ALGEBRAIC AND DIAGONABLE RINGS

M. P. DRAZIN

1. Introduction. In a well-known paper **(7)** Jacobson has shown how his structure theory for arbitrary rings can be applied to give more precise information about the so-called "algebraic" algebras. This specialization of his general theory is, however, perhaps not completely satisfying in that it deals only with algebras, i.e. rings admitting a *field* of operators, whereas neither the general structure theory nor the definition of the property of being "algebraic" seems to depend in any essential way on the precise nature of the operators.

In this paper we first show (§2) how, by suitably extending the algebraic concept to rings with arbitrary operators, Jacobson's theory of algebraic algebras can be carried over without difficulty to all "algebraic rings." Our definition of the algebraic property for arbitrary rings seems a natural one (and indeed almost inevitable if the link with π-regularity is to be preserved), and in §3 we establish some general results connected with this definition. The first of these is unspectacular, and in any case applies only to algebras; it serves chiefly as a lemma for a theorem proved later (§5). However, the second result, whose hypothesis actually excludes fields as operators, is more surprising, having the corollary that *every ring algebraic over the integers and of zero characteristic must in fact be nil*; thus the algebraic property, as defined here with respect to arbitrary operator domains, can, for some choices of the operators, and in contrast with its more usual role of "weak finite-dimensionality," be a very strong one.

In the remaining sections we investigate various related questions. Thus in §4 we generalize the familiar result that a finite-dimensional matrix algebra over an algebraically closed field must be commutative whenever every matrix in the algebra can be reduced to diagonal form by a similarity transformation (allowed in the first instance to depend on the matrix); our generalization (which is applicable to algebras over *any* field, and indeed to arbitrary rings) has a certain topical interest in view of some recent work of Motzkin and Taussky. One of the new results in §5, while again referring only to algebras, generalizes Jacobson's result that every algebraic algebra without non-zero nilpotent elements, over a finite field, is necessarily commutative; we show in particular that the conclusion remains true even if non-zero nilpotent elements exist, provided these are all central. The earlier results of §5 are of a rather curious and apparently superficial type, but do nevertheless have some unexpected implications (e.g. that, if a π-regular, or in particular algebraic, ring R has all its nilpotent elements central, then the same is true of every homomorphic image of R).

Received October 4, 1955.

We recall Herstein's result **(5)** that *if, to each element x of a given ring R, there corresponds a polynomial $p_x(\lambda)$ with integral coefficients (and possibly a constant term) such that $x - x^2p_x(x)$ lies in the centre of R, then R must be commutative.* We shall refer to this as *Herstein's theorem*, and apply it in §5 and §6, where we show how certain analogous results, and a few special cases of a related conjecture of Herstein, can be deduced from our earlier work.

2. Preliminaries. Throughout, R will denote any associative ring, not necessarily commutative or containing a unit element, admitting an arbitrary commutative ring F of operators (i.e. endomorphisms α of the additive group of R, subject to $\alpha(xy) = (\alpha x)y = x(\alpha y)$ for all $x, y \in R$); we may suppose without loss of generality that F contains the identity operator. The case of a "ring without operators" is included in this scheme on taking F to be just the ring of integers (or an appropriate quotient ring). When we refer to subrings (etc.) of R these should always be understood as sub-F-rings (etc.), i.e. as being mapped into themselves by every operator in F.

If one seeks to introduce an analogue, at this level of generality, of the property of an algebra of being "algebraic over its field of operators," one may (cf. **3)** think first of calling an element x of R algebraic over F if a positive integer n and elements $\alpha_1, \ldots, \alpha_n$ of F, not all zero, exist satisfying

$$\alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_n x^n = 0. \tag{1}$$

This of course reduces to Jacobson's definition when F is a field. However, this form is unsatisfactory from many points of view, as will become clearer below; we note for the present that it would not even enable us to carry over to rings the well-known property of algebraic algebras of having nil Jacobson radical. We therefore adopt a more stringent defining condition: we shall now call x *algebraic (over the ring F of operators)* whenever $\alpha_1, \ldots, \alpha_n$ exist as above but with the further property that the first non-vanishing α_i is the identity operator, i.e. only if x satisfies an equation of the ("lower monic") form

$$x^m + \alpha_{m+1}x^{m+1} + \ldots + \alpha_n x^n = 0; \tag{2}$$

if R happens to be an algebra, i.e. if F is a field, then this can of course always be arranged (on multiplying through by the inverse of the lowest non-zero coefficient) whenever x satisfies the formally weaker condition (1). Another equivalent form of our new definition is the following: x is algebraic if we can find a positive integer $m = m(x)$, and an element $a = a(x)$ of the subring generated by x, such that $x^m = x^m a$. We call R itself algebraic over F if each $x \in R$ is algebraic over F.

It is a straightforward matter to check that all the principal arguments and results of Jacobson's paper **(7)** on algebraic algebras are valid, with only slight verbal changes, for the wider class of algebraic rings; we omit the details.

