Available online at www.sciencedirect.com

ScienceDirect

ADVANCES IN Mathematics

Advances in Mathematics 226 (2011) 1591–1620

www.elsevier.com/locate/aim

The iterated Aluthge transforms of a matrix converge

Jorge Antezana [a,b,c,*,1], Enrique R. Pujals [d,2], Demetrio Stojanoff [b,c,3]

[a] *Departament de Matemátiques UAB, Barcelona, Spain*
[b] *Depto. de Matemática, FCE-UNLP, La Plata, Argentina*
[c] *IAM-CONICET, Argentina*
[d] *Instituto Nacional de Matemática Pura y Aplicada (IMPA), Rio de Janeiro, Brazil*

Received 11 July 2008; accepted 27 August 2010

Available online 22 September 2010

Communicated by N.G. Makarov

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 $r \times r$ matrix T. This result was conjectured by Jung, Ko and Pearcy in 2003. We also analyze the regularity of the limit function.

MSC: primary 37D10; secondary 15A60

Keywords: Aluthge transform; Stable manifold theorem; Similarity orbit; Polar decomposition

* Corresponding author at: Depto. de Matemática, FCE-UNLP, La Plata, Argentina.
E-mail addresses: antezana@mate.unlp.edu.ar (J. Antezana), enrique@impa.br (E.R. Pujals), demetrio@mate.unlp.edu.ar (D. Stojanoff).

1 Partially supported by MTM-2008-05561-C02-02/MTM, 2009 SGR 1303, Universidad de La Plata (UNLP 11 X472) and ANPCYT (PICT07-00808).

2 Partially supported by CNPq.

3 Partially supported by CONICET (PIP 4463/96), Universidad de La Plata (UNLP 11 X472) and ANPCYT (PICT03-09521).

0001-8708/$ – see front matter
doi:10.1016/j.aim.2010.08.012

Contents

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 transform was introduced in [2] to study p-hyponormal and log-hyponormal operators. Roughly speaking, the idea behind the Aluthge transform is to take an operator into another operator which shares some spectral properties with the first one, in particular the spectrum, and it is closer to be normal. It is also well known that $\Delta(\lambda VTV^*) = \lambda V\Delta(T)V^*$ for every $\lambda \in \mathbb{C}$ and unitary operator V, that it is a contraction (not necessarily strict) with respect to the spectral norm, and that

$$\Delta(T_1 \oplus T_2) = \Delta(T_1) \oplus \Delta(T_2). \tag{1}$$

The Aluthge transform has received much attention recently. One reason is its connection with the invariant subspace problem. Jung, Ko and Pearcy proved in [18] that T has a nontrivial invariant subspace if and only if $\Delta(T)$ does. On the other hand, Dykema and Schultz [13] proved that the Brown measure is preserved 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 $n \in \mathbb{N}$. In [19] Jung, Ko and Pearcy raised the following conjecture:

Conjecture 1. *For every bounded operator T on $\mathcal{H}$, the sequence of iterates $\{\Delta^n(T)\}_{n\in\mathbb{N}}$ converges.*

Although some results supported this conjecture (see for instance [3,25,27,26]), counterexamples were found in the infinite dimensional setting. In [11] for instance, Cho, Jung and Lee showed an example based in a weighted shift where the sequence of iterated does not converge even with respect to the weak operator topology.

Despite these counterexamples, the problem in the finite dimensional case remained open. In this setting, since the sequence $\{\Delta^n(T)\}_{n\in\mathbb{N}}$ is bounded, it has at least one limit point. Moreover, the following result, proved independently by Jung, Ko and Pearcy [19] and Ando [3], gives more information about the possible limit points:

Proposition 1.1. *If T is an $r \times r$ matrix, the limit points of the sequence $\{\Delta^n(T)\}_{n\in\mathbb{N}}$ are normal. Moreover, if L is a limit point, then $\sigma(L) = \sigma(T)$ counting multiplicities.*

As a simple consequence of this result, if the spectrum of T has only one point, then the iterated sequence converges. Indeed, in this case there is only one possible limit point. This allowed Ando and Yamazaki to reduce the problem for 2×2 matrices to the case where the eigenvalues are different, and in [4] they show that Conjecture 1 is true for 2×2 matrices. Later on, following Ando–Yamazaki's general ideas, Huang and Tam proved in [17] that the conjecture is true for matrices whose nonzero eigenvalues have different moduli. As far as we know, this is the most general result, in the direction of Conjecture 1, that has been proved using linear algebra techniques.

The main goal of this paper is to prove that Conjecture 1 is true for every matrix, and hence to completely solve the problem in the finite dimensional case. Our approach involves a combination of dynamical system techniques with geometrical arguments. The dynamical system techniques used here have already shown to be useful for the study of this problem. Indeed, in our previous work [7], they were used to prove the convergence of the iterated Aluthge transform for diagonalizable matrices.

This combination of dynamical and geometrical tools used in order to study an operator theoretical problem like Conjecture 1 may suggest a new possible interaction among these branches of mathematics. On one hand, it provides another field of applications of the stability theory of hyperbolic systems and invariant manifolds. In this sense, the work of Shub and Vasquez on the QR algorithm is an important precedent (see [24]). On the other hand, it provides to operator theorists a new set of tools to deal with problems where the usual techniques fail. In our case, besides the solution, it also provides a better understanding of the problem. The dynamical system perspective not only allows us to prove Conjecture 1, but it also provides further information related to the regularity of the limit function and the rate of convergence of the iterated sequence. The results of this work are stated for the standard Aluthge transform, but all of them can be generalized to the so-called λ-Aluthge transforms mutatis mutandis. See Section 6 for more details.

The paper is organized as follows: since the proof of the convergence of the iterated Aluthge transform is rather long and very technical, in Section 2 we include a description of our approach, detailing the main results and the geometrical ideas behind our strategy. In Section 3, we collect several preliminary definitions and results about the geometry of similarity and unitary orbits, the stable manifold theorem, and some known properties of the spectral projections. Section 4 contains the proof of the convergence. It is divided in subsections, each devoted to the proof of some of the steps described in Section 2. In Section 5 we study the regularity of the limit map $T \mapsto \Delta^\infty(T)$, mainly for T invertible. Section 6 contains concluding remarks about the rate of convergence, and the extension of the main results to the λ-Aluthge transforms, for every $\lambda \in (0, 1)$. Finally, we include an appendix divided in two parts. The first one is a brief review of some results in differential geometry, including the definitions and results used throughout the paper. The second part contains some comments on the stable manifold theorem and related dynamical arguments used, mainly oriented to those readers who may not be familiar with them.

We would like to thank Prof. M. Shub for comments and suggestion about the stable manifold theorems, and we would also like to thank Prof. G. Corach who introduced to us the Aluthge transform, and shared with us fruitful discussions concerning this matter.

Notation. Throughout 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 $\mathcal{M}_r^h(\mathbb{C})$ (resp. $\mathcal{M}_r^{ah}(\mathbb{C})$) the real subspace of hermitian (resp. antihermitian) matrices. We denote $\mathcal{N}(r) = \{N \in \mathcal{M}_r(\mathbb{C})\colon N \text{ is normal}\}$. If $v \in \mathbb{C}^r$, we denote by $\operatorname{diag}(v) \in \mathcal{M}_r(\mathbb{C})$ the diagonal matrix with v in its diagonal and zeroes elsewhere.

Given $T \in \mathcal{M}_r(\mathbb{C})$, $R(T)$ denotes the range or image of T, $\ker(T)$ the null space of T, $\operatorname{rk}(T) = \dim R(T)$ (i.e., the rank of T), $\sigma(T)$ the spectrum of T, $\lambda(T) \in \mathbb{C}^r$ the vector of eigenvalues of T (counted with multiplicity), $\rho(T)$ the spectral radius of T, $\operatorname{tr}(T)$ the trace of T, and T^* the adjoint of T. We shall consider the space of matrices $\mathcal{M}_r(\mathbb{C})$ as a real Hilbert space with the inner product defined by

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

The norm induced by this inner product is the Frobenius norm and is denoted by $\|\cdot\|_2$. For $T \in \mathcal{M}_r(\mathbb{C})$ and $\mathcal{A} \subseteq \mathcal{M}_r(\mathbb{C})$, by $\operatorname{dist}(T, \mathcal{A})$ we mean the distance between them with respect to the Frobenius norm.

Let M be a manifold. We denote by TM the tangent bundle of M and by T_xM the tangent space at the point $x \in M$. For any $k \geqslant 1$, we denote by $C^k(M)$ the set of C^k maps from M to $\mathbb{C}$ and by $C^k(\mathcal{U}, M)$ the set of C^k maps from an open set $\mathcal{U} \subseteq \mathbb{R}^n$ to M. Given a function $f \in C^k(M)$, we denote by $df_x(V)$ the derivative of f at the point x applied to the tangent vector $V \in T_xM$. In $C^k(\mathcal{U}, M)$ we shall consider the C^k-topology, that is the topology where two functions are close if the functions and all their derivatives until order k are uniformly close on compact subsets of $\mathcal{U}$. We denote by $\operatorname{Emb}^k(\mathcal{U}, M)$ the subset of $C^k(\mathcal{U}, M)$ consisting of the embeddings from $\mathcal{U}$ into M, endowed with the relative C^k-topology. We refer the reader to Appendix A for a more detailed description of these objects.

2. Geometrical idea of the proof

In this section we sketch the geometrical ideas behind the proof of the conjecture, the technicalities are left for Section 4.

The first step of the proof is a reduction of the problem to the invertible case. Let $T \in \mathcal{M}_r(\mathbb{C})$. In [6] it is proved that the 2×2 block matrix form of $\Delta^r(T)$ with respect to the decomposition of $\mathbb{C}^r = N(\Delta^r(T)) \oplus N(\Delta^r(T))^\perp$ is

$$\Delta^r(T) = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} = A \oplus 0,$$

where A is an invertible matrix. Since $\Delta(B_1 \oplus B_2) = \Delta(B_1) \oplus \Delta(B_2)$, this result implies that it is enough to study the convergence for invertible matrices. One of the main advantages of this reduction is that, restricted to the group of invertible matrices, the Aluthge transform is a C^∞-map, which is not true in the whole space of matrices where globally it is only continuous. Also note that in the invertible case we have that

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

Therefore the whole iterated sequence belongs to the similarity orbit of T defined by

$$\mathcal{S}(T) = \{ STS^{-1} \colon S \text{ is an } r \times r \text{ invertible matrix} \}.$$

This context seems to be the natural one to deal with the diagonalizable case. In order to introduce the geometrical ideas that will be used in the general case, we will describe in some detail the proof of convergence for diagonalizable matrices given in [7].

If T is diagonalizable and invertible, then $\mathcal{S}(T)$ coincides with $\mathcal{S}(D)$ for some invertible diagonal matrix D. Let $\mathcal{U}(D) = \{UDU^* \colon U \in \mathcal{U}(r)\}$ be the unitary orbit of D. Note that it consists precisely of those normal operators N that satisfy $\lambda(N) = \lambda(T) = \lambda(D)$. So, from this point of view, Proposition 1.1 asserts that all the limit points of the iterated sequence $\{\Delta^n(T')\}_{n\in\mathbb{N}}$ are in $\mathcal{U}(D)$, for every $T' \in \mathcal{S}(T) = \mathcal{S}(D)$. Both $\mathcal{S}(D)$ and $\mathcal{U}(D)$ have a very rich geometrical structure, and $\mathcal{U}(D)$ is a compact submanifold of $\mathcal{S}(D)$ that consists of all the fixed points for $\Delta(\cdot)$ in $\mathcal{S}(D)$. These facts motivate a dynamical approach. In this direction, and with the aim of using some dynamical tools provided by the theory of hyperbolic systems, the following theorem was proved in [7]:

Theorem 2.1. *Let $D = \operatorname{diag}(d_1, \ldots, d_r) \in \mathcal{M}_r(\mathbb{C})$ be an invertible diagonal matrix. 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 invariant for the derivative $d\Delta_N$.*
3. *$d\Delta_N|_{T_N\mathcal{U}(D)} = I_{T_N\mathcal{U}(D)}$ and $\|d\Delta_N|_{\mathcal{E}^s_N}\| \leqslant k_D < 1$, where*

$$k_D = \max_{i,j\colon d_i \neq d_j} \frac{|1 + e^{i(\arg(d_j) - \arg(d_i))}| |d_i|^{1/2} |d_j|^{1/2}}{|d_i| + |d_j|}.$$

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 $\mathcal{U}(D) \ni N \mapsto P_{\mathcal{E}^s_N \| T_N\mathcal{U}(D)}$ is smooth, where $P_{\mathcal{E}^s_N \| T_N\mathcal{U}(D)}$ denotes the projection onto $\mathcal{E}^s_N$ parallel to $T_N\mathcal{U}(D)$.

This theorem says precisely that the hypotheses of the so-called stable manifold theorem are satisfied.[4] The idea behind the stable manifold theorem is the same as in the inverse mapping theorem: the properties of the derivative are locally inherited by the function.

[4] For the sake of completeness, this theorem is stated as Theorem B.3 in Appendix B.2.

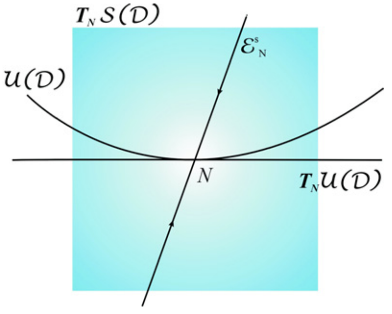

Fig. 1. $T_N\mathcal{S}(D) = \mathcal{E}^s_N \oplus T_N\mathcal{U}(D)$.

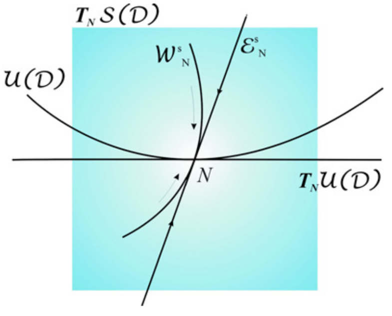

Fig. 2. The submanifolds $\mathcal{W}^s_N$.

