ANNALES SOCIETATIS MATHEMATICAE POLONAE
Series I: COMMENTATIONES MATHEMATICAE XVIII (1974)
ROCZNIKI POLSKIEGO TOWARZYSTWA MATEMATYCZNEGO
Seria I: PRACE MATEMATYCZNE XVIII (1974)

BOGDAN LIS (Poznań)

Particular spectral theory in finite-dimensional spaces

A formula is derived for the matrix-polynomial of lowest degree equivalent to a matrix function, for the class $\mathscr{F}(\mathscr{C})$ of functions holomorphic in an open set containing the spectrum of a given square matrix $\mathscr{C}$. Some applications of the formula are deduced. By a matrix function $f(\mathscr{C})$ we mean a matrix assigned to matrix $\mathscr{C}$ by means of a function f from class $\mathscr{F}(\mathscr{C})$, considered as on operator. In the literature, within the framework of general spectral theory in finite-dimensional unitary spaces, a similar formula occurs; however, the polynomial appearing there is not of lowest degree for an arbitrary matrix $f(\mathscr{C})$.

Essential to the present work was a characteristic expansion of an arbitrary polynomial of the complex variable "z" in a neighbourhood of an arbitrary finite number of points. It will be our aim to find this expansion and to apply it appropriately.

LEMMA 1. *We now proceed to a generalization of the usual Vandermonde determinant.*

DEFINITION. We define a *generalized Vandermonde determinant of degree p* as a determinant of the form:

$$\left|K_{p_1}(z_1)\;\vdots\;K_{p_2}(z_2)\;\vdots\;\ldots\;\vdots\;K_{p_s}(z_s)\right|,$$

where $K_{p_i}(z_i)$ denotes the group of p_i columns of the matric of the form:

$$\begin{bmatrix} \binom{0}{0} & 0 & \ldots & 0 \\ \binom{1}{0} z_i & \binom{1}{1} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ \binom{p_i-1}{0} z_i^{p_i-1} & \binom{p_i-1}{1} z_i^{p_i-2} & \ldots & \binom{p_i-1}{p_i-1} \\ \ldots & \ldots & \ldots & \ldots \\ \binom{p-2}{0} z_i^{p-2} & \binom{p-2}{1} z_i^{p-3} & \ldots & \binom{p-2}{p_i-1} z_i^{p-p_i-1} \\ \binom{p-1}{0} z_i^{p-1} & \binom{p-1}{1} z_i^{p-2} & \ldots & \binom{p-1}{p_i-1} z_i^{p-p_i} \end{bmatrix},$$

where $p = \sum_{i=1}^{s} p_i$ and $p_i \in \mathscr{N}$.

For a determinant thus defined, the following formula holds:

$$|K_{p_1}(z_1)|K_{p_2}(z_2)|\ldots|K_{p_s}(z_s)| = \prod_{\substack{j,i=1\\ j>i}}^{s} (z_j-z_i)^{p_jp_i}.$$

Proof. The proof proceeds by complete induction with respect to the degree of the determinant.

For $p = 1$ the formula is obvious.

Let us assume it to be correct for $p-1$, i. e.

$$|K_{p_1}(z_1)|K_{p_2}(z_2)|\ldots|K_{p_s-1}(z_s)| = \prod_{\substack{j,i=1\\ j>i}}^{s-1} (z_j-z_i)^{p_jp_i} \prod_{i=1}^{s-1} (z_s-z_i)^{(p_s-1)p_i}.$$

We now consider the generalized Vandermonde determinant of degree p and transform it in the following manner: starting with $l = p$ and ending with $l = 1$, and applying the formulas

$$\binom{l}{k} z_s^{l-k} - \binom{l-1}{k} z_s^{l-k-1} z_s = \binom{l-1}{k-1} z_s^{l-k} \qquad \text{in } K_{p_s}(z_s),$$

$$\binom{l}{k} z_i^{l-k} - \binom{l-1}{k} z_i^{l-k-1} z_s = \binom{l-1}{k} (z_i-z_s) z_i^{l-k-1} + \binom{l-1}{k-1} z_i^{l-k} \qquad \text{in } K_{p_i}(z_i)$$

for $i \neq s$, we multiply its $(l-1)$-st row by "$-z_s$" and we add it to the l-th row. Then in the s-th group of columns, in the first column with the exception of the first row in which 1 appears, we now have only zeros. We apply the Laplace expansion with respect to the $(\sum_{i=1}^{r-1} p_i+1)$-th column. Thus, $K_{p_s}(z_s)$ goes over into $K_{p_s-1}(z_s)$. Similarly, each of the $s-1$ groups of columns of the new determinant is equal to $K_{p_i}(z_i)$; after transforming it in the following manner: starting from the first column and ending by the last but one, we take out the common factor (z_i-z_s), and we add the column as a whole, with opposite sign, to next column. In turn, we take out the common factor (z_i-z_s) from the last column. Thus, the factor $(z_i-z_s)^{p_i}$ is taken out from the group as a whole.

