Operators
and
Matrices
Volume 1, Number 1 (2007), 121-141

ON THE CONVERGENCE OF ALUTHGE SEQUENCE

HUAJUN HUANG AND TIN-YAU TAM

(*communicated by Leiba Rodman*)

Abstract. For $0 < \lambda < 1$, the λ-Aluthge sequence $\{\Delta_\lambda^m(X)\}_{m\in\mathbb{N}}$ converges if the nonzero eigenvalues of $X \in \mathbb{C}_{n\times n}$ have distinct moduli, where $\Delta_\lambda(X) := P^\lambda UP^{1-\lambda}$ if $X = UP$ is a polar decomposition of X.

1. Introduction

Given $X \in \mathbb{C}_{n\times n}$, the polar decomposition [9] asserts that $X = UP$, where U is unitary and P is positive semidefinite, and the decomposition is unique if X is nonsingular. Though the polar decomposition may not be unique, the Althuge transform [1] of X:

$$\Delta(X) := P^{1/2}UP^{1/2}$$

($P^{1/2}XP^{-1/2}$ if X is nonsingular) is well defined [17, Lemma 2]. Aluthge transform has been studied extensively, for example, [1, 2, 3, 4, 5, 7, 8, 11, 12, 13, 14, 16, 17]. Recently Yamazaki [16] established the following interesting result

$$\lim_{m\to\infty} \|\Delta^m(X)\| = r(X), \tag{1.1}$$

where $r(X)$ is the spectral radius of X and

$$\|X\| := \max_{\|v\|_2=1} \|Xv\|_2$$

is the spectral norm of X. Suppose that the singular values $s_1(X),\ldots,s_n(X)$ and the eigenvalues $\lambda_1(X),\ldots,\lambda_n(X)$ of X are arranged in nonincreasing order

$$s_1(X) \geqslant s_2(X) \geqslant \cdots \geqslant s_n(X), \qquad |\lambda_1(X)| \geqslant |\lambda_2(X)| \geqslant \cdots \geqslant |\lambda_n(X)|.$$

Since $\|X\| = s_1(X)$ and $r(X) := |\lambda_1(X)|$, the following result of Ando [3] is an extension of (1.1).

THEOREM 1.1. *(Yamazaki-Ando) Let $X \in \mathbb{C}_{n\times n}$. Then*

$$\lim_{m\to\infty} s_i(\Delta^m(X)) = |\lambda_i(X)|, \quad i = 1,\ldots,n. \tag{1.2}$$

Mathematics subject classification (2000): 15A23, 15A45.
Key words and phrases: λ-Aluthge transform, polar decomposition, normal matrix.

Paper No. 01-08

Aluthge transform $\Delta(T)$ is also defined for Hilbert space bounded linear operator T [17] and (1.1) remains true [16]. Yamazaki's result (1.1) provides support for the following conjecture of Jung et al [11, Conjecture 1.11] for any $T \in B(H)$ where $B(H)$ denotes the algebra of bounded linear operators on the Hilbert space H.

CONJECTURE 1.2. *Let $T \in B(H)$. The Aluthge sequence $\{\Delta^m(T)\}_{m\in\mathbb{N}}$ is norm convergent to a quasinormal $Q \in B(H)$, that is, $\|\Delta^m(T) - Q\| \to 0$ as $m \to \infty$, where $\|\cdot\|$ is the spectral norm.*

It is known [11, Propositioin 1.10] that if the Aluthge sequence of $T \in B(H)$ converges, its limit L is quasinormal, that is, L commutes with L^*L, or equivalently, $UP = PU$ where $L = UP$ is a polar decomposition of L [9]. However very recently it is known [7] that Conjecture 1.2 is not true for infinite dimensional Hilbert space. Cho̅, Jung and Lee [7, Corollary 3.3] constructed a unilateral weighted shift operator $T : \ell_2(\mathbb{N}) \to \ell_2(\mathbb{N})$ such that the sequence $\{\Delta^m(T)\}_{m\in\mathbb{N}}$ does not converge in weak operator topology. They also constructed [7, Example 3.5] a hyponormal bilateral weighted shift $B : \ell_2(\mathbb{Z}) \to \ell_2(\mathbb{Z})$ such that $\{\Delta^m(B)\}_{m\in\mathbb{N}}$ converges in the strong operator topology, that is, for some $L : \ell_2(\mathbb{Z}) \to \ell_2(\mathbb{Z})$, $\|\Delta^m(B)x - Lx\| \to 0$ as $m \to \infty$ for all $x \in \ell_2(\mathbb{Z})$, where $\|x\|$ is the norm induced by the inner product. However $\{\Delta^m(B)\}_{m\in\mathbb{N}}$ does not converge in the norm topology. So the study of Conjecture 1.2 is reduced to the finite dimensional case $\mathbb{C}_{n\times n}$. Since the three (weak, strong, norm) topologies coincide and quasinormal and normal coincide [9] in the finite dimensional case, the limit points of the Aluthge sequence are normal [13, Proposition 3.1], [3, Theorem 1]. Also see [11, Proposition 1.14]. Moreover the eigenvalues of $\Delta(X)$ and the eigenvalues of X are identical, counting multiplicities. So the study of Conjecture 1.2 is now reduced to the finite dimensional case:

CONJECTURE 1.3. *Let $X \in \mathbb{C}_{n\times n}$. The Aluthge sequence $\{\Delta^m(X)\}_{m\in\mathbb{N}}$ is convergent to a normal matrix whose eigenvalues are $\lambda_1(X), \ldots, \lambda_n(X)$.*

Conjecture 1.3 is true when $n = 2$ [4, p.300] and the proof involves very hard computation which seems unlikely to be extended in higher dimension. It remains open for $3 \leqslant n$. It is also true for some special cases [3] [13, Corollary 3.3], for examples, (1) if the spectrum of X is a singleton set, or (2) if X is normal (then $\Delta^m(X) = X$ for all m).

In this paper we give a partial answer to Conjecture 1.3, that is, it is true if the nonzero eigenvalues of $X \in \mathbb{C}_{n\times n}$ have distinct moduli. Such matrices form a dense set in $\mathbb{C}_{n\times n}$. Indeed our result is also true for λ-Aluthge transform that we are about to mention.

From now on we only consider $X \in \mathbb{C}_{n\times n}$, the finite dimensional case.

Let $X = UP$ be a polar decomposition of $X \in \mathbb{C}_{n\times n}$ where U is unitary and P is positive semidefinite. For $0 < \lambda < 1$, Aluthge [2] introduced a generalized Aluthge transform (see [5, 11, 14]) and we call it the λ-Aluthge transform:

$$\Delta_\lambda(X) := P^\lambda U P^{1-\lambda}$$

which is also well defined. Evidently the Aluthge transform Δ is simply $\Delta_{\frac{1}{2}}$. Since $P = (X^*X)^{1/2}$, one may write

$$\Delta_\lambda(X) = (X^*X)^{\lambda/2}U(X^*X)^{(1-\lambda)/2}.$$

In addition, if X is nonsingular, then $\Delta_\lambda(X) = P^\lambda XP^{-\lambda}$ and thus similar to X. The spectrum, counting multiplicities, is invariant under Δ_λ, denoted by

$$\sigma(X) \stackrel{m}{=} \sigma(\Delta_\lambda(X)) \tag{1.3}$$

since $\sigma(XY) \stackrel{m}{=} \sigma(YX)$, where $\sigma(X)$ denotes the spectrum of X. Moreover Δ_λ respects unitary similarity:

$$\Delta_\lambda(VXV^{-1}) = V\Delta_\lambda(X)V^{-1}, \quad V \in U(n). \tag{1.4}$$

The sequence $\{\Delta_\lambda^m(X)\}_{m\in\mathbb{N}}$ is called the λ*-Aluthge sequence* of X. By the submultiplicativity of the spectral norm, it follows immediately that

$$\|\Delta_\lambda(X)\| \leqslant \|X\| \tag{1.5}$$

and thus $\{\|\Delta_\lambda^m(X)\|\}_{m\in\mathbb{N}}$ is nonincreasing. In [5, Corollary 4.2] Antezana, Massey and Stojanoff generalized Theorem 1.1: for any $X \in \mathbb{C}_{n\times n}$,

$$\lim_{m\to\infty} \|\Delta_\lambda^m(X)\| = r(X), \tag{1.6}$$

and obtained many other nice results. However (1.6) remains unknown for Hilbert space operators T.

THEOREM 1.4. *[5] Let $X \in \mathbb{C}_{n\times n}$ and $0 < \lambda < 1$.*

1. *Any limit point of the λ-Aluthge sequence $\{\Delta_\lambda^m(X)\}_{m\in\mathbb{N}}$ is normal, with eigenvalues $\lambda_1(X),\ldots,\lambda_n(X)$.*
2. $\lim_{m\to\infty} s_i(\Delta_\lambda^m(X)) = |\lambda_i(X)|$, $i = 1,\ldots,n$.
3. *If $X \in \mathbb{C}_{2\times 2}$, then $\{\Delta_\lambda^m(X)\}_{m\in\mathbb{N}}$ converges.*

Theorem 1.4(1) is [5, Proposition 4.1]. It reduces to [3, Theorem 1] and [13, Proposition 3.1] when $\lambda = 1/2$. Theorem 1.4(3) is [5, Theorem 4.6] and is an extension of [4]. Theorem 1.4(2) can be deduced from (1.6) using compound matrices via the argument in Ando [3, p.284-285].

