Choose n to be a prime q represented by C and prime to d. Then $g(q)=2$ if C is ambiguous, $g(q)=1$ if $C\neq C^{-1}$. If a form f_1 is associated with a form g_1 not in C or C^{-1}, $f_1(p^2q)=\sigma g_1(q)=0$. Hence, by (2) and (3), p^2q is represented in exactly $\eta\{p-(d'|p)\}\sigma^{-1}$ classes K, where η is 1 or 2 according as q is represented in only one (ambiguous) or two (reciprocal) primitive classes of discriminant d'.

McGILL UNIVERSITY

ON A REDUCTION OF A MATRIX BY THE GROUP OF MATRICES COMMUTATIVE WITH A GIVEN MATRIX*

BY P. L. TRUMP

1. *Introduction.* Two $n\times n$ matrices A and B, with elements in any field F, are said to be similar in F if there exists a non-singular $n\times n$ matrix S, with elements in F, such that $S^{-1}AS=B$.

Ingraham† has given a method for finding the most general solution, with elements in F, of the matrix equation

$$P(X) = A,$$

where $P(X)$ is a polynomial with coefficients in F, and A is a square matrix with elements in F. A certain set of dissimilar solutions $X_1, X_2, \cdots, X_k$ were obtained in terms of which the complete system of solutions was seen to be in the form $S^{-1}X_iS$, where S is commutative with A. The X_i's are obviously commutative with A.

The purpose of this investigation is to determine the conditions under which two $n\times n$ matrices C and D are similar under transformations of the group $[S]$ of non-singular matrices S which are commutative with a certain $n\times n$ matrix A, where the matrices C and D are also commutative with A. We then seek to describe possible canonical forms to which such matrices

---

* Presented to the Society, September 4, 1934. This paper with proofs and detail that are omitted here, is on file as a doctor's thesis at the Library of the University of Wisconsin.

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

may be reduced. At this writing the reduction has been completed with the exception of one case, the nature of which will be made clear. Procedures may in any case be given, however, for determining whether or not two particular matrices are similar under the group $[S]$. Such procedures lead to the necessity of solving a set of linear equations, a procedure which becomes involved and fails to give general results.

2. *Matrices Commutative with A.* Let T be any non-singular square matrix. The condition that $AC=CA$ is equivalent to the condition that $GH=HG$, where $G=T^{-1}AT$ and $H=T^{-1}CT$. Thus A may be considered in classical canonical form consisting of zeros except for square matrices A_i, $(i=1, 2, \cdots, r)$, along the main diagonal. The elements of A_i are zeros except for the characteristic roots α_i (not necessarily in F) of A on the main diagonal and possibly 1's in certain positions on the first diagonal below. Further, consider α_i different from α_j, if i is different from j.

The matrix C is commutative with A if and only if it is zero except for principal minor matrices C_i of the same dimension and in the same position as A_i, $(i=1, 2, \cdots, r)$, where the form of C_i is determined by the condition that $AC=CA$ if and only if $A_iC_i=C_iA_i$. Further, if S is also in the form just referred to, then $S^{-1}CS=D$ is again of the same form with $S_i^{-1}C_iS_i=D_i$.

We are thus led to a consideration of the case in which all the characteristic values of A are the same, in particular, we may assume them zero. The elementary divisors of A are then

$$\lambda^{n_1}, \lambda^{n_2}, \cdots, \lambda^{n_m},$$

where we may assume $n_i \geqq n_{i+1}$, $(i=1, 2, \cdots, m-1)$.

We will describe the frequently displayed form of the most general matrix C commutative with A in this form:*

CASE I. If A has the single elementary divisor λ^{n_1}, then C is an $n_1 \times n_1$ diagonalized matrix (c_{ij}), where c_{ij} is arbitrary except for the conditions

$$c_{ij}=0, \quad (i<j), \quad c_{ij}=c_{i+1,j+1}, \quad (i, j=1, 2, \cdots, n_1-1).$$

CASE II. In the more general case that A has m elementary divisors we will consider C to be a matrix of matrices C_{ij},

* Cullis, *Matrices and Determinoids*, vol. 3, part 1, sec. 242.

