Palestine Journal of Mathematics

Vol. 11 (1) (2022) , 6–18

ON A NEW IDENTITY FOR DIAGONAL TERMS OF 2×2 MATRIX ROOTS

Peter J. Larcombe, Eric J. Fennessey, Lee Rawlin and James Stanton

Communicated by Ayman Badawi

MSC 2010 Classifications: Primary 11C20.

Keywords and phrases: General 2×2 matrix roots, diagonal terms identity, matrix root generators.

Abstract We state and prove a condition for an identity which involves the diagonal entries of any nth root of a general 2×2 matrix, and develop a method for obtaining all roots through a so called generator pair of parameters defined by its eigenvalues.

1 Introduction

Let

$$\mathbf{M} = \mathbf{M}(A, B, C, D) = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{1.1}$$

be a general (real) 2×2 matrix. Previous work on the invariance of the anti-diagonals ratio B/C with respect to (integer) matrix power has resulted in a number of publications in which a variety of proof methodologies are described [2, 4, 5, 7] (the phenomenon also extends to multiple anti-diagonal ratios for a tri-diagonal matrix of arbitrary dimension [3, 6]). In this paper, based on a diagonalisation of $\mathbf{M}$, we establish a condition for a new identity involving the diagonal entries of any nth root of $\mathbf{M}$ (which, interestingly, as a corollary is guaranteed to hold in the square roots case $n = 2$), and in doing so create a method for securing all matrix roots (n^2 in total) through a so called generator pair of parameters defined by its eigenvalues in the case when the latter are real and distinct; some illustrative examples, with supporting computations, are provided.

2 A Preliminary Result

Assuming $\mathbf{M}$ is diagonalisable (we take A, B, C, D to be real), then

$$\mathbf{M} = \mathbf{\Omega} \mathbf{D} \mathbf{\Omega}^{-1} \tag{2.1}$$

for

$$\mathbf{\Omega} = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \tag{2.2}$$

whose entries are found to be

$$\begin{aligned} p &= p(A, B, C, D) = -2B/(A - D + K), \\ q &= q(A, B, C, D) = -2B/(A - D - K), \\ r &= 1, \\ s &= 1, \end{aligned}$$

$$\begin{aligned} \lambda &= \lambda(A, B, C, D) = (A + D - K)/2, \\ \mu &= \mu(A, B, C, D) = (A + D + K)/2, \end{aligned} \tag{2.3}$$

where the parameter K is (writing the trace of $\mathbf{M}$ as $\mathrm{Tr}\{\mathbf{M}\} = A + D$)

$$K^2 = K^2(A, B, C, D) = (A - D)^2 + 4BC = \mathrm{Tr}^2\{\mathbf{M}\} - 4|\mathbf{M}|; \tag{2.4}$$

the variables $\lambda, \mu = \frac{1}{2}(\text{Tr}\{\mathbf{M}\} \mp K)$ in $\mathbf{D}$ are the eigenvalues of $\mathbf{M}$ in terms of A, B, C, D (the columns of $\boldsymbol{\Omega}$ being their associated eigenvectors $(p,1)^T, (q,1)^T$ (with T denoting transposition)), and take distinct values for $K \neq 0$ which guarantee $|\boldsymbol{\Omega}| = p - q \neq 0$ (and so the invertibility of $\boldsymbol{\Omega}$). Our main result of the paper is based on the constraint $K^2 > 0$, so that λ and μ are real, as well as distinct.

We begin with a preliminary result (that is, Theorem 2.1)—with two lemmas as precursors—which reveals a fundamental property of all root matrices of $\mathbf{M}$, in readiness for those that we generate to use in examples accompanying our analysis.

Lemma 2.1. *We write, for $n \geq 1$,*

$$\mathbf{M}^n = \begin{pmatrix} A & B \\ C & D \end{pmatrix}^n = \begin{pmatrix} \alpha_n & \beta_n \\ \gamma_n & \delta_n \end{pmatrix}$$

for the nth integer power of $\mathbf{M}$ (where $\alpha_1 = A, \ldots, \delta_1 = D$), and, with $\rho_1 = 1$, generate $\rho_2, \rho_3, \rho_4, \ldots$, through the recursion $\rho_n = \alpha_{n-1} + D\rho_{n-1}$ ($n \geq 2$). Then

$$\begin{aligned} \beta_n &= B\rho_n, \\ \gamma_n &= C\rho_n, \\ \delta_n &= \alpha_n - (A-D)\rho_n. \end{aligned}$$

Proof. We proceed by induction, utilising a 2004 result of McLaughlin [8, Theorem 1, p. 3] (where he gives closed forms for the elements of a 2×2 integer matrix power). Our result holds for $n = 1$, recovering $\beta_1 = B\rho_1 = B$, $\gamma_1 = C\rho_1 = C$ and $\delta_1 = \alpha_1 - (A-D)\rho_1 = A-(A-D) = D$. Assuming the stated relations hold for some $n = k \geq 1$, then

$$\begin{aligned} \begin{pmatrix} \alpha_{k+1} & \beta_{k+1} \\ \gamma_{k+1} & \delta_{k+1} \end{pmatrix} = \mathbf{M}^{k+1} &= \begin{pmatrix} \alpha_k & \beta_k \\ \gamma_k & \delta_k \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} \\ &= \begin{pmatrix} A\alpha_k + C\beta_k & B\alpha_k + D\beta_k \\ A\gamma_k + C\delta_k & B\gamma_k + D\delta_k \end{pmatrix}, \end{aligned} \tag{L.1}$$

whence, noting that

$$\alpha_{k+1} = A\alpha_k + C\beta_k, \tag{L.2}$$

we find, as required,

$$\begin{aligned} \beta_{k+1} &= B\alpha_k + D\beta_k \\ &= B\alpha_k + D(B\rho_k) \qquad \text{(by assumption)} \\ &= B(\alpha_k + D\rho_k) \\ &= B\rho_{k+1}, \end{aligned}$$

$$\begin{aligned} \gamma_{k+1} &= A\gamma_k + C\delta_k \\ &= A(C\rho_k) + C[\alpha_k - (A-D)\rho_k] \qquad \text{(by assumption)} \\ &= C(\alpha_k + D\rho_k) \\ &= C\rho_{k+1}, \end{aligned}$$

$$\begin{aligned} \delta_{k+1} &= B\gamma_k + D\delta_k \\ &= B(C\rho_k) + D[\alpha_k - (A-D)\rho_k] \qquad \text{(by assumption)} \\ &= C(B\rho_k) + D[\alpha_k - (A-D)\rho_k] - A\alpha_k + A\alpha_k \\ &= A\alpha_k + C\beta_k - (A-D)(\alpha_k + D\rho_k) \qquad \text{(by assumption)} \\ &= \alpha_{k+1} - (A-D)(\alpha_k + D\rho_k) \qquad \text{(by (L.2))} \\ &= \alpha_{k+1} - (A-D)\rho_{k+1}, \end{aligned} \tag{L.3}$$

which completes the proof. □

Lemma 2.2. *If, for $n \geq 1$, the nth integer power of $\mathbf{M}$ is a diagonal matrix with distinct entries, then $\mathbf{M}$ must itself be diagonal.*

Proof. Suppose

$$\mathbf{M}^n = \begin{pmatrix} \alpha_n & \beta_n \\ \gamma_n & \delta_n \end{pmatrix} = \begin{pmatrix} m_u & 0 \\ 0 & m_l \end{pmatrix}, \tag{L.4}$$

with $m_u \neq m_l$. We argue by contradiction, assuming $\mathbf{M}$ is non-diagonal in which case either B or C is non-zero. W.l.o.g., suppose $B \neq 0$. Then we have $0 = \beta_n = B\rho_n$ (by Lemma 2.1) $\Rightarrow \rho_n = 0$. Thus, $m_l = \delta_n = \alpha_n - (A-D)\rho_n$ (by Lemma 2.1) $= \alpha_n = m_u$, which is a contradiction. □

As a point of interest for the reader, a generalised version of Lemma 2.2 (for a matrix of arbitrary dimension) appears, with proof, in Appendix A.

