ISRAEL JOURNAL OF MATHEMATICS **138** (2003), 43–60

ALGEBRAIC DERIVATIONS WITH CONSTANTS SATISFYING A POLYNOMIAL IDENTITY

BY

CHEN-LIAN CHUANG AND TSIU-KWEN LEE

Department of Mathematics, National Taiwan University
Taipei 106, Taiwan
e-mail: chuang@math.ntu.edu.tw, tklee@math.ntu.edu.tw

ABSTRACT

We prove that if a semiprime ring R possesses a derivation which is integral over its extended centroid C and whose constants satisfy a polynomial identity, then R itself is a PI-ring. This answers affirmatively a problem raised by M. Smith in 1975 and recently again by Bergen and Grzeszczuk [4].

0. Introduction

Rings considered here are always associative. We recall some notions and meanwhile fix our notation:

Definition: Given a ring S and a subring R, for $b \in S$, the **centralizer** of b in R, denoted by $C_R(b)$, is the subring $\{r \in R \mid br = rb\}$. Assume that S is an algebra with 1 over a commutative ring C. An element $b \in S$ is said to be **integral** over C, if $b^n + \alpha_1 b^{n-1} + \cdots + \alpha_n = 0$ for some $\alpha_1, \ldots, \alpha_n \in C$.

In an algebra R with 1 over a commutative ring C, if b is integral over C, then $C_R(b)$ is considered to be large in the sense that nice properties of $C_R(b)$ can be usually extended to the whole ring R. Herstein and Neumann [10] initiated this line of research by proving that the simplicity of $C_R(b)$ implies the simplicity of R, if R is semiprime and $b \in R$ is integral over the center Z of R. Under the assumption that R is a semiprime algebra over a field C and $b \in R$ is algebraic over C, Cohen [6] proved that if $C_R(b)$ is semiprime Artinian (or right Goldie respectively), then so is R. As polynomial identities (PIs) are powerful tools, one naturally asks

Received March 7, 2002

PROBLEM 1: *Let R be a semiprime ring with extended centroid C. Assume that $b \in R$ is integral over C. If $C_R(b)$ is a PI-ring, is R also a PI-ring?*

This was first formulated by Smith [21], p. 149 for the centroid instead of the extended centroid. Even partial answers to this problem are very useful. See [9] for instance. Firstly, Montgomery [18] confirmed this conjecture under the assumption that R is a simple ring with 1 and b is power central. Smith [21] improved this result by assuming that R is prime and that $b \in R$ is integral over the centroid of R. About the same time, Rowen [20] also proved the same result under the slightly different assumptions that R is prime, that $b \in R$ is algebraic over the extended centroid C of R and also that $\mu'(b)$ is invertible, where $\mu(X)$ is the minimum polynomial of b over C. The extended centroid of a prime ring, more general than the centroid, forms a field and is hence more natural to work with. But Rowen's last assumption on the minimum polynomial of b seems rather peculiar and probably redundant in view of Smith's neat statement. Anyway, neither of Smith's or Rowen's result actually implies the other. So the last word on this problem was not yet uttered in the prime case! The complete answer for prime rings is recently given by Bergen and Grzeszczuk's result [4] in a far more general context which we now explain:

Definition: (1) By a **derivation** of a ring R, we mean an additive map $\delta\colon R \to R$ such that $\delta(x+y) = \delta(x)+\delta(y)$ and $\delta(xy) = \delta(x)y+x\delta(y)$ for all $x, y \in R$. Those $x \in R$ with $\delta(x) = 0$ are called **constants** of δ. For a subring S of R, we set $S^{(\delta)} \stackrel{\text{def}}{=} \{r \in S \mid \delta(r) = 0\}$, the subring of constants of δ in S.

(2) Let R be a semiprime ring with extended centroid C. By a **continuous** derivation of R, we mean a derivation δ of the symmetric Martindale quotient ring Q of R such that $\delta(I) \subseteq R$ for an essential two-sided ideal I of R. Any derivation of R can be uniquely extended to a derivation of Q and hence is a continuous derivation of R (Proposition 1.8.1 [14] or Proposition 2.5.1 [3]). A continuous derivation δ of R is said to be **integral** over C, if there exist $\alpha_1, \ldots, \alpha_n \in C$ such that $\delta^n(x) + \alpha_1\delta^{n-1}(x) + \cdots + \alpha_n x = 0$ for all $x \in R$ (and hence for all $x \in Q$ by Theorem 8.4.1 [3]).

Bergen and Grzeszczuk's result [4] proves that, for a prime ring R with the extended centroid C and for a continuous derivation δ of R which is algebraic over C, if $R^{(\delta)}$ is a PI-ring, then so is R. Let us consider a special instance of this Theorem: Given $b \in R$ (or more generally $b \in Q$), the map $\mathrm{ad}(b)\colon x \mapsto bx - xb$ defines a continuous derivation of R and then $C_R(b)$ consists of constants of $\mathrm{ad}(b)$ in R. Obviously, $\mathrm{ad}(b)$ is algebraic over C if and only if so is b. In this context, Bergen and Grzeszczuk's result asserts that if $C_R(b)$ is PI, then so is R. This gives a complete answer to Problem 1 in the prime case. Though Problem 1 in the semiprime case is not answered yet (p. 736, [4]), it is natural to ask whether Bergen and Grzeszczuk's result for prime rings can be extended to the semiprime case, namely,

PROBLEM 2: *Let R be a semiprime ring with extended centroid C and δ a continuous derivation of R integral over C. If $R^{(\delta)}$ is a PI-ring, is R a PI-ring ?*

Our main objective here is to settle this problem affirmatively. The basic method to push results from prime rings to semiprime rings goes as follows: Consider R/P for suitable prime ideals P. Make sure that all notions considered (such as the derivation δ) induce a well-defined notion on R/P and also that all hypotheses imposed carry over to R/P. Apply then the result of the prime case to R/P. Finally, deduce the validity of the conclusion on R from its validity on these R/P. This method seems to be most natural but can be very difficult sometimes. For example, for our problem here, we need a family of prime ideals invariant under δ to insure that δ also induces a derivation on R/P. This is not obvious at all. A very powerful method, which we will adopt here and also which is particularly mentioned in Bergen and Grzeszczuk [4], is the theory of orthogonal completeness due to Beidar and Mikhalëv [2]. This theory is essentially a systematization of the above method in terms of first order logic for orthogonal complete rings, which possess very nice prime ideals defined by central idempotents. In applying this theory to our Problem 2, there are two crucial difficulties encountered:

(1) The orthogonal completion of R may have more constants of δ than the original R has. It is *not* obvious that this larger set of constants still satisfies a PI.

(2) The conclusion that each R/P satisfies a PI does *not* integrate to give a PI of R. We need here a *uniform* bound for PI-degrees of those R/P, simply because it is implied by our desired conclusion that R has a PI. In other words, we must estimate the PI-degree of R in terms of the integral degree of δ in the prime case.

