THE EQUATIONS $AX - YB = C$ AND $AX - XB = C$ IN MATRICES

WILLIAM E. ROTH

Two theorems will be established.

**Theorem I.** The necessary and sufficient condition that the equation

$$AX - YB = C,$$

where $A$, $B$, and $C$ are $m \times r$, $s \times n$, and $m \times n$ matrices respectively with elements in a field $F$, have a solution $X$, $Y$ of order $r \times n$ and $m \times s$ respectively and with elements in $F$ is that the matrices

$$
\begin{pmatrix}
A, & C \\
0, & B
\end{pmatrix}
\quad \text{and} \quad
\begin{pmatrix}
A, & 0 \\
0, & B
\end{pmatrix}
$$

be equivalent.

**Theorem II.** The necessary and sufficient condition that the equation

$$(1) \quad AX - XB = C,$$

where $A$, $B$, and $C$ are square matrices of order $n$ with elements in $F$, have a solution $X$ with elements in $F$ is that the matrices

$$
\begin{pmatrix}
A, & C \\
0, & B
\end{pmatrix}
\quad \text{and} \quad
\begin{pmatrix}
A, & 0 \\
0, & B
\end{pmatrix}
$$

be similar.

To do so we shall prove the lemma:

**The necessary and sufficient condition that the equation**

$$(3) \quad AX - YB = C,$$

where $A$, $B$, and $C$ are $n \times n$ matrices with elements in the polynomial domain $F[x]$ of the field $F$, have a solution $X$, $Y$ of order $n$ with elements in $F[x]$ is that the matrices

$$
\begin{pmatrix}
A, & C \\
0, & B
\end{pmatrix}
\quad \text{and} \quad
\begin{pmatrix}
A, & 0 \\
0, & B
\end{pmatrix}
$$

be equivalent.

The lemma holds as well if the matrices are rectangular. We shall prove it as stated.

Presented to the Society, September 2, 1949; received by the editors August 3, 1951.
The condition is necessary for, if a solution of (3) exists, we have

\[
\begin{pmatrix}
I, Y,\ A,\ C
O, I,\ O,\ B
\end{pmatrix}
\begin{pmatrix}
I, -X
O, I
\end{pmatrix}
= \begin{pmatrix}
A, C - AX + YB
O, B
\end{pmatrix}
= \begin{pmatrix}
A, O
O, B
\end{pmatrix},
\]

where \( I \) is a unit matrix; consequently matrices (4) are equivalent. Conversely, if matrices (4) are equivalent, we shall show that equation (3) has a solution.

Nonsingular matrices \( P, Q \) and \( R, S \) with elements in \( F[x] \) exist such that

\[
PAQ = A' = a_1 + a_2 + \cdots + a_\alpha + 0 + \cdots + 0,
\]

\[
RBS = B' = b_1 + b_2 + \cdots + b_\beta + 0 + \cdots + 0,
\]

where \( a_i, i = 1, 2, \cdots, \alpha \), and \( b_j, j = 1, 2, \cdots, \beta \), are the invariant factors of \( A \) and \( B \) respectively. Consequently

\[
\begin{pmatrix}
P, O \\
O, R
\end{pmatrix}
\begin{pmatrix}
A, O \\
O, B
\end{pmatrix}
\begin{pmatrix}
Q, O \\
O, T
\end{pmatrix}
= \begin{pmatrix}
A', O \\
O, B'
\end{pmatrix}
= M,
\]

\[
\begin{pmatrix}
P, O \\
O, R
\end{pmatrix}
\begin{pmatrix}
A, C \\
O, B
\end{pmatrix}
\begin{pmatrix}
Q, O \\
O, S
\end{pmatrix}
= \begin{pmatrix}
A', C' \\
O, B'
\end{pmatrix}
= N,
\]

where \( PCS = C' = (c_{ij}), i, j = 1, 2, \cdots, n \). The matrices \( M \) and \( N \) are equivalent. We shall show that matrices \( U \) and \( V \) with elements in \( F[x] \) exist such that

\[
A'U - VB' = C';
\]

in other words that elements \( u_{ij} \) and \( v_{ij} \) exist in \( F[x] \) and satisfy the equations

\[
a_iu_{ij} - v_{ij}b_j = c_{ij}, \quad i, j = 1, 2, \cdots, n.
\]

