Indian J. pure appl. Math., **19** (9) : 854-864, September 1988

ON THE MATRIX ROOTS OF $f(X) = A$

EUGENE SPIEGEL

Department of Mathematics, The University of Connecticut, Storrs, Connecticut 06268, *USA*

(*Received* 5 *March* 1987; *after revision* 8 *March* 1988)

For A an $n \times n$ complex valued matrix and $f(x)$ a non-constant polynomial with complex coefficients, we give conditions for the existence of an $n \times n$ complex matrix B with $f(B) = A$. The results are then applied to the study of the composition of polynomials.

Throughout this note $M_n(\mathbb{C})$ denotes the set of $n \times n$ complex valued matrices. $f(x)$ will denote a non-constant polynomial with complex coefficients. For $A \in M_n(\mathbb{C})$ we investigate the solution in $M_n(\mathbb{C})$ of the equation $f(X) = A$. The first section gives some necessary preliminary calculations, the second section outlines the solution of the problem and the last section shows an application of the results to the determination of the range of the composition of polynomials modulo an ideal of $\mathbb{C}[x]$.

§1. In this section the basic calculations are done.

Lemma 1.1—If $f(x) \in \mathbb{C}[x]$ and $B, C \in M_n(\mathbb{C})$ with C invertible, then $f(CBC^{-1}) = Cf(B)C^{-1}$.

PROOF : Straightforward.

Lemma 1.2—Let

$$B = \begin{pmatrix} \lambda & 1 & & & & O \\ & \lambda & 1 & & & \\ & & \lambda & 1 & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & \cdot \\ & & O & & & 1 \\ & & & & & \lambda \end{pmatrix}$$

be an $n \times n$ complex matrix and $f(x) \in \mathbb{C}[x]$. Then

$$f(B) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2!} & \frac{f'''(\lambda)}{3!} & \cdots & \\ & f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2!} & \cdots & \\ & & f(\lambda) & f'(\lambda) & \cdots & \\ & & \cdots & \cdots & \cdots & \\ O & & & & f(\lambda) & f'(\lambda) \\ & & & & & f(\lambda) \end{bmatrix}.$$

Proof : Write $f(x) = a_k x^k + a_{k-1} x^{k-1} + \ldots + a_0$, $a_i \in \mathbb{C}$, so that $f(B) = a_k B^k + a_{k-1} B^{k-1} + \ldots + a_0$. Let

$$N = \begin{bmatrix} 0 & 1 & 0 & 0 & \ldots & O \\ & 0 & 1 & 0 & \ldots & \\ & & 0 & 1 & \ldots & \\ O & & & & & \\ & & & & 0 & 1 \\ & & & & & 0 \end{bmatrix}.$$

Then

$$N^r = \begin{bmatrix} 0 & 0 & \ldots & 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & 1 & 0 & \ldots & 0 \\ . & . & . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & . & . & 0 & 1 \\ 0 & 0 & . & . & . & . & . & 0 & 0 \\ . & . & . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & . & . & 0 & 0 \end{bmatrix}$$

where the 1 in the first row appears in the $(r + 1)$st position. $B = N + \lambda$ where, for simplicity, λ implies the matrix λI.

$$B^t = (N + \lambda)^t = \sum_{j=0}^{t} \binom{t}{j} N^j \lambda^{t-j}.$$

$$f(B) = \sum_{t=0}^{k} a_t B^t = \sum_{t=0}^{k} a_t \sum_{j=0}^{t} \binom{t}{j} \lambda^{t-j} N^j$$

$$= \sum_{t=0}^{k} \sum_{j=0}^{k} a_t \binom{t}{j} \lambda^{t-j} N^j \text{ (Since } j > t \text{ implies} \binom{t}{j} = 0)$$

(*equation continued on p. 856*)

$$= \sum_{j=0}^{k} \left(\sum_{t=0}^{k} a_t \binom{t}{j} \lambda^{t-j} \right) N^j$$

$$= \sum_{j=0}^{k} \frac{f^{(j)}(\lambda)}{j!} N^j \text{ which is our required form.}$$

Let $A = (a_{ij})$ be an $n \times n$ complex matrix. For $k = 1, \dots, n$, we call the elements $a_{1,k}\, a_{2,k-1}, \dots, a_{n-k+1,n}$ the $(k-1)$ – st upper diagonal of A. The main diagonal of A is then the 0-th upper diagonal. We will say a_{ij} is below the $(k-1)$ st upper diagonal if either $j < k$ or $j \geqslant k$ and $i > (1+j) - k$.

