{V_n'}{\operatorname{Id}(A) \cap V_n'}\right) \geqslant \dim J > c_{2n}(A),$$

a contradiction. Therefore each $J \subseteq \operatorname{Id}(A) \cap V_n'$, implying $I_\lambda \subseteq \operatorname{Id}(A) \cap V_n$. □

Note that when $\operatorname{char}(F) = p > 0$, the lemma might fail unless J_λ is simple. James and Mathas [13, Main Theorem] determined when J_λ is simple for $p = 2$.

One such example is when λ is the **staircase**, which we define to be the Young tableau T_u whose u rows have length $u, u-1, \ldots, 1$. This gave rise to the James-Mathas conjecture [21] of conditions on λ characterizing when J_λ is simple in characteristic $p > 2$, which was solved by Fayers [9].

4. Kemer's Capelli Theorem for all characteristics

In this section we give a proof of Kemer's "Capelli Theorem" over a field of any characteristic. In fact in characteristic p Kemer proved a stronger result, even for non-affine algebras.

THEOREM 4.1. [17] Any PI algebra over a field F of characteristic $p > 0$ satisfies a Capelli identity Cap_n for large enough n.

This fails in characteristic 0, since the Grassmann algebra does not satisfy a Capelli identity. The proof of Theorem 4.1 given in [17] is quite complicated; an elementary proof using the "identity of algebraicity" is given in [7, §2.5.1], but still requires some computations. In the spirit of providing a full exposition which is as direct as possible, we treat only the affine case via representation theory, in which case characteristic $p > 0$ works analogously to characteristic 0. This produces a much better estimate of the degree of the sparse identity, which we obtain in Theorem 4.4.

In view of Theorem 3.3, it suffices to show that any affine PI algebra satisfies a sparse identity. Although we cannot achieve this through branching, the ideas of the previous section still apply, using [9].

4.1. Simple Specht modules in characteristic $p > 0$

In order to obtain a p-version of Proposition 3.13 in characteristic $p > 2$, first we need to find a class of partitions satisfying Fayer's criterion.

For a positive integer m, define v_p to be the p-adic valuation, i.e., $v_p(m)$ is the largest power of p dividing m. Also, temporarily write $h_{(i;j)}$ for h_x where x is the box in the i, j position. The

James-Mathas conjecture for $p \neq 2$, proved in [9], is that J_λ is simple if and only if there do not exist i, j, i', j' for which $v_p(h_{(i;j)}) > 0$ with $v_p(h_{(i;j)}), v_p(h_{(i';j)}), v_p(h_{(i;j')})$ all distinct. Of course this is automatic when each hook number is prime to p, since then every $v_p(h_{(i;j)}) = 0$.

Example 4. *A* ***wide staircase*** *is a Young tableau T_u whose u rows have all have lengths different multiples of $p-1$, the first row of length $(p-1)u$, the second of length $(p-1)(u-1)$, and so forth until the last of length $p-1$. The number of boxes is*

$$n = \sum_{j=1}^{u} (p-1)j = (p-1)\binom{u+1}{2}.$$

When $p = 2$, the wide staircase just becomes the staircase described earlier.

In analogy to Example 3, the dimension of the "wide staircase" T_u can be estimated as follows: We write $j = (p-1)j' + j''$ for $1 \leqslant j'' \leqslant p-1$. The hook of a box in the (i, j) position has length $(u+1-i)(p-1)+1-j$, and depth $u+1-j'-i$, so the hook number is

$$(u+1-i)(p-1)+1-j+u-j'-i = (u+1-i)p-j-j' = (u+1+j'-i)p-j'',$$

which is prime to p. Thus each wide staircase satisfies a stronger condition than Fayer's criterion.

The dimension can again be calculated by means of the hook formula. The first $p-1$ boxes in the first row have hook numbers

$$pu-1, pu-2, \ldots, pu-(p-1),$$

whose sum is $(p-1)pu - \binom{p}{2} = \binom{p}{2}(2u-1)$.

The next $p-1$ boxes in the first row have hook numbers

$$p(u-1)-1, p(u-1)-2, \ldots, p(u-1)-(p-1),$$

whose sum is $(p-1)p(u-1) - \binom{p}{2} = \binom{p}{2}(2u-3)$.