Theorem 2.1. *For $K \neq 0$, suppose $\mathbf{R}$ is an nth root matrix of $\mathbf{M}$. Then there exists a diagonal nth root matrix $\hat{\mathbf{D}}$ of $\mathbf{D}$ such that $\mathbf{R} = \mathbf{\Omega}\hat{\mathbf{D}}\mathbf{\Omega}^{-1}$.*

Proof. Since $\mathbf{R}^n = \mathbf{M}$ we can write, from (2.1), $\mathbf{R}^n = \mathbf{\Omega}\mathbf{D}\mathbf{\Omega}^{-1} \Rightarrow \mathbf{D} = \mathbf{\Omega}^{-1}\mathbf{R}^n\mathbf{\Omega} = (\mathbf{\Omega}^{-1}\mathbf{R}\mathbf{\Omega})^n$. Since $\mathbf{D}$ is diagonal with distinct entries, then the matrix $(\mathbf{\Omega}^{-1}\mathbf{R}\mathbf{\Omega})^n$ has these properties too, and by Lemma 2.2 $\mathbf{\Omega}^{-1}\mathbf{R}\mathbf{\Omega}$ is diagonal. It follows, therefore, writing $\mathbf{\Omega}^{-1}\mathbf{R}\mathbf{\Omega}$ as the diagonal matrix $\hat{\mathbf{D}}$, say, that $\mathbf{D} = (\mathbf{\Omega}^{-1}\mathbf{R}\mathbf{\Omega})^n = \hat{\mathbf{D}}^n$, whence $\hat{\mathbf{D}} = \mathbf{D}^{1/n}$ is an nth root of $\mathbf{D}$ for which $\mathbf{R} = \mathbf{\Omega}\hat{\mathbf{D}}\mathbf{\Omega}^{-1}$. □

We illustrate Theorem 2.1. Let $\mathbf{R}_i(n) = \mathbf{M}^{1/n}$ be an nth root matrix of $\mathbf{M}$ ($i = 1, \ldots, n$). For any given $n \geq 2$, we appeal to an algorithm of Choudhry [1] (see Remark 2.1) to extract n root matrices $\mathbf{R}_1(n), \ldots, \mathbf{R}_n(n)$ in a systematic fashion (each of which could in principle be multiplied by an nth root of unity to offer a total of n^2 root matrices). In the (square roots) case $n = 2$, for example, with $\mathbf{M}$ assigned (resp.) values $A, B, C, D = 1, 2, 3, 4$, we compute

$$\begin{aligned}
\mathbf{R}_1(2) &= \sqrt{5+2\sqrt{2}i}\begin{pmatrix} \frac{1}{11}(3+\sqrt{2}i) & \frac{1}{33}(10-4\sqrt{2}i) \\ \frac{1}{11}(5-2\sqrt{2}i) & \frac{1}{11}(8-\sqrt{2}i) \end{pmatrix} \\
&= \begin{pmatrix} 0.55368857+0.46439416i & 0.80696073-0.21242648i \\ 1.2104411-0.31863972i & 1.7641297+0.14575444i \end{pmatrix}, \\
\mathbf{R}_2(2) &= \sqrt{5-2\sqrt{2}i}\begin{pmatrix} \frac{1}{11}(3-\sqrt{2}i) & \frac{1}{33}(10+4\sqrt{2}i) \\ \frac{1}{11}(5+2\sqrt{2}i) & \frac{1}{11}(8+\sqrt{2}i) \end{pmatrix} \\
&= \begin{pmatrix} 0.55368857-0.46439416i & 0.80696073+0.21242648i \\ 1.2104411+0.31863972i & 1.7641297-0.14575444i \end{pmatrix},
\end{aligned} \tag{2.5}$$

and (noting that from (2.4) $K^2 = 33 > 0$), with

$$\mathbf{\Omega} = \begin{pmatrix} \frac{4}{3-\sqrt{33}} & \frac{4}{3+\sqrt{33}} \\ 1 & 1 \end{pmatrix}, \qquad \mathbf{D} = \begin{pmatrix} \frac{1}{2}(5-\sqrt{33}) & 0 \\ 0 & \frac{1}{2}(5+\sqrt{33}) \end{pmatrix} \tag{2.6}$$

using (2.3), we find, as expected by Theorem 2.1, $(\mathbf{\Omega}^{-1}\mathbf{R}_1(2)\mathbf{\Omega})^2 = (\mathbf{\Omega}^{-1}\mathbf{R}_2(2)\mathbf{\Omega})^2 = \mathbf{D}$.[1] The same matrix $\mathbf{M}(1,2,3,4)$ has two of its (four) Choudhry 4th root matrices delivered as

$$\mathbf{R}_3(4) = \begin{pmatrix} -0.056706113-0.42039152i & 0.72234344+0.19229847i \\ 1.0835152+0.28844771i & 1.0268090-0.13194380i \end{pmatrix} \tag{2.7}$$

[1]We thank Dr. James Clapperton for undertaking all computations related to this paper using the algebraic software package Maple.

and

$$\mathbf{R}_4(4) = \begin{pmatrix} 0.78407692 - 0.42039152i & 0.33774649 + 0.19229847i \\ 0.50661973 + 0.28844771i & 1.2906967 - 0.13194380i \end{pmatrix}, \tag{2.8}$$

for which it is easily checked that $(\mathbf{\Omega}^{-1}\mathbf{R}_3(4)\mathbf{\Omega})^4 = (\mathbf{\Omega}^{-1}\mathbf{R}_4(4)\mathbf{\Omega})^4 = \mathbf{D}$. Finally, one of the (ten) 10th root matrices of $\mathbf{M}(1, 2, 3, 4)$ is

$$\mathbf{R}_7(10) = \begin{pmatrix} -0.42711554 - 0.37918872i & 0.63319404 - 0.14464352i \\ 0.94979107 - 0.21696528i & 0.52267553 - 0.59615400i \end{pmatrix}, \tag{2.9}$$

for which $(\mathbf{\Omega}^{-1}\mathbf{R}_7(10)\mathbf{\Omega})^{10} = \mathbf{D}$.

By way of another example, we use the matrix $\mathbf{M}(3, -2, 5, 1)$ for which $K^2 = -36 < 0$ and

$$\mathbf{\Omega} = \begin{pmatrix} \frac{1}{5}(1 - 3i) & \frac{1}{5}(1 + 3i) \\ 1 & 1 \end{pmatrix}, \qquad \mathbf{D} = \begin{pmatrix} 2 - 3i & 0 \\ 0 & 2 + 3i \end{pmatrix}; \tag{2.10}$$

we confirm that one of the (five) 5th matrix roots of $\mathbf{M}$,

$$\mathbf{R}_2(5) = \begin{pmatrix} 1.3516393 & -0.16826559 \\ 0.42066398 & 1.1833737 \end{pmatrix}, \tag{2.11}$$

satisfies $(\mathbf{\Omega}^{-1}\mathbf{R}_2(5)\mathbf{\Omega})^5 = \mathbf{D}$, while the cube root matrix

$$\mathbf{R}_3(3) = \begin{pmatrix} 0.39998857 - 0.69280053i & -0.50134883 + 0.86836164i \\ 1.2533721 - 2.1709041i & -0.10136026 + 0.17556111i \end{pmatrix} \tag{2.12}$$

satisfies $(\mathbf{\Omega}^{-1}\mathbf{R}_3(3)\mathbf{\Omega})^3 = \mathbf{D}$. Many other matrix roots have been tested for the two matrices chosen here (and for other matrices where $K^2 \neq 0$).

Remark 2.1. The matrices we consider fall into the broad category of so called *non-scalar* matrices (a scalar matrix is a scalar multiple of the identity matrix, and is known to have an infinte number of roots), for which we expect to find a finite number of roots. According to Choudhry [1]—to whose paper the interested reader is directed for technical details and background reading—exactly how many roots exist (n, n^2 or none) depends on properties of $\mathbf{M}$ (its singularity, eigenvalues, and trace[2]). We follow Choudhry's methodology for the most part, particularly making use of Theorem 4.1 (pp. 189–190) therein. Much work has been conducted on matrix roots over the years, and a good overview of the various treatments of 2 × 2 matrix root extraction is given by Özdemir who includes a comparison of his approach (based on the notion of a so called hybrid number in conjunction with a 2 × 2 matrix representation of de Moivre) with others in terms of practicability, suitability, (dis)advantages, *etc.* [9, p. 23].

3 Main Result and Discussion

Theorem 3.1 (Stanton's Theorem). *For $K^2 > 0$, let α, β be (resp.) any nth roots of the eigenvalues λ, μ of a matrix $\mathbf{M}(A, B, C, D)$, with $A \neq D$. Then there exists an nth root matrix $\mathbf{R} = \mathbf{R}(\alpha, \beta; A, B, C, D)$ of $\mathbf{M}$ for which*

$$\mathrm{Re}\{R_{1,1}\}\mathrm{Im}\{R_{1,1}\} = \mathrm{Re}\{R_{2,2}\}\mathrm{Im}\{R_{2,2}\}$$

if $\mathrm{Im}\{\alpha^2\} = \mathrm{Im}\{\beta^2\}$.