It is important to bring out into the open a point which might otherwise give rise to misunderstandings later. Given a ring R over F, we may regard

any homomorphic image $R^* = R/T$ of R as again a ring over F by defining $\alpha x^* = (\alpha x)^*$ in the usual way. However, if we agree to regard two operators as equal relative to a given ring (which admits them both) if and only if they have the same effects on each element of the ring, then the operator set on R^* is, strictly, not F but the factor ring $F^* = F/G$, where G denotes the ideal of F consisting of all $\alpha \in F$ such that $\alpha R \leqslant T$. This distinction, vacuous when F is a field, can nevertheless be vital for more general operator rings F (particularly when their cardinals or characteristics are in question). Also, if we had chosen to define algebraic elements by means of (1), we could not have asserted that R being algebraic over F implies that R^* is algebraic over F^* (since some $x \in R$ might satisfy only equations (1) in which each coefficient $\alpha_i \in G$); however, using (2) ensures homomorphism-invariance for the algebraic property (since the identity element of F maps onto that of F^*).

In view of these remarks, it is not strictly true to say that every ring may be regarded as a ring over the ring I of integers: in fact this will be legitimate for a given ring R if and only if, for each positive integer k, an element x exists in R such that $kx \neq 0$. However, it is convenient and in practice not seriously confusing to be a little inexact in this connexion: we shall allow ourselves the customary liberty of regarding any ring R as a ring over I (rather than some quotient ring of I). Thus, for example, any algebra algebraic over a finite field of prime order will be regarded also as algebraic over the integers.

Our definition of the algebraic property via (2), while fulfilling most reasonable requirements, does have the slight technical disadvantage of carrying with it no immediately available concept of a minimal polynomial; for, among the polynomials satisfying (2), there is in general more than one of minimal "lower degree" m (even if we demand that $n - m$ be also minimal). However, at least when F is an integral domain, we can get something with most of the usual properties by returning to (1).

Let R be a ring with arbitrary operators F, and x any element of R algebraic over F. Then x satisfies an equation of the form (2), and *a fortiori* satisfies equations of the form (1), i.e. there are non-zero polynomials $f(\lambda)$ over F, without constant terms, such that $f(x) = 0$. Among such polynomials $f(\lambda)$, all those of minimal degree (there will in general be several, possibly infinitely many) will be called *minimal polynomials for x over F*. We note two relevant lemmas; the first is standard and leads immediately to the second.

LEMMA 2.1. *Let $f(\lambda)$, $g(\lambda)$ be arbitrary formal polynomials over a given commutative ring F, with leading terms $\alpha\lambda^n$, $\beta\lambda^k$ respectively. Then there are polynomials $q(\lambda)$, $r(\lambda)$ over F, with $r(\lambda)$ zero or of degree strictly less than n, such that*

$$\alpha^k g(\lambda) = q(\lambda) f(\lambda) + r(\lambda).$$

LEMMA 2.2. *Let x be any algebraic element of a given ring R over F, and let $f(\lambda)$ be any minimal polynomial for x over F, say with leading term $\alpha\lambda^n$. Then,*

given any polynomial $g(\lambda)$ over F, of degree k say, such that $g(x) = 0$, there is a polynomial $q(\lambda)$ over F (possibly with constant term) such that

$$\alpha^k g(\lambda) = q(\lambda) f(\lambda).$$

Of course Lemma 2.2 is of value only when we can be sure that $\alpha^k \neq 0$. If F is an integral domain we even have some measure of uniqueness ("up to scalar factors") for our minimal polynomials: for, if f, g are two such, with leading terms $\alpha\lambda^n$, $\beta\lambda^k$, then, by Lemma 2.2, polynomials $p(\lambda)$, $q(\lambda)$ exist such that $\alpha^k g = qf$, $\beta^n f = pg$, whence $\alpha^k\beta^n f = pqf$. Thus, for an integral domain F, since f is not identically zero, we have $\alpha^k\beta^n = p(\lambda)\, q(\lambda)$, and so $p(\lambda)$, $q(\lambda)$ must both be non-zero constants; in other words, any two minimal polynomials of a given element x algebraic over an integral domain must have a common non-zero scalar multiple.

3. Some general properties of algebraic rings. Our first theorem (which will find a use later) is a direct adaptation of a result from elementary algebraic number theory:

THEOREM 3.1. *Let F be a field, algebraic over a given subfield F_0. Then every algebra algebraic over F is algebraic over F_0.*

Proof. Let R be any ring over F, and x any element of R algebraic over F, say satisfying (2) above, with each $\alpha_i \in F$. Since F is a field, we can single out from the non-zero α_i that one, say α_q, with greatest index q, multiply through by α_q^{-1}, and write

$$x^q = \beta_1 x + \ldots + \beta_{q-1} x^{q-1},$$

where $q \geqslant m \geqslant 1$, each $\beta_i \in F$. Hence, if we denote the field $F_0(\beta_1, \ldots, \beta_{q-1})$ by K, then the algebra $K[x]$ is finite-dimensional over K, while also K is itself a finite extension of F_0 (since each β_i is algebraic over F_0). Thus $K[x]$ can be regarded as a finite-dimensional algebra over F_0 (its dimension as such being given by $\dim(K[x]: K)\dim(K: F_0)$), that is, x is algebraic over F_0, as required.

We come now to some of our principal results.

THEOREM 3.2. *Let F be any (commutative) integral domain, not a field but having a unit element, and let R be any ring algebraic over F. Then, given any element x of R, any equation of the form*

$$\alpha_m x^m + \alpha_{m+1} x^{m+1} + \ldots + \alpha_n x^n = 0 \tag{3}$$

with each $\alpha_i \in F$ and $\alpha_m \neq 0$ (and of course some such relation holds) implies the existence of a non-zero element α of F such that $\alpha x^m = 0$.

