POLYNOMIAL IDENTITIES

BY

S. A. AMITSUR

ABSTRACT

This article contains two surveys: (A) A historical survey of the early literature of 1920–1950 in Polynomial Identities which follows the roots and sources of PI-rings. (B) A survey of the methods of constructing identities for matrix rings, since the identity of Wagner until the central identities of Formanek and Razmyslov and the rational identity of Bergman.

A. Early History

1. *Introduction*

A PI-ring is a ring satisfying a polynomial identity $p[x_1, x_2, \cdots, x_n] = 0$ for all substitutions of the indeterminates x_i's by elements of R. It is commonly thought that the theory of PI-rings is a new chapter in the theory of associative rings and usually the survey of results on PI-rings starts with Kaplansky's paper [12] of 1948. This is absolutely wrong and the first paper in this area, known to the author, was written as early as 1922 by Dehn [1]. The modern approach was founded by Kaplansky in 1948, based upon two forerunners, namely the methods introduced by Jacobson [7] and Levitzki [9]. The first period in PI-ring theory are the years 1922–1948; the second period lasted until recently when the new central identities were founded by Formanek and Razmyslov ([25], [26]) and a new powerful tool was introduced into PI-theory, and the theory took a new turn. In this survey, we go through the first period.

Comparing the roots and the early problems of PI-rings with its recent achievements in solving old problems on division rings, one observes that these solutions could be obtained only after ring theory passed its various stages of generalizations: from division rings to zero divisors, from artinian rings to no restrictions on chain conditions and to various radicals. It is only at the last stages of the generalizations where the tools were made available to be used in connection with

PI-rings and the old problems could be solved. Part of this process is recorded here. The references at the end of this paper are listed chronologically for easy follow-up. The list may not be complete, as no serious attempt has been made to check all early literature. The list is restricted only to papers dealing with polynomial relations in associative rings and not in other algebraic structures (Birkhoff), groups (Neuman), etc.

There are three main old roots to the theory of polynomial identities, and two minor supporters. Listed in sequence of importance, as well as their historical appearance, these are: the foundation of projective geometry, theory of equations, commutativity of rings; and the two minor supporters are: the Lie structure of rings and the relatively recent and less important contribution obtained by generalizing Boolean algebras to rings satisfying the relations $px = 0$, $x^p - x = 0$ (Foster).

2. *Foundation of projective geometry*

One can recognize the seeds of PI-rings in the von-Staudt method of the last century of introducing homogeneous coordinates from a division ring in a projective plane, and later in Hilbert's example of a Desarguian geometry with the order axioms but for which the Pappus theorem fails. Hilbert's famous example was obtained by translating the problem from geometry into algebra, to the coordinate division ring and there it is solved by constructing an ordered division which is non-commutative. This led to Dehn's paper [1]. He was looking for intersection theorems in Desarguian and non-Pappian geometries.

The three main theorems in (linear) projective geometry are the complete quadrilateral theorem, the Desargue theorem, and the Pappus theorem. Dehn's attention was drawn to the gap between the Desargue theorem and the Pappus theorem.† Note that the presence of the Desargue theorem in geometry enables the introduction of homogeneous coordinates from a division ring D, hence any additional intersection theorem is equivalent to a (rational) identity for D e.g., the Pappus theorem is equivalent to the commutative law $x_1x_2 - x_2x_1 = 0$. Dehn proved in [1] that any theorem leading to a polynomial relation $\sum \alpha_{i_k} x_1^i x_2 x_1^k = 0$ (in characteristic zero) in D implies commutativity and that means that the Pappus theorem holds in this geometry. He also proved that the

† The gap between the complete quadilateral theorem and the Desargue theorem was dealt with years later by R. Moufang and her school.

same result holds for ordered Desarguian geometry (which is equivalent to the fact that D is ordered) with a theorem equivalent to a relation $p[x_1, x_2] = 0$ which is of degree $\leqq 3$ in x_1 and of degree $\leqq 2$ in x_2, i.e., D is commutative. Dehn's paper of 1922 is the first paper known to the author in polynomial identities in associative rings.

Though not stated explicitly, it seems that Dehn had the intention of showing that ordered Desarguian geometry satisfies the Pappus theorem in the presence of any other intersection theorem which does not follow from the Desargue theorem — a result which was proved many years later ([22] 1966). The partial result of Dehn led to the next paper in PI, which is Wagner's paper [4] in 1936. It took fourteen years until the first paper, which has all the flavors of a modern paper in polynomial identities, appeared. Wagner's starting point was again the foundation of projective geometry — the gap between the Desargue and the Pappus theorems, but very quickly he passed to pure algebra. In fact, he proved the first theorem in PI-rings: "An ordered ring which satisfies a polynomial identity is commutative" which extended the result of Dehn. Years later, Albert gave a very short structural proof to this theorem. Wagner's paper contained other pioneering results: the identity $[[x_1, x_2]^2, x_3] = 0$ for $M_2(F)$ where $[x_1, x_2] = x_1x_2 - x_2x_1$ and the first polynomial identity for $M_n(F)$, which will be quoted and reproved in the next part of this article. Also, the method of using generic matrices, which was rediscovered years later and became an important tool in PI-rings, appeared already in Wagner's paper [4].

Next in the line of PI-results rooted in geometry is M. Hall's paper [5] in 1943. The long time gap between the three papers mentioned, we attribute to the lack of general abstract theory of rings at that time. The paper by Hall, dealt mainly with geometry but among the many results there were a few which were the heralds of modern PI-theory. First, it was shown that a non-commutative division ring satisfying the identity $[[x_1, x_2]^2, x_3] = 0$ is a generalized quaternion, namely of dimension 4 over its center; and second, Hall raises the question whether a division ring satisfying the relation $x^{p^n} - x = 0$ is necessarily commutative. Both appeared in trying to determine Desarguian geometries which satisfy configuration theorems generated by 5 and 6 points. The second problem was connected also with the Burnside problem for finitely generated periodic groups, and of course with Wedderburn's theorem for finite division rings. Both the theorem of Hall and the question raised were important steps in the theory of polynomial identities.

