ON CERTAIN EQUATIONS IN MATRICES WHOSE ELEMENTS BELONG TO A DIVISION ALGEBRA*

M. H. INGRAHAM

1. **Introduction.** A method was given by the author† to determine all matrices X having elements in a field F satisfying the matric equation $P(X)=A$, where $P(\lambda)$ is a polynomial with coefficients in F and A is a matrix with elements in F. The result gives X to within a similarity transformation commutative with A. The purely formal generalization of allowing F to be a division algebra, possibly non-commutative, not only leads to difficulties that probably can not be handled by extensions of the methods of the above mentioned paper, but seems to be devoid of interest. However, if we consider that A defines a linear transformation, we see that the answer to the following question may be of interest: Given the constants $a_n, a_{n-1}, \cdots, a_0$, for what matrices X is $\sum_0^n X^i \xi a_i = A\xi$ for every vector ξ? If the numbers involved lie in a field, this reduces to the previously discussed problem.

After defining the necessary notation, this paper proceeds to give the solution of a slightly more general problem.

Consider a division algebra D and an $n \times n$ matrix A with elements in D. Let $g(\lambda)=\sum \lambda^i a_i$ be a polynomial in λ with coefficients a_i in D. If ξ is an $n \times 1$ matrix (vector), with elements in D, then $g(A) \odot \xi$ is defined‡ to be $\sum A^i \xi a_i$.

If g_1 and g_2 are the two polynomials $\sum \lambda^i a_{1i}$ and $\sum \lambda^i a_{2i}$, respectively, then $g_1 \odot g_2 = \sum \lambda^i g_2(\lambda) a_{1i}$.

The transformation defined by $\xi_1 = g(A) \odot \xi$ will be right linear if and only if the coefficients of g are in the centrum C of D where C denotes the totality of elements of D commutative with every element of D.

Consider two polynomials P and Q with coefficients in D. Let A be an $n \times n$ matrix with elements in D. It is the purpose of this paper to give methods for finding all solutions X of the equation

* Presented to the Society, September 8, 1937. In the preparation of this paper the author was aided by M. C. Wolf, who acted as his research assistant under appointment authorized by the Research Committee of the University of Wisconsin.

† M. H. Ingraham, *On the rational solutions of the matrix equation* $P(X)=A$, Journal of Mathematics and Physics, vol. 13 (1934), pp. 46–50.

‡ See M. H. Ingraham and M. C. Wolf, *Relative linear sets and similarity of matrices whose elements belong to a division algebra*, Transactions of this Society, vol. 42 (1937), pp. 16–31; referred to herein as *Relative linear sets.*

$$P(X) \odot \xi = Q(A) \odot \xi, \qquad (\xi),$$

where X is an $n \times n$ matrix with elements in D. The (ξ) is used throughout in displayed equations for the phrase "for every ξ." This problem reduces to the problem of factorization of polynomials and another problem described in §4.

2. **Reduction of problem to consideration of linear transformations only.** We prove the following theorem:

THEOREM 1. *If a_i, $(i=1, \cdots, k)$, elements of a division algebra D, are linearly independent as to coefficients in the centrum of D and if M_i, $(i=1, \cdots, k)$, are matrices with elements in D satisfying the relation*

$$\sum M_i \xi a_i = 0, \qquad (\xi),$$

then $M_i = 0$ for every i.

This theorem follows at once if it can be established when the ξ is limited to be of the form $\delta_j d$ where δ_j is the Kronecker delta vector and d is in D. That this is true may be established by means of the following lemma:

LEMMA A. *If a_i, $(i=1, \cdots, k)$, elements of a division algebra D, are linearly independent as to coefficients in the centrum of D and if $b_1, b_2, \cdots, b_k$ form a set of elements of D satisfying the relation $\sum b_i d a_i = 0$ for every d in D, then $b_i = 0$ for every i.*

The proof is by induction on k.

There is no loss of generality in letting $a_1 = 1$. Let e be an element of D not commutative with a_2. Both of the following relations are satisfied for every d:

$$\sum b_i d e a_i = 0, \qquad \sum b_i d a_i e = 0,$$

and hence

$$\sum_{i \neq 1} b_i d(e a_i - a_i e) = 0.$$

If the left-hand member of the above equation is expressed as a linear form in a set of the $(ea_i - a_i e)$ which are linearly independent as to coefficients in the centrum, the induction hypothesis shows that the coefficients of this set will all be zero. Hence some linear combination $\sum b_i c_i$ will be zero, where the c_i are in the centrum and $c_1 = 0$. Moreover, c_2 may be taken to be 1. Since the c_i are in the centrum, it follows that $\sum b_i d c_i a_2 = 0$ for every d, and by subtracting this from $\sum b_i d a_i = 0$ it follows that

$$b_1 d a_1 + \sum_{i>2} b_i d(a_i - c_i a_2) = 0.$$

From the induction hypothesis $b_1 = 0$, and hence by induction $b_i = 0$, $(i = 2, \cdots, k)$.