Note that Theorem 2.1 says that for every $N \in \mathcal{U}(D)$ the tangent space $T_N\mathcal{S}(D)$ can be decomposed in two $d\Delta_N$-invariant subspaces. In $T_N\mathcal{U}(D)$, the derivative $d\Delta_N$ behaves as the identity because every point in the unitary orbit of D is a fixed point for Δ. On the other hand, in $\mathcal{E}^s_N$ the derivative $d\Delta_N$ is a strict contraction. Therefore, if we take $X = E + Y \in \mathcal{E}^s_N \oplus T_N\mathcal{U}(D) = T_N\mathcal{S}(D)$, and the operator $d\Delta_N$ is applied iteratively to X, then the sequence obtained converges to $Y \in T_N\mathcal{U}(D)$, and the rate of convergence is exponential (see Fig. 1).

The stable manifold theorem assures that Δ has the same behavior locally in $\mathcal{S}(D)$. More precisely, it says that there exists a smooth submanifold $\mathcal{W}^s_N$ through each $N \in \mathcal{U}(D)$ such that $T_N(\mathcal{W}^s_N) = \mathcal{E}^s_N$ and for every $S \in \mathcal{W}^s_N$ and some $\rho \in (0, k_D)$, it holds that

$$d\bigl(N, \Delta^n(S)\bigr) \leqslant \rho^n\, d(N, S),$$

where d denotes the Riemannian distance in the similarity orbit. Note that in particular $\mathcal{W}^s_N$ is transversal to $\mathcal{U}(D)$ (see Fig. 2). Since the map $\mathcal{U}(D) \ni N \mapsto P_{\mathcal{E}^s_N \| T_N\mathcal{U}(D)}$ is smooth, the distribution of the submanifold $\mathcal{W}^s_N$ is smooth enough to show that the union of all the $\mathcal{W}^s_N$ contains an open neighborhood of $\mathcal{U}(D)$.

The existence of this neighborhood implies the convergence of $\{\Delta^n(S)\}_{n\in\mathbb{N}}$, for each $S \in \mathcal{S}(D)$. Indeed, as its limit points are in $\mathcal{U}(D)$, for n large enough the sequence enters in the neighborhood and therefore we can assure that it belongs to some $\mathcal{W}^s_N$, which implies the convergence of the sequence to N.

The non-diagonalizable case is much more complicated because the geometric framework is different. If we start with an invertible non-diagonalizable matrix T, the whole similarity orbit $\mathcal{S}(T)$ consists of non-diagonalizable matrices, including the elements of the sequence $\{\Delta^n(T)\}_{n\in\mathbb{N}}$. However, the limit points are normal, and hence they are diagonalizable. Geometrically, this can be visualized as follows: if we start with an element $T' \in \mathcal{S}(T)$ for some invertible non-diagonalizable matrix T, then the sequence $\{\Delta^n(T')\}_{n\in\mathbb{N}}$ tends to the boundary of $\mathcal{S}(T)$. But the boundary of $\mathcal{S}(T)$ not only contains those diagonalizable matrices that shares the characteristic polynomial with T, but also any matrix with "smaller" Jordan form than the Jordan form of T. So, the boundary of $\mathcal{S}(T)$ is a kind of lattice of boundaries. Therefore, it would be very complicated to pursue a similar approach, and a different strategy is needed to prove the general case. Before going on, let us introduce some definitions.

Definition 2.2. Let $\mathbb{P} \subseteq \mathcal{M}_r(\mathbb{C})$ be a compact set of fixed points for Δ, i.e., a compact set of normal matrices. Its **basin of attraction** is the set

$$B_\Delta(\mathbb{P}) = \big\{T \in \mathcal{M}_r(\mathbb{C}):\ \operatorname{dist}\big(\Delta^n(T), \mathbb{P}\big) \xrightarrow[n\to\infty]{} 0\big\}$$

and, for every $\varepsilon > 0$, the local basin of attraction is the set

$$B_\Delta(\mathbb{P})_\varepsilon = \big\{T \in B_\Delta(\mathbb{P}):\ \operatorname{dist}\big(\Delta^n(T), \mathbb{P}\big) < \varepsilon,\ n \in \mathbb{N}\big\}.$$

Note that if $\mathbb{P} = \mathcal{U}(D)$ for some diagonal invertible matrix D, Proposition 1.1 implies that the basin $B_\Delta(\mathbb{P})$ also admits the following spectral characterization

$$B_\Delta\big(\mathcal{U}(D)\big) = \big\{T \in \mathcal{M}_r(\mathbb{C}):\ \lambda(T) = \lambda(D)\big\}. \tag{2}$$

To study the general case, our framework will be the open set of invertible matrices. So, our first step is to extend the description of the derivative of the Aluthge transform at a normal matrix N given in Theorem 2.1, in order to include the new directions that appear outside of the tangent space of $\mathcal{S}(N)$.

More explicitly, the subspace $\mathcal{A}_N = \{N\}' = \{T \in \mathcal{M}_r(\mathbb{C}):\ NT = TN\}$ satisfies that $\mathcal{M}_r(\mathbb{C}) = \mathcal{A}_N \oplus T_N\mathcal{S}(N)$. The mentioned extension is contained in the following theorem, which will be proved in Section 4, following Proposition 4.1.

Theorem 2.3. *Let $D = \operatorname{diag}(d_1, \dots, d_r) \in \mathcal{M}_r(\mathbb{C})$ be an invertible diagonal matrix. For $N \in \mathcal{U}(D)$, let $F_N = \mathcal{A}_N \oplus T_N\mathcal{U}(D)$. Then, for every $N \in \mathcal{U}(D)$, the subspaces F_N and $\mathcal{E}^s_N$ satisfy $\mathcal{M}_r(\mathbb{C}) = F_N \oplus \mathcal{E}^s_N$ and they also satisfy items* 2, 3 *and* 4 *of Theorem* 2.1, *but replacing $T_N\mathcal{U}(D)$ by F_N.*

Roughly speaking, this theorem says that in those new directions that appear in the tangent space of $\mathcal{G}l_r(\mathbb{C})$, the derivative of the Aluthge transform behaves as the identity. Hence, all these directions are added to those corresponding to the tangent space of $\mathcal{U}(D)$, and this sum is what we call F_N.

The next step is to extend this decomposition to the tangent spaces of the points of some local basin $B_\Delta(\mathcal{U}(D))_\varepsilon$ (see Appendix B.3). This is a fairly standard procedure in dynamical systems. Since the points of the local basin are not necessarily fixed by the Aluthge transform, given $T \in B_\Delta(\mathcal{U}(D))_\varepsilon$, the conditions that the subspaces $\mathcal{E}^s_T$ and $\mathcal{F}_T$ have to satisfy are:

1. $\mathcal{M}_r(\mathbb{C}) = \mathcal{E}^s_T \oplus \mathcal{F}_T$.
2. The distributions $T \mapsto \mathcal{E}^s_T$ and $T \mapsto \mathcal{F}_T$ are continuous.
3. There exists $\rho \in (0, 1)$ which does not depend on T such that

$$\|d\Delta_T|_{\mathcal{E}^s_T}\| < 1 - \rho \quad \text{and} \quad \big\|(I - d\Delta_T)|_{\mathcal{F}_T}\big\| < \frac{\rho}{2}.$$

4. For every $T \in B_\Delta(\mathbb{P})_\varepsilon$, the subspace $\mathcal{E}^s_T$ is $d\Delta_T$-invariant, i.e., $d\Delta_T(\mathcal{E}^s_T) \subseteq \mathcal{E}^s_{\Delta(T)}$.

Having extended the decompositions to the local basin, we can use a version of the stable manifold theorem adapted to this context, which does not require any differentiable structure on the local basin. In this case, the conclusion is that through each $T \in B_\Delta(\mathcal{U}(D))_\varepsilon$ there is a submanifold $\mathcal{W}^{ss}_T$ of $\mathcal{G}l_r(\mathbb{C})$ such that $T_T\mathcal{W}^{ss}_T = \mathcal{E}^s_T$, and for every $S \in \mathcal{W}^{ss}_T$

$$d\big(\Delta^n(S), \Delta^n(T)\big) \leqslant \rho^n\, d(S, T), \tag{3}$$

where ρ is the same as before. Note that, as a consequence of (3), these submanifolds are entirely contained in $B_\Delta(\mathcal{U}(D))$. At this point we can highlight one of the main geometrical differences between the diagonalizable and non-diagonalizable case: in the diagonalizable case, the union of the (strong) stable manifold through each $N \in \mathcal{U}(D)$ forms a neighborhood of $\mathcal{U}(D)$ inside $\mathcal{S}(D)$. In the non-diagonalizable case, the union of the strong stable manifold do not necessarily form a neighborhood of $\mathbb{P}$ (w.r.t. the topology of $\mathcal{G}l_r(\mathbb{C})$).

Now, let us introduce the map $\Pi_E : B_\Delta(\mathcal{U}(D)) \to \mathcal{S}(D)$ defined by

$$\Pi_E(T) = \sum_{i=1}^{k} \lambda_i E_i(T) \quad \text{for every } T \in B_\Delta\big(\mathcal{U}(D)\big),$$

where each $E_i(T)$ is the spectral projection of T corresponding to the eigenvalue $\lambda_i(T) = \lambda_i(D)$. Using the Riesz functional calculus, this map can be extended to a smooth function defined open neighborhood of $\mathcal{U}(D)$ inside the set of invertible matrices (see Section 3.3).

Let $\mathcal{O}(D) = \Pi_E^{-1}(\mathcal{U}(D)) \subseteq B_\Delta(\mathcal{U}(D))$. Given $T \in B_\Delta(\mathcal{U}(D))$, T and $\Pi_E(T)$ have the same spectrum and spectral projections. So, for every $S \in \mathcal{O}(D)$ the spectral projections of S are mutually orthogonal and $S = S_1 \oplus \cdots \oplus S_k$, where each S_i is the restriction $S_i = SE_i(S)$. Therefore,

$$\Delta^n(S) = \Delta^n(S_1) \oplus \cdots \oplus \Delta^n(S_k) \quad \text{for every } n \in \mathbb{N}.$$

The advantage is that the spectrum of each S_i has only one point. Thus, as we have observed in the paragraph that follows Proposition 1.1, each sequence $\{\Delta^n(S_i)\}_{n\in\mathbb{N}}$ converges, and therefore $\{\Delta^n(S)\}_{n\in\mathbb{N}}$ also converges (to $\Pi_E(S)$). In conclusion, for every $S \in \mathcal{O}(D)$ the sequence of iterations of the Aluthge transform converges. So, to prove that it also converges for every T in $B_\Delta(\mathcal{U}(D))$, it would be enough to show that for every T in some local basin $B_\Delta(\mathcal{U}(D))_\varepsilon$, the submanifold $\mathcal{W}^{ss}_T$ contains an element $S \in \mathcal{O}(D)$. Indeed, in this case Eq. (3) forces the sequence $\{\Delta^n(T)\}_{n\in\mathbb{N}}$ to converge to the same limit of $\{\Delta^n(S)\}_{n\in\mathbb{N}}$.

The last step of the argument goes as follows: Let $T \in B_\Delta(\mathcal{U}(D))$ be a matrix close to the unitary orbit $\mathcal{U}(D)$, and project the stable manifolds $\mathcal{W}^{ss}_T$ into the orbit $\mathcal{S}(D)$, using the above mentioned function Π_E. By the properties of Π_E, this projection is also a submanifold of $\mathcal{S}(D)$. Moreover, it can be proved that $\Pi_E(\mathcal{W}^{ss}_T)$ is "close" in some sense to $\mathcal{W}^{ss}_N$, where N is certain

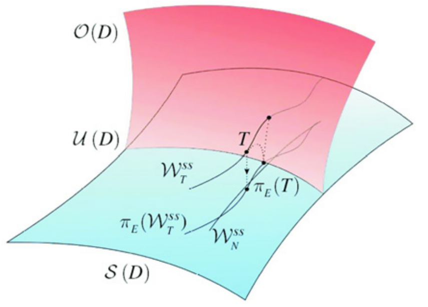

Fig. 3. The projection argument.

normal operator of $\mathcal{U}(D)$ close to T. Note that $\mathcal{W}_N^{ss}$ is one of the stable manifolds studied in the diagonalizable case. Therefore $\mathcal{W}_N^{ss}$ intersects $\mathcal{U}(D)$ transversally. These facts imply, by the well-known results about transversal intersections, that the projected submanifold $\Pi_E(\mathcal{W}_T^{ss})$ also intersects $\mathcal{U}(D)$.

Finally, if $N' \in \Pi_E(\mathcal{W}_T^{ss}) \cap \mathcal{U}(D)$, then N' is the projection of a matrix $S \in \mathcal{W}_T^{ss} \cap \mathcal{O}_D$ (see Fig. 3). Note that the rate of convergence of the sequence $\Pi_E(\Delta^n(T))$ is exponential, so that the spectral projections of the matrices $\Delta^n(T)$ become rapidly essentially orthogonal. Nevertheless, the convergence for the matrix $S \in \mathcal{W}_T^{ss} \cap \mathcal{O}_D$ can be much slower, because in this case the convergence occurs for other reasons: S has orthogonal blocks with singleton spectrum. This explains the fact, suggested by computational experiments, that the convergence is rarely slow for non-diagonalizable matrices. We shall give more details about this problem in Section 6.1.

3. Preliminaries

3.1. Similarity orbits

Definition 3.1. Let $T \in \mathcal{M}_r(\mathbb{C})$. We denote by $\mathcal{S}(T)$ the similarity orbit of T:

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

In the same fashion, $\mathcal{U}(T) = \{UTU^*:\ U \in \mathcal{U}(r)\}$ denotes the unitary orbit of T. We denote by $\pi_T : \mathcal{G}l_r(\mathbb{C}) \to \mathcal{S}(T) \subseteq \mathcal{M}_r(\mathbb{C})$ the C^∞ map defined by $\pi_T(S) = STS^{-1}$, for every $S \in \mathcal{G}l_r(\mathbb{C})$. We also use the same notation for its restriction to the unitary group: $\pi_T : \mathcal{U}(r) \to \mathcal{U}(T)$.