Hence, with regard to the equality:

$$(-1)^{\sum_{i=1}^{s-1} p_i+2} \prod_{i=1}^{s-1} (z_i-z_s)^{p_i} = \prod_{i=1}^{s-1} (z_s-z_i)^{p_i}$$

and the induction hypothesis, we obtain:

$$|K_{p_1}(z_1)|K_{p_2}(z_2)|\ldots|K_{p_s}(z_s)|$$

$$= \prod_{i=1}^{s-1} (z_s-z_i)^{p_i} \prod_{\substack{j,i=1\\ j>i}}^{s-1} (z_j-z_i)^{p_jp_i} \prod_{i=1}^{s-1} (z_s-z_i)^{(p_s-1)p_i} = \prod_{\substack{j,i=1\\ j>i}}^{s} (z_j-z_i)^{p_jp_i}.$$

Q.E.D.

In the sequel, $\mathscr{Z}$ will mean the open complex plane. Let $z_i \in \mathscr{Z}$ for $i = 1, 2, \ldots, s$ and $z_i \neq z_j$ for $i \neq j$. We define polynomials of the variable $z \in \mathscr{Z}$ by means of the formulas:

$$\mu(z) = \prod_{j=1}^{s} (z-z_j)^{p_j}, \qquad \nu_{i,k}(z) = \prod_{\substack{j=1\\ j\neq i}}^{s} (z-z_j)^{p_j}(z-z_i)^{k},$$

where $i = 1, 2, \ldots, s$, $k = 0, 1, \ldots, p_i - 1$.

We denote by $A_\nu(z)$ the determinant whose $\left(\sum_{j=1}^{i-1} p_j + k + 1\right)$-st column is given by the vector

$$\left[\nu_{i,k}(z), \left(\nu_{i,k}(z)\right)^{(1)}, \ldots, \left(\nu_{i,k}(z)\right)^{(p-1)}\right]^{tr}.$$

Then there holds the following

LEMMA 2. *For each $z \in \mathscr{Z}$ one has $A_\nu(z) \neq 0$.*

Proof. Assume $z \neq z_i$ for $i = 1, 2, \ldots, s$. Then, taking out the common factor $\mu(z)$ from each row of the determinant $A_\nu(z)$ and applying the Leibniz formula, its $\left(\sum_{j=1}^{i-1} p_j + k + 1\right)$-st column becomes:

$$\Bigg[(z-z_i)^{k-p_i}, \left((z-z_i)^{k-p_i}\right)^{(1)} + \frac{\mu^{(1)}(z)}{\mu(z)}(z-z_i)^{k-p_i}, \ldots, \left((z-z_i)^{k-p_i}\right)^{(p-1)} + \\ + \sum_{l=1}^{p-1} \binom{p-1}{l} \frac{\mu^{(l)}(z)}{\mu(z)} \left((z-z_i)^{k-p_i}\right)^{(p-1-l)}\Bigg]^{tr}.$$

Omitting in each row the linear combinations of preceding rows, it takes the form

$$\left[(z-z_i)^{k-p_i}, \left((z-z_i)^{k-p_i}\right)^{(1)}, \ldots, \left((z-z_i)^{k-p_i}\right)^{(p-1)}\right]^{tr}.$$

Taking out the common factor $(-1)^{l-1}(l-1)$ from the l-th row, the column under consideration becomes

$$\left[\binom{p_i-k-1}{0}(z-z_i)^{k-p_i}, \binom{p_i-k}{1}(z-z_i)^{k-p_i-1}, \ldots, \binom{p_i-k+p-2}{p-1}(z-z_i)^{k-p_i-p+1}\right]^{tr}.$$

