RESOLUTION INTO INVOLUTORY SUBSTITUTIONS OF THE TRANSFORMATIONS OF A NON-SINGULAR BILINEAR FORM INTO ITSELF*

BY

DUNHAM JACKSON

In a recent number of the Transactions of the Connecticut Academy, E. B. WILSON obtains a necessary and sufficient condition that a linear transformation be factorable into two involutory transformations. If this condition is translated from the language of dyadics, which WILSON uses, into that of the algebra of matrices, its form suggests at once a theorem of FROBENIUS concerning transformations of a bilinear form into itself. A mere combination of these two theorems is sufficient to establish the following, which is simpler in statement than either:

I. *A necessary and sufficient condition that a linear transformation be such as to carry some non-singular bilinear form into itself, is that it be factorable into two involutory transformations.*

It will be seen that the part of this theorem relating to the necessity of the condition is a generalization of a theorem of P. F. SMITH† which WILSON uses in proving his theorem, and which is substantially as follows:

II. (SMITH). *A necessary condition that a linear transformation be such as to carry some non-singular quadratic form into itself, is that it be factorable into two involutory transformations.*

The converse of SMITH's theorem is not true, as is remarked by WILSON; a substitution which is factorable into two involutory transformations is not necessarily capable of carrying a non-singular *quadratic* form into itself. In fact, the question of the possibility of factoring a transformation into two involutory transformations has no essential relation to the subject of symmetric bilinear forms, as distinguished from bilinear forms in general.‡

* Presented to the Society (Princeton), September 13, 1909.

† P. F. SMITH, these Transactions, vol. 6 (1905), p. 13.

‡ For a necessary and sufficient condition that a transformation be capable of carrying a non-singular quadratic form into itself, see FROBENIUS, Crelle's Journal, vol. 84 (1878), p. 41. Compare with this the theorem numbered IV below.

The theorems of WILSON and FROBENIUS mentioned at the beginning of the article may be stated thus:

III. (WILSON).* *A necessary and sufficient condition that a linear transformation be factorable into two involutory transformations, is that each elementary divisor of its characteristic matrix be paired with another, whose degree is the same, and whose root is the reciprocal of the root of the first, except that an elementary divisor which is a power of* $(\lambda - 1)$ *or* $(\lambda + 1)$ *may be paired with itself.*

IV. (FROBENIUS.)† *A necessary and sufficient condition that a linear transformation be such as to carry some non-singular bilinear form into itself,*‡ *is that each elementary divisor of its characteristic matrix be paired with another, whose degree is the same, and whose root is the reciprocal of the root of the first, except that an elementary divisor which is a power of* $(\lambda - 1)$ *or* $(\lambda + 1)$ *may be paired with itself.*

Theorem I is an obvious consequence of III and IV. But while FROBENIUS in proving IV uses the method of the algebra of matrices, WILSON's proof of III depends on GIBBS's method of dyadics and double products, and also on the work of SMITH, who uses still another method. It is possible, however, to give a proof of WILSON's theorem depending only on the algebra of matrices, and so to obtain greater uniformity of treatment. For a part of the proof it is necessary only to rewrite WILSON's dyadics in the matrix notation; the use of SMITH's theorem in establishing the sufficiency of the condition is avoided by actually writing down the factors in all, as WILSON does in some, of the special cases that are first considered.

In the course of the demonstration, reference will be made to another theorem of FROBENIUS, slightly more general than the one already given:

V. (FROBENIUS.) § *A necessary and sufficient condition that two linear transformations* α, β *be such as to carry some non-singular bilinear form into itself,*‖ *is that the elementary divisors of the characteristic matrix of* β *be obtainable from those of* α *by replacing each root of the characteristic equation by its reciprocal.*

From this, which is itself not difficult to prove, Theorem IV follows immediately. For if it is assumed that the two sets of variables in the bilinear form are to be subjected to the same transformation, then $\beta = \alpha'$, where α' denotes the conjugate of α, and the theorem just stated reduces to IV, since the elementary divisors of the characteristic matrix of α' are the same as those of α.

---

* E. B. WILSON, *Theory of Double Products and Strains in Hyperspace*, Transactions of the Connecticut Academy of Arts and Sciences, September, 1908; p. 41.

† Loc. cit., p. 34.

‡ That is, when applied to each of the two sets of cogredient variables.

§ Loc. cit., p. 31.

‖ That is, as FROBENIUS uses the terms, that a non-singular matrix ϕ exists such that $\alpha\phi\beta = \phi$.