3. *Rise of* PI-*rings*

Through the early years of the forties, following the structure theorems for artinian rings, algebraists have been looking for a general theory of rings independent of chain conditions. General nil rings and algebraic algebras became one of the main topics of associative rings. Jacobson's structure theory for associative rings with no chain conditions provided the right tools, and in fact he succeeded in [7] to solve Hall's problem and to show that a ring satisfying an identity $x^n - x = 0$ (and the general case, $n = n(x)$) is commutative, without referring to the Burnside conjecture, and thus gave the right extension to Wedderburn's theorem.

Another push to the PI-theory was given undoubtedly by the famous Kurosh problem [6]: "Are algebraic algebras locally finite?", and especially in its form raised independently by Levitzki: "Are nil rings locally nilpotent?" In [7] algebraic algebras of bounded degree were shown to have this property. Both the Kurosh and Levitzki problems, which still have no complete solution, were an incentive on the road of PI-rings. It was Kaplansky [8] who was first to prove that a nil algebra of bounded index; i.e., satisfying $x^n = 0$, over a field containing at least n elements is locally nilpotent. Kaplansky's proof was obtained by by an inductive combinatorial approach; about the same time Levitzki [9] gave a structural proof of this result for general rings: namely, rings satisfying $x^n = 0$ are locally nilpotent. The paper of Hall [5], the methods developed by Jacobson [7] and Levitzki in [9] were the parents of Kaplansky's famous paper [12] from which modern PI-theory has started.

In addition to proving that many rings, like algebraic algebras of bounded degree satisfy polynomial identities, Kaplansky proves in [12] the first basic structure theorems for PI-rings. Namely, (1) Primitive PI-rings are central simple algebras of finite dimension, and (2) Nil PI-rings are locally nilpotent. This was followed by Levitzki [16] and Kaplansky [15] in which Levitzki proved that the nil radical of a PI-ring is its lower radical and in fact can be obtained by a finite sequence of nilpotent radicals, and Kaplansky gave the (first proof) positive solution of the Kurosh problems for PI-algebras. This was the road of PI-theory which started from the foundation for projective geometry. But there were other roots which led to different aspects of PI-rings, and to them we pass now.

4. *Rank of commutativity (Roc)*

Another source of PI-theory was the attempt of generalizing the commutative law: F. W. Levi pointed out in [10] (1947) that the commutative law $x_1x_2 - x_2x_1 = 0$ has a natural generalization in the determinant identities: $S_n[x_1, \cdots, x_n] = \Sigma \pm x_{i_1}x_{i_2} \cdots x_{i_n} = 0$., where the sum ranges over all permutations with the plus sign for even permutations and the minus sign for odd permutations. These are known as the standard identities (or polynomials). A ring R was said to have rank of commutativity (Roc) $\leqq n$ if it satisfies $S_n[X]$, and Levi, who introduced this notion, proved that $\mathrm{Roc}[M_n(F)] \leqq n^2 - 1$. Independently, Kolchin had drawn to the attention of Kaplansky (see footnote in [14]) that the minimal r_n such that $S_{r_n}[X] = 0$ hold in $M_n(F)$ is bounded by $2n \leqq r_n \leqq n^2 + 1$. This immediately raised the conjecture that $r_n = 2n$, and indeed that was proved in [17]. This led also to the study of the sets of all identities satisfied by a given ring, e.g., by $M_n(F)$, a study which is, even today, far from any concrete results. A first attempt in this direction was done by Specht [18] in 1950, where identities of a general algebra were shown to constitute a T-ideal (invariant under all endomorphisms) in the free ring. Specht who introduced the notion of the T-ideal has tried to get a form of the bases of these ideals by higher commutators, when the algebras considered contain a unit element. The contribution of the Roc to the PI-theory was rather to the set of identities of a given ring, whereas the foundation of geometry contributed to the direction of structure theorems for PI-rings.

5. *Theory of equations*

A third source for PI-ring, of a very different nature can be found in a paper by A. R. Richardson [2] of 1928, and whose problems were later taken up by D. E. Littlewood [3] in 1931. Their point of view led to the notion of generalized polynomial identities, which surprisingly became an important tool in handling rational relations ([22]). The theory of equations is, in one form or another, the study of the roots of a polynomial equation $p[X]=0$ whose coefficient lies in a commutative field. Any generalization of a non-commutative division ring is faced with the complication of defining non-commutative polynomials which take the form $\Sigma a_{v_1}x^{v_1}a_{v_2}{}^{v_2} \cdots x^{v_n}a_{n+1} = p[x]$.

Richardson in [2] was the first to show that for noncommutative division rings one is led to the same situation as in finite fields; namely, that for some polynomials, [x] may vanish identically in the division ring and yet under any reasonable formal definition $p[x]$ is a non-zero polynomial, e.g., the quaternions satisfy the polynomial relation $(x^2i)+(xj)^2+(xk)^2-(ix)^2-(jx)^2-(kx)^2=0$ and other types of similar relation. The matrix ring $M_n(F)$ over a commutative field will satisfy $c_{pq}xc_{rs}xc_{uv}-c_{ps}xc_{uq}xc_{rv}=0$, where c_{ik} are the unit matrix units. The number of such independent relations is $(\frac{1}{2})n^4(n^2-1)$ and it was shown by D. E. Littlewood [3] that they form a set of generators of the vanishing polynomials for $M_n(F)$. A modern version of this result has been obtained by Procesi [24]. It was long after PI-theory was well established, that this direction was taken up and it actually led in [22] to obtain an almost complete solution of the very old project of Dehn of determining the intersection theorems between the Desargue and Pappus theorem. These theorems can be described by conditions on the characteristic and rank of commutativity (equivalently: finite dimension over the center).