It is evident from Theorem 1.4(1) that if the spectrum of X is a singleton set $\{\alpha\}$, then the λ-Aluthge sequence converges to αI_n.

The main goal of the paper is to show that if the nonzero eigenvalues of $X \in \mathbb{C}_{n\times n}$ have distinct moduli, then the λ-Aluthge sequence converges. Since such matrices X form a dense subset in $\mathbb{C}_{n\times n}$, it explains why many numerical experiments result in convergence. An example is given to show that the λ-Aluthge sequence does not converge when $\lambda = 1$.

2. Distinct moduli implies convergence

We list the following notations that we will use in the forthcoming discussion.

$\mathbb{C}_{n\times n}$ = the set of all $n\times n$ complex matrices
$\mathrm{GL}_n(\mathbb{C})$ = the general linear group of $n\times n$ nonsingular matrices
$S(n)$ = the Lie algebra of $n\times n$ skew Hermitian matrices
$H(n)$ = the real vector space of $n\times n$ Hermitian matrices
$P(n)$ = the set of $n\times n$ positive definite matrices
$U(n)$ = the group of $n\times n$ unitary matrices
$D(n)$ = the group of $n\times n$ diagonal unitary matrices
$\mathcal{D}_+(n)$ = the set of all positive diagonal matrices with diagonal entries in descending order
$\|X\|_F$ = $\sqrt{\mathrm{tr}\,(X^*X)}$, the Frobenius norm of $X\in\mathbb{C}_{n\times n}$
$\|X\|$ = $s_1(X)$, the spectral norm of $X\in\mathbb{C}_{n\times n}$
$\mathbb{N}$ = $\{1,2,\ldots,\}$, the set of natural numbers

The entire paper is to prove the following two results.

THEOREM 2.1. *Let $0<\lambda<1$. If the nonzero eigenvalues of $X\in\mathbb{C}_{n\times n}$ have distinct moduli, then the λ-Aluthge sequence $\{\Delta_\lambda^m(X)\}_{m\in\mathbb{N}}$ converges to a normal matrix with the same eigenvalues (counting multiplicity) as X.*

THEOREM 2.2. *Let $X=U^*(\oplus_{i=1}^k T_i)U$, where $U\in U(n)$ and for each $i=1,\ldots,k$, either*
1. *the nonzero eigenvalues of T_i are the same,*
2. *the nonzero eigenvalues of T_i have distinct moduli,*
3. *T_i has two nonzero eigenvalues, or*
4. *$\Delta_\lambda^q(T_i)$ is normal for some $q\in\mathbb{N}$.*

Then the λ-Aluthge sequence $\{\Delta_\lambda^m(X)\}_{m\in\mathbb{N}}$ converges.

Theorem 2.2 combines Theorem 2.1 and some known convergence results for $n\times n$ matrices in the literature.

EXAMPLE 2.3. Suppose that $0<\lambda<1$.

1. Let
$$X=\begin{bmatrix} a & * & * \\ 0 & b & * \\ 0 & 0 & c\end{bmatrix}\oplus A,$$
where $|a|,|b|,|c|$ are distinct and matrix A has a singleton spectrum. The λ-Althuge sequence $\{\Delta_\lambda^m(X)\}_{m\in\mathbb{N}}$ converges.

2. It is possible that X is not normal but $\Delta_\lambda^q(X)$ is normal for some $q\in\mathbb{N}$. Let
$$X=\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0\end{bmatrix}.$$

Then X is not normal and X is similar to $I_2 \oplus [\ 0\ ]$. By the proof of [5, Corollary 4.16],

$$\Delta_\lambda(X) = U^* \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix} U$$

for some $U \in U(3)$ and $S \in GL_2(\mathbb{C})$. By [5, Proposition 4.14(2)], S has only one eigenvalue 1 with trivial Jordan structure. So $S = I_2$ and $\Delta_\lambda(X)$ is normal. Therefore, $\Delta_\lambda(X) = \Delta_\lambda^2(X) = \cdots$, and $\{\Delta_\lambda^m(X)\}_{m\in\mathbb{N}}$ converges to $\Delta_\lambda(X)$.

The idea of proving Theorem 2.1 is to show that $\{\Delta_\lambda^m(X)\}_{m\in\mathbb{N}}$ is a Cauchy sequence via the Frobenius norm. As a finite dimensional normed space, $\mathbb{C}_{n\times n}$ is complete and thus $\{\Delta_\lambda^m(X)\}_{m\in\mathbb{N}}$ converges. The proof does not reveal the explicit form of the limit.

We will establish a few lemmas in order to prove Theorem 2.1. The following lemma can be obtained from [11, Proposition 1.10] and the remark on [11, p.445] since normal and quasinormal coincide in $\mathbb{C}_{n\times n}$.

LEMMA 2.4. *Let $0 < \lambda < 1$ and $X \in \mathbb{C}_{n\times n}$. Then X is normal if and only if $\Delta_\lambda(X) = X$.*

Given a normal matrix $A \in \mathrm{GL}_n(\mathbb{C})$, we may write the spectral decomposition of A in the following fashion

$$A = V^* D_\theta D V,$$

where $V \in U(n)$, $D_\theta \in D(n)$, and $D \in \mathcal{D}_+(n)$. Indeed,

$$D = \mathrm{diag}\,(|\lambda_1(A)|, \ldots, |\lambda_n(A)|), \quad D_\theta = \mathrm{diag}\,(e^{i\theta_1}, \ldots, e^{i\theta_n})$$

such that $\lambda_j(A) = e^{i\theta_j}|\lambda_j(A)|$, $j = 1, \ldots, n$.

The following lemma provides a representation of a sequence in $\mathrm{GL}_n(\mathbb{C})$ which converges to a normal matrix $A \in \mathrm{GL}_n(\mathbb{C})$ whose eigenvalues are the same if they have the same moduli. We will only use a special case of the lemma in the proof of our main theorem, namely, when A has distinct eigenvalue moduli.

LEMMA 2.5. *Let $\{X_m\}_{m\in\mathbb{N}} \subset \mathrm{GL}_n(\mathbb{C})$ be a sequence which converges to a normal matrix $A \in \mathrm{GL}_n(\mathbb{C})$. Write*

$$A = V^* D_\theta D V,$$

where $V \in U(n)$, $D_\theta \in D(n)$, and $D \in \mathcal{D}_+(n)$. Suppose that eigenvalues of A are identical if they have the same moduli. Then for each $m \in \mathbb{N}$, there are $V_m \in U(n)$, $B_m \in S(n)$, $D_m \in \mathcal{D}_+(n)$ such that

$$X_m = V_m^* e^{B_m} D_\theta D_m V_m, \tag{2.1}$$

satisfying

1. $\lim_{m\to\infty} D_m = D$.
2. $\lim_{m\to\infty} B_m = \mathbf{0}$.

Proof. Since $\lim_{m\to\infty} X_m = A$, we have

$$\lim_{m\to\infty} D_\theta{}^{-1} V X_m V^* = D. \tag{2.2}$$

Let

$$D_\theta{}^{-1}VX_mV^* = U_mD_mL_m \tag{2.3}$$

be a singular value decomposition of $D_\theta{}^{-1}VX_mV^*$, where $U_m, L_m \in U(n)$ and $D_m \in \mathcal{D}_+(n)$ (U_m and L_m are not unique in general). Since $D_m \in \mathcal{D}_+(n)$ contains the singular values of X_m, by the continuity of singular values

$$\lim_{m\to\infty} D_m = D. \tag{2.4}$$

Rewrite (2.3) in the fashion of polar decomposition

$$D_\theta{}^{-1}VX_mV^* = (U_mL_m)(L_m^*D_mL_m) \in \mathrm{GL}_n(\mathbb{C}) \tag{2.5}$$

where $U_mL_m \in U(n)$, $L_m^*D_mL_m \in P(n)$. The polar decomposition

$$\pi : U(n) \times H(n) \to \mathrm{GL}_n(\mathbb{C}), \qquad \pi(U,H) = U\exp H. \tag{2.6}$$

is a diffeomorphism [15, p.238]. Due to (2.2) and (2.5)

$$\lim_{m\to\infty} U_mL_m = I_n, \tag{2.7}$$

and

$$\lim_{m\to\infty} L_m^*(\log D_m)L_m = \log D$$

so that

$$\lim_{m\to\infty} L_m^*D_mL_m = D. \tag{2.8}$$

By (2.4),

$$\lim_{m\to\infty} \|L_m^*(D-D_m)L_m\| = \lim_{m\to\infty} \|D-D_m\| = 0. \tag{2.9}$$

By (2.8) and (2.9),

$$\lim_{m\to\infty} L_m^*DL_m = \lim_{m\to\infty} L_m^*D_mL_m + \lim_{m\to\infty} L_m^*(D-D_m)L_m = D. \tag{2.10}$$