Let $B_\nu(z)$ denote the determinant constructed by means of groups of columns of the form

$$\begin{bmatrix} \binom{0}{0}(z-z_i)^{-1} & \binom{1}{0}(z-z_i)^{-2} & \ldots & \binom{p_i-1}{0}(z-z_i)^{-p_i} \\ \binom{1}{1}(z-z_i)^{-2} & \binom{2}{1}(z-z_i)^{-3} & \ldots & \binom{p_i}{1}(z-z_i)^{-p_i-1} \\ \cdots & \cdots & \cdots & \cdots \\ \binom{p_i-1}{p_i-1}(z-z_i)^{-p_i} & \binom{p_i}{p_i-1}(z-z_i)^{-p_i-1} & \ldots & \binom{2p_i-2}{p_i-1}(z-z_i)^{-2p_i+1} \\ \cdots & \cdots & \cdots & \cdots \\ \binom{p-1}{p-1}(z-z_i)^{-p} & \binom{p}{p-1}(z-z_i)^{-p-1} & \ldots & \binom{p+p_i-2}{p-1}(z-z_i)^{-p-p_i+1} \end{bmatrix};$$

then

$$A_\nu(z) = \mu(z)^p\cdot(-1)^{\frac{1}{2}p(p-1)}(p-1)!(-1)^{\frac{1}{2}\sum\limits_{i=1}^{s}p_i(p_i-1)}B_\nu(z).$$

We transform the determinant $B_\nu(z)$ as follows: we begin by the last but one and we end on the first column of each group of columns, multiplying a given column by $-(z-z_i)^{-1}$ and adding to the next column. Proceeding similarly, stopping the procedure successively at the second, third, ... and finally the last but first column of a given group of columns, we finally obtain

$$\begin{bmatrix} \binom{0}{0}(z-z_i)^{-1} & 0 & \dots & 0 \\ \binom{1}{1}(z-z_i)^{-2} & \binom{1}{0}(z-z_i)^{-3} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \binom{p_i-1}{p_i-1}(z-z_i)^{-p_i} & \binom{p_i-1}{p_i-2}(z-z_i)^{-p_i-1} & \dots & \binom{p_i-1}{0}(z-z_i)^{-2p_i+1} \\ \dots & \dots & \dots & \dots \\ \binom{p-1}{p-1}(z-z_i)^{-p} & \binom{p-1}{p-2}(z-z_i)^{-p-1} & \dots & \binom{p-1}{p-p_i}(z-z_i)^{-p-p_i+1} \end{bmatrix}.$$

Taking out the common factor $(z-z_i)^{-p_i^2}$, we obtain the group of columns $K_{p_i}\big((z-z_i)^{-1}\big)$. With regard to Lemma 1, we obtain

$$B_\nu(z) = \prod_{i=1}^{s}(z-z_i)^{-p_i}\cdot\prod_{\substack{i,j=1\\ j>i}}^{s}(z-z_j)^{-p_jp_i}(z-z_i)^{-p_jp_i}(z_i-z_j)^{p_jp_i}.$$

Thus, after insertion into $A_\gamma(z)$ we have

$$A_\nu(z) = (p-1)!\prod_{\substack{i,j=1\\ j>i}}^{s}(z_j-z_i)^{p_jp_i}.$$

Since $A_\nu(z)$ is continuous, $A_\nu(z) = \mathrm{const} \neq 0$.

LEMMA 3. *For any polynomial* w_n $(n = 0, 1, \dots)$, *there always exists exactly one solution of the set of linear equations of the form*

$$w_n(z) \equiv \sum_{h=0}^{l}\mu^h(z)\sum_{\substack{i=1,2,\dots,s\\ k=0,1,\dots,p_i-1}}X_{i,k}^{h,n}\nu_{i,k}(z), \tag{1}$$

where l *is chosen in such a manner that* $n < p(l+1)$, $p = \sum\limits_{i=1}^{s}p_i$, *where*

$$X_{i,k}^{0,n} = \frac{1}{k!}\left\{\frac{w_n(z)}{\nu_{i,0}(z)}\right\}^{(k)}(z_i). \tag{2}$$

Note. By $\{\cdot\}^{(k)}(z_i)$ we understand the derivative of order k calculated at the point z_i.

Proof. We begin by proving that the set of solutions of (1) is not empty.

