Convergence of iterated Aluthge transform sequence for diagonalizable matrices

Jorge Antezana* Enrique R. Pujals † Demetrio Stojanoff ‡

July 4, 2018

Jorge Antezana and Demetrio Stojanoff
Depto. de Matemática, FCE-UNLP, La Plata, Argentina and IAM-CONICET
e-mail: antezana@mate.unlp.edu.ar and demetrio@mate.unlp.edu.ar

Enrique R. Pujals
Instituto Nacional de Matemática Pura y Aplicada (IMPA), Rio de Janeiro, Brasil.
e-mail: enrique@impa.br

Abstract

Given an $r \times r$ complex matrix T, if $T = U|T|$ is the polar decomposition of T, then, the Aluthge transform is defined by

$$\Delta(T) = |T|^{1/2} U |T|^{1/2}.$$

Let $\Delta^n(T)$ denote the n-times iterated Aluthge transform of T, i.e. $\Delta^0(T) = T$ and $\Delta^n(T) = \Delta(\Delta^{n-1}(T))$, $n \in \mathbb{N}$. We prove that the sequence $\{\Delta^n(T)\}_{n\in\mathbb{N}}$ converges for every $r \times r$ **diagonalizable** matrix T. We show that the limit $\Delta^\infty(\cdot)$ is a map of class C^∞ on the similarity orbit of a diagonalizable matrix, and on the (open and dense) set of $r \times r$ matrices with r different eigenvalues.

Keywords: Aluthge transform, Stable manifold theorem, similarity orbit, polar decomposition.

AMS Subject Classifications: Primary 37D10. Secondary 15A60.

*Partially supported by CONICET (PIP 4463/96), Universidad de La PLata (UNLP 11 X350) and ANPCYT (PICT03-09521).

†Partially supported by CNPq

‡Partially supported by CONICET (PIP 4463/96), Universidad de La PLata (UNLP 11 X350) and ANPCYT (PICT03-09521).

1 Introduction

Let $\mathcal{H}$ be a Hilbert space and T a bounded operator defined on $\mathcal{H}$ whose polar decomposition is $T = U|T|$. The *Aluthge transform* of T is the operator $\Delta(T) = |T|^{1/2}U\,|T|^{1/2}$. This was first studied in [1] in relation with the so-called p-hyponormal and log-hyponormal operators. Roughly speaking, the Aluthge transform of an operator is closer to being normal.

The Aluthge transform has received much attention in recent years. One reason is the connection of Aluthge transform with the invariant subspace problem. Jung, Ko and Pearcy proved in [8] that T has a nontrivial invariant subspace if an only if $\Delta(T)$ does. On the other hand, Dykema and Schultz proved in [6] that the Brown measures is unchanged by the Aluthge transform.

Another reason is related with the iterated Aluthge transform. Let $\Delta^0(T) = T$ and $\Delta^n(T) = \Delta(\Delta^{n-1}(T))$ for every $n \in \mathbb{N}$. It was conjectured in [8] that the sequence $\{\Delta^n(T)\}_{n\in\mathbb{N}}$ converge in the norm topology. Although this conjecture was stated for operators on an arbitrary Hilbert space, it was corrected and restated for matrices in [9] by Jung Ko and Pearcy and receantly extended to finite factors in [6] by Dykema and Schultz. In these spaces, it still remains open and there only exist some partial results. For instance, Ando and Yamazaki proved in [3] that the conjecture is true for 2×2 matrices and Dykema and Schultz in [6] proved that the conjecture is true for an operator T in a finite factor such that the unitary part of its polar decomposition normalizes an abelian subalgebra that contains $|T|$. (see [2], [14] and [15] for other results that support the conjecture in finite factors).

A result proved independently by Jung, Ko and Pearcy in [9], and by Ando in [2], states that, given an $r\times r$ matrix T, the limit points of the sequence $\{\Delta^n(T)\}_{n\in\mathbb{N}}$ are normal matrices with the same characteristic polynomial as T. In particular, if the sequence of iterated Aluthge transforms converge, the limit function, defined by $T \mapsto \lim_{n\to\infty} \Delta^n(T)$, whould be a retraction from the space of matrices onto the set of normal operators.

Another important result, concerning the finite dimensional case, states that it is enough to prove the conjecture for invertible matrices (see for example [4]). Note that, for an invertible matrix T

$$\Delta(T) = |T|^{1/2}\,T\,|T|^{-1/2}.$$

So the Aluthge transform of T belongs to the similarity orbit of T. This suggest that we can study the Aluthge transform restricted to the similarity orbit of some invertible operator.

From that point of view, the diagonalizable case has some advantages. First of all, note that the similarity orbit of a diagonalizable operator contains a compact submanifold of fixed points, and the sequence $\{\Delta^n(T)\}_{n\in\mathbb{N}}$ goes to this submanifold as $n\to\infty$. In fact, since T is diagonalizable, the similarity orbit of T coincides with the similarity orbit of some diagonal operator D, which we denote $\mathcal{S}(D)$. The unitary orbit of D, denoted by $\mathcal{U}(D)$, is a compact submanifold of $\mathcal{S}(D)$ that consists of all normal matrices in $\mathcal{S}(D)$. Hence $\mathcal{U}(D)$ is fixed by the Aluthge transform and the limits points of the sequence $\{\Delta^n(T)\}_{n\in\mathbb{N}}$ belongs to $\mathcal{U}(D)$. In contrast, for non-diagonalizable operators, the similarity orbit does not have fixed points, and the sequence of iterated Aluthge transforms goes to points that do not belong to the similarity orbit.

On the other hand, numerical computations, as well as Ando-Yamazaki's 2×2 computations (see [3]), suggest that the rate of convergence of the sequence $\{\Delta^n(T)\}_{n\in\mathbb{N}}$, for diagonalizable operators T, becomes exponential after some iterations. However, it seems that this behavior is not shared by the non-diagonalizable case.

For these reasons, we decided to study the diagonalizable case. Note that if we restrict the Aluthge transform to the similarity orbit of an invertible diagonalizable matrix T, a dynamical system approach can be performed.

In fact, we show that for any $N \in \mathcal{U}(D)$ there is a local submanifold $\mathcal{W}^s_N$ transversal to $\mathcal{U}(D)$ characterized by the matrices that converges with a exponential rate to N by the iteration of the Aluthge transform. Moreover, the union of these submanifolds form an open neighbourhood of $\mathcal{U}(D)$ (see Corollary 3.1.2). Thus, since the sequence $\{\Delta^n(T)\}_{n\in\mathbb{N}}$ goes towards $\mathcal{U}(D)$, for some n_0 large enough the sequence of iterated Aluthge tranforms enters this open neighborhood and converge exponentially.

These results follow from the classical arguments of stable manifolds (first introduced independently by Hadamard and Perron, see theorem 2.1.3; for details and general results about the stable manifold theorem see [7] or the Appendix at the end of this work). To conclude that, it is shown that the derivative of the Aluthge transform in any $N \in \mathcal{U}(D)$ has two invariant complementary directions, one tangent to $\mathcal{U}(D)$, and other transversal to it, where the derivative is a contraction (see Theorem 3.1.1). Using these results, we prove that the sequence $\{\Delta^n(T)\}_{n\in\mathbb{N}}$ converges for every $r \times r$ **diagonalizable** matrix T. We also show that the limit $\Delta^\infty(\cdot)$ is a map of class C^∞ on the similarity orbit of a diagonalizable matrix, and on the (open and dense) set of $r \times r$ matrices with r different eigenvalues.

This paper is organized as follows: in section 2, we collect several preliminary definitions and results about the the stable manifold theorem, about the geometry of similarity and unitary orbits, and about known results on Aluthge transform. In section 3, we prove the convergence results and we study the smoothness of the limit map $T \mapsto \Delta^\infty(T)$, mainly for $r \times r$ matrices with r different eigenvalues. The basic tool, to apply the stable manifold theorem to the similarity orbit of a diagonal matrix, is the mentioned Theorem 3.1.1, whose proof, somewhat technical, is done in section 4. In the Appendix, we sketch the proof of the classical version of the stable manifold theorem in order to show how it can be modified in our context, where the invariant set is a smooth submanifold consisting of fixed points, getting stronger results on the regularity conditions of the prelamination $\{\mathcal{W}^s_N\}_{N\in\mathcal{U}(D)}$.

We would like to aknowledge Prof. M. Shub for comments and suggestion about the stable manifold theorems, and Prof. G. Corach who told us about the Aluthge transform, and shared with us fruitful discussions concerning these matters.

2 Preliminaries.

In this paper $\mathcal{M}_r(\mathbb{C})$ denotes the algebra of complex $r \times r$ matrices, $\mathcal{G}l_r(\mathbb{C})$ the group of all invertible elements of $\mathcal{M}_r(\mathbb{C})$, $\mathcal{U}(r)$ the group of unitary operators, and $\mathcal{M}^h_r(\mathbb{C})$ (resp. $\mathcal{M}^{ah}_r(\mathbb{C})$) denotes the real algebra of hermitian (resp. antihermitian) matrices. Given $T \in$

$\mathcal{M}_r(\mathbb{C})$, $R(T)$ denotes the range or image of T, $\ker(T)$ the null space of T, $\sigma(T)$ the spectrum of T, $\operatorname{tr}(T)$ the trace of T, and T^* the adjoint of T. If $v \in \mathbb{C}^r$, we debote by $\operatorname{diag}(v) \in \mathcal{M}_r(\mathbb{C})$ the diagonal matrix with v in its diagonal. We shall consider the space of matrices $\mathcal{M}_r(\mathbb{C})$ as a real Hilbert space with the inner product defined by

$$\langle A,\ B\rangle = \mathbb{R}\mathrm{e}\left(\operatorname{tr}(B^*A)\right).$$

The norm induced by this inner product is the so-called Frobenius norm, that is denoted by $\|\cdot\|_2$. Along this note we also use the fact that every subspace $\mathcal{S}$ of $\mathbb{C}^n$ induces a representation of elements of $\mathcal{M}_r(\mathbb{C})$ by 2×2 block matrices, that is, we shall identify each $A \in \mathcal{M}_r(\mathbb{C})$ with a 2×2-block matrix

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{array}{c} \mathcal{S} \\ \mathcal{S}^\perp \end{array},$$

where $A_{11} = A|_{\mathcal{S},\mathcal{S}}$, $A_{12} = A|_{\mathcal{S}^\perp,\mathcal{S}}$, $A_{21} = A|_{\mathcal{S},\mathcal{S}^\perp}$ and $A_{22} = A|_{\mathcal{S}^\perp,\mathcal{S}^\perp}$.

On the other hand, let M be a manifold. By means of TM we denote the tangent bundle of M and by means of T_xM we denote the tangent space at the point $x \in M$. Given a function $f \in C^r(M)$, where $r = 1,\ldots,\infty$, $T_xf\,(v)$ denotes the derivative of f at the point x applied to the vector v.

2.1 Stable manifold theorem

In this section we state the stable manifold theorem for an invariant set of a smooth endomorphism (see 2.1.4 below). The stable set is naturally defined for a fixed point of an endomorphism, as the set of points with positive trajectories heading directly towards the fixed point. This notion is the natural extension of the stable eigenspaces of a linear transformation (the ones associated to the eigenvectors with modulus smaller than one) into the nonlinear regimen. In fact, a natural intuitive approach to the idea of the stable manifold is to consider a fixed point of a smooth differentiable map such that the derivative of the map at the fixed point has absolute value smaller than one. In this case, the linear map induced by the derivative is a map that share the same fixed point and such that any trajectory converges by forward iterate to the fixed point with a exponential rate of contraction. Using that the linear map is a "good approximation of the map in a small neighborhood of the fixed point", it follows that the map has the same dynamical behavior of its linear part.

