TRIANGULAR REPRESENTATIONS OF LINEAR ALGEBRAS

By M. P. DRAZIN

Received 5 March 1953

1. *Introduction.* It is well known that the elements of any given commutative algebra (and hence of any commutative *set*) of $n \times n$ matrices, over an algebraically closed field K, have a common eigenvector over K; indeed, the elements of such an algebra can be simultaneously reduced to triangular form (by a suitable similarity transformation). McCoy (5) has shown that a triangular reduction is always possible even for matrix algebras satisfying a condition substantially weaker than commutativity. Our aim in this note is to extend these results to more general systems (our arguments being, incidentally, simpler than some used for the matrix case even by writers subsequent to McCoy).

We consider, in fact, an arbitrary associative algebra R over K, whose elements act as linear operators on a given space U (this giving a natural analogue of the matrix situation, where U would be just the set of all n-vectors over K). To obtain any analogue for R of the matrix theorems mentioned above, we have, in the first place, to introduce some form of commutativity restriction; we deal with this question in the next paragraph. Secondly, though we do not require R to be of finite dimension, it does appear necessary to impose some finiteness restrictions on R or U (or both); those we shall adopt here are that R is finitely generated (as an algebra over K), and, further, that *either* R is algebraic (in the usual sense) *or* U is finite-dimensional.

We define the *commutator ideal* R_1 of R as the smallest (two-sided) ideal containing all the commutators $ab - ba$ (where a, b take all values in R). Next, given any property N of ideals, then, if R_1 happens to have the property N, we shall say that R is *commutator*-N. Thus a commutator-zero ring is just any commutative ring. The commutativity condition we shall use in §§ 2–3 is that of *commutator-quasinilpotence*, where we call an ideal A quasinilpotent if, given any infinite sequence $a_1, a_2, \ldots$ of (not necessarily distinct) elements of A, there is always a corresponding positive integer n (allowed to depend unboundedly on the sequence) such that $a_n a_{n-1} \ldots a_2 a_1 = 0$.

Our principal result (Theorem 1) is that the conditions set out in the two previous paragraphs are together sufficient to ensure the existence of a common eigenvector for all the elements of R. The proof is based on that given by Dungey, Gruenberg and the present writer (2) for the case in which R, U are both finite-dimensional. In § 3 we go on to obtain a generalization of one aspect of McCoy's results on triangular representations. Finally, in § 4, we show that, for algebras of finite class (in the sense of Jennings (4)), the conclusion of Theorem 1 may be replaced by something stronger.

2. *The eigenspaces of a commutator-quasinilpotent algebra.* As is customary, we shall mean by a *module of* R (or, briefly, an R-*module*) a linear space U (over the ground field

https://doi.org/10.1017/S0305004100028796 Published online by Cambridge University Press

K of R) which admits the elements of R as pre-multipliers, the product so defined being subject to the rules that, for any choices of α in K, a, b in R, and u, v in U,

$$(a+b)u = au+bu, \qquad a(u+v) = au+av,$$
$$a(bu) = (ab)u, \qquad a(\alpha u) = \alpha(au) = (\alpha a)u.\dagger$$

Given any *non-zero subspace* V of a given R-module U, and an element d of R, then we shall say that V is an *eigenspace for* d if each of its non-zero elements is an eigenvector for d, i.e. if, to each non-zero element v of V, corresponds a (unique) scalar 'eigenvalue' $\alpha = \alpha(d, v)$ such that $dv = \alpha v$.‡ If V is an eigenspace for *every* element of R, we shall call V an *eigenspace for* R; on defining the term submodule in the natural way, of course this means that V is, *a fortiori*, a submodule of U.