Therefore,

$$\begin{aligned} \lim_{m\to\infty} \|D - L_mDL_m^*\| &= \lim_{m\to\infty} \|L_m^*(D-L_mDL_m^*)L_m\| \\ &= \lim_{m\to\infty} \|L_m^*DL_m - D\| = 0 \qquad \text{by (2.10)}. \end{aligned} \tag{2.11}$$

This shows that

$$D = \lim_{m\to\infty} L_mDL_m^*. \tag{2.12}$$

Write

$$D = \mathrm{diag}\,(|\lambda_1(A)|, \ldots, |\lambda_n(A)|), \quad D_\theta = \mathrm{diag}\,(e^{i\theta_1}, \ldots, e^{i\theta_n})$$

such that $\lambda_j(A) = e^{i\theta_j}|\lambda_j(A)|$, $j = 1, \ldots, n$. Recall that eigenvalues of A are identical if they have the same moduli, that is, $|\lambda_k(A)| = |\lambda_j(A)|$ implies $e^{i\theta_k} = e^{i\theta_j}$. By Lagrange interpolation theorem, it amounts to say that $D_\theta = p(D)$ for some polynomial $p(x) \in \mathbb{C}[x]$. By (2.12),

$$\begin{aligned} \lim_{m\to\infty} L_mD_\theta L_m^* &= \lim_{m\to\infty} L_mp(D)L_m^* \\ &= \lim_{m\to\infty} p(L_mDL_m^*) = p(D) = D_\theta. \end{aligned} \tag{2.13}$$

Now

$$\begin{aligned} X_m &= V^* D_\theta U_m D_m L_m V \qquad \text{by (2.3)} \\ &= V^* L_m^* [(L_m D_\theta L_m^*)(L_m U_m) D_\theta^{-1}] D_\theta D_m L_m V. \end{aligned} \tag{2.14}$$

Denote $C_m := (L_m D_\theta L_m^*)(L_m U_m) D_\theta^{-1}$. By (2.7),

$$\begin{aligned} \lim_{m\to\infty} \|L_m U_m - I_n\| &= \lim_{m\to\infty} \|L_m^*(L_m U_m - I_n) L_m\| \\ &= \lim_{m\to\infty} \|U_m L_m - I_n\| = 0. \end{aligned} \tag{2.15}$$

So $\lim_{m\to\infty} L_m U_m = I_n$ and thus with (2.13),

$$\lim_{m\to\infty} C_m = \left(\lim_{m\to\infty} L_m D_\theta L_m^*\right)\left(\lim_{m\to\infty} L_m U_m\right) D_\theta^{-1} = I_n. \tag{2.16}$$

Notice that $C_m \in U(n)$. The exponential map $\exp : \mathbb{C}_{n\times n} \to \mathrm{GL}_n(\mathbb{C})$ [10, p.149] is onto and satisfies

$$U(n) = \exp S(n). \tag{2.17}$$

Though the exponential map $\exp : S(n) \to U(n)$ is not bijective, it gives a diffeomorphism [10, p.104]

$$\varphi : N_0 \to N_1$$

between a neighborhood N_0 of $\mathbf{0} \in S(n)$ and a neighborhood N_1 of $I_n \in U(n)$. Due to (2.17), (2.16) and the diffeomorphism φ, for each $m \in \mathbb{N}$, there exists $B_m \in S(n)$ such that

$$C_m = e^{B_m} \qquad \text{and} \qquad \lim_{m\to\infty} B_m = \mathbf{0}. \tag{2.18}$$

By (2.14),

$$X_m = V_m^* e^{B_m} D_\theta D_m V_m,$$

where $V_m := L_m V \in U(n)$, as desired. □

We now use Lemma 2.5 to establish the following lemma.

LEMMA 2.6. *Suppose that the eigenvalues of* $X \in \mathrm{GL}_n(\mathbb{C})$ *have distinct eigenvalue moduli*

$$|\lambda_1(X)| > |\lambda_2(X)| > \cdots > |\lambda_n(X)| > 0.$$

Denote

$$\begin{aligned} D_\theta &:= diag\left(\frac{\lambda_1(X)}{|\lambda_1(X)|}, \cdots, \frac{\lambda_n(X)}{|\lambda_n(X)|}\right) \\ D &:= diag\,(|\lambda_1(X)|, \cdots, |\lambda_n(X)|)\,. \end{aligned}$$

Then for a fixed $0 < \lambda < 1$,

$$\Delta_\lambda^m(X) = V_m^* e^{t_m A_m} D_\theta D_m V_m \tag{2.19}$$

for some $D_m \in \mathcal{D}_+(n)$, $A_m \in S(n)$, $V_m \in U(n)$, $t_m \geqslant 0$ *such that*

1. $\lim_{m\to\infty} D_m = D$.
2. $\lim_{m\to\infty} t_m = 0$.

3. *For each* $m \in \mathbb{N}$, $\min\{\|A_m\|, \|D_m^{1-\lambda} A_m D_m^{\lambda-1}\|\} = 1$.

Proof. We write $X_m := \Delta_\lambda^m(X)$. Notice that if X_m can be expressed in the form (2.19), then by Theorem 1.4(2), property (1) holds by the continuity of singular values since $D_m \in \mathcal{D}_+(n)$ contains the singular values of X_m.

We now consider the following two cases:

Case 1: Some element of $\{X_m\}_{m\in\mathbb{N}}$ is normal. Let X_k be the *first* normal matrix in the sequence. Then by Lemma 2.4

$$X_k = X_{k+1} = X_{k+2} = \cdots$$

Since X_k is normal and have the same spectrum of X, we may write $X_k = V^* D_\theta D V$ for some $V \in U(n)$. Hence for all $m \geqslant k$,

$$X_m = V_m^* e^{t_m A_m} D_\theta D_m V_m,$$

where $D_m = D$, $A_m = I_n$, $t_m = 0$ and $V_m = V$. It is clear that (1), (2), and (3) are true.

Case 2: None of the elements in $\{X_m\}_{m\in\mathbb{N}}$ is normal. By Theorem 1.4(1) the limit points of $\{X_m\}_{m\in\mathbb{N}}$ are normal and are located in the orbit $\mathcal{O}$ of the diagonal $D_\theta D$ under unitary similarity

$$\mathcal{O} := \{V^* D_\theta D V \mid V \in U(n)\}.$$

Let

$$X_m = U_m D_m V_m \tag{2.20}$$

be a singular value decomposition of X_m, where $D_m \in \mathcal{D}_+(n)$, $U_m, V_m \in U(n)$. We can rewrite (2.20) in the following fashion:

$$\begin{aligned} X_m &= V_m^*(V_m U_m D_\theta^{-1}) D_\theta D_m V_m \\ &= V_m^* e^{B_m} D_\theta D_m V_m \qquad \text{by (2.17)} \end{aligned} \tag{2.21}$$

where $e^{B_m} = V_m U_m D_\theta^{-1}$ for some $B_m \in S(n)$. Notice that the matrix $D_m \in \mathcal{D}_+(n)$ is uniquely defined by X_m, but $V_m \in U(n)$ and $B_m \in S(n)$ are not unique. For each $m \in \mathbb{N}$, denote

$$\mathcal{S}_m := \{B \in S(n) \mid \text{there is } V_m' \in U(n) \text{ such that } X_m = V_m'^* e^B D_\theta D_m V_m'\}.$$

The set $\mathcal{S}_m$ is closed, since if $\{B^{(i)}\}_{i\in\mathbb{N}} \subset \mathcal{S}_m$ and $\lim_{i\to\infty} B^{(i)} = B$, then

$$X_m = (V^{(i)})^* e^{B^{(i)}} D_\theta D_m V^{(i)}$$

for some $\{V^{(i)}\}_{i\in\mathbb{N}} \subset U(n)$. Since $U(n)$ is compact, the sequence $\{V^{(i)}\}_{i\in\mathbb{N}}$ has at least one limit point $V \in U(n)$. So $X_m = V^* e^B D_\theta D_m V$ and thus $B \in \mathcal{S}_m$.

Since $\mathcal{S}_m$ is closed, we choose $B_m \in \mathcal{S}_m$ in (2.21) once and for all in the way that $\|B_m\|$ is *minimal* (the choice B_m still may not be unique). Since each X_m is not normal, $B_m \neq 0$. Write $B_m = t_m A_m$, that is, $A_m := \frac{B_m}{t_m}$, and adjust $t_m > 0$ appropriately, one has

$$\min\{\|A_m\|, \|D_m^{1-\lambda} A_m D_m^{\lambda-1}\|\} = 1.$$

So property (3) is satisfied.