Proof. We can write $\alpha_m x^m = x^m a$, where a is in the subring of R generated by x; then $\alpha_m{}^j x^m = x^m a^j$ $(j = 1, 2, \ldots)$, and so, taking $j = m + 1$, we can find $b \in R$ such that $\alpha_m{}^{m+1} x^m = x^m b x^m$. Defining $e = x^m b$, we then have

$$\alpha_m^{m+1} x^m = e x^m, \quad \alpha_m^{m+1} e = e^2.$$

Now, for any $\beta \in F$, since R is algebraic, we can find a positive integer t_β and a polynomial $k_\beta(\lambda)$ over F such that $(\beta e)^{t_\beta} = (\beta e)^{t_\beta+1} k_\beta(\beta e)$. Also, by use of the relation $e^2 = {\alpha_m}^{m+1} e$, we can express $e^2 k_\beta(\beta e) = \theta_\beta e$ for some $\theta_\beta \in F$, so that $\beta^{t_\beta} e^{t_\beta} = \beta^{t_\beta+1} \theta_\beta e^{t_\beta}$, that is

$$0 = \beta^{t_\beta}(1 - \beta\theta_\beta)\, e^{t_\beta} = \beta^{t_\beta}(1 - \beta\theta_\beta)\, \alpha_m^{(m+1)(t_\beta-1)} e;$$

consequently, for each $\beta \in F$,

$$\beta^{t_\beta}(1 - \beta\theta_\beta)\, \alpha_m^{(m+1)t_\beta} x^m = 0.$$

Thus either $\alpha \equiv \beta^{t_\beta}(1 - \beta\theta_\beta){\alpha_m}^{(m+1)t_\beta}$ is non-zero for some $\beta \in F$, as required, or else $\beta^{t_\beta}(1 - \beta\theta_\beta){\alpha_m}^{(m+1)t_\beta} = 0$ for all $\beta \in F$; and in this latter case (since $\alpha_m \neq 0$ and F is an integral domain) we should have $1 = \beta\theta_\beta$ for each non-zero $\beta \in F$, contrary to our hypothesis that F is not a field.

COROLLARY 3.1. *Let R be any ring of characteristic zero. Then R is algebraic over the ring of integers if and only if R is nil.*

Proof. The ring of integers satisfies the conditions on F in Theorem 3.2, so, if R is algebraic over the integers, then, to each $x \in R$, there correspond a non-zero integer $\alpha = \alpha(x)$ and a positive integer $m = m(x)$ such that $\alpha x^m = 0$; and, since R has characteristic zero, this implies that $x^m = 0$, whence R is nil. The converse is obvious.

Theorem 3.2 may be regarded as generalizing the known fact **(7**, Theorem 11**)** that, if every element of a ring R satisfies $x^{n(x)} = x$ for some integer $n(x) \geqslant 2$, then every element of R has finite additive order; indeed, for any element x of a ring R satisfying this more stringent condition, and any admissible operator ring F, our argument shows that either an element α of F exists such that $\alpha x = 0$, $\alpha R \neq 0$, or F has the same property as R (so that, if F is an integral domain, it must be an algebraic field of prime characteristic).

The argument of Theorem 3.2 can easily be modified to show that, with F, R as before, every regular element of R has a non-zero annihilator in F. We can also, without appreciably more trouble, prove the following generalization of Theorem 3.2 **(**cf. **11)**:

THEOREM 3.3. *Let F be any integral domain, R any ring over F, and x any element of R. Suppose also that there exists a non-zero element $\pi = \pi(x)$ of F such that, to each element y of the subring of R generated by x, corresponds a non-zero polynomial $g_y(\lambda)$ over F, whose lowest non-zero coefficient is not divisible by π, such that $g_y(y) = 0$. Then any equation of the form* (3) *with each $\alpha_i \in F$ and $\alpha_m \neq 0$ implies that a non-zero element α of F exists satisfying $\alpha x^m = 0$.*

Proof. As in the proof of Theorem 3.2, we can find an element e of the subring generated by x such that ${\alpha_m}^{m+1} x^m = e x^m$, ${\alpha_m}^{m+1} e = e^2$. By our hypothesis, for each $\beta \in F$, there is a polynomial over F of the form $g_{\beta e}(\lambda) = \gamma_\beta \lambda^{t_\beta} - \lambda^{t_\beta+1} k_\beta(\lambda)$, with γ_β not divisible by π, such that $g_{\beta e}(\beta e) = 0$. As before, we deduce that, for each $\beta \in F$, an element θ_β of F exists such that

$$\beta^{t_\beta}(\gamma_\beta - \beta\theta_\beta)\, \alpha_m^{(m+1)t_\beta} x^m = 0.$$

https://doi.org/10.4153/CJM-1956-039-8 Published online by Cambridge University Press

Finally, taking $\beta = \pi$, since $\pi \neq 0$, $\alpha_m \neq 0$ and since γ_π is not divisible by π, we can be sure that $\alpha \equiv \pi^{t_\pi}(\gamma_\pi - \pi\theta_\pi)\alpha_m^{(m+1)\,t_\pi} \neq 0$.