Remark 3.2. Let $T \in \mathcal{M}_r(\mathbb{C})$ and $N \in \mathcal{N}(r)$ with $\lambda(N) = \lambda(T)$. Then

$$\mathcal{U}(N) = \{M \in \mathcal{N}(r):\ \lambda(M) = \lambda(T)\}.$$

By the Schur's triangulation theorem, there exists $N_0 \in \mathcal{U}(N)$ such that

$$\|T\|_2^2 - \sum_{i=1}^r |\lambda_i(T)|^2 = \|T - N_0\|_2^2 \geqslant \operatorname{dist}(T, \mathcal{U}(N))^2. \tag{4}$$

Since the Aluthge transform reduces the Frobenius norms, Proposition 1.1 implies that

$$\left\|\Delta^n(T)\right\|_2 \underset{n\to\infty}{\searrow} \sum_{i=1}^{r}\left|\lambda_i(T)\right|^2 = \|N\|_2.$$

Therefore, Eq. (4) assures that the sequence $\operatorname{dist}(\Delta^n(T), \mathcal{U}(N)) \underset{n\to\infty}{\longrightarrow} 0$.

A smooth and surjective map $f : M \to N$ is a submersion if df_x is surjective for every $x \in M$. The following two results are well known (see, for example, [21,12] or [5]):

Proposition 3.3. *Let $D \in \mathcal{G}l_r(\mathbb{C})$ be a diagonal matrix. Then the similarity orbit $\mathcal{S}(D)$ is a C^∞ submanifold of $\mathcal{M}_r(\mathbb{C})$, and the projection $\pi_D : \mathcal{G}l_r(\mathbb{C}) \to \mathcal{S}(D)$ is 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 also a submersion.*

Remark 3.4. For every $M \in \mathcal{S}(D)$, it is well known that

$$T_M\mathcal{S}(D) = (d\pi_M)_I\big(\mathcal{M}_r(\mathbb{C})\big) = \big\{[A, M] = AM - MA\colon\ A \in \mathcal{M}_r(\mathbb{C})\big\}.$$

If $\sigma(D) = \{\mu_1, \ldots, \mu_k\}$, and $E_i(M)$ are the spectral projections of $M \in \mathcal{S}(D)$ associated to disjoint open neighborhoods of each μ_i, then $M = \sum_{i=1}^k \mu_i E_i(M)$. Therefore,

$$\begin{aligned} T_M\mathcal{S}(D) &= \big\{AM - MA\colon\ A \in \mathcal{M}_r(\mathbb{C})\big\} \\ &= \left\{\sum_{j=1}^{k}\mu_j A E_j(M) - \sum_{i=1}^{k}\mu_i E_i(M)A\colon\ A \in \mathcal{M}_r(\mathbb{C})\right\} \\ &= \left\{\sum_{i,j=1}^{k}(\mu_j - \mu_i)E_i(M)AE_j(M)\colon\ A \in \mathcal{M}_r(\mathbb{C})\right\} \\ &= \big\{X \in \mathcal{M}_r(\mathbb{C})\colon\ E_i(M)XE_i(M) = 0,\ 1 \leqslant i \leqslant k\big\}. \end{aligned} \tag{5}$$

Indeed, these equalities can be easily justified by using that $\sum_{i=1}^k E_i(M) = I$. Throughout this paper we shall consider on $\mathcal{S}(D)$ (and on $\mathcal{U}(D)$) the Riemannian structure inherited from $\mathcal{M}_r(\mathbb{C})$ (using the usual inner product on their tangent spaces).

For every fixed $U \in \mathcal{U}(r)$, we have that $U\mathcal{S}(D)U^* = \mathcal{S}(D)$ and the map $M \mapsto UMU^*$ is isometric, on $\mathcal{S}(D)$, with respect both to the Riemannian metric and the $\|\cdot\|_2$ metric of $\mathcal{M}_r(\mathbb{C})$.

3.2. Strong stable manifold theorem for the basin of attraction

Let M be a smooth Riemannian manifold and $f : M \to M$ a smooth map. Let $N \subseteq M$ be a compact set such that $f(N) = N$. The basin of attraction of N (by f) is the set

$$B_f(N) = \big\{y \in M\colon\ \operatorname{dist}\big(f^n(y), N\big) \underset{n\to\infty}{\longrightarrow} 0\big\}.$$

Given $\varepsilon > 0$, the local basin of N is the set

$$B_f(N)_\varepsilon = \bigl\{y \in B_f(N)\colon \operatorname{dist}\bigl(f^n(y), N\bigr) < \varepsilon, \text{ for every } n \in \mathbb{N}\bigr\}.$$

The following result is standard when stated on a compact f-invariant subset N (see Appendix B, [15] or [23]). The following version, which extends the pre-lamination $\mathcal{W}^s$ to its local basin $B_f(N)_\varepsilon$ is also well known. In Remark B.5 and Appendix B.4, we shall briefly expose the principal steps of the proof of the version for N, and then explain how that proof can be extended to "its basin of attraction".

Theorem 3.5 *(Strong stable manifold theorem). Let $f : M \to M$ be a C^k map and let N be a compact f-invariant subset of M such that $f_{|N}$ is a homeomorphism. Let us assume that for some $\varepsilon > 0$ there exist two continuous subbundles of $T_{B_f(N)_\varepsilon}M$, denoted by $\mathcal{E}^s$ and $\mathcal{F}$, such that, for every $x \in B_f(N)_\varepsilon$,*

1. *$T_{B_f(N)_\varepsilon}M = \mathcal{E}^s \oplus \mathcal{F}$.*
2. *$\mathcal{E}^s_x$ is df_x-invariant in the sense that $df_x(\mathcal{E}^s_x) \subseteq \mathcal{E}^s_{f(x)}$.*
3. *$\mathcal{F}_z$ is df_z-invariant, for every $z \in N$.*
4. *There exists $\rho \in (0, 1)$ such that df_x restricted to $\mathcal{F}_x$ expands it by a factor greater than ρ, and $df_x : \mathcal{E}^s_x \to \mathcal{E}^s_{f(x)}$ has norm lower than ρ.*

Then there is a continuous f-invariant and self-coherent C^k-pre-lamination

$$\mathcal{W}^s : B_f(N)_\varepsilon \to \operatorname{Emb}^k\bigl((-1, 1)^m, M\bigr) \quad \bigl(\textit{endowed with the } C^k\textit{-topology}\bigr)$$

such that, for every $x \in B_f(N)_\varepsilon$,

1. *$\mathcal{W}^s(x)(0) = x$,*
2. *$\mathcal{W}^s_x = \mathcal{W}^s(x)((-1, 1)^m)$ is tangent to $\mathcal{E}^s_x$,*
3. *$\mathcal{W}^s_x \subseteq \{y \in M\colon \operatorname{dist}(f^n(x), f^n(y)) < \operatorname{dist}(x, y)\rho^n\}$.*

As was pointed out before, the submanifolds $\mathcal{W}^s_x$ are contained in $B_f(N)_\varepsilon$. Moreover, points y in the local basin of attraction that verify $\operatorname{dist}(f^n(x), f^n(y)) < \operatorname{dist}(x, y)\rho^n$ belong to the stable manifold $\mathcal{W}^s_x$. This is the reason why sometimes the theorem is called strong stable manifold theorem: it gives a kind of classification of the dynamical behavior of the points inside the basin of attraction. Here we have to notice that even though the local basin of attraction is foliated by the union of the (strong) stable submanifolds $\mathcal{W}^s_x$, this does not imply that all the points in the attractor converge exponentially to N (see Fig. 3 for the case $N = \mathcal{U}(D)$).

3.3. Spectral projections

In this section we state the basic properties of the spectral projections of matrices, which are constructed by using the Riesz functional calculus. A complete exposition on this theory can be found in Kato's book [20, Ch. 2]. Given $M \in \mathcal{M}_r(\mathbb{C})$ we call $\lambda = \lambda(M) \in \mathbb{C}^r$ its vector of eigenvalues. Let $\sigma(M) = \{\mu_1, \ldots, \mu_k\}$, taking one μ_i for each group of repeated $\lambda_j(M) = \mu_i$ in $\lambda(M)$ (i.e., $k \leqslant r$). Fix $D = \operatorname{diag}(\lambda) \in \mathcal{M}_r(\mathbb{C})$.

Definition 3.6. Given a diagonal matrix $D \in \mathcal{M}_r(\mathbb{C})$, denote by $\lambda = \lambda(D) \in \mathbb{C}^r$ and $\mu = (\mu_1, \dots, \mu_k) \in \mathbb{C}^k$ as before ($\mu_i \neq \mu_j$). Let

1. $\varepsilon_\mu = \frac{1}{3} \min_{i \neq j} |\mu_i - \mu_j|$ and $\Omega_\mu = \bigcup_{1 \leqslant i \leqslant k} B(\mu_i, \varepsilon_\mu)$.
2. $\widetilde{\mathcal{M}}_\mu = \{M \in \mathcal{M}_r(\mathbb{C}) \colon \sigma(M) \subseteq \Omega_\mu\}$, which is open in $\mathcal{M}_r(\mathbb{C})$.
3. Let $E : \widetilde{\mathcal{M}}_\mu \to \mathcal{M}_r(\mathbb{C})^k$ be given by
$$\widetilde{\mathcal{M}}_\mu \ni M \mapsto E(M) = \big(E_1(M), \dots, E_k(M)\big),$$
where $E_i(M) = \aleph_{B(\mu_i, \varepsilon_\mu)}(M)$ is the spectral projection of $M \in \widetilde{\mathcal{M}}_\mu$, associated to $B(\mu_i, \varepsilon_\mu)$.
4. Denote $Q = (Q_1, \dots, Q_k) = E(D)$ and consider the open set
$$\mathcal{M}_\lambda = \big\{M \in \widetilde{\mathcal{M}}_\mu \colon \operatorname{rk}\big(E_i(M)\big) = \operatorname{rk}(Q_i),\ 1 \leqslant i \leqslant k\big\}, \tag{6}$$
which is the connected component of D in $\widetilde{\mathcal{M}}_\mu$.
5. Let $\Pi_E : \mathcal{M}_\lambda \to \mathcal{M}_r(\mathbb{C})$ be given by $\Pi_E(M) = \sum_{i=1}^k \mu_i E_i(M)$, for every $M \in \mathcal{M}_\lambda$.

Remark 3.7. Given $\lambda = \lambda(D) \in \mathbb{C}^r$ and $\mu = (\mu_1, \dots, \mu_k) \in \mathbb{C}^k$ as before, the following properties hold:

1. For every $1 \leqslant i \leqslant k$, $Q_i = Q_i^*$. Also $Q_i Q_j = 0$ (if $i \neq j$) and $\sum_{i=1}^k Q_i = I$. The entries of $E(M)$ for other $M \in \mathcal{M}_\lambda$ satisfy the same properties, but they may be not self-adjoint.
2. Each map E_i (so that the map E) is of class C^∞ in $\mathcal{M}_\lambda$.
3. $E(\mathcal{M}_\lambda) = \mathcal{S}(Q) := \{SQS^{-1} = (SQ_1S^{-1}, \dots, SQ_kS^{-1}) \colon S \in \mathcal{G}l_r(\mathbb{C})\}$.
4. Moreover, if $M \in \mathcal{M}_\lambda$ and $S \in \mathcal{G}l_r(\mathbb{C})$, then $E(SMS^{-1}) = SE(M)S^{-1}$.

Then the map $\Pi_E : \mathcal{M}_\lambda \to \mathcal{M}_r(\mathbb{C})$ satisfies the following properties:

1. It is of class C^∞ on $\mathcal{M}_\lambda$.
2. For every $M \in \mathcal{S}(D)$, we have that $\Pi_E(M) = M$.
3. $\Pi_E(\mathcal{M}_\lambda) = \mathcal{S}(D)$, and the spectral radius satisfies $\rho(M - \Pi_E(M)) < \varepsilon_\mu$ for every $M \in \mathcal{M}_\lambda$.
4. If $M \in \mathcal{M}_\lambda$ and $S \in \mathcal{G}l_r(\mathbb{C})$, then $\Pi_E(SMS^{-1}) = S\Pi_E(M)S^{-1}$.

Remark 3.8. With the previous notations, for every $M \in \mathcal{M}_\lambda$, we consider the subspace

$$\mathcal{A}_M = \big\{B \in \mathcal{M}_r(\mathbb{C}) \colon BE_i(M) = E_i(M)B,\ 1 \leqslant i \leqslant k\big\}, \tag{7}$$

of *block diagonal* matrices, with respect to $E(M)$. It is easy to see that $\mathcal{A}_M = \ker(d\Pi_E)_M$ and $R((d\Pi_E)_M) = T_N\mathcal{S}(D)$, where $N = \Pi_E(M) \in \mathcal{S}(D)$. Then by Eq. (5)

$$\mathcal{M}_r(\mathbb{C}) = \mathcal{A}_N \oplus T_N\mathcal{S}(D) = \mathcal{A}_M \oplus T_N\mathcal{S}(D),$$

and the sum becomes orthogonal if $M \in \mathcal{U}(D)$. Fix $M \in \mathcal{S}(D)$. Since $\Pi_E^2 = \Pi_E$, then $(d\Pi_E)_M$ is the projector with kernel $\mathcal{A}_M$ and image $T_M\mathcal{S}(D)$. Note that

$$\mathcal{A}_M = \{M\}' := \big\{B \in \mathcal{M}_r(\mathbb{C}):\ MB = BM\big\} \quad \text{for every } M \in \mathcal{S}(D), \tag{8}$$

since in this case $M = \Pi_E(M)$.

4. Convergence of the sequence $\{\Delta^n(T)\}_{n\in\mathbb{N}}$

4.1. The derivative of Δ in $\mathcal{M}_r(\mathbb{C})$

Let $N \in \mathcal{M}_r(\mathbb{C})$ be an invertible normal matrix. Theorem 2.1 gives a description of the action of $d\Delta_N$ on $T_N\mathcal{S}(N)$. By Remark 3.4, in order to prove Theorem 2.3, it is enough to describe the action of $d\Delta_N$ on its orthogonal complement, i.e., the subspace $\mathcal{A}_N$ described in Remark 3.8.

