*Journal of*
*Linear and Topological Algebra*
*Vol.* 07*, No.* 01*,* 2018*,* 21- 38

# Transformations computations: power, roots and inverse

M. Ghorbanzadeh[a], H. Kamali[b,*], A. Mansoori[c]

[a]*Department of Basic Science, Imam Reza International University, Mashhad, Iran.*
[b]*Department of Pure Mathematics, Ferdowsi University of Mashhad, Mashhad, Iran.*
[c]*Department of Applied Mathematics, Ferdowsi University of Mashhad, Mashhad, Iran.*

Received 23 February 2017; Revised 17 December 2017; Accepted 10 February 2018.

Communicated by Hamidreza Rahimi

**Abstract.** This paper presents some results of an annihilated element in Banach algebra, and in specific case, for any square matrix. The developed method significantly improves the computational aspects of transformations calculus and especially for finding powers and roots of any annihilated element. An example is given to compare the proposed method with some other methods to show the efficiency and performance.



**Keywords:** Annihilated elements, matrix power, inverse matrix, matrix roots.

**2010 AMS Subject Classification**: 15A24, 65F60.

## 1. Introduction

Many problems in pure and applied mathematics can be solved by computing powers or roots of some square matrix $A$; for example:
(a) Differential equations and matrix exponential: the classic scalar problem $y'(t) = ay(t)$ with initial condition $y(0) = c$, has solution $y(t) = ce^t$, while the analogous vector problem

$$\frac{dY}{dt} = AY\,, \qquad Y(0) = C\,, \quad Y, C \in \mathbb{C}^n\ ,\ A \in M_n(\mathbb{C})$$

*Corresponding author.
E-mail address: ghorbanzadeh_imamreza@yahoo.com (M. Ghorbanzadeh); kamalihamid2@gmail.com (H. Kamali); am.ma7676@yahoo.com & a-mansoori@um.ac.ir (A. Mansoori).

Print ISSN: 2252-0201
Online ISSN: 2345-5934

http://jlta.iauctb.ac.ir

has solution $Y(t) = e^{At}C$, where $e^{At} = \sum_{j=0}^{\infty}(At)^j/j!$. More generally, with suitable assumptions on the smoothness of $f$, the solution to the inhomogeneous system

$$\frac{dY}{dt} = AY + f(t,Y)\,, \qquad Y(0) = C\,, \quad Y, C \in \mathbb{C}^n\ ,\ A \in M_n(\mathbb{C})$$

satisfies

$$Y(t) = e^{At}C + \int_0^t e^{A(t-s)} f(s,Y)ds$$

The matrix exponential is explicitly used in certain methods such as exponential integrators, Exponential Time Difference (ETD), Nuclear Magnetic Resonance (NMR), Markov models, Control theory ([2], Ch. 2, pp. 35-44) and continuous linear dynamical system ([9], Ch. 3). However, computing matrix exponential is depends on powers of the matrix.
(b) Trigonometric matrix functions: the matrix roots, arise in the solution of second order differential equations. For example, the problem

$$\frac{d^2Y}{dt^2} + AY = 0\ , \qquad Y(0) = C_1\ , Y'(0) = C_2 \qquad Y, C_1, C_2 \in \mathbb{C}^n\ ,\ A \in M_n(\mathbb{C})$$

has solution

$$Y(t) = \cos(\sqrt{A}t)C_1 + A^{-1/2}\sin(\sqrt{A}t)C_2$$

where $\sqrt{A}$ denotes any square root of $A$. The solution exists for all $A$. When $A$ is singular (and $\sqrt{A}$ and $A^{-1/2}$ possibly does not exist) this formula is interpreted by expanding $\cos(\sqrt{A}t)$ and $A^{-1/2}\sin(\sqrt{A}t)$ as power series in $A$. The matrix roots also are the fundamental tool for the algebraic Riccati equations and semidefinite programming ([2], Ch. 2, p. 45).
(c) Discrete linear dynamical system: linear nonsingular transformation

$$A : \mathbb{C}^n \longrightarrow \mathbb{C}^n\ , \qquad x \longmapsto Ax$$

generates a variety of dynamical systems. The operator $A$ itself generates a dynamical system acting on $\mathbb{C}^n$ by the powers $A^m$ and one gets the problem of investigating the trajectory of an arbitrary vector $x$ by the action of the linear transformation, i.e. the description of the behavior of the sequence of vectors of the form $A^m x$. In addition, linear operator $A$ generates dynamical systems on a string of other spaces related to $\mathbb{C}^n$. Hartman-Grobman's theorem explains the importance of linear systems: every diffeomorphism or vector field locally conjugate to its linear part (i.e. it's differential) at a hyperbolic fixed point ([5], Theorem 4.1, p. 60).
(d) Combinatorics and graph theory: consider the linear system of difference equations

$$\begin{cases} x_1(m+1) = a_{1,1}x_1(m) + a_{1,2}x_2(m) + \cdots + a_{1,n}x_n(m) \\ x_2(m+1) = a_{2,1}x_1(m) + a_{2,2}x_2(m) + \cdots + a_{2,n}x_n(m) \\ \quad \vdots \qquad\qquad\quad \vdots \qquad\qquad\quad \vdots \qquad\qquad\qquad\quad \vdots \\ x_n(m+1) = a_{n,1}x_1(m) + a_{n,2}x_2(m) + \cdots + a_{n,n}x_n(m) \end{cases}$$

This system may be written in the vector form $X(m+1) = AX(m)$, where

$$X(m) = [x_1(m), x_2(m), \cdots, x_n(m)]^T \in \mathbb{R}^n$$

and $A = (a_{i,j})$. In graph theory and computer science, an adjacency matrix is a square matrix used to represent a finite graph. The elements of the matrix indicate whether pairs of vertices are adjacent or not. If $A$ is the adjacency matrix of the directed or undirected graph $G$, then the matrix $A^n$ has an interesting interpretation: the element $(i, j)$ of $A^n$ gives the number of (directed or undirected) walks of length $n$ from vertex $i$ to vertex $j$. If $n$ is the smallest nonnegative integer, such that for some $i, j$, the element $(i, j)$ of $A^n$ is positive, then $n$ is the distance between vertex $i$ and vertex $j$. This implies, for example, that the number of triangles in an undirected graph $G$ is exactly the trace of $A^3$ divided by 6.

This examples show the importance of matrix powers and roots. The aim of this work is to presente a method for computing the powers and roots of any annihilated element in a Banach algebra, and in particular, for square matrices. There are many methods for computing $A^m$, and $A^{1/p}$ for $m, p \in \mathbb{N}$ (or $m, p \in \mathbb{Z}$ when $A$ is invertible) and square matrix $A$, but there are not practical approaches. By the way, some of these methods are efficient. We present a method that can be accurate and exact when the roots of minimal polynomial and their repeated order are known. Also, we introduce the Jordan and Hermite interpolation methods and compare this methods by our own way.

## 2. Preliminaries

Throughout the paper, $\mathcal{A}$ shows a Banach algebra with unity $e$, and $\varpi$ is an element of $\mathcal{A}$. If $f(z) = a_0 + a_1 z + \cdots + a_n z^n$ is a polynomial with $a_i \in \mathbb{C}$, there is not any doubt about the meaning of the symbol $f(\varpi)$; it obviously denotes the element of $\mathcal{A}$ defined by $f(\varpi) = a_0 e + a_1 \varpi + \cdots + a_n \varpi^n$. The question arises whether $f(\varpi)$ can be defined in a meaningful way for other functions $f$? In fact, if $f(z) = \sum_{j=0}^{\infty} a_j z^j$, is any entire function in $\mathbb{C}$, it is natural to define $f(\varpi) \in \mathcal{A}$ by $f(\varpi) = \sum_{j=0}^{\infty} a_j \varpi^j$; this series always converges. Another example is given by the meromorphic functions $f(z) = 1/(\alpha - z)$. In this case, the natural definition is $f(\varpi) = (\alpha e - \varpi)^{-1}$; which makes sense for all $\varpi$ whose spectrum does not contain $\alpha$.

One is thus led to the conjecture that $f(\varpi)$ should be definable, within $\mathcal{A}$, whenever $f$ is holomorphic in an open set that contains $\mathrm{Spec}(\varpi)$. This turns out to be correct and can be accomplished by a version of the Cauchy formula that converts complex functions defined in open subsets of $\mathbb{C}$ to $\mathcal{A}$-valued ones defined in certain open subsets of $\mathcal{A}$. This motivates the following definition.

**Definition 2.1** For an open set $\Omega \subset \mathbb{C}$, let $H(\Omega)$ denotes the algebra of all holomorphic functions on $\Omega$ in the complex plane. Then $\mathcal{A}_\Omega = \{\varpi \in \mathcal{A}\ ;\ Spec(\varpi) \subset \Omega\}$ is an open subset of $\mathcal{A}$ (Theorem 10.20 of [7]). Define $\widetilde{H}(\mathcal{A}_\Omega)$ to be the set of all $\mathcal{A}$-valued functions $\widetilde{f}$ with domain $\mathcal{A}_\Omega$, that arise from an $f \in H(\Omega)$ by the formula

$$\widetilde{f}(\varpi) = \frac{1}{2\pi i} \int_\Gamma f(z)(ze - \varpi)^{-1} dz \tag{1}$$

where $\Gamma$ is a closed contour that encloses the spectrum, $Spec(\varpi)$.