6. *Lie algebras and Boolean algebras*

There were two other places where PI have been mentioned, but they had really very little effect on PI-theory. The recent one which is less important from the point of view of PI-rings, was Foster's extension of the notion of a Boolean algebra to rings with the relations $px=0$, $x^p-x=0$. The ideas in this case were mainly towards applications in logic. A more important contribution to PI-ring was made by Jennings [11] in 1947, who, with no intention of studying polynomial identities, tried to determine the structure of rings with conditions imposed on the Lie algebra structure of the ring. For example, if the Lie structure of R s nilpotent, he proved that the ideal generated by the commutators is nil, and in fact locally nilpotent, and that the ideal generated by the third order commutators $[[x,y]z]$ forms a nilpotent ideal. These are actually structure theorems for rings satisfying the higher commutator identity $[[\cdots[x_1,x_2],x_3\cdots],x_n]=0$. But as stated above, these results had no effect on the advancement of PI-rings, though the problem of commutator-nil ideal has been dealt with by various authors.

Summarizing, we see that although PI-rings are relatively old in modern mathematics, it is only abstractions and generalizations which provided the tools for handling them, and in turn PI-rings became powerful tools in solving old problems.

B. Identities of matrix rings

1. *Introduction*

One of the old problems in PI-rings is to determine the set of all identities of the matrix ring $M_n(F)$, or at least to find a set of generators of these relations.

In this part we survey the known identities and the methods of obtaining new identities for $M_n(F)$.

Let $F[x_1, x_2, \cdots, x_n, \cdots] = F[x]$ be the free ring generated by a countable number of indeterminates x_i's, and $\mathcal{M}_n = \mathcal{M}_n(F)$ be the set of all polynomials $p[x] = p[x_1, \cdots, x_m] \in F[x]$ such that $p[x] = 0$ is a polynomial identity for $M_n(F)$. It is readily seen that $\mathcal{M}_1 \supset \mathcal{M}_2 \supset \cdots \supset \mathcal{M}_n$, and each $\mathcal{M}_n$ is a two-sided T-ideal in $F[x]$, in the sense that if $p[x_1, \cdots, x_m] \in \mathcal{M}_n$ then $p[g_1[x], \cdots, g_m[x]] \in \mathcal{M}_n$ for all $g_i[x] \in F[x]$. Thus, together with an identity $p[x] = 0$ we get new identities of the forms: $\Sigma f[x]p[x]g[x] = 0$, $p[g_1[x], \cdots, g_m[x]] = 0$. Furthermore, if F is an infinite field then $\mathcal{M}_n$ is homogeneous; that is if $p[x] \in \mathcal{M}_n$ then the homogeneous components of $p[x]$ also belong to $\mathcal{M}_n$. Another point to notice is that $F[x]/\mathcal{M}_n$ is isomorphic with the ring $F[X] = F[X_1, \cdots, X_n, \cdots]$ generated by generic matrices $X_i = (\xi^i_{\lambda\mu})$ of order $n \times m$ ([23]), and therefore any relation in $F[X]$ can be translated to a polynomial in $\mathcal{M}_n$. $F[X]$ is an Ore domain whose ring of quotients is a central simple algebra of dimension n^2 over its center. Discovering relations among the elements of $F[X]$ and of central elements in this ring of quotient will yield identities in $M_n(F)$. In particular, such will be the new central polynomials ([25], [26]).

2. *Algebraicity and standard polynomials*

The main tool for obtaining polynomial identities for $M_n(F)$ is the use of the properties of the standard polynomial $S_n[x_1, \cdots, x_n] = \Sigma \pm x_{i_1}x_{i_2}\cdots x_{i_n}$ described in part A. The use of properties of $M_n(F)$, like bounded algebraicity (Cayley-Hamilton theorem), or its dimension over F, or the properties of the matrix units $\{e_{\lambda\mu}\}$-together with the properties of $S_n[x]$ will yield most of the identities. The properties of $S_n[x]$ are similar to the determinant; namely, it is homogeneous, multilinear, and alternative and therefore vanishes if all x_i's are substituted by linear dependent elements.

The following is probably the first known identity for $M_n(F)$. It was discovered in 1936 by Wagner [4]; a simplified proof will be given here:

$$S_m[xD_y^2, xD_y^4, \cdots, xD_y^{2m}] = 0 \text{ where } 2m > n(n-1), \tag{1}$$

$xD_y = [x, y] = xy - yx$ and $xD_y^{k+1} = (xD_y^k)D_y$.

Indeed, let Y be a generic matrix and consider the linear transformation $\mathrm{ad}Y$ defined by $x(\mathrm{ad}Y) = xY - Yx$ on the algebra of all $n \times n$ matrices over a large extension K of the field F. Identify $M_n(K)$ with a tensor product $V \otimes V$ of two

n-dim vector spaces over K, and identify the matrix units c_{ik} with $v_i \otimes v_k$ for a fixed base $\{v_j\}$ of V. Then the matrix corresponding to $\mathrm{ad}Y$ will be $1 \otimes Y - Y \otimes 1$ in $M_{n^2}(K)$. By diagonalizing Y, we get from the equality $PYP^{-1} = \mathrm{diag}(y_1, \cdots, y_j, \cdots, y_n)$ — the relation $(P \otimes P)(1 \otimes Y - Y \otimes 1)(P \otimes P)^{-1} = \mathrm{diag}(y_i - y_j)$. The minimal polynomial of this diagonal matrix, and hence also of $1 \otimes Y - Y \otimes 1$, is the polynomial

$$q[\lambda] = \lambda \prod_{i \neq j} [\lambda - (y_i - y_j)] = \lambda \prod_{i<j} [\lambda^2 - (y_i - y_j)^2] = \lambda[\lambda^{n(n-1)} + d_1 \lambda^{n(n-1)-2} + \cdots + + d_{\binom{n}{2}}].$$

Note that

$$d_{\binom{n}{2}} = \prod (y_i - y_j)^2 = \pm \delta$$

the discriminant of the characteristic polynomial of Y. In particular, this implies that for any two matrices X, Y we have the linear dependence

$$\sum_{\nu=1}^{\binom{n}{2}+1} d_\nu X(adY)^{2\nu} = 0,$$

and so the elements $xD_y^2, \cdots, xD_y^{2\nu}$ for $\nu > \binom{n}{2}$ $(xD_y = x(\mathrm{ad}Y))$ are always linears dependent and hence (1) follows from the properties of $S_m[x]$.

