PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 66, Number 2, October 1977

ON THE NONEXISTENCE OF HYPERCOMMUTING POLYNOMIALS

AMOS KOVACS

ABSTRACT. If $f(x_1, \ldots, x_n)$ is not central for R, then the additive group generated by all specializations of f in R contains a noncentral Lie ideal of R. This is used, among other things, to prove:

THEOREM. *Let R be a semiprime algebra over an infinite field, $f_1, \ldots, f_t$ polynomials in disjoint sets of variables all noncentral for R. Then, if R satisfies $S_t[f_1, \ldots, f_t]$, R must satisfy $S_t[x_1, \ldots, x_t]$.*

Introduction. Let k be a commutative ring with 1. All rings will be unital associative k-algebras. Let $k\{X\} = k\{x_1, x_2, x_3, \ldots\}$ be the free k-algebra in a denumerable number of (noncommuting) variables. We call a polynomial $f(x_1, \ldots, x_n) \in k\{X\}$ central for a k-algebra R if all specializations of f in R lie in the center of R. Note that contrary to the common usage we do not require f to be nonvanishing on R. The ring of $n \times n$ matrices over R will be denoted by $M_n(R)$ and we denote by e_{ij} $(1 \leqslant i, j \leqslant n)$ the standard matrix units in $M_n(R)$. The standard polynomial of degree t will be denoted by $S_t = S_t[x_1, \ldots, x_t]$.

In §1 we introduce the notion of the extended range of a polynomial f on R, and we prove that this is a "large" subset of R. Namely, if f is not central for R then the extended range is an invariant submodule which (under some mild assumptions on f) contains a noncentral Lie ideal of R.

A polynomial f is central for R if $[f, y]$ is an identity for R. One is tempted to ask: Does there exist a polynomial f which is not central but such that for some $t > 2$, $S_t[f, y_1, \ldots, y_{t-1}]$ vanishes on R? It follows from [5] that the answer, at least in the manageable cases, is negative. In §2 we investigate a generalization of this problem. Let $f_1, \ldots, f_t$ be polynomials in disjoint sets of variables. We say that the f_i's are hypercommuting if none of them is central for R but $S_t[f_1, \ldots, f_t]$ is an identity for R. We prove that prime rings and semiprime algebras have no hypercommuting polynomials–except for some exceptional cases. We next specialize our results to the case $f_i = x_i^{n_i}$ to prove a "higher commutativity theorem", that is, a commutativity theorem with commutators replaced by the standard polynomial.

Received by the editors November 18, 1976 and, in revised form, January 17, 1977.

AMS (MOS) subject classifications (1970). Primary 16A38, 16A70; Secondary 16A12, 16A42.

Key words and phrases. Extended range, invariant subspace, Lie ideal, polynomial identity, hypercommuting polynomials.

1. **The extended range of a polynomial.** Let R be a k-algebra and $f(x_1, \ldots, x_n) \in k\{X\}$. We denote by $f[R]$ the (additive) k-submodule of R generated by all the specializations of f in R. We call $f[R]$ the extended range of f on R.

We start with the following simple observation:

LEMMA 1. *Let R be a k-algebra and $f \in k\{X\}$. Then,*

(i) *$f[R]$ is invariant under all homomorphisms of R into R.*

(ii) *If f is multilinear, $f[R]$ is invariant under all derivations of R, in particular, it is a Lie ideal of R.*

PROOF. The first part of the lemma is obvious. As for the second part, let $a_1, \ldots, a_n \in R$ and $a \to a'$ a k-derivation of R. It follows from the definition of a derivation that if f is multilinear

$$f(a_1, \ldots, a_n)' = \sum_{i=1}^{n} f(a_1, \ldots, a_i', \ldots, a_n) \in f[R].$$

Finally, a Lie ideal is an additive submodule invariant under all inner derivations.