To do so first consider the elements \( c_{ij}, i = 1, 2, \cdots, \alpha; j = 1, 2, \cdots, \beta \). Equations (6) are evidently satisfied for every \( c_{ij} \) where \( a_i \) and \( b_j \) are relatively prime. If \( a_i \) and \( b_j \) have the greatest common factor \( g_{ij} \), we shall now show that \( c_{ij} \) is its multiple and as a result (6) is again valid in \( F[x] \). Let \( g \) be any factor irreducible in \( F[x] \) which is common to \( a_\alpha \) and \( b_\beta \), then the invariant factors of \( A \) and \( B \) are

\[
a_i = g^{r_1}a'_i, \quad r_1 \leq r_2 \leq \cdots \leq r_\alpha,
\]

\[
b_j = g^{s_1}b'_j, \quad s_1 \leq s_2 \leq \cdots \leq s_\beta
\]

respectively, where \( a'_i \) and \( b'_j \) are polynomials in \( F[x] \) and are prime.
to $g$. Consequently the invariant factors of $M$ (of $N$) are
\[ m_k = g^{t_k}m_k', \quad t_1 \leq t_2 \leq \cdots \leq t_{\alpha+\beta}, \]
where $t_k$, $k=1, 2, \ldots, \alpha+\beta$, is a permutation of the exponents $t_i, i=1, 2, \ldots, \alpha$, and $s_j, j=1, 2, \ldots, \beta$, in non-descending order, and where $m_k'$ is prime to $g$.

The $(i+j-1)$th determinantal divisor of $M$ (and of $N$) is $M_{i+j-1} = \prod_{t=1}^{i+j-1} g^{t_k}m_k'$. It contains the factor $g^{t_k}$ of $a_i$ or the factor $g^{s_j}$ of $b_j$ for in forming the sequence of integers $t_k, k=1, 2, \ldots, i+j-1$, which we shall designate by \{t\}, either all $r_1, r_2, \ldots, r_i$ or all $s_1, s_2, \ldots, s_j$ must be taken in order to get its $i+j-1$ terms. On the other hand both $g^{t_k}$ and $g^{s_j}$ cannot be factors of $M_{i+j-1}$ for both $r_i$ and $s_j$ cannot occur in the sequence \{t\} which has only $i+j-1$ terms. Here we assume that neither $r_i$ nor $s_j$ is zero since the case where $a_i$ and $b_j$ are relatively prime was taken up above. Consequently $t_{ij}$, the lesser of $r_i$ and $s_j$, is in \{t\} and $M_{i+j-1}$ must have the factor $g^{t_{ij}}$. Now let $M'$ and $N'$ be the matrices obtained from $M$ and $N$ respectively by deleting their $i$th and $(n+j)$th rows and columns, and let $N_{i+j-2}$ be their common $(i+j-2)$th determinantal divisor. The latter will not have the factor $g^{t_{ij}}$ because the rows and columns containing $a_i$ and $b_j$ were deleted in forming $M'$ and $N'$ but will contain as factors all the remaining powers of $g$ that occur in $M_{i+j-1}$. Now $c_{ij}N_{i+j-2}$ is a minor of order $i+j-1$ of $N$ and as a consequence is a multiple of $\prod_{t=1}^{i+j-1} g^{t_{ij}}$, a factor of $M_{i+j-1}$. Therefore $c_{ij}$ must be a multiple of $g^{t_{ij}}$, the highest power of $g$ which is common to $a_i$ and $b_j$. Since $g$ is any factor irreducible in $F[x]$ and common to $a_i$ and $b_j$ and since $t_{ij}$ is the greatest power of $g$ common to $a_i$ and $b_j$ it follows that the greatest common factor, $g^{t_{ij}}$, of $a_i$ and $b_j$ is a divisor of $c_{ij}$. Consequently equations (6) are satisfied by elements $u_{ij}$ and $v_{ij}$ in $F[x]$, where $1 \leq i \leq \alpha$ and $1 \leq j \leq \beta$.

Next regard the block of elements $c_{ij}, 1 \leq i \leq \alpha, \beta < j \leq n$. Here $c_{ij} \prod_{k=1}^{i-1} a_k \prod_{k=i}^{i+j-1} a_k \prod_{k=1}^{\beta-1} b_k$ is a minor of order $\alpha+\beta$ of $N$; it must be a multiple of $\prod_{k=1}^{\alpha-1} a_k \prod_{k=1}^{\beta-1} b_k$, the $(\alpha+\beta)$th determinantal divisor of $M$ (of $N$). That is $c_{ij}$ must be a multiple of $a_i$, hence equations (6) are satisfied by $u_{ij} = c_{ij}/a_i$ and by $v_{ij}$, an arbitrary polynomial in $F[x]$, because $b_j$ may be regarded as identically equal to zero where $\beta < j \leq n$. Similarly equations (6) are valid for the block of elements $c_{ij}, \alpha < i \leq n, 1 \leq j \leq \beta$.

Finally if any $c_{ij}, \alpha < i \leq n$ and $\beta < j \leq n$, were not identically zero, $N$ would have the nonzero minor $c_{ij} \prod_{k=1}^{\alpha} a_k \prod_{k=1}^{\beta} b_k$ of order $\alpha+\beta+1$, which is impossible. Hence $c_{ij}$ is identically zero and arbitrary elements $u_{ij}$ and $v_{ij}$ in $F[x]$ will satisfy (6) for both $a_i$ and $b_j$ may be
regarded as identically zero in case \( \alpha < i \leq n \), and \( \beta < j \leq n \). Equations (6) are therefore valid for \( i, j = 1, 2, \ldots, n \).