The preceding proof shows that we can replace in (1) the even powers D_y^{2i} by odd powers $D_y^{2\ +1}$ $0 \leqq i \leqq \binom{n}{2}$ and we will get a new identity for $M_n(F)$.

The identity (1) was obtained using the property that $M_n(F)$ is an algebraic algebra of degree n, which is a consequence of the Cayley Hamilton theorem. Other identities can be obtained using the same property. One of these is the second known identity for $M_n(F)$ which holds for all algebraic algebras of degree $\leqq n$, which was discovered by Kaplansky [12]:

Define the following polynomials inductively:

$$C_1(x_1, y) = x_1 y - x_1 y = [x_1, y] \quad C_{n+1}(x_1, \cdots, x_n, x_{n+1}; y)$$
$$= \quad [C_n(x_1, \cdots, x_{n-1}, x_{n+1}; y), C_n(x_1, \cdots, x_{n-1} x_n; y)]$$

Then in $M_n(F)$, as well as in any algebraic algebra of degree $\leqq n$.

(2) $$C_n(x, x^2, \cdots, x^n; y) = 0.$$

The proof of this relation is as follows: Let $x \in M_n(F)$ then it satisfies a polynomial $\sum_{\nu=0}^{n} \alpha_\nu x^\nu = 0$, $\alpha_n = 1$. Forming the commutator with y we get the relation $\sum_{\nu=1}^{n} \alpha_\nu C_1(x^\nu, y) = 0$. Repeating this process, that is commuting with $C_1(x, y)$ we get $\sum_{\nu=1}^{n} \alpha_\nu C_2(x^\nu, y) = 0$. Continuing, we finally get $\alpha_n C_n(x, x^2, \cdots, x^n; y) = 0$ which proves (2) since $\alpha_n = 1$.

In [21], the algebraicity of $M_n(F)$ was used to obtain another identity of $M_n(F)$:

$$S_{n+1}[y, xy, x^2y, \cdots, x^ny] = 0 \tag{3}$$

Indeed, the relation $\sum_{\nu=0}^{n} \alpha_\nu x^\nu = 0$ implies the linear dependence $\sum_{\nu=0}^{n} \alpha_\nu x^\nu y = 0$, and hence the properties of S_{n+1} imply (3). Note also that all the monomials of (3) end in y, and since $F[x]/\mathscr{M}_n$ is a domain, we can cancel the common factor y, i.e., $M_n(F)$ will satisfy $S_{n+1}[y, \cdots, x^ny]y^{-1} = \sum \pm x^{\nu_0}yx^{\nu_1}y \cdots yx^{\nu_2}$, where $(\nu_0, \cdots, \nu_n)$ ranges over all permutation of $(0, 1, 2, \cdots, n)$. Multiplying the relation $\sum \alpha_\nu x^\nu = 0$ on the left does not yield a new interesting identity since

$$S_{n+1}[y, yx, \cdots, yx^n] = yS_{n+1}[y, xy, \cdots, x^n]y^{-1}.$$

Another identity, based on the same principle is obtained by noting that the trace of a matrix $[x, y]$ in $M_n(F)$ is zero; hence it satisfies the polynomial relation $[x, y]^n + c_2[x, y]^{n-2} + \cdots + c_n = 0$, and thus the following identities will hold in $M_n(F)$, for every x, y, z:

$$S_n[z, [x,y]z, \cdots, [x,y]^{n-2}z, [x,y]^{n-1}z]z^{-1} = 0 \tag{3a}$$

$$C_{n-1}([x,y], [x, y]^2, \cdots, [x,y]^{n-2}, [x, y]^n; z) = 0 \tag{3b}$$

3. *Standard polynomials and dimension; minimal identities*

Next we turn to the use of the standard polynomials and the polynomials C_n together with the fact that $[M_n(F): F] = n^2$:

For any $n^2 + 1$ matrices $x_1, \cdots, x_{n^2}, x_{n^2+1}$ we have a linear dependence $\sum_{i=1}^{n^2+1} \alpha_i x_i = 0$ (say $\alpha_1 \neq 0$). Hence, following the same procedure as in proving (2) and (3) we will get the identity in M_n.

$$C_{n^2}(x_1, x_2, \cdots, x_{n^2}, x_{n^2+1}) = 0, \quad S_{n^2+1}[x_1, x_2, \cdots, x_{n^2+1}] = 0 \tag{4}$$

The second identity was first noted by Levi [10], and in [13] he proved that $M_n(F)$ satisfies $S_{n^2-1}[x] = 0$. The conjecture that the minimal r such that $S_r[x] = 0$ holds in $M_n(F)$ is $2n$ was proved in [17] by Amitsur-Levitzki. In fact it was shown that

(5) $S_{2n}[x_1, \cdots, x_{2n}] = 0$ is the unique minimal multilinear identity of $M_n(F)$. Any multilinear identity of $M_n(F)$ is a linear combination of polynomials $S_{2n}[x_{i_1}, \cdots, x_{i_{2n}}]$. The result is still true, even if the multilinearity assumption is dropped, with the exception of two cases for $M_n(F)$: $n = 1,2$ and $F = GF[2]$ is the finite field of two element [20]. The linear base of the minimal identities

for the field $GF[2]$ will be the polynomials $x_ix_j + x_jx_i$ and $x_i^2 + x_i$, and for $M_2(GF(2))$ will be the polynomial

(5a) $$xy^3 + yxy^2 + y^2xy + y^3x + xy^2 + y^2x$$

and all its linearizations, which include $S_4[x]$; they and all the substitutions by different x_i's in them constitute the base.