It is hardly necessary to mention that the existence of an element π of F satisfying the conditions of Theorem 3.3 ensures that F cannot be a field. As a corollary of Theorem 3.2 itself (or more generally of Theorem 3.3) it is obvious that any minimal polynomial of x must have the monomial form $\alpha\lambda^n$. This is not difficult to see even under a hypothesis substantially weaker than that all elements of the *subring* generated by x be algebraic, as we show next:

THEOREM 3.4. *Let F be any integral domain, not a field but having a unit element, let R be any ring over F, and x a given element of R. Then, if every F-multiple γx of x is algebraic over F, and if*

$$h(\lambda) = \alpha_m\lambda^m + \alpha_{m+1}\lambda^{m+1} + \ldots + \alpha_n\lambda^n,$$

with $\alpha_m \neq 0$, $\alpha_n \neq 0$, is a given minimal polynomial for x over F, we must have $m = n$ (so that $\alpha_m x^m = 0$).

Proof. Suppose by way of contradiction that $m \neq n$, that is, $n - m \geqslant 1$, and let β be an arbitrary non-zero element of F (fixed throughout the ensuing argument). Then

$$0 = \beta^n\alpha_n^{n-1}h(x) = \beta^{n-m}\alpha_n^{n-m-1}\alpha_m(\alpha_n\beta x)^m + \beta^{n-m-1}\alpha_n^{n-m-2}\alpha_{m+1}(\alpha_n\beta x)^{m+1} + \ldots + \beta\alpha_{n-1}(\alpha_n\beta x)^{n-1} + (\alpha_n\beta x)^n,$$

and so, defining $y = \alpha_n\beta x$ and

$$f(\lambda) = \beta^{n-m}\alpha_n^{n-m-1}\alpha_m\lambda^m + \beta^{n-m-1}\alpha_n^{n-m-2}\alpha_{m+1}\lambda^{m+1} + \ldots + \beta\alpha_{n-1}\lambda^{n-1} + \lambda^n,$$

we have $f(y) = 0$. Indeed, $f(\lambda)$ is a *minimal* polynomial for y (since, F being an integral domain and $\alpha_n\beta$ being non-zero, if y satisfied an equation of lower degree, so would x).

Now, y being an F-multiple of x, our hypothesis assures us of the existence of a positive integer t and a polynomial $k(\lambda)$ over F such that $y^t = y^{t+1}k(y)$. Thus, by Lemma 2.2 (with $\alpha = 1$), there is a polynomial $q(\lambda)$ over F such that

$$\lambda^t - \lambda^{t+1}k(\lambda) = f(\lambda)\,q(\lambda)$$

identically. Since $\alpha_m \neq 0$, comparison of coefficients of λ^t on either side gives $1 = \beta^{n-m}\alpha_n^{n-m-1}\alpha_m\xi$, where $\xi = \xi_\beta$ is the lowest non-zero coefficient of $q(\lambda)$, that is $1 = \beta\theta_\beta$, where

$$\theta_\beta = \beta^{n-m-1}\alpha_n^{n-m-1}\alpha_m\xi_\beta.$$

But, since β was an arbitrary non-zero element of F, this would contradict our hypothesis that F is not a field; thus in fact $m = n$, as required.

There is naturally an extension of Theorem 3.4 along the lines of Theorem 3.3, but we shall not state it formally. However, we note the (trivial and known) corollary that, if z is a complex number such that z/β is an algebraic

integer (in the usual number-theoretic sense) for every positive integer β, then $z = 0$; to see this, one has only to suppose the contrary, and take $x = 1/z$ in Theorem 3.4.

4. Diagonable rings. On being given any positive integer q and on writing 1_q for the unit $q \times q$ matrix, it is customary to call a $q \times q$ matrix x, with elements in a given field F, *diagonable over* F if distinct elements $\beta_1, \ldots, \beta_s$ of F exist such that

$$(x - \beta_1 1_q) \ldots (x - \beta_s 1_q) = 0$$

(where s can be any positive integer). There are several well-known alternative forms for this definition (e.g. in terms of the existence of a non-singular $q \times q$ matrix b over F such that $b^{-1}xb$ is diagonal). We shall adopt the following (obviously equivalent) form: x is diagonable over F if and only if there are distinct elements $\gamma_1, \ldots, \gamma_t$ of F such that

$$x(\gamma_1 x + 1_q) \ldots (\gamma_t x + 1_q) = 0. \tag{4}$$

It will be noted that we have not required F to be algebraically closed; indeed, our definition remains significant for *any* (*commutative*) *ring* F. Further, since the unit matrix 1_q now occurs only in a purely formal way (i.e. can be got rid of by multiplying out the factors in (4)), we may apply the definition to *any* ring R admitting the operators F (i.e. not merely to rings of square matrices over F). If every element of a ring R over F is diagonable over F, we shall say that R is itself diagonable over F. Obviously every diagonable ring over F is algebraic over F.

Motzkin and Taussky **(10)** showed that, if x, y are given $q \times q$ matrices over an algebraically closed field F, and if also $\alpha x + \beta y$ is diagonable over F for all choices of α, β in F, then $xy = yx$ (whence it is easy to deduce the existence of a non-singular $q \times q$ matrix b over F reducing x and y simultaneously to diagonal forms $b^{-1}xb$, $b^{-1}yb$). Their proof (a geometrical one) is long; and, since hypotheses are made only about the F-module generated by x and y, ring-theoretic methods are perhaps not very suitable for dealing with the problem. However, if we are prepared to extend the diagonability hypothesis to all "non-commutative polynomials" in x and y, then the proof that x, y commute becomes almost trivial; indeed, for rings with arbitrary operators, we shall show in our next theorem that diagonability always implies commutativity. The proof depends on a familiar property of strongly regular rings; for completeness, we first derive this property, and indeed something more general, in the following lemma (which will in any case be needed later on in §6):

LEMMA 4.1. *Let R be any ring in which, to each pair of elements x, y, there corresponds a non-negative integer r such that xy^r is in the right ideal of R generated (over the given operator ring F) by y and x^2. Then, if J denotes the Jacobson radical of R, R/J is a subdirect sum of division rings.*

Proof. We know from Jacobson's structure theory that R/J is a subdirect sum of primitive rings, each of which is a homomorphic image of R/J and hence of R; and each of these primitive rings inherits the (clearly homomorphism-transitive) hypothesis on R. Thus it will be enough to show that if R is itself primitive then R must be a division ring.