A more general approach is based in the techniques known as graph transform operator. This approach can be naturally extended for invariant sets, being almost straightforward when the set consists of fixed points. An sketched version of the proof of Theorem 2.1.4, using these techniques, is done in the Appendix at the end of this work (see also [7, Thm 5.5]).

Let M be a smooth Riemann manifold and $N \subseteq M$ a submanifold (not necessarily compact). Throughout this subsection T_NM denotes the tangent bundle of M restricted to N.

Definition 2.1.1. A C^r *pre-lamination* indexed by N is a continuous choice of a C^r embedded disc $\mathcal{B}_x$ through each $x \in N$. Continuity means that N is covered by open sets $\mathcal{U}$ in which $x \to B_x$ is given by

$$\mathcal{B}_x = \sigma(x)((-\varepsilon, \varepsilon)^k)$$

where $\sigma : \mathcal{U} \cap N \to \mathrm{Emb}^r((-\varepsilon, \varepsilon)^k, M)$ is a continuous section. Note that $\mathrm{Emb}^r((-\varepsilon, \varepsilon)^k, M)$ is a C^r fiber bundle over M whose projection is $\beta \to \beta(0)$. Thus $\sigma(x)(0) = x$. If the sections mentioned above are C^s, $1 \leq s \leq r$, we say that the C^r pre-lamination is of class C^s.

Definition 2.1.2. A prelamination is *self coherent* if the interiors of each pair of its discs meet in a relatively open subset of each.

Definition 2.1.3. Let f be a smooth endomorphism of M, $\rho > 0$, and suppose that $f|_N$ is a homeomorphism. Then, N is *ρ-pseudo hyperbolic* for f if there exist two smooth subbundles of $T_N M$, denoted by $\mathcal{E}^s$ and $\mathcal{F}$, such that

1. $T_N M = \mathcal{E}^s \oplus \mathcal{F}$;

2. $TN = \mathcal{F}$;

3. Both, $\mathcal{E}^s$ and $\mathcal{F}$, are Tf-invariant;

4. $T f$ restricted to $\mathcal{F}$ is an automorphism, which expand it by a factor greater than ρ.

5. $T_x f : \mathcal{E}^s_x \to \mathcal{E}^s_{f(x)}$ has norm lower than ρ. ▲

Observe that if N is *ρ-pseudo hyperbolic* then there exists a positive constant $\lambda = \lambda(\rho) < 1$ such that

$$\frac{||Df_{\mathcal{E}^s}||}{m(Df_{|\mathcal{F}})} < \lambda \;, \tag{1}$$

where $m(.)$ means the minimum norm. If N consists of fixed points then, for example, N is *ρ-pseudo hyperbolic* (also called *normally hyperbolic*) if there is a $Tf-$invariant subbundle $\mathcal{E}^s$ (of $T_N M$) complement to TN, such that Tf contracts more sharply than any contraction in TN. In the case that $\mathcal{E}^s$ is uniformly contracted, it follows that for any point $x \in N$ it is possible to find an $f-$invariant submanifold transversal to N tangent to $\mathcal{E}^s$ and characterized as the set of points with trajectories asymptotic to the trajectory of x.

Theorem 2.1.4 (Stable manifold theorem). *Let f be a C^r endomorphism of M with a ρ-pseudo hyperbolic submanifold N with $\rho < 1$. Then, there is a f-invariant and self coherent C^r-pre-lamination of class C^0, $\mathcal{W}^s : N \to Emb^r((-1,1)^k, M)$ such that, for every $x \in N$,*

1. $\mathcal{W}^s(x)(0) = x$,

2. $\mathcal{W}^s_x = \mathcal{W}^s(x)((-1,1)^k)$ *is tangent to* $\mathcal{E}^s_x$ *at every* $x \in N$,

3. $\mathcal{W}^s_x \subseteq \Big\{ y \in M \;:\; \mathrm{dist}(f^n(x), f^n(y)) < \mathrm{dist}(x,y)\rho^n \Big\}$.

Proof. See the proof in subsection A.1 of the Appendix. ■

Corollary 2.1.5 (Smoothness of the stable lamination for a submanifold of fixed points). *Let f, M and N as in Theorem 2.1.4. Let us assume that any point p in N is a* **fixed point**. *Then C^r-pre-lamination $\mathcal{W}^s : \mathcal{N} \to Emb^r((-1,1)^k, M)$ is of class C^r.*

Proof. See Corollary A.4.1 in the Appendix. ■

Remark 2.1.6. Observe that, from Theorem 2.1.4, it holds that, for every $x \in N$

$$T_x \mathcal{W}^s_x = \mathcal{E}^s_x \ .$$

If N consists on fixed pionts, from the regularity conditions of the pre-lamination $\{\mathcal{W}^s_x\}_{x\in N}$ assured by Corollary 2.1.5, we get that, for any $x \in N$, there exists $\gamma > 0$ such that

$$B(x,\gamma) \subset \bigcup_{x\in N} \mathcal{W}^s_x \ .$$

In other words, it means that $\bigcup_{x\in N} \mathcal{W}^s_x$ contains an open neighborhood $\mathcal{W}(N)$ of N in M. Therefore, condition 3 of Theorem 2.1.4 implies that for every $x \in N$ there exists an open neighborhood $\mathcal{U}$ of x (open relative to M) such that

$$\mathcal{W}^s_x \cap \mathcal{U} = \Big\{ y \in \mathcal{U} : \ \operatorname{dist}(x, f^n(y)) < \ \operatorname{dist}(x,y)\, \rho^n \Big\}. \tag{2}$$

In particular, $\mathcal{W}^s_x \cap \mathcal{W}^s_y = \varnothing$ if $x \neq y$. Moreover, we can assure that the (well defined) map

$$p : \mathcal{W}(N) \to N \qquad \text{given by} \qquad p(a) = x \quad \text{if} \quad a \in \mathcal{W}^s_x(x) \tag{3}$$

is of class C^r. ▲

2.2 Similarity orbit of a diagonal matrix

In this subsection we recall some facts about the similarity orbit of a diagonal matrix.

Definition 2.2.1. Let $D \in \mathcal{M}_r(\mathbb{C})$. By means of $\mathcal{S}(D)$ we denote the similarity orbit of D:

$$\mathcal{S}(D) = \{ \ SDS^{-1} \ : \ S \in \mathcal{G}l_r(\mathbb{C}) \ \} \ .$$

On the other hand, $\mathcal{U}(D) = \{ \ UDU^* \ : \ U \in \mathcal{U}(r) \ \}$ denotes the unitary orbit of D. We donote by $\pi_D : \mathcal{G}l_r(\mathbb{C}) \to \mathcal{S}(D) \subseteq \mathcal{M}_r(\mathbb{C})$ the C^∞ map defined by $\pi_D(S) = SDS^{-1}$. With the same name we note its resrtiction to the unitary group: $\pi_D : \mathcal{U}(r) \to \mathcal{U}(D)$. ▲

Proposition 2.2.2. *The similarity orbit $\mathcal{S}(D)$ is a C^∞ submanifold of $\mathcal{M}_r(\mathbb{C})$, and the projection $\pi_D : \mathcal{G}l_r(\mathbb{C}) \to \mathcal{S}(D)$ becomes a submersion. Moreover, $\mathcal{U}(D)$ is a compact submanifold of $\mathcal{S}(D)$, which consists of the normal elements of $\mathcal{S}(D)$, and $\pi_D : \mathcal{U}(r) \to \mathcal{U}(D)$ is a submersion.* ■

For every $N = UDU^* \in \mathcal{U}(D)$, it is well known (and easy to see) that

$$T_N \mathcal{S}(D) = T_I(\pi_N)(\mathcal{M}_r(\mathbb{C})) = \{[A, N] = AN - NA : \ A \in \mathcal{M}_r(\mathbb{C})\}.$$

In particular

$$\begin{aligned} T_D \mathcal{S}(D) &= \{AD - DA : \ A \in \mathcal{M}_r(\mathbb{C})\} \\ &= \{X \in \mathcal{M}_r(\mathbb{C}) : \ X_{ij} = 0 \text{ for every } (i,j) \text{ such that } d_i = d_j\}. \end{aligned} \tag{4}$$

Note that,

$$\begin{aligned} T_N \mathcal{S}(D) &= \{[A, N] = AN - NA : \ A \in \mathcal{M}_r(\mathbb{C})\} \\ &= \{(UBU^*)UDU^* - UDU^*(UBU^*) : \ B \in \mathcal{M}_r(\mathbb{C})\} \\ &= \{U[B, D]U^* = BD - DB : \ B \in \mathcal{M}_r(\mathbb{C})\} = U\Big(T_D \mathcal{S}(D)\Big)U^* . \end{aligned} \tag{5}$$

On the other hand, since $T_I \mathcal{U}(r) = \mathcal{M}_r^{ah}(\mathbb{C}) = \{A \in \mathcal{M}_r(\mathbb{C}) : A^* = -A\}$, we obtain

$$\begin{aligned} &T_D \mathcal{U}(D) = T_I(\pi_D)(\mathcal{M}_r^{ah}(\mathbb{C})) = \{[A, D] = AD - DA : \ A \in \mathcal{M}_r^{ah}(\mathbb{C})\} \quad \text{and} \quad , \\ &T_N \mathcal{U}(D) = \{[A, N] = AN - NA : \ A \in \mathcal{M}_r^{ah}(\mathbb{C})\} = U\Big(T_D \mathcal{U}(D)\Big)U^* . \end{aligned} \tag{6}$$

Finally, along this paper we shall consider on $\mathcal{S}(D)$ (and in $\mathcal{U}(D)$) the Riemannian structure inherited from $\mathcal{M}_r(\mathbb{C})$ (using the usual inner product on their tangent spaces). For $S, T \in \mathcal{S}(D)$, we denote by $\operatorname{dist}(S, T)$ the Riemannian distance between S and T (in $\mathcal{S}(D)$). Observe that, for every $U \in \mathcal{U}(r)$, one has that $U\mathcal{S}(D)U^* = \mathcal{S}(D)$ and the map $T \mapsto UTU^*$ is isometric, on $\mathcal{S}(D)$, with respect to the Riemannian metric as well as with respect to the $\|\cdot\|_2$ metric of $\mathcal{M}_r(\mathbb{C})$.

2.3 Definition and basic facts about Aluthge transforms

Definition 2.3.1. Let $T \in \mathcal{M}_r(\mathbb{C})$, and suppose that $T = U|T|$ is the polar decomposition of T. Then, we define the Aluthge transform of T in the following way:

$$\Delta(T) = |T|^{1/2} U |T|^{1/2}$$

On the other hand, $\Delta^n(T)$ denotes the n-times iterated Aluthge transform of T, i.e.

$$\Delta^0(T) = T; \quad \text{and} \quad \Delta^n(T) = \Delta\left(\Delta^{n-1}(T)\right) \quad n \in \mathbb{N}.$$

The following proposition contains some properties of Aluthge transforms which follows easily from its definition.

Proposition 2.3.2. *Let $T \in \mathcal{M}_r(\mathbb{C})$. Then:*

1. $\Delta(cT) = c\Delta(T)$ *for every* $c \in \mathbb{C}$.

2. $\Delta(VTV^*) = V\Delta(T)V^*$ *for every* $V \in \mathcal{U}(r)$.

3. *If* $T = T_1 \oplus T_2$ *then* $\Delta(T) = \Delta(T_1) \oplus \Delta(T_2)$.

4. $\|\Delta(T)\|_2 \leqslant \|T\|_2$.