From Euclids' algorithm, we have

$$w_n(z) \equiv \sum_{h=0}^{l} \mu^{(h)}(z)\vartheta_{h,n}(z), \tag{3}$$

where $\vartheta_{h,n}$ is a polynomial, and degree $\vartheta_{h,n} < p$. Moreover, degree $\vartheta_{l,n} = n - lp < p$, i. e. $n < p(l+1)$.

We now compare (3) and (1). This system is satisfied if, for $h = 0, 1, \ldots \ldots, l$

$$\vartheta_{h,n}(z) = \sum_{\substack{i=1,2,\ldots,s \\ k=0,1,\ldots,p_i-1}} X_{i,k}^{h,n} v_{i,k}(z).$$

Keeping h fixed, this identity is fulfilled if

$$\{\vartheta_{h,n}(z)\}^{(j)} = \sum_{\substack{i=1,2,\ldots,s \\ k=0,1,\ldots,p_i-1}} X_{i,k}^{h,n}\{v_{i,k}(z)\}^{(j)} \quad \text{for } j = 0, 1, \ldots, p-1.$$

Since $A_\nu(z) \neq 0$ for all $z \in \mathscr{X}$, it is always possible to choose constants $X_{i,k}^{h,n}$ in a manner to fulfil the latter condition.

We now proceed to show by complete induction with respect to k at fixed i that the first p unknowns are of the form (2). For $k = 0$, we have by (1)

$$w_n(z_i) = X_{i,0}^{0,n} v_{i,0}(z_i), \qquad \text{i. e.} \qquad X_{i,0}^{0,n} = \frac{1}{0!}\left\{\frac{w_n(z)}{v_{i,0}(z)}\right\}^{(0)}(z_i).$$

Now we show that if the thesis true for $q_i - 1$, then it is also true for q_i, where $q_i < p_i$.

We note that the system (1) can be rewritten as:

$$w_n(z) = \sum_{k=0}^{q_i} X_{i,k}^{0,n} v_{i,k}(z) + (z - z_i)^{q_i+1}\vartheta(z),$$

where $\vartheta(z)$ is a polynomial.

Consequently,

$$X_{i,q_i}^{0,n}\{v_{i,q_i}(z)\}^{q_i}(z_i) = \{w_n(z)\}^{(q_i)}(z_i) - \left\{\sum_{k=0}^{q_i-1} X_{i,k}^{0,n} v_{i,k}(z)\right\}^{(q_i)}(z_i).$$

With regard to the relation

$$v_{i,k}(z) = v_{i,0}(z)\cdot(z - z_i)^k, \tag{4}$$

we have, on the one hand,

$$\{v_{i,k}(z)\}^{(q_i)}(z_i) = q!\,v_{i,0}(z_i) \neq 0$$

and, on the other (by the Leibniz formula, dropping zero terms)

$$\Big\{\sum_{k=0}^{q_i-1} X_{i,k}^{0,n} v_{i,k}(z)\Big\}^{(q_i)}(z_i) = \sum_{l=0}^{q_i-1} \binom{q_i}{l} \{v_{i,0}(z)\}^{(q_i-l)}(z_i)\cdot X_{i,l}^{0,n}\cdot l!.$$

Applying now the induction hypothesis and then the Leibniz formula, we obtain

$$q_i!\; v_{i,0}(z_i) X_{i,q_i}^{0,n} = \{w_n(z)\}^{(q_i)}(z_i) - \sum_{l=1}^{q_i-1} \binom{q_i}{l} \{v_{i,0}(z)\}^{(q_i-l)}(z_i)\cdot \left\{\frac{w_n(z)}{v_{i,0}(z)}\right\}^{(l)}(z_i)$$
$$= v_{i,0}(z_i)\left\{\frac{w_n(z)}{v_{i,0}(z)}\right\}^{(q_i)}(z_i).$$

This proves the lemma.

At the same time, this proves the uniqueness of the existence of the first p solutions. That of the remaining ones can be proved similarly. Q.E.D.

It should be noted that, with regard to the preceding lemma, the set of polynomials $v_{i,k}(z)$ is linearly independent.