Definition 2.1 calls for some comments.

(a) Since inversion is continuous in $\mathcal{A}$ (Theorem 10.12 [7]), and $\Gamma$ stays away from $\mathrm{Spec}(\varpi)$, the integral in (1) is continuous; So that the integral exists and defines $\widetilde{f}(\varpi)$ as an element of $\mathcal{A}$. The integrand is actually a holomorphic $\mathcal{A}$-valued function in the

complement of Spec($\varpi$). The Cauchy Integral Theorem implies that $\widetilde{f}(\varpi)$ is independent of the choice of $\Gamma$, provided only that $\Gamma$ surrounds Spec($\varpi$) in $\Omega$.
(b) If $\varpi = \lambda e$ for some $\lambda \in \Omega$, the equation (1) becomes $\widetilde{f}(\varpi) = f(\lambda)e$. Note that $\lambda \in \Omega$ if and only if $\lambda e \in \mathcal{A}_\Omega$. If we identify $\lambda \in \mathbb{C}$ with $\lambda e \in \mathcal{A}$, every $f \in H(\Omega)$ may be regarded as mapping a certain subset of $\mathcal{A}_\Omega$ (namely, the intersection of $\mathcal{A}_\Omega$ with the one-dimensional subspace of $\mathcal{A}$ generated by $e$) into $\mathcal{A}$, and then, the equality $\widetilde{f}(\varpi) = f(\lambda)e$ shows that $\widetilde{f}$ may be regarded as an extension of $f$. The mapping $f \longrightarrow \widetilde{f}$ is an algebra isomorphism of $H(\Omega)$ onto $\widetilde{H}(\mathcal{A}_\Omega)$ which is continuous in the following sense:

**Theorem 2.2** ([7], Theorem 10.27) If $\{f_n\}_{n=1}^{\infty} \subset H(\Omega)$ and $f_n \longrightarrow f$ uniformly on compact subsets of $\Omega$, then for every $\varpi \in \mathcal{A}_\Omega$, $\widetilde{f}(\varpi) = \lim_{n\to\infty} \widetilde{f}_n(\varpi)$. Moreover, if $g(z) = z$ and $h(z) = 1$ in $\Omega$, then $\widetilde{g}(\varpi) = \varpi$ and $\widetilde{h}(\varpi) = 1$, in $\mathcal{A}_\Omega$.

Also, there is a relation between the spectral of $\varpi$ and $\widetilde{f}(\varpi)$ as follow:

**Theorem 2.3** ([7], Theorem 10.28) Suppose $\varpi \in \mathcal{A}_\Omega$ and $f \in H(\Omega)$.
(a) $\widetilde{f}(\varpi)$ is invertible in $\mathcal{A}$, if and only if $f(\lambda) \neq 0$, for every $\lambda \in Spec(\varpi)$.
(b) $Spec\big(\widetilde{f}(\varpi)\big) = f\big(Spec(\varpi)\big)$.

An application of this symbolic calculus deals with the existence of roots and logarithms. Note that an element $\varpi \in \mathcal{A}$ has an $n$-th root in $\mathcal{A}$ if $\varpi = \vartheta^n$ for some $\vartheta \in \mathcal{A}$. If $\varpi = \exp(\vartheta)$ for some $\vartheta \in \mathcal{A}$, then $\vartheta$ is a logarithm of $\varpi$. Although $\exp(\varpi) = \sum_{j=0} \varpi^j/j!$, but the exponential function can also be defined by contour integration, as in Definition 2.1. The continuity assertion of Theorem 2.2 shows that these definitions coincide (as they do for every entire function).

**Theorem 2.4** ([7], Theorem 10.30) Assume that for some $\varpi \in \mathcal{A}$, the spectrum $Spec(\varpi)$ of $\varpi$ does not separate 0 from $\infty$. Then
(a) $\varpi$ has roots of all orders in $\mathcal{A}$,
(b) $\varpi$ has a logarithm in $\mathcal{A}$, and
(c) if $\epsilon > 0$, there is a polynomial $P$ such that $\|\varpi^{-1} - P(\varpi)\| < \epsilon$.
Moreover, if $Spec(\varpi) \cap \mathbb{R}^- = \emptyset$, the roots in (a) can be chosen so as to satisfy the same condition. In addition, for every $p \geq 2$, there is a unique $p$-th root $\vartheta$ of $\varpi$ such that $Spec(\vartheta) \subset \big\{z \in \mathbb{C}\ ;\ -\pi/p < \arg(z) < \pi/p\big\}$. We refer to $\vartheta$ as the principal $p$-th root of $\varpi$ and write $\vartheta = \varpi^{1/p}$.

These results are not quite trivial even when $\mathcal{A}$ is a finite-dimensional algebra. For example, it is a special case of (b) that a complex $n \times n$ matrix $A$ has a logarithm if and only if 0 is not an eigenvalue of $A$; that is, if and only if $A$ is invertible.

If some $\varpi \in \mathcal{A}$ satisfies a polynomial identity, i.e., if $Q(\varpi) = 0$ for some polynomial $Q$, then $\widetilde{f}(\varpi)$ can always be calculated as a polynomial in $\varpi$, without using the Cauchy integral as in Definition 2.1. If $\mathcal{A}$ is finite dimensional, then this remark applies to every $\varpi \in \mathcal{A}$. Since in this case, the set $\{\varpi^n\}_{n=1}^{\infty}$ is a linearly dependent set. Also, when $\mathcal{A}$ is finite dimensional, the polynomial $P$ which mentioned in (c) can be chosen so that $P(\varpi) = \varpi^{-1}$. Here are the details:

**Theorem 2.5** ([7], Theorem 10.31) Let $P(z) = (z-\lambda_1)^{n_1} \cdots (z-\lambda_k)^{n_k}$ be a polynomial of degree $n = n_1 + \cdots + n_k$, and $\Omega$ an open set in $\mathbb{C}$, which contains the zeros $\lambda_1, \cdots \lambda_k$ of $P$. If $\mathcal{A}$ is a Banach algebra, $\varpi \in \mathcal{A}$, and $P(\varpi) = 0$, then
(a) $Spec(\varpi) \subset \{\lambda_1, \lambda_2, \cdots, \lambda_k\}$,
(b) to every $f \in H(\Omega)$, corresponds a polynomial $Q$ of degree $< n$, and a function $g \in H(\Omega)$, so that $f(z) = P(z)g(z) + Q(z)$ and $\widetilde{f}(\varpi) = Q(\varpi)$,

(c) If $\lambda_i \neq 0$ for $i = 1, \cdots, k$ and $P(z) = a_0 + a_1 z + \cdots + a_n z^n$, then $\varpi$ is invertible and $\varpi^{-1} = -a_1/a_0 e - a_2/a_0 \varpi - \cdots - a_{n-1}/a_0 \varpi^{n-1}$.

An element $\varpi \in \mathcal{A}$ is called *annihilated*, if $\varpi$ satisfies the conclusion of Theorem 2.5; i.e., if there exist polynomial $P$ for which $P(\varpi) = 0$. *The minimal polynomial* of an annihilated element $\varpi$ is the monic polynomial $\psi$ over $\mathbb{C}$ of least degree such that, $\psi(\varpi) = 0$. By theorem 2.5, we conclude the following corollary immediately:

**Corollary 2.6** Let $\mathcal{A}$ be a Banach algebra and $\varpi$ an annihilated element, and let $f : \Omega \to \mathbb{C}$ be defined on the spectrum of $\varpi$ where $\Omega \subset \mathbb{C}$ is an open set which contains the spectrum of $\varpi$. If $\psi$ is the minimal polynomial of $\varpi$, then $f(\varpi) = p(\varpi)$, where $p(z) \in \mathbb{C}[z]$ is the polynomial of degree less than $\deg(\psi)$ that satisfies the interpolation conditions:

$$p^{(j)}(\lambda_i(\varpi)) = f^{(j)}(\lambda_i(\varpi)),$$

where $j = 0, \cdots, n_i - 1$ and $n_i$ is the repeated order of $\lambda_i(\varpi) \in \mathrm{Spec}(\varpi)$ in $\psi$.

