# ON QUASI-COMMUTATIVE MATRICES*

BY

NEAL H. McCOY

1. **Introduction.** In quantum mechanics there appear infinite matrices $p$ and $q$ with the property that

$$pq - qp = c,$$

where $c$ is a scalar matrix. It is well known† that a relation of this type can not be satisfied by finite matrices. However, in the calculation of commutation formulas for polynomials in $p$ and $q$ no use is made of the fact that $c$ is a scalar but merely that it is commutative with both $p$ and $q$.‡ And there do exist pairs of finite matrices $x$, $y$ of the same order such that $xy - yx$ is not zero and is commutative with both $x$ and $y$. Such matrices will be called *quasi-commutative matrices* and either may be said to be quasi-commutative with the other.

In a certain sense the algebra of polynomials in a pair of quasi-commutative matrices is homeomorphic to the algebra arising in quantum mechanics. It is hoped to discuss such algebras in some detail in a later paper. In the present paper we shall make a brief study of quasi-commutative matrices whose elements belong to the complex number field.

The concept of quasi-commutativity is an extension or generalization of commutativity, and as would be expected, some of the results obtained are generalizations of known theorems concerning commutative matrices.

The problem of determining quasi-commutative matrices is that of finding matrices $x$, $y$, $z$ ($\neq 0$) which satisfy the equations

$$xy - yx = z, \qquad xz = zx, \qquad yz = zy.$$

If $z$ is an assigned matrix, there may or may not exist matrices $x$ and $y$ such that $(x, y, z)$ is a solution of these equations. In §3, we shall characterize those matrices $z$ for which these equations do admit a solution.

If $x$ is a given matrix, it is clear that there always exist matrices commutative with $x$ but it is not evident whether there exist matrices quasi-commutative with $x$. It will be shown (§4) that in most cases there do exist such matrices, and a necessary and sufficient condition for their existence is ob-

---

tained. The general form of a matrix quasi-commutative with $x$ is also given for the case in which $x$ has a single elementary divisor corresponding to each root.

In §5, we prove a theorem about the roots of any scalar polynomial $\psi(x, y)$ in quasi-commutative matrices $x$ and $y$. It is shown that if the roots of $x$ and $y$ are $\lambda_i$ and $\mu_j$ respectively, then the roots of $\psi(x, y)$ are all of the form $\psi(\lambda_i, \mu_j)$. This is merely an extension of a known theorem concerning commutative matrices.*

2. **Commutative matrices.** In this section we shall make some preliminary remarks and then mention a few known properties of commutative matrices which will be needed in later sections.

Let $x$ be a given matrix of order $n$, with the elementary divisors $(\lambda - \lambda_i)^{p_i}$ $(i = 1, 2, \cdots, r)$. Then by a proper choice of basis, $x$ may be expressed in the canonical form†

$$
(1) \qquad
\begin{Vmatrix}
X_1 & 0 & \cdots & 0 \\
0 & X_2 & \cdots & 0 \\
\cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdots & \cdot \\
0 & 0 & \cdots & X_r
\end{Vmatrix},
$$

where $X_i$ is the matrix of order $p_i$,‡

$$
(2) \qquad
\begin{Vmatrix}
\lambda_i & 1 & 0 & \cdots & 0 & 0 \\
0 & \lambda_i & 1 & \cdots & 0 & 0 \\
\cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\
0 & 0 & 0 & \cdots & \lambda_i & 1 \\
0 & 0 & 0 & \cdots & 0 & \lambda_i
\end{Vmatrix}
\qquad (i = 1, 2, \cdots, r).
$$

If now we write $X_i = \lambda_i e_i + \eta_i$, the matrices $e_i$ and $\eta_i$ may be called respectively the *partial idempotent element* and the *partial nilpotent element*§ of $x$ corresponding to the elementary divisor $(\lambda - \lambda_i)^{p_i}$.

---

* Frobenius, *Über vertauschbare Matrizen*, Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin, 1896, pp. 601–614.

† Cf. Bôcher, *Higher Algebra*, p. 289.

‡ We shall think of $X_i$ as a matrix of order $p_i$ or one of order $n$ at pleasure. It will be clear from the context as which it is being considered.

§ Cf. J. H. M. Wedderburn, *The automorphic transformation of a bilinear form*, Annals of Mathematics, (2), vol. 23 (1921), pp. 122–134. In general a matrix $a$ ($\neq 0$) such that $a^2 = a$ is said to be *idempotent*, a matrix $b$ with the property that $b^\nu = 0$, $b^{\nu-1} \neq 0$, is *nilpotent* of *index $\nu$*. The index of $\eta_i$ is clearly $p_i$.

Let $\lambda_i$ $(i = 1, 2, \cdots, l)$ be the *distinct* roots of $x$ and denote by $\phi_i$ the sum of all the partial idempotent elements of $x$ corresponding to elementary divisors involving the same root $\lambda_i$. The matrix $\phi_i$ is called the *principal idempotent element* of $x$ belonging to the root $\lambda_i$. It is of considerable importance that the principal elements of a matrix $x$ may be expressed as scalar polynomials in $x$.[*]

If, in (1), we combine those blocks $X_i$ corresponding to elementary divisors with the same root $\lambda_i$ into a single block $X^{(i)}$, we may express $x$ in the form

$$
(3) \qquad
\begin{Vmatrix}
X^{(1)} & 0 & \cdots & 0 \\
0 & X^{(2)} & \cdots & 0 \\
\cdot & \cdot & \cdots & \cdot \\
0 & 0 & \cdots & X^{(l)}
\end{Vmatrix}.
$$

The principal idempotent element of $x$ belonging to $\lambda_i$ is in this case a matrix of order $n$ with 1 in each element of the principal diagonal corresponding to the position of the diagonal of the submatrix $X^{(i)}$ in $x$, and zeros elsewhere.