These difficulties show that a more careful analysis of the prime case is needed to deal with the semiprime case. Merely the assertion that $R^{(\delta)}$ being a PI-ring implies R being a PI-ring for a prime ring R does **not** seem enough for such extension to the semiprime case. Our solutions to these difficulties will be explained in detail in the next section. Our strategy is to reduce outer derivations to inner ones by Kharchenko's theory of differential identities and then to treat inner derivations by the theory of GPIs. Both of our tools, the theory of GPIs and the theory of differential identities, have their semiprime versions ([3] and [13]), which can also be considered as applications of the theory of orthogonal completeness. With this, we could also work completely in semiprime rings. But this approach seems to obscure the key points with additional complications of treating prime components all put together.

Finally, we must mention a generalization of our problems here: We may consider the centralizer of a finite dimensional subalgebra or constants of a finite dimensional (restricted) differential Lie algebra of derivations. A lot of interesting work has been done in this direction. A recent work is [15]. But results in this direction, when applied to a single element or derivation of a prime ring, do *not* yield the full strength of Bergen and Grzeszczuk's result [4], which we have quoted above. For example, Theorem 8.1 or 8.7 [15] requires, in addition to others, that the subring of constants be semiprime. This is not always the case, as can be easily seen in finite matrix rings over fields.

1. Results

Given b integral over a commutative ring C, the success for work of this type depends crucially on the construction of a generalized polynomial commuting with b as follows:

FACT A: Given b in a C-algebra and a polynomial $g(X)=\sum_{i=0}^{n}\beta_{n-i}X^i$, where $\beta_i\in C$, we define

$$\hat{g}_b(X)=\sum_{i=1}^{n}\beta_{n-i}\Big(\sum_{j=0}^{i-1}b^jXb^{i-j-1}\Big).$$

Then $[b,\hat{g}_b(X)]=[g(b),X]$. Particularly, if $g(b)=0$, then $[b,\hat{g}_b(x)]=0$ for all x.

Proof: For the first assertion, if $g(X)=X^n$, then

$$\hat{g}_b(X)=\sum_{j=0}^{n-1}b^jXb^{n-j-1}=Xb^{n-1}+bXb^{n-2}+\cdots+b^{n-2}Xb+b^{n-1}X.$$

A direct computation shows that $[b,\hat{g}_b(X)]=[b^n,X]=[g(b),X]$. The general case follows by linear extension. The last assertion is obvious.

Our solution to the first difficulty is to replace the assumption that $R^{(\delta)}$ satisfies a PI by the following weaker one for a continuous derivation δ of a semiprime ring R with symmetric Martindale quotient ring Q:

(H): *There exists a multilinear polynomial* $f(X_1,\dots,X_t)$ *with coefficients* ± 1 *satisfying the following: For any* $b\in Q$ *and* $c_i\in Q^{(\delta)}$, $i=0,\dots,m-1$, *such that* $\sum_{i=1}^m c_i\delta^i=\mathrm{ad}(b)$ *and for any linear generalized polynomial* $l(X)=\sum_{j=1}^k a_jXb_j$ *such that* $[b,l(x)]=0$ *for all* $x\in R$, *we have*

$$f\Big(\sum_{i=1}^m c_i\delta^{i-1}(l(y_1)),\dots,\sum_{i=1}^m c_i\delta^{i-1}(l(y_t))\Big)=0$$

for all $y_1,\dots,y_t\in Q$.

We observe the following

FACT B: Let R be a semiprime ring with extended centroid C and δ a continuous derivation of R which is integral over C. If $R^{(\delta)}$ is a PI-ring, then Q satisfies the hypothesis (H).

Proof: Since $R^{(\delta)}$ is a PI-ring, it satisfies a multilinear polynomial identity $f(X_1,\dots,X_t)$ with coefficients ± 1 ([1] or p. 59 Theorem 1 of [11]). Let $c_i\in Q^{(\delta)}$, $b\in Q$ and $l(X)=\sum_{j=1}^k a_jXb_j$ be as said in (H) above. For any $x\in Q$,

$$0=[b,l(x)]=\sum_{i=1}^n c_i\delta^i(l(x))=\delta\Big(\sum_{i=1}^n c_i\delta^{i-1}(l(x))\Big),$$

where the last equality follows by the assumption that $c_i\in Q^{(\delta)}$. By the continuity of δ, there exists an essential two-sided ideal I of R such that for all i and all $y\in I$, $c_i\delta^{i-1}(l(y))\in R$. So we have $\sum_{i=1}^n c_i\delta^{i-1}(l(y))\in R^{(\delta)}$ for $y\in I$. Since $f(X_1,\dots,X_t)$ is a PI for $R^{(\delta)}$, the identity displayed in (H) holds for $y_1,\dots,y_t\in I$ and hence also holds for $y_1,\dots,y_t\in Q$ by Theorem 3 of [16]. ∎

It is surprising that the weaker hypothesis (H) suffices to yield PIs for R (and hence Q also). As our work is based on the thorough analysis of algebraic derivations given in Kharchenko [12], for easy reference, we quote this result as part (1) of our Theorem 1 below:

THEOREM 1: *Let* R *be a prime ring with extended centroid* C *and* δ, *a continuous derivation of* R *which is algebraic over* C. *Assume that the subring* $R^{(\delta)}\stackrel{\text{def}}{=}\{x\in R\mid\delta(x)=0\}$ *is a PI-ring or, more precisely, that the weaker hypothesis (H) holds. Then the following holds:*

(1) *If* $\operatorname{char}R=0$, *then* $\delta=\mathrm{ad}(b)$ *for some* $b\in Q$. *If* $\operatorname{char}R=p>0$, *then there exists* $m\geq 0$ *such that* δ^{p^s}, $s=0,\dots,m-1$, *are* C*-independent modulo X-inner derivations and such that for some* $\alpha_i\in C^{(\delta)}$ *and* $b\in Q^{(\delta)}$,

$$\delta^{p^m}+\alpha_1\delta^{p^{m-1}}+\cdots+\alpha_m\delta=\mathrm{ad}(b).$$

Let $\ell(X) = X$ *if* $\operatorname{char} R = 0$ *and* $\ell(X) = X^{p^m} + \alpha_1 X^{p^{m-1}} + \cdots + \alpha_m X$ *if* $\operatorname{char} R = p > 0$.

(2) *R (and hence Q) is a PI-ring.*

(3) *For the b described in (1), $R^{(\delta)}$ and $C_R(b)$ satisfy the same PIs over C.*

(4) *The minimal polynomial of δ over C is equal to $\mu(\ell(X))$, where $\mu(X)$ is the minimal polynomial of* $\operatorname{ad}(b)$ *over C. In particular,* $\deg \ell(X) \cdot \deg_C b \leq \deg_C \delta$.

(5) *Let* $\deg_C b = \nu$ *and let q be the minimum of degrees of PIs satisfied by $R^{(\delta)}$. Then R satisfies $S_{q\nu}(X_1, \ldots, X_{q\nu})$. In particular, R satisfies $S_{qs}(X_1, \ldots, X_{qs})$, where* $s = \deg_C \delta$.