***Remark 1*** *The polynomial $p$ in Corollary 2.6 is called the Hermite interpolating polynomial of $f$, and it is given explicitly by the Lagrange-Hermite formula*

$$p(t) = \sum_{i=1}^{k} \left[ \left( \sum_{j=0}^{n_i - 1} \frac{1}{j!} \phi_i^{(j)}(\lambda_i)(t - \lambda_i)^j \right) \prod_{r \neq i} (t - \lambda_r)^{n_r} \right], \tag{2}$$

*where $\phi_i(t) = f(t)/\prod_{j \neq i}(t - \lambda_i)^{n_j}$. When the minimal polynomial $\psi$ has distinct roots ($n_i \equiv 1$, $k = n$), this formula reduces to the familiar Lagrange form*

$$p(t) = \sum_{i=1}^{n} f(\lambda_i) L_i(t) \qquad , \qquad L_i(t) = \prod_{\substack{j=1 \\ j \neq i}}^{n} \left( \frac{t - \lambda_j}{\lambda_i - \lambda_j} \right).$$

*Corollary 2.6 explicitly makes $f(\varpi)$ a polynomial in $\varpi$. It is important to note, however, that the polynomial $p$ depends on $\varpi$, through the values of $f$ on the spectrum of $\varpi$, so it is not the case that $f(\varpi) \equiv p(\varpi)$, for some fixed polynomial $p$ independent of $\varpi$.*

There is a well-known approach which gives an algorithm for computing the power of any annihilated element $\varpi \in \mathcal{A}$, and in particular, the power of any square matrix. Actually, this approach can be derived from Corollary 2.6, with $f(z) = z^m$. By this assumption, we obtain an unique and closed form for $\varpi^m$ with $m \geqslant n$, where $n$ is the degree of the minimal polynomial of $\varpi$. Hence, the $m$-th power of $\varpi$, and inductively all higher powers, are expressible as a linear combination of $e, \varpi, \cdots, \varpi^{n-1}$. Thus, any power series in $\varpi$ can be reduced to a polynomial in $\varpi$ of degree at most $n - 1$. This polynomial is rarely of an elegant form or practical interest. For every $m$, this polynomial can be presented as

$$\varpi^m = p_{n-1}(m)\varpi^{n-1} + p_{n-2}(m)\varpi^{n-2} + \ldots + p_1(m)\varpi + p_0(m) \tag{3}$$

and has several applications.

In special case, when $\mathcal{A} = M_n(\mathbb{C})$ for some $n \geq 2$, there is another approach which gives $A^m$ for $A \in M_n(\mathbb{C})$ and $m \geq n$. This approach based on the *Jordan Canonical Form.* Denoted by $\lambda_1, \cdots, \lambda_k$ the distinct eigenvalues of $A$, and let $n_i$ be the repeated times of the eigenvalue $\lambda_i$. It is a standard result that $A$ can be expressed in the Jordan canonical form

$$Z^{-1}AZ = J = \mathrm{diag}(J_1, J_2, \cdots, J_k), \tag{4}$$

where for each $i = 1, 2, \cdots, k$, we have

$$J_i = J_i(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix} \in M_{n_i}(\mathbb{C}).$$

The transforming matrix $Z$ is nonsingular and is not unique. But the Jordan matrix $J$ is unique up to the ordering of the blocks $J_i$. The function $f$ is defined on the spectrum of $A$ if the values $f^{(j)}(\lambda_i)$ for $j = 1, 2, \cdots, n_i$ and $i = 1, 2, \cdots, k$ exist. They are called the values of the function $f$ on the spectrum of $A$.

**Theorem 2.7** ([2], Theorem 1.12]) Let $f$ be defined on the spectrum of $A \in M_n(\mathbb{C})$ and let $A$ have the Jordan canonical form (4). Then

$$\widetilde{f}(A) = Z\widetilde{f}(J)Z^{-1} = Z\,diag\big(\widetilde{f}(J_i)\big)Z^{-1}, \tag{5}$$

where for every $i = 1, 2, \cdots, k$, we have

$$\widetilde{f}(J_i) = \widetilde{f}\big(J_i(\lambda_i)\big) = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \cdots & \frac{f^{(n_i-1)}(\lambda_i)}{(n_i-1)!} \\ & f(\lambda_i) & \ddots & \vdots \\ & & \ddots & f'(\lambda_i) \\ & & & f(\lambda_i) \end{bmatrix}. \tag{6}$$

To provide some insight into this theorem, we make several comments.
(a) The expression yields an $\widetilde{f}(A)$ that can be shown to be independent of the particular Jordan canonical form that is used. If $A$ is diagonalizable, then the Jordan canonical form reduces to an eigen-decomposition $A = ZDZ^{-1}$, with $D = \mathrm{diag}(\lambda_i)$, and the columns of $Z$ are the distinct eigenvectors of $A$. Then, Theorem 2.7 yields $\widetilde{f}(A) = Z\widetilde{f}(D)Z^{-1} = Z\mathrm{diag}\big(f(\lambda_i)\big)Z^{-1}$.
(b) In most cases of practical interest, $f$ is given by a formula, such as $f(z) = e^z$, $f(z) = z^m$ or any other complex valued function. However, Theorem 2.7 requires only the values of $f$ on the spectrum of $A$; it does not require any other information about $f$. Indeed any arbitrary numbers can be chosen and assigned as the values of $f^{(j)}(\lambda_i)$. It is only when we need to make statements about global properties such as continuity that we will need to assume more about $f$.
(c) Finally, we explain how (6) can be obtained from Taylor series considerations. Write $J_i = \lambda_i I_i + N_i \in M_{n_i}(\mathbb{C})$, where $N_i$ is zero except for a superdiagonal of 1s. In general, powering $N_i$ causes the superdiagonal of 1s to move a diagonal at a time towards the top right-hand corner, until at the $n_i$-th power disappears: $N_i^{n_i} = 0$; so $N_k$ is nilpotent.

Assume that $f$ has the following convergent Taylor series expansion

$$f(t) = f(\lambda_i) + f'(\lambda_i)(t-\lambda_i) + \cdots + \frac{f^{(m)}(\lambda_i)}{m!}(t-\lambda_i)^m + \cdots .$$

By substituting $J_i$ for $t$, we obtain the following finite series, since all powers of $N_i$ from the $n_i$-th onwards are zero.

$$\widetilde{f}(J_i) = f(\lambda_i)I_i + f'(\lambda_i)(J_i - \lambda_i I_i) + \cdots + \frac{f^{(n_i-1)}(\lambda_i)}{(n_i-1)!}(J_i - \lambda_i I_i)^{n_i-1}.$$

This expression is easily seen to agree with (6). In a particular case, if we take $f(z) = z^m$ for some $m \geq n$, the expression in (6) becomes

$$\big(J_i(\lambda_i)\big)^m = \begin{bmatrix} \lambda_i^m & m\lambda_i^{m-1} & \cdots & \binom{m}{n_i}\lambda_i^{m-n_i} \\ & \lambda_i^m & \ddots & \vdots \\ & & \ddots & m\lambda_i^{m-1} \\ & & & \lambda_i^m \end{bmatrix}. \tag{7}$$

Now, substituting $\big(J_i(\lambda_i)\big)^m$ in (5), we get a closed form for $A^m$. Although the expression (5) looks different than (3), but multiplying $Z$ and $Z^{-1}$ in $\widetilde{f}(J)$, the final expression for $A^m$ is agree with (3). In the next section, we give our method to conclude the formula (3); moreover, we are going to determine the coefficient functions $p_i(m)$ in the independent way from Hermite interpolation.

## 3. Main Results: Power, Inverse and $p$-th root

As we mentioned before, we give another way to obtain (3). In this new way, we use the difference equations and also we do not need to use the derivation compare with the Hermit interpolation. Moreover, we apply this closed form to find out the roots of an annihilated element in a Banach algebra. At the first, we need the following theorem:

**Theorem 3.1** [10] Let sequence $\{a_n\}$ satisfies the recurrence relation

$$c_0 a_n + c_1 a_{n-1} + \ldots + c_m a_{n-m} = 0,$$

where $c_0 \neq 0, c_m \neq 0$ and $1 \leqslant m \leqslant n$. If $\lambda_1, \ldots, \lambda_k$ be distinct roots of the equation $c_0 x^m + c_1 x^{m-1} + \ldots + c_{m-1}x + c_m = 0$. Then

$$\begin{aligned} a_n = {} & (\alpha_{11} + \alpha_{12} n + \ldots + \alpha_{1m_1} n^{m_1-1})\lambda_1^n \\ & + (\alpha_{21} + \alpha_{22} n + \ldots + \alpha_{2m_1} n^{m_2-1})\lambda_2^n \\ & + \cdots + (\alpha_{k1} + \alpha_{k2} n + \ldots + \alpha_{km_k} n^{m_k-1})\lambda_k^n, \end{aligned} \tag{8}$$

where $m_i$ is the repeated times of the root $\lambda_i$ and each $\alpha_{ij}$ is any complex number.