4. *Central elements in the ring of quotient* $F[x]$

A different method for obtaining identities, still using the properties of $S_n[x]$, was given in [21]. This method depends on identifying central elements in the quotient ring of the ring of generic matrices. Namely, if $p[x]q[x]^{-1} = c$ is a central element, then $M_n(F)$ will satisfy the identity (C) $p[x]zq[x] - q[x]zp[x] = 0$ for every substitution of the x's and z. The method we use to identify the elements c in the center is again obtained by using the properties of the standard polynomial in the following way:

Consider $n+1$ powers of x: $x^{v_0}, x^{v_n}, \cdots, x^{v}$; $v_0, v_1, \cdots, v_n$. Since x is algebraic of degree n (in the ring of generic matrices), there is a linear combination: $x^{v_n} + c_{n-1}x^{v_{n-1}} + \cdots + c_0x^{v_0} = 0$, with c_i belonging to the center of the quotient ring and those we can obtain as follows: Consider the polynomials $S_n[x^{v_0}y, x^{v_1}y, \cdots, \widehat{x^{v_i}y}, \cdots, x^{v_n}]$ where $\widehat{x^{v_i}y}$ means that this term is omitted. Since $x^{v_n}y = \sum_{i=0}^{n-1} c_ix^{v_i}y$ w: et, by substituting this relation in the left hand side, that

(6a) $$S_n[x^{v_0}y, x^{v_1}y, \cdots, \widehat{x^{v_i}y}, \cdots, x^{v_n}y] - c_iS_n[x^{v_0}y, x^{v_1}y, \cdots, x^{v_{n-1}}y] = 0.$$

From which we obtain by (C) the identity of $M_n(F)$:

(6) $$S_n[x^{v_0}y, \cdots, \widehat{x^{v_i}y}, \cdots, x^{v_n}y]zS_n[x^{v_0}y, \cdots, x^{v_{n-1}}y] - S_n[x^{v_0}y, \cdots, x^{v_{n-1}}y]zS_n[xy^{v_0}, , \cdots, \widehat{x^{v_i}y}, \cdots, x^{v_n}y] = 0$$

for $i = 0, \cdots, n-1$. Furthermore, it is well known that the coefficients c_i are symmetric polynomials in the characteristic roots of the polynomial of x. The basic symmetric functions in the roots can be obtained from (6) by taking $\{v_0, v_1, \cdots, v_n\} = \{0, 1, 2, \cdots, n\}$, since then the coefficients c_i are the coefficients of the characteristic polynomial and hence the symmetric functions in the roots. So let these functions be $S_{n-1} = P_iQ^{-1}$, where $Q = S_n[y, xy, \cdots, x^{n-1}y]$ and $\pm Q_i = S_n[y, xy, \cdots, \widehat{x^iy}, \cdots, x^ny]$. Thus, for some polynomial f_i we have the c_i of (6a) $c_i = f_i[s_1, s_2, \cdots, s_3]$. Let $\xi_0^m f[1, \xi, \xi_0^{-1}, \cdots, \xi_n\xi_0^{-1}] = \bar{f}[\xi_0, \xi_1, \cdots, \xi_n]$ be the corresponding homogeneous polynomial in commutative indeterminates ξ_i then noting that Q commutes with the P_i we get from (6) the relation:

(6b) $$Q^m S_n[x^{v_0}y,\cdots,\widehat{x^{v_i}y},\cdots,x^{v_n}_{n}y]-\bar{f}[Q,P_1,P_2,\cdots,P_n]S_n[x^{v_0}y,\cdots,x^{v_{n-1}}y]=0.$$

Incidently, our method shows how to give explicitly the symmetric functions of the characteristic roots of a generic X as quotient of two polynomials in two generic matrices, X, Y.

5. *Generalized standard polynomials*

We can repeat the procedure of the proceeding sections with the aid of any polynomial which has the same properties as the standard polynomial; namely homogeneous, multilinear, and vanishes of the set of x's are linearly dependent. Such polynomials are the following:

$$T_n[x_1,\cdots,x_n;\,y_1,y_2,\cdots,y_{n-1}] = \Sigma\,(\mathrm{sg}\sigma)x_{\sigma(1)}y_1x_{\sigma(2)}y_2\cdots y_{n-2}x_{\sigma(n-1)}y_{n-1}x_{\sigma(n)}$$

where the sum ranges over all the permutations σ of $(1,2,\cdots,n)$. These polynomials are obtained from S_n by interchanging the multiplication sign between the x's by a fixed sequence $y_1,\cdots,y_{n-1}$ of new indeterminates. It is not difficult to show that $T_n(x;y)$ as a function of x's has the same properties as the standard polynomial. Thus we can get as in (3) and (4) that $M_n(F)$ will satisfy the identities:

(7a) $$T_{n+1}[1,x,\ \cdots,\ x^n;\ y_1,\ \cdots,\ y_{n-1}]\ =0$$

(7b) $$T_{n_2+1}[x_1,x_2,\cdots,x_{n^2+1};\,y_1,\cdots,y_{n^2}]=0.$$

By using the algebraicity and the dimension of $M_n(F)$ over F, one can also get similar relations to (3a), (6a), (6), and (6b) by replacing S_n by T_n. Moreover since we will also have:

$$T_n[x^{v_0},x^{v_1},\cdots,\widehat{x^{v_i}},\cdots,x^{v_n};\,y_1,\cdots,y_{n-1}]-c_iT_n[x^{v_0},x^{v_1},\cdots,x^{v_{n-1}};\,y_n,\cdots,y_{n-1}]=0$$

for the same c_i we get the following complicated identities for matrices y,x,y_0,z_j: $i,\ j=1,2,\cdots,n-1$:

(7c) $$S_n[x^{v_0}y,\cdots,\widehat{x^{v_i}y},\cdots,x^{v_n}y]\,T_n[x^{v_0},x^{v_1},\cdots,x^{v_{n-1}};\,y_1,\cdots,y_{n-1}]-$$
$$-\,S_n[x^{v_0}y,\cdots,x^{v_{n-1}}y]\,T_n[x^{v_0},\cdots,\widehat{x^{v_i}},\cdots,x^{v_n};\,y_1,\ \cdots,y_{n-1}]=0$$

(7d) $$T_n[x^{v_0},\cdots,x^{v_{n-1}};y_1,\cdots,y_{n-1}]\,T_n[x^{v_0},\cdots,\widehat{x^{v_i}},\cdots,x^{v_n};\,z_1,\cdots,z_{n-1}]$$
$$-\,T_n[x^{v_0},\cdots,\widehat{x^{v_i}},\cdots,x^{v_n};\,y_1,\cdots,y_{n-1}]\,T_n[x^{v_0},\cdots,\widehat{x^{v_i}},\cdots,x^{v_{n-1}};z_1,\cdots,\ z_{n-1}]=0$$

6. *Central identities*

In the early days of PI-rings Kaplansky raised the question of the existence of central identities for $M_n(F)$, that is polynomials $p[x]=p[x_1,\cdots,x_n]$ such that $p[x] \not\equiv 0$ im $M_n(F)$ and for every substitution of matrices of $M_n(F)$, $p[x]$ attains only scalar values, i.e., $M_n(F)$ satisfies the identity $[p[x], z] = 0$. The first central identity was discovered by Formanek [26] and was immediately followed by Razmyslov [26]. Formanek's identity is mentioned in Goldies survey in this issue, and we recall the extension of his construction as given in [27].

