Uniqueness of matrix square roots under a numerical range condition

Charles R. Johnson a, Kazuyoshi Okubo b, *

a Department of Mathematics, College of William and Mary, P.O. Box 8795, Williamsburg, VA 23187-8795, USA
b Mathematics Laboratory, Hokkaido University of Education at Sapporo, Sapporo 002-8502, Japan

Received 1 February 2001; accepted 16 April 2001
Submitted by C.-K. Li

Dedicated to T. Ando on the occasion of his 70th birthday

Abstract

For $A \in M_n(\mathbb{C})$, let $W(A)$ denote the numerical range of $A$. It is shown that if $W(A) \cap (-\infty, 0) = \emptyset$, then $A$ has a unique square root $B \in M_n(\mathbb{C})$ with $W(B)$ in the closed right half plane. © 2002 Elsevier Science Inc. All rights reserved.

AMS classification: 15A18; 15A24; 15A60

Keywords: Matrix square root; Numerical range; Spectrum

Let $A \in M_n(\mathbb{C})$. We denote by $\sigma(A)$ the spectrum of $A$ and by $W(A)$ the numerical range of $A$, that is,

$W(A) = \{x^*Ax \mid x^*x = 1, \; x \in \mathbb{C}^n\}.$

By RHP we denote the open right half of the complex plane and by $\text{RHP}$ its closure. Let $\mathcal{I}$ be the imaginary axis of the complex plane and $\text{Re} z = \frac{1}{2}(z + \overline{z})$ for $z \in \mathbb{C}$. The convex hull of a subset $S$ of $\mathbb{C}$, denoted $\text{conv}(S)$, is the set of all convex combinations of all selections of points from $S$.

* Corresponding author.

E-mail addresses: crjohnso@math.wm.edu (C.R. Johnson), okubo@sap.hokkyodai.ac.jp (K. Okubo).
In general, if \( A \) has a square root, then it will have many square roots. But if some conditions are imposed upon \( A \) or its square root, then the uniqueness of a square root can be guaranteed. In [1], spectral conditions have been used to give such a result.

**Theorem A.** Let \( A \in M_n(\mathbb{C}) \) be such that \( \sigma(A) \cap (-\infty, 0] = \emptyset \). Then there is a unique \( B \in M_n(\mathbb{C}) \) such that \( B^2 = A \) with \( \sigma(B) \subset \text{RHP} \).

Using Theorem A, in [1], we have given an alternative proof for the following result of [3]. See also [4].

**Theorem B.** Let \( A \in M_n(\mathbb{C}) \) be such that \( W(A) \cap (-\infty, 0] = \emptyset \). Then there is a unique \( B \in M_n(\mathbb{C}) \) such that \( B^2 = A \) with \( W(B) \subset \text{RHP} \).

Related to this, a natural question arose. If \( W(A) \cap (-\infty, 0) = \emptyset \), then does it preserve the uniqueness of \( B \) by replacing \( \text{RHP} \) with \( \overline{\text{RHP}} \) in Theorem B?

We give an affirmative answer to this question. We begin with a simple example. Let
\[
A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}
\]
Then \( 0 \not\in \sigma(A) \) and \( 0 \in W(A) = \{ z \in \mathbb{C} | |z - 1| \leq 1 \} \). But if we put
\[
B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},
\]
then \( B^2 = A \) and \( W(B) = \{ z \in \mathbb{C} | |z - 1| \leq \frac{1}{2} \} \subset \text{RHP} \). It is known that \( B \) is the unique square root of \( A \) with \( W(B) \subset \text{RHP} \) by Theorem A. This example shows that there is a possibility of changing the condition of \( W(A) \cap (-\infty, 0) = \emptyset \) to \( W(A) \cap (-\infty, 0) = \emptyset \) in Theorem B. In fact, we have the following, which we use to prove our main result, Theorem 2.