Our line of attacks for Theorem 1 goes as follows: Using part (1), we can show that R is a PI-ring under the weaker hypothesis (H). That is part (2). We are now working in the nice PI-ring R and our problem is to estimate the PI-degree of R in terms of $\deg_C \delta$ and the PI-degree of $R^{(\delta)}$. To do this, we first show that the PI-degrees of $R^{(\delta)}$ and $C_R(b)$ are actually the same. This is part (3). The PI-degree of R can be easily estimated in terms of $\deg_C b$ and the PI-degree of $C_R(b)$. But we still have to relate the $\deg_C b$ to $\deg_C \delta$. This is done in part (4), where the minimal polynomials of δ and of $\operatorname{ad}(b)$ are nicely related. By combining all these, we obtain our desired bound for the PI-degree of R in part (5). The second difficulty (2) mentioned in Section 0 is thus solved and an application of theory of orthogonal completeness yields the main result:

THEOREM 2: *Let R be a semiprime ring with extended centroid C and δ a continuous derivation of R which is integral over C. If the subring $R^{(\delta)}$ of constants of δ in R is a PI-ring, then so is R. Furthermore, if $R^{(\delta)}$ satisfies a PI with coefficients ± 1 of degree t, then R satisfies $S_{st}(X_1, \ldots, X_{st})$, where s is the integral degree of δ over C.*

Strictly speaking, this is merely the semiprime version of part (5) of Theorem 1. The theory of orthogonal completeness actually yields a complete semiprime version of Theorem 1: Q can be decomposed into a direct product so that δ behaves on each factor just as in the prime case. For simplicity, let us be content with the above sketch. Granted Theorem 1, we first give the proof of Theorem 2, which is somewhat routine now.

2. Proof of Theorem 2

As our tool here is Beidar and Mikhalëv's theory of orthogonal completeness [2], we recall the related notions and the main result of this theory: Our logical symbols are: $\vee$ (*or*), $\wedge$ (*and*), $\neg$ (*not*), $\rightarrow$ (*if ..., then ...*), $\forall$ (*for all...*), $\exists$ (*there* *exists ...*) and $=$ (*equals*). We assume some familiarity with the basic notions of the first order logic with equality such as formulae and sentences (i.e., formulae without free variables). By a **language**, we mean a set of non-logical symbols (or proper symbols). The language of ring theory consists of two binary function symbols $+$ (*plus*), $\cdot$ (*times*) and a constant symbol 0. The language $\mathcal{L}$ we need here is the language of ring theory expanded by adjoining a function symbol δ intended to denote the derivation δ under consideration. For clarity, we also adopt the convention of omitting the multiplication function symbol $\cdot$ in writing formulae. The concept of **Horn** formulae is defined inductively as follows:

(1) An atomic formula is a Horn formula.

(2) A disjunction of negated atomic formulae is a Horn formula.

(3) If $\mathbf{\Theta}_1, \ldots, \mathbf{\Theta}_n, \mathbf{\Theta}$ are atomic formulae, then the formula $(\mathbf{\Theta}_1 \wedge \cdots \wedge \mathbf{\Theta}_n) \to \mathbf{\Theta}$ is a Horn formula.

(4) If v is a variable and $\mathbf{\Theta}$ is a Horn formula, then $\forall v \mathbf{\Theta}$ and $\exists v \mathbf{\Theta}$ are also Horn formulae.

(5) If $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$ are Horn formulae, then so is $\mathbf{\Theta}_1 \wedge \mathbf{\Theta}_2$.

(6) All Horn formulae are obtained in this way.

A sentence of $\mathcal{L}$ is said to be **hereditary**, if its truth on any given ring implies its truth on any direct summand of this given ring. The main result of the theory of orthogonal completeness for semiprime rings is the following:

THEOREM (Beidar and Mikhalëv [2]): *Let R be a semiprime ring with extended centroid C. Assume that R is orthogonally complete with respect to the Boolean ring B consisting of all idempotents in C. Let $\mathcal{L}$ be the language as described above.* (1) *Let $\mathbf{\Theta}$ be a sentence of $\mathcal{L}$ which is hereditary and whose negation $\neg\mathbf{\Theta}$ is logically equivalent to a Horn sentence. If $\mathbf{\Theta}$ holds on R, then $\mathbf{\Theta}$ also holds on R/P for any minimal prime ideal P of R.* (2) *Let $\mathbf{\Theta}$ be logically equivalent to a Horn sentence of $\mathcal{L}$. If $\mathbf{\Theta}$ holds on R/P for each minimal prime ideal P of R, then $\mathbf{\Theta}$ also holds on R.*

We are now ready to give the proof of Theorem 2:

Proof of Theorem 2: Let R be as stated in Theorem 2. The symmetric Martindale quotient ring Q of R is orthogonally complete and satisfies the hypothesis (H) by Fact B. Let $f(X_1, \ldots, X_t)$ be a PI of $R^{(\delta)}$ with coefficients all ± 1. Let $\mathbf{f}(\mathbf{x}_1, \ldots, \mathbf{x}_t)$ be the expression of $f(X_1, \ldots, X_t)$ in the language $\mathcal{L}$. Let $\mathbf{l}(\mathbf{x})$ denote the expression $\sum_{i=1}^{k} \mathbf{a}_i \mathbf{x} \mathbf{b}_i$. Consider the following formulae in our language:

$$\mathbf{\Theta}_k : \forall \mathbf{x}([\mathbf{b}, \mathbf{l}(\mathbf{x})] = 0),$$

$$\mathbf{\Psi}_n : \delta(\mathbf{c}_1) = 0 \wedge \cdots \wedge \delta(\mathbf{c}_n) = 0 \wedge \forall \mathbf{x}\Big(\sum_{j=1}^{n} \mathbf{c}_j \delta^j(\mathbf{x}) = [\mathbf{b}, \mathbf{x}]\Big),$$

$$\mathbf{\Phi}_{k,n} : \forall \mathbf{y}_1 \cdots \forall \mathbf{y}_t \Big(\mathbf{f}\Big(\sum_{j=1}^{n} \mathbf{c}_j \delta^{j-1}(\mathbf{l}(\mathbf{y}_1)), \ldots, \sum_{j=1}^{n} \mathbf{c}_j \delta^{j-1}(\mathbf{l}(\mathbf{y}_t))\Big) = 0\Big).$$

If $\mathbf{b}, \mathbf{c}_j, \mathbf{a}_i, \mathbf{b}_i$ denote $b, c_j, a_i, b_i \in Q$ respectively, then $\mathbf{\Theta}_k$ asserts that b commutes with $l(x) \stackrel{\text{def}}{=} \sum_{j=1}^{k} a_j x b_j$ for all $x \in Q$, $\mathbf{\Psi}_n$ asserts that each $c_j \in Q^{(\delta)}$ and $\sum_{j=1}^{n} c_j \delta^j = \operatorname{ad}(b)$, and finally $\mathbf{\Phi}_{k,n}$ asserts that the displayed identity in (H) holds on Q. Thus the hypothesis (H) is expressed in the language $\mathcal{L}$ by

$$(*) \qquad \forall \mathbf{b} \forall \mathbf{a}_1 \cdots \mathbf{a}_k \forall \mathbf{b}_1 \cdots \mathbf{b}_k \forall \mathbf{c}_1 \cdots \mathbf{c}_n (\mathbf{\Theta}_k \wedge \mathbf{\Psi}_n \to \mathbf{\Phi}_{k,n}).$$