Let $f[\xi_1, \xi_2, \cdots, \xi_{n+1}] = \sum \alpha_{(\nu)} \xi_1^{\nu_1} \xi_2^{\nu_2}, \cdots, \xi_{n+1}^{\nu_{n+1}}$ be a polynomial in $n+1$ commutative indeterminates; we transform $f[\xi]$ into apolynomial in non-commutative $p_f[x, y_1, \cdots, y_n] = \sum \alpha_{(\nu)} x^{\nu_1} y_1 x^{\nu_2} y_2, \cdots, y_n x^{\nu_{n+1}}$, and we get also the polynomial q_f by summing up the cyclic permutations of the y's

$$q_f[x, y_1, \cdots, y_n] = $$
$$p_f[x; y_1, y_2, \cdots, y_n] + p_f[x; y_n, y_1, \cdots, y_{n-1}] + \cdots + p_f[x, y_2, y_3, \cdots, y_n, y_1]$$

Properties of $f[\xi]$ yield the following properties of the non-commutative polynomials:

(8a) If $f[\xi] = 0$ whenever $\xi_i = \xi_j$ then $M_n(F)$ satisfies the identity

$$p_f[x, y_1, y_2, \cdots, y_n] = 0.$$

(8b) If $f[\xi] = 0$ when $\xi_i = \xi_j$ for all pairs $(i,j) \neq (1, n+1)$ $(i \neq j)$ then $M_n(F)$ satisfies: $p_f[x, y_1, \cdots, y_n]x - xp_f[x, y_1, \cdots, y_n] = 0$.

(8c) If $f[\xi] = 0$ when $\xi_i = \xi_j$ for all pairs $(i,j) \neq (1, n+1)$ and for $(i,j) = (1, n+1)$ $f[\xi]$ becomes a symmetric polynomial in $\xi_1, \cdots, \xi_n$ then $q_f[x: y_1, \cdots, y_n]$ is a central identity.

In the special case: $f_0[\xi] = \prod_2^n (\xi_1 - \xi_i)(\xi_{n+1} - \xi_i) \prod_{i,j}^{2,\ldots,n} (\xi_i - \xi_j)^2$ $q_0[x; y_1, \cdots, y_n] = F[x, y]$ is the Formanek central polynomial. The explicit form of this identity up to a $\pm$ sign is:

$$(9) \quad F[x; y] = \pm \sum_{i=1}^{n} \sum_{(\nu)(\mu)} [sg(\nu) sg(\mu)] x^{\nu_1} y_{i+1} x^{\nu_2 + \mu_1} y_{i+2} x^{\nu_3 + \mu_2} \cdots x^{\nu_n + \mu_{n-1}} y^{\nu_n + \mu_{n-1}} x^{\mu_n}$$

where the sum ranges over all permutations $(\nu_1, \cdots, \nu_n)$, $(\mu_1, \cdots, \mu_n)$ of $(0, 1, 2, \cdots, n-1.)$ This follows from the fact that $f_0[\xi]$ can also be written in the form

$$\pm \prod_{1 \leq i \leq j \leq n} (\xi_1 - \xi_j) \prod_{2 \leq i \leq j \leq n+1} (\xi_i - \xi_j)$$

which is up to a sign the product of two Vandermond's determinants;

$$\pm f_0[\ ] = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \xi_1 & \xi_2 & \cdots & \xi_n \\ \cdots & \cdots & \cdots & \cdots \\ \xi_1^{n-1} & \xi_2^{n-1} & \cdots & \xi_n^{n-1} \end{vmatrix} \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \xi_2 & \xi_3 & \cdots & \xi_{n+1} \\ \cdots & \cdots & \cdots & \cdots \\ \xi_2^{n-1} & \xi_3^{n-1} & \cdots & \xi_{n+1}^{n-1} \end{vmatrix}$$

$$= \sum_{(\nu)} sg(\nu)\xi_1^{\nu_1}\xi_2^{\nu_2}\cdots\xi_n^{\nu_n} \sum_{(\mu)} sg(\mu)\xi_2^{\mu_1}\xi_3^{\mu_2}\cdots\xi_n^{\mu_{n-1}}\xi_{n+1}^{\mu_n}$$

$$= \sum_{(\nu)(\mu)} [sg(\nu)sg(\mu)]\xi_1^{\nu_1}\xi_2^{\nu_2+\mu_1}\cdots\xi_n^{\nu_n+\mu_{n-1}}\xi_{n+1}^{\mu_n}$$

which yield the polynomial $q_{f_0}[x; y_1, \cdots, y_n] = F[x; y]$ given in [9].