We now discuss briefly the form of a matrix $y$ commutative with a matrix $x$ in the canonical form (1). If we set $e_i y e_j = Y_{ij}$, we may write $y$ in the form

$$
(4) \qquad y = \sum_{i,j} Y_{ij} =
\begin{Vmatrix}
Y_{11} & Y_{12} & \cdots & Y_{1r} \\
Y_{21} & Y_{22} & \cdots & Y_{2r} \\
\cdot & \cdot & \cdots & \cdot \\
Y_{r1} & Y_{r2} & \cdots & Y_{rr}
\end{Vmatrix},
$$

where $Y_{ij}$ is a rectangular matrix of $p_i$ rows and $p_j$ columns.[†] The equation $xy = yx$ is then equivalent to the set of equations,

$$
(5) \qquad\qquad X_i Y_{ij} = Y_{ij} X_j \qquad\qquad (i, j = 1, 2, \cdots, r).
$$

Let now $s$ and $t$ be fixed values of $i$ and $j$, and consider the single equation for $Y_{st}$,

$$
(6) \qquad\qquad X_s Y_{st} = Y_{st} X_t.
$$

The following facts are known concerning the equation (6).[‡] If $\lambda_s \neq \lambda_t$, the only solution is $Y_{st} = 0$. If $\lambda_s = \lambda_t$ and $p_s \geq p_t$, the general solution is of the form

---

[*] Wedderburn, loc. cit., p. 126.

[†] By actual definition, $Y_{ij}$ is a matrix of order $n$, but there will be no confusion as it has non-zero elements only in the rectangular block indicated in the array.

[‡] H. Kreis, *Contribution à la Théorie des Systèmes Linéaires*, Zurich Thesis, 1906. See also Hilton, *Homogeneous Linear Substitutions*, Oxford, 1914, chapter 5.

$$(7) \quad \begin{Vmatrix} a_1 & a_2 & a_3 & a_4 & \cdots & a_{p_t-1} & a_{p_t} \\ 0 & a_1 & a_2 & a_3 & \cdots & a_{p_t-2} & a_{p_t-1} \\ 0 & 0 & a_1 & a_2 & \cdots & a_{p_t-3} & a_{p_t-2} \\ & \cdots & & & & \cdots & \\ 0 & 0 & 0 & 0 & \cdots & a_1 & a_2 \\ 0 & 0 & 0 & 0 & \cdots & 0 & a_1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ & \cdots & & & & \cdots & \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{Vmatrix},$$

where $a_1, a_2, \cdots, a_{p_t}$ are arbitrary parameters. Similarly if $\lambda_s = \lambda_t$ and $p_s \leq p_t$, $Y_{st}$ takes the form

$$(8) \quad \begin{Vmatrix} 0 \cdots 0 & a_1 & a_2 & a_3 & \cdots & a_{p_s-1} & a_{p_s} \\ 0 \cdots 0 & 0 & a_1 & a_2 & \cdots & a_{p_s-2} & a_{p_s-1} \\ 0 \cdots 0 & 0 & 0 & a_1 & \cdots & a_{p_s-3} & a_{p_s-2} \\ & \cdots & & & & \cdots & \\ 0 \cdots 0 & 0 & 0 & 0 & \cdots & a_1 & a_2 \\ 0 \cdots 0 & 0 & 0 & 0 & \cdots & 0 & a_1 \end{Vmatrix}.$$

We may now write at once the general form of a matrix $y$ commutative with $x$, as we only need to solve each of the $r^2$ equations (5) for the $Y_{ij}$ by these rules, and substitute in (4).

The following application will be of importance in the sequel. Let $(x, y, z)$ be a solution of the equations

$$(9) \qquad xy - yx = z, \qquad xz = zx, \qquad yz = zy.$$

If $S$ is any non-singular matrix of the same order, then clearly $(SxS^{-1}, SyS^{-1}, SzS^{-1})$ is also a solution of these equations. Hence we may, without loss of generality, assume that $x$ is in canonical form (3).* Set $Y^{(ij)} = \phi_i y \phi_j$, $Z^{(ij)} = \phi_i z \phi_j$, where the $\phi_i$ are the principal idempotent elements of $x$. The first of equations (9) may be replaced by the set of equations

$$X^{(i)} Y^{(ij)} - Y^{(ij)} X^{(j)} = Z^{(ij)} \qquad (i, j = 1, 2, \cdots, l).$$

But $x$ and $z$ are commutative and hence, by the above results, $Z^{(ij)} = 0$ if $i \neq j$ and thus also $Y^{(ij)} = 0$ if $i \neq j$. Hence $y$ and $z$ must be of the types

---

* Cf. Bôcher, op. cit., p. 283.