5. T *and* $\Delta(T)$ *have the same characteristic polynomial, in particular,* $\sigma(\Delta(T)) = \sigma(T)$.

The following theorem states the regularity properties of Aluthge transforms (see [6]).

Theorem 2.3.3. *The Aluthge transform is* $(\|\cdot\|_2\,,\ \|\cdot\|_2)$*-continuous in* $\mathcal{M}_r(\mathbb{C})$ *and it is of class* C^∞ *in* $\mathcal{G}l_r(\mathbb{C})$.

Now, we recall a result proved independently by Jung, Ko and Pearcy in [9], and by Ando in [2].

Proposition 2.3.4. *If* $T \in \mathcal{M}_r(\mathbb{C})$*, the limit points of the sequence* $\{\Delta^n(T)\}_{n\in\mathbb{N}}$ *are normal. Moreover, if* L *is a limit point, then* $\sigma(L) = \sigma(T)$ *with the same algebraic multiplicity.*

Finally, we mention a result concerning the Jordan structure of Aluthge transforms proved in [4]. We need the following definitions.

Definition 2.3.5. Let $T \in \mathcal{M}_r(\mathbb{C})$ and $\mu \in \sigma(T)$. We denote

1. $m(T,\mu)$ the *algebraic multiplicity* of the eigenvalue μ for T.

2. $m_0(T,\mu) = \dim\ker(T-\mu I)$, the *geometric multiplicity* of μ.

Proposition 2.3.6. *Let* $T \in \mathcal{M}_r(\mathbb{C})$.

1. *If* $0 \in \sigma(T)$*, then, there exists* $n \in \mathbb{N}$ *such that*

$$m(T,0) = m_0(\Delta^n(T), 0) = \dim\ker(\Delta^n(T)).$$

2. *For every* $\mu \in \sigma(T)$, $m_0(T,\mu) \leqslant m_0(\Delta(T),\mu)$.

Observe that this implies that, if T is diagonalizable (i.e. $m_0(T,\mu) = m(T,\mu)$ for every μ), then also $\Delta(T)$ is diagonalizable.

3 The iterated Aluthge transform

3.1 Convergence of iterated Aluthge transform sequence for diagonalizable matrices

In this section, we prove the convergence of iterated Aluthge transforms for diagonalizable matrices. The key tool, which allows to use the stable manifold theorem 2.1.4, is the following theorem, whose proof is rather long and technical. For this reason, we postpone it until section 4, and we continue in this section with its consequences.

Theorem 3.1.1. *Let $D = \operatorname{diag}(d_1, \ldots, d_r) \in \mathcal{M}_r(\mathbb{C})$ be an invertible diagonal matrix. The Aluthge transform $\Delta(\cdot) : \mathcal{S}(D) \to \mathcal{S}(D)$ is a C^∞ map. For every $N \in \mathcal{U}(D)$, there exists a subspace $\mathcal{E}^s_N$ of the tangent space $T_N\mathcal{S}(D)$ such that*

1. *$T_N\mathcal{S}(D) = \mathcal{E}^s_N \oplus T_N\mathcal{U}(D)$;*
2. *Both, $\mathcal{E}^s_N$ and $T_N\mathcal{U}(D)$, are $T\Delta$-invariant;*
3. *$\left\| T\Delta|_{\mathcal{E}^s_N} \right\| \le k_D < 1$, where $k_D = \max\limits_{i,j:\ d_i \neq d_j} \dfrac{|1 + e^{i(\arg(d_j) - \arg(d_i))}|\,|d_i|^{1/2}|d_j|^{1/2}}{|d_i| + |d_j|}$;*
4. *If $U \in \mathcal{U}(r)$ satisfies $N = UDU^*$, then $\mathcal{E}^s_N = U(\mathcal{E}^s_D)U^*$.*

In particular, the map $\mathcal{U}(D) \ni N \mapsto \mathcal{E}^s_N$ is smooth. This fact can be formulated in terms of the projections P_N onto $\mathcal{E}^s_N$ parallel to $T_N\mathcal{U}(D)$, $N \in \mathcal{U}(D)$. ■

Corollary 3.1.2. *Let $D = \operatorname{diag}(d_1, \ldots, d_r) \in \mathcal{M}_r(\mathbb{C})$ be an invertible diagonal matrix. Let $\mathcal{E}^s_N$ and k_D as in Theorem 3.1.1. Then, in $\mathcal{S}(D)$ there exists a Δ-invariant C^∞-prelamination $\{\mathcal{W}_N\}_{N \in \mathcal{U}(D)}$ of class C^∞ such that, for every $N \in \mathcal{U}(D)$,*

1. *$\mathcal{W}_N$ is a C^∞ submanifold of $\mathcal{S}(D)$.*
2. *$T_N\mathcal{W}_N = \mathcal{E}^s_N$.*
3. *If $k_D < \rho < 1$, then $\operatorname{dist}(\Delta^n(T) - N) \le \operatorname{dist}(T, N)\rho^n$, for every $T \in \mathcal{W}_N$.*
4. *If $N_1 \neq N_2$ then $\mathcal{W}_{N_1} \cap \mathcal{W}_{N_2} = \varnothing$.*
5. *There exists an open subset $\mathcal{W}(D)$ of $\mathcal{S}(D)$ such that*
   a. *$\mathcal{U}(D) \subseteq \mathcal{W}(D) \subseteq \bigcup\limits_{N \in \mathcal{U}(D)} \mathcal{W}_N$, and*
   b. *The projection $p : \mathcal{W}(D) \to \mathcal{U}(D)$, defined by $p(T) = N$ if $T \in \mathcal{W}_N$, is of class C^∞.*

Proof. By Theorem 3.1.1, for every $k_D < \rho < 1$, $\mathcal{U}(D)$ is ρ-pseudo hyperbolic for Δ (see Definition 2.1.3), and it consists of fixed points. Thus, by Corollary 2.1.5 and Remark 2.1.6, we get a C^∞ and Δ-invariant prelamination of class C^∞, $\{\mathcal{W}_N\}_{N \in \mathcal{U}(D)}$ which satisfies all the properties of our statement. ■

In order to prove the convergence of iterated Aluthge transforms for diagonalizable matrices, we first reduce the problem to the invertible case. In [4] it was proved that if the sequence of iterated Aluthge transforms converge for every invertible matrix, then it converge for every matrix. In our case, we need to prove that if the sequence of iterated Aluthge transforms converge for every diagonalizable invertible matrix, then it does for every diagonalizable matrix. The proof of the second statement is essentially the same as the previous one, but, for a sake of completeness, we include its proof.

Lemma 3.1.3. *If the sequence $\{\Delta^m(S)\}_{m\in\mathbb{N}}$ converges for every diagonalizable invertible matrix $S \in \mathcal{M}_r(\mathbb{C})$ and every $r \in \mathbb{N}$, then the sequence $\{\Delta^m(T)\}_{m\in\mathbb{N}}$ converges for every diagonalizable matrices $T \in \mathcal{M}_r(\mathbb{C})$ and every $r \in \mathbb{N}$.*

Proof. Let $T \in \mathcal{M}_r(\mathbb{C})$. As we have observed after Proposition 2.3.6, if T is diagonalizable, then $\Delta(T)$ is also diagonalizable. So, if we begin with a diagonalizable matrix T, then every element of the sequence $\{\Delta^m(T)\}_{m\in\mathbb{N}}$ is diagonalizable. By Proposition 2.3.6, we can also assume that $m(T,0) = m_0(T,0)$. Note that, in this case, $\ker(\Delta(T)) = \ker(T)$ because $\ker(T) \subseteq \ker(\Delta(T))$ and $m(\Delta(T), 0) = m(T,0)$. On the other hand, $R(\Delta(T)) \subseteq R(|T|)$ so that $R(\Delta(T))$ and $\ker(\Delta(T))$ are orthogonal subspaces. Thus, there exists a unitary matrix U such that

$$U\Delta(T)U^* = \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix}$$

where $S \in M_s(\mathbb{C})$ is invertible and diagonalizable ($s = n - m(T,0)$). Since for every $m \geqslant 2$

$$\Delta^m(T) = U^* \begin{pmatrix} \Delta^{m-1}(S) & 0 \\ 0 & 0 \end{pmatrix} U \ ,$$

the sequence $\{\Delta^m(T)\}$ converges, because the sequence $\{\Delta^{m-1}(S)\}$ converges by hypothesis. ■

Theorem 3.1.4. *Let $T \in \mathcal{M}_r(\mathbb{C})$ be a* **diagonalizable** *matrix. Then $\{\Delta^n(T)\}_{n\in\mathbb{N}}$ converges.*

Proof. Using Lemma 3.1.3, we can assume that T is invertible. Then, $T \in \mathcal{S}(D)$ for some invertible diagonal matrix D. By Corollary 3.1.2 and Remark 2.1.6, we get on $\mathcal{S}(D)$ a C^∞ and Δ-invariant prelamination of class C^∞, denoted by $\{\mathcal{W}_N\}_{N\in\mathcal{U}(D)}$, such that

1. The set $\bigcup_{N\in\mathcal{U}(D)} \mathcal{W}_N$ contains an open neighborhood $\mathcal{W}(D)$ of $\mathcal{U}(D)$ in $\mathcal{S}(D)$.
2. If $k_D < \rho < 1$, then $\|\Delta^n(A) - N\|_2 \le \operatorname{dist}(\Delta^n(A) - N) \le \operatorname{dist}(A, N)\rho^n$, for every $A \in \mathcal{W}_N$.

On the other hand, by Proposition 2.3.4, there exists $m \in \mathbb{N}$ such that $A = \Delta^m(T) \in \bigcup_{N\in\mathcal{U}(D)} \mathcal{W}_N$. Thus, for $n > m$, $\Delta^n(T) = \Delta^{n-m}(A) \xrightarrow[n\to\infty]{} N$, where $N \in \mathcal{U}(D)$ is the unique element of $\mathcal{U}(D)$ such that $A \in \mathcal{W}_N$. ■

Remark 3.1.5. From Theorem 3.1.4 it can be deduced Ando and Yamazaki's result on the convergence of the iterated Aluthge sequence for 2×2 matrices. Indeed, in $\mathcal{M}_2(\mathbb{C})$, the spectrum of matrices uncovered by Theorem 3.1.4 must be a singleton. Therefore, by Proposition 2.3.4, the iterated Aluthge sequence for those matrices has only one limit point. So, it converges. ▲

Proposition 3.1.6. *Let $D \in \mathcal{M}_r(\mathbb{C})$ be diagonal and invertible. Then the sequence $\{\Delta^n\}_{n\in\mathbb{N}}$, resticted to the similarity orbit $\mathcal{S}(D)$, converges uniformly on compact sets to a C^∞ limit function $\Delta^\infty : \mathcal{S}(D) \to \mathcal{U}(D)$. In particular, Δ^∞ is a C^∞ retraction from $\mathcal{S}(D)$ onto $\mathcal{U}(D)$.*

Proof. Let Δ^{∞} be the limit function, which exists by Theorem 3.1.4. We can apply Corollary 3.1.2, and we shall use its notations. Fix $T \in \mathcal{S}(D)$. By Proposition 2.3.4 there exists $k \in \mathbb{N}$ such that $\Delta^k(T) \in \mathcal{W}(D)$. By the continuity of $\Delta(\cdot)$, there exists a neighborhood $\mathcal{U}$ of T such that $\Delta^k(\mathcal{U}) \subseteq \mathcal{W}(D)$. Hence, if p is the projection defined in Corollary 3.1.2, $\Delta^{\infty}|_{\mathcal{U}} = (p \circ \Delta^k)\big|_{\mathcal{U}}$, which proves that the map Δ^{∞} is C^{∞} at T.