To proceed with the proof of WILSON's theorem: Suppose at first, that α is the matrix of the product of two involutory transformations, so that

$$\alpha = \phi\psi \qquad \text{where} \qquad \phi^2 = I, \qquad \psi^2 = I.$$

Then

$$\alpha^{-1} = (\phi\psi)^{-1} = \psi^{-1}\phi^{-1} = \psi\phi,$$

since

$$\phi^{-1} = \phi, \qquad \psi^{-1} = \psi.$$

If β, γ are any two non-singular matrices, the elementary divisors of the characteristic matrices of $\beta\gamma$ and of $\gamma\beta$ are the same, since

$$\gamma\beta = \gamma(\beta\gamma)\gamma^{-1}$$

(see, for example, BÔCHER, *Introduction to Higher Algebra*, p. 286).

If γ is any non-singular matrix, the elementary divisors of the characteristic matrix of γ^{-1} can be obtained from those of γ by replacing each root by its reciprocal, by Theorem V above, for

$$\gamma\gamma\gamma^{-1} = \gamma.$$

In the present case, α, ϕ, ψ are necessarily non-singular, since $\phi^2 = \psi^2 = I$. Therefore the elementary divisors of the characteristic matrix of α^{-1} are, on the one hand, the same as those of α, and, on the other, the same as those of α with each root replaced by its reciprocal; and this amounts to saying that the condition stated in Theorem III is necessary.

Next, suppose that α is any matrix such that the elementary divisors of its characteristic matrix satisfy this condition. The form of the condition insures that α is non-singular. It is to be shown that α is the product of two square roots of the unit-matrix. It will be convenient to make use of the following

LEMMA. If α is a matrix expressible in the form $\phi\psi$, where $\phi^2 = \psi^2 = I$, then any matrix β, whose characteristic matrix has the same elementary divisors as that of α, is so expressible.

For β may be written in the form

$$\beta = \gamma\alpha\gamma^{-1} = \gamma\phi\psi\gamma^{-1} = \gamma\phi\gamma^{-1}\cdot\gamma\psi\gamma^{-1}$$

where

$$(\gamma\phi\gamma^{-1})^2 = \gamma\phi\gamma^{-1}\gamma\phi\gamma^{-1} = \gamma\phi^2\gamma^{-1} = \gamma\gamma^{-1} = I$$

and similarly

$$(\gamma\psi\gamma^{-1})^2 = I.$$

The truth of the following statements becomes obvious on writing down the formulæ:

If the n-rowed matrix β has elements different from zero only in an m-rowed principal sub-matrix $\bar{\beta}$, and the n-rowed matrix γ has elements different from zero only in the corresponding m-rowed sub-matrix $\bar{\gamma}$, then the product $\beta\gamma$ has elements different from zero only in the corresponding m-rowed principal sub-matrix, and these elements are those of the matrix $\bar{\beta}\bar{\gamma}$ obtained by multiplying together $\bar{\beta}$ and $\bar{\gamma}$ as m-rowed matrices.

If the n-rowed matrices β, γ have elements different from zero only in the principal sub-matrices $\bar{\beta}$, $\bar{\gamma}$ (not necessarily of the same order), neither of which contains any row or column corresponding to a row or column of the other, then $\beta\gamma = 0$.

Therefore, if $\beta = \beta_1 + \beta_2 + \cdots + \beta_p$, $\gamma = \gamma_1 + \gamma_2 + \cdots + \gamma_p$, where the non-vanishing elements of $\beta_1, \cdots, \beta_p$ form non-overlapping $m_1, \cdots, m_p$-rowed principal sub-matrices, and those of $\gamma_1, \cdots, \gamma_p$ occur in the corresponding sub-matrices, then

$$\beta\gamma = \beta_1\gamma_1 + \beta_2\gamma_2 + \cdots + \beta_p\gamma_p,$$

where in forming the products $\beta_i\gamma_i$, as far as the non-vanishing elements are concerned, the factors may be regarded as matrices either of order m_i or of order n.*

In particular, if $\beta = \beta_1 + \beta_2 + \cdots + \beta_p$ as before, and the sub-matrix in each β_i is factorable into two square roots $\phi_i\psi_i$ of the unit-matrix of order m_i, then β is factorable into two square roots of the unit-matrix of order n, for

$$(\phi_1 + \cdots + \phi_p)^2 = I, \qquad (\psi_1 + \cdots + \psi_p)^2 = I,$$

and

$$(\phi_1 + \cdots + \phi_p)(\psi_1 + \cdots + \psi_p) = \beta.$$

It will be shown next that the condition of the theorem is sufficient in the special cases that the characteristic matrix of the given matrix has just one elementary divisor $(\lambda \pm 1)^n$, or just two elementary divisors $(\lambda - a)^e$, $(\lambda - 1/a)^e$. By the lemma above, it will be enough to exhibit in each case a single matrix which can be factored into two square roots of the unit-matrix, and whose characteristic matrix has the elementary divisor or divisors in question. The work that has just been done will then make it easy to complete the proof of the theorem in general.