$$y = \begin{Vmatrix} Y^{(11)} & 0 & \cdots & 0 \\ 0 & Y^{(22)} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & Y^{(ll)} \end{Vmatrix}, \quad z = \begin{Vmatrix} Z^{(11)} & 0 & \cdots & 0 \\ 0 & Z^{(22)} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & Z^{(ll)} \end{Vmatrix},$$

where the elements not in the diagonal blocks are necessarily zero. Thus the equations (9) are equivalent to the equations

$$X^{(i)}Y^{(ii)} - Y^{(ii)}X^{(i)} = Z^{(ii)}, \quad X^{(i)}Z^{(ii)} = Z^{(ii)}X^{(i)}, \quad Y^{(ii)}Z^{(ii)} = Z^{(ii)}Y^{(ii)}$$

$$(i = 1, 2, \cdots, l).$$

If $l > 1$, the solution $(x, y, z)$ of the equations (9) is therefore *reducible* and may be said to be the *direct sum* of the solutions $(X^{(i)}, Y^{(ii)}, Z^{(ii)})$ $(i = 1, 2, \cdots, l)$. We have thus shown that if $(x, y, z)$ is a solution of the equations (9) and $x$ (or $y$) has more than one distinct root, the solution is reducible.

Since $\phi_i \phi_j = \phi_j \phi_i = 0$ $(i \neq j)$ and $y = \sum_i \phi_i y \phi_i$, we see that $\phi_i y = y \phi_i = Y^{(ii)}$. This proves the

THEOREM 1. *If $x$ and $y$ are quasi-commutative matrices, the principal idempotent elements of either matrix are commutative with the other.*

This theorem is of course obvious for commutative matrices as the principal elements of a matrix are polynomials in that matrix.

3. **Characterization of the matrix $z$.** Let $(x, y, z)$ be a solution, in matrices of order $n$, of the equations (9). We shall assume, as we may, that $z$ is in canonical form. If $z$ has more than one distinct root, the solution $(x, y, z)$ is reducible as $x$ and $y$ are commutative with $z$. Hence let $z$ have the single root $\alpha$. From the first of equations (9), we see that

$$\text{trace } z = 0 = n\alpha^*,$$

that is, $\alpha = 0$, and $z$ is nilpotent.

Let the elementary divisors of $z$ be

$$\lambda^{n_1}, \lambda^{n_2}, \cdots, \lambda^{n_k} \qquad \left( n_1 \geqq n_2 \geqq \cdots \geqq n_k; \ \sum_{i=1}^{k} n_i = n \right).$$

We shall find it convenient to use the symbol $[n_1, n_2, \cdots, n_k]$, called the *characteristic* of $z$, to denote the degrees of these elementary divisors when arranged in the order indicated.† We shall now prove the following theorem:

---

\* The trace of a square matrix is defined as the sum of the elements in the principal diagonal.
† Cf. Bôcher, op. cit., p. 287.

THEOREM 2. *If $z$ is a given matrix of order $n$, necessary and sufficient conditions that there exist matrices $x$ and $y$ of order $n$ satisfying the equations*

$$(9) \qquad\qquad xy - yx = z, \qquad xz = zx, \qquad yz = zy$$

*are that $z$ be nilpotent and that it have a characteristic of the type $[n_1, n_2, \cdots, n_k]$, where $n_k = 1$ and $n_i - n_{i+1} = 0$ or $1$ $(i = 1, 2, \cdots, k-1)$.*

We first establish the necessity of these conditions. It has already been shown that $z$ must be nilpotent. Assume that $z$ does not have a characteristic of the form stated in the theorem but that there exist matrices $x$ and $y$ satisfying the equations (9). We may write the characteristic of $z$ in the more explicit form

$$[n_{11}, n_{12}, \cdots, n_{1k_1}, n_{21}, n_{22}, \cdots, n_{2k_2}, \cdots, n_{l1}, n_{l2}, \cdots,$$
$$n_{lk_l}, \cdots, n_{m1}, n_{m2}, \cdots, n_{mk_m}]$$

where $n_{i1} = n_{i2} = \cdots = n_{ik_i}$ $(i = 1, 2, \cdots, m)$; $n_{i1} - n_{i+1,1} = 1$ $(i = 1, 2, \cdots, l-1)$; but either $l = m$, $n_{m1} \geq 2$, or $n_{l1} - n_{l+1,1} > 1$. In either case $n_{l1} \geq 2$.

We shall now assume that $z$ is in canonical form. Let $f_{n_{ir}}$ be the partial idempotent element of $z$ corresponding to the elementary divisor $\lambda^{n_{ir}}$, and set $f_{n_{ir}} x f_{n_{js}} = X^{n_{ir}n_{js}}$, $f_{n_{ir}} y f_{n_{js}} = Y^{n_{ir}n_{js}}$, $f_{n_{ir}} z f_{n_{js}} = Z^{n_{ir}n_{js}}$. Then $Z^{n_{ir}n_{js}} = 0$ unless $i = j$, $r = s$, and $Z^{n_{ir}n_{ir}}$ is the partial nilpotent element of $z$ corresponding to the elementary divisor $\lambda^{n_{ir}}$.

