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*

$$AX - XB = C, \tag{1}$$

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

$$\begin{pmatrix} A, & C \\ 0, & B \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} A, & 0 \\ 0, & B \end{pmatrix} \tag{2}$$

be similar.

To do so we shall prove the lemma:
The necessary and sufficient condition that the equation

$$\mathbf{A}\mathbf{X} - \mathbf{Y}\mathbf{B} = \mathbf{C}, \tag{3}$$

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

$$\begin{pmatrix} \mathbf{A}, & \mathbf{C} \\ \mathbf{O}, & \mathbf{B} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \mathbf{A}, & \mathbf{O} \\ \mathbf{O}, & \mathbf{B} \end{pmatrix} \tag{4}$$

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

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

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

$$PAQ = A' = a_1 \dotplus a_2 \dotplus \cdots \dotplus a_\alpha \dotplus 0 \dotplus \cdots \dotplus 0,$$
$$RBS = B' = b_1 \dotplus b_2 \dotplus \cdots \dotplus b_\beta \dotplus 0 \dotplus \cdots \dotplus 0,$$

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

$$A'U - VB' = C'; \tag{5}$$

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

$$a_i u_{ij} - v_{ij} b_j = c_{ij}, \qquad i, j = 1, 2, \cdots, n. \tag{6}$$

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_α and b_β, then the invariant factors of A and B are

$$a_i = g^{r_i} a_i', \qquad r_1 \leqq r_2 \leqq \cdots \leqq r_\alpha,$$
$$b_j = g^{s_j} b_j', \qquad s_1 \leqq s_2 \leqq \cdots \leqq s_\beta$$

respectively, where a_i' and b_j' are polynomials in $F[x]$ and are prime to g. Consequently the invariant factors of $\boldsymbol{M}$ (of $\boldsymbol{N}$) are

$$m_k = g^{t_k} m_k', \qquad t_1 \leqq t_2 \leqq \cdots \leqq t_{\alpha+\beta},$$

where t_k, $k=1, 2, \cdots, \alpha+\beta$, is a permutation of the exponents r_i, $i=1, 2, \cdots, \alpha$, and s_j, $j=1, 2, \cdots, \beta$, in nondescending order, and where m_k' is prime to g.

The $(i+j-1)$th determinantal divisor of $\boldsymbol{M}$ (and of $\boldsymbol{N}$) is $\boldsymbol{M}_{i+j-1} = \prod_{k=1}^{i+j-1} g^{t_k} m_k'$. It contains the factor g^{r_i} of a_i or the factor g^{s_j} of b_j for in forming the sequence of integers t_k, $k=1, 2, \cdots, i+j-1$, *which we shall designate by* $\{t\}$, either all $r_1, r_2, \cdots, r_i$ or all $s_1, s_2, \cdots, s_j$ must be taken in order to get its $i+j-1$ terms. On the other hand both g^{r_i} and g^{s_j} cannot be factors of $\boldsymbol{M}_{i+j-1}$ 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 $\boldsymbol{M}_{i+j-1}$ must have the factor $g^{t_{ij}}$. Now let $\boldsymbol{M}'$ and $\boldsymbol{N}'$ be the matrices obtained from $\boldsymbol{M}$ and $\boldsymbol{N}$ respectively by deleting their ith and $(n+j)$th rows and columns, and let $\boldsymbol{N}_{i+j-2}$ 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 $\boldsymbol{M}'$ and $\boldsymbol{N}'$ but will contain as factors all the remaining powers of g that occur in $\boldsymbol{M}_{i+j-1}$. Now $c_{ij}\boldsymbol{N}_{i+j-2}$ is a minor of order $i+j-1$ of $\boldsymbol{N}$ and as a consequence is a multiple of $\prod_{k=1}^{i+j-1} g^{t_k}$, a factor of $\boldsymbol{M}_{i+j-1}$. 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_α and b_β 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_{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 $1 \leqq i \leqq \alpha$ and $1 \leqq j \leqq \beta$.

Next regard the block of elements c_{ij}, $1 \leqq i \leqq \alpha$, $\beta < j \leqq n$. Here $c_{ij} \prod_{h=1}^{i-1} a_h \prod_{h=i+1}^{\alpha} a_h \prod_{k=1}^{\beta} b_k$ is a minor of order $\alpha+\beta$ of $\boldsymbol{N}$; it must be a multiple of $\prod_{h=1}^{\alpha} a_h \prod_{k=1}^{\beta} b_k$, the $(\alpha+\beta)$th determinantal divisor of $\boldsymbol{M}$ (of $\boldsymbol{N}$). 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 $\beta < j \leqq n$. Similarly equations (6) are valid for the block of elements c_{ij}, $\alpha < i \leqq n$, $1 \leqq j \leqq \beta$.

Finally if any c_{ij}, $\alpha < i \leqq n$ and $\beta < j \leqq n$, were not identically zero, $\boldsymbol{N}$ would have the nonzero minor $c_{ij} \prod_{h=1}^{\alpha} a_h \prod_{k=1}^{\beta} b_k$ 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 $\alpha < i \leqq n$, and $\beta < j \leqq n$. Equations (6) are therefore valid for $i, j = 1, 2, \cdots, n$.

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

$$(A - xI)\boldsymbol{X} - \boldsymbol{Y}(B - xI) = C. \tag{7}$$

Let

$$\boldsymbol{X} = X_0 + X_1 x + \cdots + X_p x^p,$$
$$\boldsymbol{Y} = Y_0 + Y_1 x + \cdots + Y_q x^q,$$

where X_i, $i = 0, 1, \cdots, p$, and Y_j, $j = 0, 1, \cdots, q$, 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

$$\begin{array}{lllll} AX_0 & & - Y_0 B & & = C, \\ AX_1 & - X_0 & - Y_1 B & + Y_0 & = 0, \\ AX_2 & - X_1 & - Y_2 B & + Y_1 & = 0, \\ \cdots & \cdots & \cdots & \cdots & \cdots , \\ AX_p & - X_{p-1} & - Y_p B & + Y_{p-1} & = 0, \\ & - X_p & & + Y_p & = 0. \end{array}$$

Multiply their members on the right by $I, B, \cdots, B^{p+1}$ 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 $1 \dotplus 1 \dotplus \cdots \dotplus 1 \dotplus 0 \dotplus \cdots \dotplus 0$. The proof of our lemma is a generalization thereof.

and add the members of the resulting equations; we then have

$$A(X_0 + X_1B + \cdots + X_pB^p) - (X_0 + X_1B + \cdots + X_pB^p)B = C.$$

Hence $X_0+X_1B+ \cdots +X_pB^p$ with elements in F is a solution of (1) where p is obviously less than n. Similarly $Y_0+AY_1+ \cdots +A^pY_p$ is also a solution. The theorem is proved.[2]

UNIVERSITY OF TULSA

[2] 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 ALGEBRAS[1]

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,

Received by the editors October 30, 1951.

[1] The preparation of this paper was sponsored in part by the Office of Naval Research.

[2] Numbers in brackets refer to the bibliography at the end of the paper.

[3] For the relevant definitions, see §3.