To call R primitive is the same as to say that R is isomorphic with a dense ring M of linear transformations of a vector space V over a division ring D. We shall denote the result of operating on $v \in V$ with $x \in M$ by vx (i.e. regard M and D as operating on V from the right), and have only to show that V cannot contain two elements v_1, v_2 independent with respect to D. But, in the contrary case, since M is dense, we could choose x, y in M so that

$$v_1x = v_2,\ v_2x = 0,\ v_1y = 0,\ v_2y = v_2;$$

then, for any α, $\beta \in F$, any a, $b \in R$, and any non-negative integer r,

$$v_1(xy^r - \alpha x^2 - x^2a - \beta y - yb) = v_1x(y^r - \alpha x - xa) - 0 = v_2y^r = v_2 \neq 0$$

(by the D-independence of v_1, v_2). But our hypothesis on R asserts that, x, y being chosen, we can find α, β, a, b, r such that $xy^r - \alpha x^2 - x^2a - \beta y - yb = 0$; thus we have our desired contradiction.

THEOREM 4.1. *Every diagonable ring is commutative.*

Proof. Given any element x of a diagonable ring R, then, on taking $\gamma_1, \ldots, \gamma_t$ as in equation (4) above and on writing

$$(\gamma_1\lambda + 1) \ldots (\gamma_t\lambda + 1) = 1 - \lambda g(\lambda),$$

$g(\lambda)$ is a polynomial over F (possibly with constant term), and $x = x^2h(x)$, where $h(\lambda) = \lambda g^2(\lambda)$ is a polynomial over F without constant term (so that $h(x)$ is well-defined). Thus, given any $x \in R$, we can find an element $a = h(x)$ of R such that $x = x^2a$. In other words, every diagonable ring R is strongly regular and hence semi-simple in Jacobson's sense, and so, by Lemma 4.1 (with $r = 0$), R is a subdirect sum of division rings, each of which is a homomorphic image of R and consequently diagonable. But a diagonable division ring is obviously commutative, so we deduce that R must in fact be a subdirect sum of fields.

5. Additive functions on π-regular rings. We recall (cf. **8**) that an element x of a ring R is said to be *π-regular in R* if a positive integer $s = s(x)$ and an element $b = b(x)$ of R exist satisfying $x^s = x^sbx^s$. Given any elements x, y of a ring, we shall use $[x, y]$ to denote their additive commutator $xy - yx$.

THEOREM 5.1. *Let R be any ring, let $\mathfrak{S}$ be any given set with a transitive binary relation $\prec$ defined on it, and let $\mathscr{M}$ be any set of mappings of R into $\mathfrak{S}$. Then, if we denote by $M(x)$ the result of operating on a typical element x of R by a typical element M of $\mathscr{M}$, the statement* (i) *to each choice of x in R and M in*

https://doi.org/10.4153/CJM-1956-039-8 Published online by Cambridge University Press

$\mathscr{M}$ there correspond $c \in R$, $N \in \mathscr{M}$ and an integer $t \geqslant 2$ such that $[x^t, c] = 0$ and $M(x) \prec N(x^t c)$, implies (ii) *$M(z) \prec M(0)$ for every $M \in \mathscr{M}$ and every nilpotent element z of R.*

Conversely, if (ii) *holds, then* (iii) *for any given π-regular element x of R, say with $x^s = x^s b x^s$, we have $M(x - x^{s+1}b) \prec M(0)$ for all $M \in \mathscr{M}$.*

Proof. Suppose first that (i) holds. Then, given any $x = x_0$ in R and any $M = M_0$ in $\mathscr{M}$, we can find a sequence of integers $t_j \geqslant 2$, a sequence c_j of elements of R and a sequence of mappings $M_j \in \mathscr{M}$ $(j = 1, 2, \ldots)$ such that

$$x_j = x_{j-1}^{t_j} c_j, \quad [x_{j-1}^{t_j}, c_j] = 0, \quad M_{j-1}(x_{j-1}) \prec M_j(x_j) \qquad (j = 1, 2, \ldots).$$

Since $\prec$ is transitive on $\mathscr{M}$, $M(x) = M_0(x_0) \prec M_j(x_j)$ $(j = 1, 2, \ldots)$, while a simple induction argument shows that

$$x_j = x^{t_1 \ldots t_j} c_1^{t_2 \ldots t_j} c_2^{t_3 \ldots t_j} \ldots c_{j-1}^{t_j} c_j \qquad (j = 1, 2, \ldots);$$

combining these two remarks we obtain (ii) on taking j sufficiently large.