Thus the sum of the hook numbers in the first row is

$$\binom{p}{2}((2u-1)+(2u-3)+\cdots+1) = \binom{p}{2}u^2.$$

Summing over all rows yields

$$\sum h_x = \binom{p}{2}\sum_{k=1}^{u} k^2 = \binom{p}{2}\frac{u(u+1)(2u+1)}{6} = \binom{p}{2}\frac{(2u+1)n}{3}.$$

Lemma 4.2. For any integer u, let μ be the wide staircase T_u of u rows. Let $n = (p-1)\binom{u}{2}$. Then

$$\left(\frac{6n}{p(p-1)(2u+1)}\right)^n \cdot \left(\frac{1}{e}\right)^n < f^\mu \qquad \text{(where} \quad e = 2.718281828\ldots).$$

In particular, if $\alpha \leqslant \frac{3n}{(2u+1)e}$, then $\alpha^n \leqslant f^\mu$.

Proof. We imitate the proof of Lemma 3.11. Since the geometric mean is bounded by the arithmetic mean,

$$\left(\prod_{x\in\mu} h_x\right)^{1/n} \leqslant \frac{1}{n}\sum_{x\in\mu} h_x \leqslant \binom{p}{2}\frac{(2u+1)n}{3} = \frac{p(p-1)(2u+1)n}{6},$$

in view of Example 4, together with $(n/e)^n < n!$, implies that

$$\left(\frac{6n}{p(p-1)(2u+1)}\right)^n \cdot \left(\frac{1}{e}\right)^n = \left(\frac{n}{e}\right)^n \cdot \left(\frac{6}{p(p-1)(2u+1)}\right)^n < \frac{n!}{\prod_{x\in\mu} h_x} = f^\mu.$$

□

Lemma 4.3. Let A be a PI algebra over a field of characteristic p, that satisfies an identity of degree d. Choose a natural number u such that, for $n=(p-1)\binom{u+1}{2}$,

$$\frac{6n}{p(p-1)(2u+1)}\cdot\frac{1}{e}\geqslant(d-1)^2.$$

Let $\lambda\vdash n$ be any partition of n corresponding to the "wide staircase" T_u. Then the elements of the corresponding $F[S_n]$-bimodule $I_\lambda\subseteq V_n'$ are sparse identities of A.

Proof. By Remark 3.12,

$$c_{2n}(A)\leqslant(d-1)^{4n}\leqslant\left(\frac{6n}{p(p-1)(2u+1)}\cdot\frac{1}{e}\right)^n<s^\lambda,$$

and we conclude from Lemma 3.18. □

4.1.1. Existence of Capelli identities

We are ready for a version of Proposition 3.15.

Theorem 4.4. [17] Any PI- algebra A over a field F of characteristic $p>0$ satisfies a Capelli identity. Explicitly:

(a) Suppose the F-algebra A satisfies an identity of degree d. Then A satisfies a sparse identity of degree $d'=(p-1)p\binom{u+1}{2}$, where $\frac{3u(u+1)}{p(2u+1)}\geqslant(d-1)^2e$.

(b) Suppose $A=F\{a_1,\dots,a_r\}$, and A satisfies an identity of degree d and take d' as in (a). Let $n=r^{d'}+d'\approx r^{4e^2(d-1)^4}$. Then A satisfies the Capelli identity Cap_n.

Proof. (a) is by Lemma 4.3. Then (b) follows from Theorem 3.3. □

For example, since $\frac{u+1}{2u+1}\geqslant\frac{1}{2}$, we could take $u\geqslant\frac{2pe(d-1)^2}{3}$.

This concludes the proof of Theorem 4.1 in the affine case.

5. Results and proofs over Noetherian base rings

We turn to the case where C is a commutative Noetherian ring. In general, we say a C-algebra is PI if it satisfies a polynomial identity having at least one coefficient equal to 1. Let us indicate the modifications that need to be made in order to obtain proofs of Theorems 1.6 and 1.7.

The method of proof of Theorem 1(2) (for the case in which the base ring C is a field) was to verify the "weak Nullstellensatz", and a similar proof works for A commutative when C is Jacobson, cf. [26, Proposition 4.4.1]. Thus we have Theorems 1.6 and 1.7 in the commutative case, which provide the base for our induction to prove Theorem 1.3. The argument is carried out using Zubrilin's methods (which were given over an arbitrary commutative base ring.)

It remains to find a way of proving Kemer's Capelli Theorem over arbitrary Noetherian base rings. One could do this directly using Young diagrams, but there also is a ring-theoretic reduction. The following observations about Capelli identities are useful.

