Angular values of nonautonomous and random linear dynamical systems: Part I – Fundamentals

Wolf-Jürgen Beyn* Gary Froyland † Thorsten Hüls*

Abstract. We introduce the notion of angular values for deterministic linear difference equations and random linear cocycles. We measure the principal angles between subspaces of fixed dimension as they evolve under nonautonomous or random linear dynamics. The focus is on long-term averages of these principal angles, which we call angular values: we demonstrate relationships between different types of angular values and prove their existence for random dynamical systems. For one-dimensional subspaces in two-dimensional systems our angular values agree with the classical theory of rotation numbers for orientation-preserving circle homeomorphisms if the matrix has positive determinant and does not rotate vectors by more than $\frac{\pi}{2}$. Because our notion of angular values ignores orientation by looking at subspaces rather than vectors, our results apply to dynamical systems of any dimension and to subspaces of arbitrary dimension. The second part of the paper delves deeper into the theory of the autonomous case. We explore the relation to (generalized) eigenspaces, provide some explicit formulas for angular values, and set up a general numerical algorithm for computing angular values via Schur decompositions.

Key words. Nonautonomous dynamical systems, random dynamical systems, angular value, ergodic average, principal angles of subspaces, numerical algorithm.

AMS subject classifications. 37C05, 37E45, 37A05, 65Q10, 15A18.

1. Introduction.

In this paper we propose and analyze suitable notions of angular values for linear nonautonomous discrete-time dynamical systems. The systems are of the form

$$u_{n+1} = A_n u_n, \quad u_0 \in \mathbb{R}^d, \quad n \in \mathbb{N}_0 \tag{1.1}$$

with $A_n \in \mathrm{GL}(\mathbb{R}^d)$, i.e. with real invertible $d \times d$ matrices $A_n, n \in \mathbb{N}_0$. Our goal is to study the average rotation of s-dimensional subspaces $V_0 \subseteq \mathbb{R}^d$ for $s = 1, \ldots, d$ when iterated as in (1.1), i.e. we consider the sequence of subspaces generated by

$$V_{n+1} = A_n V_n, \quad n \in \mathbb{N}_0, \tag{1.2}$$

so that $V_{n+1} = V_{n+1}(V_0)$ depends on V_0 via $V_{n+1} = A_n A_{n-1} \cdots A_1 A_0 V_0$. Since the matrices A_n are invertible the subspaces V_n have the same dimension s for all $n \in \mathbb{N}_0$. Their rotation is measured by the well-established notion of principal angles between subspaces which originates with C. Jordan in 1876. By $\measuredangle(V, W)$ we denote the maximum principal angle of two subspaces V, W and we recall that $0 \le \measuredangle(V, W) \le \frac{\pi}{2}$ holds. Some basics of the theory of principal angles and of their numerical computation may be found in [25], [17, Ch.6.4]. Generalizations to complex vector spaces and the triangle inequality appear in the papers [15], [21], [34]. In Section 2 we derive some specific results, tailored to our needs, such as estimates of principal angles in terms of norms and an angle bound for linear maps.

*Department of Mathematics, Bielefeld University, 33501 Bielefeld, Germany
beyn@math.uni-bielefeld.de, huels@math.uni-bielefeld.de

†School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052, Australia
g.froyland@unsw.edu.au

Using principal angles between successive spaces V_{j-1} and V_j generated by (1.2) we form the n-step average

$$\frac{1}{n}a_{1,n}(V_0), \quad \text{where} \quad a_{1,n}(V_0) = \sum_{j=1}^{n} \measuredangle(V_{j-1}, V_j), \quad n \ge 1 \tag{1.3}$$

and two types of limiting values

$$\bar{\theta}_s = \limsup_{n\to\infty} \sup_{V_0\in\mathcal{G}(s,d)} \frac{1}{n}a_{1,n}(V_0), \qquad \hat{\theta}_s = \sup_{V_0\in\mathcal{G}(s,d)} \limsup_{n\to\infty} \frac{1}{n}a_{1,n}(V_0), \tag{1.4}$$

where $\mathcal{G}(s,d)$ denotes the Grassmann manifold of s-dimensional subspaces of $\mathbb{R}^d$. We call $\bar{\theta}_s$ the s-inner and $\hat{\theta}_s$ the s-outer angular value of the system (1.1). In sections 3-4 we will discuss systems for which the lim sups in (1.4) are actually limits. More variations of these notions will be defined in Section 3.1, and some key examples will be presented in Section 3.2 which show that all types of angular values differ in general.