Now let $\mu(z)$ stand for the minimal zeroing polynomial of a square matrix $\mathscr{C}$ of degree n, and $\sigma(\mathscr{C})$ for its spectrum. By the Cayley–Hamilton theorem, the inequality: degree $\mu \leqslant n$ holds (cf. [2], p. 270). Hence, Lemma 3 leads to the following theorem:

THEOREM 1. *If f denotes an arbitrary polynomial, considered as an operator acting on the matrix $\mathscr{C}$, then*

$$f(\mathscr{C}) = \sum_{\substack{i=1,2,\ldots,s\\ k=0,1,\ldots,p-1}} \frac{1}{k!}\left\{\frac{f(z)}{v_{i,0}(z)}\right\}^{(k)}(z_i)\, v_{i,k}(\mathscr{C}), \tag{5}$$

where $z_i \in \sigma(\mathscr{C})$ and, if there exist p matrices $\mathscr{C}_{i,k}$ such that there holds

$$f(\mathscr{C}) = \sum_{\substack{i=1,2,\ldots,s\\ k=0,1,\ldots,p_i-1}} \frac{1}{k!}\left\{\frac{f(z)}{v_{i,0}(z)}\right\}^{(k)}(z_i)\,\mathscr{C}_{i,k} \tag{6}$$

for an arbitrary polynomial f, then $\mathscr{C}_{i,k} = v_{i,k}(\mathscr{C})$.

Proof. Equation (5) results from (1) on omitting terms equal to the matrix $\mathscr{C}$. Now, with regard to (4), one has:

$$\frac{1}{l!}\left\{\frac{v_{i,k}(z)}{v_{j,0}(z)}\right\}^{(l)}(z_j) = \delta_{i,j}\,\delta_{k,l}$$

whence, on putting $f = v_{i,k}$ in (6), we have $v_{i,k}(\mathscr{C}) = \mathscr{C}_{i,k}$. Q.E.D.

LEMMA 4. *If $\mathscr{C}$ is a non-singular matrix, then*

$$\mathscr{C}^{-1} = \sum_{\substack{i=1,2,\ldots,s\\ k=0,1,\ldots,p_i-1}} \frac{1}{k!}\left\{\frac{z^{-1}}{\nu_{i,0}(z)}\right\}^{(k)}(z_i)\nu_{i,k}(\mathscr{C}), \quad \textit{where } z_i \in \sigma(\mathscr{C}).$$

Proof. Let $\mu^*(z) = \prod_{i=1}^{s}(z-z_i^{-1})^{p_i}$. Then, by the equality

$$\mu(\mathscr{C}) = (-1)^p \mathscr{C}^p \prod_{i=1}^{s} z_i^{p_i}\cdot\mu^*(\mathscr{C}^{-1}),$$

$\mu^*(z)$ is the minimal polynomial zeroing the matrix $\mathscr{C}^{-1}$. We introduce the following definition:

$$\nu_{i,k}^*(z) = \prod_{\substack{j=1\\ j\neq i}}^{s}(z-z_j^{-1})^{p_j}\cdot(z-z_i^{-1})^k.$$

Note that $z^p\nu_{i,k}^*(z^{-1})$ is a polynomial in the variable z. Thus, by Theorem 1, the matrix $\mathscr{C}^p\nu_{i,k}^*(\mathscr{C}^{-1})$ is a linear combination of matrices $\nu_{i,k}(\mathscr{C})$. Also, by Theorem 1, one has the equality

$$(\mathscr{C}^{-1})^{p+1} = \sum_{\substack{i=1,2,\ldots,s\\ k=0,1,\ldots,p_i-1}} \frac{1}{k!}\left\{\frac{z^{p+1}}{\nu_{i,0}^*(z)}\right\}^{(k)}(z_i^{-1})\nu_{i,k}^*(\mathscr{C}^{-1}).$$

Consequently, $\mathscr{C}^{-1}$ is a linear combination of matrices $\nu_{i,k}(\mathscr{C})$. Hence, if $\vartheta(z)$ is a polynomial corresponding to this linear combination, then the function $f(z) = z(z^{-1}-\vartheta(z))$ is a polynomial with the property $f(\mathscr{C}) = \vartheta$. As a consequence, $z^{-1}-\vartheta(z) = \mu(z)\chi(z)z^{-1}$, where $\chi(z)$ is a polynomial, and $\{z^{-1}-\vartheta(z)\}^{(k)}(z_i) = 0$ for $i = 1, 2, \ldots, s$, $k = 0, 1, \ldots, p_i-1$, $z_i \in \sigma(\mathscr{C})$. Further steps of the proof are identical with those of (2) in Lemma 3. Q.E.D.