It remains to prove property (2), i.e., $\lim_{m\to\infty} t_m = 0$, or equivalently,

$$\lim_{m\to\infty} B_m = \mathbf{0}, \tag{2.22}$$

since $\|B_m\| = t_m\|A_m\| \geqslant t_m$ and $\lim_{m\to\infty} D_m = D$. Suppose on the contrary that (2.22) is not true. There would exist $\epsilon > 0$ and a subsequence $\{B_{m_i}\}_{i\in\mathbb{N}}$ where

$$\|B_{m_i}\| \geqslant \epsilon, \qquad \text{for all } i \in \mathbb{N}. \tag{2.23}$$

By (1.5) the subsequence $\{X_{m_i}\}_{i\in\mathbb{N}}$ is bounded above by $\|X\|$. Thus $\{X_{m_i}\}_{i\in\mathbb{N}}$ has a convergent subsequence $\{X_{m_i'}\}_{i\in\mathbb{N}}$. By Theorem 1.4(1) $\lim_{i\to\infty} X_{m_i'}$ is a normal matrix of spectrum $\sigma(X)$, that is,

$$\lim_{i\to\infty} X_{m_i'} = V^* D_\theta D V$$

for some $V \in U(n)$. By Lemma 2.5, we may write

$$X_{m_i'} = V_{m_i'}^* e^{E_{m_i'}} D_\theta D_{m_i'} V_{m_i'}$$

where $V_{m_i'} \in U(n)$, $E_{m_i'} \in \mathcal{S}_m$, and $\lim_{i\to\infty} \|E_{m_i'}\| = \mathbf{0}$. This would force $\lim_{i\to\infty} \|B_{m_i'}\| = \mathbf{0}$ because of the choice of B_m and would contradict (2.23). So (2.22) and thus property (2) are established. □

LEMMA 2.7. *Suppose* $\{T_\ell\}_{\ell=0}^m \subset \mathbb{C}_{n\times n}$. *For any* $m \in \mathbb{N}$,

$$\sum_{\ell=0}^{m} \binom{m}{\ell} (-1)^\ell T_\ell = \sum_{\ell=1}^{m} \binom{m-1}{\ell-1} (-1)^\ell (T_\ell - T_{\ell-1}).$$

Proof. Recall the combinatorial identity

$$\binom{m}{\ell} = \binom{m-1}{\ell-1} + \binom{m-1}{\ell},$$

in which we adopt the usual convention: $\binom{m}{\ell} = 0$ if $m < \ell$ or $\ell < 0$. So

$$\begin{aligned}
\sum_{\ell=0}^{m} \binom{m}{\ell} (-1)^\ell T_\ell &= \sum_{\ell=0}^{m} \binom{m-1}{\ell-1} (-1)^\ell T_\ell + \sum_{\ell=0}^{m} \binom{m-1}{\ell} (-1)^\ell T_\ell \\
&= \sum_{\ell=1}^{m} \binom{m-1}{\ell-1} (-1)^\ell T_\ell + \sum_{\ell=0}^{m-1} \binom{m-1}{\ell} (-1)^\ell T_\ell \\
&= \sum_{\ell=1}^{m} \binom{m-1}{\ell-1} (-1)^\ell T_\ell + \sum_{\ell=1}^{m} \binom{m-1}{\ell-1} (-1)^{\ell-1} T_{\ell-1} \\
&= \sum_{\ell=1}^{m} \binom{m-1}{\ell-1} (-1)^\ell (T_\ell - T_{\ell-1}). \qquad \square
\end{aligned}$$

LEMMA 2.8. *Let* $A, D \in \mathbb{C}_{n\times n}$. *For* $m \in \mathbb{N}$,

$$\left\| \sum_{\ell=0}^{m} \binom{m}{\ell} (-1)^{\ell} A^{m-\ell} D^2 A^{\ell} \right\|_F \leqslant 2^{m-1} \left\| D^2A - AD^2 \right\|_F \|A\|^{m-1}.$$

Proof. Applying Lemma 2.7 with $T_\ell = A^{m-\ell} D^2 A^\ell$, we have

$$\begin{aligned}
&\left\| \sum_{\ell=0}^{m} \binom{m}{\ell} (-1)^{\ell} A^{m-\ell} D^2 A^{\ell} \right\|_F \\
= &\left\| \sum_{\ell=1}^{m} \binom{m-1}{\ell-1} (-1)^{\ell} \left(A^{m-\ell} D^2 A^{\ell} - A^{m-\ell+1} D^2 A^{\ell-1} \right) \right\|_F \\
= &\left\| \sum_{\ell=1}^{m} \binom{m-1}{\ell-1} (-1)^{\ell} A^{m-\ell} \left(D^2 A - AD^2 \right) A^{\ell-1} \right\|_F \\
\leqslant &\sum_{\ell=1}^{m} \binom{m-1}{\ell-1} \left\| A^{m-\ell} \left(D^2 A - AD^2 \right) A^{\ell-1} \right\|_F \\
\leqslant &\sum_{\ell=1}^{m} \binom{m-1}{\ell-1} \|A\|^{m-\ell} \left\| D^2 A - AD^2 \right\|_F \|A\|^{\ell-1} \\
= &\ 2^{m-1} \left\| D^2 A - AD^2 \right\|_F \|A\|^{m-1} \qquad \text{by } \sum_{\ell=1}^{m} \binom{m-1}{\ell-1} = 2^{m-1},
\end{aligned}$$

where the last inequality is obtained by using the inequalities $\|AB\|_F \leqslant \|A\|\|B\|_F$ and $\|AB\|_F \leqslant \|A\|_F\|B\|$. □

LEMMA 2.9. *Let* $D = diag\,(d_1, \ldots, d_n)$ *with positive* $d_1, \ldots, d_n$ *and* $A \in S(n)$. *For* $m \in \mathbb{N}$,

$$\begin{aligned}
&\left\| \sum_{\ell=0}^{m} \binom{m}{\ell} (-1)^{\ell} D^{1-\lambda} A^{\ell} D^{2\lambda} A^{m-\ell} D^{1-\lambda} \right\|_F \\
\leqslant &\ 2^{m-1} \left\| D^{1-\lambda} A D^{1+\lambda} - D^{1+\lambda} A D^{1-\lambda} \right\|_F \left\| D^{\lambda-1} A D^{1-\lambda} \right\|^{m-1} \qquad (2.24) \\
\leqslant &\ 2^{m-1} \left\| D^2 A - AD^2 \right\|_F \left\| D^{\lambda-1} A D^{1-\lambda} \right\|^{m-1}. \qquad (2.25)
\end{aligned}$$

Proof. Clearly we have

$$\left\| D^{\lambda-1} A D^{1-\lambda} \right\| = \left\| -\left(D^{1-\lambda} A D^{\lambda-1} \right)^* \right\|.$$

Applying Lemma 2.7 with $T_\ell = D^{1-\lambda}A^\ell D^{2\lambda}A^{m-\ell}D^{1-\lambda}$, we have

$$\begin{aligned}
&\left\|\sum_{\ell=0}^{m}\binom{m}{\ell}(-1)^\ell D^{1-\lambda}A^\ell D^{2\lambda}A^{m-\ell}D^{1-\lambda}\right\|_F\\
=&\left\|\sum_{\ell=1}^{m}\binom{m-1}{\ell-1}(-1)^\ell\left(D^{1-\lambda}A^\ell D^{2\lambda}A^{m-\ell}D^{1-\lambda}-D^{1-\lambda}A^{\ell-1}D^{2\lambda}A^{m-\ell+1}D^{1-\lambda}\right)\right\|_F\\
=&\left\|\sum_{\ell=1}^{m}\binom{m-1}{\ell-1}(-1)^\ell D^{1-\lambda}A^{\ell-1}\left(AD^{2\lambda}-D^{2\lambda}A\right)A^{m-\ell}D^{1-\lambda}\right\|_F\\
=&\left\|\sum_{\ell=1}^{m}\binom{m-1}{\ell-1}(-1)^\ell\left(D^{1-\lambda}AD^{\lambda-1}\right)^{\ell-1}\left(D^{1-\lambda}AD^{1+\lambda}-D^{1+\lambda}AD^{1-\lambda}\right)\right.\\
&\qquad\left.\left(D^{\lambda-1}AD^{1-\lambda}\right)^{m-\ell}\right\|_F\\
\leqslant&\sum_{\ell=1}^{m}\binom{m-1}{\ell-1}\left\|D^{1-\lambda}AD^{\lambda-1}\right\|^{\ell-1}\left\|D^{1-\lambda}AD^{1+\lambda}-D^{1+\lambda}AD^{1-\lambda}\right\|_F\left\|D^{\lambda-1}AD^{1-\lambda}\right\|^{m-\ell}\\
=&\ 2^{m-1}\left\|D^{1-\lambda}AD^{1+\lambda}-D^{1+\lambda}AD^{1-\lambda}\right\|_F\left\|D^{\lambda-1}AD^{1-\lambda}\right\|^{m-1},
\end{aligned}$$

where the last inequality is obtained by using the inequalities $\|AB\|_F \leqslant \|A\|\|B\|_F$ and $\|AB\|_F \leqslant \|A\|_F\|B\|$. So we have inequality (2.24).

