ON THE MATRIC EQUATION $X^2+AX+XB+C=0$

WILLIAM E. ROTH

The equation

$$X^2 + AX + XB + C = 0, \tag{1}$$

where A, B, and C are $n\times n$ matrices with elements in F, the field of complex numbers or a subfield thereof will be solved for $n\times n$ matrices X with elements in F. Throughout the discussion which follows, the elements of all matrices and the coefficients of all polynomials which arise will be in the field F even when this is not specifically stated. Moreover the similarity of matrices and the reducibility of polynomials will be valid under the rational operations of F.

It should be remarked that equation (1) becomes unilateral if X be replaced by $Y-A$ or by $Z-B$ and in that form has been studied heretofore.[1] This fact was noted by the writer only after the essentials of the present note had been discovered.

1. Necessary conditions. Let

$$R = \begin{pmatrix} -B, & I \\ -C, & A \end{pmatrix}$$

and let X be any solution of (1) with elements in F;[2] then

$$\begin{aligned} \begin{pmatrix} X, & I \\ I, & 0 \end{pmatrix} R \begin{pmatrix} 0, & I \\ I, & -X \end{pmatrix} &= \begin{pmatrix} X+A, & -X^2-AX-XB-C \\ I, & -X-B \end{pmatrix} \\ &= \begin{pmatrix} X+A, & 0 \\ I, & -X-B \end{pmatrix}. \end{aligned} \tag{2}$$

Thus the matrices

$$R \quad \text{and} \quad \begin{pmatrix} X+A, & 0 \\ I, & -X-B \end{pmatrix}$$

are similar, and $|R-\lambda I|$ is reducible to polynomials $f(\lambda)$ and $g(\lambda)$,

Presented to the Society, September 1, 1949; received by the editors June 29, 1949.

[1] Ingraham, *Rational methods in matric equations*, Bull. Amer. Math. Soc. vol. 47 (1941) pp. 61–70; Roth, *On the unilateral equation in matrices*, Trans. Amer. Math. Soc. vol. 32 (1930) pp. 61–80. Either paper cites further references.

[2] The letter I will represent unit matrices of order n or $2n$ to agree with that of other matrices in the same expression. Thus in R above I is of order n.

namely the characteristic polynomials of $X+A$ and $-X-B$, respectively, with coefficients in F. The following theorem, which is implied by the above, will be a guide in the quest of solutions of the given equation.

THEOREM I. *If equation* (1) *has a solution with elements in F, then there exists at least one pair of polynomials $f_\alpha(\lambda)$ of degree $\alpha \leqq n$ and $g_\beta(\lambda)$ of degree $\beta \leqq n$ with coefficients in F such that $f_\alpha(R)g_\beta(R)=0$, where $f_\alpha(\lambda)g_\beta(\lambda)$ is not necessarily the minimum polynomial satisfied by R but is a divisor of $|R-\lambda I|$ and such that $f_\alpha(X+A)=0$ and $g_\beta(-X-B)=0$.*

Polynomials $f_\alpha(\lambda)$ and $g_\beta(\lambda)$, of degree $\alpha \leqq n$ and $\beta \leqq n$ respectively, such that $f_\alpha(\lambda)g_\beta(\lambda)$ is a divisor of $|R-\lambda I|$ and a multiple of the minimum polynomial satisfied by R will be designated as admissible polynomials.

Let $f_\alpha(\lambda)$ be an admissible polynomial such that $f_\alpha(X+A)=0$, where X is a solution of (1) and let

$$f_\alpha(R) = \begin{pmatrix} U, & M \\ V, & N \end{pmatrix},$$

where U, V, M, and N are polynomials in the matrices A, B, and C; hence according to (2) we have

$$\begin{pmatrix} X, & I \\ I, & 0 \end{pmatrix} f_\alpha(R) \begin{pmatrix} 0, & I \\ I, & -X \end{pmatrix} = \begin{pmatrix} f_\alpha(X+A), & 0 \\ *, & f_\alpha(-X-B) \end{pmatrix}$$
$$= \begin{pmatrix} XM+N, & XU+V-(XM+N)X \\ M, & U-MX \end{pmatrix};$$

and consequently the equations,

$$XM + N = 0, \qquad XU + V = 0, \tag{3}$$

are simultaneously satisfied whether M be singular or not, and

$$(X, I)f_\alpha(R) = (X, I)\begin{pmatrix} U, & M \\ V, & N \end{pmatrix} = 0. \tag{4}$$