On the other hand, to prove that the convergence of $\{\Delta^n(\cdot)\}_{n\in\mathbb{N}}$ is uniform on compact sets, suppose that $\mathcal{U}$ has compact closure, and denote by

$$C = \sup\{\operatorname{dist}(\Delta^k(S), \Delta^{\infty}(S)) : S \in \mathcal{U}\} \ .$$

Fix $\varepsilon > 0$ and take $m_0 > k$ such that $Ck_D^{m_0-k} < \varepsilon$. Then, using (4) of Corollary 3.1.2, for every $m \geq m_0$ and every $S \in \mathcal{U}$

$$\operatorname{dist}(\Delta^m(S) - \Delta^{\infty}(S)) = \operatorname{dist}\left(\Delta^{m-k}\left(\Delta^k(S)\right) - \Delta^{\infty}\left(\Delta^k(S)\right)\right) \leq \varepsilon.$$

This proves that for every $T \in \mathcal{S}(D)$ there exists a neighborhood of T where the convergence is uniform. Therefore, by standard arguments, it follows that the convergence is uniform on compact sets. ■

Remark 3.1.7. Let $D \in \mathcal{M}_r(\mathbb{C})$ be diagonal but not invertible. If $T \in \mathcal{S}(D)$, by arguments similar to those used in the proofs of Lemma 3.1.3 and Proposition 3.1.6 it can be proved that $\Delta(T) \in \mathcal{S}(D)$, and the map $\Delta^{\infty}\big|_{\mathcal{S}(D)} : \mathcal{S}(D) \to \mathcal{U}(D)$ is a retraction of calss C^{∞}.

3.2 Smoothness of the map $T \mapsto \Delta^{\infty}(T)$ on $\mathcal{D}_r^*(\mathbb{C})$

Let $\mathcal{D}_r^*(\mathbb{C})$ be the set of diagonalizable and invertible matrices in $\mathcal{M}_r(\mathbb{C})$ with r different eigenvalues (i.e. every eigenvalue has algebraic multiplicity equal to one). Observe that $\mathcal{D}_r^*(\mathbb{C})$ is an open dense subset of $\mathcal{M}_r(\mathbb{C})$ and it is invariant by the Aluthge transform. If $\Delta^{\infty}(\cdot)$ denotes the limit of the sequence of iterated Aluthge transforms, which is defined on the set of diagonalizable matrices by Theorem 3.1.4, we shall show that $T \mapsto \Delta^{\infty}(T)$ is of class C^{∞} on $\mathcal{D}_r^*(\mathbb{C})$. The proof of this result essentially follows the same lines as Proposition 3.1.6. For this reason, we expose a sketched version of the proof, where we only point out the main differences.

We already know that the map $\Delta^{\infty}(\cdot)$ is of class C^{∞} if it is restricted to the orbits $\mathcal{S}(T)$ for any $T \in \mathcal{D}_r^*(\mathbb{C})$. In order to study the behavior of this map outside the orbit of T, we need to define the following sets: let $D \in \mathcal{D}_r^*(\mathbb{C})$ be a diagonal matrix and let $\varepsilon > 0$; then

$$\mathcal{B}(D, \varepsilon) = \Big\{D' \in \mathcal{D}_r^*(\mathbb{C}) : \ D' \text{ is diagonal and } \|D - D'\|_2 < \varepsilon\Big\};$$

$$\mathcal{S}(D, \varepsilon) = \Big\{SD'S^{-1} : \ D' \in \mathcal{B}(D, \varepsilon) \text{ and } S \in \mathcal{G}l_r(\mathbb{C})\Big\} = \bigcup_{D' \in \mathcal{B}(D, \varepsilon)} \mathcal{S}(D');$$

$$\mathcal{U}(D, \varepsilon) = \Big\{UD'U^* : \ D' \in \mathcal{B}(D, \varepsilon) \text{ and } U \in \mathcal{U}(r)\Big\} = \bigcup_{D' \in \mathcal{B}(D, \varepsilon)} \mathcal{U}(D') \ .$$

The set $\mathcal{S}(D, \varepsilon)$ is invariant for $\Delta(\cdot)$ and it is also open in $\mathcal{G}l_r(\mathbb{C})$ for ε small enough. Since $D \in \mathcal{D}_r^*(\mathbb{C})$, it can be proved that $\mathcal{U}(D, \varepsilon)$ is a smooth submanifold of $\mathcal{M}_r(\mathbb{C})$, and it consists on the fixed points of $\mathcal{S}(D, \varepsilon)$. For each $N \in \mathcal{U}(D, \varepsilon)$, if $\{N\}'$ denotes the subspace $\{A \in \mathcal{M}_r(\mathbb{C}) : AN = NA\}$, the tangent space $T_N\mathcal{U}(D, \varepsilon)$ can be decomposed as $T_N\mathcal{U}(D, \varepsilon) = T_N\mathcal{U}(D) \oplus \{N\}'$. Then, $T_N\mathcal{S}(D, \varepsilon) = \mathcal{M}_r(\mathbb{C})$ can be decomposed as

$$T_N\mathcal{S}(D, \varepsilon) = T_N\mathcal{S}(D) \oplus \{N\}' = \Big(\mathcal{E}_N^s \oplus T_N\mathcal{U}(D)\Big) \oplus \{N\}' = \mathcal{E}_N^s \oplus T_N\mathcal{U}(D, \varepsilon) \ , \quad (7)$$

where the subspaces $\mathcal{E}_N^s$ are the same as those constructed in Theorem 3.1.1. Since $D \in \mathcal{D}_r^*(\mathbb{C})$ then, with the notations of Theorem 3.1.1, $\rho = \max\limits_{D' \in \mathcal{B}(D, \varepsilon)} k_{D'} < 1$ for ε small enought. Also, for every $N \in \mathcal{U}(D, \varepsilon)$,

1. Both $\mathcal{E}_N^s$ and $T_N\mathcal{U}(D, \varepsilon)$, are $T_N\,\Delta$-invariant;
2. $\left\| T_N\,\Delta|_{\mathcal{E}_N^s} \right\| \leq \rho < 1$, and $T_N\,\Delta\Big|_{T_N\mathcal{U}(D,\varepsilon)}$ is the identity map of $T_N\mathcal{U}(D, \varepsilon)$.

The distribution of the subspaces $\mathcal{E}_N^s$ is still smooth, since the (oblique) projection E_N onto $\mathcal{E}_N^s$ parallel to $T_N\mathcal{U}(D, \varepsilon)$ moves smoothly on $\mathcal{U}(D, \varepsilon)$. A brief justification of these facts can be found in the following Remark:

Remark 3.2.1. Let $d = \frac{1-\rho}{3}$. Consider the open discs $\mathcal{U} = \{z \in \mathbb{C} : |z| < \rho + d\}$ and $\mathcal{V} = \{z \in \mathbb{C} : |1 - z| < d\}$, which have disjoint closures. By Eq. (7), and items 1 and 2 of the previous discusion, one can deduce that the spectrum of $T_N\,\Delta$ is contained in $\mathcal{U} \cup \mathcal{V}$ for every $N \in \mathcal{U}(D, \varepsilon)$. Moreover, if $f : \mathcal{U} \cup \mathcal{V} \to \mathbb{C}$ is the holomorphic map $f = \aleph_{\mathcal{U}}$ (the charcateristic map of $\mathcal{U}$), then $E_N = f(T_N\,\Delta)$ for every $N \in \mathcal{U}(D, \varepsilon)$. If $\mathcal{M}(\mathcal{U} \cup \mathcal{V}) = \{T \in \mathcal{M}_{r^2}(\mathbb{C}) : \sigma(T) \subseteq \mathcal{U} \cup \mathcal{V}\}$, which is an open subset of $\mathcal{M}_{r^2}(\mathbb{C})$, then the map

$$\mathcal{M}(\mathcal{U} \cup \mathcal{V}) \ni T \mapsto f(T) \quad \text{is of class } C^\infty$$

(see Theorem 5.16 of Kato's book [10]). Therefore, the distribution $\mathcal{U}(D, \varepsilon) \ni N \mapsto E_N = f(T_N\,\Delta)$ is of class C^∞. A similar type of argument can be used to show that $\mathcal{U}(D, \varepsilon)$ is a smooth submanifold of $\mathcal{M}_r(\mathbb{C})$, for ε small enough. ▲

Proposition 3.2.2. *The map $\Delta^\infty(\cdot)$ is of class C^∞ on $\mathcal{D}_r^*(\mathbb{C})$, and the sequence $\{\Delta^n(\cdot)\}_{n\in\mathbb{N}}$, resticted to $\mathcal{D}_r^*(\mathbb{C})$, converges uniformly on compact sets to $\Delta^\infty(\cdot)$.*

Proof. Let $T \in \mathcal{D}_r^*(\mathbb{C})$, denote $N = \Delta^\infty(T)$ and let $D \in \mathcal{D}_r^*(\mathbb{C})$, a diagonal matrix such that $N \in \mathcal{U}(D)$. We can apply Theorem 2.1.4 to the pair $\mathcal{U}(D, \varepsilon) \subseteq \mathcal{S}(D, \varepsilon)$, for ε small. From now on, the proof follows the same steps as the proofs of Corollary 3.1.2 and Proposition 3.1.6. ■

4 Proof of Theorem 3.1.1

4.1 Matricial characterization of $T_N\Delta$

Throughout this section we fix an invertible diagonal matrix $D \in \mathcal{M}_r(\mathbb{C})$ whose diagonal entries are denoted by $(d_1, \dots, d_n)$. For every $j \in \{1, \dots, n\}$, let $d_j = e^{i\theta_j}|d_j|$ be the polar decomposition of d_j, where $\theta_j \in [0, 2\pi]$. Recall from Eq. (4) that the tangent space $T_D\mathcal{S}\,(D)$ consists on those matrices $X \in \mathcal{M}_r(\mathbb{C})$ such that $X_{ij} = 0$ if $d_i = d_j$.

Definition 4.1.1. Given $A, B \in \mathcal{M}_r(\mathbb{C})$, $A \circ B$ denotes their Hadamard product, that is, if $A = (A_{ij})$ and $B = (B_{ij})$, then $(A \circ B)_{ij} = A_{ij}B_{ij}$. With respect to this product, each matrix $A \in \mathcal{M}_r(\mathbb{C})$ induces an operator Ψ_A on $\mathcal{M}_r(\mathbb{C})$ defined by $\Psi_A(B) = A \circ B$, $B \in \mathcal{M}_r(\mathbb{C})$.

Remark 4.1.2. Note that, by Eq. (4), the subspace $T_D\mathcal{S}\,(D)$ reduces the operator Ψ_A , for every $A \in \mathcal{M}_r(\mathbb{C})$. This is the reason why, from now on, we shall consider all these operators as acting on $T_D\mathcal{S}\,(D)$. Restricted in this way, it holds that

$$\|\Psi_A\| = \sup\{\|A \circ B\|_2 : \ B \in T_D\mathcal{S}\,(D) \text{ and } \|B\|_2 = 1\} = \max_{d_i \neq d_j} |A_{ij}| \ .$$

▲

Let $P_{\mathbb{R}e}$ and $P_{\mathbb{I}m}$ be the projections defined on $T_D\mathcal{S}\,(D)$ by

$$P_{\mathbb{R}e}(B) = \frac{B + B^*}{2} \qquad \text{and} \qquad P_{\mathbb{I}m}(B) = \frac{B - B^*}{2}.$$

That is, $P_{\mathbb{R}e}$ (resp. $P_{\mathbb{I}m}$) is the restriction to $T_D\mathcal{S}\,(D)$ of the orthogonal projection onto the subspace of hermitian (resp. anti-hermitian) matrices. Observe that, for every $K \in \mathcal{M}_r^{ah}(\mathbb{C})$ (i.e., such that $K^* = -K$) and $B \in \mathcal{M}_r(\mathbb{C})$ it holds that

$$K \circ P_{\mathbb{R}e}(B) = P_{\mathbb{I}m}(K \circ B) \quad \text{and} \quad K \circ P_{\mathbb{I}m}(B) = P_{\mathbb{R}e}(K \circ B)\ . \tag{8}$$

Denote by Q_D the orthogonal projection from $T_D\mathcal{S}\,(D)$ onto $(T_D\mathcal{U}\,(D))^\perp$.

