CONCERNING MATRICES WITH ELEMENTS IN A COMMUTATIVE RING*

NEAL H. McCOY

1. **Introduction.** In a study of the algebraic properties of matrices with elements in a field, some of the most fundamental theorems are those having to do with the concepts of characteristic function and minimum function.† It is the purpose of the present note to suitably generalize some of the leading theorems in this connection to the case of matrices with elements in an arbitrary commutative ring R with unit element 1. For the most part, the theorems as well as the proofs are obtained by suitable generalizations of familiar theorems and proofs in the case in which the coefficient domain is restricted to be a field. However, the degree of generality here obtained seems to be of sufficient interest to warrant a brief account.

Let A denote a square matrix with elements in R. As will be indicated in §2, it is easy to define, in the usual way, the characteristic function $f(\lambda)$, and to show that $f(A)=0$. We shall call the principal ideal $(f(\lambda))$, in the ring $R[\lambda]$, the *characteristic ideal* of A.‡ The set of all elements $g(\lambda)$ of $R[\lambda]$ such that $g(A)=0$ is clearly an ideal in $R[\lambda]$ which may be called the *minimum ideal* of A. In general, this ideal will not be principal. The terms *characteristic ideal* and *minimum ideal* are used merely to emphasize the analogy with the characteristic and minimum functions of A in case the coefficient domain is a field. The actual determination of the minimum ideal is a fundamental problem, a solution of which is obtained in §3. The result, as stated in Theorem 3, is seen to be a generalization of the well known theorem of Frobenius concerning the minimum function. This theorem is the leading result of the present note.

One direction in which we propose to generalize certain familiar results will be sufficiently indicated by the remark that in place of irreducible factors of the characteristic (minimum) function of A we use the prime ideal divisors of the characteristic (minimum) ideal of A. For example, it is easy to show that the prime ideal divisors of the

* Presented to the Society, February 26, 1938.

† For known results concerning the characteristic and minimum functions, see J. H. M. Wedderburn, *Lectures on Matrices*, American Mathematical Society Colloquium Publications, vol. 17, 1934, chap. 2, or C. C. MacDuffee, *The Theory of Matrices*, chap 2. The former will be referred to hereafter as W, the latter as M.

‡ For fundamental definitions and properties of ideals, see van der Waerden, *Moderne Algebra*, or Krull, *Idealtheorie*.

minimum ideal coincide with those of the characteristic ideal. Some further miscellaneous results are given in §4.

2. Some preliminary remarks.

Unless otherwise specifically stated, R will always denote an arbitrary commutative ring with unit element 1, and R_n will be used to denote the ring of all matrices of order n, with elements in R. Now R_n contains a subring isomorphic to R, which we shall identify with R, so that we shall not distinguish between the unit element of R_n and that of R. Throughout the paper, λ will be used as an indeterminate, and $R[\lambda]$ is the ring of polynomials in λ with coefficients in R. Similarly $R_n[\lambda]$ is the ring of polynomials in λ with coefficients in R_n or, from another point of view, the ring of matrices of order n with elements in $R[\lambda]$.

We now make an observation concerning algebraic identities.* Let C be the ring of rational integers, and consider the ring $C'=C[x_1, x_2, \cdots, x_m]$, where $x_1, x_2, \cdots, x_m$ are indeterminates. If now $f(x_1, x_2, \cdots, x_m)$ and $g(x_1, x_2, \cdots, x_m)$ are elements of C', the statement that in C'

$$f(x_1, x_2, \cdots, x_m) = g(x_1, x_2, \cdots, x_m),$$

implies that this is an algebraic identity, so that if $a_1, a_2, \cdots, a_m$ are any elements of R, then clearly

$$f(a_1, a_2, \cdots, a_m) = g(a_1, a_2, \cdots, a_m),$$

equality now being equality in R. It is understood that an integer k occurring as a coefficient in f or g is also to be replaced by $k\cdot 1$, where 1 is the unit element of R. As an application of these remarks, let $X=\|x_{ij}\|$ be a square matrix all of whose elements are indeterminates. Then in the ring $C[x_{11}, \cdots, x_{nn}]$, we have the familiar result that

$$(\text{adj } X)X = X(\text{adj } X) = |X|.$$

Now this equation is in reality a set of equations, each being of the type discussed above. Thus it follows that if D is a matrix with elements in any commutative ring with unit element, then

$$(\text{adj } D)D = D(\text{adj } D) = |D|. \tag{1}$$

It is obvious that many other theorems of this nature may be easily extended in like manner. We shall make use of some similar results without explicitly stating that they hold in the ring R.

