

## SIMILARITY OF MATRICES IN WHICH THE ELEMENTS ARE REAL QUATERNIONS\*

BY LOUISE A. WOLF

1. *Introduction.* The purpose of this paper is to give a necessary and sufficient condition that for two matrices,  $A$  and  $B$ , of which the elements are real quaternions, there exist a non-singular matrix  $S$  whose elements are real quaternions such that  $SAS^{-1} = B$ . The matrices  $A$  and  $B$  are said to be similar if such a matrix  $S$  exists. This paper defines a set of invariant factors for any such matrix,  $A$ , in terms of the ranks of certain real polynomials in  $A$ .

2. *Definitions and Notations.* If  $A$  represents a matrix having  $m$  rows and  $n$  columns, then  $A'$  (read  $A$  transpose) is the matrix  $A$  with the rows and columns interchanged so that  $A'$  has  $n$  rows and  $m$  columns.

According to E. H. Moore a set of  $k$  vectors  $\eta_i$ , each being a matrix having  $n$  rows and one column, where  $\eta'_i = (y_{i1}, y_{i2}, \dots, y_{in})$ , ( $i = 1, 2, \dots, k$ ), whose elements  $y_{ij}$  are real quaternions, is left linearly dependent with respect to real quaternions if there exists a set of constants  $q_i$ , which are real quaternions and not all zero such that  $\sum_{i=1}^{i=k} q_i y_{ij} = 0$ , ( $j = 1, 2, \dots, n$ ). If no such set of real quaternions,  $q_i$ , exists except  $q_i \equiv 0$ , the vectors  $\eta_i$  are said to be left linearly independent. Similarly the  $k$  vectors  $\eta_i$  are right linearly dependent with respect to real quaternions if there exists a set of constants,  $q_i$ , which are real quaternions and not all zero, such that  $\sum_{i=1}^{i=k} y_{ij} q_i = 0$ , ( $j = 1, 2, \dots, n$ ). If no such set of real quaternions,  $q_i$ , exists except  $q_i \equiv 0$ , the vectors  $\eta_i$  are said to be right linearly independent.

Moore considered the columns of a matrix, whose elements are real quaternions, as vectors and defined the rank,  $r$ , of such a matrix,  $S$ , as the maximum number of columns of  $S$  which are right linearly independent with respect to real quaternions. He proved that if a matrix  $S$  is of rank  $r$ , then  $r$  is also the maximum number of rows of  $S$  that are left linearly independent with

---

\* Presented to the Society, September 13, 1935,

respect to real quaternions. He further defined a matrix  $S$  to be non-singular if the rank equals the order and proved that for every non-singular matrix  $S$  there exists a matrix  $S^{-1}$  such that  $S^{-1}S = SS^{-1} = I$ .

3. *A Correspondence.* The algebra of complex quaternions is equivalent to the total complex matrix algebra of order 4.\* A matrix  $\mathcal{A}$  corresponds to a matrix  $A$ ,  $\mathcal{A} \sim A$ , where the elements of  $A$  are real quaternions, if  $\mathcal{A}$  is obtained from  $A$  by the replacement of each element of  $A$  by a  $2 \times 2$  matrix obtained from the relations as given by Dickson. The elements of  $\mathcal{A}$  are in the complex field and  $\mathcal{A}$  is of order  $2n$  if  $A$  is of order  $n$ .

