The Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 9, pp. 32-41, April 2002

ELA
http://math.technion.ac.il/iic/ela

ON THE mth ROOTS OF A COMPLEX MATRIX*

PANAYIOTIS J. PSARRAKOS†

Abstract. If an $n \times n$ complex matrix A is nonsingular, then for every integer $m > 1$, A has an mth root B, i.e., $B^m = A$. In this paper, we present a new simple proof for the Jordan canonical form of the root B. Moreover, a necessary and sufficient condition for the existence of mth roots of a singular complex matrix A is obtained. This condition is in terms of the dimensions of the null spaces of the powers A^k $(k = 0, 1, 2, \ldots)$.

Key words. Ascent sequence, eigenvalue, eigenvector, Jordan matrix, matrix root.

AMS subject classifications. 15A18, 15A21, 15A22, 47A56

1. Introduction and preliminaries. Let $\mathcal{M}_n$ be the algebra of all $n \times n$ complex matrices and let $A \in \mathcal{M}_n$. For an integer $m > 1$, a matrix $B \in \mathcal{M}_n$ is called an *mth root* of A if $B^m = A$. If the matrix A is nonsingular, then it always has an mth root B. This root is not unique and its Jordan structure is related to the Jordan structure of A [2, pp. 231-234]. In particular, $(\lambda - \mu_0)^k$ is an elementary divisor of B if and only if $(\lambda - \mu_0^m)^k$ is an elementary divisor of A. If A is a singular complex matrix, then it may have no mth roots. For example, there is no matrix B such that $B^2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. As a consequence, the problem of characterizing the singular matrices, which have mth roots, is of interest [1], [2].

Consider the (associated) *matrix polynomial* $P(\lambda) = I_n\lambda^m - A$, where I_n is the identity matrix of order n and λ is a complex variable. A matrix $B \in \mathcal{M}_n$ is an mth root of A if and only if $P(B) = B^m - A = 0$. As a consequence, the problem of computation of mth roots of A is strongly connected with the spectral analysis of $P(\lambda)$. The suggested references for matrix polynomials are [3] and [7].

A set of vectors $\{x_0, x_1, \ldots, x_k\}$, which satisfies the equations

$$\begin{aligned} P(\omega_0)x_0 &= 0 \\ P(\omega_0)x_1 + \frac{1}{1!}P^{(1)}(\omega_0)x_0 &= 0 \\ \vdots \qquad\qquad \vdots \qquad &\quad \vdots \\ P(\omega_0)x_k + \frac{1}{1!}P^{(1)}(\omega_0)x_{k-1} + \cdots + \frac{1}{k!}P^{(k)}(\omega_0)x_0 &= 0, \end{aligned}$$

where the indices on $P(\lambda)$ denote derivatives with respect to the variable λ, is called a *Jordan chain* of length $k+1$ of $P(\lambda)$ corresponding to the *eigenvalue* $\omega_0 \in \mathbb{C}$ and the *eigenvector* $x_0 \in \mathbb{C}^n$. The vectors in a Jordan chain are not uniquely defined and for $m > 1$, they need not be linearly independent [3], [6]. If we set $m = 1$, then the

*Received by the editors on 19 February 2002. Accepted for publication on 9 April 2002. Handling Editor: Peter Lancaster.

†Department of Mathematics, National Technical University, Zografou Campus, 15780 Athens, Greece (ppsarr@math.ntua.gr).

The Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 9, pp. 32-41, April 2002

ELA
http://math.technion.ac.il/iic/ela

Jordan structure of the *linear pencil* $I_n\lambda - A$ coincides with the Jordan structure of A, and the vectors of each Jordan chain are chosen to be linearly independent [2], [6]. Moreover, there exist a matrix

$$J_A = \oplus_{j=1}^{\xi}\left(I_{k_j}\omega_j + N_{k_j}\right) \quad (k_1 + k_2 + \ldots + k_\xi = n), \tag{1.1}$$

where N_k is the nilpotent matrix of order k having ones on the super diagonal and zeros elsewhere, and an $n \times n$ nonsingular matrix

$$X_A = \left[\begin{array}{ccccccccc} x_{1,1} & \ldots & x_{1,k_1} & x_{2,1} & \ldots & x_{2,k_2} & \ldots & x_{\xi,1} & \ldots & x_{\xi,k_\xi}\end{array}\right], \tag{1.2}$$

where for every $j = 1,2,\ldots,\xi$, $\{x_{j,1}, x_{j,2},\ldots,x_{j,k_j}\}$ is a Jordan chain of A corresponding to $\omega_j \in \sigma(A)$, such that (see [2], [4], [6])

$$A = X_A\,J_A\,X_A^{-1}.$$

The matrix J_A is called the *Jordan matrix* of A, and it is unique up to permutations of the diagonal *Jordan blocks* $I_{k_j}\omega_j + N_{k_j}$ $(j = 1,2,\ldots,\xi)$ [2], [4].