This sentence is obvious hereditary. Its negation is equivalent to:

$$\exists \mathbf{b} \exists \mathbf{a}_1 \cdots \mathbf{a}_k \exists \mathbf{b}_1 \cdots \mathbf{b}_k \exists \mathbf{c}_1 \cdots \mathbf{c}_n (\mathbf{\Theta}_k \wedge \mathbf{\Psi}_n \wedge \neg \mathbf{\Phi}_{k,n}).$$

This negation is Horn, since $\mathbf{\Theta}_k$, $\mathbf{\Psi}_n$ and $\neg\mathbf{\Phi}_{k,n}$ are all Horn. By the first part of Beidar and Mikhalëv's Theorem cited above, $(*)$ also holds on any Q/P, where P is any arbitrary minimal prime ideal of Q. In other words, the hypothesis (H) is true in Q/P for any minimal prime ideal P of Q. Let q be the algebraic degree of δ over the extended centroid of Q/P. By (5) of Theorem 1, Q/P satisfies the identity $S_{qt}(X_1, \ldots, X_{qt})$. Let $s = \deg_C \delta$. Since $s \geq q$, Q/P also satisfies $S_{st}(X_1, \ldots, X_{st})$. This is true for any minimal prime ideal P of Q. Let $S_{st}(X_1, \ldots, X_{st})$ be expressed in our language by the expression $\mathbf{S}_{st}(\mathbf{x}_1, \ldots, \mathbf{x}_{st})$. The sentence

$$\forall \mathbf{x}_1 \cdots \mathbf{x}_{st} (\mathbf{S}_{st}(\mathbf{x}_1, \ldots, \mathbf{x}_{st}) = 0)$$

is obviously Horn and holds on Q/P for any minimal prime ideal P of Q. By the second part of Beidar and Mikhalëv's Theorem, this sentence also holds on Q. Thus, Q also satisfies $S_{st}(X_1, \ldots, X_{st})$ as asserted.

3. Proof of Theorem 1

We recall that the socle of R, denoted by $\operatorname{soc}(R)$, is defined to be the sum of all minimal right ideals of R. This is right-left symmetric for a prime ring R: $\operatorname{soc}(R)$ is also equal to the sum of all minimal left ideals of R. The following fact analyzes GPIs of prime rings and plays a crucial role here:

FACT C: Let R be a prime ring with extended centroid C and satisfy a multilinear GPI $f(X_1, \ldots, X_t)$ with coefficients in the symmetric Martindale quotient ring Q. Write $f(X_1, \ldots, X_t)$ in the form

$$\sum_{i=1}^{m} a_i X_1 g_i(X_2, \ldots, X_t) + h(X_1, \ldots, X_t),$$

where $a_i \in Q$ are C-independent modulo $\mathrm{soc}(Q)$ and where $h(X_1, \ldots, X_t)$ consists of a sum of monomials in the form $b_1 X_{i_1} b_2 X_{i_2} \cdots b_t X_{i_t} b_{t+1}$ with $b_1 \in \mathrm{soc}(Q)$ or $i_1 \neq 1$. Then $h(X_1, \ldots, X_t)$ and all $g_i(X_2, \ldots, X_t)$ are also GPIs of R.

Proof: By Theorem 6.4.1 [3] or Theorem 2 [5], R and Q satisfy the same GPIs with coefficients in Q. We may thus assume that $R = Q$. We also assume $t \geq 2$, for otherwise $f(X_1, \ldots, X_t)$ would be linear and hence trivial by [17]. For $x_2, \ldots, x_t \in \mathrm{soc}(R)$, since $\mathrm{soc}(R)R \subseteq R$ and also since $h(X_1, \ldots, X_t)$ consists of monomials in the form $b_1 X_{i_1} b_2 X_{i_2} \cdots b_t X_{i_t} b_{t+1}$ with $b_1 \in \mathrm{soc}(R)$ or $i_1 \neq 1$, we may write

$$h(X_1, x_2, \ldots, x_t) = \sum_{j=1}^{n} c_j X_1 d_j,$$

where $c_j \in \mathrm{soc}(R)$, $d_j \in R$. We may further assume that $c_j \in \mathrm{soc}(R)$ are C-independent by rewriting c_j as linear combinations of a C-basis of $\mathrm{soc}(R)$. Setting $b_i \overset{\mathrm{def}}{=} g_i(x_2, \ldots, x_t)$, we have

$$\begin{aligned} f(X_1, x_2, \ldots, x_t) &= \sum_{i=1}^{m} a_i X_1 g_i(x_2, \ldots, x_t) + h(X_1, x_2, \ldots, x_t) \\ &= \sum_{i=1}^{m} a_i X_1 b_i + \sum_{j=1}^{n} c_j X_1 d_j. \end{aligned}$$

This is a linear GPI for R and must be trivial by [17]. Since $c_j \in \mathrm{soc}(R)$ and since a_i are assumed to be C-independent modulo $\mathrm{soc}(R)$, we see that $b_i = g_i(x_2, \ldots, x_t) = 0$ for each i. But $x_2, \ldots, x_t \in \mathrm{soc}(R)$ are arbitrary. So each $g_i(X_2, \ldots, X_t)$ is a GPI for $\mathrm{soc}(R)$ and hence for R by Theorem 2 of [5] (or Theorem 6.4.1 [3]).

Now, we are ready for the proof of Theorem 1: From now on, R is a prime ring and δ is a continuous derivation algebraic over C such that the hypothesis (H) is satisfied. For easy reference, we reproduce the proof of part (1) in [12]:

Proof of (1) of Theorem 1: In the case of $\mathrm{char}\, R = 0$, if δ is outer, then all δ^s, $s \geq 0$, are regular words in δ. Since δ is algebraic over C, R satisfies a differential identity $\delta^m(X) + \alpha_1\delta^{m-1}(X) + \cdots + \alpha_m\delta(X)$. By Kharchenko's Theorem 1 [12], R satisfies the GPI $X_m + \alpha_1 X_{m-1} + \cdots + \alpha_m X_0$, where X_i are distinct indeterminates. This is absurd. So δ is X-inner, i.e., $\delta = \mathrm{ad}(b)$ for some $b \in Q$.

In the case of $\operatorname{char} R = p > 0$, if δ^{p^i}, $0 \le i < m$, are C-independent modulo X-inner derivations, then for given $0 \le s < p^m$, by writing $s = s_0 + s_1 p + \cdots + s_{m-1}p^{m-1}$ with each $0 \le s_i < p$, we see that $\delta^s = (\delta^{p^{m-1}})^{s_{m-1}} \cdots (\delta^p)^{s_1}\delta^{s_0}$ is a regular word in derivations δ^{p^i}, $0 \le i < m$, linearly ordered by $\delta < \delta^p < \cdots < \delta^{p^{m-1}}$. Again, by Corollaries 2 and 3 to Kharchenko's Theorem 1 [12], any differential identity of the form

$$\delta^n(X) + \alpha_1\delta^{n-1}(X) + \cdots + \alpha_n X, \quad n < p^m,$$

yields the GPI $X_n + \alpha_1 X_{n-1} + \cdots + \alpha_n X_0$, where X_i are distinct indeterminates, which implies $R = 0$, absurd. Therefore, the algebraic degree over C of δ must be $\ge p^m$. So there is a largest $m \ge 0$ such that δ^{p^i}, $0 \le i < m$, are C-independent modulo X-inner derivations. Then δ^{p^m} is a C-linear combination of δ^{p^i}, $0 \le i < m$, modulo X-inner derivations:

$$\delta^{p^m} + \alpha_1\delta^{p^{m-1}} + \cdots + \alpha_m\delta = \mathrm{ad}(b)$$

for some $\alpha_i \in C$ and $b \in Q$. By the minimality of m we see that $\alpha_i \in C^{(\delta)}$ and $b \in Q^{(\delta)}$, as asserted.