Lemma 4.1.3. *Let $J, K \in \mathcal{M}_r(\mathbb{C})$ be the matrices defined by*

$$K_{ij} = \begin{cases} |d_j - d_i|\,\mathrm{sgn}(j-i) & \text{if } d_i \neq d_j \\ 0 & \text{if } d_i = d_j \end{cases} \quad \text{and} \quad J_{ij} = \begin{cases} (d_j - d_i)K_{ij}^{-1} & \text{if } d_i \neq d_j \\ 1 & \text{if } d_i = d_j \end{cases},$$

for $1 \le i, j \le r$. Then

1. *For every $A \in \mathcal{M}_r(\mathbb{C})$, $AD - DA = J \circ K \circ A$.*

2. *It holds that $Q_D = \Psi_J P_{\mathbb{I}m} \Psi_J^{-1}$.*

3. *If $H \in \mathcal{M}_r^h(\mathbb{C})$ (i.e., if $H^* = H$), then $Q_D\Psi_H = \Psi_H Q_D$.*

Proof.

1. It is enough to note that $(J \circ K)_{ij} = d_j - d_i$ and $(AD - DA)_{ij} = (d_j - d_i)A_{ij}$.

2. Since $|J_{ij}| = 1$ for every $1 \leq i,j \leq r$, the operator Ψ_J is unitary in $(\mathcal{M}_r(\mathbb{C}), \|\cdot\|_2)$. Hence, $\Psi_J P_{\mathbb{Im}} \Psi_J^{-1}$ is an orthogonal projection. Recall that

$$T_D\mathcal{U}(D) = \{AD - DA : \ A \in \mathcal{M}_r^{ah}(\mathbb{C})\}.$$

By Eq. (8), $P_{\mathbb{Im}}\Psi_K = \Psi_K P_{\mathbb{Re}}$. Then, given $X = AD - DA \in T_D\mathcal{U}(D)$,

$$\Psi_J P_{\mathbb{Im}} \Psi_J^{-1}(X) = \Psi_J P_{\mathbb{Im}} \Psi_J^{-1}(\Psi_J \Psi_K A) = \Psi_J P_{\mathbb{Im}} \Psi_K(A) = \Psi_J \Psi_K P_{\mathbb{Re}}(A) = 0.$$

So, $T_D\mathcal{U}(D) \subseteq \ker(\Psi_J P_{\mathbb{Im}} \Psi_J^{-1})$. But, $\dim T_D\mathcal{U}(D) = \dim\ker(\Psi_J P_{\mathbb{Im}} \Psi_J^{-1})$. Therefore, we have that $Q_D = \Psi_J P_{\mathbb{Im}} \Psi_J^{-1}$.

3. It is clear that $\Psi_H\Psi_J = \Psi_J\Psi_H$. On the other hand, since H is hermitian, Ψ_H also commutes with the projection $P_{\mathbb{Im}}$. ∎

Remark 4.1.4. Let $N \in \mathcal{U}(D)$ and let Q_N be the orthogonal projection from $T_N\mathcal{S}(D)$ onto $\big(T_N\mathcal{U}(D)\big)^\perp$. Then $T_N\Delta$ has the following 2×2 matrix decomposition

$$T_N\Delta = \begin{pmatrix} A_{1N} & 0 \\ A_{2N} & I \end{pmatrix} \begin{matrix} Q_N \\ I - Q_N \end{matrix} \ , \tag{9}$$

because $T_N\Delta$ acts as the identity on $T_N\mathcal{U}(D)$. The next Proposition gives a characterization of the significative parts $A_{1N} = Q_N\big(T_N\Delta\big)Q_N$ and $A_{2N} = (I - Q_N)\big(T_N\Delta\big)Q_N$ in the case $N = D$. ▲

Proposition 4.1.5. *Let Q_D be the orthogonal projection onto $\big(T_D\mathcal{U}(D)\big)^\perp$. Then there exists $H \in \mathcal{M}_r(\mathbb{C})$ such that, if $H_1 = P_{\mathbb{Re}}(H)$ and $H_2 = P_{\mathbb{Im}}(H)$,*

$$Q_D\big(T_D\Delta\big)Q_D = Q_D \ \Psi_{H_1} \ Q_D \quad \text{and} \quad (I - Q_D)\big(T_D\Delta\big)Q_D = (I - Q_D) \ \Psi_{H_2} \ Q_D \,.$$

The matrix H_1 can be characterized as

$$(H_1)_{ij} = \frac{\big(1 + e^{\,i(\theta_j - \theta_i)}\big)|d_i|^{1/2}|d_j|^{1/2}}{|d_i| + |d_j|} \quad \text{for every} \quad 1 \leq i,j \leq r \ . \tag{10}$$

Proof. Fix a tangent vector $X = AD - DA \in T_D\mathcal{S}(D)$, for some $A \in \mathcal{M}_r(\mathbb{C})$. Then

$$T_D\Delta(X) = \frac{d}{dt}\Delta\left(e^{tA}De^{-tA}\right)\Big|_{t=0} \ .$$

Let $\gamma(t) = \big(e^{tA}De^{-tA}\big)^*\big(e^{tA}De^{-tA}\big) = e^{-tA^*}D^*e^{tA^*}e^{tA}De^{-tA}$. In terms of γ, we can write the curve $\Delta\left(e^{tA}De^{-tA}\right)$ in the following way

$$\Delta\left(e^{tA}De^{-tA}\right) = \gamma^{1/4}(t)(e^{tA}De^{-tA})\gamma^{-1/4}(t).$$

So, using that $(\gamma^{-1/4})'(0) = -\gamma^{-1/4}(0)\ (\gamma^{1/4})'(0)\ \gamma^{-1/4}(0)$ (which can be deduce from the identity $\gamma^{1/4}\gamma^{-1/4} = I$), we obtain

$$\begin{aligned}
T_D\Delta\,(X) &= (\gamma^{1/4})'(0)\ D\gamma^{-1/4}(0) + \gamma^{1/4}(0)(AD-DA)\gamma^{-1/4}(0)\\
&- \gamma^{1/4}(0)\ D\ \gamma^{-1/4}(0)\ (\gamma^{1/4})'(0)\ \gamma^{-1/4}(0)\\
&= (\gamma^{1/4})'(0)\ D|D|^{-1/2} + |D|^{1/2}(AD-DA)|D|^{-1/2}\\
&- |D|^{1/2}\ D\ |D|^{-1/2}\ (\gamma^{1/4})'(0)\ |D|^{-1/2}\\
&= \Big((\gamma^{1/4})'(0)\ D - D\ (\gamma^{1/4})'(0)\Big)|D|^{-1/2} + |D|^{1/2}(AD-DA)|D|^{-1/2}.
\end{aligned}$$

If we define the matrices $L, N \in \mathcal{M}_r(\mathbb{C})$ by

$$\begin{aligned}
N_{ij} &= |d_j|^{-1/2},\\
L_{ij} &= |d_i|^{1/2}|d_j|^{-1/2},
\end{aligned}$$

and take $J, K \in \mathcal{M}_r(\mathbb{C})$ as in Lemma 4.1.3. Then

$$T_D\Delta\,(X) = N \circ (J \circ K \circ (\gamma^{1/4})'(0)) + L \circ (J \circ K \circ A).$$

Now, we need to compute $(\gamma^{1/4})'(0)$. Firstly, we shall compute $(\gamma^{1/2})'(0)$, and then we shall repeat the procedure to get $(\gamma^{1/4})'(0)$. Using the identity $\gamma^{1/2}\gamma^{1/2} = \gamma$, we get

$$\gamma^{1/2}(\gamma^{1/2})' + (\gamma^{1/2})'\gamma^{1/2} = \gamma'$$

If $A = \gamma^{1/2}(0)$, $B = -\gamma^{1/2}(0)$ and $Y = \gamma'(0)$, we can rewrite the above identity in the following way

$$A(\gamma^{1/2})'(0) - (\gamma^{1/2})'(0)B = Y.$$

Therefore, $(\gamma^{1/2})'$ is the solution of Sylvester's equation $AX - XB = Y$. Using the well known formula for this solution (see [5, Thm. VII.2.3]), it holds that

$$(\gamma^{1/2})'(0) = \int_0^\infty e^{-tA} Y e^{tB}\ dt = \int_0^\infty e^{-t\gamma^{1/2}(0)}\ \gamma'(0)\ e^{-t\gamma^{1/2}(0)}\ dt.$$

In the same way, we get

$$\begin{aligned}
(\gamma^{1/4})'(0) &= \int_0^\infty e^{-t\gamma^{1/4}(0)}\ (\gamma^{1/2})'(0)\ e^{-t\gamma^{1/4}(0)}\ dt\\
&= \int_0^\infty e^{-t\gamma^{1/4}(0)} \left(\int_0^\infty e^{-s\gamma^{1/2}(0)}\ \gamma'(0)\ e^{-s\gamma^{1/2}(0)}\ ds\right) e^{-t\gamma^{1/4}(0)}\ dt\\
&= \int_0^\infty \int_0^\infty e^{-\left(t\gamma^{1/4}(0)+s\gamma^{1/2}(0)\right)}\ \gamma'(0)\ e^{-\left(t\gamma^{1/4}(0)+s\gamma^{1/2}(0)\right)}\ ds\,dt.
\end{aligned}$$

Finally, as $\gamma(0) = |D|^2$, we obtain

$$(\gamma^{1/2})'(0) = \int_0^\infty \int_0^\infty e^{-(t|D|^{1/2}+s|D|)}\ \gamma'(0)\ e^{-(t|D|^{1/2}+s|D|)}\ ds\, dt.$$

So, if $M \in \mathcal{M}_r(\mathbb{C})$ is the matrix defined by

$$\begin{aligned} M_{ij} &= \int_0^\infty \int_0^\infty e^{-(t|d_i|^{1/2}+s|d_i|)}\ e^{-(t|d_j|^{1/2}+s|d_j|)}\ ds\, dt \\ &= \int_0^\infty \int_0^\infty e^{-\Big(t(\ |d_i|^{1/2}+|d_j|^{1/2})\ +\ s(\ |d_i|+|d_j|)\Big)}\ ds\, dt \\ &= \int_0^\infty e^{-s\big(\ |d_i|+|d_j|\ \big)}\ ds \int_0^\infty e^{-t\big(\ |d_i|^{1/2}+|d_j|^{1/2}\big)}\ dt \\ &= \left. \frac{-e^{-s\big(\ |d_i|+|d_j|\ \big)}}{|d_i|+|d_j|} \right|_0^\infty \left. \frac{-e^{-t\big(\ |d_i|^{1/2}+|d_j|^{1/2}\big)}}{|d_i|^{1/2}+|d_j|^{1/2}} \right|_0^\infty \\ &= \frac{1}{|d_i|+|d_j|}\ \frac{1}{|d_i|^{1/2}+|d_j|^{1/2}}, \end{aligned}$$

then $(\gamma^{1/4})'(0) = M \circ \gamma'(0)$. Our next step will be to compute $\gamma'(0)$.