Now $X^{n_{ir}n_{js}}$ is a rectangular matrix of $n_{ir}$ rows and $n_{js}$ columns, and since $x$ is commutative with $z$, it will be of the general type (7) if $i \leq j$ and of the type (8) if $i \geq j$. We shall find it convenient to denote the element of $X^{n_{ir}n_{js}}$ corresponding to $a_1$ in (7) or (8) by $a_{n_{js}}^{n_{ir}}$, that corresponding to $a_2$ by $b_{n_{js}}^{n_{ir}}$. Similarly these elements in $Y^{n_{ir}n_{js}}$ will be denoted by $\alpha_{n_{js}}^{n_{ir}}$ and $\beta_{n_{js}}^{n_{ir}}$ respectively. We shall also let $X_{ab}^{n_{ir}n_{js}}$ denote the element in the $a$th row and $b$th column of $X^{n_{ir}n_{js}}$, and similarly for $y$ and $z$.

Let $i \leq l$, $p \leq k_i$ be fixed positive integers. Then from the first of equations (9) we get the equation

$$(10) \qquad \sum_{j=1}^{m} \sum_{\mu=1}^{k_j} \left[ X^{n_{ip}n_{j\mu}} Y^{n_{j\mu}n_{ip}} - Y^{n_{ip}n_{j\mu}} X^{n_{j\mu}n_{ip}} \right] = Z^{n_{ip}n_{ip}},$$

for the part of $z$ in the $(n_{ip}, n_{ip})$ block. But $n_{ip} \geq 2$ and hence $Z_{12}^{n_{ip}n_{ip}} = 1$. Thus from (10) we have

$$(11) \qquad \sum_{j=1}^{m} \sum_{\mu=1}^{k_j} \sum_{\nu=1}^{n_{j\mu}} \left[ X_{1\nu}^{n_{ip}n_{j\mu}} Y_{\nu 2}^{n_{j\mu}n_{ip}} - Y_{1\nu}^{n_{ip}n_{j\mu}} X_{\nu 2}^{n_{j\mu}n_{ip}} \right] = 1.$$

It may be verified from the form of $x$ and $y$ that $X_{\nu 2}^{n_{j\mu}n_{ip}} = Y_{\nu 2}^{n_{j\mu}n_{ip}} = 0$ if $\nu > 2$,

and thus in the sum (11) we may limit $\nu$ to the values 1, 2. If $j<i-1$, $X_{1\nu}^{n_{ip}n_{j\mu}} = Y_{1\nu}^{n_{ip}n_{j\mu}} = 0$ $(\nu=1, 2)$; and if $j>i+1$, $X_{\nu 2}^{n_{j\mu}n_{ip}} = Y_{\nu 2}^{n_{j\mu}n_{ip}} = 0$ $(\nu=1, 2)$. Hence $j$ may be restricted to the values $i-1, i, i+1$, the first clearly being omitted if $i=1$. If $i=l$, there are, under our hypotheses, two cases to be considered according as $l=m$ or $l<m$. In the first case clearly $j$ can not take the value $i+1$. If $i=l<m$, we have assumed that $n_{l1}-n_{l+1,1}>1$, hence $X_{\nu 2}^{n_{l+1},\mu^{n_{ip}}} = Y_{\nu 2}^{n_{l+1},\mu^{n_{ip}}} = 0$ $(\nu=1, 2)$. Thus in either case if $i=l$, the index $j$ may be restricted to the two values $i-1, i$.

Using the notation introduced above, we get from (11)

(12)
$$\sum_{r=1}^{k_{i-1}} \left( a_{n_{i-1,r}\ n_{ip}}^{n_{ip}\ \alpha^{n_{i-1,r}}} - a_{n_{ip}\ n_{i-1,r}}^{n_{i-1,r}\alpha^{n_{ip}}} \right)$$
$$+ \sum_{s=1}^{k_i} \left( a_{n_{is}\ n_{ip}}^{n_{ip}\beta^{n_{is}}} + b_{n_{is}\ n_{ip}}^{n_{ip}\alpha^{n_{is}}} - b_{n_{ip}\ n_{is}}^{n_{is}\alpha^{n_{ip}}} - a_{n_{ip}\ n_{is}}^{n_{is}\beta^{n_{ip}}} \right)$$
$$+ \sum_{t=1}^{k_{i+1}} \left( a_{n_{i+1,t}\ n_{ip}}^{n_{ip}\ \alpha^{n_{i+1,t}}} - a_{n_{ip}\ n_{i+1,t}}^{n_{i+1,t}\ \alpha^{n_{ip}}} \right) = 1$$

$$(i = 1, 2, \cdots, l; p = 1, 2, \cdots, k_i),$$

with the understanding that the first sum is to be omitted when $i=1$, the last when $i=l$. Let us now sum the equations (12) with respect to the index $p$. The middle sum then vanishes. Hence if $l=1$, the sum of the left members of the equations (12) vanishes, which is clearly impossible. Suppose $l>1$. If we denote by $K(\mu, \nu)$ the double sum

$$\sum_{p=1}^{k_\mu} \sum_{r=1}^{k_\nu} \left( a_{n_{\nu r}\ n_{\mu p}}^{n_{\mu p}\alpha^{n_{\nu r}}} - a_{n_{\mu p}\ n_{\nu r}}^{n_{\nu r}\alpha^{n_{\mu p}}} \right),$$

the resulting $l$ equations may be written in order as follows:

(13)
$$K(1, 2) = k_1,$$
$$K(2, 1) + K(2, 3) = k_2,$$
$$\cdots \cdots \cdots \cdots \cdots \cdots \cdots$$
$$K(l - 1, l - 2) + K(l - 1, l) = k_{l-1},$$
$$K(l, l - 1) = k_l.$$

But $K(\mu, \nu) = -K(\nu, \mu)$, and hence if we add the equations (13) we get $0 = \sum_{i=1}^{l} k_i$, which is impossible as $l \geq 1$, $k_i \geq 1$. Therefore there can exist no matrices $x$ and $y$ satisfying the equations (9). We have thus established the necessity of the conditions stated in the theorem.