We now would like to see how much of the second part of the lemma remains valid for a general polynomial f. To that purpose we examine the effect of multilinearization on the extended range $f[R]$. For details on the multilinearization process and the terms used the reader is referred to [4, pp. 15–19].

Recall that a polynomial f is called blended in some variable x_i if each monomial of f has a positive degree in x_i. f is called blended if it is blended in each variable appearing in it. Every polynomial may be written as a sum of blended polynomials. To be more precise: let $f = f(x_1, \ldots, x_n)$, for each subset $A \subseteq \{x_1, \ldots, x_n\}$ define f_A to be the sum of the monomials in f having positive degree exactly in the variables appearing in A. We have then $f = \sum_{A \subseteq \{x_1, \ldots, x_n\}} f_A$, the f_A's are blended and uniquely determined by f. We call them the blended components of f. With the notation as above we have:

LEMMA 2. $f[R] = \sum_A f_A[R]$.

PROOF. $\subseteq$ is clear. For the converse we have to show that each $f_A[R]$ is contained in $f[R]$. Assume inductively that $f_B[R] \subseteq f[R]$ whenever $B \subset A$. Denote by g the result of substituting 0 in f for each variable not in A. Clearly $g = \sum_{B \subseteq A} f_B = \sum_{B \subset A} f_B + f_A$. By the definition of g, $g[R] \subseteq f[R]$ and so, since by our induction hypothesis the $f_B[R]$'s are included in $f[R]$, we conclude that $f_A[R] \subseteq f[R]$.

THEOREM 3. *Let R be a k-algebra, $f(x_1, \ldots, x_n) \in k\{X\}$. $f[R]$ contains a nonzero (noncentral) Lie ideal of R–unless all the multilinearizations of all blended components of f are identities (central) for R.*

If k is a field of characteristic zero and f is not an identity (central) for R, then $f[R]$ contains a nonzero (central) Lie ideal.

Proof. Let G be either $\{0\}$ or the center of R–as the case may be. By our assumption f has a blended component g with a multilinearization h such that h is not G valued on R (that is, $h[R] \nsubseteq G$). By Lemma 2, $g[R] \subseteq f[R]$ and clearly, since h is a multilinearization of g, $h[R] \subseteq g[R]$. It now follows from part (ii) of Lemma 1 that $h[R]$ is a nonzero (noncentral) Lie ideal in $f[R]$.

Assume now k is a field, $\chi(k) = 0$ and f is not G valued. By Lemma 2, f must have a blended component g which is not G valued. As before, let h be a multilinearization of g. It is well known that the assumption on the characteristic of k implies $g[R] = h[R]$, and so $h[R]$ is a nonzero (central) Lie ideal in $f[R]$.

Remark. It is not difficult to see that k does not really have to be of characteristic zero. It is enough that $\chi(k) = p > t$ where t is the maximum of the degrees of f in each variable.

In view of the theorem we shall say that a polynomial f is strongly noncentral (strongly nonvanishing) on R if it has a blended component with a multilinearization which is not central (vanishing) for R.

Our previous observations make it rather easy to classify all possible extended ranges for a matrix ring $R = M_n(k)$ over a field k. We first note that the only subspaces of R invariant under all inner automorphisms are: 0, the scalar matrices, the trace zero matrices–$[R, R]$, and the whole space R, except when $R = M_2(GF[2])$. This fact is not too difficult to prove directly, or one could apply theorems of Baxter [2, Theorem 2] and Sah [6, Proposition 2] which taken together imply the desired result. Either way, we deduce that if f is a polynomial which is not central for $R = M_n(k)$, then $f[R] = R$ or $f[R] = [R, R]$–unless $R = M_2(GF[2])$. In order to handle this last exceptional case we need to assume that f is strongly noncentral for R. Assuming this, we know by Theorem 3 that $f[R]$ is not only an invariant subspace of R but contains a noncentral Lie ideal of R. A direct computation will show that while $R = M_2(GF[2])$ does contain exceptional invariant subspaces and Lie ideals, the only invariant subspaces containing a noncentral Lie ideal are again R and $[R, R]$. This proves:

Corollary 4. *Let $R = M_n(k)$, $f \in k\{X\}$ a polynomial not central on R, then $f[R] = [R, R]$ or $f[R] = R$, except when $R = M_2(GF[2])$. The conclusion will still hold in the last case provided f is strongly noncentral for R.*

Remarks. 1. Most of the previous discussion can be done in a wider context. For that purpose one would use Amitsur's results about invariant subspaces of simple rings [1], rather than Baxter's.

2. It is easy to see that for $R = M_2(GF[2])$ the requirement that f be strongly noncentral is indeed necessary. The extended range of $f(x) = x^5 + x^3$ on $M_2(GF[2])$ is the subspace of matrices of the form $\left(\begin{smallmatrix}\alpha & \alpha+\beta\\ \alpha+\beta & \beta\end{smallmatrix}\right)$. This happens since the multilinearization of f (in characteristic 2) is S_5 which of course is an identity of $M_2(GF[2])$.

2. **Hypercommuting polynomials.** Let R be a k-algebra, f and g two polynomials in $k\{X\}$. We shall say that f dominates g on R if $f[R] \supseteq g[R]$.

LEMMA 5. *Let R be a k-algebra, $T(x_1, \ldots, x_t)$ a multilinear polynomial in $k\{X\}$ (e.g. S_t), and $f_1, \ldots, f_t$, $g_1, \ldots, g_t$ arbitrary elements of $k\{X\}$ in disjoint sets of variables such that each f_i dominates g_i on R. Then $T(f_1, \ldots, f_t)$ dominates $T(g_1, \ldots, g_t)$ on R. In particular, if $T(f_1, \ldots, f_t)$ is an identity (central) for R, so is $T(g_1, \ldots, g_t)$.*

PROOF. It is easy to see from the multilinearity of T, that for each i, $T(g_1, \ldots, g_{i-1}, f_i, \ldots, f_t)$ dominates $T(g_1, \ldots, g_i, f_{i+1}, \ldots, f_t)$. The result now follows by the transitivity of dominance.

LEMMA 6. *For $n > 1$, $M_n(k)$ satisfies*

$$S_t[[x_1, x_2], [x_3, x_4], \ldots, [x_{2t-1}, x_{2t}]],$$

if and only if $t \geqslant 2n$ or $n = 2$, $t = 3$ and $3 = 0$ in k.

PROOF. The fact that for $t \geqslant 2n$, $M_n(k)$ satisfies

$$S_t[\ \ldots [x_{2i-1}, x_{2i}] \ldots\ ]$$

follows from the Amitsur-Lewitski theorem. Since for $l < t$, S_t can be expressed in terms of S_l, to prove the other direction it is enough to show that $M_n(k)$ does not satisfy

$$S_{2n-1}[\ \ldots, [x_{2i-1}, x_{2i}], \ldots\ ],$$

except in the one exceptional case. Let us first note that

$$(1) \quad S_{2n-1}[e_{11}, e_{12}, e_{23}, \ldots, e_{n-1n}, e_{nn-1}, e_{n-1,n-2}, \ldots, e_{21}] = 2 - e_{nn}.$$

Indeed, the only orders in which these matrix units yield nonvanishing products are the cyclic permutations of the given order. Among the $2n - 1$ possible products each e_{ii} occurs twice except e_{nn} which occurs only once. This, together with the fact that all these permutations are even, yields (1).