$$\begin{aligned} \gamma'(0) &= -A^*D^*D + D^*A^*D + D^*AD - D^*DA = 2D^*P_{\mathbb{R}e}(A)D - (D^*DA + A^*D^*D) \\ &= 2D^*P_{\mathbb{R}e}(A)D - (D^*DP_{\mathbb{R}e}(A) + P_{\mathbb{R}e}(A)D^*D) - (D^*DP_{\mathbb{I}m}(A) - P_{\mathbb{I}m}(A)D^*D) \end{aligned}$$

Let $R, T^+, T^- \in \mathcal{M}_r(\mathbb{C})$ be the matrices defined by

$$R_{ij} = 2\bar{d}_i d_j\ , \quad T^+_{ij} = |d_i|^2 + |d_j|^2\ , \quad \text{and} \quad T^-_{ij} = |d_j|^2 - |d_i|^2\ , \quad 1 \le i,j \le r\ .$$

Then, $\gamma'(0)$ can be rewritten in the following way

$$\gamma'(0) = R \circ P_{\mathbb{R}e}(A) - T^+ \circ P_{\mathbb{R}e}(A) + T^- \circ P_{\mathbb{I}m}(A).$$

In consequence, $T_D\Delta\,(AD - DA)$ can be characterized (in terms of A) as

$$T_D\Delta\,(X) = N \circ J \circ K \circ M \circ \Big[(R - T^+) \circ P_{\mathbb{R}e}(A) + T^- \circ P_{\mathbb{I}m}(A)\Big] + L \circ J \circ K \circ A.$$

Now, we shall express $T_D\Delta\,(X)$ in terms of $X = J \circ K \circ A$. Recall that, since $K^* = -K$, then $P_{\mathbb{I}m}\Psi_K = \Psi_K P_{\mathbb{R}e}$, by Eq. (8). Therefore,

$$\begin{aligned} T_D\Delta\,(X) &= M \circ N \circ (R - T^+) \circ J \circ P_{\mathbb{I}m}(K \circ A) \\ &\quad + M \circ N \circ T^- \circ J \circ P_{\mathbb{R}e}(K \circ A) + L \circ (J \circ K \circ A) \\ &= M \circ N \circ (R - T^+) \circ (\Psi_J P_{\mathbb{I}m} \Psi_J^{-1})(X) \\ &\quad + M \circ N \circ T^- \circ (\Psi_J P_{\mathbb{R}e} \Psi_J^{-1})(X) + L \circ (X) \end{aligned}$$

Then, using Lemma 4.1.3

$$
\begin{aligned}
T_D\Delta\,(X) &= \Big(M\circ N\circ (R-T^+)+L\Big)\circ Q_D(X)\\
&\quad + \Big(M\circ N\circ T^- + L\Big)\circ (I-Q_D)(X).
\end{aligned}
$$

If $H = M\circ N\circ (R-T^+)+L$, then $H_{i,j} =$

$$
\begin{aligned}
&= |d_i|^{1/2}|d_j|^{-1/2} + |d_j|^{-1/2}\frac{2\bar{d}_i d_j - (|d_i|^2+|d_j|^2)}{(|d_i|^{1/2}+|d_j|^{1/2})(|d_i|+|d_j|)}\\
&= \frac{|d_i|^{1/2}|d_j|^{-1/2}(|d_i|^{1/2}+|d_j|^{1/2})(|d_i|+|d_j|)\ +\ 2\bar{d}_i d_j|d_j|^{-1/2} - |d_i|^2|d_j|^{-1/2} - |d_j|^{3/2}}{(|d_i|^{1/2}+|d_j|^{1/2})(|d_i|+|d_j|)}\\
&= \frac{|d_i||d_j|^{1/2} + |d_i|^{3/2} + |d_i|^{1/2}|d_j|\ +\ 2\bar{d}_i d_j|d_j|^{-1/2} - |d_j|^{3/2}}{(|d_i|^{1/2}+|d_j|^{1/2})(|d_i|+|d_j|)}\\
&= \frac{|d_i||d_j|^{1/2} + |d_i|^{3/2} + |d_i|^{1/2}|d_j| + |d_j|^{3/2}\ +\ 2\bar{d}_i d_j|d_j|^{-1/2} - 2|d_j|^{3/2}}{(|d_i|^{1/2}+|d_j|^{1/2})(|d_i|+|d_j|)}\\
&= 1 + 2\frac{\bar{d}_i d_j|d_j|^{-1/2} - |d_j|^{3/2}}{(|d_i|^{1/2}+|d_j|^{1/2})(|d_i|+|d_j|)}.
\end{aligned}
$$

On the other hand

$$
\begin{aligned}
(M\circ N\circ T^- + L) &= |d_i|^{1/2}|d_j|^{-1/2} + |d_j|^{-1/2}\frac{|d_j|^2-|d_i|^2}{(|d_i|^{1/2}+|d_j|^{1/2})(|d_i|+|d_j|)}\\
&= |d_j|^{-1/2}\Big(|d_i|^{1/2}+|d_j|^{1/2}-|d_i|^{1/2}\Big) = 1
\end{aligned}
$$

Therefore, we get that $T_D\Delta\,(X) = \Big(\ HQ_D + (I-Q_D)\ \Big)(X)$. Given $Y\in R(Q_D)$,

$$
\begin{aligned}
Q_D\big(T_D\Delta\big)Q_D(Y) &= Q_D(H\circ Y) = (\Psi_J P_{\mathbb{Im}}\Psi_J^{-1})(H\circ Y)\\
&= J\circ\Big(P_{\mathbb{Im}}(H\circ\Psi_J^{-1}Y)\Big)\\
&= \frac{1}{2}\ J\circ\Big(H\circ\Psi_J^{-1}(Y) - \Big(H\circ\Psi_J^{-1}(Y)\Big)^*\Big)\\
&= \frac{1}{2}\ J\circ\Big(H\circ\Psi_J^{-1}(Y)\ +\ H^*\circ\Psi_J^{-1}(Y)\Big)\\
&= J\circ P_{\mathbb{Re}}(H)\circ\Psi_J^{-1}(Y) = P_{\mathbb{Re}}(H)\circ Y = Q_D\Psi_{P_{\mathbb{Re}(H)}}(Y)\ .
\end{aligned}
$$

Analogously

$$
\begin{aligned}
(I-Q_D)\big(T_D\Delta\big)Q_D(Y) &= (I-Q_D)(H\circ Y) = (\Psi_J P_{\mathbb{Re}}\Psi_J^{-1})(H\circ Y)\\
&= J\circ\Big(P_{\mathbb{Re}}(H\circ\Psi_J^{-1}Y)\Big)\\
&= \frac{1}{2}\ J\circ\Big(H\circ\Psi_J^{-1}(Y) + \Big(H\circ\Psi_J^{-1}(Y)\Big)^*\Big)\\
&= \frac{1}{2}\ J\circ\Big(H\circ\Psi_J^{-1}(Y)\ -\ H^*\circ\Psi_J^{-1}(Y)\Big)\\
&= J\circ P_{\mathbb{Im}}(H)\circ\Psi_J^{-1}(Y) = P_{\mathbb{Im}}(H)\circ Y = (I-Q_D)\Psi_{P_{\mathbb{Im}(H)}}(Y)\ .
\end{aligned}
$$

So, Eq. (10) holds. Moreover,

$$\begin{aligned}
(H_1)_{ij} &= \frac{1}{2}\left(1+2\frac{\bar{d}_i d_j |d_j|^{-1/2} - |d_j|^{3/2}}{(|d_i|^{1/2}+|d_j|^{1/2})(|d_i|+|d_j|)} + 1 + 2\frac{\bar{d}_i d_j |d_i|^{-1/2} - |d_i|^{3/2}}{(|d_i|^{1/2}+|d_j|^{1/2})(|d_i|+|d_j|)}\right) \\
&= 1 + \frac{\bar{d}_i d_j |d_j|^{-1/2} - |d_j|^{3/2} + \bar{d}_i d_j |d_i|^{-1/2} - |d_i|^{3/2}}{(|d_i|^{1/2}+|d_j|^{1/2})(|d_i|+|d_j|)} \\
&= \frac{|d_i||d_j|^{1/2} + |d_j||d_i|^{1/2} + \bar{d}_i d_j |d_j|^{-1/2} + \bar{d}_i d_j |d_i|^{-1/2}}{(|d_i|^{1/2}+|d_j|^{1/2})(|d_i|+|d_j|)} \\
&= \frac{|d_i|^{1/2}|d_j|^{1/2}\left(|d_i|^{1/2}+|d_j|^{1/2}+e^{i(\theta_j-\theta_i)}|d_i|^{1/2}+e^{i(\theta_j-\theta_i)}|d_j|^{1/2}\right)}{(|d_i|^{1/2}+|d_j|^{1/2})(|d_i|+|d_j|)} \\
&= \frac{\left(1+e^{i(\theta_j-\theta_i)}\right)|d_i|^{1/2}|d_j|^{1/2}}{|d_i|+|d_j|}\ ,
\end{aligned}$$

which completes the proof. ∎

Corollary 4.1.6. *Given $N \in \mathcal{U}(D)$, consider the matrix decomposition*

$$T_N\Delta = \begin{pmatrix} A_{1N} & 0 \\ A_{2N} & I \end{pmatrix} \begin{matrix} Q_N \\ I - Q_N \end{matrix}\ ,$$

as in Remark 4.1.4. Then $\|A_{1N}\| \leq \max_{i,j:\ d_i \neq d_j} \dfrac{|1+e^{i(\theta_j-\theta_i)}|\,|d_i|^{1/2}|d_j|^{1/2}}{|d_i|+|d_j|} < 1.$

Proof. Let $N = UDU^* \in \mathcal{U}(D)$, for some $U \in \mathcal{U}(r)$. Then,

$$T_N\Delta = Ad_U\Big(T_D\Delta\Big)Ad_U^{-1} \quad \text{and} \quad Q_N = Ad_U\Big(Q_D\Big)Ad_U^{-1}\ .$$

Since $Ad_U : T_D\mathcal{S}(D) \to T_N\mathcal{S}(D)$ is an isometric isomorphism, it holds that

$$\|A_{1N}\| = \|Q_N(T_N\Delta)Q_N\| = \left\|Ad_U\Big(Q_D(T_D\Delta)Q_D\Big)Ad_U^{-1}\right\| = \|Q_D(T_D\Delta)Q_D\| = \|A_{1D}\|.$$

Take the selfadjoint matrix H_1 given by Proposition 4.1.5. Hence,

$$\|A_{1D}\| \leq \|\Psi_{H_1}\| = \max_{i,j:\ d_i\neq d_j} \frac{|1+e^{i(\theta_j-\theta_i)}|\,|d_i|^{1/2}|d_j|^{1/2}}{|d_i|+|d_j|}\ .$$

Finally, this maximum is strictly lower than one because, by the triangle inequality and the arithmetic-geometric inequality,

$$\frac{|1+e^{i(\theta_j-\theta_i)}|\,|d_i|^{1/2}|d_j|^{1/2}}{|d_i|+|d_j|} \ \leq\ \frac{2\,|d_i|^{1/2}|d_j|^{1/2}}{|d_i|+|d_j|} \ \leq\ 1\ .$$

But the equality holds only if $\theta_j = \theta_i \mod (2\pi)$ and $|d_i| = |d_j|$, that is, if $d_i = d_j$. ∎

Remark 4.1.7. It is easy to see, using Lemma 4.1.3 and Eq. (10), that $T_D\Delta$ is invertible, and therefore Δ is a local diffeomorphism near D, if and only if $e^{i(\theta_j-\theta_i)} \neq -1$ for every i, j. This means that there are not pairs d_i, d_j such that $d_i \cdot d_j \in \mathbb{R}_{<0}$. ▲

4.2 The proof

Now we rewrite the statement of Theorem 3.1.1 and conclude its proof:

Theorem. *The Aluthge transform* $\Delta(\cdot) : \mathcal{S}(D) \to \mathcal{S}(D)$ *is a* C^∞ *map, and for every* $N \in \mathcal{U}(D)$*, there exists a subspace* $\mathcal{E}_N^s$ *in the tangent space* $T_N\mathcal{S}(D)$ *such that*

1. $T_N\mathcal{S}(D) = \mathcal{E}_N^s \oplus T_N\mathcal{U}(D)$;
2. *Both,* $\mathcal{E}_N^s$ *and* $T_N\mathcal{U}(D)$*, are* $T_N\,\Delta$*-invariant;*
3. $\left\| T_N\,\Delta|_{\mathcal{E}_N^s} \right\| \le k_D < 1$*, where* $k_D = \max\limits_{i,j:\ d_i \ne d_j} \dfrac{\left|1 + e^{i(\arg(d_j) - \arg(d_i))}\right| |d_i|^{1/2}|d_j|^{1/2}}{|d_i| + |d_j|}$;
4. *If* $U \in \mathcal{U}(r)$ *satisfies* $N = UDU^*$*, then* $\mathcal{E}_N^s = U(\mathcal{E}_D^s)U^*$.