The basic lemma for the solution of our problem is the following.

Lemma 1.3—Let $1 \leqslant k \leqslant n-1$ and $B \in M_n(\mathbb{C})$. Suppose each element of the k-th upper diagonal of B is non-zero and each element below the kth upper diagonal is zero. Write $n = qk + r$ with $1 \leqslant r \leqslant k$. Then the Jordan form of B consists of k nilpotent blocks, r of them $(q+1) \times (q+1)$ and the remaining $k - r$, $q \times q$.

Proof : Write $B = (b_{ij})$ and let B be the matrix of a transformation T from V to V (V denotes an n-dimensional complex vector space) with respect to a basis $v_1, v_2, \dots, v_n$ of V ; i. e. $Tv_l = \sum_{i=1}^{n} b_{il} v_i$ for $l = 1, 2, \dots, n$. Because each element below the kth diagonal of B is zero, coupled with the assumption that each element on the kth upper diagonal is non-zero, we have that $Tv_{k+1}, Tv_{k+2}, \dots, Tv_n, v_{n-k+1}, v_{n-k+2}, \dots, v_n$ is a basis of V. The result will follow if we can find k T-invariant subspaces $U_1, U_2, \dots, U_k$ with the properties :

(a) $\dim(U_i) = q + 1,\ 1 \leqslant i \leqslant r$

$\dim(U_i) = q,\ \ r + 1 \leqslant i \leqslant k$

(b) $V = U_1 \oplus U_2 \oplus \dots \oplus U_k$

(c) For each i, $1 \leqslant i \leqslant k$, the degree of the minimum polynomial to T restricted to U_i equals the dimension of U_i.

For $1 \leqslant i \leqslant r$, set U_i to be the subspace spanned by $v_{n-i+1}, Tv_{n-i+1}, T^2 v_{n-i+1}, \dots, T^q v_{n-i+1}$, and for $r + 1 \leqslant i \leqslant k$, let U_i be the subspace spanned by $v_{n-i+1}, Tv_{n-i+1}, \dots, T^{q-1} v_{n-i+1}$ where we have interpreted T^{q-1} to be the identity if $q = 1$.

Since Tv_i is a linear combination of $v_1, v_2, \dots, v_{i-k}$ for $i \geqslant k + 1$, with the coefficient of v_{i-k} non-zero, while $Tv_i = 0$ if $i \leqslant k$, we easily deduce by induction that $T^j v_i$ is a linear combination of $v_1, v_2, \dots, v_{i-jk}$, with $v_l = 0$ if $l \leqslant 0$. So $T^j v_i$ is zero if and only if $i - jk \leqslant 0$. If $n - r + 1 \leqslant i \leqslant n$, $T^q v_i$ is non-zero since $n + 1 > qk + r$

while $T^{q+1} v_i = 0$ because $n + 1 \leqslant (q + 1) k + r$. Similarly, if $n - k + 1 \leqslant i \leqslant n - r$, $T^{q-1} v_i \neq 0$ but $T^q v_i = 0$. This implies that the subspaces U_i, $i = 1, \ldots, k$, are T-invariant cyclic subspaces. But it is easy to check (using the basis $Tv_{k+1}, \ldots, Tv_n, v_{n-k+1}, v_{n-k+2}, \ldots v_n$) that the collection $v_n, Tv_n, \ldots, T^q v_n, v_{n-1}, Tv_{n-1}, \ldots, T^q v_{n-1}, \ldots, v_{n-r+1}, Tv_{n-r+1}, \ldots, T^q v_{n-r+1}, v_{n-r}, Tv_{n-r}, \ldots, T^{q-1} v_{n-r}, \ldots, v_{n-k+1}, Tv_{n-k+1}, \ldots, T^{q-1} v_{n-k+1}$ is also a basis of V. The result now follows.