Proof. First, define a diagonal matrix

$$\mathbf{D}^*(\alpha, \beta) = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}. \tag{P.1}$$

[2]We summarise as follows: If $\mathbf{M}$ is singular with zero trace then there are no roots; if it is non-singular with distinct eigenvalues then there are n^2 roots (a set of n that is multiplied up to one of n^2 by roots of unity (or else delivered separately, in full)); otherwise, a singular matrix with non-zero trace, or a non-singular matrix with equal eigenvalues, has n roots.

For $K^2 > 0$ the eigenvalues λ, μ of $\mathbf{M}$ are real and distinct. Setting $\lambda = \alpha^n$, $\mu = \beta^n$, then α, β are (resp.) nth roots of λ, μ, and $(\mathbf{D}^*)^n = \mathbf{D}$, the eigenvalues matrix of (2.2). Thus, the decomposition (2.1) gives

$$\begin{aligned} \mathbf{R} &= \mathbf{R}(\alpha, \beta; A, B, C, D) \\ &= \mathbf{M}^{1/n} \\ &= (\mathbf{\Omega D \Omega}^{-1})^{1/n} \\ &= \mathbf{\Omega D}^{1/n}\mathbf{\Omega}^{-1} \\ &= \mathbf{\Omega D}^*\mathbf{\Omega}^{-1}. \end{aligned} \tag{P.2}$$

Accordingly, we see that

$$\begin{aligned} \mathbf{R} &= \frac{1}{|\mathbf{\Omega}|}\begin{pmatrix} ps\alpha - qr\beta & pq(\beta-\alpha) \\ rs(\alpha-\beta) & ps\beta - qr\alpha \end{pmatrix} \\ &= \begin{pmatrix} R_{1,1} & R_{1,2} \\ R_{2,1} & R_{2,2} \end{pmatrix}, \end{aligned} \tag{P.3}$$

say, where, by (2.3), a little algebra delivers explicit forms

$$\begin{aligned} R_{1,1} &= [\alpha + \beta + (\beta - \alpha)(A-D)/K]/2, \\ R_{1,2} &= B(\beta-\alpha)/K, \\ R_{2,1} &= C(\beta-\alpha)/K, \\ R_{2,2} &= [\alpha + \beta - (\beta - \alpha)(A-D)/K]/2. \end{aligned} \tag{P.4}$$

Let us write $\sigma = (A-D)/K \neq 0$, and express the complex form of each eigenvalue root pair α, β as $\alpha = \alpha_r + i\alpha_c$ and $\beta = \beta_r + i\beta_c$, say. Noting that

$$\begin{aligned} \alpha + \beta &= \alpha_r + \beta_r + i(\alpha_c + \beta_c), \\ \beta - \alpha &= \beta_r - \alpha_r + i(\beta_c - \alpha_c), \end{aligned} \tag{P.5}$$

then

$$\begin{aligned} 2R_{1,1} &= \alpha + \beta + (\beta-\alpha)\sigma \\ &= \alpha_r(1-\sigma) + \beta_r(1+\sigma) + i[\alpha_c(1-\sigma) + \beta_c(1+\sigma)] \end{aligned} \tag{P.6}$$

and

$$\begin{aligned} 2R_{2,2} &= \alpha + \beta - (\beta-\alpha)\sigma \\ &= \alpha_r(1+\sigma) + \beta_r(1-\sigma) + i[\alpha_c(1+\sigma) + \beta_c(1-\sigma)], \end{aligned} \tag{P.7}$$

yielding

$$\begin{aligned} &\text{Re}\{R_{1,1}\}\text{Im}\{R_{1,1}\} \\ &\quad = [\alpha_r(1-\sigma) + \beta_r(1+\sigma)][\alpha_c(1-\sigma) + \beta_c(1+\sigma)]/4 \\ &\quad = [\alpha_r\alpha_c(1-\sigma)^2 + (\alpha_r\beta_c + \alpha_c\beta_r)(1-\sigma)(1+\sigma) + \beta_r\beta_c(1+\sigma)^2]/4 \end{aligned} \tag{P.8}$$

and

$$\begin{aligned} &\text{Re}\{R_{2,2}\}\text{Im}\{R_{2,2}\} \\ &\quad = [\alpha_r\alpha_c(1+\sigma)^2 + (\alpha_r\beta_c + \alpha_c\beta_r)(1+\sigma)(1-\sigma) + \beta_r\beta_c(1-\sigma)^2]/4, \end{aligned} \tag{P.9}$$

and in turn

$$\begin{aligned} \text{Re}\{R_{1,1}\}\text{Im}\{R_{1,1}\} - \text{Re}\{R_{2,2}\}\text{Im}\{R_{2,2}\} &= -\sigma(\alpha_r\alpha_c - \beta_r\beta_c) \\ &= 0 \end{aligned} \tag{P.10}$$

if $\alpha_r\alpha_c = \beta_r\beta_c$, that is, if $\text{Im}\{\alpha^2\} = \text{Im}\{\beta^2\}$. □

Corollary 3.1. In the (square roots) case $n = 2$, $\alpha^2 = \lambda \in \mathbb{R}$ and $\beta^2 = \mu \in \mathbb{R}$, $\Rightarrow \text{Im}\{\alpha^2\} = \text{Im}\{\beta^2\} = 0$ and the condition holds trivially; thus, Theorem 3.1 is guaranteed to be true for any square root matrices of $\mathbf{M}$.

We note also that, given nth root matrices $\mathbf{R}_1(n), \mathbf{R}_2(n), \ldots, \mathbf{R}_n(n)$ of a matrix $\mathbf{M}$ ($n \geq 2$), any product of $p \leq n$ of these root matrices constitutes an nth root of $\mathbf{M}^p$. Recalling that any pair of nth root matrices commute (an elementary reader exercise[3]), this follows by considering, w.l.o.g., $[\mathbf{R}_1(n) \cdot \mathbf{R}_2(n) \cdot \cdots \cdot \mathbf{R}_p(n)]^n = [\mathbf{R}_1(n)]^n \cdot [\mathbf{R}_2(n)]^n \cdot \cdots \cdot [\mathbf{R}_p(n)]^n = \mathbf{M} \cdot \mathbf{M} \cdot \cdots \cdot \mathbf{M} = \mathbf{M}^p$, whence $\mathbf{R}_1(n) \cdot \mathbf{R}_2(n) \cdot \cdots \cdot \mathbf{R}_p(n) = (\mathbf{M}^p)^{1/n}$.

Remark 3.1. We remark that the general nature of the elements of $\mathbf{R}(\alpha, \beta; A, B, C, D)$ in (P.4) confirms the invariance of the anti-diagonals ratio B/C of any matrix root, since $R_{1,2}/R_{2,1} = B/C$ (which is independent of eigenvalue nth roots $\alpha = \lambda^{1/n}$, $\beta = \mu^{1/n}$, and so n). This extends the anti-diagonals ratio invariance of 2×2 (integer) matrix powers [2, 4, 5, 7] to now cover inverse powers (that is, roots), a fact that is also available directly from Özdemir's work [9].[4]