In the following theorems and definition  $A$  is a matrix in which the elements are real quaternions. The minimum function of  $A$  in the field of real numbers,  $R$ , is the polynomial of lowest degree,  $g(\lambda)$ , in which the coefficients are real and the leading coefficient unity such that  $g(A) = 0$ . The following theorems can readily be established.†

**THEOREM 1.** *The minimum function of  $A$  in the field  $R$  is unique.*

**THEOREM 2.** *If  $\mathcal{A} \sim A$ , then the minimum function of  $A$  in  $R$  is also the minimum function of  $\mathcal{A}$  in  $R$ .*

**THEOREM 3.** *If  $A$  is non-singular, then  $\mathcal{A}$  is non-singular and conversely.*

**THEOREM 4.** *A necessary and sufficient condition that  $A$  which corresponds to  $\mathcal{A}$  be similar to  $B$  which corresponds to  $\mathcal{B}$  is that there exist a non-singular matrix  $S$  with elements in the complex field such that  $S\mathcal{A}S^{-1} = \mathcal{B}$ .*

4. *A Special Type of Matrix with Elements in the Complex Field.* If  $\mathcal{A}$  is a matrix obtained from  $A$  by the replacement of the real quaternionic elements of  $A$  by  $2 \times 2$  matrices by means of the correspondence of § 3, then  $\mathcal{A}$  has the following properties.

\* See L. E. Dickson, *Linear Algebras*, p. 13.

† The proofs omitted in this article may be found in a thesis for the doctorate which is on file under the same title at the University of Wisconsin Library.

THEOREM 5. *The determinant of  $\mathcal{A}$  is real.*

THEOREM 6. *The determinants of the  $k \times k$  minors of  $(\mathcal{A} - \lambda I)$  are real or occur in pairs that are either conjugates of one another or negative conjugates.*

THEOREM 7. *The invariant factors of  $(\mathcal{A} - \lambda I)$  are real.*

5. *A Comparison of the Ranks of  $A$  and  $\mathcal{A}$ .* The statements given in this paragraph concerning the rank of matrices with quaternionic elements are taken from the Lectures of E. H. Moore: The vectors  $\eta_i$ , where  $\eta_i' = (\delta_{i1}, \delta_{i2}, \dots, \delta_{in})$ , ( $i = 1, 2, \dots, n$ ), and where  $\delta_{ji} = 0$  for  $j \neq i$  and  $\delta_{ii} = 1$ , form a base with respect to quaternionic coefficients for the vector space of which  $\eta$  is an element, where  $\eta' = (q_1, q_2, \dots, q_n)$ , and where the  $q_i$  are real quaternions. If a set of  $k$  vectors, such as  $\eta_i$ , where  $\eta_i' = (q_{i1}, q_{i2}, \dots, q_{in})$ , ( $i = 1, 2, \dots, k$ ), is right linearly independent with respect to real quaternions, it is always possible to complete the base for the vector space by adding  $n - k$  right linearly independent vectors such that all  $n$  vectors are right linearly independent with respect to real quaternions. The rank of a matrix  $A$  whose elements are real quaternions is  $n - s$  if  $s$  is the maximum number of right linearly independent vectors  $\eta_i$  such that  $A\eta_i = 0$ , ( $i = 1, 2, \dots, s$ ).

In considering the rank of the matrix  $\mathcal{A}$  which corresponds to the matrix  $A$ , it is convenient to consider the columns of these matrices as vectors. Let  $A = (\eta_1, \eta_2, \dots, \eta_n)$ , where  $\eta_i' = (q_{i1}, q_{i2}, \dots, q_{in})$  and  $q_{ji} = q_{ji0} + q_{ji1}i + q_{ji2}j + q_{ji3}k$ .

If  $\mathcal{A} \sim A$ , then  $\mathcal{A} = (\xi_1, \xi_2, \dots, \xi_{2n-1}, \xi_{2n})$ , where  $\eta_i \sim (\xi_{2i-1}, \xi_{2i})$  and the elements of  $\xi_i$  are in the complex field  $C$ .

THEOREM 8. *If the rank of  $A$  is  $k$ , then the rank of  $\mathcal{A}$  is  $2k$ .*

If  $A$  is of order  $n$ , then  $\mathcal{A}$  is of order  $2n$ . Suppose that the rank of  $A$  is  $n$ , then  $A$  is non-singular and  $\mathcal{A}$  is also non-singular due to the equivalence and hence  $\mathcal{A}$  is of rank  $2n$ . If  $A$  is of rank zero,  $\eta_i \equiv 0$  and hence  $\xi_i \equiv 0$  and  $\mathcal{A} \equiv 0$ . Therefore  $\mathcal{A}$  is of rank zero.