Let $\varpi$ be an annihilated element in Banach algebra $\mathcal{A}$, and it's minimal polynomial be as follow:

$$\psi(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0 \tag{9}$$

By the definition, $\varpi^n = b_{n-1}\varpi^{n-1} + \ldots + b_1\varpi + b_0 e$, where $e$ is the unity of $\mathcal{A}$, and $b_i = -a_i$ for $i = 0, 1, \ldots, n-1$. Our aim is to present an approach to calculate $\varpi^m$ for every $m \in \mathbb{N}$ and $m \geqslant n$ (also for every $m \in \mathbb{Z}$ whenever $\varpi$ is invertible) in general. Considering (3) and multiplying by $\varpi$, yields:

$$\varpi^{m+1} = p_{n-1}(m)\varpi^n + p_{n-2}(m)\varpi^{n-1} + \ldots + p_1(m)\varpi^2 + p_0(m)\varpi \tag{10}$$

By substituting (9) in (10) provides:

$$\begin{aligned}\varpi^{m+1} &= (b_{n-1}p_{n-1}(m) + p_{n-2}(m))\varpi^{n-1} \\ &\quad + (b_{n-2}p_{n-1}(m) + p_{n-3}(m))\varpi^{n-2} \\ &\quad + \cdots + (b_1 p_{n-1}(m) + p_0(m))\varpi + b_0 p_{n-1}(m)e. \end{aligned} \tag{11}$$

In other hand, changing index $m$ to $m+1$ in (3) gives:

$$\varpi^{m+1} = p_{n-1}(m+1)\varpi^{n-1} + p_{n-2}(m+1)\varpi^{n-2} + \ldots + p_1(m+1)\varpi + p_0(m+1)e. \tag{12}$$

By comparing (11) and (12), one can give the following equations:

$$p_{n-1}(m+1) = b_{n-1}p_{n-1}(m) + p_{n-2}(m), \tag{13}$$

$$\vdots$$

$$p_{n-i}(m+1) = b_{n-i}p_{n-1}(m) + p_{n-i-1}(m), \tag{14}$$

$$\vdots$$

$$p_2(m+1) = b_2 p_{n-1}(m) + p_1(m), \tag{15}$$

$$p_1(m+1) = b_1 p_{n-1}(m) + p_0(m), \tag{16}$$

$$p_0(m+1) = b_0 p_{n-1}(m). \tag{17}$$

From (17), $p_0(m) = b_0 p_{n-1}(m-1)$. Substituting this relation in (16) gives:

$$p_1(m) = b_1 p_{n-1}(m-1) + b_0 p_{n-1}(m-2). \tag{18}$$

By substituting (18) in (15) yields:

$$p_2(m) = b_2 p_{n-1}(m-1) + b_1 p_{n-1}(m-2) + b_0 p_{n-1}(m-3).$$

Similarly, for the index $i$ one can easily get the following equation:

$$p_i(m) = b_i p_{n-1}(m-1) + \ldots + b_1 p_{n-1}(m-i) + b_0 p_{n-1}(m-i-1)$$

and for the index $n-1$,

$$p_{n-1}(m) = b_{n-1}p_{n-1}(m-1) + \ldots + b_1 p_{n-1}(m-n-1) + b_0 p_{n-1}(m-n).$$

According to the above discussion, the equations (13)-(17) can be reformulated as follow:

$$\begin{aligned} p_0(m) &= b_0 p_{n-1}(m-1),\\ p_1(m) &= b_1 p_{n-1}(m-1) + b_0 p_{n-1}(m-2),\\ p_2(m) &= b_2 p_{n-1}(m-1) + b_1 p_{n-1}(m-2) + b_0 p_{n-1}(m-3),\\ &\vdots\\ p_i(m) &= b_i p_{n-1}(m-1) + \ldots + b_1 p_{n-1}(m-i) + b_0 p_{n-1}(m-i-1),\\ &\vdots\\ p_{n-1}(m) &= b_{n-1}p_{n-1}(m-1) + \ldots + b_1 p_{n-1}(m-n-1) + b_0 p_{n-1}(m-n). \end{aligned} \tag{19}$$

By using the last equation of (19) and the fact that $b_i = -a_i$ $(i = 0, 1, \ldots, n-1)$, we have

$$p_{n-1}(m) + a_{n-1}p_{n-1}(m-1) + \ldots + a_1 p_{n-1}(m-n-1) + a_0 p_{n-1}(m-n) = 0. \tag{20}$$

By computing the $p_{n-1}(m)$ of (20) and the other $p_j(m)$ for $j = 0, 1, \ldots, n-2$ of (19), the $\varpi^m$ is obtained. In fact, the characteristic polynomial for the recurrence relation (20) is

$$z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0 = 0. \tag{21}$$

One can see (21) is equal to the minimal polynomial of $\varpi$ as in (9). Now, based on Theorem 3.1, $p_{n-1}(i)$ for $i \geqslant 1$ can be found. Here, if $\lambda_1, \ldots, \lambda_k$ for $1 \leqslant k \leqslant n$ are distinct roots of (21) and each $\lambda_i$ has multiplicity $n_i$, then Theorem 3.1 implies

$$\begin{aligned} p_{n-1}(m) = {} & (\alpha_{11}^{(n-1)} + \alpha_{12}^{(n-1)}m + \ldots + \alpha_{1n_1}^{(n-1)}m^{n_1-1})\lambda_1^m\\ & + (\alpha_{21}^{(n-1)} + \alpha_{22}^{(n-1)}m + \ldots + \alpha_{2n_2}^{(n-1)}m^{n_2-1})\lambda_2^m\\ & + \cdots + (\alpha_{k1}^{(n-1)} + \alpha_{k2}^{(n-1)}m + \ldots + \alpha_{kn_k}^{(n-1)}m^{n_k-1})\lambda_k^m, \end{aligned} \tag{22}$$

where each $\alpha_{ij}^{(n-1)}$ is a complex number and determines from the initial condition of recurrence relation. In fact, $p_{n-1}(n) = -a_{n-1}$, $p_{n-1}(n-1) = 1$, and $p_{n-1}(i) = 0$ for $i = 1, 2, \ldots, n-2$. Now, by substituting $p_{n-1}(m)$ into (19), the other $p_j(m)$ for $j = 0, 1, \ldots, n-2$ can be found. Therefore, by using (3), $\varpi^m$ for $m \in \mathbb{N}$ is computed. The other $p_j(m)$ for $j = 0, 1, \ldots, n-2$ are obtained as the same as $p_{n-1}(m)$ in (22), i.e.,

$$\begin{aligned} p_j(m) = {} & (\alpha_{11}^{(j)} + \alpha_{12}^{(j)}m + \ldots + \alpha_{1n_1}^{(j)}m^{n_1-1})\lambda_1^m\\ & + (\alpha_{21}^{(j)} + \alpha_{22}^{(j)}m + \ldots + \alpha_{2n_2}^{(j)}m^{n_2-1})\lambda_2^m\\ & + \cdots + (\alpha_{k1}^{(j)} + \alpha_{k2}^{(j)}m + \ldots + \alpha_{kn_k}^{(j)}m^{n_k-1})\lambda_k^m. \end{aligned} \tag{23}$$

$\varpi^m$ can be shown in another way different from (3). In fact, if we substitute $p_{n-1}(m)$ and the other $p_j(m)$ for $j = 0, 1, \ldots, n-2$ into (3), then

$$\begin{aligned}\varpi^m = &[(\alpha_{11}^{(n-1)} + \alpha_{12}^{(n-1)} m + \ldots + \alpha_{1n_1}^{(n-1)} m^{n_1-1})\varpi^{n-1} + \ldots \\ &+ (\alpha_{11}^{(0)} + \alpha_{12}^{(0)} m + \ldots + \alpha_{1n_1}^{(0)} m^{n_1-1})e]\lambda_1^m \\ &+ [(\alpha_{21}^{(n-1)} + \alpha_{22}^{(n-1)} m + \ldots + \alpha_{2n_2}^{(n-1)} m^{n_2-1})\varpi^{n-1} + \ldots \\ &+ (\alpha_{21}^{(0)} + \alpha_{22}^{(0)} m + \ldots + \alpha_{2n_2}^{(0)} m^{n_2-1})]\lambda_2^m \\ &\vdots \\ &+ [(\alpha_{k1}^{(n-1)} + \alpha_{k2}^{(n-1)} m + \ldots + \alpha_{kn_k}^{(n-1)} m^{n_k-1})\varpi^{n-1} + \ldots \\ &+ (\alpha_{k1}^{(0)} + \alpha_{k2}^{(0)} m + \ldots + \alpha_{kn_k}^{(0)} m^{n_k-1})e]\lambda_k^m.\end{aligned}$$

By classifying and summarizing the above equation, $\varpi^m$ can be written briefly as follows:

$$\begin{aligned}\varpi^m = &(m^{n_1-1}\lambda_1^m)S_{11} + (m^{n_1-2}\lambda_1^m)S_{12} + \ldots + \lambda_1^m S_{1n_1} \\ &+ (m^{n_2-1}\lambda_2^m)S_{21} + (m^{n_2-2}\lambda_2^m)S_{22} + \ldots + \lambda_2^m S_{2n_2} \\ &+ \cdots + (m^{n_k-1}\lambda_k^m)S_{k1} + (m^{n_k-2}\lambda_k^m)S_{k2} + \ldots + \lambda_k^m S_{kn_k}, \qquad (24)\end{aligned}$$

where $S_{ij}$ are the linear combinations of $\{\varpi^i\}_{i=0}^{n-1}$, i.e.,

$$S_{ij} = \alpha_{in_i-j+1}^{(n-1)}\varpi^{n-1} + \alpha_{in_i-j+1}^{(n-2)}\varpi^{n-2} + \ldots + \alpha_{in_i-j+1}^{(0)} e. \qquad (25)$$