In particular, the map $\mathcal{U}(D) \ni N \mapsto \mathcal{E}_N^s$ *is smooth. This fact can be formulated in terms of the projections* P_N *onto* $\mathcal{E}_N^s$ *parallel to* $T_N\mathcal{U}(D)$*,* $N \in \mathcal{U}(D)$*.*

Proof. Fix $N = UDU^* \in \mathcal{U}(D)$. By Corollary 4.1.6 $\|A_{1N}\| < 1$, so the operator $I - A_{1N}$ acting on $R(Q_N)$ is invertible. Let $\mathcal{E}_N^s$ be the subspace defined by

$$\mathcal{E}_N^s = \left\{ \begin{pmatrix} y \\ -A_{2N}(I - A_{1N})^{-1}y \end{pmatrix} : \ y \in R(Q_N) \right\},$$

where Q_N, as in Corollary 4.1.6, is the orthogonal projection onto $\big(T_N\mathcal{U}(D)\big)^\perp$. A straightforward computation shows that

$$P_N = \begin{pmatrix} I & 0 \\ -A_{2N}(I - A_{1N})^{-1} & 0 \end{pmatrix} \begin{matrix} Q_N \\ I - Q_N \end{matrix}$$

is a projection onto $\mathcal{E}_N^s$ parallel to $T_N\mathcal{U}(D)$. Therefore

$$T_N\mathcal{U}(D) = \mathcal{E}_N^s \oplus T_N\mathcal{U}(D).$$

Moreover, since $T_N\Delta = Ad_U\Big(T_D\Delta\Big)Ad_U^{-1}$, $Q_N = Ad_U\Big(Q_D\Big)Ad_U^{-1}$, and P_N can be written as

$$P_N = Q_N - (I - Q_N)(T_N\Delta)Q_N\big(I - Q_N(T_N\Delta)Q_N\big)^{-1}Q_N,$$

it holds that

$$P_N = Ad_U(P_D)Ad_U^{-1}.$$

This shows that $\mathcal{E}_N^s = U(\mathcal{E}_D^s)U^*$ as we desired. On the other hand

$$\begin{aligned} Q_N(T_N\Delta) &= \begin{pmatrix} A_{1N} & 0 \\ A_{2N} & I \end{pmatrix} \begin{pmatrix} I & 0 \\ -A_{2N}(I - A_{1N})^{-1} & 0 \end{pmatrix} = \begin{pmatrix} A_{1N} & 0 \\ A_{2N}\Big(I - (I - A_{1N})^{-1}\Big) & 0 \end{pmatrix} \\ &= \begin{pmatrix} A_{1N} & 0 \\ A_{2N}\Big(-A_{1N}\Big)(I - A_{1N})^{-1} & 0 \end{pmatrix} = \begin{pmatrix} A_{1N} & 0 \\ -A_{2N}(I - A_{1N})^{-1}A_{1N} & 0 \end{pmatrix}. \end{aligned}$$

and

$$(T_N\Delta)Q_N = \begin{pmatrix} I & 0 \\ -A_{2N}(I-A_{1N})^{-1} & 0 \end{pmatrix}\begin{pmatrix} A_{1N} & 0 \\ A_{2N} & I \end{pmatrix} = \begin{pmatrix} A_{1N} & 0 \\ -A_{2N}(I-A_{1N})^{-1}A_{1N} & 0 \end{pmatrix}.$$

So, $Q_N T_N\Delta = T_N\Delta Q_N$. This implies that both, $\mathcal{E}^s_N$ and $T_N\mathcal{U}(D)$, are invariant for $T_N\Delta$. Clearly, $T_N\Delta$ restricted to $T_N\mathcal{U}(D)$ is the identity. Hence, it only remains to prove that $\big(T_N\Delta\big)\big|_{\mathcal{E}^s_N}$ has norm lower or equal to k_D. Observe that it is enough to make the estimation at $T_D\mathcal{S}(D)$. Indeed, for every $X \in \mathcal{E}^s_N$, it holds that $T_N\Delta(X) = Ad_U\big(T_D\Delta\big)Ad_U^{-1}(X)$, $Ad_U^{-1}(X) \in \mathcal{E}^s_D$, and Ad_U is an isometric isomorphism from $T_D\mathcal{S}(D)$ onto $T_N\mathcal{S}(D)$.

So, let $Y = \begin{pmatrix} y \\ -A_{2D}(I-A_{1D})^{-1}y \end{pmatrix} \in \mathcal{E}^s_D$. Then

$$\begin{aligned}
\|(T_D\Delta)(Y)\|_2^2 &= \left\|\begin{pmatrix} A_{1D} & 0 \\ A_{2D} & I \end{pmatrix}\begin{pmatrix} y \\ -A_{2D}(I-A_{1D})^{-1}y \end{pmatrix}\right\|_2^2 \\
&= \left\|\begin{pmatrix} A_{1D}(y) \\ A_{2D}(y) - A_{2D}(I-A_{1D})^{-1}(y) \end{pmatrix}\right\|_2^2 \\
&= \|A_{1D}(y)\|_2^2 + \left\|A_{2D}(y) - A_{2D}(I-A_{1D})^{-1}(y)\right\|_2^2 \\
&\le k_D^2\,\|y\|_2^2 + \left\|-A_{2D}A_{1D}(I-A_{1D})^{-1}(y)\right\|_2^2 .
\end{aligned}$$

where the inequality holds because, by Corollary 4.1.6, $\|A_{1D}\| \le k_D$. On the other hand, by Lemma 4.1.3, we know that $\Psi_{H_1}Q_D = Q_D\Psi_{H_1}$. So, using Proposition 4.1.5, we obtain

$$\begin{aligned}
\left\|-A_{2D}A_{1D}(I-A_{1D})^{-1}(y)\right\|_2^2 &= \left\|-(I-Q_D)\,\Psi_{H_2}\,Q_D\,\Psi_{H_1}\,Q_D\Big((I-A_{1D})^{-1}(y)\Big)\right\|_2^2 \\
&= \left\|-\Psi_{H_1}(I-Q_D)\,\Psi_{H_2}\,Q_D\Big((I-A_{1D})^{-1}(y)\Big)\right\|_2^2 \\
&\le \|\Psi_{H_1}\|^2\left\|-(I-Q_D)\,\Psi_{H_2}\,Q_D\Big((I-A_{1D})^{-1}(y)\Big)\right\|_2^2 \\
&= k_D^2\left\|-A_{2D}(I-A_{1D})^{-1}(y)\right\|_2^2 .
\end{aligned}$$

Therefore

$$\|(T_D\Delta)(Y)\|_2^2 \le k_D^2\ \|y\|_2^2 + k_D^2\ \left\|-A_{2D}(I-A_{1D})^{-1}(y)\right\|_2^2 = k_D^2\|Y\|_2^2\ .$$

The smoothness of the map $\mathcal{U}(D) \ni N \mapsto \mathcal{E}^s_N$ follows from item (4) and the existence of C^∞ local cross sections for the map $\pi_D : \mathcal{U}(r) \to \mathcal{U}(D)$, which exist by Proposition 2.2.2. For example, if $\sigma_D : \mathcal{U} \to \mathcal{U}(r)$ is such a section near D, then by item (4) and Eq. (6)

$$P_N = Ad_{\sigma_D(N)}\,P_D Ad_{\sigma_D(N)^*} \quad , \quad N \in \mathcal{U}\ .$$

This completes the proof. ■

A Appendix: Stable manifold Theorem

Let f be a smooth endomorphism of a Riemannian manifold and let N be an f-invariant submanifold of M. Under the conditions of Theorem 2.1.4 we can suppose that the tangent bundle at N can be splitted in two $Df-$invariant subbundles, one given by the tangent bundle of N and the other being contracted by Df (see Definition 2.1.3). In this case, as it holds for fixed points, it is proved that for each point x in N there is a transversal smooth submanifold to N containing x and characterized by the points that converges assymptoticaly to the orbit of x. The union of these submanifolds conforms a foliation in a neighborhood of N (also called pre-lamination). This is the statement of theorem 2.1.4, which is obtained using a classical technique in dynamical systems known as *graph transform operator* (see definition (11)). This stable foliation has smooth leaves but in general is only continuous. However, if certain conditions over the $Df-$invariant splitting are also satisfied, then it can be proved that the foliation is smooth. This result, is consequences of the C^r-section theorem (stated here as theorem A.2.3 in subsection A.2). Moreover, the C^r-section theorem can be reformulated in a suitable version useful for our goals. This version is stated in theorem A.3.1; in particular, in the statement is explicite which condition should be satisfied by the $Df-$invariant splitting (see inequality (14)). To obtain this reformulation it is necessary to show that the graph transform operator introduced as a tool in the proof of the stable manifold theorem verifies certain properties. Therefore, and also for the sake of understanding for the reader, we give a sketch of the proof of the stable manifold theorem.

In our context, we want to apply the previous result for the case that the invariant submanifold is formed by fixed points. Therefore, we need to show that the hypothesis of theorem A.3.1 are full filed when we deal with a submanifold of fixed points. This is done in theorem A.4.1.

A.1 Proof of theorem 2.1.4.