The set of all eigenvalues of $P(\lambda)$, that is, $\sigma(P) = \{\mu \in \mathbb{C} : \det P(\mu) = 0\}$, is called the *spectrum* of $P(\lambda)$. Denoting by $\sigma(A) = \sigma(I_n\lambda - A)$ the spectrum of the matrix A, it is clear that $\sigma(P) = \{\mu \in \mathbb{C} : \mu^m \in \sigma(A)\}$. If J_A is the Jordan matrix of A in (1.1), then it will be convenient to define the *J-spectrum* of A, $\sigma_J(A) = \{\omega_1,\omega_2,\ldots,\omega_\xi\}$, where the eigenvalues of A follow exactly the order of their appearance in J_A (obviously, repetitions are allowed). For example, the J-spectrum of the matrix $M = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 \end{bmatrix} \oplus \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is $\sigma_J(M) = \{0,0,1\}$.

In this article, we study the Jordan structure of the mth roots $(m > 1)$ of a complex matrix. In Section 2, we consider a nonsingular matrix and present a new constructive proof for the Jordan canonical form of its mth roots. This proof is simple and based on spectral analysis of matrix polynomials [2], [3], [7]. Furthermore, it yields directly the Jordan chains of the mth roots. We also generalize a known uniqueness statement [5]. In Section 3, using a methodology of Cross and Lancaster [1], we obtain a necessary and sufficient condition for the existence of mth roots of a singular matrix.

2. The nonsingular case. Consider a nonsingular matrix $A \in \mathcal{M}_n$ and an integer $m > 1$. If A is diagonalizable and $S \in \mathcal{M}_n$ is a nonsingular matrix such that

$$A = S\,\mathrm{diag}\{r_1 e^{\mathrm{i}\phi_1}, r_2 e^{\mathrm{i}\phi_2},\ldots,r_n e^{\mathrm{i}\phi_n}\}\,S^{-1},$$

where $r_j > 0$, $\phi_j \in [0,2\pi)$ $(j = 1,2,\ldots,n)$, then for every n-tuple $(s_1,s_2,\ldots,s_n)$, $s_j \in \{1,2,\ldots,m\}$ $(j = 1,2,\ldots,n)$, the matrix

$$B = S\,\mathrm{diag}\{r_1^{\frac{1}{m}} e^{\mathrm{i}\frac{\phi_1+2(s_1-1)\pi}{m}}, r_2^{\frac{1}{m}} e^{\mathrm{i}\frac{\phi_2+2(s_2-1)\pi}{m}},\ldots,r_n^{\frac{1}{m}} e^{\mathrm{i}\frac{\phi_n+2(s_n-1)\pi}{m}}\}\,S^{-1}$$

is an mth root of A. Hence, the investigation of the mth roots of a nonsingular (and not diagonalizable) matrix A via the Jordan canonical form of A arises in a natural way [2]. The following lemma is necessary and of independent interest.

The Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 9, pp. 32-41, April 2002

ELA
http://math.technion.ac.il/iic/ela

Lemma 2.1. *Let* $\{x_0, x_1, \ldots, x_k\}$ *be a Jordan chain of* $A \in \mathcal{M}_n$ *(with linearly independent terms) corresponding to a nonzero eigenvalue* $\omega_0 = r_0 e^{\mathrm{i}\,\phi_0} \in \sigma(A)$ $(r_0 > 0,\ \phi_0 \in [0, 2\pi))$, *and let* $P(\lambda) = I_n\lambda^m - A$. *Then for every eigenvalue*

$$r_0^{\frac{1}{m}}\, e^{\mathrm{i}\,\frac{\phi_0+2(t-1)\pi}{m}} \in \sigma(P) \ ; \quad t = 1, 2, \ldots, m,$$

the matrix polynomial $P(\lambda)$ *has a Jordan chain of the form*

$$\begin{aligned} y_0 &= x_0 \\ y_1 &= a_{1,1}x_1 \\ y_2 &= a_{2,1}x_1 + a_{2,2}x_2 \\ \vdots \quad & \quad \vdots \qquad\quad \vdots \\ y_k &= a_{k,1}x_1 + a_{k,2}x_2 + \cdots + a_{k,k}x_k, \end{aligned} \tag{2.1}$$

where the coefficients $a_{i,j}$ $(1 \le j \le i \le k)$ *depend on the integer* t *and for every* $i = 1, 2, \ldots, k$, $a_{i,i} = (m\, r_0^{\frac{m-1}{m}}\, e^{\mathrm{i}(m-1)\,\frac{\phi_0+2(t-1)\pi}{m}})^i \neq 0$. *Moreover, the vectors* $y_0, y_1, \ldots, y_k$ *are linearly independent.*