***Remark 2*** *For some annihilated element $\varpi$ and applying the following minimal polynomial*

$$\psi(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0,$$

*we have $a_0 = a_1 = \cdots = a_r$ for some $0 \le r \le n-2$. Hence, the relations* (13)-(17) *imply $p_0 = p_1 = \cdots = p_r \equiv 0$ and so*

$$\varpi^m = p_{n-1}(m)\varpi^{n-1} + p_{n-2}(m)\varpi^{n-2} + \ldots + p_{r+1}(m)\varpi^{r+1}.$$

*In such a situation, $p_{n-1}(m)$ can be derived by the initial values $p_{n-1}(r) = p_{n-1}(r+1) = p_{n-1}(n-2) = 0$, $p_{n-1}(n-1) = 1$ and $p_{n-1}(n) = b_{n-1}$. Actually, in this case, there is no polynomial $Q(z)$ so that represent $\varpi^m$ as superposition of $e, \varpi, \cdots, \varpi^{n-1}$, for which $e, \varpi, \cdots, \varpi^r$ having non-zero coefficient (since the minimal polynomial of $\varpi$ has the factor $z^r$, and any other polynomial $Q$ with $Q(\varpi) = 0$ should have this factor).*

***Remark 3*** *Although* (23) *gives the canonical form of $p_j(m)$ for $j = 1, 2, \cdots, k$ and $m \ge n$, but in most situations, we do not need to determine the exact values of $a_{i,r}^{(j)}$. We can use the identity*

$$p_j(m) = b_j p_{n-1}(m-1) + \ldots + b_1 p_{n-1}(m-j) + b_0 p_{n-1}(m-j-1), \qquad (1 \le j \le n-1) \quad (26)$$

*and $p_j(m)$ can be derived from $p_{n-1}(m-1), p_{n-1}(m-2), \cdots, p_{n-1}(m-j), p_{n-1}(m-j-1)$. In other word, to compute $\varpi^m$, we just need to compute $p_{n-1}(m), p_{n-1}(m-1), p_{n-1}(m-2), \cdots, p_{n-1}(m-j), p_{n-1}(m-j-1)$. Then, (26) gives other $p_j(m)$s. Actually, to compute $p_j(m)$ by (23) ( for simplicity, ignore the necessary calculations to obtain $a^{(j)}_{i,r}$ ), we must have $k(m+1) + \sum_{i=1}^{n} \frac{n_i(n_i+1)}{2}$ summers and multiplications; and taking sum over $j$, at least we have $n\left(k(m+1) + \sum_{i=1}^{n} \frac{n_i(n_i+1)}{2}\right)$ operation. But using the recurrent relation (26) to drive $p_j(m)$, we need $2(j+1)$ operation; and $\varpi^m$ obtain by $(n+1)(n+2)$ operation. This reduction in computations is due to recurrent relationships between $p_j(m)$s i.e. (19). However, in the Hermite interpolation or Jordan canonical method, our results are the existence of $p_j(m)$ and Corollary 2.6 or Theorem 5 does not give any information about the relations between $p_j(m)$s.*

***Remark 4*** *If $\varpi \in \mathcal{A}$ satisfies the conclusions of Theorem 2.5 (c), then*

$$\varpi^{-1} = -a_1/a_0 e - a_2/a_0 \varpi - \cdots - 1/a_0 \varpi^{n-1}.$$

*In this case, we have $p_0(0) = 1$ and $p_i(0) = 0$ for $i = 1, 2, \ldots, n-1$. Thus, (13)-(17) change as follow:*

$$\begin{aligned} p_0(0) &= b_0 p_{n-1}(-1) = 1 \\ p_1(0) &= b_1 p_{n-1}(-1) + p_0(-1) = 0 \\ &\vdots \\ p_{n-i}(0) &= b_{n-i} p_{n-1}(-1) + p_{n-i-1}(-1) = 0 \\ &\vdots \\ p_{n-1}(0) &= b_{n-1} p_{n-1}(-1) + p_{n-2}(-1) = 0. \end{aligned} \tag{27}$$

*This implies $p_j(-1) = -a_{j+1}/a_0$ with assumption $a_n = 1$. Thus,*

$$\varpi^{-1} = p_{n-1}(-1)\varpi^{n-1} + p_{n-2}(-1)\varpi^{n-2} + \cdots + p_1(-1)\varpi + p_0(-1)e, \tag{28}$$

*which means the expression (3) is also valid for $m = -1$. In fact, when $\varpi$ is invertible, (3) is valid for all $m \in \mathbb{Z}$. This can be done by induction; let for some $m < 0$, we have*

$$\varpi^m = p_{n-1}(m)\varpi^{n-1} + p_{n-2}(m)\varpi^{n-2} + \cdots + p_1(m)\varpi + p_0(m)e. \tag{29}$$

*By multiplying (29) in $\varpi^{-1}$, we get that*

$$\varpi^{m-1} = p_{n-1}(m)\varpi^{n-2} + p_{n-2}(m)\varpi^{n-3} + \cdots + p_1(m)e + p_0(m)\varpi^{-1}. \tag{30}$$

*By substituting (28) in (30), we have*

$$\begin{aligned} \varpi^{m-1} = &\frac{1}{b_0} p_0(m)\varpi^{n-1} + \big(\frac{-b_{n-1}}{b_0} p_0(m) + p_{n-1}(m)\big)\varpi^{n-2} \\ &+ \cdots + \big(\frac{-b_2}{b_0} p_0(m) + p_2(m)\big)\varpi + \big(\frac{-b_1}{b_0} p_0(m) + p_1(m)\big)e. \end{aligned} \tag{31}$$

*Now,* (13)-(17) *imply that* $\frac{-b_i}{b_0}p_0(m) + p_i(m) = -b_i p_{n-1}(m-1) + p_i(m) = p_{i-1}(m-1)$, *which means* (31) *can be written as*

$$\varpi^{m-1} = p_{n-1}(m-1)\varpi^{n-1} + p_{n-2}(m-1)\varpi^{n-2} + \cdots + p_1(m-1)\varpi + p_0(m-1)e.$$

*This statement about* $m < 0$ *is the consequences of equations* (13)-(17)*; i.e. these equations are valid even for* $m < 0$*, enabled us to conclude the expression* (30)*. The functional equations* (13)-(17) *are valid for every* $m \in \mathbb{R}$*. So, one would expect that for every* $x \in \mathbb{Q}$, $\varpi^x = \sum_{i=0}^{n-1} p_i(x)\varpi^i$*, and this is what the next theorem says:*

**Theorem 3.2** Let $\varpi$ be an annihilated invertible element of Banach algebra $\mathcal{A}$ and consider the closed form of $\varpi^m$ in (29). Then, for every $p/q \in \mathbb{Q}$, $\varpi^{p/q}$ has the following representation:

$$\varpi^{q/p} = \sum_{i=0}^{n-1} p_i(q/p)\varpi^i \tag{32}$$

**Proof.** Since $\varpi$ is an annihilated invertible element, by Theorem 2.4 (a), $\varpi$ has roots of all order in $\mathcal{A}$. Suppose that $\vartheta$ is one of this $p$-th roots. Also, Theorem 2.3 and 2.5 implies that each eigenvalue of $\vartheta$ is an $p$-th root of some eigenvalue of $\varpi$. Therefore, using the expansion (24), $\vartheta^q$ for every $q \in \mathbb{Z}$ can be written as the following form:

$$\begin{aligned}\vartheta^q = &\left(q^{d_1-1}(\lambda_1^{\frac{1}{p}})^q\right)U_{11} + \left(q^{d_1-2}(\lambda_1^{\frac{1}{p}})^q\right)U_{12} + \ldots + (\lambda_1^{\frac{1}{p}})^q U_{1d_1}\\ &+ \left(q^{d_2-1}(\lambda_2^{\frac{1}{p}})^q\right)U_{21} + \left(q^{d_2-2}(\lambda_2^{\frac{1}{p}})^q\right)U_{22} + \ldots + (\lambda_2^{\frac{1}{p}})^q U_{2d_2}\\ &+ \cdots + \left(q^{d_k-1}(\lambda_k^{\frac{1}{p}})^q\right)U_{k1} + \left(q^{d_k-2}(\lambda_k^{\frac{1}{p}})^q\right)U_{k2} + \ldots + (\lambda_k^{\frac{1}{p}})^q U_{kd_k}\end{aligned} \tag{33}$$

where $U_{ij}$ has the same definition as (25). Put $q = mp$ for $m \in \mathbb{Z}$. Thus, we have

$$\begin{aligned}\vartheta^{mp} = &\left((mp)^{d_1-1}(\lambda_1^{\frac{1}{p}})^{mp}\right)U_{11} + \left((mp)^{d_1-2}(\lambda_1^{\frac{1}{p}})^{mp}\right)U_{12} + \ldots + (\lambda_1^{\frac{1}{p}})^{mp} U_{1d_1} \qquad (34)\\ &+ \left((mp)^{d_2-1}(\lambda_2^{\frac{1}{p}})^{mp}\right)U_{21} + \left((mp)^{d_2-2}(\lambda_2^{\frac{1}{p}})^{mp}\right)U_{22} + \ldots + (\lambda_2^{\frac{1}{p}})^{mp} U_{2d_2}\\ &+ \ldots + \left((mp)^{d_k-1}(\lambda_k^{\frac{1}{p}})^{mp}\right)U_{k1} + \left((mp)^{d_k-2}(\lambda_k^{\frac{1}{p}})^{mp}\right)U_{k2} + \ldots + (\lambda_k^{\frac{1}{p}})^{mp} U_{kd_k}\end{aligned}$$