In the sequel, we will need the fact that δ^s, $0 \le s < p^m$, are distinct regular words in derivations δ^{p^j}, $0 \le j < m$ (in the case of $\operatorname{char} R = p > 0$). It is also interesting to observe that any powers of δ may be uniquely written as a C-linear combination of δ^s, $0 \le s < p^m$.

Proof of (2) of Theorem 1: Since R and its symmetric Martindale quotient ring Q satisfy the same differential identities by Theorem 2 of [16], Q also satisfies the hypothesis (H). We may thus assume $Q = R$ here. For notational convenience, we set

$$H \stackrel{\text{def}}{=} \begin{cases} \text{the identity map} & \text{if } \operatorname{char} R = 0, \\ \delta^{p^m-1} + \alpha_1\delta^{p^{m-1}-1} + \cdots + \alpha_m & \text{if } \operatorname{char} R = p > 0. \end{cases}$$

So $\delta H = \mathrm{ad}(b)$ always holds no matter whether $\operatorname{char} R = 0$ or $\operatorname{char} R = p > 0$. Let $f(X_1, \ldots, X_t)$ denote the polynomial asserted to exist in (H). In terms of H, the hypothesis (H) says that if b commutes with a linear GP $l(X)$ in the sense that $[b, l(x)] = 0$ for all $x \in R$, then R satisfies the differential identity $f\Big(H(l(Y_1)), \ldots, H(l(Y_t))\Big)$. Let $g(X)$ be the minimal polynomial of b over C.

Claim 1: $f(\hat{g}_b(X_1),\ldots,\hat{g}_b(X_t))$ is a GPI of R: By Fact A, b commutes with the linear polynomial $\hat{g}_b(X)$. By the hypothesis (H), the differential identity

$$(\dagger)\qquad f(H(\hat{g}_b(X_1)),\ldots,H(\hat{g}_b(X_t)))$$

holds for Q. If $\operatorname{char} R = 0$, then H is the identity map and ($\dagger$) is already the asserted GPI. So let $\operatorname{char} R = p > 0$ and hence $H = \delta^{p^m-1}+\alpha_1\delta^{p^{m-1}-1}+\cdots+\alpha_m$. Since $\hat{g}_b(X_i)$ is linear in X_i, $H(\hat{g}_b(X_i))$ is equal to $\hat{h}_b(\delta^{p^m-1}(X_i))$ plus terms with $\delta^j(X_i)$, $j < p^m - 1$. So ($\dagger$) is equal to

$$f(\hat{g}_b(\delta^{p^m-1}(X_1)),\ldots,\hat{g}_b(\delta^{p^m-1}(X_t)))$$

plus terms with $\delta^j(X_i)$, $j < p^m - 1$. As said in part (1), δ^s, $0 \le s < p^m$, are regular words in δ^{p^i}, $0 \le i \le m-1$. Our expansion of ($\dagger$) above is thus reduced. We apply Kharchenko's Theorem 2 [13] to ($\dagger$) by substituting new indeterminates Y_i for $\delta^{p^m-1}(X_i)$ and 0 for all other $\delta^j(X_i)$, $0 < j < p^m - 1$. The asserted GPI (with X_i replaced by Y_i) follows. This proves our Claim 1.

If $g(X)$ is of degree n, then $1, b, \ldots, b^{n-1}$ are C-independent and the asserted GPI of Claim 1, involving the term $X_1 b^{n-1}\cdots X_t b^{n-1}$ nontrivially, must be nontrivial. It follows from Martindale's Theorem [17] that R has nonzero socle and its skew field is finite C-dimensional.

Let $h(X)$ be the monic polynomial, with coefficients in C, of minimal degree $n \ge 1$ such that $h(b) \in \operatorname{soc}(R)$. Since b is algebraic over C, such a polynomial $h(X)$ does exist. We may call $h(X)$ the **minimal** polynomial over C of b modulo $\operatorname{soc}(R)$. By Litoff's Theorem [7], there exists an idempotent $e \in \operatorname{soc}(R)$ such that $h(b) \in eRe$. We refine our first claim as follows:

Claim 2: $f(\hat{h}_b((1-e)X_1(1-e)),\ldots,\hat{h}_b((1-e)X_t(1-e)))$ is a GPI of R: For $y \in R$, we have

$$[b, \hat{h}_b((1-e)y(1-e))] = [h(b), (1-e)y(1-e)] = [eh(b)e, (1-e)y(1-e)] = 0.$$

So the linear polynomial $\hat{h}_b((1-e)Y(1-e))$ commutes with b. By the hypothesis (H), the differential identity

$$(\ddagger)\qquad f(H(\hat{h}_b((1-e)Y_1(1-e))),\ldots,H(\hat{h}_b((1-e)Y_t(1-e))))$$

holds for Q. If $\operatorname{char} R = 0$, then H is the identity map and ($\ddagger$) is the asserted GPI (with X_i replaced by Y_i). So we assume $\operatorname{char} R = p > 0$. In this case, $H = \delta^{p^m-1}+\alpha_1\delta^{p^{m-1}-1}+\cdots+\alpha_m$. Since $\hat{h}_b(X_i)$ is linear in X_i, $H(\hat{h}_b((1-e)Y_i(1-e)))$

is equal to $\hat{h}_b((1-e)\delta^{p^m-1}(Y_i)(1-e))$ plus terms with $\delta^j(Y_i)$, $j < p^m - 1$. So (‡) is equal to

$$f(\hat{h}_b((1-e)\delta^{p^m-1}(Y_1)(1-e)), \ldots, \hat{h}_b((1-e)\delta^{p^m-1}(Y_t)(1-e)))$$

plus terms with some $\delta^j(Y_i)$, $j < p^m - 1$. As said in part (1), δ^s, $0 \le s < p^m$, are regular words in δ^{p^i}, $0 \le i \le m-1$. So our expansion of (‡) above is reduced and Kharchenko's Theorem may be applied. Replacing $\delta^{p^m-1}(Y_i)$ by new indeterminates X_i and all other $\delta^j(Y_i)$, $0 < j < p^m - 1$ by 0, we obtain the asserted GPI. This proves our claim.