LEMMA 5. *For any $\lambda \in \sigma(\mathscr{C})'$, the following formula holds:*

$$(\lambda I-\mathscr{C})^{-1} = \sum_{\substack{i=1,2,\ldots,s\\ k=0,1,\ldots,p_i-1}} \frac{1}{k!}\left\{\frac{(\lambda-z)^{-1}}{\nu_{i,0}(z)}\right\}^{(k)}(z_i)\nu_{i,k}(\mathscr{C}), \quad \textit{where } z_i \in \sigma(\mathscr{C}). \tag{7}$$

Proof. Note that $\mu^*(z) = \prod_{i=1}^{s}(z-z_i+\lambda)^{p_i}$ is the minimal polynomial zeroing the matrix $\mathscr{C}-\lambda I$. We use the definition

$$\nu_{i,k}^*(z) = \prod_{\substack{j=1\\ j\neq i}}^{s}(z-z_j+\lambda)^{p_j}\cdot(z-z_i+\lambda)^k.$$

We now have, by Lemma 4,

$$(\mathscr{C}-\lambda I)^{-1} = \sum_{\substack{i=1,2,\ldots,s\\ k=0,1,\ldots,p_i-1}} \frac{1}{k!}\left\{\frac{z^{-1}}{\nu_{i,0}^*(z)}\right\}^{(k)}(z_i-\lambda)\nu_{i,k}^*(\mathscr{C}-\lambda\mathscr{I}).$$

Since $v^*_{i,k}(z) = v_{i,k}(z+\lambda)$ and $\{(\cdot)(z)\}^{(k)}(z_i-\lambda) = \{(\cdot)(z-\lambda)\}^{(k)}(z_i)$, we obtain formula (7) which is equivalent to the thesis of this lemma. Q.E.D.

Let $\mathscr{F}(\mathscr{C})$ denote the class of functions holomorphic in an open set containing $\sigma(\mathscr{C})$. If $f\in\mathscr{F}(\mathscr{C})$, we define:

$$f(\mathscr{C}) = \frac{1}{2\pi i}\int_{\Gamma}\frac{f(\lambda)}{\lambda I-\mathscr{C}}\,d\lambda, \tag{8}$$

where Γ is a contour consisting of a finite number of curves Γ_j, enclosing $\sigma(\mathscr{C})$ and satisfying the assumptions needed for the Cauchy formula. This definition is of use in the following theorem.

THEOREM 2. *If $f\in\mathscr{F}(\mathscr{C})$, then the matrix $f(\mathscr{C})$ is uniquely determined by the formula*:

$$f(\mathscr{C}) = \sum_{\substack{j=1,2,\ldots,s\\ k=0,1,\ldots,p_j-1}}\frac{1}{k!}\left\{\frac{f(z)}{v_{j,0}(z)}\right\}^{(k)}(z_j)v_{j,k}(\mathscr{C}), \quad \textit{where } z_j\in\sigma(\mathscr{C}). \tag{9}$$

Proof. By Lemma 5, the right-hand term of equation (8) is an integral of a matrix polynomial. We note that, with regard to the uniform convergence of the integrals, the following equality holds:

$$\frac{1}{2\pi i}\int_{\Gamma}\left\{\frac{(\lambda-z)^{-1}f(\lambda)}{v_{j,0}(z)}\right\}^{(k)}(z_j)\,d\lambda = \left\{\frac{f(z)}{v_{j,0}(z)}\right\}^{(k)}(z_j),$$

its right-hand term being independent of the contour Γ. Hence, on placing the sum before the integration sign, we obtain (9). Uniqueness of this representation follows from the linear independence of the system of matrices $v_{j,k}(\mathscr{C})$, $j = 1, 2, \ldots, s$, $k = 0, 1, \ldots, p_j-1$, immediately.

THEOREM 3. *If $f\in\mathscr{F}(\mathscr{C})$, then equation* (9) *holds, if and only if, there exists a polynomial $g(z)$ such that*

$$f(\mathscr{C}) = g(\mathscr{C}) \quad \textit{and} \quad f^{(k)}(z_i) = g^{(k)}(z_i) \tag{10}$$

$$\textit{for } k = 0, 1, \ldots, p_i-1,\ z_i\in\sigma(\mathscr{C}).$$

If, moreover, degree of g is $\leqslant p$, then the polynomial g is determined uniquely.