Sketch of the proof: The proof consist in to use the graph transform operator. Basically consists in the following: In a neighborhood of any points $x \in N$ we consider the exponential map $\exp_x : (T_xM)_r \to M$ where $(T_xM)_r$ is the ball of radius r in T_xM, and we take the sets

$$\hat{\mathcal{E}^s}_x(r) = \exp(\mathcal{E}^s_x \cap (T_xM)_r), \quad \hat{\mathcal{F}}_x(r) = \exp(\mathcal{F}_x \cap (T_xM)_r).$$

Then it is taken r small and the space of pre-lamination σ such that for each $x \in N$ follows that σ_x is a smooth map $\sigma_x : \hat{\mathcal{E}^s}_x(r) \to \hat{\mathcal{F}}_x(r)$ (in what follows, to avoid notation we simple note these subbundles with $\hat{\mathcal{E}^s}_x$ and $\hat{\mathcal{F}}_x$). Then it is taken the operator which roughly speaking transform one pre-lamination into another one such that its images are related in the following way (see (11) for details):

$$\sigma \to \tilde{\sigma}, \text{ such that } image(\tilde{\sigma}_x) = f^{-1}(image(\sigma_{f(x)})) \cap B_r(x).$$

The goal is to prove that this operator is a contractive operator and so it has a fixed point. Latter it is shown that this fixed point corresponds to the stable lamination. Coming back to the sketch of the proof, first it is considered the maps

$$f^1_x = p^1_x \circ f : M \to \hat{\mathcal{E}^s}_x \quad \text{and} \quad f^2_x = p^2_x \circ f : M \to \hat{\mathcal{F}}_x \ ,$$

where p^1_x is the projection on $\hat{\mathcal{E}^s}_x$ and p^2_x is the projection on $\hat{\mathcal{F}}_x$. We take

$$C^r(\hat{\mathcal{E}^s}_x, \hat{\mathcal{F}}_x)$$

the set of C^r maps from $\hat{\mathcal{E}^s}_x$ to $\hat{\mathcal{F}}_x$ and we consider the space

$$C^{r,0}(\hat{\mathcal{E}^s}, \hat{\mathcal{F}}) = \{\sigma : N \to C^r(\hat{\mathcal{E}^s}_x, \hat{\mathcal{F}}_x)\}$$

i.e.: for each $x \in N$ we take $\sigma_x \in C^r(\hat{\mathcal{E}^s}_x, \hat{\mathcal{F}}_x)$ and we assume that $x \to \sigma_x$ moves continuously with x. We can represent $C^{r,0}(\hat{\mathcal{E}^s}, \hat{\mathcal{F}})$ as a vector bundle over N given by $N \times \{C^r(\hat{\mathcal{E}^s}_x, \hat{\mathcal{F}})\}_{x \in X}$. Now we take the *graph transform operator*

$$\Gamma_f(\sigma_x) = \big(f^2_x \circ (id, \sigma_{f(x)})\big)^{-1} \circ \big(f^1_x \circ (id, \sigma_x)\big)\big|_{\hat{\mathcal{E}^s}_x}. \tag{11}$$

It is proved that the graph transform operator is a contractive map and therefore it has a fixed point. In fact, to show that is contractive operator it is used the following remark

Remark A.1.1. The Lipschitz constant of the graph transform operator is smaller that λ where λ is the constant that bounds $\frac{||Df_{\mathcal{E}^s}||}{m(Df_{|\mathcal{F}})}$ (see inequality (1) in Definition 2.1.3). In fact, to prove that it is enough to show that graph transform operator associated to f is close to the graph transform operator Γ_{Df} associated to $Df$and that λ is an upper bound for $Lip(\Gamma_{Df})$. The graph transform operator associated to the derivative of f, acts on the space $L(\mathcal{E}^s, \mathcal{F})$ which is the bundle of linear maps from $\mathcal{E}^s$ into $\mathcal{F}$. Using the splitting $\mathcal{E}^s \oplus \mathcal{F}$, we can write Df in the following way:

$$Df = \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix} \ ,$$

where $A = Df_{|\mathcal{E}^s}$ and $D = Df_{|\mathcal{F}}$. Hence, if $P \in L(\mathcal{E}^s, \mathcal{F})$, then $\Gamma_{Df}(P)$ is defined as

$$\Gamma_{Df}(P) = D^{-1} \circ P \circ A. \tag{12}$$

In particular, it follows that

$$Lip(\Gamma_{Df}) = \frac{||A||}{m(D)} = \frac{||Df_{|\mathcal{E}^s}||}{m(Df_{|\mathcal{F}})} < \lambda < 1.$$

Later, it is shown that the graph transform Γ_f is close to Γ_{Df} and so the remark follows. To see that " Γ_f is close to Γ_{Df}" observe that D^{-1} in $x \in N$ is the derivative of $f^{2^{-1}}_x$ and A in $x \in N$ is the derivative of f^1_x. ▲

From the remark A.1.1, we conclude that Γ_f is a contractive operator with Lipschitz constant bounded by λ.

A.2 C^r-section theorem.

The goal is to prove that the pre-lamination obtained in Theorem 2.1.4 is smooth. To do that, it is a used the following general theorem and latter we show how to adapt to prove the smoothness of the pre-lamination and we will address the particular case of a submanifold of fixed points.

Definition A.2.1. Let $\Pi : E \to X$ be a vector bundle with a metric space base X. We say that d is an admissible on E when:

1. it induces a norm on each fiber;
2. there is a Banach space A such that the product metric on $X \times A$ induced d on E;
3. the projection of $X \times A$ onto E is of norm 1.

Without loss of generality we can assume that $E = X \times A$.

Definition A.2.2. Let $\Pi : E \to X$ be a vector bundle with a metric space base X, with an admissible metric on E. Let X_0 be a subset of X and D be the disc bundle of radius C in E, where $C > 0$ is a finite constant. Let D_0 be the restriction of D to X_0; $D_0 = D \cap \Pi^{-1}(X_0)$. Let h be a continuous map of X_0 into X. We say that $F : D_0 \to D$ is a map which covers h, if

$$\Pi \circ F = h.$$

Theorem A.2.3 (**C^r-section theorem.**). *Let $\Pi : E \to X$ be a vector bundle over the metric space X, with an admissible metric on E. Let X_0 be a subset of X and D be the disc bundle of radius C in E, where $C > 0$ is a finite constant. Let D_0 be the restriction of D to X_0; $D_0 = D \cap \Pi^{-1}(X_0)$. Let h be an overflowing continuous map of X_0 into X, that is $X_0 \subset h(X_0)$. Let $F : D_0 \to D$ be a map which covers h. Suppose that there is a constant k, $0 \leq k_0 < 1$ such that for all $x \in X_0$, the restriction of F to the fiber over X, $F_x : D_x \to D_{h(x)}$, is Lipschitz with constant at most k. Then:*

1. *There is a unique section $\sigma : X_0 \to D_0$ such that $F(\text{Image of}\sigma) \cap D_0 = \text{Image of}\sigma$.*
2. *If, X, X_0 and E are C^r-manifolds with bounded derivatives, if $\mu = Lip(h^{-1})$ be the Lipschitz constant of h^{-1} and it is satisfied*

$$k\mu^r < 1 \tag{13}$$

then follows that σ is C^r. ■

The previous theorem corresponds to theorem 5.18 of [12] (see page 58) and [13] (see page 44).

Remark A.2.4. Observe that in the previous Theorem, it is not assumed that the manifolds have to be compact. ▲

A.3 Application to the smoothness of the stable lamination.

Theorem A.3.1 (Smoothness of the stable lamination). *Let f be a C^r endomorphism of M with a ρ-pseudo hyperbolic submanifold N with $\rho < 1$. Let $\mathcal{W}^s : N \to Emb^r((-1,1)^k, M)$ be the C^r-pre-lamination of class C^0, introduced in Theorem 2.1.4. If $m(\cdot)$ denotes the minimum norm, and*

$$\frac{||Df_{/\mathcal{E}^s}||}{m(Df_{/\mathcal{F}})}||Df_{/\mathcal{F}}||^r < \lambda < 1, \tag{14}$$

then $\mathcal{W}^s : \mathcal{U} \cap N \to Emb^r((-1,1)^k, M)$ is a C^r-pre-lamination of class C^r.

Sketch of the proof: In the hypothesis of Theorem A.2.3 we consider $X = M$, $X_0 = N$, $E = M \times \{C^r(\hat{\mathcal{E}^s}_x, \hat{\mathcal{F}}_x)\}_{x\in N}$ (i.e.: the pairs (x, σ_x) such that $\sigma_x : \hat{\mathcal{E}^s}_x \to \hat{\mathcal{F}}_x$), $h = f^{-1}$, $D_0 = N \times \{C^r(\hat{\mathcal{E}^s}_x, \hat{\mathcal{F}})\}_{x\in N}$ and $F(x, \sigma) = (f(x), \Gamma_f)$ where Γ_f is the graph transform operator associated to f. From remark A.1.1 follows that $Lip(F)$ is close to $\frac{||Df_{/\mathcal{E}^s}||}{m(Df_{/\mathcal{F}})}$ and it is immediate that $Lip(h^{-1}) = Lip(f) = ||Df||$. Therefore, if (14) holds, then

$$Lip(f)^r Lip(\Gamma_f) < 1,$$

and therefore the inequality (13) holds and so we can apply Theorem A.2.3. ■

A.4 Application to a compact submanifold of fixed points.

Now we shows that we can apply A.3.1 to the case of a submanifold of fixed points.

Corollary A.4.1 (Smoothness of the stable lamination for a submanifold of fixed points). *Let f, M and N as in Theorem 2.1.4. Let us assume that any point p in N is a* **fixed point**. *Then C^r-pre-lamination $\mathcal{W}^s : \mathcal{N} \to Emb^r((-1,1)^k, M)$ is of class C^r.*

Proof. Observe that $Df_{/\mathcal{F}} = Id$. Therefore

$$\frac{||Df_{/\mathcal{E}^s}||}{m(Df_{/\mathcal{F}})}||Df_{/\mathcal{F}}||^k = ||Df_{/\mathcal{E}^s}|| < \lambda < 1$$

and so it follows that $\mathcal{W}^s : \mathcal{U} \cap N \to \mathrm{Emb}^r((-1,1)^k, M)$ is a C^r-pre-lamination of class C^r, by Theorem A.3.1. □

Remark A.4.2. Similar results to the one obtained in theorem A.4.1 are obtained in [11]. In this paper, it is shown that the stable foliation is C^1 assuming a similar condition to (14) for the context of partial hyperbolic systems.

References

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

[2] T. Ando, *Aluthge Transforms and the Convex Hull of the Eigenvalues of a Matrix*, Linear Multilinear Algebra 52 (2004), 281-292.

[3] T. Ando and T. Yamazaki, *The iterated Aluthge transforms of a 2-by-2 matrix converge*, Linear Algebra Appl. 375 (2003), 299-309.

[4] J. Antezana, P. Massey and D. Stojanoff, λ*-Aluthge transforms and Schatten ideals*, Linear Algebra Appl. 405 (2005), 177-199.

[5] R. Bhatia, Matrix Analysis, Berlin-Heildelberg-New York , Springer 1997.

[6] K. Dykema and H. Schultz, *On Aluthge Transforms: continuity properties and Brown measure*, Available online at www.arxiv.org/PS_cache/math/pdf/0512/0512197.pdf

[7] M. W. Hirsch, C. C. Pugh, and M. Shub, *Invariant manifolds*, Lecture Notes in Mathematics, Vol. 583. Springer-Verlag, Berlin-New York, 1977.

[8] I. Jung, E. Ko, and C. Pearcy, *Aluthge transform of operators*, Integral Equations Operator Theory 37 (2000), 437-448.

[9] I. Jung, E. Ko, and C. Pearcy, *The Iterated Aluthge Transform of an operator*, Integral Equations Operator Theory 45 (2003), 375-387.

[10] T. Kato, Perturbation theory for linear operators. Reprint of the 1980 edition. Classics in Mathematics. Springer-Verlag, Berlin, 1995.

[11] C. Pugh, M. Shub, A. Wilkinson *Holder foliations* Duke Math. J. 86, No.3 (1997), pp.517-546.

[12] M. Shub, *Stabilite Globale des Systemes Dynamiques*, Asterisque, Vol. (56), 211 pages, Societe Mathematique de France, Paris, 1978.

[13] M. Shub, *Global Stability of Dynamical Systems, Springer*, 1986 (this book is a translation of the French edition with added chapters and corrections).

[14] P. Y. Wu, *Numerical range of Aluthge transform of operator*, Linear Algebra and Appl. 357(2002), 295-298.

[15] T. Yamazaki, *An expression of the spectral radius via Aluthge tranformation*, Proc. Amer. Math. Soc. 130 (2002), 1131-1137.