Proof. Since $\{x_0, x_1, \ldots, x_k\}$ is a Jordan chain of the matrix A corresponding to the eigenvalue $\omega_0 \neq 0$, we have

$$(A - I_n\omega_0)x_0 \ = \ 0$$

and

$$(A - I_n\omega_0)x_i \ = \ x_{i-1} \ ; \qquad i = 1, 2, \ldots, k.$$

Let μ_0 be an eigenvalue of $P(\lambda)$ such that $\mu_0^m = \omega_0$. By the equation

$$(I_n\omega_0 - A)x_0 \ = \ (I_n\mu_0^m - A)x_0 \ = \ 0,$$

it is obvious that $y_0 = x_0$ is an eigenvector of $P(\lambda)$ corresponding to $\mu_0 \in \sigma(P)$. Assume now that there exists a vector $y_1 \in \mathbb{C}^n$ such that

$$P(\mu_0)y_1 + \frac{P^{(1)}(\mu_0)}{1!}\, y_0 \ = \ 0.$$

Then

$$(I_n\mu_0^m - A)y_1 \ = \ -\, m\mu_0^{m-1}y_0,$$

or equivalently,

$$(I_n\omega_0 - A)y_1 \ = \ m\mu_0^{m-1}(I_n\omega_0 - A)x_1.$$

Hence, we can choose $y_1 = a_{1,1}x_1$, where $a_{1,1} = m\mu^{m-1} \neq 0$. Similarly, if we consider the equation

$$P(\mu_0)y_2 + \frac{P^{(1)}(\mu_0)}{1!}\, y_1 + \frac{P^{(2)}(\mu_0)}{2!}\, y_0 \ = \ 0,$$

The Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 9, pp. 32-41, April 2002

ELA
http://math.technion.ac.il/iic/ela

then it follows

$$(I_n\mu_0^m - A)y_2 \;=\; -\,m\mu_0^{m-1}y_1 - \frac{m(m-1)}{2}\,\mu_0^{m-2}y_0,$$

or equivalently,

$$(I_n\omega_0 - A)y_2 \;=\; (I_n\omega_0 - A)\left((m\mu_0^{m-1})^2x_2 + \frac{m(m-1)}{2}\,\mu_0^{m-2}x_1\right).$$

Thus, we can choose $y_2 = a_{2,1}x_1 + a_{2,2}x_2$, where $a_{2,1} = \frac{m(m-1)}{2}\,\mu_0^{m-2}$ and $a_{2,2} = \left(m\mu^{m-1}\right)^2 \neq 0$. Repeating the same steps implies that the matrix polynomial $P(\lambda)$ has a Jordan chain $\{y_0, y_1, \dots, y_k\}$ as in (2.1).

Define the $n \times (k+1)$ matrices

$$X_0 \;=\; \left[\begin{array}{cccc} x_0 & x_1 & \cdots & x_k \end{array}\right] \quad\text{and}\quad Y_0 \;=\; \left[\begin{array}{cccc} y_0 & y_1 & \cdots & y_k \end{array}\right].$$

Since the vectors $x_0, x_1, \dots, x_k \in \mathbb{C}^n$ are linearly independent, $\operatorname{rank}(X_0) = k+1$. Moreover, $Y_0 = X_0\,T_0$, where the $(k+1)\times(k+1)$ upper triangular matrix

$$T_0 \;=\; \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & a_{1,1} & a_{2,1} & \cdots & a_{k,1} \\ 0 & 0 & a_{2,2} & \cdots & a_{k,2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{k,k} \end{bmatrix}$$

is nonsingular. As a consequence, $\operatorname{rank}(Y_0) = k+1$, and the proof is complete. ∎