Proposition 4.1. *Let $D = \operatorname{diag}(d_1, \ldots, d_r) \in \mathcal{G}l_r(\mathbb{C})$ be a diagonal matrix with k different eigenvalues. Fix $N \in \mathcal{U}(D)$ and consider the subspace $\mathcal{A}_N \subseteq \mathcal{M}_r(\mathbb{C})$ defined in Eq.* (7)*. Then $d\Delta_N|_{\mathcal{A}_N} = I_{\mathcal{A}_N}$.*

Proof. If $N \in \mathcal{U}(D)$, then N is normal, and $\mathcal{A}_N = \{N\}'$. Hence, if $Y \in \mathcal{A}_N$ is normal, then $N + tY$ is also normal for every $t \in \mathbb{R}$ (see for example p. 103 of [16]). Therefore

$$d\Delta_N(Y) = \frac{d}{dt}\Delta(N + tY)|_{t=0} = \frac{d}{dt}(N + tY)|_{t=0} = Y.$$

On the other hand, since $\mathcal{A}_N$ is closed under taking adjoints, if $X \in \mathcal{A}_N$, then $A = \frac{X+X^*}{2}$ and $B = \frac{X-X^*}{2}$ are still in $\mathcal{A}_N$. Therefore $d\Delta_N(X) = d\Delta_N(A + B) = A + B = X$. $\square$

Now we can restate and prove the following result:

Theorem 2.3. *Let $D \in \mathcal{M}_r(\mathbb{C})$ be an invertible diagonal matrix. For every $N \in \mathcal{U}(D)$, the subspaces $F_N = \mathcal{A}_N \oplus T_N\mathcal{U}(D)$ and $\mathcal{E}^s_N$ (defined in Theorem* 2.1*) satisfy that*

1. $\mathcal{M}_r(\mathbb{C}) = F_N \oplus \mathcal{E}^s_N$.
2. *Both subspaces are $d\Delta_N$ invariant.*
3. $d\Delta_N|_{F_N} = I_{F_N}$ *and* $\|d\Delta_N|_{\mathcal{E}^s_N}\| \leqslant k_D < 1$*, where k_D is defined as in Theorem* 2.1.
4. *The distributions $N \mapsto F_N$ and $N \mapsto \mathcal{E}^s_N$ are smooth.*

Proof. Using Theorem 2.1, and the descriptions given in Eqs. (5) and (7) we can easily deduce that $\mathcal{M}_r(\mathbb{C}) = \mathcal{A}_N \oplus T_N\mathcal{S}(N) = \mathcal{A}_N \oplus T_N\mathcal{U}(N) \oplus \mathcal{E}^s_N$ for every $N \in \mathcal{U}(D)$. The other statements follow directly from Proposition 4.1 and Theorem 2.1. $\square$

Remark 4.2. With the notations of Theorem 2.3, the subspaces F_N and $\mathcal{E}^s_N$ can be characterized by means of the functional calculus applied to the linear maps $d\Delta_N$, for every $N \in \mathcal{U}(D)$. Indeed,

$$F_N = R\big(\aleph_{B(1,\varepsilon)}(d\Delta_N)\big) \quad \text{and} \quad \mathcal{E}^s_N = R\big(\aleph_{B(0,k_D+\varepsilon)}(d\Delta_N)\big),$$

for every $\varepsilon > 0$ sufficiently small. In particular, this implies that the distribution of subspaces F_N and $\mathcal{E}^s_N$ can be extended smoothly to an open neighborhood of $\mathcal{U}(D)$.

4.2. Strong stable manifolds

Let $\mathbb{P} \subseteq \mathcal{M}_r(\mathbb{C})$ be a compact set of fixed points for Δ, i.e., a compact set of normal matrices. Recall that its basin of attraction is the set

$$B_\Delta(\mathbb{P}) = \bigl\{T \in \mathcal{M}_r(\mathbb{C}):\ \operatorname{dist}\bigl(\Delta^n(T), \mathbb{P}\bigr) \xrightarrow[n\to\infty]{} 0\bigr\}$$

and, for every $\varepsilon > 0$, the local basin is the set

$$B_\Delta(\mathbb{P})_\varepsilon = \bigl\{T \in B_\Delta(\mathbb{P}):\ \operatorname{dist}\bigl(\Delta^n(T), \mathbb{P}\bigr) < \varepsilon,\ n \in \mathbb{N}\bigr\}.$$

In this subsection, using the strong stable manifold Theorem 3.5, we shall prove that, if $\mathbb{P}$ has a distribution of subspaces with good properties (like the distribution of Theorem 2.3 for $\mathbb{P} = \mathcal{U}(D)$), then through each $T \in B_\Delta(\mathbb{P})$ closed enough to $\mathbb{P}$ there is a stable manifold $\mathcal{W}_T^{ss}$ with the property

$$\mathcal{W}_T^{ss} \subseteq \bigl\{B \in \mathcal{M}_r(\mathbb{C}):\ \bigl\|\Delta^n(T) - \Delta^n(B)\bigr\| < C\gamma^n\bigr\},$$

where $\gamma < 1$ and C is a positive constant. With this aim, firstly we need to extend the distribution of subspaces given on $\mathbb{P}$ to some local basin. This extension is a quite standard procedure in dynamical systems. For completeness, we include a sketch of its proof (adapted to our case) in Appendix B.3.

Proposition 4.3. *Let $\mathbb{P}$ be a compact set consisting of fixed points of Δ. Suppose that, for every $N \in \mathbb{P}$, there are subspaces $\mathcal{E}_N^s$ and $\mathcal{F}_N$ of $\mathcal{M}_r(\mathbb{C})$ with the following properties:*

1. $\mathcal{M}_r(\mathbb{C}) = \mathcal{E}_N^s \oplus \mathcal{F}_N$.
2. *The distributions $N \mapsto \mathcal{E}_N^s$ and $N \mapsto \mathcal{F}_N$ are continuous.*
3. *There exists $\rho \in (0,1)$ which does not depend on N such that*

$$\|d\Delta_N|_{\mathcal{E}_N^s}\| < 1 - \rho \quad \text{and} \quad \bigl\|(I - d\Delta_N)|_{\mathcal{F}_N}\bigr\| < \frac{\rho}{2}. \tag{9}$$

4. *Both subspaces $\mathcal{E}_N^s$ and $\mathcal{F}_N$ are $d\Delta_N$ invariant.*

Then, there exists $\varepsilon > 0$ such that the distributions $N \mapsto \mathcal{E}_N^s$ and $N \mapsto \mathcal{F}_N$ can be extended to the local basin $B_\Delta(\mathbb{P})_\varepsilon$, verifying conditions 1, 2, 3 and the following new condition:

4′. *For every $T \in B_\Delta(\mathbb{P})_\varepsilon$, the subspace $\mathcal{E}_T^s$ is $d\Delta_T$-invariant, i.e., $d\Delta_T(\mathcal{E}_T^s) \subseteq \mathcal{E}_{\Delta(T)}^s$.*

Remark 4.4. As we pointed out in Section 2, to prove the convergence it is enough to establish the above result for $\mathbb{P} = \mathcal{U}(D)$. We include a more general statement because it will be useful to study the regularity of the limit function.

Now we are ready to state and prove the announced result on stable manifolds.

Proposition 4.5. *Let $\mathbb{P}$ be a compact set consisting of fixed points of Δ, with two distributions $N \mapsto \mathcal{E}^s_N$ and $N \mapsto \mathcal{F}_N$ which satisfy the hypothesis of Proposition* 4.3*. Then, there exist $\varepsilon > 0$ and a C^2-pre-lamination $\mathcal{W}^s : B_\Delta(\mathbb{P})_\varepsilon \to \mathrm{Emb}^2((-1,1)^m, B_\Delta(\mathbb{P}))$ (endowed with the C^2-topology) of class C^0 such that, for every $T \in B_\Delta(\mathbb{P})_\varepsilon$,*

1. *$\mathcal{W}^s(T)(0) = T$.*
2. *If $\mathcal{W}^{ss}_T$ is the submanifold $\mathcal{W}^s(T)((-1,1)^m)$, then $T_T\mathcal{W}^{ss}_T = \mathcal{E}^s_T$.*
3. *There are constants $\gamma < 1$ and $C > 0$ such that*

$$\mathcal{W}^{ss}_T \subseteq \bigl\{B \in \mathcal{M}_r(\mathbb{C}):\ \bigl\|\Delta^n(T) - \Delta^n(B)\bigr\| < C\gamma^n\bigr\}. \tag{10}$$

Proof. By Proposition 4.3, the distributions $\mathbb{P} \ni N \mapsto \mathcal{E}^s_N, \mathcal{F}_N$ can be extended to a local basin $B_\Delta(\mathbb{P})_\varepsilon$, satisfying the hypothesis of Theorem 3.5. Note that the condition $\|(I - d\Delta_T)|_{\mathcal{F}_T}\| < \frac{\rho}{2}$ implies that $\|d\Delta_T(Y)\| > (1 - \frac{\rho}{2})\|Y\|$ for every $Y \in \mathcal{F}_T$. $\square$

4.3. The case $\mathbb{P} = \mathcal{U}(D)$

Let $D \in \mathcal{M}_r(\mathbb{C})$ be an invertible diagonal matrix. In this section we shall consider the compact invariant set $\mathbb{P} = \mathcal{U}(D)$. Note that the distributions $N \mapsto F_N$ and $N \mapsto \mathcal{E}^s_N$ ($N \in \mathbb{P}$) given by Theorem 2.3 clearly verify the hypothesis of Propositions 4.3 and 4.5. Thus, by Proposition 4.5, there exist a continuous pre-lamination $\mathcal{W}^s : B_\Delta(\mathbb{P})_\varepsilon \to \mathrm{Emb}^2((-1,1)^k, B_\Delta(\mathbb{P}))$ and submanifolds $\mathcal{W}^{ss}_T$ (for every $T \in B_\Delta(\mathbb{P})_\varepsilon$), which will be also used throughout this section. In this setting we can give simple characterizations of the basins of $\mathbb{P}$. Indeed, by Proposition 1.1 and Remark 3.2,

$$B_\Delta\bigl(\mathcal{U}(D)\bigr) = \bigl\{T \in \mathcal{M}_r(\mathbb{C}):\ \lambda(T) = \lambda(D)\bigr\}. \tag{11}$$

Given $T \in B_\Delta(\mathcal{U}(D))$, we denote $d_n(T) = \|\Delta^n(T)\|_2^2 - \|D\|_2$, for every $n \in \mathbb{N}$. By Remark 3.2, we know that $\mathrm{dist}(\Delta^n(T), \mathcal{U}(N)) \leqslant d_n(T) \underset{n\to\infty}{\searrow} 0$. Therefore

$$\bigl\{T \in B_\Delta\bigl(\mathcal{U}(D)\bigr):\ d_1(T) < \varepsilon\bigr\} \subseteq B_\Delta\bigl(\mathcal{U}(D)\bigr)_\varepsilon, \tag{12}$$

and it is also an open neighborhood of $\mathcal{U}(D)$ in $B_\Delta(\mathcal{U}(D))$. Therefore, if $T \in B_\Delta(\mathcal{U}(D))$ is close enough to $\mathcal{U}(D)$, then $T \in B_\Delta(\mathcal{U}(D))_\varepsilon$ (and we do not need to check that $\mathrm{dist}(\Delta^n(T), \mathcal{U}(N))$ for $n > 1$). Note that if $T \in B_\Delta(\mathcal{U}(D))$, despite the equality $\lambda(T) = \lambda(D)$, the matrix T can have any Jordan form.

4.4. The sets $\mathcal{O}_D$

In this subsection we identify some convenient sets of matrices where the iterated Aluthge transform sequence converges (possibly slowly). By their properties, these sets will play a key role in the proof of the convergence of the iterated Aluthge transform sequence. Let D be an invertible diagonal matrix, $\lambda = \lambda(D)$, and $\Pi_E : \mathcal{M}_\lambda \to \mathcal{S}(D)$ the map defined in Section 2.4. If $\mathbb{P} = \mathcal{U}(D)$, consider the following subset of $B_\Delta(\mathbb{P})$:

$$\mathcal{O}_D = \bigl\{T \in B_\Delta(\mathbb{P}):\ \Pi_E(T) \in \mathbb{P}\bigr\} = \Pi_E^{-1}(\mathbb{P}) \cap B_\Delta(\mathbb{P}). \tag{13}$$

Note that, if $T \in \mathcal{O}_D$, then the system of projectors $E(T)$ is orthogonal. Hence we get the following simple consequence.

Proposition 4.6. *If $T \in \mathcal{O}_D$, then $\Delta^n(T) \xrightarrow[n\to\infty]{} \Pi_E(T) \in \mathcal{U}(D)$.*

Proof. If $T \in \mathcal{O}_D \subseteq B_\Delta(\mathcal{U}(D))$ then $\lambda(T) = \lambda(D)$, by Eq. (11). On the other hand, if $N = \Pi_E(T)$, then $E(T) = E(N)$ is an **orthogonal** system of projectors, and $T \in \mathcal{A}_N$, the subspace defined in Eq. (7). Write $T = T_1 \oplus \cdots \oplus T_k$, where $T_i = T|_{R(E_i(T))}$. Then, by the properties of the Aluthge transform, $\Delta^n(T) = \Delta^n(T_1) \oplus \cdots \oplus \Delta^n(T_k)$, for every $n \in \mathbb{N}$. Since $\lambda(T) = \lambda(D)$, then each $\sigma(T_i) = \{\mu_i\}$. Hence, by Proposition 1.1,

$$\Delta^n(T_i) \xrightarrow[n\to\infty]{} \mu_i I_{R(E_i(T))}, \quad 1 \leqslant i \leqslant k.$$

Therefore $\Delta^n(T) \xrightarrow[n\to\infty]{} \sum_{i=1}^k \mu_i E_i(T) = \Pi_E(T)$. $\square$

Another important characteristic of the sets $\mathcal{O}_D$ is that each element of $B_\Delta(\mathbb{P})$ "close enough" to $\mathbb{P}$ is exponentially attracted towards $\mathcal{O}_D$. This property is precisely described in the following statement.

Proposition 4.7. *Let $D \in \mathcal{M}_r(\mathbb{C})$ be an invertible diagonal matrix, $\mathbb{P} = \mathcal{U}(D)$ and $\mathcal{W}^s : B_\Delta(\mathbb{P})_\varepsilon \to \mathrm{Emb}^2((-1,1)^k, B_\Delta(\mathbb{P}))$ the pre-lamination given by Proposition 4.5. Then, there exists $\eta < \varepsilon$ such that $\mathcal{W}_T^{ss} \cap \mathcal{O}_D \neq \emptyset$ for every $T \in B_\Delta(\mathbb{P})_\eta$.*

The proof is rather long and technical, so it is divided in several parts. Throughout the rest of this subsection we shall use the following notations: If $D \in \mathcal{M}_r(\mathbb{C})$ is an invertible diagonal matrix then $\mathbb{P} = \mathcal{U}(D)$ and, for every $T \in B_\Delta(\mathbb{P})_\varepsilon$, by means of $\mathcal{W}_T^{ss} = \mathcal{W}^{ss}(T) : (-1,1)^m \to B_\Delta(\mathbb{P})$ we denote the maps given by Proposition 4.5. For the sake of simplicity, for every $t > 0$, $\mathcal{Q}_t$ denotes the m-dimensional cube $(-t,t)^m$. The invariant manifolds will be denoted by $\mathcal{W}_T^{ss}(\mathcal{Q}_1)$.