3.1 On Square Roots

If a root matrix is purely real (or purely imaginary), then Theorem 3.1 is satisfied trivially, as in the Choudhry square roots

$$\mathbf{R}_1(2) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{and} \quad \mathbf{R}_2(2) = \frac{\sqrt{33}}{11} \begin{pmatrix} 3 & 10/3 \\ 5 & 8 \end{pmatrix} \tag{3.1}$$

of the matrix $\mathbf{M}(7, 10, 15, 22)$ (for which $K^2 = 825 > 0$), and roots

$$\mathbf{R}_1(2) = \begin{pmatrix} 2 & 5 \\ 14 & 42 \end{pmatrix} \quad \text{and} \quad \mathbf{R}_2(2) = \frac{\sqrt{470}}{47} \begin{pmatrix} 3 & 11 \\ 154/5 & 91 \end{pmatrix} \tag{3.2}$$

of the matrix $\mathbf{M}(74, 220, 616, 1834)$ (for which $K^2 = 3,639,680 > 0$)—we also include roots

$$\mathbf{R}_1(2) = i \begin{pmatrix} -1 & 1 \\ 4 & -3 \end{pmatrix} \quad \text{and} \quad \mathbf{R}_2(2) = \frac{i}{\sqrt{5}} \begin{pmatrix} 3 & -2 \\ -8 & 7 \end{pmatrix} \tag{3.3}$$

of the matrix $\mathbf{M}(-5, 4, 16, -13)$ (for which $K^2 = 320 > 0$). Where square root matrices are fully complex, they are conjugates of each other as seen in (2.5); another example is the square root matrix pair

$$\begin{aligned} \mathbf{R}_1(2) &= \begin{pmatrix} 1.5267117 + 2.8981618i & -1.8400672 + 2.0439247i \\ -2.1647850 + 2.4046173i & 2.6091042 + 1.6958531i \end{pmatrix}, \\ \mathbf{R}_2(2) &= \begin{pmatrix} 1.5267117 - 2.8981618i & -1.8400672 - 2.0439247i \\ -2.1647850 - 2.4046173i & 2.6091042 - 1.6958531i \end{pmatrix}, \end{aligned} \tag{3.4}$$

derived from the matrix $\mathbf{M}(-7, -17, -20, 3)$ for which $K^2 = 1460 > 0$; these illustrate the statement of Corollary 3.1.

As an aside, the same comments apply if we set $A = D$ (that is, $\sigma = 0$) with the additional observation that each matrix of any square root pair has its diagonal entries equal, as in the conjugate pair

$$\begin{aligned} \mathbf{R}_1(2) &= \begin{pmatrix} 0.93482766 + 0.84097329i & 1.4780923 - 1.3296955i \\ 0.59123692 - 0.53187821i & 0.93482766 + 0.84097329i \end{pmatrix}, \\ \mathbf{R}_2(2) &= \begin{pmatrix} 0.93482766 - 0.84097329i & 1.4780923 + 1.3296955i \\ 0.59123692 + 0.53187821i & 0.93482766 - 0.84097329i \end{pmatrix}, \end{aligned} \tag{3.5}$$

[3]Using the decomposition (P.2) of any nth root matrix, combined with the commutativity of diagonal matrices.

[4]This is clearly true for any non-zero $K(A, B, C, D)$, over and above $K^2 > 0$.

derived from the matrix $\mathbf{M}(\frac{1}{3}, 5, 2, \frac{1}{3})$ for which $K^2 = 40 > 0$—this is clear from (P.4) where, for *any* nth root matrix, $\sigma = 0 \Rightarrow R_{1,1} = R_{2,2} = (\alpha + \beta)/2$ (and so the root matrix has a diagonals ratio of unity, in which case Theorem 3.1 is automatically self-satisfying).

3.2 On Higher Order Roots

Example 1. Consider the matrix $\mathbf{M}(2, 2, 7, 15)$, with $K^2 = \text{Tr}^2\{\mathbf{M}\} - 4|\mathbf{M}| = 17^2 - 4(16) = 225 = 15^2 > 0$. Only two of the sixteen 16th Choudhry root matrices $\mathbf{R}_1(16), \ldots, \mathbf{R}_{16}(16)$ of $\mathbf{M}$ satisfy Theorem 3.1, and they do so trivially because they are real matrices; the other seven conjugate root matrix pairs fail the condition. It is easy enough, however, to construct a (non-trivial) 16th root matrix for which Theorem 3.1 does hold, described as follows. Noting that $\mathbf{M}$ has eigenvalues $\lambda = 1$, $\mu = 16$, then any solutions α, β of $\alpha^{16} = \lambda = 1$, $\beta^{16} = \mu = 16$ will, in defining a diagonal matrix $\mathbf{D}^*(\alpha, \beta)$ (P.1), produce a 16th root matrix of general form

$$\mathbf{R}(\alpha, \beta) = \frac{1}{15}\begin{pmatrix} 14\alpha + \beta & 2(\beta - \alpha) \\ 7(\beta - \alpha) & \alpha + 14\beta \end{pmatrix} \tag{3.6}$$

using (P.4). Choosing specific solutions

$$\begin{aligned} \alpha_s &= \exp(5\pi i/4) = -(1+i)/\sqrt{2}, \\ \beta_s &= 2^{1/4}\exp(3\pi i/8) = \left(\sqrt{2-\sqrt{2}} + \sqrt{2+\sqrt{2}}i\right)/2^{3/4}, \end{aligned} \tag{3.7}$$

yields a particular 16th root matrix

$$\mathbf{R}(\alpha_s, \beta_s) = \begin{pmatrix} -0.62962701 - 0.586720721i & 0.15495955 + 0.24077212i \\ 0.54235843 + 0.84270242i & 0.37761008 + 0.97829805i \end{pmatrix} \tag{3.8}$$

of $\mathbf{M}$ for which Theorem 3.1 indeed holds because

$$\text{Im}\{\alpha_s^2\} = \text{Im}\{i\} = 1 = \text{Im}\{-1 + i\} = \text{Im}\{\beta_s^2\}. \tag{3.9}$$

Moreover, the 16 choices for each of α, β combine to offer 16^2 root matrices $R(\alpha, \beta)$ (3.6) which are found to correspond *precisely* to those 16 Choudhry root matrices $\mathbf{R}_1(16), \ldots, \mathbf{R}_{16}(16)$ that are each multiplied by a 16th root of unity (that is to say, the collection of composites $\omega_j \mathbf{R}_i(16)$ $(i, j = 1, \ldots, 16)$, where $\omega_1, \ldots, \omega_{16}$ are the 16th roots of unity), of which 32, in fact, satisfy Theorem 3.1 inclusive of the matrix $\mathbf{R}(\alpha_s, \beta_s)$ (3.8) (this represents $\frac{1}{8}$ of the full set of root matrices, with the real, imaginary and fully complex matrix roots numbering (resp.) 4,4 and 24).[5]

Clearly, it is possible to access n^2 nth root matrices of an arbitrary matrix $\mathbf{M}(A, B, C, D)$ based on all pairwise combinations of those nth roots α, β of the eigenvalues of $\mathbf{M}$ which act as a *generating pair* of system parameters; the general form entries (P.4) of the root matrix (P.3) are evidently all dependent on a linear combination of these generators in connection with the matrix variables A, B, C, D and $K(A, B, C, D)$ of $\mathbf{M}$ as shown. For the purpose of illustration we finish with a second example where a root matrix is constructed for which Theorem 3.1 holds.

Example 2. Consider now the matrix $\mathbf{M}(5, 16, 118, 32)$, with $K^2 = \text{Tr}^2\{\mathbf{M}\} - 4|\mathbf{M}| = 37^2 - 4(-1728) = 8,281 = 91^2 > 0$. This time we choose 12th roots

$$\begin{aligned} \alpha_s &= 3^{1/4}\exp(\pi i/4) = 3^{1/4}(1+i)/\sqrt{2}, \\ \beta_s &= \sqrt{2}\exp(\pi i/6) = (\sqrt{3}+i)/\sqrt{2}, \end{aligned} \tag{3.10}$$

of its (resp.) eigenvalues $\lambda = -27$, $\mu = 64$ which define $\mathbf{D}^*(\alpha_s, \beta_s)$ and identify a particular 12th root matrix (from (P.4))

$$\mathbf{R}(\alpha_s, \beta_s) = \begin{pmatrix} 1.03403871 + 0.8520121282i & 0.0517169253 - 0.0392963653i \\ 0.381412324 - 0.2898106949i & 1.121311021 + 0.7856995118i \end{pmatrix} \tag{3.11}$$

[5] Of course the five pairs of square root matrices seen in (3.1)-(3.5) can each be multiplied by -1 (the (non-trivial) square root of of unity) to produce a complete group of $4 = 2^2$ in each example.

of $\mathbf{M}$ for which Theorem 3.1 holds (in this case $\text{Im}\{\alpha_s^2\} = \text{Im}\{\sqrt{3}i\} = \sqrt{3} = \text{Im}\{1+\sqrt{3}i\} = \text{Im}\{\beta_s^2\}$). Linear combinations of the generator pair α, β offer up all 12th root matrices which assume a general form

$$\mathbf{R}(\alpha,\beta) \quad = \quad \frac{1}{91}\begin{pmatrix} 59\alpha+32\beta & 16(\beta-\alpha) \\ 118(\beta-\alpha) & 32\alpha+59\beta \end{pmatrix}, \tag{3.12}$$

of which $\mathbf{R}(\alpha_s, \beta_s)$ (3.11) is one instance, with all 12^2 matrix roots of $\mathbf{M}$ matching those extracted via Choudhry's algorithm accordingly ($\frac{1}{9}$ of which satisfy Theorem 3.1, all fully complex ones).