To prove that (ii) implies (iii) we notice that $(x - x^{s+1}b)^s$ can be written in the form $x^s + x^s d$ (for a suitably chosen $d \in R$), so that

$$(x - x^{s+1}b)^{s+1} = (x - x^{s+1}b)\, x^s + (x - x^{s+1}b)\, x^s \,.\, d;$$

also, if $x^s = x^s b x^s$, then

$$(x - x^{s+1}b)\, x^s = x^{s+1} - x^{s+1} b x^s = x^{s+1} - x \,.\, x^s = 0,$$

so that $x - x^{s+1}b$ is nilpotent, and (ii) gives $M(x - x^{s+1}b) \prec M(0)$.

It is perhaps worth remarking that, if $x^s = x^s b x^s$ and we write $s = 2r - 1 + \delta$, where r is a positive integer and $\delta = 0$ or 1, then one can show (only slightly less easily than in the second part of the proof above) that $(x - x^{r+\delta} b x^r)^s = 0$, so that (ii) also implies $M(x - x^{r+\delta} b x^r) \prec M(0)$; however, this fact seems to be less useful in applications.

We have set out Theorem 5.1 in the very general (and accordingly rather bogus-looking) form above in order to highlight the essential argument, which will be successively more and more obscured in our next theorems (where we return to earth, and make the "converse" more worthy of the name, by specializing $\mathfrak{S}$, $\mathscr{M}$). We shall mean by an *additive function on R* any mapping, say $f: x \to f(x)$, of R into itself such that $f(x + y) = f(x) + f(y)$ for all $x, y \in R$; in particular, $f(0) = 0$. We do not require that $f(\alpha x) = \alpha f(x)$ for admissible operators α.

THEOREM 5.2. *Let R be any ring, and $\mathscr{L}$ any set of additive functions on R. Then the statement* (i) *to each choice of x in R and f in $\mathscr{L}$ there correspond $c \in R$, $g \in \mathscr{L}$ and an integer $t \geqslant 2$ such that $[x^t, c] = 0$ and such that $f(x)$ is in the two-sided ideal of R generated by $g(x^t c)$, implies* (ii) *$f(z) = 0$ for every $f \in \mathscr{L}$ and every nilpotent element z of R.*

Conversely, if (ii) *holds, then* (iii) *for any given π-regular element x of R, say with $x^s = x^s b x^s$, we have $f(x) = f(x^{s+1}b)$ for every $f \in \mathscr{L}$.*

https://doi.org/10.4153/CJM-1956-039-8 Published online by Cambridge University Press

Proof. For any $x \in R$, $f \in \mathscr{L}$, let $M_f(x)$ denote the two-sided ideal of R generated by the element $f(x)$ of R. Then Theorem 5.2 is just the special case of Theorem 5.1 with $\mathfrak{S}$ chosen as the set of all two-sided ideals of R, ordered in the natural way by inclusion, and with $\mathscr{M}$ chosen as the set of all mappings $M_f: x \to M_f(x)$.

We recall that an element x of a ring R is said to be *strongly regular in* R if an element $a = a(x)$ of R exists such that $x = x^2a$.

THEOREM 5.3. *Every π-regular ring without non-zero nilpotent elements is strongly regular.*

Proof. This follows at once from the second part of Theorem 5.2 on taking $\mathscr{L}$ to consist of the single function $f: x \to f(x) = x$. Alternatively and more directly, going back to the proof of Theorem 5.1, we have merely to observe that $x^s = x^sbx^s$ implies $(x - x^{s+1}b)^{s+1} = 0$, so that, if R has no non-zero nilpotent elements, then $x - x^{s+1}b = 0$, that is

$$x = x^2(x^{s-1}b).$$

Conversely, if R is strongly regular, then (independently of the π-regularity hypothesis) of course R can obtain no non-zero nilpotent element. Thus we see that, *among π-regular rings, the property of having no non-zero nilpotent elements is homomorphism-invariant.*

From now on all we shall need of what has already been proved in this section is the following consequence of Theorem 5.2:

THEOREM 5.4. *For any ring R, the statement* (i) *to each choice of x, y in R there correspond $c \in R$ and an integer $t \geqslant 2$ such that $[x^t, c] = [x - x^tc, y] = 0$, implies* (ii) *every nilpotent element of R is central.*

Conversely, if (ii) *holds, and x is any given π-regular element of R, say with $x^s = x^sbx^s$, then* (i) *holds, for this x and all y, with $c = b$ and $t = 2$ if $s = 1$, and with $c = xb$ and $t = s$ otherwise.*

Proof. The first part is essentially the special case of the corresponding part of Theorem 5.2 with $\mathscr{L}$ chosen as the set of "commutator functions" $f_y: x \to [x, y]$ (one such function being associated with each $y \in R$); indeed, we have thrown away some generality elsewhere by writing $[x - x^tc, y] = 0$ in (i) rather than the weaker "for some z in R, $[x, y]$ lies in the two-sided ideal of R generated by $[x^tc, z]$."

To prove the converse, we quote from **(2)** that (ii) implies that every idempotent element e of R is central. Taking $e = x^sb$, we deduce that $x^s = x^{2s}b$; in a similar way, we see that $x^s = bx^{2s}$. Hence

$$x^sb = bx^{2s} \cdot b = b \cdot x^{2s}b = bx^s;$$

also, by the converse part of Theorem 5.2, $x - x^{s+1}b$ is central, so the proof is complete.

