ON THE MATRICES A AND f(A)

R. C. Thompson*

(received March 26, 1968)

In a recent note [1], M. R. Embry proved that if A is an operator on a Banach space then, under a certain condition on the spectrum of A, each operator commuting with A^n also commutes with A, where n is a fixed positive integer. It turns out that, when A is a finite matrix, Embry's conditions imply that A is a polynomial in A^n and hence plainly each operator commuting with A^n also commutes with A. Since A^n is a polynomial in A and since any matrix commuting with A commutes also with A^n, we see that for finite matrices Embry's problem is a special case of each of the following more general problems:

I. Let A and $B = f(A)$ be $m \times m$ matrices over a field $\mathfrak{F}$, where $f(\lambda)$ is a polynomial over $\mathfrak{F}$. Under what circumstances does a polynomial $g(\lambda)$ exist over $\mathfrak{F}$ such that $A = g(B)$?

II. Let A be an $m \times m$ matrix over $\mathfrak{F}$. Characterize those matrices B over $\mathfrak{F}$ such that the algebra of matrices over $\mathfrak{F}$ commuting with A coincides with the algebra of matrices over $\mathfrak{F}$ commuting with B.

* The preparation of this paper was supported in part by the U.S. Air Force Office of Scientific Research, under Grant 698-67.

For finite matrices over $\mathfrak{F}$, these two problems are equivalent and both will be solved here. Special cases of problem I are often encountered in the literature, for example, in the study of powers and roots of matrices.

THEOREM 1. *Let* A *be a square matrix over* $\mathfrak{F}$ *and let* $B = f(A)$ *where* $f(\lambda)$ *is a polynomial over* $\mathfrak{F}$. *Let* $m(\lambda)$ *be the minimal polynomial of* A. *Then a polynomial* $g(\lambda)$ *over* $\mathfrak{F}$ *exists such that* $A = g(B)$ *if, and only if, the following two conditions are satisfied*:

(i) $f(\lambda)$ *is one to one on the roots of* $m(\lambda)$;

(ii) $f'(\alpha) \neq 0$ *whenever* α *is an root of* $m(\lambda)$ *of multiplicity* > 1. *Here* $f'(\lambda)$ *is the derivative of* $f(\lambda)$.

An equivalent formulation of condition (ii) is

$$(m(\lambda), m'(\lambda), f'(\lambda)) = 1.$$

Proof. Suppose $A = g(f(A))$. Then the polynomial $\lambda - g(f(\lambda))$ must be divisible by $m(\lambda)$, and hence

$$\lambda = g(f(\lambda)) + G(\lambda)m(\lambda) \,, \tag{1}$$

where $G(\lambda)$ is some polynomial. If α_1 and α_2 are distinct zeros of $m(\lambda)$ and $f(\alpha_1) = f(\alpha_2)$, we have $\alpha_1 = g(f(\alpha_1) = g(f(\alpha_2)) = \alpha_2$, a contradiction. Therefore f is one to one on the zeros of $m(\lambda)$. Now let α be a multiple root of $m(\lambda)$. From (1) we get

$$1 = g'(f(\lambda))f'(\lambda) + G'(\lambda)m(\lambda) + G(\lambda)m'(\lambda)$$

and hence setting $\lambda = \alpha$ yields $1 = g'(f(\alpha))f'(\alpha)$. Therefore $f'(\alpha) \neq 0$.

Suppose now that $B = f(A)$ and that polynomial $f(\lambda)$ satisfies conditions (i) and (ii). We construct a polynomial $g(\lambda)$ over the algebraic closure $\bar{\mathfrak{F}}$ of $\mathfrak{F}$ such that $A = g(B)$. Later we will get a polynomial $g(\lambda)$ over $\mathfrak{F}$.

Let $\alpha_1, \ldots, \alpha_s$ be the distinct roots of $m(\lambda)$ with multiplicity > 1 and let $\beta_1, \ldots, \beta_t$ be the distinct simple roots of $m(\lambda)$. Let

$$F(\lambda) = (\lambda - \alpha_1) \ldots (\lambda - \alpha_s) ,$$

$$\hat{F}(\lambda) = (\lambda - f(\alpha_1)) \ldots (\lambda - f(\alpha_s)) ,$$

$$H(\lambda) = (\lambda - \beta_1) \ldots (\lambda - \beta_t) .$$

Since $f(\lambda)$ is one to one on the zeros of $m(\lambda)$, $f(\alpha_1), \ldots, f(\alpha_s)$, $f(\beta_1), \ldots, f(\beta_t)$ are distinct, so that (by an interpolation formula) a polynomial $p_o(\lambda)$ exists such that