Remark 3.2. It is immediate from (P.3) (or (P.4)) that every root matrix $\mathbf{R} = \mathbf{R}(\alpha,\beta)$ has the property that $R_{1,1} + R_{2,2} = \alpha + \beta$ (as the matrices in (3.6) and (3.12) bear out);[6] a similar feature is exhibited by the sum of the diagonal terms of the square of every such root matrix, and this observation can be generalised (see Remark B.1 of Appendix B). For interest, Appendix B states and proves a couple of results cast in terms of $\mathbf{R}^2$ when Theorem 3.1 holds.

Remark 3.3. In fact, with $R_{1,1}+R_{2,2} = \alpha+\beta$ we can derive Theorem 3.1 in a slightly different manner (included here for completeness), for noting (from (P.4)) that $R_{1,1} - R_{2,2} = (\beta-\alpha)\sigma$, we can write $\text{Re}\{R_{1,1}\}\text{Im}\{R_{1,1}\} - \text{Re}\{R_{2,2}\}\text{Im}\{R_{2,2}\} = \frac{1}{2}[\text{Im}\{(R_{1,1})^2\} - \text{Im}\{(R_{2,2})^2\}] = \frac{1}{2}\text{Im}\{(R_{1,1})^2 - (R_{2,2})^2\} = \frac{1}{2}\text{Im}\{[R_{1,1}+R_{2,2}][R_{1,1}-R_{2,2}]\} = \frac{1}{2}\text{Im}\{(\alpha+\beta)\cdot(\beta-\alpha)\sigma\} = \frac{1}{2}\sigma\text{Im}\{\beta^2-\alpha^2\} = 0$ if $0 = \text{Im}\{\beta^2-\alpha^2\} = \text{Im}\{\beta^2\} - \text{Im}\{\alpha^2\}$. Our thanks go to Prof. Sam Northshield (Plattsburgh State University of New York) for pointing this out to the author P.J.L. in a private communication.

4 Condition Counts

We have seen that the sets $\{\mathbf{R}(\alpha,\beta) : \alpha^n = \lambda, \beta^n = \mu\}$ and $\{\omega_j \mathbf{R}_i(n) : i,j = 1,\ldots,n,\ \omega^n = 1\}$ contain all (that is, n^2) nth roots of any given matrix $\mathbf{M}(A,B,C,D)$. We now turn our attention to counting those root matrices for which Theorem 3.1 holds, defining $\mathcal{C}_n(A,B,C,D)$ to be the number of nth root matrices of $\mathbf{M}(A,B,C,D)$ that satisfy the result.

4.1 Tr{M} ≠ 0

For $\mathbf{M}(2,2,7,15)$ of Example 1, it is found that $\{\mathcal{C}_n(2,2,7,15)\}_{n=2}^{\infty}$ forms the sequence $\{4,1,16,1,4,1,16,1,4,1,16,1,4,1,32,1,4,1,16,1,4,1,16,1,\ldots\}$, while Example 2 gives $\{\mathcal{C}_n(5,16,118,32)\}_{n=2}^{\infty} = \{4,1,0,1,4,1,0,1,4,1,16,1,4,1,0,1,4,1,0,1,\ldots\}$. The anomaly values 32 (Example 1) and 16 (Example 2) prevent each sequence looking like period 4 ones (we have taken computations as far as $n = 50$, at which point $50^2 = 2,500$ 50th root matrices are being tested against Theorem 3.1). Periodic instances do seem to occur, however, as seen in $\{\mathcal{C}_n(4,1,3,-2)\}_{n=2}^{\infty} = \{4,1,0,1,4,1,0,1,\ldots\}$ ($K^2 = 48$; $\lambda = 1-2\sqrt{3} < 0$, $\mu = 1+2\sqrt{3} > 0$) and $\{\mathcal{C}_n(-3,2,8,-10)\}_{n=2}^{\infty} = \{4,1,0,1,4,1,0,1,\ldots\}$ ($K^2 = 113$; $\lambda = \frac{1}{2}(-13-\sqrt{113}) < 0$, $\mu = \frac{1}{2}(-13+\sqrt{113}) < 0$) (the same sequence appears to arise in other examples such as $\{\mathcal{C}_n(-2,2,1,10)\}_{n=2}^{\infty}$ ($K^2 = 152$; $\lambda = 4-\sqrt{38} < 0$, $\mu = 4+\sqrt{38} > 0$) and $\{\mathcal{C}_n(-11,7,14,-13)\}_{n=2}^{\infty}$ ($K^2 = 396$; $\lambda = -12-3\sqrt{11} < 0$, $\mu = -12+3\sqrt{11} < 0$)). We also find $\{\mathcal{C}_n(3,3/2,6,4)\}_{n=2}^{\infty} = \{4,1,16,1,4,1,16,1,\ldots\}$ ($K^2 = 37$; $\lambda = \frac{1}{2}(7-\sqrt{37}) > 0$, $\mu = \frac{1}{2}(7+\sqrt{37}) > 0$) (repeated as $\{\mathcal{C}_n(1,1,-6,6)\}_{n=2}^{\infty}$ ($K^2 = 1$; $\lambda = 3 > 0$, $\mu = 4 > 0$) and $\{\mathcal{C}_n(10,5,4,7)\}_{n=2}^{\infty}$ ($K^2 = 89$; $\lambda = \frac{1}{2}(17-\sqrt{89}) > 0$, $\mu = \frac{1}{2}(17+\sqrt{89}) > 0$)). *From the computational tests conducted, the conclusion to which one is led is that $\{\mathcal{C}_n(A,B,C,D)\}_{n=2}^{\infty}$ is, for the most part (that is, with a relatively small number of exceptions through occasional 'rogue' elements), the period 4 sequence $\{4,1,0,1,\ldots\}$ for eigenvalues λ,μ of different sign (resp., $<0, >0$) or both negative, or else it is $\{4,1,16,1,\ldots\}$ if they are both positive; the eigenvalues must have different modulus values.*

[6] Alternatively, we can argue that $R_{1,1} + R_{2,2} = \text{Tr}\{\mathbf{R}\} = \text{Tr}\{\mathbf{\Omega}\mathbf{D}^*\mathbf{\Omega}^{-1}\}$ (by (P.2)) $= \text{Tr}\{\mathbf{D}^*\}$ (by the invariance of the trace function under cyclic permutation) $= \alpha+\beta$.

4.2 $\mathrm{Tr}\{\mathbf{M}\} = 0$

We have a rather different result when $\mathbf{M}$ has zero trace (and $|\lambda| = |\mu|$). The matrix $\mathbf{M}(1,2,3,-1)$, for instance (with $K^2 = 28$; $\lambda, \mu = \mp\sqrt{7}$), returns a count sequence $\{\mathcal{C}_n(1,2,3,-1)\}_{n=2}^{\infty} = \{4,3,0,5,12,7,0,9,20,11,0,13,28,15,0,17,36,19,0,21,44,23,0,25,52,\ldots\}$ with $\mathcal{C}_n = 0$ ($n = 4,8,12,16,\ldots$), $\mathcal{C}_n = 2n$ ($n = 2,6,10,14,\ldots$), and for n (odd) ≥ 3 $\mathcal{C}_n = n$ (see Appendix C for further context for this sequence $\{\mathcal{C}_n(A,B,C,-A)\}_{n=2}^{\infty}$); the same count sequence is delivered for the matrices $\mathbf{M}(-7,8,-6,7)$ ($K^2 = 4$; $\lambda, \mu = \mp 1$), $\mathbf{M}(-8,13,-3,8)$ ($K^2 = 100$; $\lambda, \mu = \mp 5$), $\mathbf{M}(-6,7,-3,6)$ ($K^2 = 60$; $\lambda, \mu = \mp\sqrt{15}$), $\mathbf{M}(3,-5,1,-3)$ ($K^2 = 16$; $\lambda, \mu = \mp 2$), $\mathbf{M}(8,1,6,-8)$ ($K^2 = 280$; $\lambda, \mu = \mp\sqrt{70}$) and $\mathbf{M}(14,-16,-18,-14)$ ($K^2 = 1936$; $\lambda, \mu = \mp 22$).[7]

Lemma 4.1. *The diagonal entries of any nth root matrix of* $\mathbf{M}$ *have the same imaginary part iff those in* $\mathbf{D}^*(\alpha,\beta)$ *do also. Furthermore, this is the same for both pairs of diagonal terms.*

Proof. We base our proof on the form of $R_{1,1} = [\alpha+\beta+(\beta-\alpha)\sigma]/2$ and $R_{2,2} = [\alpha+\beta-(\beta-\alpha)\sigma]/2$, from (P.4), with $\sigma = (A-D)/K$.