We similarly find that:

$$(2) \quad S_{2n-1}[e_{nn}, e_{12}, e_{23}, \ldots, e_{n-1n}, e_{nn-1}, \ldots, e_{21}] = (-1)^{n-1}(2 - e_{11}).$$

Subtracting (2) from (1) we get on the left a specialization of S_{2n-1} in which all the arguments are commutators, while on the right we get $4 - e_{11} - e_{nn}$ when n is even and $e_{11} - e_{nn}$ when n is odd. These matrices do not vanish unless $n = 1$, or $n = 2$ and $3 = 0$ in k. To see that 2×2 matrices over a ring of characteristic 3 actually satisfy

$$S_3[[x_1, x_2], [x_3, x_4], [x_5, x_6]],$$

note that $e_{22} - e_{11}$, e_{12} and e_{21} span $[R, R]$ for $R = M_2(k)$.

If R is a prime PI ring we denote by p.i. deg R the least integer n such that R satisfies S_{2n}. The next theorem is well known and we include it for completeness.

THEOREM 7. *Let R be a prime PI algebra over k,* p.i. deg $R = n$, *Z the* *center of R and Q the field of quotients of Z. Then R and $M_n(Q)$ satisfy the same identities over k.*

PROOF. Let $R_0 = R \otimes_Z Q$ be the algebra of central quotients of R. R_0 is central simple and n^2 dimensional over Q and satisfies the same identities as R [4, Theorem 2, p. 57]. If Q is a finite field then by the two Wedderburn theorems $R_0 \cong M_n(Q)$ and we are done. If Q is infinite, let L be a splitting field for R_0; $R_0 \otimes_Q L \cong M_n(L) \cong M_n(Q) \otimes_Q L$. Tensoring over infinite fields preserves identities [4, Lemma 1, p. 89] therefore R_0 and $M_n(Q)$ have the same identities.

We can now prove our result about the nonexistence of hypercommuting polynomials for prime rings:

THEOREM 8. *Let R be a prime k-algebra, $f_1, \ldots, f_t$ polynomials in disjoint sets of variables all noncentral for R. If R satisfies $S_t[f_1, \ldots, f_t]$ then R satisfies $S_t[x_1, \ldots, x_t]$ except when $\chi(R) = 3$ and* p.i. $\deg R = 2$ *or when $R = M_n(GF[2])$. In the last case the theorem still holds provided the f_i's are strongly noncentral on R.*

PROOF. Since the f_i's are in disjoint sets of variables $S_t[f_1, \ldots, f_t]$ is a nontrivial identity for R and R is a prime PI algebra. Let Z and Q be as in Theorem 7 and p.i. $\deg R = n > 1$. By Theorem 7 we may assume $R = M_n(Q)$. Note that $M_n(Q) = M_2(GF[2])$ if and only if $R = M_2(GF[2])$. Since none of the f_i's is central for R, by Corollary 4, $f_i[R] \supseteq [R, R]$–that is, f_i dominates $[x_{2i-1}, x_{2i}]$ on R. By Lemma 5 this implies that $S_t[f_1, \ldots, f_t]$ dominates $S_t[\ldots, [x_{2i-1}, x_{2i}], \ldots]$ on R and therefore, by our assumption, the latter is an identity for R. This implies, via Lemma 6, that S_t is an identity for R–unless $n = 2$ and $\chi(R) = 3$.

The following example indicates where we may run into trouble when trying to generalize Theorem 8 to semiprime rings.

EXAMPLE. Let $R = M_2(Z) \oplus M_3(Z_p)$, p a prime. Let f_1 be a (multilinear) polynomial which is an identity for $M_2(Z)$ but is not central on $M_3(Z_p)$ (e.g. $f_1 = S_4[x_1, x_2, x_3, x_4]$). Let f_2 be a (multilinear) polynomial which is an identity for $M_3(Z_p)$ but not central for $M_2(Z)$ (px_5 will do). $[f_1, f_2]$ is now an identity of R (as a matter of fact $f_1 f_2$ is already an identity), but neither f_1 nor f_2 are central for R–and of course R does not satisfy $S_2[x_1, x_2]$.

We can however avoid these difficulties if we restrict ourselves to algebras over a field.