$(i, j=1, 2, \cdots, m)$, where the block C_{ii}, of order n_i, is as described in Case I. The block C_{ij}, $i<j$, will be an $n_i \times n_j$ matrix in which the first $n_i - n_j$ rows are zero and the remaining $n_j \times n_j$ matrix is of the same form as C_{jj}. If $i>j$, C_{ij} will be an $n_i \times n_j$ matrix in which the last $n_j - n_i$ columns are zero and the remaining $n_i \times n_i$ matrix is of the same form as C_{ii}. Define the ρ's as follows, if $i<j$: $\rho_{ij}=\rho_{ji}=n_i-n_j$, $\rho_{ii}=0$. The elements of C_{ij} will be indicated by c_{ijk}. If the element occurs in the rth diagonal of the indicated square portion of C_{ij}, then $k=2(r-1)+\rho_{ij}$.

In order to illustrate, let the elementary divisors of A be λ^3 and λ^2. We will then assume A and C in the following forms

$$A = \begin{Vmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{Vmatrix}, \quad C = \begin{Vmatrix} c_{110} & 0 & 0 & 0 & 0 \\ c_{112} & c_{110} & 0 & c_{121} & 0 \\ c_{114} & c_{112} & c_{110} & c_{123} & c_{121} \\ c_{211} & 0 & 0 & c_{220} & 0 \\ c_{213} & c_{211} & 0 & c_{222} & c_{220} \end{Vmatrix}.$$

We may now define certain unit matrices in terms of which C may be expressed. Let e_{ijk} represent the matrix of the same dimension as C with $c_{ijk}=1$ and all other elements zero. We may then write C in the form

$$C = \sum_{i,j,t} c_{ijt} e_{ijt}, \qquad (i, j = 1, 2, \cdots, m), \tag{1}$$

and $t-\rho_{ij}=0, 2, \cdots, 2(n_l-1)$, where l is the greater of i and j. The multiplication table for these basal units e_{ijk} is as follows:

$$e_{ijk}e_{rst} = \begin{cases} 0, & j \neq r, \\ e_{i,s,k+t}, & j = r. \end{cases}$$

If $k+t>2(n_l-1)+\rho_{ij}$, where l is the greater of i and s, then $e_{i,s,k+t}=0$.

3. *An Isomorphism.* At this point we will establish an isomorphism through which we are led to a consideration of matrices of reduced dimension whose elements are certain polynomials corresponding to the blocks C_{ij}. To C as expressed in (1) we make correspond a matrix P of the form

$$P = (P_{ij}(\gamma^2)\gamma^{\rho_{ij}}), \qquad (i, j = 1, 2, \cdots, m), \tag{2}$$

where

$$P_{ij}(\gamma^2) = c_{ijw} + c_{ijw+2}\gamma^2 + c_{ijw+4}\gamma^4 + \cdots,$$

with $w=\rho_{ij}$. The product of P and $Q=(Q_{ij}(\gamma^2)\gamma^{\rho_{ij}})$ is determined as usual and then the polynomial in the ith row and jth column is reduced modulo $\gamma^{2(n_l-1)+\rho_{ij}}$, where l is the greater of i and j.

The correspondence is obviously preserved under addition and scalar multiplication. To establish the preservation of the correspondence under multiplication, it is convenient to consider all $m\times m$ matrices of the type P with γ^k in the ith row and jth column, where $k-\rho_{ij}=0, 2, \cdots, 2(n_l-1)$, ($l$ the greater of i and j), and zeros elsewhere, as a set of basal units in terms of which matrices of type P may be expressed with scalar coefficients. Indicate such a matrix by f_{ijk}. Let e_{ijk} correspond to f_{ijk}. Since the multiplication table for these basal units is the same in each case the isomorphism is established.

