A THEOREM ON MATRICES OVER A COMMUTATIVE RING

NEAL H. McCOY

1. **Introduction.** Let R be an arbitrary commutative ring with unit element 1, and $R[\lambda]$ the ring of polynomials in the indeterminate λ, with coefficients in R. If A is a matrix of order n, with elements in R, the set of all elements $g(\lambda)$ of $R[\lambda]$, such that $g(A)=0$, is an ideal which we shall call the *minimum ideal* of A. The element $f(\lambda)=|\lambda-A|$ of $R[\lambda]$ is the *characteristic function* of A, and the principal ideal $(f(\lambda))$ may be called the *characteristic ideal* of A.* In a recent note,† it was shown that the minimum ideal of a matrix can be characterized in a manner generalizing Frobenius' characterization of the minimum function of a matrix for the case in which the coefficient domain is a field.‡ It was also shown that, in $R[\lambda]$, the prime ideal divisors of the minimum ideal coincide with those of the characteristic ideal. If R is specialized to be an algebraically closed field, this result yields the familiar theorem to the effect that the distinct linear factors of the characteristic function of A coincide with the distinct linear factors of the minimum function of A. It is the primary purpose of the present note to generalize, in a similar way, the well known theorem of Frobenius concerning the characteristic roots of a polynomial in two or more commutative matrices—or, more precisely, an extension of this theorem which we shall now describe in some detail.

Let K denote an algebraically closed field, and let us say that the matrices A_i, $(i=1, 2, \cdots, m)$, with elements in K, have property P, if the characteristic roots of every scalar polynomial $f(A_1, A_2, \cdots, A_m)$, with coefficients in K, are all of the form $f(\lambda_1, \lambda_2, \cdots, \lambda_m)$ where λ_i is a characteristic root of A_i, $(i=1, 2, \cdots, m)$.

In a previous paper,§ the following statements were shown to be equivalent:

I. The matrices A_i, $(i=1, 2, \cdots, m)$, have property P.

* The terms *minimum ideal* and *characteristic ideal* are used merely to emphasize that they generalize the usual notions of minimum and characteristic functions, respectively.

† Neal H. McCoy, *Concerning matrices with elements in a commutative ring*, this Bulletin, vol. 45 (1939), pp. 280–284.

‡ For the classical theorems concerning the characteristic and minimum functions and related topics, see C. C. MacDuffee, *The Theory of Matrices*, chap. 2, or J. H. M. Wedderburn, *Lectures on Matrices*, chap. 2.

§ N. H. McCoy, *On the characteristic roots of matric polynomials*, this Bulletin, vol. 42 (1936), pp. 592–600. Hereafter this will be referred to as M.

II. All matrices $A_iA_j-A_jA_i$, $(i, j=1, 2, \cdots, m)$, are contained in the radical of the algebra of polynomials in the A's with coefficients in K.*

If the matrices A_i are commutative in pairs, clearly condition II is satisfied, and thus I is true. Hence Frobenius' theorem, which states that a set of commutative matrices always has property P, is a special case of the equivalence of I and II. Other interesting special cases of the general result stated above, or examples of matrices having property P, have been obtained by Bruton, Ingraham, Roth, and Williamson.†

If now the matrices A_i, $(i=1, 2, \cdots, m)$, have elements in an arbitrary commutative ring R with unit element, it is obvious that the above definition of property P no longer has any meaning. However, we shall give below a suitable definition of property P which is equivalent to the above, if R is specialized to be an algebraically closed field. The principal result of the present note is then a proof of the equivalence of I and II in this generalized sense.