Let us now assume that $z$ has a characteristic of the type prescribed by

the theorem and exhibit a pair of matrices $x$ and $y$ satisfying the equations (9). There is no loss of generality in assuming that the elementary divisors of $z$ are not all linear, as in this case any pair of commutative matrices will satisfy the equations. Let the characteristic of $z$ be $[n_1, n_2, \cdots, n_t, n_{t+1}, \cdots, n_k]$, where $k \geq t+1$, $n_t = 2$, $n_{t+1} = \cdots = n_k = 1$, $n_i - n_{i+1} = 0$ or $1$ $(i = 1, 2, \cdots, t-1)$. If $f_{n_i}$ is the partial idempotent element of $z$ corresponding to the elementary divisor $\lambda^{n_i}$, we may again break any matrix $A$ of order $n$ into rectangular sub-matrices $A^{n_i n_j} = f_{n_i} A f_{n_j}$, and may call this the $(n_i, n_j)$ block of $A$.

Let $e_{pq}^{n_i n_j}$ denote the matrix of order $n$ for which the element in the $p$th row and $q$th column of the $(n_i, n_j)$ block is 1 and all other elements are zero. Then

$$e_{pq}^{n_i n_j} e_{rs}^{n_l n_m} = \begin{cases} e_{ps}^{n_i n_m}, & \text{if } j = l, q = r, \\ 0, & \text{if } j \neq l \text{ or } q \neq r. \end{cases}$$

With this notation we have

$$z = \sum_{i=1}^{t} Z^{n_i n_i} = \sum_{i=1}^{t} \sum_{j=1}^{n_i - 1} e_{j, j+1}^{n_i n_i}.$$

If now we set

$$A_i = \sum_{j=1}^{n_i - 1} e_{j, j+n_{i+1}-n_i+1}^{n_i n_{i+1}} \qquad (i = 1, 2, \cdots, t),$$

$$B_i = \sum_{j=1}^{n_{i+1}} e_{j, j+n_i-n_{i+1}}^{n_{i+1} n_i} \qquad (i = 1, 2, \cdots, t),$$

it may be verified that each $A_i$ and $B_i$ is commutative with $z$ and further that

$$A_i B_j = \begin{cases} Z^{n_i n_i}, & i = j, \\ 0, & i \neq j, \end{cases}$$

(14)

$$B_i A_j = \begin{cases} Z^{n_{i+1} n_{i+1}}, & i = j \neq t, \\ 0, & i \neq j \text{ or } i = j = t. \end{cases}$$

Hence if we set

$$x = \sum_{j=1}^{t} j A_j, \qquad y = \sum_{j=1}^{t} B_j,$$

then $x$ and $y$ are commutative with $z$ and by the relations (14) we find

$$xy - yx = \sum_{j=1}^{t} j Z^{n_j n_j} - \sum_{j=1}^{t-1} j Z^{n_{j+1} n_{j+1}} = \sum_{j=1}^{t} Z^{n_j n_j} = z.$$

Thus we have exhibited a solution of the equations (9), and the proof of Theorem 2 is completed.

The following corollary follows readily.

COROLLARY. *There exist no quasi-commutative matrices of order two.*

For if $n=2$, the only choice of a characteristic for $z$ which satisfies the conditions of Theorem 2 is $[1, 1]$, that is, $z=0$, and $x$ and $y$ are commutative.

**4. Matrices quasi-commutative with a given matrix.** Let $x$ be an assigned matrix of order $n$. We shall in this section find necessary and sufficient conditions that there exist a matrix $v$ quasi-commutative with $x$, and shall determine the general form of all such matrices for the special case in which $x$ has a single elementary divisor corresponding to each root.

We first consider the case in which $x$ has just one elementary divisor. There is no loss of generality in assuming that the root of $x$ is zero,* hence if $x$ is put in canonical form we have

$$(15) \qquad\qquad x = \sum_{i=1}^{n-1} e_{i,i+1}.\dagger$$

If $n=1$ or 2, there exist no matrices quasi-commutative with $x$. Hence assume $n \geq 3$. Now if $y$ and $z$ are matrices such that $(x, y, z)$ is a solution of equations (9), $z$ is commutative with $x$ and must also be nilpotent. Hence it must be of the form

$$(16) \qquad\qquad z = a_1 x + a_2 x^2 + \cdots + a_{n-1} x^{n-1},$$

where the $a_i$ are scalars.

If we write $y = \sum y_{ij} e_{ij}$, the equation $xy - yx = z$ becomes

$$\sum_i e_{i,i+1} \sum_{j,k} y_{jk} e_{jk} - \sum_{j,k} y_{jk} e_{jk} \sum_i e_{i,i+1}$$

$$= \sum_r y_{r+1,1} e_{r1} + \sum_{r,s} (y_{r+1,s+1} - y_{rs}) e_{r,s+1}$$

$$= a_1 \sum_i e_{i,i+1} + a_2 \sum_i e_{i,i+2} + \cdots + a_{n-1} e_{1n}.$$

From this it follows that $y_{21} = y_{31} = \cdots = y_{n1} = 0$. Thus if $i > j$, $y_{i+1,j+1} = y_{ij} = 0$. If $i \leq j$, $y_{i+1,j+1} = y_{ij} + a_{j-i+1}$, from which we see that

$$y_{ij} = y_{1,j-i+1} + (i-1)a_{j-i+1} \qquad\qquad (i \leq j).$$

---

* For if $x$ and $y$ are quasi-commutative, so are $x-\lambda$ and $y$, where $\lambda$ is any scalar matrix.