§2. *Proposition* 2 1—Let $A \in M_n(\mathbb{C})$, $f(x) \in C[x]$ and λ_i, $i = 1, \ldots, k$ be distinct complex numbers. Suppose A is similar to

$$\begin{bmatrix} A_1 & & & & O \\ & A_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ O & & & & A_k \end{bmatrix}$$

where A is an $n_i \times n_i$ matrix having λ_i as its only eigenvalue. There exists a $B \in M_n(\mathbb{C})$ such that $f(B) = A$ if and only if there exists for $i = 1, \ldots, k$, $B_i \in M_{n_i}(\mathbb{C})$ such that $f(B_i) = A_i$.

PROOF : If there exist $B_1, B_2, \ldots, B_k$ with $B_i \in M_{n_i}(\mathbb{C})$ such that $f(B_i) = A_i$, then

$$B = \begin{bmatrix} B_1 & & & & \\ & B_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & B_k \end{bmatrix}$$

is an $n \times n$ complex matrix for which $f(B)$ is similar to A. By Lemma 1.1, there is then a solution to $f(X) = A$.

Conversely, suppose there is $B \in M_n(\mathbb{C})$ such that $f(B) = A$. Let J denote the Jordan form of B. By Lemma 1.1, $f(J)$ is a matrix similar to A and so, by Lemma 1.2, the eigenvalues of B are solutions of $f(x) = \lambda_i$, $i = 1, \ldots, k$. Let D_i be the sum of all Jordan blocks of J which have eigenvalues obtained by solving $f(x) = \lambda_i$. We can then assume

$$J = \begin{bmatrix} D_1 & & & & \\ & D_2 & & & \\ & & , & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & D_k \end{bmatrix}.$$

Because

$$f(J) = \begin{bmatrix} f(D_1) & & & & \\ & f(D_2) & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & f(D_k) \end{bmatrix}$$

the Jordan form of it, being the sum of the Jordan forms of $f(D_i)$, $i = 1, \ldots, k$, must have the sum of all blocks with λ_i on the main diagonal as the Jordan form of $f(D_i)$. Like reasoning for A_i implies that A_i and $f(D_i)$ for $i = 1, \ldots, k$, are similar as they have the same Jordan form. For $i = 1, \ldots, k$. let $C_i \in M_{n_i}(\mathbb{C})$ be such that $C_i\, f(D_i)\, C_i^{-1} = A_i$ and $B_i = C_i D_i C_i^{-1}$. By Lemma 1.1, the result now follows.

In view of Proposition 2.1, we may assume that A has only one eigenvalue λ, in finding solutions to $f(X) = A$.

Suppose A is an $n \times n$ complex matrix with λ its only eigenvalue. We say that A is of type $(a_1, a_2, \ldots, a_n)$ if the Jordan form of A consists of a_1 1×1 blocks, a_2 2×2 blocks, $\ldots$, a_n $n \times n$ blocks, where we have called the $i \times i$ matrix

$$\begin{bmatrix} \lambda & 1 & & & & \\ & \lambda & 1 & & O & \\ & & \lambda & 1 & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & \cdot \\ O & & & & & 1 \\ & & & & & \lambda \end{bmatrix}$$

an $i \times i$ block. Note that $a_i \geqslant 0$ and $\sum_{i=1}^{n} i a_i = n$.

We will call the $n \times n$ matrix B an irreducible matrix if B is not similar to a matrix of the form

$$\begin{pmatrix} B_1 & O \\ O & B_2 \end{pmatrix}$$

where B_1 and B_2 are square matrices.

We say $f(X) = A$ has an irreducible solution if there exists an irreducible matrix $B \in M_n(f)$ such that $f(B) = A$. By Proposition 2.1, in order for $f(X)$ A to have an irreducible solution. A must have a single eigenvalue. Of course, if A is a 1×1 matrix, there exists an irreducible solution to $f(X) = A$.

Proposition 2.2— Let $n \geqslant 2$, and A an $n \times n$ complex matrix having a single eigenvalue λ. Suppose that A is of type $(a_1, a_2, \ldots, a_n)$. Let $t = \max\limits_{1 \leqslant i \leqslant n} \{i \mid a_i \neq 0\}$ and write $k = a_t + a_{t-1}$. (If $t = 1$, call $a_0 = 0$.) The equation $f(X) = A$ has an irreducible solution if and only if

(i) $n = kt - a_{t-1}$

(ii) there exists a $\gamma \in \mathbb{C}$ such that $f(\gamma) = \lambda$, $f'(\gamma) = \dots = f^{(k-1)}(\gamma) \neq 0$, and $f^{(k)}(\gamma) \neq 0$.