The Jordan chain $\{y_0, y_1, \dots, y_k\}$ of $P(\lambda)$ in the above lemma, is said to be *associated* to the Jordan chain $\{x_0, x_1, \dots x_k\}$ of A, and clearly depends on the choice of $t \in \{1, 2, \dots, m\}$. Consider now the nonsingular matrix $X_A \in \mathcal{M}_n$ in (1.2), and for any $(s_1, s_2, \dots, s_\xi)$, $s_j \in \{1, 2, \dots, m\}$ $(j = 1, 2, \dots, \xi)$, define the matrix

$$Y_A(s_1, s_2, \dots, s_\xi) \;=\; \left[\begin{array}{ccccccccc} y_{1,1} & \dots & y_{1,k_1} & y_{2,1} & \cdots & y_{\xi,1} & \cdots & y_{\xi,k_\xi} \end{array}\right],$$

where for every $j = 1, 2, \dots, \xi$, the set $\{y_{j,1}, y_{j,2}, \dots, y_{j,k_j}\}$ is the associated Jordan chain of $P(\lambda)$ corresponding to the Jordan chain $\{x_{j,1}, x_{j,2}, \dots, x_{j,k_j}\}$ of A and the integer s_j.

COROLLARY 2.2. *For every $(s_1, s_2, \dots, s_\xi)$, $s_j \in \{1, 2, \dots, m\}$ $(j = 1, 2, \dots, \xi)$, the associated matrix $Y_A(s_1, s_2, \dots, s_\xi)$ is nonsingular.*

Proof. By Lemma 2.1, there exist upper triangular matrices $T_1, T_2, \dots, T_\xi$, which depend on the choice of the ξ-tuple $(s_1, s_2, \dots, s_\xi)$ and have nonzero diagonal elements, such that

$$Y_A(s_1, s_2, \dots, s_\xi) \;=\; X_A\left(\oplus_{j=1}^{\xi} T_j\right).$$

Since X_A is also nonsingular the proof is complete. ∎

The Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 9, pp. 32-41, April 2002

ELA
http://math.technion.ac.il/iic/ela

Theorem 2.3. *Let $A \in \mathcal{M}_n$ be a nonsingular complex matrix with Jordan matrix $J_A = \oplus_{j=1}^{\xi} \left(I_{k_j}\omega_j + N_{k_j} \right)$ as in (1.1) and J-spectrum*

$$\sigma_J(A) \,=\, \{\, \omega_1 = r_1 e^{\mathrm{i}\phi_1}, \omega_2 = r_2 e^{\mathrm{i}\phi_2}, \ldots, \omega_\xi = r_\xi e^{\mathrm{i}\phi_\xi} \}.$$

Consider an integer $m > 1$, the nonsingular matrix $X_A \in \mathcal{M}_n$ in (1.2) such that $A = X_A J_A X_A^{-1}$, a ξ-tuple $(s_1, s_2, \ldots, s_\xi)$, $s_j \in \{1, 2, \ldots, m\}$ $(j = 1, 2, \ldots, \xi)$ and the associated matrix $Y_A(s_1, s_2, \ldots, s_\xi)$. Then the matrix

$$B \,=\, Y_A(s_1, s_2, \ldots, s_\xi) \left(\oplus_{j=1}^{\xi} \left(I_{k_j} r_j^{\frac{1}{m}} e^{\mathrm{i}\frac{\phi_j + 2(s_j - 1)\pi}{m}} + N_{k_j} \right) \right) Y_A(s_1, s_2, \ldots, s_\xi)^{-1} \tag{2.2}$$

is an mth root of A.

Proof. Since the associated matrix $Y_A(s_1, s_2, \ldots, s_\xi)$ is nonsingular, by Corollary 7.11 in [3], the linear pencil $I_n\lambda - B$ is a *right divisor* of $P(\lambda) = I_n\lambda^m - A$, i.e., there exists an $n \times n$ matrix polynomial $Q(\lambda)$ of degree $m - 1$ such that

$$P(\lambda) \,=\, Q(\lambda)\,(I_n\lambda - B).$$

Consequently, by [2, pp. 81-82] (see also Lemma 22.9 in [7]), $P(B) = B^m - A = 0$, and hence B is an mth root of the matrix A. □

At this point, we remark that the associated matrix $Y_A(s_1, s_2, \ldots, s_n)$ can be computed directly by the method described in the proof of Lemma 2.1. Moreover, it is clear that a nonsingular matrix may have mth roots with common eigenvalues. Motivated by [5], we obtain a spectral condition that implies the uniqueness of an mth root.