It can be shown that the determinant of P is a polynomial in γ^2 and that P^{-1} exists and is of type (2) if and only if $|P| \not\equiv 0 \bmod \gamma$.

4. *Reduction to a Semi-Canonical Form.* We proceed with the reduction of matrices defined by the isomorphism by transformations with non-singular matrices of the same type. Consider

$$Q = (Q_{ij}(\gamma^2)\gamma^{\rho_{ij}}),$$

where

$$Q_{ij}(\gamma^2) = q_{ijt} + q_{ijt+2}\gamma^2 + \cdots, \qquad (t = \rho_{ij}).$$

Suppose

$$\rho_{uv} = 0, \text{ but } \rho_{u-1,u} \neq 0 \text{ and } \rho_{v,v+1} \neq 0,$$

if defined for some $u<v\leqq m$. Let

$$q = (q_{ij0}), \qquad (i, j = u, u+1, \cdots, u+v). \tag{3}$$

Let F_1 be the field obtained by enlarging F to include the roots of the characteristic equation of q. We will from now on restrict ourselves to the field F_1 instead of F. For some suitably chosen matrix $t=(t_{ij0})$ we may reduce q to its classical canonical form

$$q' = t^{-1}qt = (q'_{ij0}).$$

Using

$$T = (T_{ij}(\gamma^2)\gamma^{\rho_{ij}}), \qquad (i, j = 1, 2, \cdots, m),$$

where

$$T_{ij}(\gamma^2) = t_{ij0}, \qquad (i, j = u, u+1, \cdots, u+v),$$

and otherwise

$$T_{ij}(\gamma^2) \text{ is zero for } i \neq j \text{ and 1 for } i = j,$$

we obtain

$$P = T^{-1}QT = (P_{ij}(\gamma^2)\gamma^{\rho_{ij}}),$$

where

$$P_{ij}(\gamma^2) = p_{ijw} + p_{ijw+2}\gamma^2 + \cdots, \qquad (i, j = 1, \cdots, m;\ w = \rho_{ij}),$$

$$p_{ij0} = q'_{ij0}, \qquad (i, j = u, u+1, \cdots, u+v).$$

This argument may be repeated for any u for which (3) is true. All elements above the main diagonal are now congruent to zero modulo γ. Let $G(\lambda, \gamma^2)$ be the characteristic equation of P,

$$G(\lambda, \gamma^2) = |P - \lambda I| = g_0(\lambda) + g_2(\lambda)\gamma^2 + \cdots,$$

where

$$g_0(\lambda) = \prod_{i=1}^{m} (p_{ii0} - \lambda) = g_{10}(\lambda)g_{20}(\lambda)\cdots g_{s0}(\lambda),$$

$$g_{i0}(\lambda) = (\lambda_i - \lambda)^{r_i}, \qquad (\lambda_i \neq \lambda_j \text{ when } i \neq j),$$

that is, the λ_i $(i=1, 2, \cdots, s)$, represent the distinct values of p_{jj0}, $(j=1, 2, \cdots, m)$. We will denote by the set r_i those j for which $p_{jj0}=\lambda_i$.

$G(\lambda, \gamma^2)$ may be factored as follows.

$$G(\lambda, \gamma^2) = G_i(\lambda, \gamma^2)H_i(\lambda, \gamma^2), \qquad (i = 1, 2, \cdots, s),$$

where

$$G_i(\lambda, 0) = g_{i0}(\lambda), \qquad H_i(\lambda, 0) = h_{i0}(\lambda) = \prod_{j=1, j\neq i}^{s} g_{j0}(\lambda).$$

Such factorization may be made to depend upon the fact that the resultant* of $h_{i0}(\lambda)$ and $g_{i0}(\lambda)$ is not zero.