As a physical motivation of angular values consider some object, such as a small massless rod or a sheet, carried materially by a time-varying fluid flow, and assume that data about its position and orientation are available at discrete time instances. The task then is to measure the maximum average rotation of the object. In mathematical terms we think of a continuous time dynamical system determining its trajectory, and we assume that the system (1.1) describes its linearization about the trajectory when sampled at discrete times. Then the first and second outer angular values measure the maximum average angle of rotation exerted by the flow on a line ($s = 1$) or on a plane ($s = 2$). Rotations of subspaces $s \ge 3$ may be relevant in higher-dimensional phase spaces. In view of such applications it is natural to extend the quantities (1.4) to continuous-time systems. A short discussion of such an extension is given in the outlook of this article.

Perhaps the simplest example is a 2×2 orthogonal matrix, where $d = 2$, $s = 1$ and

$$A_n \equiv A = T_\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}, \quad 0 \le \varphi \le \frac{\pi}{2}. \tag{1.5}$$

All summands in (1.3) are φ and $a_{1,n}(V_0) = n\varphi$ for all one-dimensional $V_0 \subset \mathbb{R}^2$. Hence we find $\bar{\theta}_1 = \hat{\theta}_1 = \varphi$ in this case.

A first motivating example is the following randomized version of (1.5). Let $(\Omega, \mathbb{P})$ be a probability space, $\tau : \Omega \to \Omega$ be an ergodic transformation preserving $\mathbb{P}$ and $\varphi : \Omega \to [0, \frac{\pi}{2}]$ be a random variable. Setting $A(\omega) = T_{\varphi(\omega)}$ and $A_n = A(\tau^n\omega_0)$ for some $\omega_0 \in \Omega$ we see that $a_{1,n}(V_0) = \sum_{j=0}^{n-1} \varphi(\tau^j\omega_0)$ for every V_0. By Birkhoff's ergodic theorem, for $\mathbb{P}$-almost every ω_0, one has $\lim_{n\to\infty} \frac{1}{n}a_{1,n}(V_0) = \int \varphi(\omega)\, \mathrm{d}\mathbb{P}(\omega)$. The above general formula holds for driving systems τ modeling any ergodic stationary deterministic or stochastic process. In Section 4 we generalize the various notions of angular values to the general setting of random dynamical systems (cf. [2]). We establish their existence via ergodic theorems and prove inequalities between the various types; see Theorem 4.2.

A second motivating example abandons orthogonality and changes (1.5) by a skewing factor $0 < \rho \le 1$ to

$$A_n \equiv A(\rho, \varphi) = \begin{pmatrix} \cos(\varphi) & -\rho^{-1}\sin(\varphi) \\ \rho\sin(\varphi) & \cos(\varphi) \end{pmatrix}, \quad 0 \le \varphi \le \frac{\pi}{2}.$$

This matrix turns out to be a kind of normal form with regard to measuring angles between a one-dimensional subspace and its image (see Proposition 5.2). The angular values $\hat{\theta}_1$ and $\bar{\theta}_1$ agree in this case, but they differ from φ in general and depend critically on the value of ρ (see Proposition 5.2 and Theorem 6.1).

There is a weak analogy of first angular values to Lyapunov exponents which measure the maximum average exponential growth of a linear nonautonomous system (1.1); see e.g. [2, Ch.3.2], [6], [24, Suppl.2]). For the latter purpose it is enough to compare the norm of the last iterate with the first one and average the logarithm. However, in the angular direction one expects only linear growth which requires one to calculate an arithmetic average over every single time step.

For certain systems, the above definition of angular values is related to existing concepts of measuring rotations in dynamical systems, which we now discuss. We first mention the classical theory of rotation numbers for orientation-preserving homeomorphisms of the circle, cf. [9], [24, Ch.11], [27]. If the system (1.1) is two-dimensional and autonomous (i.e. $A_n \equiv A \in \mathrm{GL}(\mathbb{R}^2)$), then it generates a homeomorphism of the unit circle, which is orientation-preserving for $\det(A) > 0$. If, in addition, no vector rotates by an angle greater than $\frac{\pi}{2}$, then the rotation number agrees (up to a factor of 2π) with the first angular value; see Section 5.1, Remark 5.3 and Proposition 5.2 for more details. However, such a comparison is no longer possible for a reflection or for matrices which generate rotations of vectors with angles larger than $\frac{\pi}{2}$. By contrast to rotation numbers, our definition (1.4) avoids assuming or specifying any orientation, even when one observes the motion of one-dimensional subspaces (rather than vectors) in a two-dimensional space. Including orientation typically leads to complications in discrete-time systems. For example, for rotations that are close to reflections one needs extra analytic information from the system (such as $\det(A_n)$), which we consider as inaccessible to observation. When rotations of vectors larger than $\frac{\pi}{2}$ occur, our definition takes the smaller of both possible angles; Figure 1.1 illustrates this for a sequence of subspaces. Note that angles between successive subspaces are indicated by black arcs with time progressing outward.