The (i,j)-entry of $D^{1-\lambda}AD^{1+\lambda}-D^{1+\lambda}AD^{1-\lambda}$ is $a_{ij}(d_i^{1-\lambda}d_j^{1+\lambda}-d_i^{1+\lambda}d_j^{1-\lambda})$ and the (i,j)-entry of D^2A-AD^2 is $a_{ij}(d_i^2-d_j^2)$. We claim that

$$|d_i^{1-\lambda}d_j^{1+\lambda}-d_i^{1+\lambda}d_j^{1-\lambda}| \leqslant |d_i^2-d_j^2|. \tag{2.26}$$

For definiteness, suppose $d_i \geqslant d_j (>0)$. Then $|d_i^{1-\lambda}d_j^{1+\lambda}-d_i^{1+\lambda}d_j^{1-\lambda}| = d_i^{1+\lambda}d_j^{1-\lambda} - d_i^{1-\lambda}d_j^{1+\lambda}$ for $0<\lambda<1$ and $|d_i^2-d_j^2| = d_i^2-d_j^2$, and

$$d_i^2-d_j^2-(d_i^{1+\lambda}d_j^{1-\lambda}-d_i^{1-\lambda}d_j^{1+\lambda}) = (d_i^{1+\lambda}+d_j^{1+\lambda})(d_i^{1-\lambda}-d_j^{1-\lambda}) \geqslant 0.$$

Hence (2.26) is established and

$$\left\|D^{1-\lambda}AD^{1+\lambda}-D^{1+\lambda}AD^{1-\lambda}\right\|_F \leqslant \left\|D^2A-AD^2\right\|_F$$

so that (2.25) follows. □

Given $X \in \mathbb{C}_{n\times n}$, define

$$f(X) := \|X^*X - XX^*\|_F$$

which is interpreted as a measure of how close X to a normal matrix. For example, $f(X)=0$ if and only if X is normal. We interpret that X is close to a normal matrix if $f(X)$ is small. Notice that f is constant on the orbit of X under unitary similarity, that is,

$$f(X) = f(UXU^*), \quad U \in U(n). \tag{2.27}$$

The notation $g(t) = \mathrm{O}(t^k)$ for a real value function g means

$$\overline{\lim_{t\to 0}} \left|\frac{g(t)}{t^k}\right| \leqslant M$$

for some constant M.

LEMMA 2.10. *Let $0 < \lambda < 1$. Suppose that*

$$X = V^* e^{tA} D_\theta D V \in \mathrm{GL}_n(\mathbb{C})$$

is not normal, where $A \in S(n)$, $V \in U(n)$, $D = diag\,(d_1, \cdots, d_n) \in \mathcal{D}_+(n)$, and

$$D_\theta = diag\,(e^{i\theta_1}, \cdots, e^{i\theta_n}), \qquad \theta_1, \cdots, \theta_n \in \mathbb{R}.$$

Suppose further $0 < t < 1$ and $\min\{\|A\|, \|D^{1-\lambda} A D^{\lambda-1}\|\} \leqslant 1$. Then

$$\begin{aligned} \frac{f(\Delta_\lambda(X))}{f(X)} &\leqslant \sqrt{\frac{\sum_{i,j=1}^n |a_{ij}|^2 (d_i - d_j)^2 (d_i^\lambda d_j^{1-\lambda} + d_i^{1-\lambda} d_j^\lambda)^2}{\sum_{i,j=1}^n |a_{ij}|^2 (d_i^2 - d_j^2)^2}} + O(t) &\quad (2.28)\\ &\leqslant \alpha + O(t) &\quad (2.29) \end{aligned}$$

where the bounds for $O(t)$'s in (2.28) and (2.29) are independent of X, and

$$\alpha := \max_{1\leqslant i<j\leqslant n} \frac{d_i^\lambda d_j^{1-\lambda} + d_i^{1-\lambda} d_j^\lambda}{d_i + d_j}. \tag{2.30}$$

Moreover, $\alpha < 1$ whenever $d_1, \cdots, d_n$ are distinct.

Proof. By (1.4) and (2.27)

$$\frac{f(\Delta_\lambda(X))}{f(X)} = \frac{f(\Delta_\lambda(VXV^*))}{f(VXV^*)} = \frac{f(\Delta_\lambda(e^{tA} D_\theta D))}{f(e^{tA} D_\theta D)}. \tag{2.31}$$

Since X is not normal, the denominator

$$f(e^{tA} D_\theta D) = f(X) > 0. \tag{2.32}$$

Since $D_\theta \in D(n)$ and $D \in \mathcal{D}_+(n)$ commute, we have

$$\begin{aligned} f(e^{tA} D_\theta D) &= \left\| D^2 - e^{tA} D^2 e^{-tA} \right\|_F \\ &= \left\| D^2 - \left(\sum_{k=0}^\infty \frac{t^k A^k}{k!}\right) D^2 \left[\sum_{k=0}^\infty \frac{(-1)^k t^k A^k}{k!}\right] \right\|_F \\ &= \left\| t\left(D^2 A - A D^2\right) - \sum_{m=2}^\infty \frac{t^m}{m!} \left[\sum_{\ell=0}^m \binom{m}{\ell} (-1)^\ell A^{m-\ell} D^2 A^\ell\right] \right\|_F. \end{aligned}$$

We consider the second term of the last expression. Since $0 < t < 1$, one has $t^2 \geqslant t^m$ for all $m \geqslant 2$. Since $\|A\| \leqslant 1$,

$$\begin{aligned}
&\left\| \sum_{m=2}^{\infty} \frac{t^m}{m!} \left[\sum_{\ell=0}^{m} \binom{m}{\ell} (-1)^{\ell} A^{m-\ell} D^2 A^{\ell} \right] \right\|_F \\
\leqslant\ & \sum_{m=2}^{\infty} \frac{t^2}{m!} \left\| \sum_{\ell=0}^{m} \binom{m}{\ell} (-1)^{\ell} A^{m-\ell} D^2 A^{\ell} \right\|_F && \text{by } t^2 \geqslant t^m \\
\leqslant\ & t^2 \sum_{m=2}^{\infty} \frac{2^{m-1}}{m!} \left\| D^2 A - A D^2 \right\|_F && \text{by Lemma 2.8} \\
=\ & t^2 \frac{(e^2-3)}{2} \left\| D^2 A - A D^2 \right\|_F \\
=\ & \mathrm{O}\left(t^2\right) \left\| D^2 A - A D^2 \right\|_F .
\end{aligned}$$

Since

$$\|B\|_F - \|C\|_F \leqslant \|B + C\|_F \leqslant \|B\|_F + \|C\|_F, \qquad B, C \in \mathbb{C}_{n \times n}, \tag{2.33}$$

the denominator can be written as

$$f\left(e^{tA} D_{\theta} D\right) = \left(t + \mathrm{O}(t^2)\right) \left\| D^2 A - A D^2 \right\|_F . \tag{2.34}$$

On the other hand, the numerator is

$$\begin{aligned}
f\left(\Delta_{\lambda}\left(e^{tA} D_{\theta} D\right)\right) &= f\left(D^{\lambda} e^{tA} D_{\theta} D^{1-\lambda}\right) \\
&= \left\| D^{1-\lambda} D_{\theta}^{-1} e^{-tA} D^{2\lambda} e^{tA} D_{\theta} D^{1-\lambda} - D^{\lambda} e^{tA} D^{2-2\lambda} e^{-tA} D^{\lambda} \right\|_F \\
&= \Bigg\| D^{1-\lambda} D_{\theta}^{-1} \left[\sum_{k=0}^{\infty} \frac{(-1)^k t^k A^k}{k!} \right] D^{2\lambda} \left(\sum_{k=0}^{\infty} \frac{t^k A^k}{k!} \right) D_{\theta} D^{1-\lambda} \\
&\qquad - D^{\lambda} \left(\sum_{k=0}^{\infty} \frac{t^k A^k}{k!} \right) D^{2-2\lambda} \left[\sum_{k=0}^{\infty} \frac{(-1)^k t^k A^k}{k!} \right] D^{\lambda} \Bigg\|_F .
\end{aligned}$$

Set $B := D_{\theta}^{-1} A D_{\theta}$. Then

$$\begin{aligned}
f\left(\Delta_{\lambda}\left(e^{tA} D_{\theta} D\right)\right) &= \Bigg\| D^{1-\lambda} \left[\sum_{k=0}^{\infty} \frac{(-1)^k t^k B^k}{k!} \right] D^{2\lambda} \left(\sum_{k=0}^{\infty} \frac{t^k B^k}{k!} \right) D^{1-\lambda} \\
&\qquad - D^{\lambda} \left(\sum_{k=0}^{\infty} \frac{t^k A^k}{k!} \right) D^{2-2\lambda} \left[\sum_{k=0}^{\infty} \frac{(-1)^k t^k A^k}{k!} \right] D^{\lambda} \Bigg\|_F \\
&= \Bigg\| \sum_{m=0}^{\infty} \frac{t^m}{m!} \left[\sum_{\ell=0}^{m} \binom{m}{\ell} (-1)^{\ell} D^{1-\lambda} B^{\ell} D^{2\lambda} B^{m-\ell} D^{1-\lambda} \right] \\
&\qquad - \sum_{m=0}^{\infty} \frac{t^m}{m!} \left[\sum_{\ell=0}^{m} \binom{m}{\ell} (-1)^{\ell} D^{\lambda} A^{m-\ell} D^{2-2\lambda} A^{\ell} D^{\lambda} \right] \Bigg\|_F
\end{aligned}$$

$$= \Big\| t \left(D^{1+\lambda} B D^{1-\lambda} - D^{1-\lambda} B D^{1+\lambda} - D^{\lambda} A D^{2-\lambda} + D^{2-\lambda} A D^{\lambda} \right) + \sum_{m=2}^{\infty} \frac{t^m}{m!} \left[\sum_{\ell=0}^{m} \binom{m}{\ell} (-1)^{\ell} D^{1-\lambda} B^{\ell} D^{2\lambda} B^{m-\ell} D^{1-\lambda} \right] - \sum_{m=2}^{\infty} \frac{t^m}{m!} \left[\sum_{\ell=0}^{m} \binom{m}{\ell} (-1)^{\ell} D^{\lambda} A^{m-\ell} D^{2-2\lambda} A^{\ell} D^{\lambda} \right] \Big\|_F .$$