Proof. Sufficiency of condition (10) follows from Theorem 2. In order to prove necessity of (10), it suffices to take

$$g(z) = \sum_{\substack{i=1,2,\ldots,s\\ k=0,1,\ldots,p_i-1}}\frac{1}{k!}\left\{\frac{f(z)}{v_{i,0}(z)}\right\}^{(k)}(z_i)v_{i,k}(z).$$

Then, $f(\mathscr{C}) = g(\mathscr{C})$, and formula (9) holds. Since the representation is unique, we get

$$\left\{\frac{f(z)}{\nu_{i,0}(z)}\right\}^{(k)}(z_i) = \left\{\frac{g(z)}{\nu_{i,0}(z)}\right\}^{(k)}(z_i) \quad \text{for } k = 0, 1, \ldots, p_i-1,\ z_i \in \sigma(\mathscr{C}).$$

From the Leibniz formula we obtain after a rearrangement

$$\sum_{l=0}^{k} \binom{k}{l} \{\nu_{i,0}(z)^{-1}\}^{(k-l)}(z_i)\{f(z)-g(z)\}^{(l)}(z_i) = 0$$

for $k = 0, 1, \ldots, p_i - 1$, $z_i \in \sigma(\mathscr{C})$. Since for fixed i the determinant of this system is $\nu_{i,0}(z_i)^{-p_i} \neq 0$, relations (10) are necessary. Q.E.D.

We now proceed to formulate theorems concerning the minimal zeroing polynomial.

THEOREM 4. *Let* $\mu(z) = \prod_{i=1}^{s} (z-z_i)^{p_i}$ *be the minimal polynomial zeroing the matrix* $\mathscr{C}$ *and let* $f \in \mathscr{F}(\mathscr{C})$. *If* f *possesses the following properties:*

1° *Let* Q_l *be the set of those* i*-s corresponding to* $z_i \in \sigma(\mathscr{C})$, *for which the function* f *takes the same value (denoted by* u_l*) and* r *is the number of all distinct numbers* u_l,

2° *the point* $z_i \in \sigma(\mathscr{C})$ *is an* r_i*-fold point of the function* f, *then the minimal polynomial zeroing the matrix* $f(\mathscr{C})$ *is of the form:*

$$\mu^f(z) = \prod_{l=1}^{r} (z-u_l)_{q_l}, \quad \textit{where } q_l = \min\{q \in \mathscr{N}\colon\ r_i q \geqslant p_i,\ i \in Q_l\}.$$

Proof. The function f can by dealt with as a polynomial. Since, in neighbourhood of an r_i-fold point z_i, the function f can be represented in the form: $f(z) = u_l + (z-z_i)^{r_i} g_i(z)$, where $g_i(z_i) \neq 0$, and $r_i q_i \geqslant p_i > k$, by applying equation (9) to the polynomial $\prod_{m=1}^{r} [f(z)-u_m]^{q_m}$ it is seen that the coefficient by the matrix $\nu_{i,k}(\mathscr{C})$ is equal to

$$\frac{1}{k!}\left\{(z-z_i)^{r_i q_l} g_i(z)^{q_l} \nu_{i,0}(z)^{-1} \prod_{\substack{m=1\\ m\neq l}}^{r} [f(z)-u_m]^{q_m}\right\}^{(k)}(z_i) = 0.$$

Hence, $\prod_{m=1}^{r} [f(\mathscr{C}) - u_m I]^{q_m} = \mathscr{O}$.

If $\mu^f(z)$ denotes the minimal polynomial zeroing the matrix $f(\mathscr{C})$, then $\mu^f(z) \mid \prod_{m=1}^{r} (z-u_m)^{q_m}$. Since the matrix $f(\mathscr{C})$ is not zero for the polynomial

$\prod\limits_{\substack{m=1\\ m\neq l}}^{r}(z-u_m)^{q_m}(z-u_l)^{q_l-1}$, because for $h = r_i(q_l-1) \leqslant (p_i-1)$ the coefficient by the matrix $\nu_{i,h}(\mathscr{C})$ is

$$g_i(z_i)^{q_l-1}\nu_{i,0}(z_i)^{-1}\prod_{\substack{m=1\\ m\neq l}}^{r}(u-u_m)^{q_m}\neq 0,$$

we have $\mu^f(z) = \prod\limits_{m=1}^{r}(z-u_m)^{q_m}$. Q.E.D.