THEOREM 9. *Let R be a semiprime algebra over an infinite field k of characteristic $\neq 3$. Let $f_1, \ldots, f_t$ be polynomials in disjoint sets of variables all noncentral for R. If R satisfies $S_t[f_1, \ldots, f_t]$ then R satisfies $S_t[x_1, \ldots, x_t]$. If k is finite $x(k) \neq 3$ the theorem still holds provided the f_i's are multilinear.*

PROOF. Assume R does not satisfy S_t. Pick a prime $P_0 \lhd R$ with p.i. deg R/P_0 maximal. By Theorem 8 some f_i, say f_1, is central for R/P_0. Let $P \lhd R$ be any prime, then p.i. deg $R/P \leqslant$ p.i. deg R/P_0 and it follows easily from the general theory (see [4]) that under our assumptions f_1 is central for R/P. Since R is semiprime, R is a subdirect sum of its prime homomorphic images, hence f_1 is central for R.

We now restrict the f_i's in order to get a generalized commutativity theorem, namely:

THEOREM 10. *Let R be a semiprime k-algebra satisfying $S_t[x_1^{n_1}, \dots, x_t^{n_t}]$. Then R satisfies $S_t[x_1, \dots, x_t]$.*

PROOF. Assume first R is prime and p.i. deg $R = n > 1$. Since $x_i^{n_i}$ is clearly not central on R, the result follows immediately from Theorem 8, except when $\chi(R) = 3$ and $n = 2$ or when $R = M_2(GF[2])$. To prove the theorem for these two exceptional cases it is enough to show that 2×2 matrices (over any commutative ring with $1 \neq 0$) do not satisfy an identity of the form $S_2[x_1^{n_1}, x_2^{n_2}]$ or $S_3[x_1^{n_1}, x_2^{n_2}, x_3^{n_3}]$. This is easily verified by the substitution $x_1 = e_{11} + e_{12}$, $x_2 = e_{11} + e_{21}$, $x_3 = 1$. If R is semiprime choose again a prime $P_0 \lhd R$ with p.i. deg R/P_0 maximal. By the previous discussion R/P_0 satisfies $S_t[x_1, \dots, x_t]$, our choice of P_0 now implies that R satisfies S_t.

Theorem 10 may naturally be considered as a commutativity theorem for higher degrees of commutativity. It would be interesting to see what can be said in this situation when the exponents are not fixed, i.e.:

Question. What can be said about a (say prime) ring R satisfying

$$S_t[a_1^{n_1}, a_2^{n_2}, \dots, a_t^{n_t}] = 0$$

where $n_i = n_i(a_1, a_2, \dots, a_t)$?

This was only recently solved for the case $t = 2$ (see [3]). A general solution will certainly necessitate a different approach than the one employed here.

ACKNOWLEDGEMENT. The author wishes to thank the referee Professor George M. Bergman for his many helpful suggestions which improved the results and the presentation of this paper.

REFERENCES

1. S. A. Amitsur, *Invariant submodules of simple rings*, Proc. Amer. Math. Soc. **7** (1956), 987–989.

2. W. E. Baxter, *Lie simplicity of a special class of associative rings*, Proc. Amer. Math. Soc. **7** (1956), 855–863.

3. I. N. Herstein, *A commutativity theorem*, J. Algebra **38** (1976), 112–118.

4. N. Jacobson, *PI-algebras, an introduction*, Lecture Notes in Math., vol. 441, Springer-Verlag, Berlin and New York, 1975.

5. A. Kovacs, *The n-center of a ring*, J. Algebra **40** (1976), 107–124.

6. C. H. Sah, *Cohomology of split group extensions*, J. Algebra **29** (1974), 255–302.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MASSACHUSETTS, AMHERST, MASSACHUSETTS 01002

Current address: Faculty of Mathematics, Technion, Technion City, Haifa, Israel