Hence \( U = (u_{ij}) \) and \( V = (v_{ij}) \) are matrices of order \( n \) with elements in \( F[x] \) which satisfy equation (5). Now \( A' = PAR, B' = RBS, \) and \( C' = PCS \) and since \( P \) and \( S \) are nonsingular, we find from (5) that \( X = QUS^{-1}, Y = P^{-1}VR \) with elements in \( F[x] \) exist such that equation (3) is satisfied and the lemma is proved.\(^1\)

Theorem I follows as an immediate consequence of this lemma, for the latter is valid when the matrices are rectangular.

The necessary condition for Theorem II is proved as for the lemma. The proof that the condition is also sufficient follows.

Since the matrices (2) are similar,

\[
\begin{pmatrix}
A - xI, & C \\
0, & B - xI
\end{pmatrix}
\text{ and }
\begin{pmatrix}
A - xI, & 0 \\
0, & B - xI
\end{pmatrix}
\]

are equivalent and have elements in \( F[x] \); hence according to the lemma, matrices \( X \) and \( Y \) with elements in \( F[x] \) exist such that

\[(7) \quad (A - xI)X - Y(B - xI) = C.
\]

Let

\[
X = X_0 + X_1x + \cdots + X_px^p,
\]

\[
Y = Y_0 + Y_1x + \cdots + Y_qx^q,
\]

where \( X_i, i = 0, 1, \ldots, p \), and \( Y_j, j = 0, 1, \ldots, q \), are \( n \times n \) matrices with elements in \( F \) and where neither \( p \) nor \( q \) exceeds \( n^2 - 1 \). Evidently \( q = p \) because \( A, B, \) and \( C \) are independent of the indeterminate \( x \). Upon equating the coefficients of like powers of \( x \) in (7), we obtain the following \( p + 2 \) equations

\[
\begin{align*}
AX_0 - Y_0B &= C, \\
AX_1 - X_0 - Y_1B + Y_0 &= 0, \\
AX_2 - X_1 - Y_2B + Y_1 &= 0, \\
\cdots & \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \\
AX_p - X_{p-1} - Y_pB + Y_{p-1} &= 0, \\
- X_p + Y_p &= 0.
\end{align*}
\]

Multiply their members on the right by \( I, B, \ldots, B^{p+1} \) respectively

---

\( ^1 \) The writer is indebted to Professor MacDuffee for a simple proof of the sufficient condition of Theorem I based on the normal form \( 1+1+\cdots+1+0+\cdots+0 \). The proof of our lemma is a generalization thereof.
and add the members of the resulting equations; we then have
\[ A(X_0 + X_1 B + \cdots + X_p B^p) - (X_0 + X_1 B + \cdots + X_p B^p)B = C. \]
Hence \( X_0 + X_1 B + \cdots + X_p B^p \) with elements in \( F \) is a solution of (1) where \( p \) is obviously less than \( n \). Similarly \( Y_0 + A Y_1 + \cdots + A^p Y_p \) is also a solution. The theorem is proved.²

University of Tulsa

Professor Jacobson, Editor of these Proceedings, in a letter to me dated June 15, 1951, says in part: The referee “noted that the second theorem can be formulated in terms of groups with operators from a principal ideal domain.” And he expressed the opinion that a formulation in finite abelian groups may serve the same purpose. Subsequently, the writer has read, perhaps with insufficient understanding, Chatelet’s dissertation, University of Paris, 1911, and has not found therein the group theoretic formulation noted by the referee as resulting from Theorem II above, and hopes the referee’s observation is new.

---

SYMMETRY OF BANACH ALGEBRAS¹

IRVING KAPLANSKY

1. Introduction. Gelfand and Neumark [2]² raised the question as to whether a \( B^* \)-algebra is necessarily symmetric.³ This question remains open. Also open, of course, is the more general question propounded by the author in [3]: is a \( C \)-symmetric algebra symmetric?

The latter question is of some independent interest, since an affirmative answer would ease the task of proving that a specific algebra (such as a group algebra) is symmetric. Consequently we shall devote this note to pushing forward the second question to the same point that the first has reached. We shall prove the following generalization of [4, Lemma 7.2].

**Theorem.** Let \( A \) be a \( C \)-symmetric algebra with continuous involution. Then for no element \( x \) in \( A \) can \( -x^*x \) be a nonzero idempotent.

The proof rests largely on a purely algebraic result (Lemma 5) which appears to be new even in the finite-dimensional case.

2. Definitions. Our Banach algebras will admit complex scalars,