Necessity: Assuming $R_{1,1}, R_{2,2}$ have the same imaginary part, then the difference $R_{1,1} - R_{2,2} = (\beta-\alpha)\sigma$ is real. Thus, since σ is real, so must $\beta - \alpha$ be real, and so $\mathrm{Im}\{\alpha\} = \mathrm{Im}\{\beta\}$ (we also trivially deduce the inference $\mathrm{Im}\{R_{2,2}\} = \mathrm{Im}\{R_{1,1}\} = \frac{1}{2}\mathrm{Im}\{\alpha+\beta+(\beta-\alpha)\sigma\} = \frac{1}{2}\mathrm{Im}\{\alpha+\beta\} = \frac{1}{2}[\mathrm{Im}\{\alpha\}+\mathrm{Im}\{\beta\}] = \frac{1}{2}[\mathrm{Im}\{\alpha\}+\mathrm{Im}\{\alpha\}] = \mathrm{Im}\{\alpha\}$ $(= \mathrm{Im}\{\beta\})$).

Sufficiency: Assuming $\mathrm{Im}\{\alpha\} = \mathrm{Im}\{\beta\}$, then $\beta - \alpha$ is real $\Rightarrow$ $(\beta-\alpha)\sigma$ is real, whence $\mathrm{Im}\{R_{1,1}\} = \frac{1}{2}\mathrm{Im}\{\alpha+\beta+(\beta-\alpha)\sigma\} = \frac{1}{2}\mathrm{Im}\{\alpha+\beta\} = \frac{1}{2}\mathrm{Im}\{\alpha+\beta-(\beta-\alpha)\sigma\} = \mathrm{Im}\{R_{2,2}\}$ (again, it is trivial to see that $\mathrm{Im}\{R_{1,1}\} = \mathrm{Im}\{R_{2,2}\} = \frac{1}{2}\mathrm{Im}\{\alpha+\beta\} = \cdots = \mathrm{Im}\{\alpha\} = \mathrm{Im}\{\beta\}$).

This completes the proof. □

We also have a further result through which Lemma 4.1 follows independently. Let

$$\mathbf{L} = \begin{pmatrix} w & x \\ y & z \end{pmatrix} \tag{4.1}$$

$\in \mathcal{M}_2[\mathbb{R}]$ be a real valued 2×2 matrix and, for ϵ a given real, define $\Delta_2(\epsilon) = \{\mathbf{L} + \epsilon i \mathbf{I}_2 : \mathbf{L} \in \mathcal{M}_2[\mathbb{R}]\}$ as the set of all 2×2 matrices whose anti-diagonal terms are real and whose complex diagonal terms have the same imaginary part ϵ ($\mathbf{I}_2$ is the 2-square identity matrix here).

Lemma 4.2. *If* $\mathbf{\Omega} \in \mathcal{M}_2[\mathbb{R}]$ *is invertible, and* $\mathbf{Q} \in \Delta_2(\epsilon)$*, then both* $\mathbf{\Omega Q \Omega}^{-1}$ *and* $\mathbf{\Omega}^{-1}\mathbf{Q\Omega} \in \Delta_2(\epsilon)$.

Proof. Consider first $\mathbf{\Omega Q \Omega}^{-1} = \mathbf{\Omega}(\mathbf{L} + \epsilon i \mathbf{I}_2)\mathbf{\Omega}^{-1}$ (for some $\mathbf{L} \in \mathcal{M}_2[\mathbb{R}]$) $= \cdots = \mathbf{\Omega L \Omega}^{-1} + \epsilon i \mathbf{I}_2 \in \Delta_2(\epsilon)$ since $\mathbf{\Omega L \Omega}^{-1} \in \mathcal{M}_2[\mathbb{R}]$. Likewise, $\mathbf{\Omega}^{-1}\mathbf{Q\Omega} \in \Delta_2(\epsilon)$ follows similarly. □

The application of Lemma 4.2 to establish Lemma 4.1 is given in Appendix D.

In attempting to quantify those root matrices of a general matrix $\mathbf{M}$ that satisfy Theorem 3.1, we note, by Theorem 2.1, that for each successful root matrix $\mathbf{R}_p$ that 'passes' the condition their exists a corresponding nth root matrix $\mathbf{D}_p^*(\alpha,\beta)$ of the eigenvalues matrix $\mathbf{D}$ that also does, and moreover the diagonal entries of the pair of squared matrices $(\mathbf{R}_p)^2$, $(\mathbf{D}_p^*)^2$ each have equal imaginary parts (Theorem B.1 of Appendix B). Lemma 4.1 now establishes a one-to-one correspondence between $\mathbf{R}_p$ and $\mathbf{D}_p^*$. Because $(\mathbf{R}_p)^2 = (\mathbf{\Omega D}_p^*\mathbf{\Omega}^{-1})^2 = \mathbf{\Omega}(\mathbf{D}_p^*)^2\mathbf{\Omega}^{-1}$, Lemma 4.1 also applies to the *square* of any nth root matrix $\mathbf{R}_p$ of $\mathbf{M}$ and the *square* of the corresponding matrix $\mathbf{D}_p^*$, and so the count of successful matrices $\mathbf{R}_p$ is matched by those associated nth root matrices $\mathbf{D}_p^*$ of simply the eigenvalues matrix $\mathbf{D}$ itself; in other words, sets of nth root matrices for which Theorem 3.1 holds have the same cardinality when seeking roots of either $\mathbf{M}(A,B,C,D)$ or its diagonalising matrix of eigenvalues.

[7] In the case $A = D$ then we know (see the end of Section 3.1) that all nth root matrices satisfy Theorem 3.1. Computations on some test matrices of this type confirm that $\mathcal{C}_n = n^2$.

Theorem 4.1. *Given any matrix* $\mathbf{M}(A,B,C,D)$*, with eigenvalues* $\lambda(A,B,C,D), \mu(A,B,C,D)$ *as in (2.3), then for* $n \geq 2$ $\mathcal{C}_n(A,B,C,D) = \mathcal{C}_n(\lambda,0,0,\mu)$ *provided* $A \neq D$*.*

We have tested Theorem 4.1 on all of the explicit matrices listed in this section, and confirm its validity. Computations reveal that the result does not hold when $A = D$, the reasons for which (being quite subtle) are omitted here.

5 Summary

In this paper a condition for an identity associated with the diagonal terms of 2×2 matrix roots has been formulated theoretically, and tested based on an established method to generate such roots; we believe this result is new to the literature. Additionally, it has been shown that a set of n^2 nth roots of a matrix is available from pairwise combinations of all nth roots of its (real) eigenvalues.

We have mentioned [9] as a source of references to previous articles on matrix roots. The study of Tam and Huang [10]—adopting a mainly graph theoretic approach to analyse (non-negative) pth roots of (non-negative) matrices which includes the $p = 2$ square roots case—cites historical works that reveal interest in a topic dating back to the 19th century. Despite this, it is hoped that the results and observations detailed here are viewed as contributing further to the topic in a positive way.[8]

Appendix A: Generalisation of Lemma 2.2

Here we generalise, with proof, Lemma 2.2.