Theorem 2.4. *Suppose $A \in \mathcal{M}_n$ is a nonsingular complex matrix and its spectrum $\sigma(A)$ lies in a cone*

$$\mathcal{K}_0 \,=\, \{z \in \mathbb{C} : \theta_1 \le \operatorname{Arg} z \le \theta_2,\ 0 < \theta_2 - \theta_1 \le \vartheta_0 < 2\pi\}.$$

Then for every $k = 1, 2, \ldots, m$, A has a unique mth root B_k such that

$$\sigma(B_k) \,\subset\, \left\{ z \in \mathbb{C} : \frac{\theta_1 + 2(k-1)\pi}{m} \le \operatorname{Arg} z \le \frac{\theta_2 + 2(k-1)\pi}{m} \right\}.$$

In particular, for every $k = 2, 3, \ldots, m$, $B_k = e^{\mathrm{i}\frac{2(k-1)\pi}{m}} B_1$.

Proof. Observe that the spectrum of $P(\lambda)$, $\sigma(P) \,=\, \{\mu \in \mathbb{C} : \mu^m \in \sigma(A)\}$, lies in the union

$$\bigcup_{k=1}^{m} \left\{ z \in \mathbb{C} : \frac{\theta_1 + 2(k-1)\pi}{m} \le \operatorname{Arg} z \le \frac{\theta_2 + 2(k-1)\pi}{m} \right\},$$

and for every $k = 1, 2, \ldots, m$, denote

$$\begin{aligned} \Sigma_k &= \sigma(P) \cap \left\{ z \in \mathbb{C} : \frac{\theta_1 + 2(k-1)\pi}{m} \le \operatorname{Arg} z \le \frac{\theta_2 + 2(k-1)\pi}{m} \right\} \\ &= \{\, r^{\frac{1}{m}} e^{\mathrm{i}\frac{\phi + 2(k-1)\pi}{m}} \,:\, re^{\mathrm{i}\phi} \in \sigma(A),\ r > 0,\ \phi \in [\theta_1, \theta_2] \}. \end{aligned}$$

The Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 9, pp. 32-41, April 2002

ELA
http://math.technion.ac.il/iic/ela

Then by Theorem 2.3, for every $k = 1, 2, \ldots, m,$ the matrix A has an mth root B_k such that $\sigma(B_k) = \Sigma_k$. Since the sets $\Sigma_1, \Sigma_2, \ldots, \Sigma_m$ are mutually disjoint, Lemma 22.8 in [7] completes the proof. $\square$

It is worth noting that if $J_A = \oplus_{j=1}^{\xi}\left(I_{k_j}\omega_j + N_{k_j}\right)$ is the Jordan matrix of A in (1.1), and $\sigma_J(A) = \{\omega_1 = r_1 e^{\mathrm{i}\phi_1}, \omega_2 = r_2 e^{\mathrm{i}\phi_2}, \ldots, \omega_\xi = r_\xi e^{\mathrm{i}\phi_\xi}\}$ is the J-spectrum of A, then for every $k = 1, 2, \ldots, m,$ the mth root B_k in the above theorem is given by (2.2) for $s_1 = s_2 = \cdots = s_\xi = k.$ Furthermore, if we allow $\theta_1 \longrightarrow -\pi^+$ and $\theta_2 \longrightarrow \pi^-,$ then for $m = 2,$ we have the following corollary.

Corollary 2.5. (Theorem 5 in [5]) *Let $A \in \mathcal{M}_n$ be a complex matrix with $\sigma(A) \cap (-\infty, 0] = \emptyset$. Then A has a unique square root B such that $\sigma(B) \subset \{z \in \mathbb{C} : \operatorname{Re} z > 0\}$.*

Example 2.6. Consider a 5×5 complex matrix $A = X_A J_A X_A^{-1}$, where $X_A \in \mathcal{M}_5$ is nonsingular and

$$J_A = \begin{bmatrix} \mathrm{i} & 1 & 0 & 0 & 0 \\ 0 & \mathrm{i} & 1 & 0 & 0 \\ 0 & 0 & \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Suppose $m = 3$ and for a pair $(s_1, s_2),\ s_j \in \{1, 2, 3\}\ (j = 1, 2),$ denote

$$\alpha = e^{\mathrm{i}\frac{\pi/2 + 2(s_1 - 1)\pi}{3}} \quad \text{and} \quad \beta = e^{\mathrm{i}\frac{2(s_2 - 1)\pi}{3}}.$$

Then the associated matrix of X_A, corresponding to (s_1, s_2), is

$$Y_A(s_1, s_2) = X_A \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3\alpha^2 & 3\alpha \\ 0 & 0 & 9\alpha^4 \end{bmatrix} \oplus \begin{bmatrix} 1 & 0 \\ 0 & 3\beta^2 \end{bmatrix} \right).$$