Now let

$$f_\alpha(R)\begin{pmatrix} I \\ -X_1 \end{pmatrix} = \begin{pmatrix} U, & M \\ V, & N \end{pmatrix}\begin{pmatrix} I \\ -X_1 \end{pmatrix} = 0,$$

and M be nonsingular.[3] Then $f_\alpha(R)$ and M have the same rank n

[3] It can be shown that M is singular only if, but not necessarily if, $f_\alpha(\lambda)$ and $g_\beta(\lambda)$ have a common factor.

and a unique solution of this equation exists such that

$$U - MX_1 = 0, \qquad V - NX_1 = 0. \tag{5}$$

We shall show that X_1 so obtained is also a solution of (1) and that

$$X_1 = - M^{-1}(X + B)M - A, \tag{6}$$

where X is a solution of (1) and as such satisfies (4).

Because R and $f_\alpha(R)$ are commutative, we readily obtain the relations:

$$\begin{aligned} N &= U + MA + BM, \\ V &= BU - UB - MC = AN - NA - CM, \end{aligned} \tag{7}$$

and by (5)

$$\begin{aligned} V - NX_1 &= BU - UB - MC - (U + MA + BM)X_1 \\ &= BMX_1 - MX_1B - MC - (MX_1 + MA + BM)X_1 \\ &= - M(X_1^2 + AX_1 + X_1B + C) = 0. \end{aligned}$$

Now M is nonsingular and X_1 is therefore a solution of (1). Equation (6) follows from (3), (5), and the first equation of (7).

The solutions X and X_1 of equation (1) are such that $f_\alpha(X+A)=0$ and $f_\alpha(-X_1-B)=0$ and arise in case $f_\alpha(R)$ and

$$\begin{pmatrix} 0, & 0 \\ M, & 0 \end{pmatrix}$$

are equivalent under elementary transformations since

$$\begin{pmatrix} X, & I \\ I, & 0 \end{pmatrix} f_\alpha(R) \begin{pmatrix} 0, & I \\ I, & -X_1 \end{pmatrix} = \begin{pmatrix} X, & I \\ I, & 0 \end{pmatrix}\begin{pmatrix} U, & M \\ V, & N \end{pmatrix}\begin{pmatrix} 0, & I \\ I, & -X_1 \end{pmatrix} = \begin{pmatrix} 0, & 0 \\ M, & 0 \end{pmatrix}.$$

The following theorem is proved.

THEOREM II. *If X is a solution of equation* (1) *with elements in F, then there exists a polynomial $f_\alpha(\lambda)$ of degree $\alpha \leqq n$ with coefficients in F such that equation* (4) *is satisfied. If in addition M is nonsingular, $X_1 = -M^{-1}(X+B)M-A$ is also a solution of* (1) *with elements in F.*

2. **The converse problem.** We shall state the theorem:

THEOREM III. *If $f_\alpha(\lambda)$ is a polynomial of degree $\alpha \leqq n$ with coefficients in F such that*

$$f_\alpha(R) = \begin{pmatrix} U, & M \\ V, & N \end{pmatrix}$$

is of rank n and if M is nonsingular, then X such that $(X, I)f_\alpha(R)=0$ *is a solution of equation* (1) *with elements in F and* $X_1=-M^{-1}(X+B)M-A$ *is likewise such a solution.*

The proof of this theorem is entirely analogous to that above by which X_1 was shown to be a solution of (1) and will not be given here. Though Theorem III holds for any polynomial $f_\alpha(\lambda)$ of whatever degree in λ; those of lowest degree which may satisfy the hypotheses will be found only among the admissible polynomials defined above.

Addendum. The referee's suggestion that the results above might be extended to the equation

$$XDX + AX + XB + C = 0,$$

where D is nonsingular, was anticipated by Professor James H. Bell in a letter to the writer dated March 3, 1949. This generalization appeared of slight significance because the product of this equation on the right (on the left) by D can readily be transformed to (1) in case D is nonsingular. However, the writer has subsequently noted that if I be replaced by D in R and consequently if $X+A$ be replaced by $XD+A$ and $X+B$ by $DX+B$, as suggested by the referee, the content of the above paper after appropriate minor changes is still valid even when D is singular. Moreover, in case D is singular, this equation is not transformable to the unilateral form and consequently is of particular interest. The writer wishes to express his indebtedness to Professor Bell and to the referee.

UNIVERSITY OF TULSA