If $e = 1$, then R is finite C-dimensional and is hence a PI-ring. We are done in this case. We thus assume towards a contradiction that $e \ne 1$. Let $k \overset{\text{def}}{=} \deg h(X)$. Recall that $f(X_1, \ldots, X_t)$ has coefficients ± 1. By reindexing, we may assume that the monomial $X_1 X_2 \cdots X_t$ occurs in $f(X_1, \ldots, X_t)$. Write

$$f(X_1, \ldots, X_t) = X_1 f_1(X_2, \ldots, X_t) + h_1(X_1, \ldots, X_t),$$

where $h_1(X_1, \ldots, X_t)$ consists of monomials starting with some X_j, $j > 1$. Claim 2 says that

$$\begin{aligned}\hat{h}_b((1-e)X_1(1-e)) f_1(\hat{h}_b((1-e)X_2(1-e)), \ldots, \hat{h}_b((1-e)X_t(1-e))) \\ + h_1(\hat{h}_b((1-e)X_1(1-e)), \ldots, \hat{h}_b((1-e)X_t(1-e)))\end{aligned}$$

is a GPI for R. The expression $\hat{h}_b((1-e)X_1(1-e))$ is a sum of terms starting with $b^j(1-e)X_1$, $j = 0, \ldots, k-1$, and the only term starting with $b^{k-1}(1-e)X_1$ is $b^{k-1}(1-e)X_1(1-e)$. Observe that $b^{k-1}(1-e), \ldots, b(1-e), (1-e)$ are C-independent modulo $\mathrm{soc}(R)$, for $\sum_{i=1}^k \gamma_i b^{k-i}(1-e) \in \mathrm{soc}(R)$, $\gamma_i \in C$, implies $\sum_{i=1}^k \gamma_i b^{k-i} = \sum_{i=1}^k \gamma_i b^{k-i}(1-e) + \sum_{i=1}^k \gamma_i b^{k-i} e \in \mathrm{soc}(R)$ and hence all $\gamma_i = 0$ by the minimality of $k = \deg h(X)$. Apply Fact C to the identity of Claim 2 thus expressed above. The expression after $b^{k-1}(1-e)X_1$ yields the GPI of R:

$$(1-e) f_1(\hat{h}_b((1-e)X_2(1-e)), \ldots, \hat{h}_b((1-e)X_t(1-e))).$$

The polynomial $f_1(X_2, \ldots, X_t)$ has the degree $t-1$ and contains the monomial $X_2 X_3 \cdots X_t$. Write

$$f_1(X_2, \ldots, X_t) = X_2 f_2(X_3, \ldots, X_t) + h_2(X_2, \ldots, X_t),$$

where $h_2(X_2, \ldots, X_t)$ consists of monomials starting with some X_j, $j > 2$. We observe similarly the C-independence of $(1-e)b^{k-1}(1-e), \ldots, (1-e)b(1-e), (1-e)$

modulo $\mathrm{soc}(R)$. An analogous application of Fact C yields the following GPI of R:

$$f_2(\hat{h}_b((1-e)X_3(1-e)), \ldots, \hat{h}_b((1-e)X_t(1-e))).$$

Continuing in this manner, we finally obtain the GPI $\hat{h}_b((1-e)X_t(1-e))$ of R. But Fact C applies to this GPI as well and yields a desired contradiction.

We digress here for the following fact, which is implicit in [13]:

FACT D: Let R be a prime PI-ring with center Z and extended centroid C. Suppose that $\sum_{i=1}^{n} \phi_i(\Delta_j(z_k))a_i = 0$ for all $z_k \in Z$, where $a_i \in Q$, where $\phi_i(Z_{jk})$ are polynomials over C in commuting variables Z_{jk} and where Δ_j are distinct regular words in an ordered set of independent outer derivations of Q (modulo X-inner derivations). Then $\sum_{i=1}^{n} \phi_i(z_{jk})a_i = 0$ for all $z_{jk} \in C$.

Proof: A derivation of Q vanishing on C must be X-inner ([13], p. 68]). So the restriction to Z of an independent set of X-outer derivations of Q (modulo X-inner derivations) yields a C-independent set of derivations of C, which can be ordered in the same way. In this sense, restrictions to C of regular derivation words for Q give rise to regular derivation words for C.

Pick a C-basis $\{b_1, \ldots, b_t\}$ of Q and write $a_i = \sum_{m=1}^{t} \beta_{im} b_m$, where $\beta_{im} \in C$. By the C-independence of b_i, our assumption that

$$\sum_{i=1}^{n} \sum_{m=1}^{t} \beta_{im} b_m \phi_i(\Delta_j(z_k)) = 0$$

for all $z_k \in Z$ implies that for each m, $\sum_{i=1}^{n} \beta_{im}\phi_i(\Delta_j(z_k)) = 0$ for all $z_k \in Z$. Note that C is the quotient field of Z. Applying Kharchenko's Theorem to these differential identities of Z, we see that each $\sum_{i=1}^{n} \beta_{im}\phi_i(Z_{jk})$ is a GPI for C. Thus $\sum_{i=1}^{n} \phi_i(z_{jk})a_i = 0$ for all $z_{jk} \in C$, as asserted. ∎

We are now ready to give the

Proof of (3) of Theorem 1: If $\mathrm{char}\, R = 0$, then $\delta = \mathrm{ad}(b)$ and hence $R^{(\delta)} = C_R(b)$. The conclusion is trivially true in this case. We may thus assume $\mathrm{char}\, R = p > 0$. It is clear that $R^{(\delta)} \subseteq C_R(b)$, implying that each PI for $C_R(b)$ is also satisfied by $R^{(\delta)}$. For the converse, let $h(X_1, \ldots, X_s)$ be a PI for $R^{(\delta)}$ and fix $y_1, \ldots, y_s \in C_R(b)$. We aim to show $h(y_1, \ldots, y_s) = 0$: The polynomial identity $h(X_1, \ldots, X_s)$ of $R^{(\delta)}$ is nontrivial, for otherwise, there will be nothing to prove. It follows from part (2) of Theorem 1 that R is a prime PI-ring. By continuity of δ, there exists a two-sided ideal $I \neq 0$ of R such that $\delta(I) \subseteq R$. A simple induction shows that

$\delta(I^k) \subseteq I^{k-1}$ and hence $\delta^k(I^k) \subseteq R$ for $k \geq 1$. By [19], C is the quotient field of the center $Z(I)$ of I. Pick $\xi \in Z(I)$ such that $\xi\alpha_i \in Z(I)$ for all α_i in the expression $H \overset{\text{def}}{=} \delta^{p^m-1} + \alpha_1\delta^{p^{m-1}-1} + \cdots + \alpha_m$. Note that $\delta(\xi^p) = 0$. Set $\gamma \overset{\text{def}}{=} \xi^{2p^m}$. For arbitrary β in the center $Z(R)$ of R, we have

$$\begin{aligned}\gamma H(\beta y_i) =& \xi^{2p^m} H(\beta y_i)\\ =& \xi^{p^m}\delta^{p^m-1}(\xi^{p^m}\beta y_i) + \xi^{p^m}\alpha_1\delta^{p^{m-1}-1}(\xi^{p^m}\beta y_i) + \cdots + \xi^{2p^m}\alpha_m\beta y_i \in R.\end{aligned}$$

Also, $\delta(\gamma H(\beta y_i)) = \gamma\delta H(\beta y_i) = \gamma[b, \beta y_i] = 0$. This implies that $\gamma H(\beta y_i) \in R^{(\delta)}$. Thus we see that $h(\gamma H(\beta_1 y_1), \ldots, \gamma H(\beta_s y_s)) = 0$ for arbitrarily given $\beta_1, \ldots, \beta_s \in Z(R)$. Apply Fact D to this identity by substituting distinct indeterminates Z_i for $\delta^{p^m-1}(\beta_i)$ and 0 for all other $\delta^j(\beta_i)$, $j \neq p^m - 1$. We obtain the identity $f(\gamma Z_1 y_1, \ldots, \gamma Z_s y_s)$ for C. By setting $Z_i = \gamma^{-1}$, $f(y_1, \ldots, y_s) = 0$ follows as asserted.