Lemma 5.1. (i) Suppose $n=n_1n_2\cdots n_t$. If A satisfies the identity $\mathrm{Cap}_{n_1}\times\cdots\times\mathrm{Cap}_{n_t}$, then A satisfies the Capelli identity Cap_n.

(ii) If $I\triangleleft A$ and A/I satisfies Cap_m for m odd, with $I^k=0$, then A satisfies Cap_{km}.

(iii) If $I\triangleleft A$ and A/I satisfies Cap_m with $I^k=0$, then A satisfies $\mathrm{Cap}_{k(m+1)}$.

Proof. (i) Viewing the symmetric group $S_{n_1} \times \cdots \times S_{n_m} \hookrightarrow S_n$, we partition S_n into orbits under the subgroup $S_{n_1} \times \cdots \times S_{n_m}$ and match the permutations in Cap_n.

(ii) This time we note that any interchange of two odd-order sets of letters has negative sign, so we partition S_{km} into k parts each with m letters.

(iii) Any algebra satisfying Cap_m for m even, also satisfies Cap_{m+1}, and $m+1$ is odd. □

Thus, it suffices to prove that A satisfies a product of Capelli identities.

Theorem 5.2. Any affine PI algebra over a commutative Noetherian base ring C satisfies some Capelli identity.

Proof. By Noetherian induction, we may assume that the theorem holds for every affine PI-algebra over a proper homomorphic image of C.

First we do do the case where C is an integral domain, and $A = C\{a_1, \dots, a_\ell\}$ satisfies some multilinear PI f. It is enough to assume that A is the relatively free algebra $C\{x_1, \dots, x_n\}/I$ (where I is the T-ideal generated by f). Let F be the field of fractions of C. Then $A_F := A \otimes_C F$ is also a PI-algebra, and thus, by Theorem 4.1 satisfies some Capelli identity $f_1 = \mathrm{Cap}_n$. Thus the image $\bar{f}_1$ of f_1 in A becomes 0 when we tensor by F, which means that there is some $s \in C$ for which $sf_1 = 0$. Letting I' denote the T-ideal of A generated by the image of f_1, we see that $sI' = 0$. If $s = 1$ then we are done, so we may assume that $s \in C$ is not invertible. Then A/sA is an affine PI-algebra over the proper homomorphic image C/sC of C, and by Noetherian induction, satisfies some Capelli identity Cap_m, so $A/(sA \cap I')$ satisfies $\mathrm{Cap}_{\max\{m,n\}}$. But $sA \cap I'$ is nilpotent modulo $sAI' = AsI' = 0$, implying by Lemma 5.1 that A satisfies some Capelli identity.

For the general case, the nilpotent radical N of C is a finite intersection $P_1 \cap \cdots \cap P_t$ of prime ideals. By the previous paragraph, A/P_jA, being an affine PI-algebra over the integral domain C/P_j, satisfies a suitable Capelli identity Cap_{n_j}, for $1 \leqslant j \leqslant t$, so $A/\cap(P_jA)$ satisfies Cap_n, where $n = \max\{n_1, \dots, n_t\}$. But $\cap(P_jA)$ is nilpotent modulo NA, so, by Lemma 5.1, A/NA satisfies a suitable Capelli identity Cap_n. Furthermore, $N^m = 0$ for some m, implying again by Lemma 5.1 that A satisfies C^{mn}. □

СПИСОК ЦИТИРОВАННОЙ ЛИТЕРАТУРЫ

1. S.A. Amitsur, A generalization of Hilbert Nullstellensatz, Proc. Amer. Math. Soc. **8** (1957) 649-656.

2. S.A. Amitsur, A note on P.I. rings, Israel J. Math. **10** (1971) 210–211.

3. S.A. Amitsur and C. Procesi, Jacobson rings and Hilbert algebras with polynomial identities, Ann. Mat. Pura Appl. **71** (1966) 67–72.

4. S.A. Amitsur and A. Regev: P.I. algebras and their cocharacters, J. of Algebra **78** (1982) 248–254.

5. S.A. Amitsur and L. Small, *Affine algebras with polynomial identities,* Supplemento ai Rendiconti del Circolo Matematico di Palermo **31** (1993).

6. A. Belov, L. Bokut, L.H. Rowen, and J.T. Yu, *The Jacobian Conjecture, together with Specht and Burnside-type problems*, Proc. Groups of Automorphisms in Birational and Affine Geometry, Springer, editors M.Zaidenberg, M. Rich, and M. Reizakis, to appear.