Let two matrices ϕ, ψ be defined as follows: The elements in the principal diagonal of each shall be alternately 1 and -1, and those just above the principal diagonal alternately 1 and 0, the 1's of ϕ, above the diagonal, correspond- to the 0's of ψ; all the other elements are zero. If n, the order of the matrices, is 4, to take a definite case, the formulæ are

$$\phi = \left\|\begin{matrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \end{matrix}\right\|, \qquad \psi = \left\|\begin{matrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{matrix}\right\|.$$

* FROBENIUS, loc. cit., p. 18.

On multiplication, it appears that $\phi^2 = I$, $\psi^2 = I$, and $\phi\psi$ is a matrix whose characteristic determinant is equal to $(\lambda - 1)^n$; when $n = 4$, as above,

$$\phi\psi = \begin{Vmatrix} 1 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{Vmatrix}.$$

The first minor in the upper right-hand corner of the characteristic determinant does not vanish for $\lambda = 1$, so that there is just one elementary divisor $(\lambda - 1)^n$. Therefore any matrix whose characteristic matrix has only a single elementary divisor, of the form $(\lambda - 1)^n$, is factorable into two square roots of the unit-matrix. A similar result may be deduced for the case that there is a single elementary divisor $(\lambda + 1)^n$, by considering the product $(-\phi)\psi$.

Let the notation

$$\begin{Vmatrix} \alpha & \beta \\ \gamma & \delta \end{Vmatrix}$$

represent the matrix of order $2e$ having the e-rowed matrices α, β, γ, δ, in its corners, arranged as indicated. A similar notation might be employed for matrices of order ne. By actually following through the process of multiplication, it is fairly easy to see that the ordinary rule for matrix multiplication still holds, if the elements of the matrices are regarded, not as scalars, but as themselves representing matrices, all of the same order.*

Let $\bar\phi$, $\bar\psi$ be two square roots of the unit-matrix of order e, whose product $\bar\phi\bar\psi$ is such that its characteristic matrix has just one elementary divisor, $(\lambda - 1)^e$. It has been shown that such matrices exist. Let a be any number different from zero. Let

$$\phi = \begin{Vmatrix} 0 & \sqrt{a}\,\bar\phi \\ \frac{1}{\sqrt{a}}\bar\phi & 0 \end{Vmatrix}, \qquad \psi = \begin{Vmatrix} 0 & \frac{1}{\sqrt{a}}\bar\psi \\ \sqrt{a}\,\bar\psi & 0 \end{Vmatrix},$$

where $\sqrt{a}$ may denote either square root, provided it is given the same value throughout. Form the product of ϕ and ψ:

$$\phi\psi = \begin{Vmatrix} a\bar\phi\bar\psi & 0 \\ 0 & \frac{1}{a}\bar\phi\bar\psi \end{Vmatrix}.$$

* Regard each matrix as a sum of matrices, in each of which all the matrix-elements but one are zero.

The characteristic matrix of $a\bar{\phi}\bar{\psi}$ has just one elementary divisor $(\lambda - a)^e$, and that of $1/a\bar{\phi}\bar{\psi}$ has just one elementary divisor $(\lambda - 1/a)^e$. Therefore the characteristic matrix of $\phi\psi$ has just two elementary divisors, $(\lambda - a)^e$ and $(\lambda - 1/a)^e$. Furthermore,

$$\phi^2 = \psi^2 = \left\| \begin{matrix} I_e & 0 \\ 0 & I_e \end{matrix} \right\| = I,$$

where I_e represents the unit-matrix of order e, and I, that of order $2e$. Consequently any matrix whose characteristic matrix has just two elementary divisors $(\lambda - a)^e$ and $(\lambda - 1/a)^e$ is factorable into two square roots of the unit-matrix.

Return now to the matrix α, which was assumed to be such that the elementary divisors of its characteristic matrix satisfy the conditions of the theorem. It is possible to write down a matrix β, having these same elementary divisors in its characteristic matrix, and expressible in the form $\beta_1 + \beta_2 + \cdots + \beta_p$, where the non-vanishing elements of $\beta_1, \cdots, \beta_p$ form non-overlapping principal minors, and the characteristic matrix of the sub-matrix corresponding to each β_i has either just two elementary divisors, $(\lambda - a)^{e_i}$ and $(\lambda - 1/a)^{e_i}$, or just one elementary divisor, $(\lambda \pm 1)^{m_i}$. In either case, the sub-matrix is factorable into two square roots of the corresponding unit-matrix. Therefore β, regarded as a linear substitution, is the product of two involutory transformations, and hence the same is true of α.

HARVARD UNIVERSITY, CAMBRIDGE, MASS.,
July 1 1909.