2. **Preliminary remarks and notation.** We first recall a few properties of ideals which will be of importance in the sequel.‡ If $\mathfrak{a}$ is an ideal in the commutative ring R, the set of all elements of R, of which some finite power belongs to $\mathfrak{a}$, is an ideal called the *radical* of $\mathfrak{a}$. The radical of R is the radical of the null ideal, that is, the set of all nilpotent elements. Clearly $\mathfrak{a}$ and its radical have the same prime ideal divisors. A *minimal prime ideal divisor* of $\mathfrak{a}$ is one containing§ no other prime ideal divisor of $\mathfrak{a}$. *Each ideal* $\mathfrak{a}$ *has minimal prime ideal divisors, and the radical of* $\mathfrak{a}$ *is the intersection of all minimal prime ideal divisors of* $\mathfrak{a}$. Each prime ideal divisor of $\mathfrak{a}$ contains at least one minimal prime ideal divisor, so that, in fact, the radical of $\mathfrak{a}$ is the intersection of *all* prime ideal divisors of $\mathfrak{a}$.

Henceforth R will denote an arbitrary commutative ring with

* In other words, this may be described as follows, using the notation of the present paper. If S is the ring of polynomials in the A's over the ring R, and $\mathfrak{S}$ denotes the two-sided ideal in S generated by all elements $A_iA_j-A_jA_i$, then II states that all elements of $\mathfrak{S}$ are nilpotent.

† G. S. Bruton, *Certain aspects of the theory of equations for a pair of matrices*, this Bulletin, abstract 38-9-196; M. H. Ingraham, *A study of certain related pairs of square matrices*, ibid., abstract 38-9-197; W. E. Roth, *On the characteristic values of the matrix* $f(A, B)$, Transactions of this Society, vol. 39 (1936), pp. 234–243; J. Williamson, *The simultaneous reduction of two matrices to triangle form*, American Journal of Mathematics, vol. 57 (1935), pp. 281–293.

‡ For definitions and fundamental theorems, see W. Krull, *Idealtheorie*, particularly p. 9.

§ In the set-theoretic sense.

unit element, and A_i, $(i=1, 2, \cdots, m)$, fixed matrices of order n with elements in R. Let R_n denote the ring of all matrices of order n over R, and S the subring of R_n generated by the elements A_i, $(i=1, 2, \cdots, m)$, together with the unit element of R_n.* We shall denote by $\mathfrak{s}$ the two-sided ideal in S generated by the elements $A_{ij}=A_iA_j-A_jA_i$, $(i, j=1, 2, \cdots, m)$. Each element of $\mathfrak{s}$ is, therefore, expressible as a finite sum of terms of the form $gA_{ij}h$, where g and h are elements of S.

If y is any element of S, let $\bar{y}$ be the corresponding element of the ring $S/\mathfrak{s}$ under the homomorphic correspondence $S\rightarrow S/\mathfrak{s}$. We now introduce the ring $R'=R[x_1, x_2, \cdots, x_m]$ of polynomials in the commutative indeterminates $x_1, x_2, \cdots, x_m$ over R. To each polynomial $g(A)=g(A_1, A_2, \cdots, A_m)$ in the A's we may, therefore, make correspond the element $g(x)=g(x_1, x_2, \cdots, x_m)$ of R', obtained by formally replacing A_i by x_i, $(i=1, 2, \cdots, m)$.† For example, if $g(A)=A_1A_2A_1$, then $g(x)=x_1^2x_2$. Since $A_iA_j\equiv A_jA_i$ $(\mathfrak{s})$, it follows that multiplication is commutative in $S/\mathfrak{s}$. Thus the correspondence

$$f(x) = \sum a_{i_1i_2\cdots i_m}x_1^{i_1}x_2^{i_2}\cdots x_m^{i_m} \rightarrow \sum \bar{a}_{i_1i_2\cdots i_m}\bar{A}_1^{i_1}\bar{A}_2^{i_2}\cdots \bar{A}_m^{i_m},$$

which we may abbreviate in the form $f(x)\rightarrow\bar{f}(\bar{A})$, is a homomorphism between R' and $S/\mathfrak{s}$, and thus

$$S/\mathfrak{s}\cong R'/\mathfrak{m},$$

where $\mathfrak{m}$ is the ideal in R' of all elements $f(x)$ such that $\bar{f}(\bar{A})=0$. Clearly, $\mathfrak{m}$ contains the minimum ideal of each of the individual matrices A_i, $(i=1, 2, \cdots, m)$.