Suppose  $A$  is of rank  $k$ , where  $0 < k < n$ , then the following consideration shows that  $\mathcal{A}$  is at least of rank  $2k$ . Take  $k$  right linearly independent columns of  $A$  and designate them as  $\eta_1, \eta_2, \dots, \eta_k$ . Take  $\xi_1, \xi_2, \dots, \xi_{2k-1}, \xi_{2k}$  of  $\mathcal{A}$  so that  $\eta_i$

$\sim(\xi_{2i-1}, \xi_{2i})$ . Complete a base for the vector space  $\eta_i$  by adding  $\eta_{k+1}, \dots, \eta_n$  to  $\eta_1, \eta_2, \dots, \eta_k$ . Let  $\xi_{2k+1}, \xi_{2k+2}, \dots, \xi_{2n-1}, \xi_{2n}$  be the vectors that correspond to  $\eta_{k+1}, \dots, \eta_n$ . Then  $B = (\eta_1, \dots, \eta_k, \eta_{k+1}, \dots, \eta_n)$  and  $\mathcal{B} = (\xi_1, \xi_2, \dots, \xi_{2k-1}, \xi_{2k}, \xi_{2k+1}, \xi_{2k+2}, \dots, \xi_{2n-1}, \xi_{2n})$  are both non-singular matrices.  $B$  is non-singular by construction and  $\mathcal{B} \sim B$ . Thus  $\mathcal{B}$  is of rank  $2n$  and  $\xi_1, \xi_2, \dots, \xi_{2k-1}, \xi_{2k}$  are linearly independent. But  $\xi_1, \xi_2, \dots, \xi_{2k-1}, \xi_{2k}$  are columns of  $\mathcal{A}$  and hence  $\mathcal{A}$  has at least  $2k$  linearly independent columns and the rank of  $\mathcal{A}$  is equal to or greater than  $2k$ .

We can also prove as follows that the rank of  $\mathcal{A}$  is equal to or less than  $2k$ . Suppose  $A$  is of rank  $k = n - r$ , where  $r < n$ . Then there exist  $r$  right linearly independent vectors  $\eta_i \neq 0$  such that  $A\eta_i = 0, (i = 1, 2, \dots, r)$ . Because of the correspondence there are  $2r$  right linearly independent vectors  $\xi_1, \xi_2, \dots, \xi_{2r-1}, \xi_{2r}$  such that we have  $\mathcal{A}\xi_i = 0, (i = 1, 2, \dots, 2r)$ , where  $\eta_i \sim (\xi_{2i-1}, \xi_{2i}), (i = 1, 2, \dots, r)$ . The linear independence of the  $2r$  vectors,  $\xi_i$ , was established in the first part of the proof of this theorem. But if  $2r$  linearly independent vectors,  $\xi_i$ , exist such that  $\mathcal{A}\xi_i = 0$ , then the rank of  $\mathcal{A}$  is less than or equal to  $2n - 2r$ ; that is, the rank of  $\mathcal{A}$  is less than or equal to  $2k$ . Hence the rank of  $\mathcal{A}$  is  $2k$  if the rank of  $A$  is  $k$ .