We now examine the middle term of the last expression. When $0 < t < 1$,

$$\begin{aligned} & \left\| \sum_{m=2}^{\infty} \frac{t^m}{m!} \left[\sum_{\ell=0}^{m} \binom{m}{\ell} (-1)^{\ell} D^{1-\lambda} B^{\ell} D^{2\lambda} B^{m-\ell} D^{1-\lambda} \right] \right\|_F \\ \leqslant\ & \sum_{m=2}^{\infty} \frac{t^2}{m!} \left\| \sum_{\ell=0}^{m} \binom{m}{\ell} (-1)^{\ell} D^{1-\lambda} B^{\ell} D^{2\lambda} B^{m-\ell} D^{1-\lambda} \right\|_F \\ \leqslant\ & \sum_{m=2}^{\infty} \frac{t^2}{m!} 2^{m-1} \left\| D^2 B - B D^2 \right\|_F \qquad \text{by Lemma2.9 and} \quad \|D^{\lambda-1} A D^{1-\lambda}\| \leqslant 1 \\ =\ & t^2 \frac{(e^2-3)}{2} \left\| D_{\theta}^{-1} D^2 A D_{\theta} - D_{\theta}^{-1} A D^2 D_{\theta} \right\|_F \qquad \text{since} \quad B := D_{\theta}^{-1} A D_{\theta} \\ =\ & \mathrm{O}(t^2) \left\| D^2 A - A D^2 \right\|_F . \end{aligned}$$

Likewise we examine the last term. Replacing λ by $1-\lambda$ in Lemma 2.9 and using the identity $\binom{m}{\ell} = \binom{m}{m-\ell}$, we get

$$\left\| \sum_{m=2}^{\infty} \frac{t^m}{m!} \left[\sum_{\ell=0}^{m} \binom{m}{\ell} (-1)^{\ell} D^{\lambda} A^{m-\ell} D^{2-2\lambda} A^{\ell} D^{\lambda} \right] \right\|_F = \mathrm{O}(t^2) \left\| D^2 A - A D^2 \right\|_F .$$

From the above computations,

$$\begin{aligned} & f\left(\Delta_{\lambda}\left(e^{tA} D_{\theta} D\right)\right) \\ =\ & t \left\| D^{1+\lambda} B D^{1-\lambda} - D^{1-\lambda} B D^{1+\lambda} - D^{\lambda} A D^{2-\lambda} + D^{2-\lambda} A D^{\lambda} \right\|_F \\ & \qquad + \mathrm{O}(t^2) \left\| D^2 A - A D^2 \right\|_F \\ =\ & t \left\| D^{1+\lambda} D_{\theta}^{*} A D_{\theta} D^{1-\lambda} - D^{1-\lambda} D_{\theta}^{*} A D_{\theta} D^{1+\lambda} - D^{\lambda} A D^{2-\lambda} + D^{2-\lambda} A D^{\lambda} \right\|_F \\ & \qquad + \mathrm{O}(t^2) \left\| D^2 A - A D^2 \right\|_F . \end{aligned} \tag{2.35}$$

Denote

$$\begin{aligned} P &:= \left\| D^{1+\lambda} D_{\theta}^{*} A D_{\theta} D^{1-\lambda} - D^{1-\lambda} D_{\theta}^{*} A D_{\theta} D^{1+\lambda} - D^{\lambda} A D^{2-\lambda} + D^{2-\lambda} A D^{\lambda} \right\|_F \\ Q &:= \left\| D^2 A - A D^2 \right\|_F . \end{aligned}$$

Then $Q > 0$ in view of (2.32) and (2.34). Substituting (2.34) and (2.35) into (2.31),

$$\frac{f\left(\Delta_{\lambda}(X)\right)}{f(X)} = \frac{tP + \mathrm{O}(t^2)Q}{(t+\mathrm{O}(t^2))Q} = \frac{P}{Q} + \frac{-\mathrm{O}(t^2)P + \mathrm{O}(t^2)Q}{(t+\mathrm{O}(t^2))Q}. \tag{2.36}$$

By direct computation,

$$\begin{aligned}
\frac{P}{Q} &= \frac{\left\| \left[\ e^{i(\theta_j-\theta_i)} d_i^{1+\lambda} a_{ij} d_j^{1-\lambda} - e^{i(\theta_j-\theta_i)} d_i^{1-\lambda} a_{ij} d_j^{1+\lambda} - d_i^{\lambda} a_{ij} d_j^{2-\lambda} + d_i^{2-\lambda} a_{ij} d_j^{\lambda}\ \right]_{n\times n} \right\|_F}{\left\| \left[\ d_i^2 a_{ij} - a_{ij} d_j^2\ \right]_{n\times n} \right\|_F} \\
&= \sqrt{\frac{\sum_{i,j=1}^n |a_{ij}|^2 \left| e^{i(\theta_j-\theta_i)} (d_i^{1+\lambda} d_j^{1-\lambda} - d_i^{1-\lambda} d_j^{1+\lambda}) + d_i^{2-\lambda} d_j^{\lambda} - d_i^{\lambda} d_j^{2-\lambda} \right|^2}{\sum_{i,j=1}^n |a_{ij}|^2 (d_i^2 - d_j^2)^2}}
\end{aligned}$$

Notice that the two terms in the above expressions

$$\begin{aligned}
d_i^{1+\lambda} d_j^{1-\lambda} - d_i^{1-\lambda} d_j^{1+\lambda} &= d_i d_j \left[\left(\frac{d_i}{d_j}\right)^{\lambda} - \left(\frac{d_j}{d_i}\right)^{\lambda} \right] \\
d_i^{2-\lambda} d_j^{\lambda} - d_i^{\lambda} d_j^{2-\lambda} &= d_i d_j \left[\left(\frac{d_i}{d_j}\right)^{1-\lambda} - \left(\frac{d_j}{d_i}\right)^{1-\lambda} \right]
\end{aligned}$$

are of the same sign, that is, both positive, negative, or zero. Thus

$$\begin{aligned}
\frac{P}{Q} &\leqslant \sqrt{\frac{\sum_{i,j=1}^n |a_{ij}|^2 (d_i^{1+\lambda} d_j^{1-\lambda} - d_i^{1-\lambda} d_j^{1+\lambda} + d_i^{2-\lambda} d_j^{\lambda} - d_i^{\lambda} d_j^{2-\lambda})^2}{\sum_{i,j=1}^n |a_{ij}|^2 (d_i^2 - d_j^2)^2}} \\
&= \sqrt{\frac{\sum_{i,j=1}^n |a_{ij}|^2 (d_i - d_j)^2 (d_i^{\lambda} d_j^{1-\lambda} + d_i^{1-\lambda} d_j^{\lambda})^2}{\sum_{i,j=1}^n |a_{ij}|^2 (d_i - d_j)^2 (d_i + d_j)^2}} \qquad (2.37) \\
&\leqslant \sqrt{\max_{\substack{1\leqslant i,j\leqslant n \\ d_i \neq d_j \\ a_{ij}\neq 0}} \frac{(d_i^{\lambda} d_j^{1-\lambda} + d_i^{1-\lambda} d_j^{\lambda})^2}{(d_i + d_j)^2}} \qquad (2.38) \\
&\leqslant \max_{1\leqslant i<j\leqslant n} \frac{d_i^{\lambda} d_j^{1-\lambda} + d_i^{1-\lambda} d_j^{\lambda}}{d_i + d_j} = \alpha. \qquad (2.39)
\end{aligned}$$

The inequality (2.38) comes from the fact that

$$\frac{a_1 + \cdots + a_k}{b_1 + \cdots + b_k} \leqslant \max_{1\leqslant i\leqslant k} \frac{a_i}{b_i} \qquad \text{if } a_i > 0 \text{ and } b_i > 0 \text{ for } 1 \leqslant i \leqslant k.$$

The expression (2.39) is due to symmetry. The constant $\alpha \leqslant 1$ since

$$d_i + d_j - d_i^{\lambda} d_j^{1-\lambda} - d_i^{1-\lambda} d_j^{\lambda} = (d_i^{\lambda} - d_j^{\lambda})(d_i^{1-\lambda} - d_j^{1-\lambda}) \geqslant 0.$$

Moreover, $\alpha < 1$ whenever $d_1, \cdots, d_n$ are distinct. Now $P/Q \leqslant \alpha \leqslant 1$. By (2.36), $\frac{O(t^2)}{t+O(t^2)} = O(t)$, (2.37) and (2.39),

$$\begin{aligned}\frac{f(\Delta_\lambda(X))}{f(X)} &= \frac{P}{Q} + O(t)\\ &\leqslant \sqrt{\frac{\sum_{i,j=1}^n |a_{ij}|^2 (d_i - d_j)^2 (d_i^\lambda d_j^{1-\lambda} + d_i^{1-\lambda} d_j^\lambda)^2}{\sum_{i,j=1}^n |a_{ij}|^2 (d_i - d_j)^2 (d_i + d_j)^2}} + O(t)\\ &\leqslant \alpha + O(t).\end{aligned}$$