Note that $\mathcal{W}_T^{ss}(\mathcal{Q}_1)$ intersects $\mathcal{O}_D$ if and only if $\Pi_E(\mathcal{W}_T^{ss}(\mathcal{Q}_1))$ intersects $\mathbb{P}$. The proof of Proposition 4.7 uses this remark and it is based on some well-known results about transversal intersections, using that $\Pi_E(\mathcal{W}_T^{ss}(\mathcal{Q}_1))$ is "C^2-close" to another manifold ($\mathcal{W}_N^{ss}(\mathcal{Q}_1)$ for some $N \in \mathbb{P}$ near T) which intersecs transversally $\mathbb{P}$, both contained in $\mathcal{S}(D)$. We give a proof adapted to our case, divided into three lemmas: We begin with the following classical result (see for example [14, p. 36]).

Lemma 4.8. *Let $U \subseteq \mathbb{R}^m$ be an open set and $W \subseteq U$ an open set with compact closure $\overline{W} \subseteq U$. Let $M \subseteq \mathbb{R}^n$ be a smooth submanifold and $f : U \to M$ a C^1 embedding. There exists $\varepsilon > 0$ such that, if*

$$g : U \to M \text{ is } C^1, \qquad \|dg_x - df_x\| < \varepsilon \quad \text{and} \quad \|g(x) - f(x)\| < \varepsilon$$

for every $x \in W$, then $g|_W$ is an embedding.

Lemma 4.9. *Let D and $\mathbb{P}$ be as in Proposition* 4.7. *Then there is $\eta < \varepsilon$ such that the map*

$$\mathcal{V} : B_\Delta(\mathbb{P})_\eta \to \mathrm{Emb}^2\big(\mathcal{Q}_{\frac{1}{2}}, \mathcal{S}(D)\big), \quad \text{given by } \mathcal{V}_T = \Pi_E \circ \mathcal{W}_T^{ss}\big|_{\mathcal{Q}_{\frac{1}{2}}},$$

is well defined and continuous with respect to the C^2 topology of $\mathrm{Emb}^2(\mathcal{Q}_{\frac{1}{2}}, \mathcal{S}(D))$.

Proof. Consider the map $\widetilde{\mathcal{V}} : B_\Delta(\mathbb{P})_\varepsilon \to C^2(\mathcal{Q}_1, \mathcal{S}(D))$ given by $\widetilde{\mathcal{V}}_T = \Pi_E \circ \mathcal{W}_T^{ss}$. By Proposition 4.5 and Remark 3.7, $\widetilde{\mathcal{V}}$ is well defined and continuous, if $C^2(\mathcal{Q}_1, \mathcal{S}(D))$ is endowed with the C^2 topology. Note that $\mathcal{W}_N^{ss}$ takes values in $\mathcal{S}(D)$ for every $N \in \mathbb{P}$. Indeed, this follows by Corollary 3.1.2 of [7], or by rewriting the proof of Proposition 4.5 inside $\mathcal{S}(D)$ in this case. Therefore, $\widetilde{\mathcal{V}}_N = \mathcal{W}_N^{ss}$ for every $N \in \mathbb{P}$, because the map Π_E acts as the identity on the manifold $\mathcal{S}(D)$.

Given $N \in \mathbb{P}$, Lemma 4.8 assures that there exists ε_N such that, if $\mathcal{T} : \mathcal{Q}_1 \to \mathcal{M}_r(\mathbb{C})$ is a C^1 map which satisfies that

$$\big\| \big(d\mathcal{W}_N^s\big)_x - d\mathcal{T}_x \big\| < \varepsilon_N \quad \text{and} \quad \big\| \mathcal{W}_N^s(x) - \mathcal{T}(x) \big\| < \varepsilon_N, \tag{14}$$

for every $x \in \mathcal{Q}_{\frac{1}{2}}$, then $\mathcal{T}|_{\mathcal{Q}_{\frac{1}{2}}}$ is an embedding. By the continuity of $\widetilde{\mathcal{V}}$, there is a neighborhood $\mathcal{U}_N$ of N in $B_\Delta(\mathbb{P})_\varepsilon$ such that, for every $T \in \mathcal{U}_N$, the map $\widetilde{\mathcal{V}}_T$ satisfies (14). Take $\eta > 0$ such that $B_\Delta(\mathbb{P})_\eta \subseteq \bigcup_{N \in \mathbb{P}} \mathcal{U}_N$. Then, $\mathcal{V}(T) = \widetilde{\mathcal{V}}_T|_{\mathcal{Q}_{\frac{1}{2}}} \in \mathrm{Emb}^2(\mathcal{Q}_{\frac{1}{2}}, \mathcal{S}(D))$, for every $T \in B_\Delta(\mathbb{P})_\eta$, i.e., $\mathcal{V}$ is well defined. The continuity of $\mathcal{V}$ follows from the fact that both $\widetilde{\mathcal{V}}$ and the restriction map $\mathcal{T} \mapsto \mathcal{T}|_{\mathcal{Q}_{\frac{1}{2}}}$ are continuous with respect to the C^2 topology. □

Lemma 4.10. *Let D and $\mathbb{P}$ be as in Proposition* 4.7. *Given $N_0 \in \mathbb{P}$ and $\varepsilon > 0$, there exists a C^2-neighborhood Ω of $\mathcal{W}_{N_0}^{ss}|_{\mathcal{Q}_{\frac{1}{2}}}$ in the space* $\mathrm{Emb}^2(\mathcal{Q}_{\frac{1}{2}}, \mathcal{S}(D))$, *such that $\mathcal{T}(\mathcal{Q}_{\frac{1}{2}})$ intersects the submanifold $\mathbb{P}$ at a point $N \in B(N_0, \varepsilon)$, for every $\mathcal{T} \in \Omega$.*

Proof. Let $(\mathcal{U}_{N_0}, \varphi)$ be a chart in $\mathcal{S}(D)$ such that $N_0 \in \mathcal{U}_{N_0} \subseteq B(\varepsilon, N_0)$, $\varphi(N_0) = 0$, and $\varphi(\mathbb{P} \cap \mathcal{U}_{N_0}) = \varphi(\mathcal{U}_{N_0}) \cap (\{0\} \oplus \mathbb{R}^{n-m})$, where $\mathbb{R}^n \simeq \mathbb{R}^m \oplus \mathbb{R}^{n-m}$. Let P denote the orthogonal projection from $\mathbb{R}^n$ onto $\mathbb{R}^m \oplus \{0\}$.

By Proposition 4.5, the intersection $\mathcal{W}_{N_0}^{ss}(\mathcal{Q}_{\frac{1}{2}}) \cap \mathcal{U}(D) = \{N_0\}$ is transversal. Then, there exist $\delta \in (0, 1/2)$ and a C^2-neighborhood Ω_0 of $\mathcal{W}_{N_0}^{ss}|_{\mathcal{Q}_{\frac{1}{2}}}$ in $\mathrm{Emb}^2(\mathcal{Q}_{\frac{1}{2}}, \mathcal{S}(D))$ such that, for every $\mathcal{T} \in \Omega_0$,

1. $\mathcal{T}(\overline{\mathcal{Q}_\delta}) \subseteq \mathcal{U}_{N_0}$.
2. $\ker P \oplus d\widetilde{\mathcal{T}}_M(\mathcal{Q}_\delta) = \mathbb{R}^n$, where $\widetilde{\mathcal{T}} = \varphi \circ \mathcal{T}|_{\mathcal{Q}_\delta}$ and $M \in \widetilde{\mathcal{T}}(\mathcal{Q}_\delta)$.
3. The angle between $\ker P$ and $d\widetilde{\mathcal{T}}_M(\mathcal{Q}_\delta)$ is uniformly bounded from below.

Note that items 2 and 3 imply that $\mathbb{R}^k \oplus \{0\}$. On the other hand, item 2 also implies that for every $\mathcal{T} \in \Omega_0$ and every $M \in \widetilde{\mathcal{T}}(\mathcal{Q}_\delta)$, the linear map P acting on $d\widetilde{\mathcal{T}}_M(\mathcal{Q}_\delta)$ is uniformly bounded from below. Since the norm of the second derivative of $P \circ \widetilde{\mathcal{T}}$ is bounded on $\mathcal{Q}_{\frac{\delta}{2}}$, there exists $\mu > 0$ so that, for every $M \in \widetilde{\mathcal{T}}(\mathcal{Q}_{\frac{\delta}{2}})$,

$$B\big(P(M), \mu\big) \subseteq P\big(\widetilde{\mathcal{T}}(\mathcal{Q}_\delta)\big). \tag{15}$$

Take $\Omega \subseteq \Omega_0$ such that $\|\widetilde{\mathcal{W}}^{ss}_{N_0}(x) - \widetilde{\mathcal{T}}(x)\| < \mu/2$ for every $\mathcal{T} \in \Omega$ and every $x \in \mathcal{Q}_\delta$, where $\widetilde{\mathcal{W}}^{ss}_{N_0} = \varphi \circ \mathcal{W}^{ss}_{N_0}|_{\mathcal{Q}_\delta}$. As $\widetilde{\mathcal{W}}^{ss}_{N_0}(0) = 0$, Eq. (15) implies that $0 \in P(\widetilde{\mathcal{T}}(\mathcal{Q}_\delta))$, for every $\mathcal{T} \in \Omega$. Thus $\mathcal{T} \cap \mathcal{U}(D) \cap \mathcal{U}_{N_0} \neq \emptyset$, because $\mathcal{T}(\overline{\mathcal{Q}_\delta}) \subseteq \mathcal{U}_{N_0}$. In particular, $\mathcal{T}(\mathcal{Q}_{\frac{1}{2}})$ intersects the submanifold $\mathbb{P}$ transversally at a point $N \in \mathcal{U}_{N_0} \subseteq B(N_0, \varepsilon)$. $\square$

Proof of Proposition 4.7. Given $N \in \mathbb{P}$, Lemma 4.10 assures that there is a C^2-neighborhood Ω_N of $\mathcal{W}^{ss}_N|_{\mathcal{Q}_{\frac{1}{2}}}$ in $\mathrm{Emb}^2(\mathcal{Q}_{\frac{1}{2}}, \mathcal{S}(D))$ such that $\mathcal{T}(\mathcal{Q}_{\frac{1}{2}}) \cap \mathbb{P} \neq \emptyset$ for every $\mathcal{T} \in \Omega_N$. Let $\mathcal{V}$ be the function defined in Lemma 4.9, and let $\mathcal{U}_N = \mathcal{V}^{-1}(\Omega_N)$. Define $\mathcal{U}_{\mathbb{P}} = \bigcup_{N \in \mathbb{P}} \mathcal{U}_N$. Therefore, $\mathcal{U}_{\mathbb{P}}$ is an open neighborhood of $\mathbb{P}$ contained in $B_\Delta(\mathbb{P})$. Since $\mathbb{P}$ is compact, there exists $0 < \eta < \varepsilon$ such that $B_\Delta(\mathbb{P})_\eta \subseteq \mathcal{U}_{\mathbb{P}}$. Then, for every $T \in B_\Delta(\mathbb{P})_\eta$, $\Pi_E(\mathcal{W}^{ss}_T(\mathcal{Q}_{\frac{1}{2}}))$ intersects $\mathbb{P}$. By Proposition 4.5, $\mathcal{W}^{ss}_T(\mathcal{Q}_{\frac{1}{2}}) \subseteq B_\Delta(\mathbb{P})$. Therefore $\mathcal{W}^{ss}_T(\mathcal{Q}_{\frac{1}{2}}) \cap \mathcal{O}_D \neq \emptyset$. $\square$

The proof of the next result, which is used in the proof of the continuity of the limit function Δ^∞, follows the same lines as the proof of Lemma 4.9.

Lemma 4.11. *Let $N_0 \in \mathcal{M}_r(\mathbb{C})$ be a normal matrix, $\mathbb{P}_\beta$ as in Definition* 5.4 *and $\mathcal{W}^{ss}: B_\Delta(\mathbb{P}_\beta)_\varepsilon \to \mathrm{Emb}^2(\mathcal{Q}_1, B_\Delta(\mathbb{P}))$ the pre-lamination given by Proposition* 4.5. *If Π_E is defined with respect to the spectrum of N_0, then there exists $\eta < \beta$ so that the map*

$$\mathcal{V} : B_\Delta(\mathbb{P}_\eta)_\eta \to \mathrm{Emb}^2\big(\mathcal{Q}_{\frac{1}{2}}, \mathcal{S}(N_0)\big)$$

given by $\mathcal{V}_T = \Pi_E \circ \mathcal{W}^{ss}_T|_{\mathcal{Q}_{\frac{1}{2}}}$ is well defined and continuous with the C^2 topology of $\mathrm{Emb}^2(\mathcal{Q}_{\frac{1}{2}}, \mathcal{S}(N_0))$.

4.5. The proof of Jung, Ko and Pearcy's conjecture

Now, we are in position to prove the main result of this paper:

Theorem 4.12. *For every $T \in \mathcal{M}_r(\mathbb{C})$, the sequence $\Delta^n(T)$ converges.*

Proof. By Corollary 4.16 of [6], we can assume that $T \in \mathcal{G}l_r(\mathbb{C})$. Let $D \in \mathcal{G}l_r(\mathbb{C})$ be a diagonal matrix such that $\lambda(T) = \lambda(D)$, and let $\mathbb{P} = \mathcal{U}(D)$. By Eq. (11), $T \in B_\Delta(\mathbb{P})$. By Eq. (12), replacing T by $\Delta^n(T)$ for some n large enough, we can assume that $T \in B_\Delta(\mathbb{P})_\rho$, for any fixed $\rho > 0$.