6. *Similarity of Matrices Whose Elements Are Real Quaternions in Terms of Invariant Factors of These Matrices.* The matrix  $\mathcal{A}$ , which corresponds to  $A$  of order  $n$ , is a matrix of order  $2n$  with elements in the complex field,  $C$ . The irreducible factors of  $|\mathcal{A} - \lambda I|$  are of the form  $(\lambda - \alpha_i)$ , where the  $\alpha_i$  are in  $C$ . If  $r_{ij}$  is the rank of  $(\mathcal{A} - \alpha_i I)^j$  and  $r_{i0} \equiv 2n$ , it can be seen from the classical canonical form for  $\mathcal{A}$  that  $r_{ij} = 2n - \sum_{l=1}^j t_i^{(l)}$ , where  $t_i^{(l)}$  is the number of invariant factors that contain  $(\lambda - \alpha_i)^l$ . The set of positive integers  $r_{i0} - r_{i1} \geq r_{i1} - r_{i2} \geq r_{i2} - r_{i3} \geq \dots$  is monotonic non-increasing. If  $e_{ij} = r_{ij-1} - 2r_{ij} + r_{ij+1}$ , then  $(\lambda - \alpha_i)^j$  appears among the elementary divisors of  $(\mathcal{A} - \lambda I)$  a number of times equal to  $e_{ij}$  and the first invariant factor is obtained by taking the product of the highest powers of  $(\lambda - \alpha_i)$  which appear among the elementary divisors; the second invariant factor is obtained by taking the product of the second highest powers  $(\lambda - \alpha_i)$ , and so on.

A second method of defining the invariant factors of  $(\mathcal{A} - \lambda I)$  will be given which uses the fact that all the invariant factors

of  $(\mathcal{A}-\lambda I)$  are in the real field,  $\mathcal{R}$ . Let  $f_1(\lambda), f_2(\lambda), \dots, f_k(\lambda)$  be the irreducible factors of  $|\mathcal{A}-\lambda I| = f(\lambda)$  in  $\mathcal{R}$ . Let  $w_{ij}$  be the rank of  $[f_i(\mathcal{A})]^j$  and  $w_{i0} \equiv 2n$ .

It can easily be established that  $w_{ij} = 2n - \sum_{l=1}^j \mu_i t_i^{(l)}$ , where  $\mu_i$  is the degree of  $f_i(\lambda)$  and  $t_i^{(l)}$  is the number of invariant factors of  $(\mathcal{A}-\lambda I)$  that are divisible by  $[f_i(\lambda)]^l$ , where the invariant factors of  $(\mathcal{A}-\lambda I)$  are defined in terms of the elementary divisors in the conventional manner. It was previously proved that  $f(\lambda)$  is in the real field and hence the irreducible factors  $f_i(\lambda)$  of  $f(\lambda)$  in  $\mathcal{R}$  are either linear or quadratic. The two cases are considered separately as follows:

Case 1. Suppose  $\mu_i = 1$ , then  $f_i(\lambda) = \lambda - \alpha_i$ , where  $\alpha_i$  is real. It is clear that in this case  $w_{ij} = r_{ij} = 2n - \sum_{l=1}^j t_i^{(l)}$ , where  $t_i^{(l)}$  is the number of invariant factors divisible by  $(\lambda - \alpha_i)^l$  as before.

Case 2. Suppose  $\mu_i = 2$ , then  $f_i(\lambda) = (\lambda - \alpha_i)(\lambda - \bar{\alpha}_i)$ , where  $\alpha_i$  is complex. All the invariant factors of  $(\mathcal{A}-\lambda I)$  are real, hence  $(\lambda - \alpha_i)^k$  and  $(\lambda - \bar{\alpha}_i)^k$  are divisors of the same  $t_i^{(k)}$  invariant factors when the complex field is considered and therefore\*

$$w_{ij} = 2n - \sum_{l=1}^j t_i^{(l)} - \sum_{l=1}^j t_i^{(l)} = 2n - \sum_{l=1}^j 2t_i^{(l)}.$$

Thus  $w_{ij} = 2n - \sum_{l=1}^j \mu_i t_i^{(l)}$ , where  $\mu_i$  is the degree of  $f_i(\lambda)$ .  $w_{i0} - w_{i1} \geq w_{i1} - w_{i2} \geq w_{i2} - w_{i3}, \dots$ , is a set of monotonic non-increasing positive integers.