† The matrix unit $e_{pq}$ is a matrix with 1 at the intersection of the $p$th row and $q$th column, and zeros elsewhere. The rule for multiplying these units is $e_{pq} e_{rs} = 0 \ (q \neq r)$, $e_{pq} e_{qs} = e_{ps}$. It will be convenient to let $e_{pq} = 0$ if either $p$ or $q$ is greater than $n$ or less than 1.

If then we set $y_{1j} = y_j$, $y$ is of the form*

$$(17) \quad \begin{Vmatrix} y_1 & y_2 & y_3 & y_4 & \cdots y_n \\ 0 & y_1 + a_1 & y_2 + a_2 & y_3 + a_3 & \cdots y_{n-1} + a_{n-1} \\ 0 & 0 & y_1 + 2a_1 & y_2 + 2a_2 & \cdots y_{n-2} + 2a_{n-2} \\ \cdot & \cdot & \cdot & \cdot & \cdots \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdots \cdot \\ 0 & 0 & 0 & 0 & \cdots y_1 + (n-1)a_1 \end{Vmatrix}.$$

If $x$ is the matrix (15), then $y$ and $z$ defined by (17) and (16), where $y_1, y_2, \cdots, y_n, a_1, a_2, \cdots, a_{n-1}$ are arbitrary parameters, are the most general matrices satisfying the first two of equations (9). We wish now to make necessary restrictions on $y$ and $z$ in order that they may be commutative. Let $y$ given by (17) be expressed in the form $y = Y_1 + Y_2$, where $Y_1 = y_1 + y_2 x + y_3 x^2 + \cdots + y_n x^{n-1}$. Then $y$ will be commutative with $z$ if and only if $Y_2$ and $z$ are commutative. Since the elements of $Y_2$ depend only upon the parameters $a_i$, the only restriction will be on these parameters, that is, on the matrix $z$.

In (16) suppose $a_k \neq 0$ ($1 \leq k \leq n-1$) but $a_i = 0$ ($i < k$). If positive integers $p$, $q$ are defined by

$$n = pk + q, \qquad 0 \leq q < k,$$

then it is known† that $z$ has $q$ elementary divisors of degree $p+1$ and $k-q$ of degree $p$. Hence by Theorem 2, there can exist matrices $x$ and $y$ satisfying equations (9) only if $p = 1$, that is, $k \geq (n+1)/2$. Let $k$ be the smallest integer satisfying this inequality. Then the elements of $z$ are all zero except in the block of $n-k$ rows and $n-k$ columns in the upper right hand corner. Similarly $Y_2$ has non-zero elements only in the block of $n-k+1$ rows and $n-k$ columns in the upper right hand corner. Hence $zY_2 = Y_2 z = 0$, and we have thus established the following theorem:

THEOREM 3. *If $x$ is the matrix (15) of order $n \geq 3$ and $k$ is the smallest integer not less than $(n+1)/2$, then the general form of a matrix $y$ quasi-commutative with $x$ is given by (17) where $a_i = 0$ ($i = 1, 2, \cdots, k-1$), and $a_k, a_{k+1}, \cdots, a_{n-1}, y_1, y_2, \cdots, y_n$ are arbitrary parameters.*

---

* Cf. R. Weitzenböck, *Über die Matrixgleichung $Ax + xB = C$*, Akademie van Wetenschappen te Amsterdam, Proceedings, vol. 35 (1932), pp. 54–59. The form (17) obtained for $y$ is a special case of a formula given by Weitzenböck.

† H. Kreis, op. cit., p. 47. See also D. E. Rutherford, *On the canonical form of a rational integral function of a matrix*, Proceedings of the Edinburgh Mathematical Society, (2), vol. 3 (1932), pp. 135–143.

This theorem is sufficient to give the form of $y$ even if $x$ has more than one root, provided it has only one elementary divisor corresponding to each root For by the results of §2, the general form of such a $y$ is the direct sum of matrices of the type prescribed by this theorem. If $x$ is not of this simple type it seems to be difficult to give the general form of a matrix quasi-commutative with $x$ and we shall now limit ourselves to a consideration of the conditions under which such matrices exist.

Let $x$ be a matrix of order $n$ and let us separate the elementary divisors of $x$ into two sets $A_1$ and $A_2$. By a proper choice of basis $x$ may then be expressed as the direct sum of two matrices, thus

$$x = \left\|\begin{array}{cc} X_1 & 0 \\ 0 & X_2 \end{array}\right\|$$

where the matrix $X_1$ has the elementary divisors $A_1$ and $X_2$ the elementary divisors $A_2$. If now $Y_1$ is a matrix quasi-commutative with $X_1$, then

$$y = \left\|\begin{array}{cc} Y_1 & 0 \\ 0 & 0 \end{array}\right\|$$

will be quasi-commutative with $x$. Hence there will exist a matrix quasi-commutative with a given matrix $x$ if there exists a matrix quasi-commutative with a matrix the set of whose elementary divisors is a subset of the elementary divisors of $x$. We may now prove the following lemma:

LEMMA. *Let $x$ be a matrix of order $n$ with a single root $\lambda_1$. A necessary and sufficient condition that there exist a matrix $y$ quasi-commutative with $x$ is that $n$ be greater than two and the elementary divisors of $x$ be not all linear.*

The necessity of these conditions follows from the Corollary to Theorem 2 and from the fact that if the elementary divisors of $x$ are all linear, $x$ is a scalar matrix and is thus commutative with all matrices of order $n$.