Theorem A.1. *If, for* $n \geq 1$*, the* n*th integer power of an* $s \times s$ *matrix* $\mathbf{N}$ *is a diagonal matrix with distinct entries, then* $\mathbf{N}$ *must itself be diagonal.*

Proof. Suppose $\mathbf{N}^n = \mathbf{D}$ for some diagonal ($s \times s$) matrix $\mathbf{D}$ whose entries are distinct. If $\mathbf{N}$ has eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_s$, say, then those of $\mathbf{D}$ are $\lambda_1^n, \lambda_2^n, \ldots, \lambda_s^n$. Since $\mathbf{D}$ is diagonal its eigenvalues are simply its diagonal entries, which latter being distinct means that $\lambda_1^n, \lambda_2^n, \ldots, \lambda_s^n$ are all distinct and in turn that the set of eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_s$ of $\mathbf{N}$ is a distinct one also (this is a straightforward inference, for consider any arbitrary pair of eigenvalues λ_i, λ_j, where $i \neq j$—writing the difference $\lambda_i^n - \lambda_j^n$ as $(\lambda_i - \lambda_j)(\lambda_i^{n-1} + \lambda_i^{n-2}\lambda_j + \cdots + \lambda_i\lambda_j^{n-2} + \lambda_j^{n-1})$, then by inspection $\lambda_i^n - \lambda_j^n \neq 0 \Rightarrow \lambda_i - \lambda_j \neq 0$). Thus, $\mathbf{N}$ is expressible in diagonalised form as

$$\mathbf{N} = \mathbf{V}\hat{\mathbf{N}}\mathbf{V}^{-1}, \tag{AP.1}$$

where $\hat{\mathbf{N}}$ is the diagonal matrix containing the (distinct) eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_s$ of $\mathbf{N}$, and $\mathbf{V}$ has its associated eigenvectors as columns. Moreover, since $\mathbf{D} = \mathbf{N}^n$ it is well known (by standard theory) that $\mathbf{N}$ and $\mathbf{D}$ share the same set of eigenvectors, so that the matrix $\mathbf{V}$ which diagonalises $\mathbf{N}$ must also diagonalise $\mathbf{D}$. However, since $\mathbf{D}$ is diagonal itself, $\mathbf{V}$ must be the ($s \times s$) identity matrix, whence (AP.1) reads $\mathbf{N} = \hat{\mathbf{N}}$, and so $\mathbf{N}$ is diagonal. □

Appendix B: The Square of an nth Root Matrix

Theorem B.1. *For* $K^2 > 0$*, let* α, β *be (resp.) any* n*th roots of the eigenvalues* λ, μ *of a matrix* $\mathbf{M}(A,B,C,D)$ $(A \neq D)$*, and suppose* $\mathbf{R} = \mathbf{R}(\alpha,\beta;A,B,C,D)$ *is an* n*th root matrix of* $\mathbf{M}$ *with* $\mathbf{S} = \mathbf{R}^2$*. If Theorem 3.1 holds (i.e.,* $\mathrm{Im}\{\alpha^2\} = \mathrm{Im}\{\beta^2\}$*), then*

$$\mathrm{Im}\{\alpha^2\} = \mathrm{Im}\{\beta^2\} = \mathrm{Im}\{S_{1,1}\} = \mathrm{Im}\{S_{2,2}\}.$$

[8] The author Lee Rawlin was involved in this work as part of his undergraduate final year dissertation project at the University of Derby (supervised by P.J.L.) during the 2020-21 academic year. The motivation for thinking about properties of matrix roots arose from some extensive computational experiments conducted by James Stanton who first raised the possibilty of the result given here as Theorem 3.1.

Proof. Consider, using (P.1)-(P.3),

$$\begin{aligned} \mathbf{S} &= \mathbf{R}^2 \\ &= \mathbf{\Omega}(\mathbf{D}^*)^2\mathbf{\Omega}^{-1} \\ &= \frac{1}{|\mathbf{\Omega}|}\begin{pmatrix} ps\alpha^2 - qr\beta^2 & pq(\beta^2-\alpha^2) \\ rs(\alpha^2-\beta^2) & ps\beta^2 - qr\alpha^2 \end{pmatrix} \\ &= \begin{pmatrix} S_{1,1} & S_{1,2} \\ S_{2,1} & S_{2,2} \end{pmatrix}, \end{aligned} \tag{BP.1}$$

say. If Theorem 3.1 holds then α^2 and β^2 have a common imaginary part. Denoting this by c', and their (resp.) real parts as α'_r, β'_r, we write

$$\alpha^2 = \alpha'_r + ic' \qquad \text{and} \qquad \beta^2 = \beta'_r + ic', \tag{BP.2}$$

whereupon

$$\begin{aligned} |\mathbf{\Omega}|S_{1,1} &= ps\alpha^2 - qr\beta^2 \\ &= ps(\alpha'_r + ic') - qr(\beta'_r + ic') \\ &= ps\alpha'_r - qr\beta'_r + (ps-qr)ic' \\ &= ps\alpha'_r - qr\beta'_r + |\mathbf{\Omega}|ic', \end{aligned} \tag{BP.3}$$

which is to say,

$$S_{1,1} = (ps\alpha'_r - qr\beta'_r)/|\mathbf{\Omega}| + ic'. \tag{BP.4}$$

In a similar fashion, we find that

$$S_{2,2} = (ps\beta^2 - qr\alpha^2)/|\mathbf{\Omega}| = \cdots = (ps\beta'_r - qr\alpha'_r)/|\mathbf{\Omega}| + ic', \tag{BP.5}$$

and the result is established. □

Remark B.1. It is immediate—from (BP.1) or consideration of Tr$\{\mathbf{S}\}$—that $S_{1,1}+S_{2,2} = \alpha^2+\beta^2$ for every matrix $\mathbf{S}$ which is the square of a root matrix $\mathbf{R}$ of $\mathbf{M}$ (1.1). Extending this observation, then $\mathbf{T} = \mathbf{R}^n$ is such that $T_{1,1} + T_{2,2} = \alpha^n + \beta^n = \lambda + \mu$, but since $\mathbf{T}$ co-incides with $\mathbf{M}$ we merely recover the relation $\lambda + \mu = M_{1,1} + M_{2,2} = A + D = \text{Tr}\{\mathbf{M}\}$ offered by (2.3).

Theorem B.1, as a consequence of Theorem 3.1, is no great surprise, being a slight contextual enhancement of the observation that for any complex 2×2 matrix $\mathbf{N}$, say, and its square $\mathbf{P} = \mathbf{N}^2$, the following statements are equivalent (showing that Statement A ⇔ Statement B is a simple algebraic exercise which is left for the reader to confirm):

Statement A : Re$\{N_{1,1}\}$Im$\{N_{1,1}\}$ = Re$\{N_{2,2}\}$Im$\{N_{2,2}\}$;
Statement B : Im$\{P_{1,1}\}$ = Im$\{P_{2,2}\}$.

Theorem B.2. *For* $K^2 > 0$*, let* α, β *be (resp.) any* n*th roots of the eigenvalues* λ, μ *of a matrix* $\mathbf{M}(A,B,C,D)$ *(*$A \neq D$*), and suppose* $\mathbf{R} = \mathbf{R}(\alpha,\beta;A,B,C,D)$ *is an* n*th root matrix of* $\mathbf{M}$ *with* $\mathbf{S} = \mathbf{R}^2$*. If Theorem 3.1 holds (i.e.,* Im$\{\alpha^2\}$ = Im$\{\beta^2\}$*), then*

$$\text{Re}\{S_{1,1}\}\text{Im}\{S_{1,1}\} = \text{Re}\{S_{2,2}\}\text{Im}\{S_{2,2}\}$$

if either $\alpha^2, \beta^2 \in \mathbb{R}$ *or* $\alpha^2 = \beta^2$.

Proof. Using (BP.4),(BP.5) (or else deploying the line of argument described in Remark 3.3), we see that Re$\{S_{1,1}\}$Im$\{S_{1,1}\}$ − Re$\{S_{2,2}\}$Im$\{S_{2,2}\}$ $= c'(ps+qr)(\alpha'_r - \beta'_r)/|\mathbf{\Omega}|$ which, using (2.3), reduces to $-\sigma c'(\alpha'_r - \beta'_r)$ (with σ (as previously defined) $\neq 0$) and is zero if either $c' = 0$ or $\alpha'_r = \beta'_r$. In the case when $c' = 0$ (in other words, if $\alpha^2, \beta^2 \in \mathbb{R}$) then Theorem B.2 holds trivially since Im$\{\alpha^2\}$ = Im$\{\beta^2\}$ = 0 (a special case of Theorem 3.1) and so Im$\{S_{1,1}\}$ = Im$\{S_{2,2}\}$ = 0 by Theorem B.1. The case $\alpha'_r = \beta'_r$ arises if $\alpha^2 = \beta^2$, whence $S_{1,1} = S_{2,2}$ by (BP.4),(BP.5) and again Theorem B.2 holds automatically. □

Remark B.2. As an aside, a necessary (but not sufficient) condition for equality of α^2 and β^2 is that the matrix $\mathbf{M}$ has zero trace, for $\alpha^2 = \beta^2 \Rightarrow (\alpha^2)^n = (\beta^2)^n \Rightarrow (\alpha^n)^2 = (\beta^n)^2 \Rightarrow \lambda^2 = \mu^2 \Rightarrow \lambda = -\mu$ (the paper is based on the assumption that $\lambda \neq \mu$) iff $\mathrm{Tr}\{\mathbf{M}\} = 0$ (in which case $\lambda = -K/2 = -\mu$).[9] Note also that $\alpha^2 \neq \beta^2$ for any even n, for this would otherwise give (for some even $n = 2m$ (integer $m \geq 1$)) $\lambda = \alpha^n = \alpha^{2m} = (\alpha^2)^m = (\beta^2)^m = \beta^{2m} = \beta^n = \mu$; we see a particular instance of this in the $n = 2$ square roots case where (referring to Corollary 3.1) $\alpha^2 = \lambda \neq \mu = \beta^2$.