Let  $m_{ij} = w_{ij-1} - 2w_{ij} + w_{ij+1}$ . Define the characteristic divisors of  $(\mathcal{A}-\lambda I)$  to be the polynomials  $[f_i(\lambda)]^k$ , where  $[f_i(\lambda)]^k$  appears a number of times equal to  $m_{ik}/\mu_i$ .

The invariant factors of  $(\mathcal{A}-\lambda I)$  are defined to be in turn the product of the highest powers of the  $f_i(\lambda)$  appearing among the characteristic divisors, the product of the second highest powers of the  $f_i(\lambda)$  appearing among the characteristic divisors, and so on.

The elementary divisors of  $(\mathcal{A}-\lambda I)$  are the distinct linear factors of the characteristic divisors raised to the powers to which the characteristic divisors occur.

**THEOREM 9.** *The invariant factors of  $(\mathcal{A}-\lambda I)$  as defined in terms of characteristic divisors are the same as those defined in the conventional manner.*

\* See Menge, this Bulletin, vol. 38 (1932), pp. 88-94.

If  $[f_i(\lambda)]^k = (\lambda - \alpha_i)^k$ , where  $\alpha_i$  is real,  $[f_i(\lambda)]^k$  appears  $m_{ik}$  times among the characteristic divisors of  $(\mathcal{A} - \lambda I)$ , where  $m_{ik} = w_{ik-1} - 2w_{ij} + w_{ik+1}$  and  $w_{ik} = 2n - \sum_{l=1}^k t_i^{(l)}$  and hence  $e_{ik} = m_{ik}$ , since  $e_{ik} = r_{ik-1} - 2r_{ik} + r_{ik+1}$  and  $r_{ik} = 2n - \sum_{l=1}^k t_i^{(l)}$ .

If  $[f_i(\lambda)]^k = (\lambda - \alpha_i)^k (\lambda - \bar{\alpha}_i)^k$ , where  $\alpha_i$  is complex,  $[f_i(\lambda)]^k$  appears  $m_{ik}/2$  times among the characteristic divisors of  $(\mathcal{A} - \lambda I)$ , where  $m_{ik} = w_{ik-1} - 2w_{ik} + w_{ik+1}$ .

In this case  $w_{ik} = 2n - \sum_{l=1}^k 2t_i^{(l)}$ , and hence  $m_{ik} = 2t_i^{(k)} - 2t_i^{(k+1)}$  if the last two equations are combined. The invariant factors by the original definition are also real and if  $(\lambda - \alpha_i)^k$  appears  $e_{ik}$  times among the elementary divisors,  $(\lambda - \bar{\alpha}_i)^k$  also appears  $e_{ik}$  times and thus  $f_i(\lambda) = (\lambda - \alpha_i)^k (\lambda - \bar{\alpha}_i)^k$  appears  $e_{ik}$  times, where  $e_{ik} = r_{ik-1} - 2r_{ik} + r_{ik+1}$  and  $r_{ik} = 2n - \sum_{l=1}^k t_i^{(l)}$ . Thus  $e_{ik} = t_i^{(k)} - t_i^{(k+1)}$ , and it is evident that  $m_{ik} = 2e_{ik}$ .

This is sufficient to establish the theorem because the invariant factors defined in terms of characteristic divisors will be the products of the same factors since  $(\lambda - \alpha_i)^k$  appears the same number of times among the characteristic divisors as it appears among the elementary divisors.

To consider  $A$  whose elements are real quaternions, let  $g(\lambda)$  be the minimum function of  $A$  in  $R$  and let  $f_i(\lambda)$  be the distinct irreducible factors of  $g(\lambda)$  in  $R$ . Let  $\rho_{ij}$  be the rank of  $[f_i(A)]^j$  and define  $\rho_{i0}$  to be  $n$ . Then  $\rho_{i0} - \rho_{i1} \geq \rho_{i1} - \rho_{i2} \geq \rho_{i2} - \rho_{i3} \geq \dots$ .