We now prove that these conditions are sufficient. If $n>2$ and the elementary divisors of $x$ are not all linear then (i) $x$ has an elementary divisor of degree $\geq 3$, or (ii) $x$ has at least two elementary divisors of degree 2, or (iii) $x$ has one elementary divisor of degree 2 and at least one of degree 1. In the first case, Theorem 3 establishes the existence of a matrix $y$ quasi-commutative with $x$. The existence of such a matrix $y$ in the cases (ii) and (iii) is shown by the two examples

$$x = \lambda_1 + e_{12} + e_{34}, \qquad y = e_{24},$$

and

$$x = \lambda_1 + e_{12}, \qquad y = e_{23},$$

respectively.

The following theorem follows readily.

**THEOREM 4.** *A necessary and sufficient condition that there exist a matrix $y$ quasi-commutative with a given matrix $x$ is that for some root $\lambda_i$ of $x$ the sum of the degrees of the elementary divisors of $x$ associated with $\lambda_i$ be greater than two, and at least one of these elementary divisors be not linear.*

**5. Roots of a polynomial in quasi-commutative matrices.** In this section we shall prove the following theorem which is well known for the case of commutative matrices.*

**THEOREM 5.** *If $x$ and $y$ are quasi-commutative matrices with principal idempotent elements $R_i$ and $S_j$ $(i, j = 1, 2, \cdots)$ and corresponding roots $\lambda_i$ and $\mu_j$ respectively, then the roots of any scalar polynomial $\psi(x, y)$ in $x$ and $y$ are $\psi(\lambda_i, \mu_j)$, where $i$ and $j$ take only those values for which $R_iS_j \neq 0$.†*

We first prove two lemmas.

**LEMMA 1.** *If $\psi(x, y)$ is any scalar polynomial in the quasi-commutative matrices $x$ and $y$, and $z = xy - yx$, then*

$$\psi(x, y) = \psi_1(x, y) + z\,\psi_2(x, y, z),$$

*where $\psi_1$ is of the form $\sum a_{ij}x^iy^j$, the $a_{ij}$ being scalars.*

It is clear that by repeated substitutions of the type $yx = xy - z$, $\psi(x, y)$ can be reduced to the form $\sum a_{ijk}z^ix^jy^k$. Hence we only need to set $\psi_1 = \sum_{j,k} a_{0jk}x^jy^k$, $z\psi_2 = \psi - \psi_1$.

**LEMMA 2.** *If $x$ and $y$ are quasi-commutative matrices and $z = xy - yx$, then*

$$x^{m_1}y^{n_1}x^{m_2}y^{n_2} \cdots x^{m_k}y^{n_k} = \sum_{l=0} a_l z^l x^{P-l} y^{Q-l},$$

*where the $m_i$ and $n_i$ are any positive integers and $P = \sum_{i=1}^k m_i$, $Q = \sum_{i=1}^k n_i$.*

The sum of the exponents of $x$ which appear explicitly in a given term may be called the *degree* of the term in $x$, and similarly for $y$. For example, the term $z^4x^2y^4x^2y$ is of degree 4 in $x$ and 5 in $y$. If we think of $z$ as being replaced by $xy - yx$, then the total degree in $x$ and $y$ would be 17. For convenience, let us call this the *weight* of the term.

As above, we may express $x^{m_1}y^{n_1}x^{m_2}y^{n_2} \cdots x^{m_k}y^{n_k}$ in the form

(18)                                    $\sum b_{ijk}z^ix^jy^k,$

* Frobenius, loc. cit. The method of this section is a modification of that used by Wedderburn, loc. cit., p. 127.

† It is not necessary that $x$ and $y$ be quasi-commutative in order that the roots of $\psi(x, y)$ shall be of the form $\psi(\lambda_i, \mu_j)$. Cf. G. S. Bruton, *Certain aspects of the theory of equations for a pair of matrices*, and M. H. Ingraham, *A study of certain related pairs of square matrices.* Abstracts of these papers appear in the Bulletin of the American Mathematical Society, vol. 38 (1932), p. 633.

by repeated substitutions of $xy-z$ for $yx$. Each time this substitution is made in a term we get two terms; in one the degree in $x$ and in $y$ is the same as before the substitution, in the other the degree of each has been reduced by one. The weight is invariant under a substitution of this form. Hence each term of (18) is of the type $b_{ijk}z^ix^jy^k$, where $j=P-l$, $k=Q-l$, $2i+j+k=P+Q$, that is, $i=l$.