Proof : Suppose there is an irreducible solution B of $f(X) = A$. Then the Jordan form $\bar{B}$ of B is

$$\begin{bmatrix} \gamma & 1 & & & & O \\ & \gamma & 1 & & & \\ & & \gamma & 1 & & \\ & & & \ddots & & \\ O & & & & & 1 \\ & & & & & \gamma \end{bmatrix}$$

for some $\gamma \in \mathbb{C}$. $f(\bar{B})$ is similar to A, and $f(\gamma) = \lambda$. Suppose γ has the property that $f^{(j)}(\gamma) \neq 0$ but $f'(\gamma) = \dots = f^{(j-1)}(\gamma) = 0$. By Lemmas 1.2 and 1.3, the Jordan form of $f(B)$ has Jordan blocks only of size $(q+1) \times (q+1)$ and $q \times q$ where q satisfies the condition $n = qj + r$ with $1 \leqslant r \leqslant j$. There are r blocks of size $(q+1) \times (q+1)$ and $j - r$ blocks of size $q \times q$. But $f(B)$ and A have the same Jordan form and so $j = k$, $q = t - 1$, $r = a_t$ and the result follows.

Conversely suppose $n = kt - a_{t-1}$ and $\gamma \in \mathbb{C}$ with $f(\gamma) = \lambda$.

$$f'(\gamma) = \dots = f^{(k-1)}(\gamma) = 0, \; f^{(k)}(\gamma) \neq 0.$$

Let

$$B = \begin{bmatrix} \gamma & 1 & & & & O \\ & \gamma & 1 & & & \\ & & \gamma & 1 & & \\ & & & \ddots & & \\ & & & & \gamma & 1 \\ O & & & & & \gamma \end{bmatrix}$$

From Lemmas 1.2 and 1.3 we see that $f(B)$ is similar to A and from Lemma 1.1 the proposition follows.

Corollary 2.3—Let $A \in M_n(\mathbb{C})$ and suppose A is similar to

$$\begin{pmatrix} \lambda & 1 & & & & O \\ & \lambda & 1 & & & \\ & & \lambda & 1 & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & \cdot & \\ & & & & \lambda & 1 \\ O & & & & & \lambda \end{pmatrix} \text{ with } \lambda \in \mathbb{C}$$

There exists $B \in M_n(\mathbb{C})$ such that $f(B) = A$ if and only if

(i) $n = 1$ or

(ii) there exists $\gamma \in \mathbb{C}$, such that $f(\gamma) = \lambda$ and $f'(\gamma) \neq 0$.

Proof : If $n = 1$, set B to be a root of $f(x) = \lambda$. Suppose $n \geqslant 2$. Then A is of type $(0, 0, \ldots, 0, 1)$. If there exists a solution $B \in M_n(\mathbb{C})$ such that $f(B) = A$, B is irreducible as A is. By Proposition 2.2, $k = 1$ and there exists $\gamma \in \mathbb{C}$ with $f(\gamma) = \lambda$ and $f'(\gamma) \neq 0$. Conversely, if there is a $\gamma \in \mathbb{C}$ with $f(\gamma) = \lambda$ and $f'(\gamma) \neq 0$ Proposition 2.2 implies there is a solution to $f(X) = A$.

Corollary 2.4—Let $A \in M_n(\mathbb{C})$ and suppose λ is the only eigenvalue for A. If there is a $\gamma \in \mathbb{C}$ with $f(\gamma) = \lambda$ and $f'(\gamma) \neq 0$ then there exists $B \in M_n(\mathbb{C})$ such that $f(B) = A$.

Proof : The Jordan form of A consists of a sum of blocks A_i of the type in Corollary 2.3. Each block gives a solution B_i of $f(X) = A_i$ and a matrix similar to the diagonal sum of B_i's is a solution of $f(X) = A$.

Corollary 2.5—Let $A \in M_n(\mathbb{C})$ and suppose A is diagonalizable. Then there exists $B \in M_n(\mathbb{C})$ such that $f(X) = A$.