$$p_o(f(\alpha_i)) = \alpha_i, \qquad 1 \leq i \leq s ,$$

$$p_o(f(\beta_i)) = \beta_i, \qquad 1 \leq i \leq t .$$

Hence all of $\alpha_1, \ldots, \alpha_s$, $\beta_1, \ldots, \beta_t$ are zeros of the polynomial $\lambda - p_o(f(\lambda))$ and therefore

$$\lambda \equiv p_o(f(\lambda)) \pmod{F(\lambda)H(\lambda)}.$$

We now construct by induction a sequence $p_0(\lambda), p_1(\lambda), \ldots, p_r(\lambda), \ldots$ of polynomials such that

(2) $$\lambda \equiv p_r(f(\lambda)) \pmod{F(\lambda)^{r+1}H(\lambda)} .$$

Given $p_r(\lambda)$, we proceed to construct $p_{r+1}(\lambda)$. We put

(3) $$p_{r+1}(\lambda) = p_r(\lambda) + x(\lambda)\hat{F}(\lambda)^{r+1} .$$

Here $x(\lambda)$ is a polynomial to be determined shortly. Because of (2),

(4) $$\lambda = p_r(f(\lambda)) + G(\lambda)F(\lambda)^{r+1}H(\lambda) \ ,$$

where $G(\lambda)$ is some polynomial. Substituting (3) into (4), we get

(5) $$\lambda = p_{r+1}(f(\lambda)) + F(\lambda)^{r+1}\left\{ G(\lambda)H(\lambda) - x(f(\lambda)) \left(\frac{\hat{F}(f(\lambda))}{F(\lambda)}\right)^{r+1} \right\} .$$

Observe that

(6) $$\frac{\hat{F}(f(\lambda))}{F(\lambda)} = \prod_{i=1}^{s} \frac{f(\lambda) - f(\alpha_i)}{\lambda - \alpha_i} \quad .$$

Therefore $\hat{F}(f(\lambda))/F(\lambda)$ is a polynomial and $\hat{F}(f(\lambda))/F(\lambda)\big|_{\lambda=\alpha_j} \neq 0$ for $j = 1, 2,\ldots,s$, since $f(\alpha_j) \neq f(\alpha_i)$ if $j \neq i$ and $f'(\alpha_j) \neq 0$. Since $f(\alpha_1),\ldots,f(\alpha_s)$, $f(\beta_1),\ldots,f(\beta_t)$ are distinct, we may find a polynomial $x(\lambda)$ such that

$$x(f(\alpha_i)) = G(\lambda)H(\lambda) \left(\frac{\hat{F}(f(\lambda))}{F(\lambda)}\right)^{-(r+1)}\Bigg|_{\lambda=\alpha_i} \ , \quad i = 1,\ldots,s,$$

$$x(f(\beta_i)) = 0, \quad i = 1, 2,\ldots,t.$$

The expression in { } in (5) is now a polynomial vanishing when λ is $\alpha_1,\ldots,\alpha_s$, $\beta_1,\ldots,\beta_t$. Therefore this polynomial is divisible by $F(\lambda)H(\lambda)$. Hence

$$\lambda \equiv p_{r+1}(f(\lambda) \ (\text{mod } F(\lambda)^{r+2}H(\lambda)).$$

This completes the inductive step.

Now, for sufficiently large r, $F(\lambda)^{r+1}H(\lambda)$ is divisible by $m(\lambda)$. For this large r, $F(A)^{r+1}H(A) = 0$. But then

$$A = p_{r+1}(f(A)) \ .$$

Thus we have a polynomial $g(\lambda)$ over $\mathfrak{F}$ such that $A = g(B)$. But it is an elementary fact that if A is a polynomial in B, where the polynomial has coefficients in an extension field of $\mathfrak{F}$, then A is also a polynomial in B where the polynomial has coefficients in $\mathfrak{F}$.

The proof of Theorem 1 is now complete.

THEOREM 2. *Let* A *and* B *be* $m \times m$ *matrices over* $\mathfrak{F}$, *and let* C *and* D *be* $n \times n$ *matrices over* $\mathfrak{F}$. *Suppose* A *and* C *have a common minimal polynomial* $m(\lambda)$. *Then the set of all* $m \times n$ *matrices* X *over* $\mathfrak{F}$ *for which*

$$AX = XC \tag{6}$$

coincides with the set of all $m \times n$ *matrices* X *over* $\mathfrak{F}$ *for which*

$$BX = XD \tag{7}$$

if and only if:

(i) $B = f(A)$ *and* $D = f(C)$ *for some polynomial* $f(\lambda)$ *over* $\mathfrak{F}$;

(ii) $f(\lambda)$ *is one to one on the roots of* $m(\lambda)$;

(iii) $f'(\alpha) \neq 0$ *for each nonsimple root* α *of* $m(\lambda)$.