Consider now the stable manifolds $\mathcal{W}^{ss}_{T'}$ constructed in Proposition 4.5, for every $T' \in B_\Delta(\mathbb{P})_\varepsilon$. By Proposition 4.7, there exists $0 < \eta < \varepsilon$ such that $\mathcal{W}^{ss}_{T'} \cap \mathcal{O}_D \neq \emptyset$, for every $T' \in B_\Delta(\mathbb{P})_\eta$. By the previous discusion, we can assume that our $T \in B_\Delta(\mathbb{P})_\eta$. So, there exists $M \in \mathcal{W}^{ss}_T \cap \mathcal{O}_D$. Then, by Proposition 4.6 and Eq. (10) of Proposition 4.5, we deduce that $\Pi_E(M) = \lim_{n\to\infty} \Delta^n(M) = \lim_{n\to\infty} \Delta^n(T)$. $\square$

5. Regularity of the map Δ^∞

Given $T \in \mathcal{M}_r(\mathbb{C})$ we denote $\Delta^\infty(T) = \lim_{n\to\infty} \Delta^n(T)$, which is a normal matrix. Note that the map $\Delta^\infty : \mathcal{M}_r(\mathbb{C}) \to \mathcal{N}(r)$ is a retraction. In this section we study the regularity of this retraction.

5.1. Differentiability vs. continuity

In [7] we proved that the map Δ^∞ is of class C^∞ when restricted to the open dense set of those matrices in $\mathcal{M}_r(\mathbb{C})$ with r different eigenvalues. The following proposition shows that this cannot be extended globally to the set of all matrices.

Proposition 5.1. *The map Δ^∞ cannot be C^1 in a neighborhood of the identity.*

Proof. Suppose that Δ^∞ is C^1 in a neighborhood of the identity. By the same argument used in the proof of Proposition 4.1, it follows that $d\Delta_I^\infty$ is the identity map (in this case, $\mathcal{A}_I = \mathcal{M}_r(\mathbb{C})$). This implies that Δ^∞ is a local diffeomorphism. However, this is impossible because it takes values in the set of normal operators. $\square$

Remark 5.2. The map Δ also fails to be differentiable at some matrices $T \in \mathcal{M}_r(\mathbb{C}) \setminus \mathcal{G}l_r(\mathbb{C})$. Indeed, suppose that Δ were differentiable at $T = 0$. In this case, given $X \in \mathcal{M}_r(\mathbb{C})$,

$$d\Delta_0(X) = \frac{d}{dt}\Delta(tX)|_{t=0} = \frac{d}{dt}t\Delta(X)|_{t=0} = \Delta(X).$$

But this is impossible, because the map $X \mapsto \Delta(X)$ is not linear. Using (1), this fact can be easily extended to any $T \in \mathcal{M}_r(\mathbb{C}) \setminus \mathcal{G}l_r(\mathbb{C})$ such that $\ker T$ is orthogonal to $R(T)$ (for example, every non-invertible normal matrix).

5.2. Continuity of Δ^∞ on $\mathcal{G}l_r(\mathbb{C})$

For the sake of convenience, throughout this subsection we shall use the spectral norm, instead of the Frobenius norm, to measure distances in $\mathcal{M}_r(\mathbb{C})$.

Remark 5.3. Since Δ^∞ is the limit of continuous maps and it is a retraction, in order to show that it is continuous on $\mathcal{G}l_r(\mathbb{C})$, it is enough to prove the continuity at the normal matrices of $\mathcal{G}l_r(\mathbb{C})$. Indeed, we have that Δ^n is continuous for every $n \in \mathbb{N}$. Then, for every $T \in \mathcal{G}l_r(\mathbb{C})$, and every neighborhood $\mathcal{W}$ of $\Delta^\infty(T)$, there exists $n \in \mathbb{N}$ and a neighborhood $\mathcal{U}$ of T such that $\Delta^n(\mathcal{U}) \subseteq \mathcal{W}$. Then, note that $\Delta^\infty \circ \Delta^n = \Delta^\infty$.

From now on, let $N_0 \in \mathcal{N}(r)$ be a fixed normal invertible matrix such that $\lambda(N_0) = \lambda$ and $\sigma(N_0) = (\mu_1, \ldots, \mu_k)$. Let $\varepsilon_\mu = \frac{1}{3}\min_{i \neq j}|\mu_i - \mu_j|$. Consider the open set $\mathcal{M}_\lambda$ defined in Eq. (6). Recall that, for $T \in \mathcal{M}_\lambda$, we call $\Pi_E(T) = \sum_{i=1}^k \mu_i E_i(T) \in \mathcal{S}(N_0)$.

Definition 5.4. With the previous notations, let

1. $\mathcal{M}_{\lambda,\eta}$ the open subset of $\mathcal{M}_\lambda$ obtained in the same way, but by replacing ε_μ by $\eta \in (0, \varepsilon_\mu)$. Note that $\rho(T - \Pi_E(T)) < \eta$ for every $T \in \mathcal{M}_{\lambda,\eta}$.
2. Given $\beta > 0$, let $\mathbb{P}_\beta = \{N \in \mathcal{N}(r)\colon \operatorname{dist}(N, \mathcal{U}(N_0)) \leqslant \beta\}$. Note that $\mathbb{P}_\beta$ is compact.

Lemma 5.5. *With the previous notations, let $\beta > 0$ be such that the closed ball $\overline{B(N_0, \beta)}$ is contained in $\mathcal{M}_\lambda$. Then*

1. $\mathbb{P}_\beta \subseteq \mathcal{M}_\lambda$.
2. *If* $\eta < \min\{\beta, \varepsilon_\mu\}$, *then* $\mathcal{M}_{\lambda,\eta} \subseteq B_\Delta(\mathbb{P}_\beta)$.
3. *Moreover, if* $T \in \mathcal{M}_{\lambda,\eta}$ *and* $N = \Delta^\infty(T)$, *then* $N \in \mathcal{M}_{\lambda,\eta}$ *and* $\|N - \Pi_E(N)\| < \eta$.

Proof. If $N \in \mathbb{P}_\beta$, let $U \in \mathcal{U}(r)$ such that $\|N - UN_0U^*\| \leqslant \beta$ (recall that $\mathcal{U}(N_0)$ is compact). Then $U^*NU \in \overline{B(N_0,\beta)} \subseteq \mathcal{M}_\lambda$, so that also $N \in \mathcal{M}_\lambda$.

Let $T \in \mathcal{M}_{\lambda,\eta}$ and denote $N = \Delta^\infty(T)$. Since $\lambda(N) = \lambda(T)$, then also $N \in \mathcal{M}_{\lambda,\eta}$. Since N is normal, we conclude that $\Pi_E(N) \in \mathcal{U}(N_0)$ and

$$\eta > \rho\big(N - \Pi_E(N)\big) = \big\|N - \Pi_E(N)\big\|,$$

because N commutes with $\Pi_E(N)$, so that $N - \Pi_E(N)$ is normal. Therefore, $N \in \mathbb{P}_\eta \subseteq \mathbb{P}_\beta$ and $T \in B_\Delta(\mathbb{P}_\beta)$. □

Theorem 5.6. *The map* Δ^∞ *is continuous on* $\mathcal{G}l_r(\mathbb{C})$.

Proof. By Remark 5.3, it is enough to prove continuity at the normal matrices of $\mathcal{G}l_r(\mathbb{C})$. Fix $N_0 \in \mathcal{G}l_r(\mathbb{C})$ a normal matrix. Let $\varepsilon > 0$, such that $B(N_0,\varepsilon) \subseteq \mathcal{G}l_r(\mathbb{C})$. We shall use the notations of the previous statements relative to N_0. For $\beta < \min\{\frac{\varepsilon}{2}, \varepsilon_\mu\}$ small enough, we can extend the distribution of subspaces $N \mapsto F_N$ and $\mathcal{E}^s_N$ given by Theorem 2.3 (for $\mathbb{P} = \mathcal{U}(N_0)$) to the compact set $\mathbb{P}_\beta$, by using the functional calculus on the derivatives $d\Delta_M$, for $M \in \mathbb{P}_\beta$ (see Remark 4.2). In this case, the subspaces F_M and $\mathcal{E}^s_M$ are $d\Delta_M$-invariant, $d\Delta_M|_{F_M}$ is near I_{F_M} and $\|d\Delta_M|_{\mathcal{E}^s_M}\| \leqslant k'_{N_0} < 1$, for some $k_{N_0} < k'_{N_0} < 1$.

The set $\mathbb{P}_\beta$ consists of fixed points for Δ. Hence this distribution satisfies the hypothesis of Proposition 4.5. Let $\rho > 0$ such that $B(N_0,\rho) \subseteq \mathcal{M}_{\lambda,\beta} \subseteq B_\Delta(\mathbb{P}_\beta)$. Following the same steps of the proof of Proposition 4.7, but using Lemma 4.11 instead of Lemma 4.9, we obtain that, if ρ is small enough, then for every $T \in B(N_0,\rho)$, there exists

$$N_1 \in \Pi_E\big(\mathcal{W}^{ss}_T\big) \cap \mathcal{U}(N_0) \cap B\left(N_0, \frac{\varepsilon}{2}\right).$$

Let $S \in \mathcal{W}^{ss}_T$ such that $\Pi_E(S) = N_1$. Since $E(S) = E(N_1)$ is an orthogonal system of projectors, then Eqs. (1) and (10) of Proposition 4.5 assure that

$$\Delta^\infty(T) = \Delta^\infty(S) = N_2 \quad \text{and} \quad \Pi_E(N_2) = \Pi_E(S) = N_1.$$

Since $T \in \mathcal{M}_{\lambda,\beta}$, Lemma 5.5 assures that

$$\|N_2 - N_1\| = \big\|N_2 - \Pi_E(N_2)\big\| < \beta < \frac{\varepsilon}{2}.$$

This shows that $\Delta^\infty(B(N_0,\rho)) \subseteq B(N_0,\varepsilon)$, i.e., that Δ^∞ is continuous at N_0. □

Remark 5.7. By Remark 5.2, the Aluthge transform fails to be differentiable at every non-invertible normal matrix. Because of this, we cannot use the previous techniques for proving continuity of Δ^∞ on $\mathcal{M}_r(\mathbb{C}) \setminus \mathcal{G}l_r(\mathbb{C})$. We conjecture that it is, indeed, continuous on $\mathcal{M}_r(\mathbb{C})$, but we have no proof for non-invertible matrices.

6. Concluding remarks

6.1. Rate of convergence

In [7] we proved that, if $T \in \mathcal{M}_r(\mathbb{C})$ is diagonalizable, then after some iterations the rate of convergence of the sequence $\Delta^n(T)$ becomes exponential. More precisely, for some $n_0 \in \mathbb{N}$ and every $n \geqslant n_0$, there exist $C > 0$ and $0 < \gamma < 1$ such that $\|\Delta^n(T) - \Delta^\infty(T)\| < C\gamma^n$. This exponential rate depends on the spectrum of T. Actually, if $\lambda(T) = \lambda(D)$ for some diagonal matrix D, then $\gamma = k_D$, the constant that appears in Theorem 2.1. Using the formula for k_D, one can see that it is closer to 1 (so that the rate of convergence becomes slower) if the different eigenvalues are closer one to each other.

These facts are not longer true if T is not diagonalizable, since the rate of convergence for such a T depends on the rate of convergence for some $M \in \mathcal{W}_T^{ss} \cap \mathcal{O}_D$ (with the notation of Proposition 4.7), which can be much slower (and not exponential). Note that the proof of the convergence of the sequence $\{\Delta^n(M)\}$, given in Proposition 4.6, does not study the rate of convergence. It only shows that there exists a unique possible limit point for the sequence.

Nevertheless, using Proposition 4.7 and Eq. (10), it is easy to see that the system of projections $E(\Delta^n(T))$ converges to $E(\Delta^\infty(T))$ exponentially, because $E(M) = E(\Delta^\infty(T))$. As in the case of diagonalizable matrices the rate of convergence of the spectral projections depends on the spectrum of T, which agrees with the spectrum of M. Note that the spectrum of T and the spectral projections of M completely characterize the limit $\Delta^\infty(T)$. Indeed, if $\sigma(T) = \{\mu_1, \ldots, \mu_k\}$, then

$$\Delta^\infty(T) = \Delta^\infty(M) = \Pi_E(M) = \sum_{j=1}^{k} \mu_j E_j(M).$$

6.2. λ-Aluthge transform

Given $\lambda \in (0, 1)$ and a matrix $T \in \mathcal{M}_r(\mathbb{C})$ whose polar decomposition is $T = U|T|$, the λ-Aluthge transform of T is defined by

$$\Delta_\lambda(T) = |T|^\lambda U |T|^{1-\lambda}.$$

All the results obtained in this paper are also true for the λ-Aluthge transform for every $\lambda \in (0, 1)$, with almost the same proofs. Indeed, the basic results about Aluthge transform used throughout Sections 3 and 4 are Theorem 2.1 and those stated in Section 2.1. All these results were extended to every λ-Aluthge transform (see [6] and [8]). The unique difference is that the constant k_D of Theorem 2.1 now depends on λ (see Theorem 3.2.1 of [8]). Anyway, the new constants are still smaller than one for every $\lambda \in (0, 1)$. Moreover, they are uniformly lower than one on compact subsets of $(0, 1)$.

Another result which depends particularly on the Aluthge transform is Proposition 4.1, which is used to prove Theorem 2.3. Nevertheless, it is easy to see that both results are still true for every $\lambda \in (0, 1)$. On the other hand, the proof of Theorem 5.6 uses the same facts about the Aluthge transform. So, it also remains true for Δ_λ, for every $\lambda \in (0, 1)$. We resume all these remarks in the following statement:

Theorem 6.1. *For every $T \in \mathcal{M}_r(\mathbb{C})$ and $\lambda \in (0, 1)$, the sequence $\Delta_\lambda^n(T)$ converges to a normal matrix $\Delta_\lambda^\infty(T)$. The map $T \mapsto \Delta_\lambda^\infty(T)$ is continuous on $\mathcal{G}l_r(\mathbb{C})$.*

We extend the conjecture given in Remark 5.7 to the following:

Conjecture 2. *The map* $(0,1) \times \mathcal{M}_r(\mathbb{C}) \ni (\lambda, T) \mapsto \Delta_\lambda^\infty(T)$ *is continuous.*

Using the same ideas as in Section 4 of [8], it can be proved that the above map is continuous if it is restricted to $(0,1) \times \mathcal{G}l_r(\mathbb{C})$.