Proof : Since A is similar to a diagonal sum of 1×1 matrices by Corollary 2.3 we have a solution to $f(X) = A$.

Now let A be an $n \times n$ complex matrix and $f(x)$ a non-constant complex coefficient polynomial. To see if $f(x) = A$ has a solution $B \in M_n(\mathbb{C})$ we employ the following procedure. Break up A, or a matrix similar to A, into the diagonal sum of matrices, $A_1, A_2, \ldots, A_k$ where A_i is a matrix having λ_i as its only eigenvalue and $\lambda_i \neq \lambda_j$ for $i \neq j$. By Proposition 2.1, $f(X) = A$ has a solution if and only if $f(X) = A_i$, $i = 1, \ldots, k$, each have one. To solve $f(X) = A_i$, we must find, say l_i, irreducible matrices $B_{i_1}, B_{i_2}, \ldots, B_{i_{l_i}}$ such that $\sum_{j=1}^{l_i} \dim (B_{i_j})$-dim A_i and with the additional property that

$$\begin{bmatrix} f(B_{i_1}) & & & \\ & f(B_{i_2}) & & \\ & & \ddots & \\ & & & f(B_{i_l}) \end{bmatrix} \text{ is similar to } A_i.$$

Conditions for this are implied by Proposition 2.2 and Corollary 2.3. We now illustrate the technique outlined above.

Proposition 2.6—Suppose $1 \geq t$, $a, \gamma \in \mathbb{C}$ with $a \neq 0$ and A is a non-scalar matrix in $M_n(\mathbb{C})$ with λ its only eigenvalue. Let $(a_1, a_2, \ldots, a_n)$ be the type of A, and write $b_i = \sum_{j=1}^{n} a_j$, $i = 1, \ldots, n$. Suppose further $f(x) = a(x-\gamma)^r + \lambda$. Then $f(X) = A$ has a solution in $M_n(\mathbb{C})$ if and only if for each i

$$b_i \equiv 0 \pmod{r} \text{ or } i > 1 \text{ and } \left[\frac{b_{i-1}}{r}\right] > \left[\frac{b_i}{r}\right].$$

($[x]$ denotes the greatest integer in x).

PROOF : If $n = 2$, then the theorem follows by Proposition 2.2. We can thus assume, via induction, that the theorem is true for smaller values of n. Assume there is a solution B of $f(X) = A$. If B is irreducible the result will follow from Proposition 2.2. Otherwise, there exist square matrices A_1 and A_2 such that A is similar to

$$\begin{pmatrix} A_1 & O \\ O & A_2 \end{pmatrix}$$

and square matrices B_1 and B_2 with B similar to

$$\begin{pmatrix} B_1 & O \\ O & B_2 \end{pmatrix}$$

and $f(B_k) = A_k$, $k = 1.2$. Suppose the type of A_k, $k = 1.2$, is $(a_{k_1}, a_{k_2}, \ldots, a_{k_n})$ and

$b_k{}' = \sum^{n} a_{k_j}$. Then

$$a_i = a_{1_i} + a_{2_i} \text{ and } b_i = b_{1_i} + b_{2_i}$$

for $i = 1, \ldots, n$. By induction for $k = 1.2$, $b_{k_i} \equiv 0 \pmod{r}$ or

$i > 1$ and $\left[\frac{b_{k_{i-1}}}{r}\right] > \left[\frac{b_{k_i}}{r}\right]$ and it is easy to check that the conditions of the theorem are valid for the sequence b_i.

Conversely, if the conditions of the proposition on the b_i sequence hold and $t = \{\max i \mid a_i \neq 0\}$ then either $a_t \geqslant r$ or $a_t < r$ and $a_t + a_{t-1} \geqslant r$. In the first case let A_1 be an $(n - rt) \times (n - rt)$ matrix with type $(a_1, a_2, .., a_{t-1}, a_t - r, a_{t+1}, ..., a_n)$ and A_2 an $rt \times rt$ matrix with type $(0, 0, ..., 0, r, 0, .., 0)$ where the entry r is in the tth position, while in the second case let A_1 be a matrix with type