We now examine the matrix product

$$G_i(P, \gamma^2)H_i(P, \gamma^2) \equiv 0.$$

The columns of $H_i(P, \gamma^2)$ are vectors orthogonal to the matrix $G_i(P, \gamma^2)$. The elements in the main diagonal of $G_i(P, 0)$ are congruent modulo γ to

$$g_{i0}(\lambda_j)\begin{cases} = 0, & j \text{ in the set } \tau_i, \\ \not\equiv 0, & j \text{ not in the set } \tau_i; \end{cases}$$

above the main diagonal the elements are congruent to zero. The elements in the main diagonal of $H_i(P, 0)$ are congruent modulo γ to

$$h_{i0}(\lambda_j)\begin{cases} \not\equiv 0, & j \text{ in the set } \tau_i, \\ = 0, & j \text{ not in the set } \tau_i; \end{cases}$$

above the main diagonal they are congruent to zero.

Pick vectors $\xi_{i1}, \xi_{i2}, \cdots, \xi_{ir_i}$, $(i=1, 2, \cdots, s)$, as the columns numbered j of $H_i(P, \gamma^2)$, where j is in the set τ_i. This will determine $\sum_{j=1}^{s} r_j = m$ vectors ξ_{ij}. We propose to use ξ_{ij}, $(j=1, \cdots, r_i)$, as the τ_i, $(i=1, 2, \cdots, s)$, columns of the transforming matrix S. This insures that S is of admissible type. Further,

$$|S| \equiv \prod h_{i0}(\lambda_j) \not\equiv 0 \bmod \gamma^2, \qquad (i = 1, 2, \cdots, s; j \text{ in } \tau_i).$$

Therefore S^{-1} exists, and the vectors ξ_{ij}, $(i=1, \cdots, s; j=1, \cdots, r_i)$, form a complete vector space.

It can be proved that the following lemma holds.

LEMMA. *Every vector η satisfying*

$$G_i(P, \gamma^2)\eta = 0,$$

where i is any one of the numbers 1, 2, $\cdots$, s, is expressible as follows:

$$\eta = \sum_{j=1}^{r_i} a_j \xi_{ij}, \qquad (a_j \text{ are polynomials in } \gamma^2).$$

where the a_j are polynomials in γ^2.

* Bôcher, *Introduction to Higher Algebra*, pp. 195, 196.

We refer to the columns of S, picked and ordered as above, as vectors ξ_i, $(i=1, 2, \cdots, m)$.

Let η_i', $(i=1, \cdots, m)$, represent the rows of S^{-1}. Then the vectors η_i' and ξ_j form a biorthogonal system.

$$S^{-1}PS = \begin{Vmatrix} \eta_1' \\ \eta_2' \\ \cdot \\ \cdot \\ \eta_m' \end{Vmatrix} P(\xi_1, \xi_2, \cdots, \xi_m) = R = (R_{kj}),$$

where $R_{kj}=\eta_k P\xi_j$. If ξ_j is one of the set $(\xi_{i0}, \xi_{i1}, \cdots, \xi_{ir_i})$, then

$$G_i(P, \gamma^2)P\xi_j = PG_i(P, \gamma^2)\xi_j = 0,$$

and $P\xi_j$ is orthogonal to $G_i(P, \gamma^2)$. Hence, by the lemma above,

$$P\xi_j = \sum_{l=1}^{r_i} a_l\xi_i,$$

and $\eta_k P\xi_j=0$ for any k not one of the set τ_i.

Thus in R, any element R_{ij} corresponding to the intersection of P_{ii} and P_{jj}, where $p_{ii0}\not\equiv p_{jj0}$, has been reduced to zero.

The complete solution of the problem may now be shown to rest upon the reduction to canonical forms of matrices M in which the elements along the main diagonal are congruent to each other modulo γ and those above the main diagonal are congruent to zero. If the elements $m_{ij}(\gamma^2)\gamma^{\rho_{ij}}$ of M are congruent to zero modulo $\gamma^{\rho_{ij}+2}$ for all i greater than j, we may repeat the preceding reduction using $\gamma^2(M-m_{110}I)$. Otherwise at this writing no general reduction has been obtained in this case. The problem, however, is still under consideration.

THE UNIVERSITY OF WISCONSIN