One can verify that the matrix

$$\begin{aligned} B = Y_A(s_1, s_2) & \begin{bmatrix} \alpha & 1 & 0 & 0 & 0 \\ 0 & \alpha & 1 & 0 & 0 \\ 0 & 0 & \alpha & 0 & 0 \\ 0 & 0 & 0 & \beta & 1 \\ 0 & 0 & 0 & 0 & \beta \end{bmatrix} Y_A(s_1, s_2)^{-1} \\ = X_A & \begin{bmatrix} \alpha & \frac{1}{3}\alpha^{-2} & -\frac{1}{9}\alpha^{-5} & 0 & 0 \\ 0 & \alpha & \frac{1}{3}\alpha^{-2} & 0 & 0 \\ 0 & 0 & \alpha & 0 & 0 \\ 0 & 0 & 0 & \beta & \frac{1}{3}\beta^{-2} \\ 0 & 0 & 0 & 0 & \beta \end{bmatrix} X_A^{-1} \end{aligned}$$

is a 3rd root of A (see also the equation (58) in [2, p. 232]). Moreover, if we choose $s_1 = s_2 = 1,$ then

$$Y_A(1, 1) = X_A \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3e^{\mathrm{i}\frac{\pi}{3}} & 3e^{\mathrm{i}\frac{\pi}{6}} \\ 0 & 0 & 9e^{\mathrm{i}\frac{2\pi}{3}} \end{bmatrix} \oplus \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \right),$$

The Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 9, pp. 32-41, April 2002

ELA
http://math.technion.ac.il/iic/ela

and the matrix

$$B_1 = Y_A(1,1)\begin{bmatrix} e^{\mathrm{i}\frac{\pi}{6}} & 1 & 0 & 0 & 0 \\ 0 & e^{\mathrm{i}\frac{\pi}{6}} & 1 & 0 & 0 \\ 0 & 0 & e^{\mathrm{i}\frac{\pi}{6}} & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} Y_A(1,1)^{-1}$$

$$= X_A \begin{bmatrix} e^{\mathrm{i}\frac{\pi}{6}} & \frac{1}{3}e^{-\mathrm{i}\frac{\pi}{3}} & -\frac{1}{9}e^{-\mathrm{i}\frac{5\pi}{6}} & 0 & 0 \\ 0 & e^{\mathrm{i}\frac{\pi}{6}} & \frac{1}{3}e^{-\mathrm{i}\frac{\pi}{3}} & 0 & 0 \\ 0 & 0 & e^{\mathrm{i}\frac{\pi}{6}} & 0 & 0 \\ 0 & 0 & 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} X_A^{-1}$$

is the unique 3rd root of A with $\sigma(B_1) \subset \{z \in \mathbb{C} : 0 \le \operatorname{Arg} z \le \pi/6\}$.

3. The singular case. Let $A \in \mathcal{M}_n$ be a singular matrix, and let $J_A = \oplus_{j=1}^{\xi}\left(I_{k_j}\omega_j + N_{k_j}\right)$ be its Jordan matrix in (1.1). For the remainder and without loss of generality, we assume that $\omega_j = 0$ for $j = 1, 2, \ldots, \psi$ $(1 \le \psi \le \xi)$, with $\omega_j \ne 0$ otherwise, and $k_1 \ge k_2 \ge \cdots \ge k_\psi$ [1], [2]. We also denote by

$$J_0 \;=\; \oplus_{j=1}^{\psi}\left(I_{k_j}\omega_j + N_{k_j}\right) \;=\; \oplus_{j=1}^{\psi} N_{k_j}$$

the diagonal block of J_A corresponding to the zero eigenvalue. Then by [2, pp. 234-239], we have the following lemma.

LEMMA 3.1. *The matrix $A \in \mathcal{M}_n$ has an mth root if and only if J_0 has an mth root.*

The *ascent sequence* of A is said to be the sequence

$$d_i \;=\; \dim \operatorname{Null} A^i - \dim \operatorname{Null} A^{i-1} \;\;;\qquad i = 1, 2, \ldots$$

By [1], we have the following properties:

(P1) The ascent sequences of A and J_0 are equal.

(P2) For every $i = 1, 2, \ldots$, d_i is the number of the diagonal blocks of J_0 of order at least i. Thus, if $d_0 = \sum_{j=1}^{\psi} k_j$ is the order of J_0, then $d_0 \ge d_1 \ge d_2 \ge \cdots \ge d_{k_1} \ge 0$ and $d_{k_1+1} = d_{k_1+2} = \cdots = 0$.

THEOREM 3.2. *The complex matrix $A \in M_n$ has an mth root if and only if for every integer $\nu \ge 0$, the ascent sequence of A has no more than one element between $m\,\nu$ and $m\,(\nu + 1)$.*

(Note that the result is obvious when the matrix A is nonsingular.)