THEOREM 1. *Let R be a given finitely generated commutator-quasinilpotent algebra, over an algebraically closed field K. Let U be any R-module, and suppose that either* (i) *R is algebraic, or* (ii) *U is finite-dimensional. Then, given any subspace V of U with $RV \neq 0$, there is an element h of R such that hV is an eigenspace for R.*

Proof. Since $RV \neq 0$, we can certainly find $d \in R$ such that $dV \neq 0$; we shall first show that

(*) *If $R_1 dV \neq 0$, then there is an element c of R such that $cdV \neq 0$, $R_1 cdV = 0$.*

If $R_1 dV \neq 0$, then there is an element c_1 of R_1 such that $c_1 dV \neq 0$. If $R_1 c_1 dV = 0$, we may simply take $c = c_1$; if not, then we can find an element c_2 of R_1 such that $c_2 c_1 dV \neq 0$. Again, if $R_1 c_2 c_1 dV = 0$, then we may take $c = c_2 c_1$, while otherwise there is an element c_3 of R_1 such that $c_3 c_2 c_1 dV \neq 0$. Proceeding in this way, we can construct a *sequence* of elements c_k of R_1; i.e., supposing $c_1, c_2, \ldots, c_{k-1}$ to be already defined, then, *if* $R_1 c_{k-1} \ldots c_2 c_1 dV \neq 0$, we take for c_k any element of R_1 such that $c_k c_{k-1} \ldots c_2 c_1 dV \neq 0$. Since R_1 is quasinilpotent, this process must *terminate* at some stage, and so we can find $c_1, c_2, \ldots, c_j$ such that

$$c_j c_{j-1} \ldots c_2 c_1 dV \neq 0, \quad R_1 c_j c_{j-1} \ldots c_2 c_1 dV = 0.$$

Then (*) holds with $c = c_j c_{j-1} \ldots c_2 c_1$.

For the next part of the proof, we write $V' = dV$ if $R_1 dV = 0$, while, if $R_1 dV \neq 0$, we choose c as in (*), and take $V' = cdV$. In either case, V' is a non-zero subspace of U, and $R_1 V' = 0$.

Let $s_1, s_2, \ldots, s_m$ be any given finite set of generators for R (over K). We shall prove (by induction) that, for each integer r in the range $1 \leqslant r \leqslant m$, there is a corresponding element h_r of R such that $h_r V'$ is a common eigenspace for $s_1, s_2, \ldots, s_r$ (provided in each case that V' does not itself satisfy this condition).

Suppose, for a given $r \leqslant m$, that we have a common eigenspace V'' (of the form $V'' = V'$ or $V'' = h_{r-1} V'$) for $s_1, s_2, \ldots, s_{r-1}$, and that V'' is not already an eigenspace for s_r.

Under either of the hypotheses (i), (ii) of the statement of the theorem (of which (ii) makes R, effectively, finite-dimensional on the left of U), there is non-zero polynomial

† We do not require the existence of a unit element for R, or of an element in R which operates as a (left) identity on the elements of U.

‡ It is easy to verify that in these circumstances $\alpha(d, v)$ is in fact always independent of v, i.e. we may write $\alpha = \alpha(d, V)$, and refer to α as *the eigenvalue of V for d.*

https://doi.org/10.1017/S0305004100028796 Published online by Cambridge University Press

$f(x)$ over K, with no constant term, such that $f(s_r)U = 0$, and, *a fortiori*, $f(s_r)\,V'' = 0$; we can therefore choose a non-zero polynomial $g(x)$ (still with no constant term) *of lowest possible degree* satisfying $g(s_r)\,V'' = 0$.† Since V'' is not an eigenspace for s_r, $g(x)$ must have degree at least 2, and our hypothesis that K is algebraically closed then ensures that we can choose a scalar α, and a polynomial $h(x)$ *with no constant term*, such that $g(x) \equiv (x-\alpha)\,h(x)$. Then $h(s_r)$ exists as an element of R, and, since $h(x)$ has degree strictly less than that of $g(x)$, $h(s_r)\,V''$ is a *non-zero* subspace of U. Also