In particular, we have proved that (i) and (ii) are equivalent in any π-regular ring. For the special case of rings algebraic over the integers, the two parts of Theorem 5.4 are implicit in Herstein's papers **(5**, Lemma 3; 6**)** respectively. We note also the following immediate consequence of Theorem 5.4:

COROLLARY 5.1. *Among π-regular rings, the property of having all nilpotent elements central is preserved under homomorphism (even under homomorphisms which do not commute with the given operators).*

This corollary cannot of course be extended to arbitrary rings (consider for example the free ring R generated over the integers by two non-commutative indeterminates, and the natural homomorphism of R onto $R/(4R)$).

Combining Herstein's theorem (quoted in the Introduction) with the converse part of Theorem 5.4, we have (since every algebraic ring is clearly π-regular)

THEOREM 5.5. *Let R be a given ring algebraic over the integers (or any quotient ring), and suppose that every nilpotent element of R is central. Then R is commutative.*

This was previously pointed out by Herstein **(5)**, and generalizes a result of Arens and Kaplansky **(1**, Theorem 4.2**)**. They proved commutativity for any ring R, necessarily without non-zero nilpotent elements, in which each element x has finite non-zero additive order and satisfies an equation of the form

$$\alpha_1 x + \ldots + \alpha_n x^n = 0, \tag{1}$$

with $\alpha_1, \ldots, \alpha_n$ integral and $\alpha_n x^n \neq 0$. For in these circumstances every element has squarefree characteristic, so that R is the restricted direct sum of $R_{(p)}$, where p takes all prime values and $R_{(p)}$ denotes the set of all $x \in R$ with $px = 0$; and it is easy to see that each $R_{(p)}$ is algebraic over the integers (in our sense) and without non-zero nilpotent elements. Restrictions on the additive orders of elements of R are no longer in evidence in the statement of Theorem 5.5; however, Theorem 3.2 shows that this aspect of generalization of the result of Arens and Kaplansky is illusory.

We should naturally like to have something similar to Theorem 5.5 valid for rings with more general operators than the integers. Such a generalization would of course follow for a given operator ring F if Herstein's theorem could be extended to allow elements of F as coefficients in $p_x(\lambda)$. Consideration of the quaternion algebra over the reals sets a limit on such hopes, but, by Theorems 3.1 and 5.5, we do have at least the following generalization of Jacobson's result **(7**, Theorem 9**)** mentioned in the Introduction:

THEOREM 5.6. *Let F be any field of non-zero characteristic algebraic over its prime subfield (in particular, any finite field). Then, if a given algebra R algebraic over F has all its nilpotent elements central, R is commutative.*

It will be noted that the hypotheses on F imply that F is a perfect field; however, the quaternions show that the result does not hold for all perfect fields F.

We note also the following analogous, and more elementary, result:

THEOREM 5.7. *Let F be any algebraically closed field. Then, if a given algebra R algebraic over F has all its nilpotent elements central, R is commutative.*

Proof. Given any element x of R, then, since R is algebraic, x generates a finite-dimensional subalgebra over F, and, since F is algebraically closed, consequently, by the theory of the classical canonical form, we can write

$$x = f + \sum_i \alpha_i e_i,$$

where f is a nilpotent (and hence central) element of R, the e_i are idempotent elements of R, and the α_i are in F. Thus, to prove R commutative, it would be enough to show that all idempotent elements of R commute with one another. But in fact, by **(2)** again, the hypothesis that all nilpotent elements are central implies (in any ring) that every idempotent element is central, so the result follows.

6. H-rings. We now turn to some questions arising from Herstein's theorem. Herstein's method of proof was to settle first the division ring case (which he succeeded in doing by a comparatively short argument), and then to show (by a rather lengthy sequence of lemmas) how the result for arbitrary rings can be reduced to this special case. Herstein has conjectured (in a letter to the writer) that if, to each element x of a given ring R, there corresponds an element a of R such that $x - x^2a$ is central, then R is a subdirect sum of a commutative ring and a (possibly vacuous) set of division rings; we shall refer to this as *Herstein's conjecture*.

This conjecture can reasonably be thought of as generalizing Herstein's theorem, since any division ring D occurring as a subdirect summand of R is necessarily a homomorphic image of R, so that, if a is always a polynomial in the $x \in R$ to which it corresponds, then a similar statement holds for D (while, as we have noted, the division ring case of Herstein's theorem takes up only a small part of the proof). Further, the conjecture, if true, would have over the theorem the advantage that its (much weaker) hypothesis does not involve any restriction on the operators, whereas the quaternion ring shows that the theorem as originally stated definitely does not extend to rings with arbitrary operators (rather than the integers). Thus the conjecture embodies as much as one could hope to be true in the general case and also, essentially, in the case of integer operators first considered by Herstein; if the conjecture could be substantiated, the theorem (and most of its subsequent ramifications) could be deduced from it in a comparatively trivial way.