Proof of (4) of Theorem 1: Let $p(X)$ and $\mu(X)$ be respectively the minimal polynomials over C of δ and of $\mathrm{ad}(b)$. Clearly, both polynomials have their coefficients in the subfield $C^{(\delta)} \overset{\text{def}}{=} \{\alpha \in C \mid \delta(\alpha) = 0\}$. Write $p(X) = \sum_{i=0}^{s} \ell(X)^i q_i(X)$ with each $\deg q_i(X) < \deg \ell(X) = p^m - 1$. (Divide $p(X)$ by $\ell(X)$. The remainder is $q_0(X)$. Divide the quotient obtained by $\ell(X)$ again. The remainder is $q_1(X)$ and so on.) Since both $p(X)$ and $\ell(X)$ have their coefficients in $C^{(\delta)}$, so has each $q_i(X)$. For $x \in R$,

$$0 = p(\delta)(x) = \sum_{i=0}^{s} \ell(\delta)^i q_i(\delta)(x) = \sum_{i=0}^{s} \mathrm{ad}(b)^i q_i(\delta)(x).$$

That is,

$$\sum_{i=0}^{s} \mathrm{ad}(b)^i (q_i(\delta)(X))$$

is a differential identity for R. Let k be the degree of $q_s(X)$ and write $q_s(X) = \beta_s X^k + \cdots$ with $\beta_s \neq 0$. Suppose that $q_j(X) = \cdots + \beta_j X^k + \cdots$ for $j < s$. Since δ^s, $0 \leq s < p^m - 1$, are distinct regular words in δ^{p^i}, $0 \leq i \leq m - 1$, Kharchenko's Theorem applied to the above differential identity yields the GPI $\sum_{j=0}^{s} \beta_j \,\mathrm{ad}(b)^j(X)$ for R. So $\mathrm{ad}(b)$ also satisfies the polynomial $\sum_{j=0}^{s} \beta_j X^j$. Hence $\deg \mu(X) \leq s$ by the minimality of $\mu(X)$. Thus

$$\deg p(X) \geq s(\deg \ell(X)) \geq \deg \mu(X) \deg \ell(X).$$

On the other hand, $\mu(\ell(\delta)) = \mu(\mathrm{ad}(b)) = 0$. By the minimality of $p(X)$, $p(X)$ divides $\mu(\ell(X))$ and hence $\deg p(X) \leq \deg \mu(X) \deg \ell(X)$. By comparing degrees, $p(X) = \mu(\ell(X))$ follows.

For the last statement, it suffices to show that $\deg_C b \leq \deg_C \mathrm{ad}(b)$: Set $k = \deg_C \mathrm{ad}(b)$. Then there exist $\beta_1, \ldots, \beta_{k-1} \in C$ such that

$$\mathrm{ad}(b)^k(x) + \beta_1 \mathrm{ad}(b)^{k-1}(x) + \cdots + \beta_{k-1} \mathrm{ad}(b)(x) = 0$$

for all $x \in R$. Expanding this and collecting terms according to their left coefficients $1, b, \ldots, b^k$, we obtain the linear GPI $b^k X + \sum_{i=1}^k b^{k-i} X a_i$ for R, where $a_i \in Q$. In view of Lemma 1 [17], the elements $1, b, \ldots, b^k$ are C-dependent, implying that $\deg_C b \leq \deg_C \mathrm{ad}(b)$, as asserted.

Proof of (5) of Theorem 1: Firstly, since $R^{(\delta)}$ and $C_R(b)$ satisfy the same PIs over C by part (3), we may simply assume $\delta = \mathrm{ad}(b)$. By part (2), R is a PI-ring. By Posner's Theorem (p. 57 [11]), the center $Z(R)$ of R is nonzero and, moreover, both Q and C are merely the localizations of R and $Z(R)$ at $Z(R) \setminus \{0\}$ respectively. A direct computation shows that $C_Q(b) = C_R(b)C$. Obviously, $C_R(b)$ and $C_R(b)C$ satisfy the same multilinear PIs. Via linearization, if $C_R(b)$ (or $C_R(b)C$ respectively) satisfies a PI, then it also satisfies a multilinear PI of the same degree. Therefore, the minimum of degrees of PIs satisfied by $C_R(b)$ is actually equal to the PI-degree of $C_Q(b) = C_R(b)C$. We may thus assume that $R = Q = RC$.

Let F denote the algebraic closure of C. Then $Q \otimes_C F$ is the ring of n by n matrices over F for some finite $n \geq 1$. A direct computation shows that $C_{Q \otimes_C F}(b \otimes 1) = C_Q(b) \otimes_C F$. Again, Q, $Q \otimes_C F$ are of the same PI-degree and $C_Q(b)$, $C_{Q \otimes_C F}(b \otimes 1)$ are also of the same PI-degree. Replacing R by $Q \otimes_C F$, we shall assume that R is the ring of $n \times n$ matrices over an algebraically closed field C. By Amitsur-Levitzki's theorem (p. 21 [11]), R satisfies $S_{2n}(X_1, \ldots, X_{2n})$. Let $q \stackrel{\text{def}}{=}$ the PI-degree of $C_R(b)$. It suffices to show that $q \cdot \deg_C b \geq 2n$.

For simplicity, we adopt the following terminology from Herstein's *Topics in Algebra* (p. 273 [8]): A $k \times n$ matrix is said to be of size k by n. A k by k square matrix is said to be of size k only. By the principal diagonal of a $k \times n$ matrix, we mean the entries in the positions (i, i), and by the super diagonal of a $k \times n$ matrix, we mean the entries in the positions $(i, i+1)$, wherever they are defined. Consider matrices over a field C. By a Jordan block belonging to $\lambda \in C$, we mean a square matrix with λ on the principal diagonal, with 1 on the super diagonal and with 0 elsewhere:

$$J = \begin{pmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix}.$$

A square matrix is said to be in the Jordan normal form, if it is of the form

$$\begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k \end{pmatrix}, \tag{1}$$

where each J_i is a Jordan block belonging to some $\lambda_i \in C$.