By considering $\vartheta^{mp} = (\vartheta^p)^m = \varpi^m$ and $(\lambda_j^{\frac{1}{p}})^{mp} = \lambda_j^m$ for $j = 1, 2, \ldots, k$, we have

$$\begin{aligned}\varpi^m = &(m^{d_1-1}\lambda_1^m)(p^{d_1-1})U_{11} + (m^{d_1-2}\lambda_1^m)(p^{d_1-2})U_{12} + \ldots + \lambda_1^m U_{1d_1}\\ &+ (m^{d_2-1}\lambda_2^m)(p^{d_2-1})U_{21} + (m^{d_2-2}\lambda_2^m)(p^{d_2-2})U_{22} + \ldots + \lambda_2^m U_{2d_2}\\ &+ \cdots + (m^{d_k-1}\lambda_k^m)(p^{d_k-1})U_{k1} + (m^{d_k-2}\lambda_k^m)(p^{d_k-2})U_{k2} + \ldots + \lambda_k^m U_{kd_k}\end{aligned} \tag{35}$$

Here, by comparing $\varpi^m$ in (24), (35) and the uniqueness representation of $\varpi^m$ in the terms of its eigenvalues, we have $d_i = n_i$ and $U_{ij} = \frac{1}{p^{d_i-j}}S_{ij}$. Substituting (35) in (33) and summarizing yields:

$$\begin{aligned}
\varpi^{\frac{q}{p}} = \vartheta^q = & \left( (\frac{q}{p})^{d_1-1} S_{11} + (\frac{q}{p})^{d_1-2} S_{12} + \ldots + S_{1d_1} \right) \lambda_1^{\frac{q}{p}} \\
& + \left( (\frac{q}{p})^{d_2-1} S_{21} + (\frac{q}{p})^{d_2-2} S_{22} + \ldots + S_{2d_2} \right) \lambda_2^{\frac{q}{p}} \\
& + \ldots + \left( (\frac{q}{p})^{d_k-1} S_{k1} + (\frac{q}{p})^{d_k-2} S_{k2} + \ldots + S_{kd_k} \right) \lambda_k^{\frac{q}{p}} \\
= & \sum_{i=0}^{n-1} p_i(\frac{q}{p}) \varpi^i. \qquad (36)
\end{aligned}$$

■

***Remark 5*** *Theorem 3.2 requires some comments.*
*(a) In Remark 3, if $m$ is an integer, then the recurrent relations (26) can be used to reduce the computations. But in the case $m \in \mathbb{Q}$, we have to calculate $p_{n-1}(\cdot)$ as in (22) and then the other $p_j(\cdot)$s can be derived from (26). However, (26) still is useful in reducing the computations in comparison with Hermit and Jordan methods for p-th roots.*
*(b) In the case $\mathcal{A} = M_n(\mathbb{C})$, where $M_n(\mathbb{C})$ is an $n \times n$ nonsingular matrix with $k$ distinct eigenvalues. Then it has $p^k$ distinct p-th roots ([2], Theorem 7.1 p. 173). The principle p-th root of the non-singular matrix $A$ can be obtained by taking the principle branch of $\lambda_j^{\frac{1}{p}}$ in (36) for $j = 1, 2, \ldots, k$. Furthermore, the other p-th roots are given by taking the other branches of $\lambda_j^{\frac{1}{p}}$ in (36). Hence, all the $p^k$ distinct p-th roots are accessible.*

***Example* 3.3** In this example, we consider the following non-diagonalizable $10 \times 10$ matrix $A$, and calculate $A^m$ using all three methods which we described.

$$A = \begin{bmatrix}
1 & 1 & 1 & -2 & 1 & -1 & 2 & -2 & 4 & -3 \\
-1 & 2 & 3 & -4 & 2 & -2 & 4 & -4 & 8 & -6 \\
-1 & 0 & 5 & -5 & 3 & -3 & 6 & -6 & 12 & -9 \\
-1 & 0 & 3 & -4 & 4 & -4 & 8 & -8 & 16 & -12 \\
-1 & 0 & 3 & -6 & 5 & -4 & 10 & -10 & 20 & -15 \\
-1 & 0 & 3 & -6 & 2 & -2 & 12 & -12 & 24 & -18 \\
-1 & 0 & 3 & -6 & 2 & -5 & 15 & -13 & 28 & -21 \\
-1 & 0 & 3 & -6 & 2 & -5 & 12 & -11 & 32 & -24 \\
-1 & 0 & 3 & -6 & 2 & -5 & 12 & -14 & 37 & -26 \\
-1 & 0 & 3 & -6 & 2 & -5 & 12 & -14 & 36 & -25
\end{bmatrix}.$$

(a) The Jordan canonical method: the Jordan block structure of a matrix is difficult to determine, since the set of $n \times n$ diagonalizable matrices is dense in $M_n(\mathbb{C})$, and thus, small changes in a matrix can radically alter its Jordan form. Golub and Loan [[1], p. 248] illustrate the difficulty of calculating the Jordan canonical form of the matrix $A$. Also, Li et al. [3] used the algorithm (based on symbolic computation) to obtain the Jordan canonical form of $A$. They obtained the minimal polynomial of $A$ as follow

$$\psi(z) = z^6 - 13z^5 + 69z^4 - 191z^3 + 290z^2 - 228z + 72 \qquad (37)$$

and then, using the symbolic computations (which are very difficult and long), they

found the Jordan canonical form $J = diag(2I_3 + N_3, 3I_2 + N_2, 1, 2I_2 + N_2, 3I_2 + N_2)$ of $A$. However, their method is unable to compute the matrix $Z \in M_{10}(\mathbb{C})$ which $Z^{-1}AZ = J$. To conclude $Z$, assume that we are written $J = \text{diag}\big(J(\lambda_1), \cdots, J(\lambda_k)\big)$, where for $i = 1, \cdots, k$ and $J(\lambda_i) \in M_{d_i \times d_i}(\mathbb{C})$ denotes the Jordan block corresponding to the eigenvalue $\lambda_i$ and $d_i$ is the repeated order of $\lambda_i$ in the characteristic polynomial of $A$. Moreover, suppose that $Z = [Z_1, Z_2, \cdots, Z_k]$, where $Z_i \in M_{n \times d_i}(\mathbb{C})$ are the columns of $Z$ associated with $i$-th Jordan block $J(\lambda_i)$, which $AZ_i = Z_iJ_i$. Let $v_{i,j}$s are the columns of $Z_i$ and $Z_i = [v_{i,1}, v_{i,2}, \cdots, v_{i,n_i}]$. Then $Av_{i,1} = \lambda_i v_{i,1}$ and $Av_{i,j} = v_{i,j-1} + \lambda_i v_{i,j}$ for $i = 1, 2, \cdots, k$ and $j = 1, 2, \cdots, n_i$. The vectors $v_{i,j}$ are called *generalized eigenvectors*. $v_{i,j}$s can be found by following algorithm:
(1) Solve $(A - \lambda_i I_{10})v_{i,1} = 0$. This step finds all the eigenvectors associated with $\lambda_i$. The number of eigenvectors depends on $rank(A - \lambda_i I_{10})$. For example, if $rank(A - \lambda_i I_{10}) = n - 1$, there is only one eigenvector.
(2) For each independent $v_{i,1}$ from (1), solve $(A - \lambda_i I_{10})v_{i,2} = v_{i,1}$. The number of linearly independent solutions at this step depends on $rank(A - \lambda_i I_{10})^2$. If, for example $rank(A - \lambda_i I_{10})^2 = n - 2$, there are two linearly independent solutions to the homogeneous equation $(A - \lambda_i I_{10})^2 v_{i,2} = 0$. One of these solutions is $v_{i,1}$; since $(A - \lambda_i I_{10})^2 v_{i,1} = (A - \lambda_i I_{10})0 = 0$. The other solution is the desired generalized eigenvectors.
(3) For each independent $v_{i,2}$ from step (2), solve $(A - \lambda_i I_{10})v_{i,3} = v_{i,2}$ and continue until the total number of independent generalized eigenvectors $v_{i,1}, \cdots, v_{i,n_i}$ found.

Unfortunately, this natural-looking procedure can fail to find all Jordan vectors. For more extensive treatments, see, for example, [4] and [6]. Determination of eigenvectors and generalized eigenvectors is obviously very tedious for anything beyond simple problems ($n = 2$ or 3, say). Attempts to do such calculations in finite-precision floating-point arithmetic, generally prove unreliable. There are significant numerical difficulties inherent in attempting to compute the Jordan canonical form.