**Theorem 1.** Let \( A \in M_n(\mathbb{C}) \) be nonsingular and \( W(A) \cap (-\infty, 0) = \emptyset \). Then there is a unique \( B \in M_n(\mathbb{C}) \) such that \( B^2 = A \) and \( W(B) \subset \text{RHP} \).

**Proof.** Since \( W(A) \cap (-\infty, 0) = \emptyset \), there is \( \theta \) such that \( |\theta| \leq \frac{1}{2} \pi \) and \( W(e^{i\theta} A) \subset \overline{\text{RHP}} \). Put \( C = e^{i\theta} A \). Then this is equivalent to \( C + C^* \geq 0 \) (see [2, 1.2.5b]). If \( \sigma(C) \cap \mathcal{I} \neq \emptyset \), \( \sigma(C) \cap \mathcal{I} = \{ \lambda_{m+1}, \ldots, \lambda_n \} \) is included in \( \partial W(C) \), the boundary of \( W(C) \) since \( \sigma(C) \subset \overline{\text{RHP}} \). Then \( \lambda_j, j = m + 1, \ldots, n \), are normal eigenvalues of \( A \) (see [2, Theorem 1.6.6]); hence, after unitary similarity we can write
\[
C = \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix}
\]
with \( \sigma(C_1) \subset \text{RHP} \) and \( C_2 = \text{diag}(\lambda_{m+1}, \ldots, \lambda_n) \). From Theorem A there is unique \( D_1 \in M_m(\mathbb{C}) \) such that \( D_1^2 = C_1 \) and \( \sigma(D_1) \subset \{ z \in \mathbb{C} \setminus \{0\} | \text{Arg}(z) < \frac{1}{2} \pi \} \). Then we have
\[
D_1(D_1 + D_1^*) + (D_1 + D_1^*)D_1^* = C_1 + C_1^* + 2D_1D_1^*.
\]
Since $C_1 + C_1^* \geq 0$ and $D_1 D_1^* > 0$ the right-hand side of the above equation is positive definite and $D_1$ is positive stable, Lyapunov’s Theorem [2, Theorem 2.2.1] implies that $D_1 + D_1^* > 0$. Let $B_1 = e^{-i\theta/2} D_1$. Then, $B_1^2 = e^{-i\theta} C_1$ and $B_1 + B_1^* > 0$ as the proof of Theorem B. This implies that $W(B_1) \subseteq \text{RHP}$. Now, put $B_2 = \text{diag}(e^{-i\theta/2\sqrt{\lambda_{m+1}}}, \ldots, e^{-i\theta/2\sqrt{\lambda_n}})$. Here, for $m + 1 \leq j \leq n$, since $e^{-i\theta_1} \lambda_j \notin (-\infty, 0]$, we take $e^{-i\theta/2\sqrt{\lambda_j}} \in \text{RHP}$. Since $W(B_2) = \text{conv}(e^{-i\theta/2\sqrt{\lambda_{m+1}}}, \ldots, e^{-i\theta/2\sqrt{\lambda_n}})$, $W(B_2) \subseteq \text{RHP}$. Let $B = B_1 \oplus B_2$. Then it is easy to see $B^2 = A$ and $W(B) = \text{conv}(W(B_1) \cup W(B_2)) \subseteq \text{RHP}$. The uniqueness of $B$ follows from Theorem A. □

**Theorem 2.** Let $A \in M_n(\mathbb{C})$ be such that $W(A) \cap (-\infty, 0) = \emptyset$. Then there is a unique $B \in M_n(\mathbb{C})$ such that $B^2 = A$ with $W(B) \subseteq \overline{\text{RHP}}$.