$$\{s_r h(s_r) - \alpha h(s_r)\}\,V'' = g(s_r)\,V'' = 0,$$

i.e., for each element w of V'', we have $s_r h(s_r)\,w = \alpha h(s_r)\,w$; thus we have converted V'' into an eigenspace $h(s_r)\,V''$ for s_r. But this is also an eigenspace for each of $s_1, s_2, \ldots, s_{r-1}$; for, by the definition of V'', we can find scalars $\alpha_i = \alpha_i(w)$ such that $s_i w = \alpha_i w$ $(i = 1, 2, \ldots, r-1)$, whence, for the same range of i,

$$s_i h(s_r)\,w - \alpha_i h(s_r)\,w = \{s_i h(s_r) - h(s_r)\,s_i\}\,w \in R_1\,V'' \leqslant R_1\,V' = 0.$$

Thus, if V'' is not already a common eigenspace for $s_1, s_2, \ldots, s_{r-1}$ *and* s_r, then we can always find a subspace $h(s_r)\,V''$ of U which is, i.e. our induction hypothesis carries over from $r-1$ to r (with $h_r = h(s_r)\,h_{r-1}, h_{r-1}$ or $h(s_r)$ in the various cases); and the argument above of course incidentally proves the truth of the induction hypothesis for $r = 1$.

Hence the induction hypothesis holds also for $r = m$, i.e. we can construct a common eigenspace, of the type $h_m V'$ or simply V', for $s_1, s_2, \ldots, s_m$; and, since $s_1, s_2, \ldots, s_m$ generate R, hV (where $h = h_m d$, $h_m cd$ or cd) is then an eigenspace for the whole of R, as required.‡

It is easy to produce examples to show that (even when R is generated by a single element, and is therefore certainly commutator-quasinilpotent) the hypothesis of algebraic closure for K, and the alternatives (i) and (ii), are essential in Theorem 1; also the result would no longer be true if we dropped the assumption that R is finitely generated (or, of course, the assumption that R is commutator-quasinilpotent).

3. *Representations of commutator-quasinilpotent algebras.*

THEOREM 2. *If R, K, U are as in Theorem* 1, *and if W is a given proper submodule of U, then we can find a strictly increasing sequence*

$$W = V_0 < V_1 < V_2 < \ldots$$

of submodules V_k such that each quotient V_k/V_{k-1} is one-dimensional (the sequence terminating, if at all, only when $V_k = U$).

Proof. It will obviously be enough to prove that V_1 can be constructed (for we can then obtain V_2 by applying this result to V_1 in place of W, and so on). We consider the quotient space $U^* = U/W$. Since W is a submodule of U, we may regard U^* as an R-module by (unambiguously) defining $au^* = (au)^*$ for each $a \in R$, $u \in U$ (where u^* denotes the coset of W which contains u).

† We have always to insist that the polynomials are without constant terms, since we have not required R to contain a unit element.

‡ We have not, explicitly, used the full force of our hypothesis (i), having applied it only to the *generators* s_i; however, it is easy to see that every commutator-nil algebra which can be generated from a finite set of algebraic elements is in fact an algebraic algebra, so we have not sacrificed any generality by this.

https://doi.org/10.1017/S0305004100028796 Published online by Cambridge University Press

Since W does not exhaust U, therefore U^* is a non-zero space. In particular, U^* certainly contains a one-dimensional subspace, say $V^* = Kv^*$, and so, by Theorem 1, *U^* contains a one-dimensional eigenspace Khv^* (or Kv^*) for R*; in other words, we can find an element v_1 ($= hv$ or v) of U, not lying in W, such that $Rv_1^* \leqslant Kv_1^*$ (or, equivalently, $Rv_1 \leqslant W + Kv_1$). On defining $V_1 = W + Kv_1$, then of course V_1/W is one-dimensional, while also

$$RV_1 = RW + RKv_1 \leqslant W + K(W + Kv_1) = V_1,$$

i.e. V_1 *is* a submodule, as required.

We shall now show how Theorem 2 (for the case $W = 0$) may be used to construct certain (infinite) matrix representations for R.