In the best situation, for $n \times n$ matrix, we can find $Z$ with the algorithm described above, by solving at least $n$ non-homogeneous linear system of $n$ variables and $n$ equation, which their coefficient matrix is not invertible. If we use the $L - U$-factorization method or Gaussian elimination method to solving each system of equation in the above algorithm, it can be shown that the number of computations has the order $\text{O}(n^3)$ [[8], page 56]. And so, the number of whole computations for finding $Z$, is at least from order $\text{O}(n^4)$. In addition, we did not concluded the computations for finding the Jordan canonical form.
(b) Hermite interpolation method: using the minimal polynomial of $A$ in (37), we have

$$\begin{cases} \lambda_1 = 1,\ n_1 = 1 \quad and \quad \phi_1(t) = \frac{t^m}{(t-2)^3(t-3)^2} \\ \lambda_2 = 2,\ n_2 = 3 \quad and \quad \phi_2(t) = \frac{t^m}{(t-1)(t-3)^2} \\ \lambda_3 = 3,\ n_3 = 2 \quad and \quad \phi_3(t) = \frac{t^m}{(t-2)^3(t-1)} \end{cases}$$

Doing the arithmetic computations, we obtain

$$\begin{cases} \phi_1(1) = -\frac{1}{4} \\ \phi_3(3) = \frac{3^m}{2} \quad and \quad \phi_3'(3) = 3^m(\frac{m}{6} - \frac{7}{4}) \\ \phi_2(2) = 2^m\ , \quad \phi_2'(2) = 2^m(\frac{m}{2} + 1) \quad and \quad \phi_2^{(2)}(2) = 2^m(\frac{m^2}{4} + \frac{3m}{4} + 4) \end{cases}$$

Assume that for $i = 1, 2, 3$,

$$\Phi_i(t) = \left(\sum_{j=0}^{n_i-1} \frac{1}{j!}\phi_i^{(j)}(\lambda_i)(t-\lambda_i)^j\right)\left(\prod_{r\neq i}(t-\lambda_r)^{n_r}\right)$$

Then

$$\begin{cases} \Phi_1(t) = -\frac{1}{4}(t-2)^3(t-3)^2 \\ \Phi_3(t) = 3^m\left(\frac{1}{2} + (\frac{m}{6} - \frac{7}{4})(t-3)\right)(t-1)(t-2)^3 \\ \Phi_2(t) = 2^m\left(1 + (\frac{m}{2}+1)(t-2) + (\frac{m^2}{8} + \frac{3m}{8} + 2)(t-2)^2\right)(t-1)(t-3)^2 \end{cases}$$

which can be written as:

$$\begin{aligned} \Phi_1(t) &= -\frac{1}{4}\left(t^5 - 12t^4 + 57t^3 - 134t^2 + 156t - 72\right), \\ \Phi_3(t) &= 3^m\left((\frac{m}{6} - \frac{7}{4})t^5 + (18 - \frac{5m}{3})t^4 + (\frac{13m}{2} - \frac{287}{4})t^3 - (\frac{37m}{3} - \frac{277}{2})t^2\right) \\ &\quad + 3^m\left((\frac{34m}{3} - 129)t - 4m + 46\right), \\ \Phi_2(t) &= 2^m\left((\frac{m^2}{8} + \frac{3m}{8} + 2)t^5 - (\frac{11m^2}{8} + \frac{29m}{8} + 21)t^4 + (\frac{47m^2}{8} + \frac{105m}{8} + 86)t^3\right) \\ &\quad - 2^m\left((\frac{97m^2}{8} + \frac{175m}{4} + 172)t^2 - (12m^2 + \frac{33m}{2} + 168)t + \frac{9}{2}m^2 + \frac{9}{2}m + 63\right). \end{aligned}$$

Now, the Hermite interpolation polynomial is

$$\begin{aligned} p(t) &= \Phi_1(t) + \Phi_2(t) + \Phi_3(t) \\ &= p_5(m)t^5 + p_4(m)t^4 + p_3(m)t^3 + p_2(m)t^2 + p_1(m)t + p_0(m) \end{aligned}$$

where

$$p_5(m) = (2 + \frac{3m}{8} + \frac{m^2}{8})2^m + (\frac{m}{6} - \frac{7}{4})3^m - \frac{1}{4}, \tag{38}$$

$$p_4(m) = (-21 - \frac{29m}{8} - \frac{11m^2}{8})2^m + (18 - \frac{5m}{3})3^m + 3, \tag{39}$$

$$p_3(m) = (86 + \frac{105m}{8} + \frac{47m^2}{8})2^m + (\frac{13m}{2} - \frac{287}{4})3^m - \frac{57}{4}, \tag{40}$$

$$p_2(m) = (-172 - \frac{175m}{8} - \frac{97m^2}{8})2^m + (\frac{277}{2} - \frac{37m}{3})3^m + \frac{67}{2}, \tag{41}$$

$$p_1(m) = (168 + \frac{33m}{2} + 12m^2)2^m + (\frac{34m}{3} - 129)3^m - 39, \tag{42}$$

$$p_0(m) = (-63 - \frac{9m}{2} - \frac{9m^2}{2})2^m + (46 - 4m)3^m + 18. \tag{43}$$

Thus,

$$A^m = p_5(m)A^5 + p_4(m)A^4 + p_3(m)A^3 + p_2(m)A^2 + p_1(m)A + p_0(m)I_{10}. \tag{44}$$

Rewriting (44), we can find the representation (25) which we described in Theorem 3.1

$$A^m = \big(m^2(2)^m\big)S_{11} + \big(m(2)^m\big)S_{12} + (2^m)S_{13} + \big(m(3)^m\big)S_{21} + (3^m)S_{22} + S_{31},$$

where

$$\begin{aligned}
S_{11} &= \frac{1}{8}A^5 - \frac{11}{8}A^4 + \frac{47}{8}A^3 - \frac{97}{8}A^2 + 12A - \frac{9}{2}I_{10}\\
S_{12} &= \frac{3}{8}A^5 - \frac{29}{8}A^4 + \frac{105}{8}A^3 - \frac{175}{8}A^2 + \frac{33}{2}A - \frac{9}{2}I_{10}\\
S_{13} &= 2A^5 - 21A^4 + 86A^3 - 172A^2 + 168A - 63I_{10}\\
S_{21} &= \frac{1}{6}A^5 - \frac{5}{3}A^4 + \frac{13}{2}A^3 - \frac{37}{3}A^2 + \frac{34}{3}A - 4I_{10}\\
S_{22} &= -\frac{7}{4}A^5 + 18A^4 - \frac{287}{4}A^3 + \frac{277}{2}A^2 - 129A + 46I_{10}\\
S_{31} &= -\frac{1}{4}A^5 + 3A^4 - \frac{57}{4}A^3 + \frac{67}{2}A^2 - 39A + 18I_{10}
\end{aligned}$$

One of difficulties of this method is calculating $\phi_i^{(j)}$s for $j = 1, \cdots, n_i$ and $i = 1, \cdots, k$. For each $i = 1, 2, \cdots, k$, we can write

$$\phi_i(t) = t^m \prod_{r \neq i} (t - \lambda_r)^{-n_r}.$$

Using multinomial formula of general Leibnitz rule, we get that

$$\phi_i^{(j)}(t) = \sum_{r_1+\cdots+r_n=j} \binom{j}{r_1, \cdots, r_n} (t^m)^{(r_1)} \prod_{\substack{q\neq i\\ t\neq 1}} \big((t-\lambda_q)^{-n_q}\big)^{(r_t)},$$

where $\binom{j}{r_1,\cdots,r_n} = \frac{j!}{r_1!\cdots r_n!}$. The number of calculation works for obtaining $\Phi_i(t)$ is at least from order $O(n_i^4)$; and the whole calculation is at least from order $O(s^4)$, where $s = n_1 + n_2 + \cdots + n_4$.

The other difficulty is involving symbolic calculations with polynomials, rational functions of polynomials and their product or differentials, which are boring and nerve-racking.