Proof. By Lemma 3.1 and Property (P1), it is enough to prove the result for J_0. First assume that J_0 has an mth root Z, and that there exist a nonnegative integer ν and two terms of the ascent sequence of J_0, say d_t and d_{t+1}, such that

$$m\,\nu \;<\; d_{t+1} \;\le\; d_t \;<\; m\,(\nu+1).$$

The Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 9, pp. 32-41, April 2002
ELA
http://math.technion.ac.il/iic/ela

For every $i = 1, 2, \ldots, k_1$, $\operatorname{Null} Z^{mi} = \operatorname{Null} J_0^i$, and consequently, if $c_1, c_2, \ldots$ is the ascent sequence of Z, then

$$\sum_{j=1}^{mi} c_j = \sum_{j=1}^{i} d_j.$$

Thus, we have

$$d_t = c_{mt} + c_{mt-1} + \cdots + c_{mt-(m-1)}$$

and

$$d_{t+1} = c_{m(t+1)} + c_{m(t+1)-1} + \cdots + c_{mt+1},$$

where

$$c_{mt-(m-1)} \geq \cdots \geq c_{mt} \geq c_{mt+1} \geq \cdots \geq c_{m(t+1)}.$$

If $c_{mt} \geq \nu + 1$, then $d_t \geq m\, c_{mt} \geq m\,(\nu+1)$, a contradiction. On the other hand, if $c_{mt} \leq \nu$, then $d_{t+1} \leq m\, c_{mt} \leq m\,\nu$, which is also a contradiction. Hence, we conclude that if the matrix A has an mth root, then for every integer $\nu \geq 0$, the ascent sequence of A has no more than one element between $m\,\nu$ and $m\,(\nu+1)$.

Conversely, we indicate a constructive proof for the existence of an mth root of J_0 given that between two successive nonnegative multiplies of m there is at most one term of the ascent sequence of J_0. Denote $n_i = \sum_{j=1}^{i} k_j$ for $i = 1, 2, \ldots, \psi$, and let $\{\mathrm{e}_1, \mathrm{e}_2, \ldots, \mathrm{e}_{n_\psi}\}$ be the standard basis of $\mathbb{C}^{n_\psi}$. By [1], for the vectors $x_i = \mathrm{e}_{n_i}$ $(i = 1, 2, \ldots, \psi)$, we can write the standard basis of $\mathbb{C}^{n_\psi}$ in the following scheme:

$$\begin{array}{cccccccc}
J_0^{k_1-1}x_1 & J_0^{k_1-2}x_1 & J_0^{k_1-3}x_1 & J_0^{k_1-4}x_1 & \cdots & J_0x_1 & x_1 \\
J_0^{k_2-1}x_2 & J_0^{k_2-2}x_2 & J_0^{k_2-3}x_2 & \cdots & \cdots & x_2 & \\
J_0^{k_3-1}x_3 & J_0^{k_3-2}x_3 & \cdots & \cdots & x_3 & & \\
\vdots & \vdots & & \vdots & & & \\
J_0^{k_\psi-1}x_\psi & J_0^{k_\psi-2}x_\psi & \cdots & x_\psi\,. & & &
\end{array} \tag{3.1}$$

In this scheme, there are ψ rows of vectors, such that the jth row contains k_j vectors $(j = 1, 2, \ldots, \psi)$. Recall that $k_1 \geq k_2 \geq \cdots \geq k_\psi$, and hence the rows are of non-increasing length. Moreover, the above scheme has k_1 columns and by Property (P2), for $t = 1, 2, \ldots, k_1$, the length of the tth column is equal to the tth term of the ascent sequence of J_0, d_t.

With respect to the above scheme (and the order of its elements), we define a linear transformation $\mathcal{F}$ on $\mathbb{C}^{n_\psi}$ by

$$(j,t)\text{th element} \;\begin{array}{c}\nearrow\\ \to\\ \searrow\end{array}\; \begin{array}{ll}
(j-1,t)\text{th element} & \text{if } j \neq 1 \,(\operatorname{mod} m), \\
(j+m-1,t-1)\text{th element} & \text{if } j = 1 \,(\operatorname{mod} m) \text{ and } t \neq 1, \\
0 & \text{if } j = 1 \,(\operatorname{mod} m) \text{ and } t = 1.
\end{array}$$

The Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 9, pp. 32-41, April 2002

ELA
http://math.technion.ac.il/iic/ela

Separating the rows of the scheme in m-tuples, one can see that our assumption that for every nonnegative integer ν, the ascent sequence has no more than one element between $m\nu$ and $m(\nu+1)$ ensures the existence of $\mathcal{F}$. If B is the $n_\psi \times n_\psi$ matrix whose jth column is $\mathcal{F}(\mathrm{e}_j)$, then $B^m = J_0$ and the proof is complete. ◻

Since between two successive even integers there is exactly one odd integer, for $m = 2$, Theorem 3.2 yields the main result of [1].