We may now proceed with the proof of the theorem. The principal elements $R_i$ and $S_j$ are polynomials in $x$ and $y$ respectively, and by Theorem 1 are commutative with each other and with $x$, $y$ and $z=xy-yx$. If we set $T_{ij}=R_iS_j$, then $T_{ij}T_{pq}=0$ if $i\neq p$ or $j\neq q$, $T_{ij}^2=T_{ij}$, $\sum T_{ij}=1$. Further those $T_{ij}$ which are not zero are linearly independent, for if $\sum a_{ij}T_{ij}=0$, then $R_p\sum a_{ij}T_{ij}R_q=a_{pq}T_{pq}=0$ and thus $T_{pq}=0$ unless $a_{pq}=0$.

Let us now write

$$x = \sum_{i,j}[\lambda_i - (x - \lambda_i)]T_{ij}, \qquad y = \sum_{i,j}[\mu_j - (y - \mu_j)]T_{ij}.$$

The matrices $(x-\lambda_i)T_{ij}$ and $(y-\mu_j)T_{ij}$ are then nilpotent. If $\psi(x, y)$ is a polynomial in $x$ and $y$, we have, by the first lemma,

$$\psi(x, y) = \psi_1(x, y) + z\psi_2(x, y, z),$$

where $\psi_1(x, y) = \sum a_{ij}x^iy^j$. Since no interchange of order is necessary, we may write, as in the commutative case,

$$\psi_1(x, y) = \sum_{r,s}\psi_{rs}^{ij}(x - \lambda_i)^r(y - \mu_j)^s,$$

where the $\psi_{rs}^{ij}$ are scalar constants. Thus

$$\psi(x, y) = \sum_{i,j}\left[\psi_1(\lambda_i, \mu_j)T_{ij} + \sum_{r,s}\psi_{rs}^{ij}(x - \lambda_i)^r(y - \mu_j)^sT_{ij}\right] + z\psi_2,$$

with the understanding that $r$ and $s$ are not to be zero simultaneously.

Let us set

$$A = \sum_{i,j}\psi_1(\lambda_i, \mu_j)T_{ij}, \quad B = \sum_{i,j}\sum_{r,s}[\psi_{rs}^{ij}(x - \lambda_i)^r(y - \mu_j)^sT_{ij}] + z\psi_2(x, y, z).$$

Then $\psi(x, y) = A + B$, and $A$ and $B$ are commutative. It will be shown below that $B$ is nilpotent. Hence the roots of $\psi(x, y)$ are the roots of $A$ and these are of the form $\psi_1(\lambda_i, \mu_j)$ where $T_{ij}\neq 0$.[*] But $\psi_1(\lambda_i, \mu_j) = \psi(\lambda_i, \mu_j)$ and the theorem is established. We shall now complete this proof by showing that $B$ is nilpotent.

---

[*] Wedderburn, loc. cit.

Let

$$B_1 = \sum_{i,j} \sum_{r,s} \psi_{rs}^{ij}(x - \lambda_i)^r(y - \mu_j)^s T_{ij} = \sum_{i,j} A_{ij}T_{ij}.$$

Since $T_{ij}$ is commutative with each $A_{ij}$, we see that

$$B_1^k = \sum_{i,j} A_{ij}^k T_{ij}.$$

We can thus show that $B_1$ is nilpotent by showing that each $A_{ij}T_{ij}$ is nilpotent. Let $x_1 = x - \lambda_i$, $y_1 = y - \mu_j$; then $x_1$ and $y_1$ are quasi-commutative and $x_1 y_1 - y_1 x_1 = z$. Let $\rho_1$ be the index of the nilpotent matrix $x_1 R_i$, $\rho_2$ that of $y_1 S_j$, and $t$ that of $z$. Let $N = t + \max(\rho_1, \rho_2)$ and consider

$$(A_{ij}T_{ij})^{2N} = \left[ \sum_{r,s} \psi_{rs}^{ij} x_1^r y_1^s \right]^{2N} T_{ij}.$$

The right hand side will, when expanded, consist of a sum of terms of the general type $a x_1^{m_1} y_1^{n_1} x_1^{m_2} y_1^{n_2} \cdots x_1^{m_k} y_1^{n_k} T_{ij}$. But by Lemma 2, this may be put in the form

$$\sum_l a_l z^l x_1^{P-l} y_1^{Q-l} T_{ij} = \sum_l a_l z^l (x_1 R_i)^{P-l}(y_1 S_j)^{Q-l} T_{ij}.$$

Here $P = \sum m_i$, $Q = \sum n_i$, and as $r$ and $s$ are not both zero, either $P \geqq N$ or $Q \geqq N$. Now the term in this last sum of which $a_l$ is the coefficient is zero provided $l \geqq t$ or $P - l \geqq \rho_1$ or $Q - l \geqq \rho_2$. Suppose, for example, that $P \geqq N$. Then the term containing $a_l$ is zero provided $l \geqq t$ or $l \leqq P - \rho_1$. But $P - \rho_1 \geqq N - \rho_1 \geqq t$, and thus all $l$ are included. Hence $(A_{ij}T_{ij})^{2N} = 0$ and $B_1$ is nilpotent of index, say, $r$. It follows that

$$B^{rt} = [(B_1 + z\psi_2)^r]^t = [B_1^r + z(\ )]^t = 0.$$

Hence $B$ is nilpotent and the proof of the theorem is completed.

It may be noted that the fact that $z = xy - yx$ is nilpotent is a special case of this theorem. For if we choose $\psi(x, y) = xy - yx$, the roots of $\psi$ must be of the form $\lambda_i \mu_j - \mu_j \lambda_i = 0$.

SMITH COLLEGE,
NORTHAMPTON, MASS.