**Proof.** If $W(A) \cap (-\infty, 0) = \emptyset$, then the statement of Theorem 2 follows from Theorem B. So, consider the case of $0 \notin W(A)$. If $0 \notin \sigma(A)$, Theorem 1 implies the claim of Theorem 2. We treat only the case $0 \in \sigma(A)$. Since $W(A) \cap (-\infty, 0) = \emptyset$, $0 \in \sigma(A) \cap \partial W(A)$, so $0$ is a normal eigenvalue of $A$. Thus, up to unitary similarity

$$A = \begin{pmatrix} A' & 0 \\ 0 & 0 \end{pmatrix}$$

in which $A' \in M_m(\mathbb{C})$ and $0 \notin \sigma(A')$. By Theorem 1, there is a square root $B'$ of $A'$ with $W(B') \subseteq \text{RHP}$. If we put $B = B' \oplus 0 \in M_n(\mathbb{C})$, $B^2 = B'^2 \oplus 0 = A' \oplus 0 = A$ and $W(B) = \text{conv}(W(B') \cup \{0\}) \subseteq \overline{\text{RHP}}$. Thus, we have only to verify the uniqueness of $B$. Let $B^2 = A$ and

$$B = (b_{ij}) = \begin{pmatrix} B_{11} & \beta \\ \alpha & b_{nn} \end{pmatrix},$$

with $B_{11} \in M_{n-1}$,

$$\beta = (b_{1n} \ b_{2n} \cdots \ b_{n-1,n})^t \quad \text{and} \quad \alpha = (b_{n1} \ b_{n2} \cdots \ b_{n,n-1}).$$

Since

$$A = B^2 = \begin{pmatrix} B_{11}^2 + \beta \alpha & (B_{11} + b_{nn} I) \beta \\ \alpha (B_{11} + b_{nn} I) & \alpha \beta + b_{nn}^2 \end{pmatrix},$$

we have

1. $B_{11}^2 + \beta \alpha = A' \oplus 0 \in M_{n-1}(\mathbb{C})$,
2. $(B_{11} + b_{nn} I) \beta = 0$,
3. $\alpha (B_{11} + b_{nn} I) = 0$,
4. $\alpha \beta + b_{nn}^2 = 0$.

By (1), (2) and (4),

$$(A' \oplus 0) \beta = B_{11}^2 \beta + \beta \alpha \beta = (b_{nn}^2 - b_{nn}^2) \beta = 0,$$

that is, $\beta$ belongs to the kernel of $A' \oplus 0$. This means $b_{1n} = b_{2n} = \cdots = b_{nn} = 0$. Similarly, using (1), (3) and (4), that must have $b_{n1} = b_{n2} = \cdots = b_{nn} = 0$. If $m =$
n − 1, then it is easy to see that B must have the form $B = B_{11} \oplus B_{nn}$ and $B^2 = A$ implies $B = B_{11} \oplus 0$. By Theorem 1 there is unique $B_{11}$ such that $B_{11}^2 = A'$. If $m < n − 1$, the same argument may be applied to $B_{11}$ to produce another 0 direct summand, and the procedure may be continued for as many 0 eigenvalues as there are of A. Thus, B must be of the form $B = B' \oplus 0$, with $B' \in M_m(\mathbb{C})$ and $B'^2 = A'$. The uniqueness of $B'$ follows from Theorem 1, so that $B$ is unique, completing the proof. □

The following is a natural question.

**Question.** Which matrix A has a unique square root $A^{1/2}$ such that $W(A^{1/2}) \subset \mathbb{R}^+$?

We do not have complete solution for this question. But it is easy to show that a normal matrix A with $\sigma(A) \cap (-\infty, 0) = \emptyset$ is one of the solutions of above question. For $2 \times 2$ matrices, we can give a complete solution. By Schur’s triangularization we may assume

$$A = \begin{pmatrix} \lambda_1 & \alpha \\ 0 & \lambda_2 \end{pmatrix} \in M_2(\mathbb{C}).$$

Let $\sqrt{\lambda_1}$, $\sqrt{\lambda_2}$ be square roots of $\lambda_1$, and $\lambda_2$ with $\text{Re} \sqrt{\lambda_1} \geq 0$ and $\text{Re} \sqrt{\lambda_2} \geq 0$, respectively. Then we have the following;

**Theorem 3.** A has a unique square root $B$ such that $W(B) \subset \mathbb{R}^+$ if and only if $|\alpha|^2 \leq 4 \text{Re} \sqrt{\lambda_1} \cdot \text{Re} \sqrt{\lambda_2} |\sqrt{\lambda_1} + \sqrt{\lambda_2}|^2$.