For any given $a \in R$, $v \in V_k$, we have $av \in RV_k \leqslant V_k$, and so the rule $v \to av$ defines a (clearly linear) transformation $\lambda^{(k)}(a)$ of V_k into itself. Thus the mapping $a \to \lambda^{(k)}(a)$ *is a homomorphism of R into the algebra of all linear transformations of V_k into itself.* Also, by Theorem 2, we can choose a sequence $v_1, v_2, \ldots$ of elements of U such that, for each positive integer k, the first k terms of the sequence form a *basis* of V_k. In the usual way, we may conveniently identify the transformation $\lambda^{(k)}(a)$ of V_k with its matrix relative to such a basis, i.e. with the $k \times k$ matrix whose (i, j) element, $\lambda_{ij}^{(k)}(a)$, is the coefficient of v_i in the (unique) expression of av_j as a linear combination of the basis elements $v_1, v_2, \ldots, v_k$; and, since each V_j is itself a submodule, we have $\lambda_{ij}^{(k)}(a) = 0$ $(1 \leqslant j < i \leqslant k)$, i.e. the matrix $(\lambda_{ij}^{(k)}(a))$ is *triangular*.

Thus, for each positive integer k, we have a homomorphism, or *representation*, $a \to (\lambda_{ij}^{(k)}(a))$ of R into the algebra $T_k(K)$ of all triangular $k \times k$ matrices over K. Clearly this representation will be *faithful* (i.e. will provide an *isomorphism* of R with a subalgebra of $T_k(K)$) if, and only if, no non-zero element of R annihilates the whole of V_k. Thus, if we are to have any chance of obtaining a faithful representation of R by linear transformations of U (or of its submodules), we certainly require that the left ideal of U-annihilators in R be zero; also, if U happens to be of *finite* dimension (n, say), then we have $V_k = U$ for $k = n$, and so here this necessary condition is actually sufficient for the existence of a faithful representation of R (in $T_n(K)$), as is of course well known.

However, the point of interest for us now is that, in the general case, i.e. when U (and, possibly, R) is *not* finite-dimensional, we can legitimately take $k = \infty$ in the argument above, and obtain a homomorphism of R into the algebra $T_\infty(K)$ of all (enumerably) infinite triangular matrices over K.† But, unfortunately, even if U has an enumerable basis, and even if R contains no non-trivial U-annihilator, we cannot always choose the sequence V_k in Theorem 2 so that the corresponding representation $a \to \lambda^{(\infty)}(a)$ is faithful (i.e. so that U is the infinite union of the terms of the sequence); thus, in general, we must resort to transfinite methods to obtain faithful representations.

4. *Algebras of finite class.* We recall that the lower central series of R is the series $R = R_0 \geqslant R_1 \geqslant \ldots$, where R_{k+1} is the smallest ideal of R containing all the commutators $ab - ba$ with $a \in R_k$, $b \in R$. If the lower central series of R terminates in the zero ideal,

† This is allowable only in virtue of the fact that the elements of $T_\infty(K)$ are *column-finite* (so that associative multiplication is always possible, i.e. $T_\infty(K)$ *is* an algebra).

https://doi.org/10.1017/S0305004100028796 Published online by Cambridge University Press

then R is said to be of finite class; as Jennings (4) showed, this property always implies commutator-nilpotence (and hence of course commutator-quasinilpotence).

THEOREM 3. *If, in Theorem* 1, R *is of finite class, then we can choose* h *so that not only is* hV *an eigenspace for* R*, but also, in addition,* $(ah-ha)\,V = 0$ *for every* $a \in R$.

Proof. We have $R_0 V = RV \neq 0$, while $R_k V = 0$ for some k, so we can certainly find a (non-negative) integer j such that

$$R_j V \neq 0, \quad R_{j+1} V = 0.$$

Since $R_j V \neq 0$, we can choose an element g of R_j such that $gV \neq 0$, i.e. such that gV is a non-zero subspace of U. If gV is not already an eigenspace for R, then, by Theorem 1, there is an element f of R such that fgV is an eigenspace for R. Thus, in either case, we can find an element h of R_j such that hV is an eigenspace for R (by taking $h = g$ or fg according as gV is, or is not, an eigenspace for R).