Our method: we first give some description about computation works of our method, and some explanations for obtaining the functions $p_j(m)$s. Assume that the conditions of Theorem 3.1 are satisfied. By relation (23), for $j = 0, 1, \cdots, n-1$, we can write:

$$\begin{aligned}
p_j(m) = &(\alpha_{j,1,1} + \alpha_{j,1,2}m + \ldots + \alpha_{j,1,n_1}m^{n_1-1})\lambda_1^m\\
&+ (\alpha_{j,2,1} + \alpha_{j,2,2}m + \ldots + \alpha_{j,2,n_2}m^{n_2-1})\lambda_2^m\\
&+ \cdots + (\alpha_{j,k,1} + \alpha_{j,k,2}m + \ldots + \alpha_{j,k,n_k}m^{n_k-1})\lambda_k^m
\end{aligned} \tag{45}$$

According to (44), $p_j(q) = 0$ for $q \neq j, n-1$; $p_j(j) = 1$ and $p_j(n-1) = -a_j$. By this initial values, $\alpha_{j,r,q}$s must satisfies the linear system of $s$ equations in $s$ variables:

$$\begin{bmatrix} \lambda_1 & \lambda_1 & \cdots & \lambda_1 & \cdots\cdots & \lambda_k & \lambda_k & \cdots & \lambda_k & \lambda_k \\ \lambda_1^2 & 2\lambda_1^2 & \cdots & 2^{d_1}\lambda_1^2 & \cdots\cdots & \lambda_k^2 & 2\lambda_k^2 & \cdots & 2^{d_k-1}\lambda_k^2 & 2^{d_k}\lambda_k^2 \\ \lambda_1^3 & 3\lambda_1^3 & \cdots & 3^{d_1}\lambda_1^3 & \cdots\cdots & \lambda_k^3 & 3\lambda_k^3 & \cdots & 3^{d_k-1}\lambda_k^3 & 3^{d_k}\lambda_k^3 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda_1^j & j\lambda_1^j & \cdots & j^{d_1}\lambda_1^j & \cdots\cdots & \lambda_k^j & j\lambda_k^j & \cdots & j^{d_k-1}\lambda_k^j & j^{d_k}\lambda_k^j \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda_1^d & d\lambda_1^d & \cdots & d^{d_1}\lambda_1^d & \cdots\cdots & \lambda_k^d & d\lambda_k^d & \cdots & d^{d_k-1}\lambda_k^d & d^{d_k}\lambda_k^d \\ \lambda_1^n & n\lambda_1^n & \cdots & n^{d_1}\lambda_1^n & \cdots\cdots & \lambda_k^n & n\lambda_k^n & \cdots & n^{d_k-1}\lambda_k^n & n^{d_k}\lambda_k^n \end{bmatrix} \begin{bmatrix} \alpha_{j,1,1} \\ \alpha_{j,1,2} \\ \vdots \\ \alpha_{j,1,n_1} \\ \vdots \\ \alpha_{j,k,1} \\ \alpha_{j,k,2} \\ \vdots \\ \alpha_{j,k,n_1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ -a_{n-1} \end{bmatrix}.$$

where $d = n-1$ and $d_i = n_i - 1$ for $i = 1, 2, \cdots, k$, and 1 appears in the $j$-th coordinate in the right hand of equality. Invertibility of coefficient matrix follows from the uniqueness of Hermite interpolation polynomial. Assume that we are written the above system in the form $C\underline{\alpha}_j = \underline{b}_j$, where $C$ is the coefficient matrix, $\underline{\alpha}_j$ is the matrix of unknown $\alpha_{j,i,r}$s, and $\underline{b}_j$ is the constant terms matrix. Moreover, let $C^{-1} = (c_{i,r})_{i,r=1}^n$ be the inverse of $C$. Then

$$\underline{\alpha}_j = \begin{bmatrix} \alpha_{j,1,1} \\ \alpha_{j,1,2} \\ \vdots \\ \alpha_{j,1,n_1} \\ \vdots \\ \alpha_{j,k,1} \\ \alpha_{j,k,2} \\ \vdots \\ \alpha_{j,k,n_k} \end{bmatrix} = C^{-1}\underline{b}_j = \begin{bmatrix} c_{1,j} - a_{n-1}c_{1,n} \\ c_{2,j} - a_{n-1}c_{2,n} \\ \vdots \\ c_{n_1,j} - a_{n-1}c_{n_1,n} \\ \vdots \\ c_{n-n_k,j} - a_{n-1}c_{n-n_k,n} \\ c_{n-n_k+1,j} - a_{n-1}c_{n-n_k+1,n} \\ \vdots \\ c_{n,j} - a_{n-1}c_{n,n} \end{bmatrix}$$

It can be shown that for every $s \times s$ matrix, the number of computations for obtaining it's inverse by Gaussian elimination method, has the order $O(s^3)$ ([8], page 57). Then, each $\underline{\alpha}_j$ can be concluded by $3s$ computation and whole $\underline{\alpha}_j$s are derived by $3s^2$ computation. Thus, all the computations which we need to drive $p_j(m)$s has the order $O(s^3)$, which is so less than Hermite interpolation and Jordan methods.

In the special case, for the matrix $A$ which introduced above, the coefficient matrix $C$ and the constant term matrix $\underline{b}$ are as follow:

$$C = \begin{bmatrix} 2 & 2 & 2 & 3 & 3 & 1 \\ 4 & 8 & 16 & 9 & 18 & 1 \\ 8 & 24 & 72 & 27 & 81 & 1 \\ 16 & 64 & 256 & 81 & 324 & 1 \\ 32 & 160 & 800 & 243 & 1215 & 1 \\ 64 & 384 & 2304 & 729 & 4374 & 1 \end{bmatrix}, \quad \underline{b} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -72 & 228 & -290 & 191 & -69 & 13 \end{bmatrix}.$$

Then, by the relation $\underline{\alpha} = C^{-1}\underline{b}$, we obtain

$$\begin{bmatrix} \alpha_{0,1,1} & \alpha_{1,1,1} & \alpha_{2,1,1} & \alpha_{3,1,1} & \alpha_{4,1,1} & \alpha_{5,1,1} \\ \alpha_{0,1,2} & \alpha_{1,1,2} & \alpha_{2,1,2} & \alpha_{3,1,2} & \alpha_{4,1,2} & \alpha_{5,1,2} \\ \alpha_{0,1,3} & \alpha_{1,1,3} & \alpha_{2,1,3} & \alpha_{3,1,3} & \alpha_{4,1,3} & \alpha_{5,1,3} \\ \alpha_{0,2,1} & \alpha_{1,2,1} & \alpha_{2,2,1} & \alpha_{3,2,1} & \alpha_{4,2,1} & \alpha_{5,2,1} \\ \alpha_{0,2,2} & \alpha_{1,2,2} & \alpha_{2,2,2} & \alpha_{3,2,2} & \alpha_{4,2,2} & \alpha_{5,2,2} \\ \alpha_{0,3,1} & \alpha_{1,3,1} & \alpha_{2,3,1} & \alpha_{3,3,1} & \alpha_{4,3,1} & \alpha_{5,3,1} \end{bmatrix} = \begin{bmatrix} -63 & 168 & -172 & 86 & -21 & 2 \\ -\frac{9}{2} & \frac{33}{2} & -\frac{175}{8} & \frac{105}{8} & -\frac{29}{8} & \frac{3}{8} \\ -\frac{9}{2} & 12 & -\frac{97}{8} & \frac{47}{8} & -\frac{11}{8} & \frac{1}{8} \\ 46 & -129 & \frac{277}{2} & -\frac{287}{4} & 18 & -\frac{7}{4} \\ -4 & \frac{34}{3} & -\frac{37}{3} & \frac{13}{2} & -\frac{5}{3} & \frac{1}{6} \\ 18 & -39 & \frac{67}{2} & -\frac{57}{4} & 3 & -\frac{1}{4} \end{bmatrix}$$

which is agree with equations (38)-(43). As we mentioned before, the computation works in the Hermite method is at least $O(s^4)$, and in Jordan method (in the best case, when the numerical problems dose not occur) is at least $O(n^4)$. But in our method, we reduced the number of computation works to $O(s^3)$. Another advantage of this method to the Jordan method, is that we do not need to obtain the Jordan canonical form. Also, we get the same result just by solving one non-homogeneous linear system which the coefficient matrix is invertible, instead of solving $n$ non-homogeneous linear system of $n$ variables and $n$ equation, which their coefficient matrix is not invertible. The other preference than Hermite interpolation method, we can say that in this method, we removed the boring symbolic calculation with polynomials and rational functions of polynomials and their differentials; and we obtained the same answer as Hermite interpolation method. In other word, we replaced the nonlinear computation of Hermite method, by the linear computations, and we reduced the question to finding out the answer of a linear system of equations.

## Acknowledgements

The authors would like to thank from the anonymous reviewers for carefully reading of the manuscript and giving useful comments, which will help to improve the paper.

## References


[1] G. H. Golub, C. F. Van Loan, Matrix computations, JHU Press, 1983.
[2] N. J. Higham, Functions of matrices: theory and computation, SIAM, 2008.
[3] T. Li, Z. Zhang, T. Wang, Determining the structure of the Jordan normal form of a matrix by symbolic computation, Linear Algebra and its Applications. 252 (1-3) (1997), 221-259.
[4] J. M. Ortega, Matrix theory: A second course, Springer Science & Business Media, 2013.
[5] J. J. Palis, W. De Melo, Geometric theory of dynamical systems: an introduction, Springer Science & Business Media, 2012.
[6] R. Penrose, A generalized inverse for matrices, Math. Proc. Camb. Philos. Soc. 51 (3) (1955), 406-413.
[7] W. Rudin, Functional Analysis, Internat. Ser. Pure. Appl. Math., 1991.
[8] E. Süli, D. F. Mayers, An introduction to numerical analysis, Cambridge University Press, 2003.
[9] M.Tomás-Rodriguez, S. P. Banks, Linear time-varying approximations to nonlinear dynamical systems: with applications in control and optimization, Springer Science & Business Media, 2010.
[10] J. H. Van Lint, R. M. Wilson, A course in combinatorics, Cambridge University Press, 2001.