Let  $\epsilon_{ij} = \rho_{ij-1} - 2\rho_{ij} + \rho_{ij+1}$  and let  $[f_i(\lambda)]^k$  appear among the characteristic divisors of  $A$  a number of times equal to  $2\epsilon_{ik}/\mu_i$ , where  $\mu_i$  is the degree of  $f_i(\lambda)$ .

The invariant factors of  $A$  are defined to be in turn the product of the highest powers of the  $f_i(\lambda)$  appearing among the characteristic divisors, the product of the second highest powers of the  $f_i(\lambda)$  appearing among the characteristic divisors,  $\dots$ , in a manner analogous to the above definition in terms of the elementary divisors.

Theorem 4 proves that  $A$  is similar to  $B$  if and only if the matrices  $\mathcal{A}$  and  $\mathcal{B}$  are similar, where  $\mathcal{A} \sim A$  and  $\mathcal{B} \sim B$ .  $\mathcal{A}$  is similar to  $\mathcal{B}$  if and only if the invariant factors of  $(\mathcal{B} - \lambda I)$  and  $(\mathcal{A} - \lambda I)$  are the same. The invariant factors of  $A$  as defined above are the same as the invariant factors of  $(\mathcal{A} - \lambda I)$ . This is evident when it is recalled that the rank of  $f_i(\mathcal{A})$  is  $2k$  if the rank of  $f_i(A)$  is  $k$ , and hence  $m_{ik} = 2\epsilon_{ik}$ . Furthermore,  $A$  and  $\mathcal{A}$  have the same minimum function in  $R$ , namely,  $g(\lambda)$ . The distinct

irreducible factors of  $f(\lambda) = |\mathcal{A} - \lambda I|$  in  $R$  are likewise the distinct irreducible factors of  $g(\lambda)$ , the minimum function of  $\mathcal{A}$  in  $R$ . Thus the characteristic divisors of  $(\mathcal{A} - \lambda I)$  and  $\mathcal{A}$  are the same and hence their invariant factors are the same. The same is true of  $(\mathcal{B} - \lambda I)$  and  $\mathcal{B}$  because  $\mathcal{B} \sim \mathcal{B}$ . The following theorem has thus been established.

**THEOREM 10.** *A is similar to B if and only if the invariant factors of A are the same as the invariant factors of B.*

THE UNIVERSITY OF WISCONSIN

---

THE THEOREM " $p \rightarrow q = .pq = p$ " AND  
HUNTINGTON'S RELATION BETWEEN LEWIS'S  
STRICT IMPLICATION AND BOOLEAN  
ALGEBRA

BY TANG TSAO-CHEN

In this Bulletin, vol. 40 (1934), p. 729, E. V. Huntington pointed out that the relation called "strict implication" in C. I. Lewis's system of logic can be shown to be substantially equivalent to the relation called subsumption in ordinary Boolean algebra. His main result is as follows:

Whenever we find the formula " $p \rightarrow q$ " asserted, we may thereupon write down the formula " $p = pq$ "; and conversely, whenever we find the formula " $p = pq$ " established, we may write down that the formula " $p \rightarrow q$ " is asserted.

That is, Huntington's relation is

$$("p \rightarrow q" \text{ is asserted}) \Leftrightarrow ("pq = p" \text{ is established}).$$

This relation is not the same as the following theorem:

$$p \rightarrow q = .pq = p,$$

where "=" takes the meaning of logical equivalence given in Lewis's *Symbolic Logic*.

This theorem being not explicitly mentioned in Lewis's *Symbolic Logic*, I shall prove it here.

Throughout this paper we shall follow Lewis's practice of ignoring the distinction, which is characteristic of Huntington's