Remark B.3. Evidently, when Theorem 3.1 holds for a root matrix $\mathbf{R}$ of $\mathbf{M}$, it follows that $\mathbf{S} = \mathbf{R}^2 \in \Delta_2(c')$ with an explicit compositional matrix $\mathbf{L}$ (4.1) for which

$$\begin{aligned} w &= (ps\alpha_r' - qr\beta_r')/|\mathbf{\Omega}|, \\ z &= (ps\beta_r' - qr\alpha_r')/|\mathbf{\Omega}|, \end{aligned} \tag{B.1}$$

directly from (BP.4),(BP.5), and

$$\begin{aligned} x &= pq(\beta_r' - \alpha_r')/|\mathbf{\Omega}|, \\ y &= rs(\alpha_r' - \beta_r')/|\mathbf{\Omega}|, \end{aligned} \tag{B.2}$$

given readily by (BP.1),(BP.2), where $\mathbf{\Omega} \in \mathcal{M}_2[\mathbb{R}]$ is the (invertible) matrix of (2.2). Thus, $\mathbf{\Omega S \Omega}^{-1} \in \Delta_2(c')$ by Lemma 4.2.

Appendix C

The counting sequence $\{\mathcal{C}_n(A, B, C, -A)\}_{n=2}^{\infty} = \{4, 3, 0, 5, 12, 7, 0, 9, \ldots\}$ associated with a zero trace matrix $\mathbf{M}$ (Section 4.2) is contained within Sequence No. A251610 (that is, $\{1, 4, 3, 0, 5, 12, 7, 0, 9, \ldots\}$) on the well known On-Line Encyclopaedia of Integer Sequences;[10] we see that the latter sequence is listed as having an enumerative interpretation in relation to determinants of spiral knots—it (i) has (ordinary) generating function

$$\frac{x^4 + 2x^3 - 2x^2 + 2x + 1}{(x-1)^2(x^2+1)^2} = 1 + 4x + 3x^2 + 5x^4 + 12x^5 + 7x^6 + 9x^8 + \cdots, \tag{C.1}$$

and (ii) is generated (a) as $\{k(s_k)^2\}_{k=1}^{\infty}$ through an intermediate sequence $\{s_k\}_{k=1}^{\infty}$ delivered by the linear recursion $s_k = \sqrt{2}s_{k-1} - s_{k-2}$ ($k \geq 3$; $s_1 = 1$, $s_2 = \sqrt{2}$),[11] or (b) directly, via a closed form version for its general term, as

$$\{k[2 - (-i)^k - i^k]/2\}_{k=1}^{\infty}, \tag{C.2}$$

or equivalently,

$$\{k[1 - \cos(k\pi/2)]\}_{k=1}^{\infty}. \tag{C.3}$$

Appendix D: Alternative Proof of Lemma 4.1

Here we establish Lemma 4.1 by application of Lemma 4.2.

Proof. We again refer to the elements (P.4) of any nth root matrix $\mathbf{R}$ of a matrix $\mathbf{M}$.

Necessity: Assuming $R_{1,1}, R_{2,2}$ have the same imaginary part, then (by the previous proof of Lemma 4.1) $\beta - \alpha$ is real. Writing $\alpha = \alpha_r + i\alpha_c$ and $\beta = \beta_r + i\beta_c$, it follows that $\alpha_c = \beta_c$ and in turn that $\mathrm{Im}\{R_{1,1}\} = \mathrm{Im}\{R_{2,2}\} = \alpha_c$. Thus, with $R_{1,2}$ and $R_{2,1}$ real by inspection, $\mathbf{R} \in \Delta_2(\alpha_c)$ and so (using (P.2)) $\mathbf{D}^*(\alpha, \beta) = \mathbf{\Omega}^{-1}\mathbf{R}\mathbf{\Omega} \in \Delta_2(\alpha_c)$ by Lemma 4.2.

Sufficiency: We assume $\mathrm{Im}\{\alpha\} = \mathrm{Im}\{\beta\}$ (so that $\alpha_c = \beta_c$). Then $\mathbf{D}^*(\alpha, \beta) \in \Delta_2(\alpha_c)$ trivially (its anti-diagonal elements are real, being zero), whence (using (P.2)) $\mathbf{R} = \mathbf{\Omega D}^*\mathbf{\Omega}^{-1} \in \Delta_2(\alpha_c)$ by Lemma 4.2. □

[9] The condition is not a sufficient one, for $\mathrm{Tr}\{\mathbf{M}\} = 0 \Rightarrow \lambda = -\mu \Rightarrow \lambda^2 = \mu^2 \Rightarrow \cdots \Rightarrow (\alpha^2)^n = (\beta^2)^n$, but it is not possible to infer from this that $\alpha^2 = \beta^2$.

[10] See https://oeis.org/.

[11] Which yields $\{s_k\}_{k=1}^{\infty}$ as the period 8 sequence $\{1, \sqrt{2}, 1, 0, -1, -\sqrt{2}, -1, 0, \ldots\}$.

References

[1] A. Choudhry, Extraction of nth roots of 2×2 matrices, *Lin. Alg. Appl.* **387**, 183–192 (2004).

[2] P. J. Larcombe, A note on the invariance of the general 2×2 matrix anti-diagonals ratio with increasing matrix power: four proofs, *Fib. Quart.* **53**, 360–364 (2015).

[3] P. J. Larcombe and E. J. Fennessey, A new tri-diagonal matrix invariance property, *Palest. J. Math.* **7**, 9–13 (2018).

[4] P. J. Larcombe and E. J. Fennessey, A note on two rational invariants for a particular 2×2 matrix, *Palest. J. Math.* **7**, 410–413 (2018).

[5] P. J. Larcombe and E. J. Fennessey, On anti-diagonals ratio invariance with exponentiation of a 2×2 matrix: two new proofs, *Palest. J. Math.* **9**, 12–14 (2020).

[6] P. J. Larcombe and E. J. Fennessey, A formalised inductive approach to establish the invariance of anti-diagonal ratios with exponentiation for a tri-diagonal matrix of fixed dimension, *Palest. J. Math.* **9**, 670–672 (2020).

[7] P. J. Larcombe and E. J. Fennessey, A short graph-theoretic proof of the 2×2 matrix anti-diagonals ratio invariance with exponentiation, *Palest. J. Math.* **10**, 102–103 (2021).

[8] J. McLaughlin, Combinatorial identities deriving from the n-th power of a 2×2 matrix, *Int.: Elec. J. Comb. Num. Theor.* **4**, Art. No. A19, 15 pp. (2004).

[9] M. Özdemir, Finding n-th roots of a 2×2 real matrix using de Moivre's formula, *Adv. Appl. Cliff. Alg.* **29** (Issue 1), 25 pp. (2019).

[10] B.-S. Tam and P.-R. Huang, Nonnegative square roots of matrices, *Lin. Alg. Appl.* **498**, 404-440 (2016).

Author information

Peter J. Larcombe, Department of Computing and Mathematics, College of Engineering and Technology, University of Derby, Kedleston Road, Derby DE22 1GB, U.K.
E-mail: `p.j.larcombe@derby.ac.uk`

Eric J. Fennessey, BAE Systems Surface Ships Ltd., HM Naval Base, Portsmouth PO1 3NJ, U.K.
E-mail: `eric.fennessey@baesystems.com`

Lee Rawlin, 4, Scarbrough Avenue, Skegness PE25 2SY, U.K.
E-mail: `leerawlin@btinternet.com`

James Stanton, 3, Brackley Drive, Spondon, Derby DE21 7SA, U.K.
E-mail: `mrstanton@hotmail.co.uk`

Received: July 15, 2020.
Accepted: October 23, 2020