Proof. It was proved in [2, Theorem 1] that if A and C have a common minimal polynomial, and if $BX = XD$ whenever $AX = XC$, then $B = f(A)$ and $D = f(C)$ for some polynomial $f(\lambda)$ over $\mathfrak{F}$. We claim that $B = f(A)$ and $D = f(C)$ must have the same minimal polynomial. This follows from the fact that the coincidence of the minimal polynomials of A and C forces the algebra of polynomials (over $\mathfrak{F}$) in A to be isomorphic to the algebra of polynomials (over $\mathfrak{F}$) in C.

https://doi.org/10.4153/CMB-1969-075-7 Published online by Cambridge University Press

In this algebra isomorphism B and D are isomorphic elements and hence must have the same minimal polynomial. Now it follows that we can reverse the argument in the first part of the proof and deduce from the implication $BX = XD \Rightarrow AX = XC$ that $A = g(B)$ and $C = g(D)$ for some polynomial $g(\lambda)$ over $\mathfrak{F}$. Therefore $A = g(f(A))$ and $C = g(f(C))$, and hence by Theorem 1 $f(\lambda)$ must satisfy conditions (ii) and (iii).

Suppose now that $B = f(A)$ and $D = f(C)$, where $f(\lambda)$ satisfies conditions (ii) and (iii). In the proof of Theorem 1 polynomial $p_r(\lambda)$ was obtained such that $\lambda \equiv p_r(f(\lambda)) \pmod{F(\lambda)^{r+1}H(\lambda)}$. For sufficiently large r we obtain $A = p_{r+1}(f(A))$ and $C = p_{r+1}(f(C))$. Thus $A = g(B)$ and $C = g(D)$, where $g(\lambda) = p_{r+1}(\lambda)$. But it follows easily (see [2]) from $B = f(A)$ and $D = f(C)$ that $AX = XC$ implies $BX = XD$; and similarly from $A = g(B)$ and $C = g(D)$ we see that $BX = XD$ implies $AX = XC$. The proof is complete.

When $C = A$ and $D = B$ Theorem 2 provides the solution of problem II.

COROLLARY. *Let* A *be an* $m \times m$ *matrix over* $\mathfrak{F}$ *with minimal polynomial* $m(\lambda)$. *Let* B *be an* $m \times m$ *matrix over* $\mathfrak{F}$. *Then the algebra of matrices over* $\mathfrak{F}$ *commuting with* A *coincides with the algebra of matrices over* $\mathfrak{F}$ *commuting with* B *if and only if* $B = f(A)$, *where* $f(\lambda)$ *is a polynomial satisfying the conditions* (i) *and* (ii) *of Theorem 1.*

As special cases, we obtain the following theorems bearing on Embry's result.

THEOREM 3. *Let* n *be a positive integer,* $n > 1$, *and let the*

https://doi.org/10.4153/CMB-1969-075-7 Published online by Cambridge University Press

characteristic of $\mathfrak{F}$ not divide n. Let A be an $m \times m$ matrix over $\mathfrak{F}$ with minimal polynomial $m(\lambda)$. Then the algebra of matrices over $\mathfrak{F}$ commuting with A coincides with the algebra of matrices over $\mathfrak{F}$ commuting with A^n if and only if:

(i) when α is a nonzero eigenvalue of A, $\zeta_n\alpha$ is not an eigenvalue of A for each n^{th} root of unity $\zeta_n \neq 1$;

(ii) if A is singular, $\lambda = 0$ is a simple root of $m(\lambda)$.

THEOREM 4. Let p be the characteristic of $\mathfrak{F}$. The the matrix algebra over $\mathfrak{F}$ commuting with A coincides with the matrix algebra over $\mathfrak{F}$ commuting with A^p if and only if A is similar to a diagonal matrix.

REFERENCES

1. M.R. Embry, N^{th} roots of operators. Proc. Amer. Math. Soc. 19 (1968) 63-68.

2. R.C. Thompson, Generalization of a well-known result in matrix theory. Proc. Glasgow Math. Association 7 (1965) 29-31.

University of California
Santa Barbara

https://doi.org/10.4153/CMB-1969-075-7 Published online by Cambridge University Press