Let $\mathfrak{p}_\alpha$‡ denote an arbitrary minimal prime ideal divisor of $\mathfrak{m}$, and $\mathfrak{r}$ the radical of $\mathfrak{m}$, so that $\mathfrak{r}$ is the intersection of all $\mathfrak{p}_\alpha$. With each $\mathfrak{p}_\alpha$ we may associate, by means of a given polynomial $f(A)$, a prime ideal $\mathfrak{p}_\alpha'$ in $R[\lambda]$, whose elements are the polynomials $t(\lambda)$, such that $t[f(x)]\equiv 0$ $(\mathfrak{p}_\alpha)$. For convenience, we may indicate the intersection of all $\mathfrak{p}_\alpha'$ by $\mathfrak{k}$. We remark that $h[f(x)]\equiv 0$ $(\mathfrak{r})$, if and only if $h(\lambda)\equiv 0$ $(\mathfrak{k})$.

* Elements of S are, therefore, expressible as polynomials in the A_i, that is, as finite sums of terms of the form $aA_{i_1}A_{i_2}\cdots A_{i_k}$, where a is in R and each A_{i_j} is some one of the matrices $A_1, A_2, \cdots, A_m$. It may happen that no A_i appears in a term, in which case the term will be simply of the form a, as we shall not distinguish between the unit element of R and the unit element of R_n.

† This use of the symbols "A" and "x" will cause no confusion, as they do not appear without subscripts in any other connection. However, λ will always denote a single indeterminate.

‡ The use of a subscript is not meant to imply that the number of minimal prime ideal divisors is finite or even enumerable. The range of α may be an arbitrary set.

Let $\mathfrak{n}$ be the minimum ideal of the matrix $f(A)$, that is, the set of all polynomials $g(\lambda)$ such that $g[f(A)]=0$. We now make the following definition:

DEFINITION. *The matrices A_i, $(i=1, 2, \cdots, m)$, with elements in R are said to have property P, if for every polynomial $f(A)$, the radical of $\mathfrak{n}$ is $\mathfrak{k}$.*

Before proceeding, we pause to point out briefly the meaning of property P, if R is specialized to be an algebraically closed field. Let $f(A)$ be a given polynomial in the A_i. Since $\mathfrak{m}$ contains the characteristic function $g_j(x_j)$ of A_j, $(j=1, 2, \cdots, m)$, it follows* that each prime ideal divisor of $\mathfrak{m}$ is necessarily of the form

$$\mathfrak{p}_\alpha = (x_1 - \lambda_1^{(\alpha)}, x_2 - \lambda_2^{(\alpha)}, \cdots, x_m - \lambda_m^{(\alpha)}),$$

where $\lambda_j^{(\alpha)}$ is a characteristic root of A_j, $(j=1, 2, \cdots, m)$, and further that each $\mathfrak{p}_\alpha$ is minimal. By a Taylor's series expansion we see at once that

$$\mathfrak{p}_\alpha' = (\lambda - f(\lambda_1^{(\alpha)}, \lambda_2^{(\alpha)}, \cdots, \lambda_m^{(\alpha)}),$$

and that the prime ideal divisors of $\mathfrak{k}$ are precisely these $\mathfrak{p}_\alpha'$. Now the prime ideal divisors of $\mathfrak{n}$ are of the form $\lambda - a_j$, where a_j varies over the distinct characteristic roots of $f(A)$. Thus, if the given matrices have property P according to the definition above, the radical of $\mathfrak{n}$ is $\mathfrak{k}$ and, therefore, the prime ideal divisors of $\mathfrak{n}$ coincide with the prime ideal divisors of $\mathfrak{k}$. This means that the characteristic roots of $f(A_1, A_2, \cdots, A_m)$ are all of the form $f(\lambda_1^{(\alpha)}, \lambda_2^{(\alpha)}, \cdots, \lambda_m^{(\alpha)})$, and thus that the matrices have property P as defined in the introduction. Conversely, it is not difficult to show,† although we shall omit the proof, that if the matrices have property P as defined in the introduction, they also have property P as defined here. Thus, if R is an algebraically closed field, the two definitions are equivalent.