The last theorem leads to the following corollaries:

COROLLARY 1. *If* $f\in\mathscr{F}(\mathscr{C})$, *then* degree $\mu^f =$ degree μ, *if and only if*,

1. *the function* f *is one-to-one on* $\sigma(\mathscr{C})$,
2. *for each* $z_i\in\sigma(\mathscr{C})$ *for which* $p_i\geqslant 2$, $f'(z_i)\neq 0$.

COROLLARY 2. *If* $f\in\mathscr{F}(\mathscr{C})$, *then* degree $\mu^f\leqslant$ degree μ.

Moreover, the following theorem holds:

THEOREM 5 *If* $f\in\mathscr{F}(\mathscr{C})$, *there exists a function* $g\in\mathscr{F}\big(f(\mathscr{C})\big)$ *of the property* $(g\circ f)(\mathscr{C}) = \mathscr{C}$ *if and only if*

$$\text{degree } \mu^f = \text{degree } \mu.$$

Proof. Since, by Corollary 2, we have

$$\text{degree}\,\mu = \text{degree}\,\mu^{g\circ f}\leqslant \text{degree}\,\mu^f\leqslant \text{degree}\,\mu,$$

it results that the condition degree$\,\mu^f =$ degree$\,\mu$ is necessary. Inversely, from Corollary 1 and Theorem 3 follows that the function f can be considered as a polynomial with the property $f'(z_i)\neq 0$ for $i = 1, 2, \ldots, s$. Then from the fact that zeros of a holomorphic are isolated, inverse function exists in some neighbourhood of $\sigma\big(f(\mathscr{C})\big)$. Q.E.D.

Assume r, q_l, u_l, Q_l as having the same meaning as in Theorem 4.

THEOREM 6. *If* $f\in\mathscr{F}(\mathscr{C})$ *and if* $|\mathscr{C}-zI| = (-1)^n\prod\limits_{i=1}^{s}(z-z_i)^{a_i}$ *denotes the characteristic polynomial of matrix* $\mathscr{C}$, *then the characteristic polynomial of the matrix* $f(\mathscr{C})$ *is of the form*

$$|f(\mathscr{C})-zI| = (-1)^n\prod_{l=1}^{r}(z-u_l)^{\beta_l},\quad \textit{where } \beta_l = \sum_{i\in Q_l}a_i \textit{ and } \sum_{l=1}^{r}\beta_l = n.$$

Proof. The function f can be considered as a polynomial. Applying equation (9) to the function $[f(z)-u_l]^{q_l}$ and making use of the relations

$$\sum_{\substack{i\in Q_l\\ k=0,1,\ldots,p_i-1}}\frac{1}{k!}\{\nu_{i,0}(z)^{-1}[f(z)-u_l]^{q_l}\}^{(k)}(z_i)\,\nu_{i,k}(\mathscr{C}) = \mathscr{O},$$

$v_{i,k}(\mathscr{C}) = \mathscr{A}\cdot\prod_{j\in Q_l}(\mathscr{C}-z_jI)^{p_j}$ for $i\notin Q_l$, where $\mathscr{A}$ is a matrix, we obtain

$$[f(\mathscr{C})-u_lI]^{q_l} = \mathscr{B}\cdot\prod_{i\in Q_l}(\mathscr{C}-z_iI)^{p_i},$$

where $\mathscr{B}$ is a matrix. Hence any vector which zeroes the matrix $(\mathscr{C}-z_iI)^{p_i}$, where $i\in Q_l$, also zeroes the matrix $[f(\mathscr{C})-u_lI]^{q_l}$. By a theorem of [2] (p. H, p. 273), the multiplicity of the eigenvalue u_l which is β_l, fulfils the inequality $\beta_l \geqslant \sum_{i\in Q_l}\alpha_i$. Since, moreover, the inequality $n = \sum_{l=1}^{r}\beta_l \geqslant \sum_{l=1}^{r}\sum_{i\in Q_l}\alpha_i = \sum_{i=1}^{s}\alpha_i = n$ holds, we obtain $\beta_l = \sum_{i\in Q_l}\alpha_i$, proving the theorem.

References

[1] Nelson Dunford and Jacob T. Schwartz, *Linear operators*, Part I (1958).

[2] L. S. Pontriagin, *Równania różniczkowe zwyczajne*, PWN, Warszawa 1964 (in Polish).