We shall in fact consider here only the case in which $x^2a = x^tc = cx^t$, where $c \in R$ and t is some integer with $t \geqslant 2$, but we can afford to weaken the centrality condition slightly. Formally, we call a given ring R an *H-ring* if, to each pair x, y in R, there correspond $c = c(x, y) \in R$ and an integer $t = t(x, y) \geqslant 2$ such that

$$[x^t, c] = [x - x^tc, y] = 0.$$

Certain of Herstein's arguments can be straightforwardly generalized to apply to these rings; since every division ring is an H-ring (e.g. with $c = x^{-1}$ for $x \neq 0$ and otherwise arbitrary) we cannot hope to prove all H-rings commutative, and we shall be chiefly concerned with side-conditions sufficient to ensure commutativity (cf. Theorem 4.1 above). Our next result shows that all H-rings have a certain property which would follow as an immediate consequence of the truth of Herstein's conjecture; and, conversely, that, when we restrict attention to π-regular rings, this property actually characterizes the H-rings:

THEOREM 6.1. *Every H-ring has all its nilpotent elements central. Conversely, if a given ring R is π-regular and all its nilpotent elements are central, then R is an H-ring and c, t can be chosen independently of y; also, if R is algebraic (over some given ring F of operators), then, corresponding to each x ∈ R, there is a polynomial $p_x(\lambda)$ over F such that $x - x^2p_x(x)$ is central.*

This is, essentially, just a partial restatement of a special case of Theorem 5.4 in H-ring terminology.

Extending slightly concepts which have been used by Goldhaber and Whaples **(4)** and by McLaughlin and Rosenberg **(9)**, we shall say that a commutative ring F is *quasi-algebraically closed* if every division ring algebraic over F (or over a factor ring of F) is commutative. Obviously every algebraically closed field is quasi-algebraically closed; and, by Theorems 5.5 and 5.6, the property of being quasi-algebraically closed is also shared by the ring of integers (with all its quotient rings), and by every finite field.

If, in Lemma 4.1, F is quasi-algebraically closed and R is algebraic over F, then clearly R/J is a subdirect sum of fields, and so (since J is nil in any π-regular ring) R is commutator-nil, i.e. the two-sided ideal of R generated by all the commutators $[x, y]$ with $x, y \in R$ is a nil ideal;[1] and clearly every H-ring satisfies the hypothesis of Lemma 4.1 (with $r = 1$). If Herstein's conjecture were true, we should even have commutativity for all algebraic H-rings over quasi-algebraically closed operator rings F. Not every algebraic H-ring is commutative (consider again the quaternions), but, by combining Theorem 6.1 with Herstein's theorem, and also with Theorems 3.1 and 5.7, we find

[1]Clearly this conclusion still holds good even if R is given as only π-regular (rather than actually algebraic) provided that *every* division ring over F is commutative; a variety of analogous results can be obtained by weakening the hypothesis on either R or F and correspondingly strengthening the hypothesis on the other.

THEOREM 6.2. *Every H-ring algebraic over the integers, or over any finite or algebraically closed field, is commutative.*

More generally, if Herstein's theorem could be extended to allow given operators F as coefficients in $p_x(\lambda)$, then we could similarly show that every H-ring algebraic over this particular F is commutative. Commutativity (and hence local finiteness) would then of course follow for every division ring algebraic over F; and this is the same as to say that F is quasi-algebraically closed. Thus Herstein's theorem definitely cannot be extended to any non-quasi-algebraically closed F.

Without prejudging how far Herstein's theorem does extend, or whether his conjecture is in fact true, we can at least show that every H-ring algebraic over a quasi-algebraically closed *field* F must be locally finite. For the algebraic condition on R makes J nil and consequently (by Theorem 6.1) central, while we have previously seen (from Lemma 4.1) that R/J must be commutative. Then R/J and J, being commutative algebraic algebras over F, are both locally finite over F, whence, by **(7**, Theorem 15**)**, R is itself locally finite, as we asserted.

Now commutativity for R is equivalent to that of all its doubly generated subrings; also these are finite-dimensional over F by what we have just proved, and are H-rings in view of the last part of Theorem 6.1. Thus, to prove commutativity for all H-rings algebraic over a given quasi-algebraically closed field F, it is enough to do so only for finite-dimensional R; this is easy when F is also perfect (and, more generally, whenever R can be expressed as a supplementary sum $R = S + J$ with $S \cong R/J$).

REFERENCES

1. R. Arens and I. Kaplansky, *Topological representation of algebras*, Trans. Amer. Math. Soc. *63* (1948), 457–481.
2. M. P. Drazin, *Rings with central idempotent or nilpotent elements* (to appear).
3. M. P. Drazin and K. W. Gruenberg, *Commutators in associative rings*, Proc. Cambridge Phil. Soc., *49* (1953), 590–594.
4. J. K. Goldhaber and G. Whaples, *On some matrix theorems of Frobenius and McCoy*, Can. J. Math., *5* (1953), 332–335.
5. I. N. Herstein, *The structure of a certain class of rings*, Amer. J. Math., *75* (1953), 864–871.
6. ———, *A note on rings with central nilpotent elements*, Proc. Amer. Math. Soc., *5* (1954), 620.
7. N. Jacobson, *Structure theory for algebraic algebras of bounded degree*, Ann. Math., *46* (1945), 695–707.
8. N. H. McCoy, *Generalized regular rings*, Bull. Amer. Math. Soc., *45* (1939), 175–178.
9. J. E. McLaughlin and A. Rosenberg, *Zero divisors and commutativity of rings*, Proc. Amer. Math. Soc., *4* (1953), 203–212.
10. T. Motzkin and O. Taussky, *Pairs of matrices with property L.* II, Trans. Amer. Math. Soc., *80* (1955), 387–401.
11. A. Rosenberg and D. Zelinsky, *On Nakayama's extension of the* $x^{n(x)} = x$ *theorem*, Proc. Amer. Math. Soc., *5* (1954), 484–486.

Trinity College, Cambridge