The bounds for $O(t)$'s are independent of X by scrutinizing the process. □

COROLLARY 2.11. *Suppose that $X \in \mathrm{GL}_n(\mathbb{C})$ has distinct eigenvalue moduli*

$$|\lambda_1(X)| > \cdots > |\lambda_n(X)| > 0.$$

Suppose that $X_m := \Delta_\lambda^m(X)$ is not normal for all $m \in \mathbb{N}$. Then

$$\overline{\lim_{m\to\infty}} \frac{f(\Delta_\lambda(X_m))}{f(X_m)} \leqslant \alpha, \tag{2.40}$$

where

$$\alpha := \max_{1\leqslant i<j\leqslant n} \frac{|\lambda_i(X)|^\lambda |\lambda_j(X)|^{1-\lambda} + |\lambda_i(X)|^{1-\lambda} |\lambda_j(X)|^\lambda}{|\lambda_i(X)| + |\lambda_j(X)|} < 1. \tag{2.41}$$

Proof. Let D_θ and D be denoted as in Lemma 2.6, that is,

$$\begin{aligned} D_\theta &:= \mathrm{diag}\left(\frac{\lambda_1(X)}{|\lambda_1(X)|}, \cdots, \frac{\lambda_n(X)}{|\lambda_n(X)|}\right)\\ D &:= \mathrm{diag}\,(|\lambda_1(X)|, \cdots, |\lambda_n(X)|).\end{aligned}$$

Then by Lemma 2.6,

$$X_m = V_m^* e^{t_m A_m} D_\theta D_m V_m,$$

where $D_m \in D_+(n)$, $V_m \in U(n)$, $A_m \in S(n)$, $t_m \geqslant 0$ such that

$$\begin{cases} \lim_{m\to\infty} D_m = D\\ \lim_{m\to\infty} t_m = 0\\ \min\{\|A_m\|, \|D_m^{1-\lambda} A_m D_m^{\lambda-1}\|\} = 1.\end{cases} \tag{2.42}$$

Denote

$$D_m := \mathrm{diag}\,(d_1^{(m)}, \cdots, d_n^{(m)}), \tag{2.43}$$

$$\alpha_m := \max_{1\leqslant i<j\leqslant n} \frac{(d_i^{(m)})^\lambda (d_j^{(m)})^{1-\lambda} + (d_i^{(m)})^{1-\lambda} (d_j^{(m)})^\lambda}{d_i^{(m)} + d_j^{(m)}}. \tag{2.44}$$

Since X_m is not normal for all $m \in \mathbb{N}$, we have $f(X_m) > 0$ for all $m \in \mathbb{N}$. By Lemma 2.10,

$$\frac{f(\Delta_\lambda(X_m))}{f(X_m)} \leqslant \alpha_m + O(t_m),$$

where the bound for $\mathrm{O}(t_m)$ is independent of X_m. So by (2.42),

$$\overline{\lim_{m\to\infty}} \frac{f(\Delta_\lambda(X_m))}{f(X_m)} \leqslant \overline{\lim_{m\to\infty}} \alpha_m + \overline{\lim_{m\to\infty}} \mathrm{O}(t_m) = \alpha,$$

where α is given in (2.41), and $\alpha < 1$ since X has distinct eigenvalue moduli. $\square$

LEMMA 2.12. *If $X \in \mathrm{GL}_n(\mathbb{C})$ and $0 < \lambda < 1$, then*

$$\|\Delta_\lambda(X) - X\|_F \leqslant (n^{1/2-\lambda/4}\|X\|^{1-\lambda}) f(X)^{\lambda/2}. \tag{2.45}$$

Proof. The idea comes from the proof of [5, Theorem 4.6] for the 2×2 case. Let $X = UP$ be the polar decomposition of X, where $U \in U(n)$ and $P \in P(n)$. Then

$$\begin{aligned}
\|\Delta_\lambda(X) - X\|_F &= \|(P^\lambda U - UP^\lambda)P^{1-\lambda}\|_F \\
&\leqslant \|P^\lambda U - UP^\lambda\|_F \|P^{1-\lambda}\| &(2.46)\\
&= \|P^\lambda - UP^\lambda U^*\|_F \|P\|^{1-\lambda} \\
&= \|(P^2)^{\lambda/2} - (UP^2U^*)^{\lambda/2}\|_F \|P\|^{1-\lambda} \\
&= \|(X^*X)^{\lambda/2} - (XX^*)^{\lambda/2}\|_F \|X\|^{1-\lambda} \\
&\leqslant \|I_n\|_F^{1-\lambda/2} \|X^*X - XX^*\|_F^{\lambda/2} \|X\|^{1-\lambda} &(2.47)\\
&= (n^{1/2-\lambda/4}\|X\|^{1-\lambda}) f(X)^{\lambda/2},
\end{aligned}$$

where the inequality (2.46) follows from $\|AB\|_F \leqslant \|A\|_F\|B\|$ and the inequality (2.47) follows from an inequality of Bhatia and Kittaneh [6] (see [5, Proposition 2.5]). $\square$

Proof of Theorem 2.1.

The proof adopts some nice ideas in the proofs of [5, Theorem 4.6 and Corollary 4.16]. Let $X_m := \Delta_\lambda^m(X)$. There are two cases:

Case 1: X is nonsingular with distinct eigenvalue moduli.

We now consider two possibilities:

(i) X_m is normal for some $m \in \mathbb{N}$. Then by Lemma 2.4 we have the convergence.

(ii) X_m is not normal for all $m \in \mathbb{N}$. Then $f(X_m) > 0$ for all $m \in \mathbb{N}$. We will show that the sequence $\{X_m\}_{m\in\mathbb{N}}$ is a Cauchy sequence. By Corollary 2.11 for each $\epsilon > 0$ with $\alpha + \epsilon < 1$, there is $N_\epsilon \in \mathbb{N}$ such that whenever $m > N_\epsilon$,

$$\frac{f(\Delta(X_m))}{f(X_m)} < \alpha + \epsilon < 1.$$

So

$$f(X_m) = f(X_{N_\epsilon}) \prod_{i=N_\epsilon}^{m-1} \frac{f(X_{i+1})}{f(X_i)} \leqslant (\alpha+\epsilon)^{m-N_\epsilon} f(X_{N_\epsilon}). \tag{2.48}$$

Given $m_2 > m_1 > N_\epsilon$,

$$\begin{aligned}
& \|X_{m_2} - X_{m_1}\|_F \\
\leqslant\ & \sum_{i=m_1}^{m_2-1} \|X_{i+1} - X_i\|_F \\
\leqslant\ & \sum_{i=m_1}^{m_2-1} \left(n^{1/2-\lambda/4}\|X_i\|^{1-\lambda}\right) f(X_i)^{\lambda/2} && \text{by Lemma 2.12} \\
\leqslant\ & \left(n^{1/2-\lambda/4}\|X\|^{1-\lambda}\right) \sum_{i=m_1}^{m_2-1} f(X_i)^{\lambda/2} && \text{by (1.5)} \\
\leqslant\ & \left(n^{1/2-\lambda/4}\|X\|^{1-\lambda}\right) \sum_{i=m_1}^{m_2-1} (\alpha+\epsilon)^{(i-N_\epsilon)\lambda/2} f(X_{N_\epsilon})^{\lambda/2} && \text{by (2.48)} \\
=\ & \left[n^{1/2-\lambda/4}\|X\|^{1-\lambda}(\alpha+\epsilon)^{-N_\epsilon\lambda/2} f(X_{N_\epsilon})^{\lambda/2}\right] \sum_{i=m_1}^{m_2-1} (\alpha+\epsilon)^{i\lambda/2} \\
\leqslant\ & M(\alpha+\epsilon)^{m_1\lambda/2} \to 0 \qquad \text{as } m_1 \to \infty,
\end{aligned}$$

where M is a constant independent of m_1 and m_2:

$$\begin{aligned}
M &:= [n^{1/2-\lambda/4}\|X\|^{1-\lambda}(\alpha+\epsilon)^{-N_\epsilon\lambda/2} f(X_{N_\epsilon})^{\lambda/2}] \sum_{i=0}^{\infty} (\alpha+\epsilon)^{i\lambda/2} \\
&= [n^{1/2-\lambda/4}\|X\|^{1-\lambda}(\alpha+\epsilon)^{-N_\epsilon\lambda/2} f(X_{N_\epsilon})^{\lambda/2}] \frac{1}{1-(\alpha+\epsilon)^{\lambda/2}}.
\end{aligned}$$

So $\{X_m\}_{m\in\mathbb{N}}$ is a Cauchy sequence and thus convergent.

Case 2: X is singular whose nonzero eigenvalues are of distinct moduli.