$(a_1, a_2, ..., a_{t-2}, a_{t-1}, + a_t - r, 0, a_{t+1}, ..., a_n)$ and A_2 a matrix with type $(0, 0, ...$ $0, r - a_t, a_t\ 0, ..., 0)$. In either case A is similar to

$$\begin{pmatrix} A_1 & O \\ O & A_2 \end{pmatrix}.$$

If $A_2 = A$, then A has an irreducible solution by Proposition 2.2. Otherwise the b sequences for A_1 and A_2 satisfy the hypothesis of the proposition and so there exists B_1 and B_2 square matrices such that $f(B_k) = A_k$, $k = 1,2$. This implies that a matrix similar to

$$\begin{pmatrix} B_1 & O \\ O & B_2 \end{pmatrix}$$

gives the required result.

Corollary 2.7—Let $A \in M_n(\mathbb{C})$. A is diagonalizable if and only if each non-constant polynomial $f(x) \in \mathbb{C}[x]$ has a solution $B = B_f \in M_n(\mathbb{C})$ of $f(X) = A$.

Proof : If A is diagonalizable and $f(x) \in \mathbb{C}[x]$ with $f(x)$ of degree at least one, by Corollary 2.5, the equation $f(x) = A$ has a solution. If A is not diagonalizable, for some $\lambda \in \mathbb{C}$, $(x - \lambda)^2$ divides the minimal polynomial for A. By Proposition 2.6, if $\gamma \in \mathbb{C}$, the equation $(x - \gamma)^{n+1} + \lambda = A$ has no solution $B \in M_n(\mathbb{C})$.

Corollary 2.8—Suppose $f(x) \in \mathbb{C}[x]$ is a polynomial of degree 2 and A is an $n \times n$ matrix with λ its only eigenvalue and type $(a_1, a_2, ..., a_n)$. Let $b_i = \sum_{j=i}^{n} a_j$ for $i = 1, ..., n$. Then there is an $n \times n$ matrix B such that $f(B) = A$ if and only if either

(i) $f(x) - \lambda = 0$ has distinct roots or

(ii) $f(x) - \lambda = 0$ has a double root and for $i = 1, ..., n$, $b_i \equiv 0 \pmod 2$ or $i > 1$ and $\left[\frac{b_{i-1}}{2}\right] > \left[\frac{b_i}{2}\right]$.

Proof : If $f(x) = A$ has a solution and $f(x) - \lambda = 0$ does not have distinct roots, then $f(x) = a(x - \gamma)^2 + \lambda$ for $a, \gamma \in \mathbb{C}$ with $a \neq 0$. By Proposition 2.6 $b_i \equiv 0 \pmod 2$ or $i > 1$ and $\left[\frac{b_{i-1}}{2}\right] > \left[\frac{b_i}{2}\right]$ for $i = 1, ..., n$.

Conversely, if $f(x) - \lambda = 0$ has distinct roots then $f(X) = A$ has a solution by Corollary 2.4 while if condition (ii) is met, then there is a solution to $f(X) = A$ via Proposition 2.6.

Corollary 2.9—Suppose $f(x) \in \mathbb{C}[x]$ is a polynomial of degree 3 and A is an $n \times n$ matrix with λ its only eigenvalue and type $(a_1, a_2, \ldots, a_n)$. Let $b_i = \sum_{j=1}^{n} a_j$ for $i = 1, \ldots, n$. Then there is an $n \times n$ matrix B such that $f(B) = A$ if and only if

(i) $f(x) - \lambda = 0$ does not have a triple root, or

(ii) $f(x) - \lambda = 0$ has a triple root and for $i = 1, \ldots, n$,

$$b_i \equiv 0 \pmod 3) \text{ or } i > 1 \text{ and } \left[\frac{b_{i-1}}{3}\right] > \left[\frac{b_i}{3}\right].$$

Proof : If $f(x) = A$ has a solution and $f(x) - \lambda = 0$ does have a triple root, the Proposition 2.6 implies condition (ii) holds. Conversely, if $f(x) - \lambda = 0$ does not have a triple root, then there must be a root of $f(x) - \lambda = 0$ such that $f'(\gamma) \neq 0$. By Corollary 2.4, $f(X) = A$ then has a solution. If condition (ii) is met, then there is a solution to $f(X) = A$ by Proposition 2.6.