Then, for this h, and any $a \in R$, since $h \in R_j$, clearly

$$(ah-ha)\,V \leqslant R_{j+1} V = 0;$$

thus we have constructed an h with the required properties.

Theorem 3 entails that, if V happens to be already an eigenspace for some element d of R, with eigenvalue α, say, then, for any element v of V, we have

$$d(hv) = (hd)\,v = h(\alpha v) = \alpha(hv).$$

In other words, under the hypothesis of finite class, we can, by choosing h suitably, convert V into an eigenspace hV for the whole of R having the *same* eigenvalue for d as V has. However, a stronger result than this can be easily obtained directly, as follows:

THEOREM 4. *If, in Theorem* 1, V *happens to be already an eigenspace for some given element* d *of* R*, and if* R *is of finite class, then, whatever choice of* h *in* R *we make (subject always to the condition that* hV *is to be an eigenspace for* R*),* hV *will necessarily have the same eigenvalue for* d *as* V *has.*

Proof. Let v be any given element of V with $hv \neq 0$; since V, hV are both eigenspaces for d, there are scalars α, β such that $dv = \alpha v$, $dhv = \beta hv$. We have only to prove that $\alpha = \beta$.

Define a sequence h_k of elements of R, by the rules

$$h_0 = h, \quad h_k = h_{k-1} d - d h_{k-1} \quad (k = 1, 2, \ldots);$$

then we easily find (by induction) that

$$h_k v = (\alpha-\beta)^k hv \quad (k = 0, 1, 2, \ldots).$$

Also, since R is of finite class, there is a (positive) integer j such that $h_j = 0$, whence of course $(\alpha-\beta)^j hv = 0$; since v was chosen so that $hv \neq 0$, we deduce $\alpha = \beta$, as required.

By virtue of the joint results of Gruenberg and the present writer (3), Theorems 3 and 4 above supersede Theorem 2 of (1). They also suggest the conjecture that, in Theorem 3, we must even have $(ah-ha)\,V = 0$ for every $a \in R$ and *every* permissible choice of h (which result would include both Theorem 3 and Theorem 4); however, this conjecture is in fact false.

https://doi.org/10.1017/S0305004100028796 Published online by Cambridge University Press

To see this, consider the algebra R generated (over, say, the complex field) by the pair of 3×3 matrices

$$a = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix};$$

clearly R satisfies all the conditions of Theorem 3 (with $R_2 \leqslant R^3 = 0$). Also the set of all complex 3-vectors is a permissible R-module U, and we may take for V the set Kv of all complex multiples of the vector $v = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$. Then $bv = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, so that $h = b$ is a valid choice for h in Theorem 1; however, this choice gives $(ah - ha)\, v = -bav \neq 0$.

REFERENCES

(1) DRAZIN, M. P. Some generalizations of matrix commutativity. *Proc. Lond. math. Soc.* (3), 1 (1951), 222–31.

(2) DRAZIN, M. P., DUNGEY, J. W. and GRUENBERG, K. W. Some theorems on commutative matrices. *J. Lond. math. Soc.* 26 (1951), 221–8.

(3) DRAZIN, M. P. and GRUENBERG, K. W. Commutators in associative rings. *Proc. Camb. phil. Soc.* 49 (1953), 590.

(4) JENNINGS, S. A. Central chains of ideals in an associative ring. *Duke math. J.* 9 (1942), 341–55.

(5) MCCOY, N. H. On the characteristic roots of matric polynomials. *Bull. Amer. math. Soc.* 42 (1936), 592–600.

TRINITY COLLEGE
CAMBRIDGE

https://doi.org/10.1017/S0305004100028796 Published online by Cambridge University Press