Let r be the size of the largest Jordan block of X corresponding to the zero eigenvalue. By [5, Proposition 4.14(1)], the Jordan structure for the zero eigenvalue in X_{r-1} is trivial, that is, all the Jordan blocks of X_{r-1} corresponding to the zero eigenvalue are 1×1. By the proof of [5, Corollary 4.16], there is $U \in U(n)$ such that

$$X_r = U^* \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix} U$$

where $S \in \mathrm{GL}_{n-r}(\mathbb{C})$. The eigenvalues of S are the nonzero eigenvalues of X. So S has distinct eigenvalue moduli and thus $\{\Delta_\lambda^m(S)\}_{m\in\mathbb{N}}$ converges by Case 1. By (1.4) and the fact that $\Delta_\lambda(A \oplus B) = \Delta_\lambda(A) \oplus \Delta_\lambda(B)$,

$$X_{m+r} = U^* \begin{bmatrix} \Delta_\lambda^m(S) & 0 \\ 0 & 0 \end{bmatrix} U.$$

So $\{X_m\}_{m\in\mathbb{N}}$ converges. $\square$

Proof of Theorem 2.2.

Using (1.4) and $\Delta_\lambda(A \oplus B) = \Delta_\lambda(A) \oplus \Delta_\lambda(B)$, it is sufficient to consider $X = T$ where T is of one of the four forms. As in the proof of Theorem 2.1, it is further reduced to the nonsingular T. Then use Theorem 2.1 to handle (2), Theorem 1.4(1) and (3) to handle (1) and (3), respectively. As to (4), if $\Delta_\lambda^q(T)$ is normal for some $q \in \mathbb{N}$, then $\Delta_\lambda^{q+m}(T) = \Delta_\lambda^q(T)$ for all $m \in \mathbb{N}$ and so $\{\Delta_\lambda^m(T)\}_{m\in\mathbb{N}}$ converges. □

3. Some remarks

In general when $\lambda \notin [0,1)$ (the case $\lambda = 0$ is trivial), the λ-Aluthge sequence may not converge. In particular we consider $\lambda = 1$ and $D(X) := \Delta_1(X)$ is called the Duggal transform [8] of X.

EXAMPLE 3.1. The Duggal sequence $\{X_m\}_{m\in\mathbb{N}} := \{D^m(X)\}_{m\in\mathbb{N}}$ does not converge in general. Indeed $\{P_m\}_{m\in\mathbb{N}}$ may not converge where $X_m = U_m P_m$ is the polar decomposition of X_m. For example,

$$\begin{aligned} X &:= \begin{bmatrix} -1 & -1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \\ X_1 &= \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \\ X_2 &= \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = X, \ldots \end{aligned}$$

So $\{P_m\}_{m\in\mathbb{N}}$ and $\{X_m\}_{m\in\mathbb{N}}$ are alternating.

REMARK 3.2. Though the nonlinear map $\Delta_\lambda : \mathbb{C}_{n\times n} \to \mathbb{C}_{n\times n}$ is continuous [5, Theorem 3.6] for each $0 < \lambda < 1$, it is neither injective or surjective. For example, let $N = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Then $\Delta_\lambda(N) = \mathbf{0}$ but there is no $A \in \mathbb{C}_{2\times 2}$ such that $\Delta_\lambda(A) = N$ by [5, Proposition 4.14].

Numerical experiences suggest the following

CONJECTURE 3.3. *Let* $0 < \lambda < 1$.

$$\|X^*X - XX^*\|_F \geqslant \|\Delta_\lambda(X)^*\Delta_\lambda(X) - \Delta_\lambda(X)\Delta_\lambda(X)^*\|_F \tag{3.1}$$

for all $X \in \mathbb{C}_{n\times n}$.

If the conjecture is true, then $\{\|X_m^*X_m - X_mX_m^*\|_F\}_{m\in\mathbb{N}}$ is always a nonincreasing sequence convergent to 0 by Theorem 1.4 where $X_m := \Delta_\lambda^m(X)$.

REMARK 3.4. One may want to have the representation (2.1) of X_m in Lemma 2.5 for all normal $A \in \mathrm{GL}_n(\mathbb{C})$:

$$X_m = V_m^* e^{B_m} D_\theta D_m V_m,$$

such that $\lim_{m\to\infty} B_m = \mathbf{0}$. But this is not true in general. The assumption that eigenvalues of A are identical if they have the same moduli in Lemma 2.5 is equivalent to $D_\theta = p(D)$ for some polynomial $p \in \mathbb{C}[x]$. It is not hard to see that it amounts to say that D_θ commutes with every permutation matrix commuting with D. In Lemma 2.5, if D_θ is not a polynomial of D, then the statement does not hold. In such case, there is a permutation matrix V such that $DV = VD$ but $D_\theta V \neq VD_\theta$. There is $\{D_m\}_{m\in\mathbb{N}} \subset \mathcal{D}_+(n)$ such that each D_m has distinct diagonal entries and $\lim_{m\to\infty} D_m = D$. Denote $X_m = D_\theta V^* D_m V$. Then

$$\lim_{m\to\infty} X_m = D_\theta V^* DV = D_\theta D.$$

We show by contradiction that $X_m \in \mathrm{GL}_n(\mathbb{C})$ cannot be expressed in the form (2.1). If (2.1) were true, then X_m would have two polar decompositions

$$X_m = D_\theta(V^* D_m V) = (V_m^* e^{B_m} D_\theta V_m)(V_m^* D_m V_m).$$

By the uniqueness of polar decomposition of $\mathrm{GL}_n(\mathbb{C})$,

$$D_\theta = V_m^* e^{B_m} D_\theta V_m \qquad V^* D_m V = V_m^* D_m V_m. \tag{3.2}$$

By the second equality of (3.2), $V'_m := V_m V^*$ commutes with D_m. So $V'_m \in D(n)$ since D_m has distinct diagonal entries. Then D_θ and V'_m commute. From the first equality of (3.2) we get

$$e^{B_m} = V_m D_\theta V_m^* {D_\theta}^{-1} = V'_m V D_\theta V^* V_m'^* {D_\theta}^{-1} = V'_m (V D_\theta V^* {D_\theta}^{-1}) V_m'^*.$$

Then we get

$$\lim_{m\to\infty} V'_m (V D_\theta V^* {D_\theta}^{-1}) V_m'^* = \lim_{m\to\infty} e^{B_m} = I_n.$$

So $V D_\theta V^* {D_\theta}^{-1} = I_n$. This contradicts $VD_\theta \neq D_\theta V$. So the desired representation in Lemma 2.5 does not hold in this situation.

REFERENCES

[1] A. ALUTHGE, *On p-hyponormal operators for $0 < p < 1$*, Integral Equations Operator Theory, **13** (1990), 307–315.

[2] A. ALUTHGE, *Some generalized theorems on p-hyponormal operators*, Integral Equations Operator Theory, **24** (1996), 497–501.

[3] T. ANDO, *Aluthge transforms and the convex hull of the eigenvalues of a matrix*, Linear Multilinear Algebra, **52** (2004), 281–292.

[4] T. ANDO AND T. YAMAZAKI, *The iterated Aluthge transforms of a 2-by-2 matrix converge*, Linear Algebra Appl., **375** (2003), 299–309.

[5] J. ANTEZANA, P. MASSEY AND D. STOJANOFF, *λ-Aluthge transforms and Schatten ideals*, Linear Algebra Appl., **405** (2005) 177–199.

[6] R. BHATIA AND F. KITTANEH, *Some inequalities for norms of commutators*, SIAM J. Matrix Anal. Appl., **18** (1997) 258–263.

[7] M. CHŌ, I. B. JUNG AND W. Y. LEE, *On Aluthge transform of p-hyponormal operators*, Integral Equations Operator Theory, **53** (2005), 321–329.

[8] C. FOIAŞ, I. B. JUNG, E. KO AND C. PEARCY, *Complete contractivity of maps associated with the Aluthge and Duggal transforms*, Pacific J. Math., **209** (2003), 249–259.

[9] P. R. HALMOS, *A Hilbert Space Problem Book*, Springer-Verlag, New York, 1974.

[10] S. HELGASON, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Academic Press, New York, 1978.

[11] I. B. JUNG, E. KO AND C. PEARCY, *Aluthge transforms of operators*, Integral Equations Operator Theory, **37** (2000), 437–448.

[12] I. B. JUNG, E. KO AND C. PEARCY, *Spectral pictures of Aluthge transforms of operators*, Integral Equations Operator Theory, **40** (2001), 52–60.

[13] I. B. JUNG, E. KO AND C. PEARCY, *The iterated Aluthge transform of an operator*, Integral Equations Operator Theory, **45** (2003), 375–387.

[14] K. OKUBO, *On weakly unitarily invariant norm and the Aluthge transformation*, Linear Algebra Appl., **371** (2003), 369–375.

[15] A. L. ONISHCHIK AND E. B. VINBERG, *Lie groups and algebraic groups*, Springer-Verlag, Berlin, 1990.

[16] T. YAMAZAKI, *An expression of spectral radius via Aluthge transformation*, Proc. Amer. Math. Soc., **130** (2002), 1131–1137.

[17] T. YAMAZAKI, *On numerical range of the Aluthge transformation*, Linear Algebra Appl., **341** (2002) 111–117.

(Received September 18, 2006)

Department of Mathematics and Statistics
Auburn University
AL 36849–5310, USA
e-mail: huanghu@auburn.edu, tamtiny@auburn.edu