Appendices

Appendix A. Brief review of differential geometry

In this section we recall some basic facts about differential geometry that have been used throughout the article. This resume is based on the books of Hirsch [14] and Abraham, Marsden and Ratiu [1]. We assume that the reader is familiar with the basic definitions of manifold, maps between manifolds, and tangent bundle. Only finite dimensional manifolds will be considered, because it is enough for our purposes.

We start by recalling the definition of some special classes of maps:

Definition A.1. Let M and N be two smooth manifolds and $f : M \to N$ a C^1 map. We say that f is immersive at $x \in M$ if $dT_x : T_xM \to T_{f(x)}N$ is injective, and it is an immersion if it is immersive for every $x \in M$. Analogously, we say that f is submersive at $x \in M$ if $dT_x : T_xM \to T_{f(x)}N$ is surjective, and it is a submersion if it is submersive for every $x \in M$. A map f is called embedding if f is an immersion and it maps M homeomorphically onto its image.

Roughly speaking, the next result states that the property of being an embedding is preserved under small perturbations:

Proposition A.2. *Let* $U \subseteq \mathbb{R}^m$ *be an open set and* $W \subseteq U$ *an open set with compact closure* $\overline{W} \subseteq U$*. Let* $M \subseteq \mathbb{R}^n$ *be a smooth submanifold and* $f : U \to M$ *a* C^1 *embedding. There exists* $\varepsilon > 0$ *such that, if*

$$g : U \to M \text{ is } C^1, \qquad \|dg_x - df_x\| < \varepsilon \quad \text{and} \quad \big\|g(x) - f(x)\big\| < \varepsilon$$

for every $x \in W$*, then* $g|_W$ *is an embedding.*

One of the most important applications of submersions is the next theorem:

Proposition A.3. *Let* M *and* N *be two smooth manifolds and* $f : M \to N$ *a* C^1 *map. Given* $y \in N$*, if* f *is submersive at every point* $x \in f^{-1}(\{y\})$*, then* $f^{-1}(\{y\})$ *is a submanifold of* M *(the regularity of* $f^{-1}(\{y\})$ *depends on the regularity of* f*).*

The following result provides a useful trick to prove that a function is a submersion, and it is a consequence of the inverse function theorem:

Proposition A.4. *Let* M *and* N *be two smooth manifolds and* $f : M \to N$ *a* C^1 *map. Then* f *is a submersion if and only of for any* $x \in N$ *there exist an open neighborhood* U_x *of* x *and a* C^1 *function* $s : U_x \to M$ *such that* $f(s(y)) = y$ *for every* $y \in U_x$*.*

In Proposition A.3 we state conditions that assure that the pre-image of a point is a submanifold. Sometimes, it is required to prove that the image by a function is a submanifold. The next result, which is a consequence of the inverse function theorem, goes in that direction:

Proposition A.5. *Let M and N be two smooth manifolds and $f : M \to N$ a C^1 map. If $\widetilde{M} \subseteq M$ is a submanifold of M such that $T_x\widetilde{M}$ is contained in a supplement of the nullspace of df_x for every $x \in \widetilde{M}$, then, $f(\widetilde{M})$ is a submanifold of N.*

Now we recall the definition of the weak topology of $C^k(M,N)$, also called C^r compact-open topology.

Definition A.6 *(C^r topology).* Let M and N be two smooth manifolds. Given $f \in C^r(M,N)$ and charts of M and N, $\varphi : U \to \mathbb{R}^n$ and $\psi : V \to \mathbb{R}^m$ respectively, consider $K \subseteq U$ compact such that $f(K) \subseteq V$, $\varepsilon > 0$ and define $\mathcal{N}(f,(\varphi,U),(\psi,V),K,\varepsilon)$ as the set of functions $g \in C^r(M,N)$ such that for every $x \in \psi(K)$ and every $j = 1,\ldots,k$

$$\left\| d^j\left(\varphi f \psi^{-1}\right)(x) - d^j\left(\varphi g \psi^{-1}\right)(x)\right\| \leqslant \varepsilon,$$

where $d^j(\cdot)$ denotes the j-derivative operator. The (weak) C^k topology on $C^r(M,N)$ is the topology generated by the sets $\mathcal{N}(f,(\varphi,U),(\psi,V),K,\varepsilon)$.

Remark A.7. In order to get the idea behind this topology, suppose that M and N are open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Then, a sequence of functions $\{f_n\}_{n\in\mathbb{N}}$ belonging to $C^r(M,N)$ converges to $f \in C^r(M,N)$ with respect to the C^r topology if, for every $j = 1,\ldots,k$, $d^j(f_n) \underset{n\to\infty}{\longrightarrow} d^j(f)$ uniformly on each compact subset of M.

To conclude this section of the appendix we give the definition of pre-lamination.

Definition A.8. A C^r *pre-lamination* indexed by N is a continuous function $\mathcal{B} : N \to \mathrm{Emb}^r((-\varepsilon,\varepsilon)^k, M)$, where in $\mathrm{Emb}^r((-\varepsilon,\varepsilon)^k, M)$ the C^r topology is considered. A pre-lamination is *self-coherent* if the interiors of each pair of its embedded discs meet in a relatively open subset of each.

Appendix B. Stable manifold theorem and related techniques

In this section we review some definitions related with hyperbolic dynamical systems, and in particular, we state the stable manifold theorem for an invariant set of a smooth map. The stable set is naturally defined for a fixed point of a map, as the set of points with positive trajectories heading directly toward 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 shares the same fixed point and such that any trajectory converges by forward iterate to the fixed point with an 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. The main references for these results are [15] (see especially Theorem 5.1 there). Also Chapter 5 of [22] can be checked (stable and unstable sets for non-bijective maps). In [9] and [10] the theory of persistence of normally submanifolds has been extended to non-bijective maps.

B.1. General remarks

First of all, let us review some terminology commonly used in dynamical systems:

Definition B.1. Let X be a topological space, $f : X \to X$. Then:

1. The *forward trajectories or forward orbits* are the sequences $\{f^n(x)\}_{n\in\mathbb{N}}$ for any $x \in X$, where f^n denotes n-composition of the map f.
2. A subset Λ is called forward invariant if $f(\Lambda) \subseteq \Lambda$.
3. A compact forward invariant subset Λ is called attractor if there is an open neighborhood U of Λ such that $\bigcap_{n\in\mathbb{N}} f^n(U) = \Lambda$. This is equivalent to say that for any $x \in U$, the accumulation points of the forward orbit $\{f^n(x)\}_{n\in\mathbb{N}}$ belong to Λ.

The idea of stable sets gives a formal mathematical definition to the general notions embodied in the idea of an attractor. The *stable set* of a compact invariant set Λ is defined as

$$W^s(\Lambda) := \bigl\{x :\in X\colon\ f^n(x) \underset{n\to+\infty}{\longrightarrow} \Lambda\bigr\}.$$

The *local stable* sets of Λ, in the case of metric spaces, is defined as

$$W^s(\Lambda)_\varepsilon := \bigl\{x :\in U\colon\ \operatorname{dist}\bigl(f^n(x), \Lambda\bigr) < \varepsilon,\ \forall n > 0,\ f^n(x) \underset{n\to+\infty}{\longrightarrow} \Lambda\bigr\}.$$

With these definitions, note that Λ is an attractor if and only if $W^s(\Lambda)$ contains an open neighborhood of Λ. Moreover, it follows that the local stable set is a neighborhood of Λ. For the case of an attractor, the stable set is usually called the basin of attraction, and the local stable set is called the local basin of attraction, denoted in this work by $B_f(\Lambda)$ and $B_f(\Lambda)_\varepsilon$ respectively.

A particularly interesting case is that of stable sets associated to hyperbolic fixed points. Given a fixed point p, i.e., $f(p) = p$, it is called hyperbolic if the spectrum of df_p does not intersect the unitarian circle. In this case, the local stable set is a connected submanifold whose tangent space is the subspace given by the eigenvalues with modulus smaller than one (they could be zero) and it is called *the local stable manifold*. Moreover, the points in the local stable manifold can also be characterized as the set of points that converge exponentially fast to the fixed point; i.e., there are $\varepsilon > 0$, $\lambda < 1$ and $C > 0$ such that

$$W^s(p)_\varepsilon := \bigl\{x :\in U\colon\ \operatorname{dist}\bigl(f^n(x), p\bigr) < \varepsilon,\ \operatorname{dist}\bigl(f^n(x), p\bigr) < C\lambda^n,\ \forall n > 0\bigr\}.$$

Recall that since f is not necessarily invertible, points can converge to the set in finite iterates (meaning that these points belong to the pre-image of p at distance at most ε from p). Recasting this notion, note that to say p is a hyperbolic fixed point is equivalent to say that the tangent space T_pM is decomposed in two complementary subspaces of df, one contracted by the action of df_p. In this sense, the local stable manifold plays the role "of the stable subspace" for the map f.

B.2. The (strong) stable manifold theorem

The notion of hyperbolicity of a fixed point can be naturally extended to the case of forward invariant sets of a map, in particular when the invariant set is a submanifold.

Definition B.2. Let f be a smooth map from M into itself, $\rho > 0$, and suppose that $f|_N$ is a homeomorphism. Then, N is *ρ-pseudo hyperbolic* for f if there exist two continuous 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. Both, $\mathcal{E}^s$ and $\mathcal{F}$, are df-invariant, in the sense that $df_x(\mathcal{E}^s_x) \subseteq \mathcal{E}^s_{f(x)}$ and $df_x(\mathcal{F}_x) \subseteq \mathcal{F}_{f(x)}$.
3. Tf restricted to $\mathcal{F}$ is an automorphism, which expands it by a factor greater than ρ.
4. $df_x : \mathcal{E}^s_x \to \mathcal{E}^s_{f(x)}$ has norm lower than ρ.

Under this hypothesis, for any point in N a local stable submanifold is "attached", i.e., the following standard version of the stable manifold theorem holds:

Theorem B.3 *(Stable manifold theorem). Let f be a C^r map from M into itself, and N a ρ-pseudo hyperbolic set with $\rho < 1$. Then, there is an f-invariant and self-coherent C^r-pre-lamination, $\mathcal{W}^s : N \to \operatorname{Emb}^r((-1,1)^k, M)$, 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. *$\mathcal{W}^s_x \subseteq \{y \in M\colon \operatorname{dist}(f^n(x), f^n(y)) < \operatorname{dist}(x,y)\rho^n$ for every $n \in \mathbb{N}\}$.*

A particular case of a compact ρ-pseudo hyperbolic invariant submanifold (and relevant for the present paper) is given when N is formed by fixed points and it is boundaryless. Under this hypothesis, N becomes an attractor and it is possible to put all the local stable manifolds together in such a way they form a lamination and provide a neighborhood of N (recall the definitions in Appendix B.1). The correct framework to prove this statement consists in dealing with all the local stable manifolds at the same time. This is the context and approach in the setting of Theorem 3.5. The points in those submanifolds are characterized as the points that their distance decrease exponentially fast. Theorem 2.1 follows from the previous result.

So, in what follows, we assume that N is a forward invariant attracting set and we consider its local basin of attraction $B_f(N)_\varepsilon$. This set is also forward invariant. The goal, is to show that under similar hypothesis of those in Theorem B.3, the local basin *is laminated* by submanifolds with the property that the distance between points in each such submanifold decrease exponentially fast.

To illustrate this, consider the following example: let $f(x,y) = (\frac{1}{3}x, y - y^3)$; note that $(0,0)$ is an attracting fixed point and $\mathbb{R} \times (-1,1)$ is contained in the basin of attraction; moreover, we have that the horizontal lines are invariant and the distance between forward iterates of points in the same horizontal lines decrease exponentially; after that, the lamination given by horizontal lines is called the strong stable foliation.

The first step is the extension of the invariant subbundles of $T_N M$, $\mathcal{E}^s$ and $\mathcal{F}$ to its local basin $B_f(N)_\varepsilon$ for some $\varepsilon > 0$ small enough. This is a classical argument in dynamical systems, and in the next subsection we sketch the proof of this extension in our particular case, just to show the idea of the techniques involved. The precise statement of the stable manifold theorem for local basins is the following:

Theorem B.4 *(Strong stable manifold theorem for local basins). Let $f: M \to M$ be a C^k map and let N be a ρ-pseudo hyperbolic attracting f-invariant subset of M with $\rho < 1$. Let us assume that for some $\varepsilon > 0$ there exist two continuous subbundles of $T_{B_f(N)_\varepsilon} M$, denoted by $\mathcal{E}^s$ and $\mathcal{F}$, such that, for every $x \in B_f(N)_\varepsilon$,*

1. $T_{B_f(N)_\varepsilon} M = \mathcal{E}^s \oplus \mathcal{F}$.
2. *$\mathcal{E}^s_x$ is df_x-invariant in the sense that $df_x(\mathcal{E}^s_x) \subseteq \mathcal{E}^s_{f(x)}$.*
3. *$\mathcal{F}_z$ is df_z-invariant, for every $z \in N$.*
4. *There exists $\rho \in (0,1)$ such that df_x restricted to $\mathcal{F}_x$ expand it by a factor greater than ρ, and $df_x : \mathcal{E}^s_x \to \mathcal{E}^s_{f(x)}$ has norm lower than ρ.*

Then, there is a continuous, f-invariant and self-coherent C^k-pre-lamination

$$\mathcal{W}^s : B_f(N)_\varepsilon \to \operatorname{Emb}^k\big((-1,1)^m, M\big) \quad \big(\textit{endowed with the } C^k\textit{-topology}\big)$$

such that, for every $x \in B_f(N)_\varepsilon$,

1. $\mathcal{W}^s(x)(0) = x$,
2. *$\mathcal{W}^s_x = \mathcal{W}^s(x)((-1,1)^m)$ is tangent to $\mathcal{E}^s_x$,*
3. $\mathcal{W}^s_x \subseteq \{y \in M\colon \operatorname{dist}(f^n(x), f^n(y)) < \operatorname{dist}(x,y)\rho^n\}$.