§3. As an illustration of the preceding results, we investigate in this section the composition of two polynomials modulo a given ideal.

Let I denote a non-zero ideal in $\mathbb{C}[x]$ which is generated by $m(x)$ a monic non-constant polynomial. Suppose $f(x), h(x) \in \mathbb{C}[x]$ with $f(x)$ of degree at least one. Does there exist $g(x) \in \mathbb{C}[x]$ such that $f(g(x)) \equiv h(x) \pmod I$? Two contrasting possibilities for $h(x)$ are considered, namely $h(x) = 0$ and $h(x) = x$.

Proposition 3.1—If $f(x)$ is a non-constant polynomial in $\mathbb{C}[x]$ and I is a proper ideal in $\mathbb{C}[x]$, then there exists $g(x) \in \mathbb{C}[x]$ such that $f(g(x)) \equiv O \pmod I$.

Proof : As $\mathbb{C}[x]$ is a principal ideal domain we can suppose I is generated by a monic polynomial $m(x)$. Let c be a root of $f(x) = 0$. If $s(x) \in \mathbb{C}[x]$, then $g(x) = s(x)\, m(x) + c$ satisfies the proposition.

Suppose now $m(x) = (x - \lambda_1)^{e_1} (x - \lambda_2)^{e_2} \ldots (x - \lambda_k)^{e_k}$ with $e_i > 0$ for $i = 1, \ldots, k$ and $\lambda_i \neq \lambda_j$ for $i \neq j$. Let $n = e_1 + e_2 + \ldots + e_k$. Finally let A be an $n \times n$ matrix with minimal polynomial $m(x)$. In terms of this notation we show

Lemma 3.2—There exists $g(x) \in \mathbb{C}[x]$ such that $f(g(x)) \equiv x \pmod I$ if and only if there exists $B \in M_n(\mathbb{C})$ such that $f(B) = A$.

Proof : Suppose there is a $B \in M_n(\mathbb{C})$ such that $f(B) = A$. Then A, being a polynomial in B, commutes with B. As the characteristic polynomial of A agrees with its minimal polynomial, A is non-derogatory and so there exists a $g(x) \in \mathbb{C}[x]$ such that $B = g(A)$ (see Marcus[2], p. 9), Theorem 2.14). $f(g(A)) - A = 0$ implies that $m(x)$ divides $f(g(x)) - x$ or $f(g(x)) \equiv x \pmod{\mathrm{I}}$.

Conversely, if $f(g(x)) \equiv x \pmod{\mathrm{I}}$ then $m(x)$ divides $(f(g(x)) - x)$ and $f(g(A)) - A = 0$. Let $B = g(A)$.

Proposition 3.3—Let $f(x)$ be a non-constant polynomial in $\mathbb{C}[x]$, and $m(x) = (x - \lambda_1)^{e_1} \ldots (x - \lambda_k)^{e_k}$ with $e_i > 0$ for $i = 1, \ldots, k$ and $\lambda_i \neq \lambda_j$ for $i \neq j$. Also let I be the ideal generated by $m(x)$. Then there exists $g(x) \in \mathbb{C}[x]$ such that $f(g(x)) \equiv x \pmod{\mathrm{I}}$ if and only if for each $i = 1, \ldots, k$ either $e_i = 1$ or there exists $\gamma_i \in \mathbb{C}$ such that $f(\gamma_i) = \lambda_i$ and $f'(\gamma_i) \neq 0$.

PROOF : A is similar to

$$\begin{bmatrix} A_1 & & & & O \\ & A_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ O & & & & A_R \end{bmatrix}$$

where

$$A_i = \begin{bmatrix} \lambda_i & 1 & & O \\ & \lambda_i & 1 & \\ & & \lambda_i & 1 \\ O & & & \lambda_i \end{bmatrix}$$

is an $e_i \times e_i$ matrix,

since A is non-derogatory. Lemma 3.2 together with Corollary 2.3 implies the result.

REFERENCES

1. Nathan Jacobson, *Lectures in Abstract Algebra* (vol. II). Princeton, N. J., Van Nostrand Company, 1953.
2. Marvin Marcus, 'Basic Theorems in Matrix Theory' *Applied Mathematics Series* 57. National Bureau of Standards, 1960.