**Proof.** Let $\lambda_1, \lambda_2 \notin (-\infty, 0]$. Then the unique square root $B$ of $A$ with $\sigma(B) \subset \mathbb{R}^+$ is given by

$$B = \begin{pmatrix} \sqrt{\lambda_1} & \alpha \\ 0 & \sqrt{\lambda_1 + \lambda_2} \end{pmatrix} \begin{pmatrix} \sqrt{\lambda_2} \\ 0 \end{pmatrix}$$

with $\text{Re} \sqrt{\lambda_1} > 0$ and $\text{Re} \sqrt{\lambda_2} > 0$. Then the numerical range of $B$ is exactly given by using $W(A)$. In fact, since, for $x, y \in \mathbb{C}$ we can calculate

$$\langle B \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \rangle = \sqrt{\lambda_1} |x|^2 + \frac{\alpha}{\sqrt{\lambda_1 + \lambda_2}} y \bar{x} + \sqrt{\lambda_2} |y|^2 = \frac{1}{\sqrt{\lambda_1} + \sqrt{\lambda_2}} \langle A \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \rangle + \frac{\sqrt{\lambda_1} \sqrt{\lambda_2}}{\sqrt{\lambda_1} + \sqrt{\lambda_2}}.$$

Hence we have

$$W(B) = \frac{1}{\sqrt{\lambda_1} + \sqrt{\lambda_2}} W(A) + \frac{\sqrt{\lambda_1} \sqrt{\lambda_2}}{\sqrt{\lambda_1} + \sqrt{\lambda_2}}.$$

Moreover, as we mentioned in the proof of Theorem 2, it is known that $W(B) \subset \mathbb{R}^+$ if and only if
\( B + B^* = \begin{pmatrix} 2 \text{Re} \sqrt{\lambda_1} & \frac{\alpha}{\sqrt{\lambda_1 + \sqrt{\lambda_2}}} \\ \frac{\alpha}{\sqrt{\lambda_1 + \sqrt{\lambda_2}}} & 2 \text{Re} \sqrt{\lambda_2} \end{pmatrix} \)

is positive semidefinite. Since \( \text{Re} \sqrt{\lambda_1}, \text{Re} \sqrt{\lambda_2} > 0 \), this is equivalent to

\[
0 \leq \det(B + B^*) = 4 \text{Re} \sqrt{\lambda_1} \cdot \text{Re} \sqrt{\lambda_2} - \left| \frac{\alpha}{\sqrt{\lambda_1 + \sqrt{\lambda_2}}} \right|^2.
\]

If one of \( \lambda_1 \) or \( \lambda_2 \) is zero (we may assume \( \lambda_2 = 0 \)), then the square root \( B \) of \( A \) has a form

\[
B = \begin{pmatrix} \sqrt{\lambda_1} & \frac{\alpha}{\sqrt{\lambda_1}} \\ 0 & 0 \end{pmatrix}
\]

by simple computation. Hence if \( \alpha \neq 0 \), 0 is an interior point of \( W(B) \), so we can show \( W(B) \subset \text{RHP} \) if and only if \( \alpha = 0 \). If \( \lambda_1 = 0 \) and \( \lambda_2 = 0 \), then it is easy to see that \( A \) has square roots if and only if \( \alpha = 0 \). Then the only square root \( B \) of \( A \) with \( W(B) \subset \text{RHP} \) is zero matrix. Thus, the proof of Theorem 3 is complete. \( \square \)

References