7. A.K. Belov and L.H. Rowen, *Computational aspects of Polynomial Identities,* A. K. Peters (2005).

8. A. Braun, The nilpotency of the radical in a finitely generated PI-ring, J. Algebra **89** (1984), 375-396.

9. M. Fayers, Irreducible Specht modules for Hecke algebras of type A, Advances in Math. **193** (2005), 438–452.

10. M. Fayers, S. Lyle, S. Martin, p-restriction of partitions and homomorphisms between Specht modules, J. Algebra **306** (2006), 175–190.

11. N. Jacobson, *Basic Algebra II*, second edition, Freeman and company (1989).

12. G. D. James, *The Representation Theory of the Symmetric Groups*, Lecture Notes in Math, Vol. 682, Springer–Verlag, New York, NY, (1978).

13. G. James and A. Mathas, *The irreducible Specht modules in characteristic* 2, Bull. London Math. Soc. **31** (1999), 457–62.

14. G. D. James and A. Kerber, *The Representation Theory of the Symmetric group*, Encyclopedia of Mathematics and its Applications, Vol. 16, Addison–Wesley, Reading, MA, (1981).

15. A.R Kemer, Capelli identities and nilpotence of the radical of a finitely generated PI-algebra, Dokl. Akad. Nauk SSSR **255** (1980), 793-797 (Russian). English translation: Soviet Math. Dokl. **22** (1980), 750-753.

16. A.R. Kemer, Ideals of identities of associative algebras, Amer. Math. Soc. Translations of monographs **87** (1991).

17. A.R. Kemer, Multilinear identities of the algebras over a field of characteristic p, Internat. J. Algebra Comput. **5** no. 2, (1995), 189–197.

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

19. L'vov, Unpublished (Russian).

20. I.G. Macdonald, *Symmetric Functions and Hall Polynomials*, 2nd edition, Oxford University Press, Oxford, (1995).

21. A. Mathas, *Iwahori–Hecke algebras and Schur algebras of the symmetric group*, University Lecture Series 15, American Mathematical Society, Providence, RI (1999).

22. Yu.P. Razmyslov, The Jacobson radical in PI-algebras (Russian), Algebra i Logika **13** (1974), 337-360. English translation: Algebra and Logika **13** (1974), 192-204.

23. A. Regev, Existence of identities in $A \otimes B$, Israel J. Math. **11** (1972), 131–152.

24. A. Regev, The representations of S_n and explicit identities for P.I. algebras, J. Algebra **51** (1978), 25–40.

25. Richard Resco, Lance W. Small and J. T. Stafford, Krull and Global Dimensions of Semiprime Noetherian PI-Rings Transactions of the American Mathematical Society **274, No. 1** (1982), 285–295

26. L.H. Rowen, *Polynomial Identities in Ring Theory*, Pure and Applied Mathematics, 84. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, (1980).

27. L.H. Rowen, *Ring Theory*, Vol. II. Pure and Applied Mathematics, 128. Academic Press, Inc., Boston, MA, (1988).

28. L.H. Rowen, *Graduate algebra: Commutative View*, AMS Graduate Studies in Mathematics **73**, 2006.

29. L.H. Rowen, *Graduate algebra: Noncommutative View*, AMS Graduate Studies in Mathematics **91**, 2008.

30. B.E. Sagan, *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions*, 2nd edition, Graduate Texts in Mathematics 203, Springer-Verlag (2000).

31. A.I. Shirshov, On certain non associative nil rings and algebraic algebras (Russian), Mat. Sb. **41** (1957), 381-394.

32. A.I. Shirshov, On rings with identity relations (Russian), Mat. Sb. **43** (1957), 277-283.

33. Small, L.W., *An example in PI rings*, J. Algebra **17** (1971), 434–436.

34. K.A. Zubrilin, Algebras satisfying Capelli identities, Sbornic Math. **186** no. 3 (1995) 359-370.

REFERENCES

1. S.A. Amitsur, A generalization of Hilbert Nullstellensatz, Proc. Amer. Math. Soc. **8** (1957) 649-656.

2. S.A. Amitsur, A note on P.I. rings, Israel J. Math. **10** (1971) 210–211.

3. S.A. Amitsur and C. Procesi, Jacobson rings and Hilbert algebras with polynomial identities, Ann. Mat. Pura Appl. **71** (1966) 67–72.