CASE 1: The minimum polynomial of b is of the form $(x-\lambda)^\nu$: We may assume that b is in the Jordan normal form (1), where each J_i is a Jordan block of size m_i belonging to λ. By reordering, we may assume that $\nu = m_1 \geq m_2 \geq \cdots \geq m_k$. Any element $a \in R$ may be written in the form

$$a = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ A_{21} & A_{22} & \cdots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} & A_{k2} & \cdots & A_{kk} \end{pmatrix}, \tag{2}$$

where each A_{ij} is a matrix of size m_i by m_j. For each $1 \leq i \leq j \leq k$, let $e_{ij} \in R$ be the matrix, written in the form (2) above, where $A_{st} = 0$ for $(s,t) \neq (i,j)$ and where A_{ij} is the m_i by m_j matrix with 1 on its principal diagonal and with 0 elsewhere. Using the assumption $m_i \geq m_j$ for $i \leq j$, we verify easily that $e_{ij} \in C_R(b)$ for each $1 \leq i \leq j \leq k$. Also, for $i \leq j$ and $s \leq t$, a direct computation shows that

$$e_{ij}e_{st} = \begin{cases} 0, & \text{if } j \neq s; \\ e_{it}, & \text{if } j = s. \end{cases}$$

Assume towards a contradiction that the PI-degree q of $C_R(b)$ is $< 2k$. Without loss of generality, we may assume that $C_R(b)$ satisfies a multilinear PI of the form

$$\mu(X_1, X_2, \ldots, X_{2k-1}) = X_1X_2\cdots X_{2k-1} + \cdots,$$

where the dots denote the sum of terms different from $X_1X_2\cdots X_{2k-1}$. We set

$$X_1 = e_{11}, X_2 = e_{12}, X_3 = e_{22}, \ldots, X_{2k-1} = e_{k,k}.$$

We have $\mu(e_{11}, e_{12}, e_{22}, \ldots, e_{k,k}) = e_{1k} \neq 0$, a contradiction. So the PI-degree q of $C_R(b)$ is $\geq 2k$. With this, we have $q\nu \geq 2k\nu \geq 2(m_1 + m_2 + \cdots + m_k) = 2n$, as asserted.

CASE 2: The minimum polynomial of b is of the form $(x-\lambda_1)^{m_1}\cdots(x-\lambda_s)^{m_s}$, where the λ_i's are distinct. Set $k_0 = 0$. We may assume that $J_{k_{i-1}+1},\ldots,J_{k_{i-1}+k_i}$ consist of all Jordan blocks belonging to λ_i and n_i is the sum of their sizes. We may thus write b in the form

$$b = \begin{pmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_s \end{pmatrix},$$

where each B_i is the square matrix of size n_i:

$$B_i = \begin{pmatrix} J_{k_{i-1}+1} & 0 & \cdots & 0 \\ 0 & J_{k_{i-1}+2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{k_{i-1}+k_i} \end{pmatrix}.$$

Let R_i be the ring of n_i by n_i matrices over C and let q_i be the PI-degree of $C_{R_i}(B_i)$. The minimal polynomial of B_i is obviously $(X-\lambda_i)^{m_i}$, where m_i is the maximum of the sizes of J_s, $k_{i-1} < s < k_{i-1}+k_i$. We have $q_i m_i \geq 2n_i$ by Case 1. Any $a \in R$ may be written in the form

$$a = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ A_{s1} & A_{s2} & \cdots & A_{ss} \end{pmatrix},$$

where A_{ij} is the matrix of size n_i by n_j. If $A_{ii} \in C_{R_i}(B_i)$ and all other $A_{st} = 0$, then $a \in C_R(b)$. In this manner, $C_{R_i}(B_i)$ is naturally embedded as a subring of $C_R(b)$. So the PI-degree q of $C_R(b)$ is $\geq$ the PI-degree q_i of each $C_{R_i}(B_i)$. Note $\nu \overset{\text{def}}{=} \deg_C b = m_1 + \cdots + m_s$. Thus $q\nu = q(m_1+\cdots+m_s) \geq q_1 m_1 + \cdots + q_s m_s \geq 2(n_1+\cdots+n_s) = 2n$, as asserted. This completes the proof of (5).

References

[1] S. A. Amitsur, *A note on PI-rings*, Israel Journal of Mathematics **10** (1971), 210–211.

[2] K. I. Beidar and A. V. Mikhälev, *Orthogonal completeness and algebraic systems*, Uspekhi Matematicheskikh Nauk **40** (1985), 79–115.

[3] K. I. Beidar, W. S. Martindale 3rd and A. V. Mikhälev, *Rings with Generalized Identities*, Marcel Dekker, Inc., New York–Basel–Hong Kong, 1996.

[4] J. Bergen and P. Grzeszczuk, *Skew derivations whose invariants satisfy a polynomial identity*, Journal of Algebra **228** (2000), 710–737.

[5] C.-L. Chuang, *GPIs having coefficients in Utumi quotient rings*, Proceedings of the American Mathematical Society **103** (1988), 723–728.

[6] M. Cohen, *Centralizers of algebraic elements*, Communications in Algebra **6** (1978), 1505–1519.

[7] C. Faith and Y. Utumi, *On a new proof of Litoff's theorem*, Acta Mathematica Academiae Scientiarum Hungaricae **14** (1963), 369–371.

[8] I. N. Herstein, *Topics in Algebra*, 2nd ed., Xerox College Publishing, 1975.

[9] I. N. Herstein, *Rings with involution*, University of Chicago Press, Chicago, 1976.

[10] I. N. Herstein and L. Neumann, *Centralizers in rings*, Annali di Matematica Pura ed Applicata (4) **102** (1975), 37–44.

[11] N. Jacobson, *PI-algebras: An Introduction*, Lecture Notes in Mathematics **441**, Springer-Verlag, Berlin and New York, 1975.

[12] V. K. Kharchenko, *Differential identities of prime rings*, Algebra i Logika **17** (1978), 220–238; English translation: Algebra and Logika **17** (1978), 154-168.

[13] V. K. Kharchenko, *Differential identities of semiprime rings*, Algebra i Logika **18** (1979), 86–119; English translation: Algebra and Logika **18** (1979), 58–80.

[14] V. K. Kharchenko, *Automorphisms and Derivations of Associative Rings*, Kluwer Academic Publishers, Dordrecht, 1991.

[15] V. K. Kharchenko, J. Keller and S. Rodrigues-Romo, *Prime rings with PI rings of constants*, Israel Journal of Mathematics **96** (1996), 357–377.

[16] T.-K. Lee, *Semiprime rings with differential identities*, Bulletin of the Institute of Mathematics. Academia Sinica **20** (1992), 27–38.

[17] W. S. Martindale 3rd, *Prime rings satisfying a generalized polynomial identity*, Journal of Algebra **12** (1969), 576–584.

[18] S. Montgomery, *Centralizers satisfying polynomial identities*, Israel Journal of Mathematics **18** (1974), 209–219.

[19] L. H. Rowen, *Some results on the center of a ring with polynomial identity*, Proceedings of the American Mathematical Society **79** (1973), 219–223.

[20] L. H. Rowen, *General polynomial identities. II.*, Journal of Algebra **38** (1976), 380–392.

[21] M. K. Smith, *Rings with an integral element whose centralizer satisfies a polynomial identity*, Duke Mathematical Journal **42** (1975), 137–149.