One can replace $f_0[\xi]$ by any product $g[\xi_1, \cdots, \xi_n]f_0[\xi]$ where $g[\xi_1, \cdots, \xi_n]$ is a symmetric polynomial, then there exists a relation between q_{f_0} and q_{gf_0}: $q_{gf_0}[x, y] = g(x)q_{f_0}[x,y]$ with $g(x)$ a scalar whose value is $g[\alpha_1, \cdots, \alpha_n]$ where α_i are the characteristic roots of x. With this property we can get many new identities for $M_n(F)$ following the method described preceding (6b). In particular, if $g_1, \cdots, g_{n+1}$ are any $n+1$ symmetric functions, then there is a polynomial $P[g_1, g_2, \cdots, g_{n+1}] = 0$. Let $\bar{P}[\xi_0, \xi_1, \cdots, \xi_n]$ be the homogenized polynomial obtained from $P[\xi]$, then noting that $q_{g_i f_0}[x; y] = g_i(x)q_{f_0}[x; y]$, we get the identity:

(9) $$\bar{P}[q_{f_0}[x; y], q_{g_i f_0}[x; y], \cdots, q_{g_{n+1} f_0}[x; y]] = 0.$$

If $s_0 = 1, s_1, \cdots, s_n$ are the basic symmetric functions and set $q_i[x, y] = q_{s_i f_0}$, then the Cayley-Hamilton theorem yields that $\sum_0^n(-1)s_i x^{n-i} = 0$ and hence the following identity of $M_n(F)$:

(10) $$q_G = q_{f_0}[x, y]x^n - q_1[x;y]x^{n-1} + \cdots + (-1)^n q_n[x; y] = 0.$$

where $G = f_0[\xi_1, \cdots, \xi_n] \prod_{i=1}^n (\xi_{n+1} - \xi_1)$.

Next we review Razymslov central identity [26] which has the advantage of being multilinear. Let $y_1, \cdots, y_{n^2}; x_1 x_2, \cdots, x_{2(n^2-1)}$ and let $z_i = [x_{2i-1}, x_{2i}]$ $i = 1, 2, \cdots, n^2 - 1$, then the Razsmylov identity is the polynomial:

(11) $$R(z; y) = \sum_{k=1}^{n_2} \sum_{\sigma(k)=n^2} (sg\sigma)y_k z_{\sigma(k+1)}y_{k+1}z_{\sigma(k+2)}\cdots z_{\sigma(n^2)}y_{n^2}z_{\sigma(1)}y_1\cdots z_{\sigma(k-1)}y_{k-1}$$

where σ ranges over all permutations of $(1, 2, \cdots, n^2)$ and the monomials start with the index $k = \sigma^{-1}(n^2)$. Razymslov proves that $R(z; y)$ is a central polynomial; a simplified proof was given by Markov. It is interesting to note that there is a relation between $R(x; y)$ and the generalized standard identity:

$$R(x;y) = \sum_{k=1}^{n} (-1)^{(n^2-1)k} y_k T_{n^2}(z_1, \cdots, z_{n^2-1}, y_{k+1}, \cdots, y_{k-2}) y_{k-1}.$$

To conclude this part on central identities we point out an observation by Procesi that if $p[x]$ is a central identity for $M_n(F)$ then $p[x] = 0$ is an identity for $M_{n-1}(F)$. Indeed, if instead of substituting $(n-1) \times (n-1)$ matrices in $p[x]$ we substitute $n \times n$ matrices with zeros in the last row and last column. This will yield a scalar matrix with zero in the (n, n) place and hence $p[x] = 0$.

7. *Generators for* $\mathcal{M}_n$

The minimal identity in $\mathcal{M}_n$—the set of identities of $M_n(F)$ is the standard polynomial $S_{2n}[x]$. Leron (thesis, 1970) proved the identities of degree $2n + 1$ are consequences of $S_{2n}[x]$, i.e., they belong to the T-ideal generated by S_{2n}, but nevertheless M_n is not generated as a T-ideal by S_{2n}. In fact, we are going to prove:

The polynomial $S_{n+1}[y, xy, \cdots, x^n y]y^{-1}$ *of* (3) *does not belong to the T-idea generated by* $S_{2n}[x]$ *for* $n \geqq 2$.

Indeed if it were not the case then: $S_{n+1}[y, xy, \cdots, x^n y]y^{-1} = \sum a[x;y] \cdot S_{2n}[m_1, m_2, \cdots, m_{2n}] b[x,y]$ and we may assume that m_i's are monomials $x^{\nu_1} y^{\nu_1} x^{\nu_3} \cdots$ and each product $aS_{2n}b$ is a non zero polynomial in the ring of polynomials $F[x,y]$. Furthermore, the fact that $\mathcal{M}_n$ is homogeneous implies that we can assume that the degree of each term in x and y respectively are the same, i.e., of degree n in y and of degree $\binom{n}{2}$ in x. This means that at least n of the monomials m; will not have a y in them and so they are powers x^i. The total degree of n different powers of this type is $\geqq \binom{n}{2}$ and so the rest of the m_j will be powers of y, but as the total degree in y is n and $n \geqq 2$ we must have all $m_j = y$, and hence $aS_{2n}b$ is a zero polynomial as $S_{2n}[m_i]$ has at least two equal entries m_j. But this contradicts the assumption that none of these terms are zeros.

It is trivial to show that $\mathcal{M}_1$ is generated as a T-ideal by the commutative low $x_1, x_2 - x_2 x_1$, and recently Razmyslow has proved [29] that the ideal $\mathcal{M}_2$ of the identities of ןhe 2×2 matrices over a field of characteristic zero has a finite basis.

8. *Rational identities for* $M_n(F)$

In the previous sections we dealt with polynomial identities obtained only by the ring operators. If we allow taking inverses, we get additional rational identities

P. J. Albada (thesis, Utrecht, 1955) shows that the quaternions and hence matrix ring $M_2(F)$ satisfy identities like:

$$(12)\qquad [xyx^{-1}y^{-1}+yxy^{-1}x^{-1},x]=0;\; x[xy]^{-1}x[xy]-[xy]^{-1}x[xy]x=0$$

One can obtain other rational identities by identifying elements in the ring of quotent of ring of generic matrices. Thus, if $p[x]q[x]^{-1}$ is a central element, one gets an identity like $pq^{-1}z-zpq^{-1}=0$. It was shown in [22] that a rational identity which holds in all finite matrices $M_n(F)$, for infinite F, are those and only those which hold for all division algebras over F and there are such non-trivial. But for polynomial identities, $\cap\ \mathcal{M}_n=0$ and no such polynomial identity exists. Another difference between the polynomial and rational identities is that $\mathcal{M}_{n-1}\supset\mathcal{M}_n$ for polynomial identity, but (not as stated in [22]) it fails for rational identities. Bergman has recently given [28] a rational identity which holds for 3×3 matrices and does not hold for 2×2 matrices:

Let $y'=yx-xy$, the following is Bergman rational's identity

$$(13)\qquad \theta(y)\theta(y'')\theta(y''^{-1})\theta(y'''^{-1})=B(x,y)\ \text{ where }\ \theta(y)=\frac{(y'^2)'}{(y'^{-1})'}$$

9. *Identities of subalgebras of $M_n(F)$*

We conclude this survey with the interesting observation that identities can serve to distinguish between $M_n(F)$ and its proper subalgebra. The result we quote is:

A subalgebra $S\subseteq M_n(F)$ is $\neq M_n(F)$ if and only if it satisfies a polynomial identity which does not hold in $M_n(F)$, i.e. $\notin\mathcal{M}_n$.