This lemma is also a consequence of a theorem proved by the author* that a set of finite order which is both right and left linear relative to coefficients in D has a commutative base.

Another form of the above lemma is as follows:

If ξ is a vector whose elements belong to a division algebra D and are linearly independent as to coefficients in the centrum of D, there exists a vector η with elements in D such that the transpose $(\xi\eta')'$ of the dyad $\xi\eta'$ is non-singular.

Consider

$$P(X) \odot \xi = Q(A) \odot \xi, \qquad (\xi),$$

and let $1, \alpha_2, \cdots, \alpha_k$ be a proper base for the division algebra D over its centrum C. This equation may be written in the form

$$\sum P_i(X)\xi\alpha_i = \sum A_i\xi\alpha_i, \qquad (\xi),$$

where the P_i have coefficients in the centrum, and hence, from Theorem 1, X must satisfy the equations

$$P_i(X) = A_i, \qquad (i).$$

3. **The solution of $P(X) = A$ where P has coefficients in the centrum.** Let P be a polynomial with coefficients in the centrum C of the division algebra D. If Y is a matrix such that $P(Y)$ is similar to A, then the transformation that takes $P(Y)$ into A takes Y into a matrix X such that $P(X) = A$. All solutions of $P(X) = A$ similar to X are transforms of X by non-singular matrices commutative with A. The problem of the similarity of two matrices whose elements belong to a division algebra has been discussed by Jacobson† and also by the author‡ in collaboration with M. C. Wolf. The following statements are based upon the results of these papers.

If g is a polynomial in D, then there exists a polynomial h of least degree with coefficients in the centrum C and leading coefficient unity for which g is an interior (left-hand) factor. If g is irreducible in D,

* M. H. Ingraham, *General theory of linear sets*, Transactions of this Society, vol. 27 (1925), pp. 163–196.

† N. Jacobson, *Pseudo-linear transformations*, Annals of Mathematics, (2), vol. 38 (1937), p. 485.

‡ *Relative linear sets.*

then h is irreducible in C. In this case any other polynomial irreducible in D which is an interior factor of h is of the same degree as g. The degree of g, $\langle g\rangle$, is also symbolized by $\langle\langle h\rangle\rangle$, that is, the degree of the minimum polynomial defining h.

The nullity of a matrix M is the order of the vector space orthogonal to M, that is, the nullity is equal to the number of right linearly independent vectors ξ satisfying $M\xi=0$. It may be remarked that, since the rank plus the nullity of a matrix is equal to its order, a knowledge of the nullity of M is equivalent to a knowledge of the rank of M.

For any matrix M there exists a minimum polynomial h with coefficients in the centrum C for which $h(M)=0$. Consider $h=\prod h_i^{k_i}$ where the h_i are polynomials over C irreducible in C. If also $h(N)=0$ and if the nullity of $h_i^t(N)$ is the same as $h_i^t(M)$ for every i and $t\leqq k_i$, then M is similar to N. This condition is also necessary. Moreover, the second difference of the nullity of $h_i^t(M)$ as a function of the exponent t is always zero or negative. Since the nullity of $h_i^0(M)$ is zero and since for sufficiently large t the first and second difference of the nullity remains zero, it is clear that a knowledge of the second difference of the nullity of $h_i^t(M)$ for all i and $t\leqq k_i$ is sufficient to determine the class of similar matrices to which M belongs.