4. S.A. Amitsur and A. Regev: P.I. algebras and their cocharacters, J. of Algebra **78** (1982) 248–254.

5. S.A. Amitsur and L. Small, *Affine algebras with polynomial identities,* Supplemento ai Rendiconti del Circolo Matematico di Palermo **31** (1993).

6. A. Belov, L. Bokut, L.H. Rowen, and J.T. Yu, *The Jacobian Conjecture, together with Specht and Burnside-type problems*, Proc. Groups of Automorphisms in Birational and Affine Geometry, Springer, editors M.Zaidenberg, M. Rich, and M. Reizakis, to appear.

7. A.K. Belov and L.H. Rowen, *Computational aspects of Polynomial Identities,* A. K. Peters (2005).

8. A. Braun, The nilpotency of the radical in a finitely generated PI-ring, J. Algebra **89** (1984), 375-396.

9. M. Fayers, Irreducible Specht modules for Hecke algebras of type A, Advances in Math. **193** (2005), 438–452.

10. M. Fayers, S. Lyle, S. Martin, p-restriction of partitions and homomorphisms between Specht modules, J. Algebra **306** (2006), 175–190.

11. N. Jacobson, *Basic Algebra II*, second edition, Freeman and company (1989).

12. G. D. James, *The Representation Theory of the Symmetric Groups*, Lecture Notes in Math, Vol. 682, Springer–Verlag, New York, NY, (1978).

13. G. James and A. Mathas, *The irreducible Specht modules in characteristic* 2, Bull. London Math. Soc. **31** (1999), 457–62.

14. G. D. James and A. Kerber, *The Representation Theory of the Symmetric group*, Encyclopedia of Mathematics and its Applications, Vol. 16, Addison–Wesley, Reading, MA, (1981).

15. A.R Kemer, Capelli identities and nilpotence of the radical of a finitely generated PI-algebra, Dokl. Akad. Nauk SSSR **255** (1980), 793-797 (Russian). English translation: Soviet Math. Dokl. **22** (1980), 750-753.

16. A.R. Kemer, Ideals of identities of associative algebras, Amer. Math. Soc. Translations of monographs **87** (1991).

17. A.R. Kemer, Multilinear identities of the algebras over a field of characteristic p, Internat. J. Algebra Comput. **5** no. 2, (1995), 189–197.

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

19. L'vov, Unpublished (Russian).

20. I.G. Macdonald, *Symmetric Functions and Hall Polynomials*, 2nd edition, Oxford University Press, Oxford, (1995).

21. A. Mathas, *Iwahori–Hecke algebras and Schur algebras of the symmetric group*, University Lecture Series 15, American Mathematical Society, Providence, RI (1999).

22. Yu.P. Razmyslov, The Jacobson radical in PI-algebras (Russian), Algebra i Logika **13** (1974), 337-360. English translation: Algebra and Logika **13** (1974), 192-204.

23. A. Regev, Existence of identities in $A \otimes B$, Israel J. Math. **11** (1972), 131–152.

24. A. Regev, The representations of S_n and explicit identities for P.I. algebras, J. Algebra **51** (1978), 25–40.

25. Richard Resco, Lance W. Small and J. T. Stafford, Krull and Global Dimensions of Semiprime Noetherian PI-Rings Transactions of the American Mathematical Society **274, No. 1** (1982), 285–295

26. L.H. Rowen, *Polynomial Identities in Ring Theory*, Pure and Applied Mathematics, 84. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, (1980).

27. L.H. Rowen, *Ring Theory*, Vol. II. Pure and Applied Mathematics, 128. Academic Press, Inc., Boston, MA, (1988).

28. L.H. Rowen, *Graduate algebra: Commutative View*, AMS Graduate Studies in Mathematics **73**, 2006.

29. L.H. Rowen, *Graduate algebra: Noncommutative View*, AMS Graduate Studies in Mathematics **91**, 2008.

30. B.E. Sagan, *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions*, 2nd edition, Graduate Texts in Mathematics 203, Springer-Verlag (2000).

31. A.I. Shirshov, On certain non associative nil rings and algebraic algebras (Russian), Mat. Sb. **41** (1957), 381-394.

32. A.I. Shirshov, On rings with identity relations (Russian), Mat. Sb. **43** (1957), 277-283.

33. Small, L.W., *An example in PI rings*, J. Algebra **17** (1971), 434–436.

34. K.A. Zubrilin, Algebras satisfying Capelli identities, Sbornic Math. **186** no. 3 (1995) 359-370.

Получено 16.05.2020 г.
Принято в печать 22.10.2020 г.