It is now easy to prove the following theorem:

* Cf. K. Rychlik, *Eine Bemerkung zur Determinantentheorie*, Journal für die reine und angewandte Mathematik, vol. 167 (1932), p. 197.

THEOREM 1. *An element A of R_n has an inverse if and only if $|A|$ has an inverse in R.*

Suppose $|A|$ has an inverse b in R. Then from (1) it follows that $b(\text{adj } A)$ is the inverse of A in R_n. Conversely, if A has an inverse B, then from $AB = BA = 1$, it follows by taking determinants that $|B|$ is an inverse of $|A|$ in R.

If A is an element of R_n, we call the polynomial

$$f(\lambda) = |\lambda - A| = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n,$$

the characteristic function of A, and the principal ideal $(f(\lambda))$ in $R[\lambda]$ the *characteristic ideal* of A. It will be noted that $f(\lambda)$ is an element of $R[\lambda]$, the leading coefficient being 1 and the constant term being $\pm|A|$. Since the leading coefficient is 1, it is easy to see that, in $R[\lambda]$, $f(\lambda)$ is not a divisor of zero—a fact which will be used in §3.

Since, by (1), with $D = \lambda - A$, we have

$$(\lambda - A) \text{ adj } (\lambda - A) = f(\lambda),$$

the factor theorem* shows that $f(A) = 0$. We have thus proved the next theorem:

THEOREM 2. *If $f(\lambda)$ is the characteristic function of A, then $f(A) = 0$.*

It is now easy to prove the following corollary:

COROLLARY.† *If A has an inverse in R_n, it is expressible as a polynomial in A with coefficients in R.*

3. **Determination of the minimum ideal.** As above, we let A denote a fixed element of R_n, the characteristic function of A being $f(\lambda)$. The *minimum ideal* of A is the ideal $\mathfrak{m}$ in $R[\lambda]$ of all polynomials $g(\lambda)$ such that $g(A) = 0$.

Let the first minors of the matrix $\lambda - A$ be denoted by $h_{ij}(\lambda)$, $(i, j = 1, 2, \cdots, n)$. We may now establish the following fundamental result:

THEOREM 3. *An element $g(\lambda)$ of $R[\lambda]$ is an element of $\mathfrak{m}$ if and only if*

$$(2) \qquad g(\lambda)h_{ij}(\lambda) \equiv 0 \ (f(\lambda)), \qquad i, j = 1, 2, \cdots, n.$$

Before proving the theorem, we may point out what it means in the case in which R is specialized to be a field. In that case, each ideal in $R[\lambda]$ is principal and thus $\mathfrak{m} = (m(\lambda))$, where $m(\lambda)$ is the minimum

* See A. A. Albert, *Modern Higher Algebra*, p. 26.

† The proof follows readily from Theorems 1 and 2 by the method of M, p. 21.

function of A. If $h(\lambda)$ is the g.c.d. of the $h_{ij}(\lambda)$, the theorem merely states that $m(\lambda)=f(\lambda)/h(\lambda)$, which is the theorem of Frobenius. We may remark also that in the language of ideal quotients,* Theorem 3 states that $\mathfrak{m}=(f(\lambda)):\mathfrak{h}$, where $\mathfrak{h}$ is the ideal with basis elements $h_{ij}(\lambda)$, $(i,j=1,2,\cdots,n)$.

Our proof is a modification of a method of Perron.† Suppose $g(\lambda)$ satisfies all the conditions (2). Now the elements of adj $(\lambda-A)$ are precisely the $h_{ij}(\lambda)$ except possibly for sign. Thus we have

$$(3)\qquad g(\lambda)\text{ adj }(\lambda-A)=\|f(\lambda)k_{ij}(\lambda)\|=f(\lambda)K(\lambda),$$

where $K(\lambda)$ is a matrix with elements in $R[\lambda]$, in other words an element of $R_n[\lambda]$. If we multiply (3) by $\lambda-A$ on the left, and apply relation (1), we get

$$g(\lambda)f(\lambda)=f(\lambda)(\lambda-A)K(\lambda).$$

Now it was pointed out above that $f(\lambda)$ is not a divisor of zero, and we thus have $g(\lambda)=(\lambda-A)K(\lambda)$. The factor theorem then shows at once that $g(A)=0$, and thus that $g(\lambda)\equiv 0\ (\mathfrak{m})$.