3. **The main theorem.** We now prove the following theorem which is the principal result of this note:

THEOREM. *Let A_i, $(i=1, 2, \cdots, m)$, be matrices of order n with elements in an arbitrary commutative ring R with unit element, and denote by S the ring of polynomials in the A_i with coefficients in R. If $\mathfrak{s}$ is the two-sided ideal in S generated by the matrices $A_iA_j - A_jA_i$,*

* For theorems on polynomial ideals see van der Waerden, *Moderne Algebra*, vol. 2.

† Cf. M, pp. 598–599.

$(i, j=1, 2, \cdots, m)$, then a necessary and sufficient condition that the given matrices have property P is that all elements of $\mathfrak{s}$ be nilpotent.

First, we assume that all elements of $\mathfrak{s}$ are nilpotent. If $f(A)$ is an arbitrary polynomial in the A_i, we shall show that $\mathfrak{k}$ is the radical of $\mathfrak{n}$. Let $g(\lambda)$ be any element of $\mathfrak{k}$. Then it follows that $g[f(x)] \equiv 0$ $(\mathfrak{r})$, where $\mathfrak{r}$ is the radical of $\mathfrak{m}$. Thus, for some positive integer k, $\{g[f(x)]\}^k \equiv 0$ $(\mathfrak{m})$. This means, however, that $\{g[f(A)]\}^k \equiv 0$ $(\mathfrak{s})$, and, since all elements of $\mathfrak{s}$ are nilpotent, there exists a positive integer l such that $\{g[f(A)]\}^{kl} = 0$. This implies that $[g(\lambda)]^{kl} \equiv 0$ $(\mathfrak{n})$, that is, that $g(\lambda)$ is in the radical of $\mathfrak{n}$. Thus $\mathfrak{k}$ is contained in the radical of $\mathfrak{n}$.

Now let $h(\lambda)$ be an arbitrary element of the radical of $\mathfrak{n}$, that is, $[h(\lambda)]^\alpha \equiv 0$ $(\mathfrak{n})$, for some positive integer α. Then $\{h[f(A)]\}^\alpha = 0$, from which it follows that $\{\bar{h}[\bar{f}(\bar{A})]\}^\alpha = 0$, and thus that

$$\{h[f(x)]\}^\alpha \equiv 0 \ (\mathfrak{m}).$$

This means that $h[f(x)] \equiv 0$ $(\mathfrak{r})$, and this, in turn, implies that $h(\lambda) \equiv 0$ $(\mathfrak{k})$. We have therefore shown that if all elements of $\mathfrak{s}$ are nilpotent, the radical of $\mathfrak{n}$ is $\mathfrak{k}$.

Conversely, let us now assume that for every polynomial $f(A)$, the radical of $\mathfrak{n}$ is $\mathfrak{k}$. Select any element of $\mathfrak{s}$ and write it, in any way, in the form of a sum of terms

$$F(A)(A_iA_j - A_jA_i)G(A),$$

and denote by $f(A)$ the polynomial thus obtained. Then, clearly, $f(x) = 0$, and thus $f(x) \equiv 0$ $(\mathfrak{r})$. Hence $\lambda \equiv 0$ $(\mathfrak{k})$ and therefore, by hypothesis, there exists a positive integer β such that $\lambda^\beta \equiv 0$ $(\mathfrak{n})$. It follows that $[f(A)]^\beta = 0$, and thus that $f(A)$ is nilpotent. The theorem is therefore established.

SMITH COLLEGE