Remark B.5. Note that in the statement of Theorem B.3 it is assumed that f is a homeomorphism when restricted to N. This is not the case for Theorem B.4. The main idea that shows how to get the strong stable manifolds in the basin of attraction is very similar to the one used to extend the invariant subbundle in N to its basin. Therefore, the sketch of above theorems is provided in Appendix B.4 after proving Proposition 4.3.

B.3. Sketch of the proof of Proposition 4.3

Recall that $\mathbb{P}$ is a compact set consisting of fixed points of Δ with two complementary, continuous and $d\Delta_N$ invariant distributions $N \mapsto \mathcal{E}^s_N$ and $N \mapsto \mathcal{F}_N$ such that

$$\|d\Delta_N|_{\mathcal{E}^s_N}\| < 1-\rho \quad \text{and} \quad \big\|(I - d\Delta_N)|_{\mathcal{F}_N}\big\| < \frac{\rho}{2}, \quad N \in \mathbb{P}, \tag{16}$$

for some $\rho \in (0,1)$ which does not depend on N. The aim of the proposition is to extend them to distributions defined in some local basin of $\mathbb{P}$ with almost the same properties.

The first step is to extend these distributions using functional calculus: Fix $\varepsilon > 0$ such that $\sigma(d\Delta_T) \subseteq B(1, \frac{\rho}{2}) \cup B(0, 1-\rho)$, for every $T \in B_\Delta(\mathbb{P})_\varepsilon$. As in Remark 4.2, consider the spectral subspaces

$$F_T = R\big(\aleph_{B(1,\frac{\rho}{2})}(d\Delta_T)\big) \quad \text{and} \quad E_T = R\big(\aleph_{B(0,1-\rho)}(d\Delta_T)\big).$$

Note that Eq. (16) assures that $E_N = \mathcal{E}_N^s$ and $F_N = \mathcal{F}_N$ for every $N \in \mathbb{P}$. Since the functional calculus is smooth and $\mathbb{P}$ is compact, we can assume that, for every $T \in B_\Delta(\mathbb{P})_\varepsilon$, the angle between F_T and E_T is uniformly bounded from below, and $d\Delta_T$ satisfies inequalities as in Eq. (16), when it is restricted to E_T and F_T. Let us take the cones $C_T = C(\alpha, E_T)$ of size α in the direction E_T. For every small α, we can assume that:

(a) There exists $\gamma > 0$ such that $C_T \cap C(\gamma, F_T) = \{0\}$ for every $T \in B_\Delta(\mathbb{P})_\varepsilon$.
(b) Every subspace $E'_T \subseteq C_T$ with $\dim E'_T = \dim E_T$ satisfies inequalities as in Eq. (16).

Claim B.6. *There exist positive constants $\lambda_0 < 1$ and $\alpha > 0$ such that, if ε is a small enough, then for every $T \in B_\Delta(\mathbb{P})_\varepsilon$ we have that $[d\Delta_T]^{-1}(C_{\Delta(T)})$ is a cone of size not greater than $\lambda_0\alpha$ inside C_T.*

Proof. First note that, by the properties of the subspaces E_T and F_T, there exist $\lambda_1 < 1$ and $\alpha > 0$ such that, for every $T \in B_\Delta(\mathbb{P})_\varepsilon$ it holds that $[d\Delta_T]^{-1}(C_T) \subseteq C(\lambda_1\alpha, E_T)$ which is a cone of size $\lambda_1\alpha$ inside C_T.

Take $T \in B_\Delta(\mathbb{P})_\varepsilon$ and its image $\Delta(T) \in B_\Delta(\mathbb{P})_\varepsilon$. Note that Δ commutes with unitary conjugations, and it is uniformly continuous on compact sets. Hence, if ε is taken small enough, then T is arbitrarily (and uniformly) close to $\Delta(T)$ for every $T \in B_\Delta(\mathbb{P})_\varepsilon$. Therefore, E_T is arbitrarily and uniformly close to $E_{\Delta(T)}$, and the same occurs between $C_{\Delta(T)}$ and C_T. Putting all together, it follows that

$$[d\Delta_T]^{-1}(C_{\Delta(T)}) \sim [d\Delta_T]^{-1}(C_T) \subseteq C(\lambda_1\alpha, E_T).$$

Therefore, there exists $\lambda_1 < \lambda_0 < 1$ such that $[T\Delta]^{-1}(C_{\Delta(T)})$ is a cone of size not greater than $\lambda_0\alpha$ inside C_T. This completes the proof of the claim. $\square$

It is easy to see that Claim B.6 implies that, if C is a cone of size $\beta < \alpha$ inside $C_{\Delta(T)}$ and of the same dimension, then $[d\Delta_{\Delta(T)}]^{-1}(C)$ is a cone of size not greater than $\lambda_0\beta$ inside C_T. For each $T \in B_\Delta(\mathbb{P})_\varepsilon$, consider the sequence $\{\Delta^n(T)\}_{n\in\mathbb{N}}$ and the sequence of cones

$$C_1 = [d\Delta_T]^{-1}(C_{\Delta(T)}) \quad \text{and} \quad \{C_n\}_{n\in\mathbb{N}} = \big\{\big[d\Delta_T^n\big]^{-1}(C_{\Delta^n(T)})\big\}_{n\in\mathbb{N}}$$

in $T_T\mathcal{M}_r(\mathbb{C})$. The following facts hold: For every $n \in \mathbb{N}$,

$$\begin{aligned} C_{n+1} &= \big[d\Delta_T^n\big]^{-1}\big([d\Delta_{\Delta^n(T)}]^{-1}C_{\Delta^{n+1}(T)}\big) \\ &\subseteq \big[d\Delta_T^n\big]^{-1}(C_{\Delta^n(T)}) = C_n. \end{aligned}$$

Therefore $C_{n+1} \subseteq C_n \subseteq C_1 \subseteq C_T$ and every C_n is a cone of size not greater than $\lambda_0^n\alpha$. An easy argument of dimensions shows that every set C_n contains a subspace of dimension equal to $\dim E_T$ (even if the derivatives $d\Delta_{\Delta^n(T)}$ are not bijective). Therefore,

$$\mathcal{E}^s_T := \bigcap_{n\in\mathbb{N}} C_n = \bigcap_{n\in\mathbb{N}} [d\Delta_{\Delta^n(T)}]^{-n}(C_{\Delta^n(T)})$$

is a well-defined unique direction, and $\dim \mathcal{E}^s_T = \dim E_T$. Note that the direction is invariant and $\mathcal{E}^s_T \subseteq C_T$, and so it is contracted by $T\Delta$. Take $\mathcal{F}_T = F_T$, $T \in B_\Delta(\mathbb{P})_\varepsilon$, which is continuous by construction. The continuity of $\mathcal{E}^s_T$ follows from the fact that this subbundle is invariant and uniformly contracted for any forward iterate and from the uniqueness of a subbundle (with maximal dimension) exhibiting these properties. Finally, the subspaces $\mathcal{E}^s_T$ and $\mathcal{F}_T$ satisfy Eq. (16) by construction.

Remark B.7. In a general context, the previous proof can be recasted in the following steps:

1. Let N be a ρ-pseudo hyperbolic attracting subset of f;
2. let $T_N M = E^s \oplus F$ the df-invariant splitting in N;
3. let $x \to C(E^s_x)$ be a cone field defined on N such that $F_x \cap C(E^s_x) = \{0\}$ and note that this cone field is contracted, i.e., $[d_x f]^{-1}(C(E^s_{f(x)})) \subset C(E^s_x)$;
4. extend the previous cone field to a cone field $y \to C(y)$ defined in $B(N)_\varepsilon$ for some ε small in such a way that keeps the contraction property;
5. for any $y \in B(N)_\varepsilon \setminus N$ and any positive integer n consider the following family of cones $\{C_n(y) = [d_y f^n]^{-1}(C(f^n(y)))\}$ and note that $C_{n+1}(y) \subset C_n(y)$;
6. define $E^s_y = \bigcap_n C_n(y)$.

B.4. Sketch of the stable manifold's theorems

The proof of the existence of a map $\mathcal{W}^s : N \to \mathrm{Emb}^k((-1,1)^m, M)$ which satisfies all the mentioned conditions, consists in using the graph transform operator. We shall see that it is well defined if we only consider forward iterates. Therefore, since the basin of attraction of N is properly mapped inside by f, the graph transform operator is well defined on $B_f(N)_\varepsilon$, allowing us to extend the proof of stable manifolds to the whole local basin. Recall that to define the graph transform operator, first we consider $C^k(\widehat{\mathcal{E}}^s_x, \widehat{\mathcal{F}}_x)$, the set of C^k maps from $\widehat{\mathcal{E}}^s_x$ to $\widehat{\mathcal{F}}_x$, where

$$\widehat{\mathcal{E}}^s_x(\mu) = \exp\bigl(\mathcal{E}^s_x \cap (T_x M)_\mu\bigr), \qquad \widehat{\mathcal{F}}_x(\mu) = \exp\bigl(\mathcal{F}_x \cap (T_x M)_\mu\bigr)$$

and $\exp_x : (T_x M)_\mu \to M$ is the exponential map acting on $(T_x M)_\mu$, the ball of radius μ in $T_x M$. Later we consider the space

$$C^{k,0}\bigl(\widehat{\mathcal{E}}^s, \widehat{\mathcal{F}}\bigr) = \bigl\{\sigma : N \to C^k\bigl(\widehat{\mathcal{E}}^s_x, \widehat{\mathcal{F}}_x\bigr)\bigr\},$$

i.e., for each $x \in N$ we take $\sigma_x \in C^k(\widehat{\mathcal{E}}^s_x, \widehat{\mathcal{F}}_x)$ and we assume that the $x \mapsto \sigma_x$ is continuously. We can represent $C^{k,0}(\widehat{\mathcal{E}}^s, \widehat{\mathcal{F}})$ as a vector bundle over N given by $N \times \{C^k(\widehat{\mathcal{E}}^s_x, \widehat{\mathcal{F}})\}_{x\in X}$. Then, we take the maps

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

where p^1_x is the projection on $\widehat{\mathcal{E}}^s_x$ and p^2_x is the projection on $\widehat{\mathcal{F}}_x$. Now we take the *graph transform operator*. If f is a diffeomorphism, then we can obtain an explicit formula for the graph transform:

$$\Gamma_f(\sigma_x) = \left(f_x^2 \circ (id, \sigma_{f(x)})\right)^{-1} \circ \left(f_x^1 \circ (id, \sigma_x)\right)\big|_{\widehat{\mathcal{E}}_x^s}. \tag{17}$$

If f is not necessarily bijective (but a homeomorphism when restricted to N), the graph transform can be defined implicitly. Even though f^{-1} may not exist as a point-valued map, it does exist as a valued map. Moreover, the fact that $z \neq z'$ implies that $f^{-1}(z) \cap f^{-1}(z') = \emptyset$ is nearly as useful as injectivity of a point-valued map. In fact, the graph transform operator can be recasted using the relation that $f^{-1}(\mathrm{image}(\sigma_{f(x)})) = \mathrm{image}(\Gamma_f(\sigma_x))$. In this way, the local stable manifolds of N are obtained.

Arguing in the same way, this map is well defined in $B_f(N)_\varepsilon$ and therefore the whole proof can be carried out, in the sense of proving that the graph transform operator is a contractive map and therefore it has a fixed point.

References

[1] R. Abraham, J.E. Marsden, T. Ratiu, Manifolds, Tensor Analysis, and Applications, second ed., Appl. Math. Sci., Springer-Verlag, New York, 1988.

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

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

[5] E. Andruchow, D. Stojanoff, Differentiable structure of similarity orbits, J. Operator Theory 21 (1989) 349–366.

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

[7] J. Antezana, E. Pujals, D. Stojanoff, Convergence of iterated Aluthge transform sequence for diagonalizable matrices, Adv. Math. 216 (2007) 255–278.

[8] J. Antezana, E. Pujals, D. Stojanoff, Convergence of iterated Aluthge transform sequence for diagonalizable matrices II: λ-Aluthge transform, Integral Equations Operator Theory 62 (2008) 465–488.

[9] P. Berger, Persistence of stratifications of normally expanded laminations, C. R. Math. Acad. Sci. Paris 346 (13–14) (2008) 767–772.

[10] P. Berger, Persistence of laminations, preprint available in arXiv:0807.4952.

[11] M. Cho, I.B. Jung, W.Y. Lee, On Aluthge transform of p-hyponormal operators, Integral Equations Operator Theory 53 (2005) 321–329.

[12] G. Corach, H. Porta, L. Recht, The geometry of spaces of projections in C^*-algebras, Adv. Math. 101 (1993) 59–77.

[13] K. Dykema, H. Schultz, On Aluthge transforms: continuity properties and Brown measure, preprint available in arXiv:math/0512197.

[14] M.W. Hirsch, Differential Topology, Grad. Texts in Math., vol. 33, Springer-Verlag, New York, 1994.

[15] M.W. Hirsch, C.C. Pugh, M. Shub, Invariant Manifolds, Lecture Notes in Math., vol. 583, Springer-Verlag, Berlin, New York, 1977.

[16] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, 1985.

[17] Huajun Huang, Tin-Yau Tam, On the convergence of the Aluthge sequence, Oper. Matrices 1 (1) (2007) 121–141.

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

[19] I. Jung, E. Ko, C. Pearcy, The iterated Aluthge transform of an operator, Integral Equations Operator Theory 45 (2003) 375–387.

[20] T. Kato, Perturbation Theory for Linear Operators, reprint of the 1980 edition, Classics Math., Springer-Verlag, Berlin, 1995.

[21] J.E. Marsden, T.S. Ratiu, Introduction to Mechanics and Symmetry, Springer-Verlag, 1994.

[22] M. Quian, J.-S. Xie, S. Zhu, Smooth Ergodic Theory for Endomorphisms, Lecture Notes in Math., Springer-Verlag, Berlin, 2009.

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

[24] M. Shub, A.T. Vasquez, Some linearly induced Morse–Smale systems, the QR algorithm and the Toda lattice, in: The Legacy of Sonya Kovalevskaya, Proc. Symp., Radcliffe Coll., 1985, in: Contemp. Math., vol. 64, 1987, pp. 181–194.

[25] T.Y. Tam, Aluthge iteration and spectral radius, Integral Equations Operator Theory 60 (2008) 591–596.

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

[27] D. Wang, Heinz and McIntosh inequalities, Aluthge transformation and the spectral radius, Math. Inequal. Appl. 6 (2003) 121–124.