Let us now assume that $g(\lambda)\equiv 0\ (\mathfrak{m})$. Then clearly

$$g(\lambda)=g(\lambda)-g(A)=(\lambda-A)G(\lambda),$$

where $G(\lambda)$ is an element of $R_n[\lambda]$. Multiplication by adj $(\lambda-A)$ yields the result

$$g(\lambda)\text{ adj }(\lambda-A)=f(\lambda)G(\lambda),$$

and hence

$$g(\lambda)h_{ij}(\lambda)=\pm f(\lambda)g_{ij}(\lambda)\equiv 0\ (f(\lambda)),\qquad i,j=1,2,\cdots,n.$$

Thus the conditions (2) are satisfied, and the proof of the theorem is completed.

We may now prove the following corollary:

COROLLARY. *The prime ideal divisors of the minimum ideal of A coincide with those of the characteristic ideal of A.*

Since $f(\lambda)\equiv 0\ (\mathfrak{m})$, it is only necessary to prove that if $\mathfrak{p}$ is a prime ideal divisor of $(f(\lambda))$, it also divides $\mathfrak{m}$. Let $g(\lambda)$ be any element of $\mathfrak{m}$; we show that $g(\lambda)\equiv 0\ (\mathfrak{p})$. Taking determinants of both sides of relation (3), we get

* See W. Krull, *Ein neuer Beweis für die Hauptsätze der allgemeinen Idealtheorie*, Mathematische Annalen, vol. 90 (1923), pp. 55–64.

† See M, p. 20.

$$[g(\lambda)]^n[f(\lambda)]^{n-1} = [f(\lambda)]^n | K(\lambda) |,$$

from which it follows that

$$[g(\lambda)]^n = f(\lambda) | K(\lambda) | \equiv 0 \ (\mathfrak{p}).$$

But since $\mathfrak{p}$ is a prime ideal, this implies that $g(\lambda) \equiv 0 \ (\mathfrak{p})$.

4. **Further miscellaneous results.** We conclude with a few results which follow readily from the definition of the minimum ideal $\mathfrak{m}$.

THEOREM 4. *If $h(\lambda)$ is an element of $R[\lambda]$, then $h(A)$ has an inverse if and only if $(h(\lambda), \mathfrak{m}) = (1)$.*

The sufficiency of the condition is almost obvious. To prove the necessity of the condition, let us suppose that $h(A)$ has an inverse. Then, by the corollary to Theorem 2, it follows that there exists a polynomial $t(\lambda)$ in $R[\lambda]$ such that $t[h(A)]$ is the inverse of $h(A)$. That is, $h(\lambda)t[h(\lambda)] - 1 \equiv 0 \ (\mathfrak{m})$, and thus $(h(\lambda), \mathfrak{m}) = (1)$ as required.

Let $f'(\lambda)$ denote the formal derivative of the characteristic function $f(\lambda)$ of A. We now prove the following theorem:

THEOREM 5. *If $(f'(\lambda), \mathfrak{m}) = (1)$, the only matrices of R_n commutative with A are polynomials in A.*

Suppose B is an element of R_n commutative with A. An examination of the Sylvester identities* shows that they hold for the case in which the coefficient domain is our ring R. In particular, we have

$$f'(A)B = g(A),$$

where $g(A)$ is a polynomial in A with coefficients in R. Now since, by hypothesis, $(f'(\lambda), \mathfrak{m}) = (1)$, the preceding theorem shows that $f'(A)$ has an inverse, which is of necessity a polynomial in A. Thus B is of the required form.

THEOREM 6. *If $h(\lambda)$ is an element of $R[\lambda]$, then $h(A)$ is nilpotent if and only if $h(\lambda)$ is divisible by all prime ideal divisors of $\mathfrak{m}$.*

For clearly $h(A)$ is nilpotent if and only if some power of $h(\lambda)$ is in $\mathfrak{m}$, and a result of Krull† shows that this is the case if and only if $h(\lambda)$ is an element of each prime ideal divisor of $\mathfrak{m}$.

INSTITUTE FOR ADVANCED STUDY AND
SMITH COLLEGE

* W, p. 27.

† *Idealtheorie*, p. 9.