Suppose f is a polynomial over C irreducible in C, and suppose that $f^{k_1}, f^{k_2}, \cdots, f^{k_r}$ are the powers of f in the characteristic divisors of M, that is, the highest powers of f in the various invariant factors of M when the latter are factored into polynomials irreducible in C. Let $n(l)$ be the number of k_i equal to l. For every pair of positive integers s and t the nullity of $f^{st}(M)$ is equal to

$$\langle\langle f\rangle\rangle\left[\sum_{l=1}^{st} ln(l) + st\sum_{l>st} n(l)\right].$$

The second difference of this nullity considered as a function of t is

$$-\langle\langle f\rangle\rangle\left\{\sum_{l=1}^{s}[ln(st+l)+(s-l)n(st+s+l)]\right\}. \tag{1}$$

If $s=1$, this reduces to

$$-\langle\langle f\rangle\rangle n(t+1). \tag{1'}$$

Returning to the problem of finding a matrix Y such that $P(Y)$ is similar to A, let h be the minimum polynomial of A with coefficients in the centrum, and let $h=\prod h_i^{k_i}$, where the h_i are distinct polynomials with coefficients in the centrum and irreducible in the

centrum. Let the characteristic divisors of A be $h_i^{q_{ik}}$. Moreover, let

$$h_i(P) = \prod f_j^{l_{ij}}$$

and let Y be a matrix with characteristic divisors $f_j^{p_{jk}}$. Let $m_j(s)$ and $n_j(s)$ equal, respectively, the number of the q_{jk} and p_{jk} which are equal to s. If $P(Y)$ is to be similar to A, the second difference of the nullity of $[h_i(P(Y))]^t$ must be the same as the second difference of the nullity of $[h_i(A)]^t$. Using (1) and (1′) and the fact that the nullity of the product of two relatively prime polynomials in a matrix is the sum of the nullities of the two polynomials in the given matrix taken separately, we see that this yields the diophantine equations

$$(2)\quad \sum_j \langle\langle f_j\rangle\rangle \sum_{l=1}^{l_{ij}} [ln_j(tl_{ij}+l) + (l_{ij}-l)n_j(tl_{ij}+l_{ij}+l)] = \langle\langle h_i\rangle\rangle m_i(t+1), \qquad (i,\ 0 \leqq t \leqq k_i).$$

This system of diophantine equations for the $n_j(s)$ can have only a finite number of positive integral solutions, each of which will define a matrix Y similar to a solution of $P(X)=A$, and any two solutions define dissimilar Y's. Moreover, every solution X will be similar to some Y thus defined.

4. **The equation** $P(X)\odot\xi = Q(A)\odot\xi$. It has been proved that this equation may be reduced to a system of equations $P_i(X)=A_i$, where the coefficients of the P_i are in the centrum. The solution of these equations having been reduced to the factorization of polynomials over a division algebra, there still remains the problem of picking out their simultaneous solutions. Consider the case of two equations. Let $X_{11}, X_{12}, \cdots, X_{1k}$ be a complete set of dissimilar solutions of $P_1(X)=A_1$. Any transform of X_{1i} by a non-singular matrix commutative with A is a solution of this equation, and all solutions are of this form. Let $X_{21}, X_{22}, \cdots, X_{2l}$ be a complete set of dissimilar solutions of $P_2(X)=A_2$. If X_{1i} is dissimilar to all the X_{2j}, then there is no simultaneous solution of the two equations similar to X_{1i}. Each of $P_1(X)=A_1$ and $P_2(X)=A_2$ defines an equation (2), and, of course, only such $n_j(l)$ as are simultaneous solutions for these equations (2) need be used. Suppose X_{11} is similar to X_{21}. One must still determine whether or not there exists a matrix X similar to X_{11} under a transformation commutative with A_1 and similar to X_{21} under a transformation commutative with A_2. If such an X exists, then the simultaneous solutions similar to X of the two equations are the transforms of X by non-singular matrices simultaneously commuta-

tive with A_1 and A_2. The problem of determining under what conditions such a matrix X exists and the nature of its transforms by matrices commutative with A_1 and A_2 is being considered by H. C. Trimble, G. Whaples, and the author.

5. **Two theorems on right linear transformations.** The following two theorems, though not necessary to the chief purpose of this paper, seem of interest. They were derived from certain suggestions of M. C. Wolf and F. A. Kiokemeister.

THEOREM 2. *The solutions of* $P(X) \odot \xi = 0$ *for all vectors* ξ *are all matrices* X *for which* $h(X) = 0$ *where* h *is the polynomial of maximum degree with coefficients in the centrum which divides* P.

Clearly such an X is a solution.

If $P = \sum P_i \alpha_i$, where the α's are basal elements of the division algebra D over its centrum C and the P_i have coefficients in C, then if $P(X) \odot \xi = 0$ for every ξ, it follows (from Theorem 2) that $P_i(X) = 0$ for every i. Hence each P_i is divisible by the minimum polynomial h_1 in C for which $h_1(X) = 0$.