The result is a simple consequence of the fact that if $S\nsubseteq M_n(F)$, then S/N is a direct sum of simple central algebras of dim $\leqq(n-1)^2$, where $N=N(S)$ is the nilpotent radical of S. Hence, S/N will satisfy all identities of $\mathcal{M}_{n-1}$. Now N is a nil subalgebra of $M_n(F)$ then $N^n=0$, and, consequently, S satisfies all polynomials of $\mathcal{M}^n_{n-1}$. In particular,

$$(14)\ \mathrm{Sub}_n[x]=S_{2n-2}[x_{11},\cdots,x_{1\,2n-2}]S_{2n-2}[x_{21},\cdots,x_{2\,2n-2}]\cdots S_{2n-2}[x_{n1},\cdots,x_{n2n-2}]=0$$

is an identity in $\mathcal{M}^n_{n-1}$ but does not belong to $\mathcal{M}_n$; indeed, since $F[x]/\mathcal{M}_n$ is a domain and $\mathrm{Sub}_n[x]\in\mathcal{M}_n$ then $S_{2n-2}[x]\in\mathcal{M}_n$ which is impossible. The converse is trivial, if $S \subseteq M_n(F)$ satisfies an identity $\notin \mathcal{M}_n$ and it is a subalgebra of $M_n(F)$, then it cannot be $= M_n(F)$.

References

1. M. Dehn, *Über die Grundlagen der projektiven Geometrie und allgemeine Zahlsysteme*, Math. Ann. **85** (1922), 184–193.

2. A. R. Richardson, *Equations over a division algebra*, Messeng. Math. **57** (1928), 1–6.

3. D. E. Littlewood, *Identical relations in algebra*, Proc. London Math. Soc. **32** (2) (1931), 312–320.

4. W. Wagner *Über die Grundlagen der projektiven Geometrie and allgemeine Zahlsysteme*, Math. Z. **113** (1936–37), 528–567.

5. M. Hall, *Projective planes*, Trans. Amer. Math. Soc. **54** (1943), 229–277.

6. M. Kurosh, *Ringtheoretische Probleme, die mit dem Burnsideschen Probleme über periodiche Gruppen in Zusammenhang stehen*, Bull. Acad. Sci. URSS, Ser. Math. **5** (1941), 223–240.

7. N. Jacobson, *Structure theory for algebraic algebras of bounded degree*, Ann. of Math. **46** (1945), 695–707.

8. I. Kaplansky, *On a problem of Kurosh and Jacobson*, Bull. Amer. Math. Soc. **52** (1946), 496–500.

9. J. Levitski, *On a problem of Kurosh*, Bull. Amer. Math. Soc. **52** (1946), 1033–1035.

10. F. W. Levi, *On Skewfields of a given degree*, J. Indian. Math. Soc. **II** (1947), 85–86.

11. S. A. Jennings, *On rings whose associated Lie ring is nilpotent*, Bull. Amer. Math. Soc. **53** (1947) 593–598.

12. I. Kaplansky, *Rings with a polynomial identity*, Bull. Amer. Math. Soc. **54** (1948), 575–580.

13. F. W. Levi, *Uber der Kommutativitatsrang in einem Ring*, Math. Ann. **121** (1949), 184–190.

14. I. Kaplansky, *Group presentation of algebras*, Canad. J. Math. **1** (1949), 105–112.

15. I. Kaplansky, *Topological representation of algebras*, Trans. Amer. Math. Soc. **68** (1950), 334–336.

16. J. Levitzki, *A theorem on polynomial identities*, Proc. Amer. Math. Soc. **1** (1950), 449–463.

17. S. A. Amitsur and J. Levitski, *Minimal identities for algebras*, Proc. Amer. Math. Soc. **1** (1950), 449–463.

18. W. Specht, *Gesetz in Ringen I*, Math. Z. **52** (1949/50), 557–584.

19. A. I. Malcev, *On algebras satisfying defining identical relations*, Mat. Sb. **68** (1950), 11–33.

20. S. A. Amitsur and J. Levitzki, *Remarks on minimal identities for algebras*, Proc. Amer. Soc. **2** (1951), 320–327.

21. S. A. Amitsur, *Identities and generators of matrix rings* Bull. Res. Counc. Israel **5a** (1955), 1–10.

22. S. A. Amitsur, *Rational identities and applications to algebras and geometry*, J. Algebra 3 (1966), 304–359.

23. S. A. Amitsur and C. Procesi, *Jacobson rings and Hilbert algebras with polynomial identities*, Ann. di Mat., IV **71** (1966), 61–72.

24. C. Procesi, *Sulle identita delle algebra semplice*, Rend. Circ. Mat. Palermo, II **17** (1968), 13–18.

25. E. Formanek, *Central polynomials for matrix rings*, J. Algebra **23** (1972), 129–132.

26. P. Razymslov, *On the Kaplansky problem* (Russian), Izv. Acad Nauk. SSSR (Ser. Mat. **37**(3), (1973).

27. S. A. Amitsur, *Polynomial identities and Azumaya algebras* J. Algebra **27** (1973), 117–125.

28. S. Bergman, *Rational relations and rational identities in division rings*, J. Algebra (to be published).

29. P. Razmyslov, *Finite basis of the identities of a matrix algebra at second order over a field of characteristic zero* (Russian), Algebra i. Logika **12** (1973), 83–113.

Institute of Mathematics
The Hebrew University of Jerusalem
Jerusalem, Israel