Corollary 3.3. (Theorem 2 in [1]) *The matrix $A \in \mathcal{M}_n$ has a square root if and only if no two terms of its ascent sequence are the same odd integer.*

The ascent sequence of the $k \times k$ nilpotent matrix N_k is $1, 1, \ldots, 1, 0, \ldots$ with its first k terms equal to 1. Thus, for every integer $m > 1$, it has k terms (i.e., the k ones) between 0 and m. Hence, it is verified that there is no matrix $M \in \mathcal{M}_n$ such that $M^m = N_k$ (see also [2]).

Corollary 3.4. *Let $d_1, d_2, d_3, \ldots$ be the ascent sequence of a singular complex matrix $A \in \mathcal{M}_n$.*

(i) *If $d_2 = 0$ (i.e., $J_0 = 0$), then for every integer $m > 1$, A has an mth root.*

(ii) *If $d_2 > 0$, then for every integer $m > d_1$, A has no mth roots.*

Our methodology is illustrated in the following example.

Example 3.5. Consider the Jordan matrix

$$J_0 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

and let $m = 3$. The ascent sequence of J_0 is $3, 3, 1, 0, \ldots,$ and if $\{\mathrm{e}_1, \mathrm{e}_2, \ldots, \mathrm{e}_7\}$ is the standard basis of $\mathbb{C}^7$, then the scheme in (3.1) is

$$\begin{array}{lll} J_0^2 x_1 = \mathrm{e}_1 & J_0 x_1 = \mathrm{e}_2 & x_1 = \mathrm{e}_3 \\ J_0 x_2 = \mathrm{e}_4 & x_2 = \mathrm{e}_5 & \\ J_0 x_3 = \mathrm{e}_6 & x_3 = \mathrm{e}_7. & \end{array}$$

As in the proof of Theorem 3.2, we define the linear transformation $\mathcal{F}$ on $\mathbb{C}^7$ by

$$\mathcal{F}(\mathrm{e}_1) = 0,\ \ \mathcal{F}(\mathrm{e}_2) = \mathrm{e}_6,\ \ \mathcal{F}(\mathrm{e}_3) = \mathrm{e}_7,\ \ \mathcal{F}(\mathrm{e}_4) = \mathrm{e}_1,$$

$$\mathcal{F}(\mathrm{e}_5) = \mathrm{e}_2,\ \ \mathcal{F}(\mathrm{e}_6) = \mathrm{e}_4 \ \ \text{and} \ \ \mathcal{F}(\mathrm{e}_7) = \mathrm{e}_5.$$

One can see that the 7×7 matrix

$$B = \begin{bmatrix} 0 & \mathrm{e}_6 & \mathrm{e}_7 & \mathrm{e}_1 & \mathrm{e}_2 & \mathrm{e}_4 & \mathrm{e}_5 \end{bmatrix} = \begin{bmatrix} 0&0&0&1&0&0&0 \\ 0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&0 \\ 0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&1 \\ 0&1&0&0&0&0&0 \\ 0&0&1&0&0&0&0 \end{bmatrix}$$

is a 3rd root of J_0. Finally, observe that for every integer $m > 1$ different than 3, the matrix J_0 has no mth roots.

The Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 9, pp. 32-41, April 2002

ELA
http://math.technion.ac.il/iic/ela

REFERENCES

[1] G.W. Cross and P. Lancaster. Square roots of complex matrices. *Linear and Multilinear Algebra*, 1:289-293, 1974.
[2] F.R. Gantmacher. *The Theory of Matrices*, Vol. I. Chelsea Publishing Company, New York, 1959.
[3] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. Academic Press, New York, 1982.
[4] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.
[5] C.R. Johnson, K. Okubo, and R. Reams. Uniqueness of matrix square roots and an application. *Linear Algebra and Its Applications*, 323:51-60, 2001.
[6] P. Lancaster and M. Tismenetsky. *The Theory of Matrices*. Academic Press, Orlando, 1985.
[7] A.S. Markus. *Introduction to the Spectral Theory of Polynomial Operator Pencils*. Amer. Math. Society, Providence, Translations of Math. Monographs, Vol. 71, 1988.