If $g = \sum \lambda^i a_i$ and α is an element of D, then g_α, the transform* of g, is defined to be $\sum \lambda^i \alpha^{-1} a_i \alpha$.

THEOREM 3. *A necessary and sufficient condition that* $P(X) \odot \xi$ *for every* ξ *represents a right linear transformation is that* $h(X) = 0$ *where* h *is the polynomial of maximum degree with coefficients in the centrum* C *which divides* $P - P_\alpha$ *for every* α.

We have $P_\alpha(X) \odot (\xi\alpha) = (P(X) \odot \xi)\alpha$, and if $P(X) \odot \xi$ for every ξ represents a right linear transformation, then $P(X) \odot (\xi\alpha) = (P(X) \odot \xi)\alpha$ and it follows that $[P(X) - P_\alpha(X)]\xi = 0$ for every ξ. Hence Theorem 3 is a consequence of Theorem 2.

6. **Example.** The following example illustrates much of the theory of this paper.

Consider the equation $X^2 = A$, where

$$A = \begin{pmatrix} i & k \\ j & i - 1 \end{pmatrix}.$$

It is desired to find the matrices with rational quaternionic elements which are solutions of this equation.

---

* See O. Ore, *Theory of non-commutative polynomials*, Annals of Mathematics, (2), vol. 34 (1933), pp. 480–508.

The minimum equation with rational coefficients which is satisfied by A is

$$\lambda^2 + \lambda + 1 = 0,$$

and is irreducible rationally.

Hence

$$h_1 = \lambda^2 + \lambda + 1,$$

where

$$\lambda^2 + \lambda + 1 = [\lambda + (1 + i + j + k)/2][\lambda + (1 - i - j - k)/2].$$

Therefore $m_1(1) = 2$, and $\langle\langle h_1 \rangle\rangle = 1$.

Moreover,

$$h_1(\lambda^2) = (\lambda^2 + \lambda + 1)(\lambda^2 - \lambda + 1),$$

and with

$$f_1 = h_1, \qquad f_2 = \lambda^2 - \lambda + 1,$$

it follows that $l_{11} = l_{12} = 1$ and $\langle\langle f_1 \rangle\rangle = \langle\langle f_2 \rangle\rangle = 1$. Equation (2) of §3 becomes

$$n_1(1) + n_2(1) = 2$$

and has solutions

$$\begin{aligned} n_1(1) &= 2, \quad & n_2(1) &= 0, \\ n_1(1) &= 0, \quad & n_2(1) &= 2, \\ n_1(1) &= n_2(1) = 1. \end{aligned}$$

Hence if

$$Y_1 = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \qquad Y_2 = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix},$$

$$Y_3 = \begin{pmatrix} -(1 + i + j + k)/2 & 0 \\ 0 & (1 + i + j + k)/2 \end{pmatrix},$$

then Y_1, Y_2, and Y_3 will be similar to solutions of $X^2 = A$.

Since $T^{-1}Y_3^2T = A$ when

$$T = \begin{pmatrix} 1 & (1 + i + j + k)/2 \\ i & (1 - i - j + k)/2 \end{pmatrix},$$

it follows that

$$X_3 = T^{-1}Y_3T = \begin{pmatrix} 0 & -i \\ -1 & -j \end{pmatrix}$$

is a solution.

Similarly

$$X_1 = \begin{pmatrix} -1-i & -k \\ -j & -i \end{pmatrix}$$

and $X_2 = -X_1$ are solutions of the above equations similar to Y_1 and Y_2 respectively. All matrices commutative with A can be seen to be of the form $Q = PSP^{-1}$ where S is commutative with

$$P^{-1}AP = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}$$

and

$$P = \begin{pmatrix} 1 & -1-i \\ 0 & -j \end{pmatrix}.$$

The matrix Q can be shown to be of the form

$$\begin{pmatrix} q_1 - iq_2 & q_1k - iq_1j - iq_2k - q_2j - iq_2j \\ -jq_2 & -jq_2k - jq_1j - jq_2j \end{pmatrix}.$$

This is commutative with X_1 and X_2 but not with X_3 except for particular values of the quaternions q_1 and q_2. It follows that all solutions of $X^2 = A$ are X_1, X_2, and the infinite family of matrices $Q^{-1}X_3Q$, where Q is a non-singular matrix of the type given above.

UNIVERSITY OF WISCONSIN