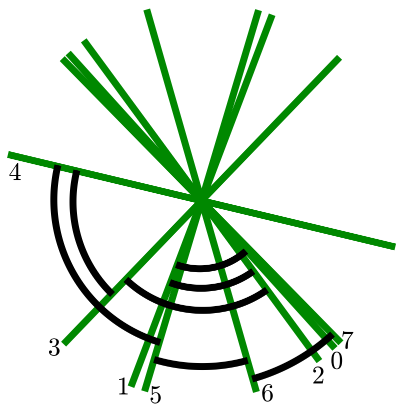

Figure 1.1: Angles between successive subspaces in the Hénon system (6.10).

The theory of rotation numbers for homeomorphisms of the circle has been generalized to so-called rotation sets of toral automorphisms in [26], and a numerical approach appears in [29]. However, there seems to be no connection to the definition (1.4) in higher dimensions.

Another far-reaching extension of rotation numbers to nonautonomous continuous time systems of arbitrary dimension has been proposed and investigated in [3], [2, Ch.6.5]. The average rotation of vectors is measured within all two-dimensional subspaces (more generally within tangent planes of a manifold) mapped by the system. Orientation is taken into account where counterclockwise refers to positive values. In essence one studies the flow induced by the given system on the Grassmannian $\mathcal{G}(2,d)$. The concept generalizes to nonlinear random dynamical systems and even leads to a multiplicative ergodic theorem, see [2, Th.6.5.14]. However, the conceptual difference to angular values remains the same as for the classical rotation numbers.

Yet another concept of rotation numbers has been developed for continuous time linear Hamiltonian systems of arbitrary dimension; see [23], [22], with a route from the theory to numerical results provided in [13]. The notion is based on a suitable generalization of the arg-function from a scalar complex system to the even dimensional real case. Then the rotation number appears as the limit of the time average of the arg-function when applied to a symplectic fundamental matrix. The setting is similar to the random dynamical systems mentioned above. The resulting rotation number has interesting relations to the dichotomy spectrum of a parametrically perturbed Hamiltonian system; see [13, Theorem 4-6]. This notion differs from the first angular values of this paper since time is continuous and orientation is taken into account by the choice of the arg-function.

Let us also mention the notion of antieigenvalues and antieigenvectors developed in [18]. They are determined by the maximum angle $\measuredangle(v, Av)$ by which a given matrix A can turn a vector $v \in \mathbb{R}^d$. This corresponds to maximizing the first summand in (1.3), but ergodic averages seem not to have been considered in this theory.

In the following we summarize some further results of this paper. In Section 3.1 we collect elementary properties of angular values, such as inequalities among them and invariance under special kinematic similarities, see [16] for this notion. Section 5 presents an in-depth study of the autonomous case $A_n \equiv A$. The main theoretical result is Theorem 5.7 which reduces the computation of angular values to the case of a block-diagonal matrix. Theorem 5.7 builds on a spectral decomposition (Blocking Lemma 5.5), on a special treatment of multiple real eigenvalues (Proposition 5.4), and on a detailed analysis of the two-dimensional case (Proposition 5.2). In the two-dimensional case we show that all types of first angular values coincide and provide a rather explicit formula (Proposition 5.2, Theorem 6.1). While real eigenvalues of the matrix lead to a vanishing angular value, complex conjugate ones lead to interesting resonances depending on a skewness parameter; see Figures 6.1, 6.2. In the latter case we use ergodic theory to derive an integral expression for the first angular value when rotation occurs with irrational multiples of π, and we reduce the computation to maximizing a finite sum in the rational case. In Section 6 we present a numerical algorithm for the autonomous case based on eigenvalue computations and one-dimensional optimization which avoids failure caused by simple forward iteration. We apply the algorithm to study various systems up to dimension 10^4, and we confirm numerically the rather subtle behavior in the two-dimensional complex conjugate case.

2. Angles of subspaces.

In this section we collect some useful results about principal angles between subspaces. In the following, let $\|v\| = \sqrt{v^\top v}$ denote the Euclidean norm for $v \in \mathbb{R}^d$ and let $\mathcal{R}(A)$, $\mathcal{N}(A)$ and $\sigma(A)$ denote the range, the kernel and the spectrum of a matrix A. Recall the definition of principal angles and principal vectors of two subspaces

V, W of $\mathbb{R}^d$ of equal dimension from [17, Ch.6.4.3].

Definition 2.1. *Let V, W be subspaces of $\mathbb{R}^d$ of dimension s. Then the principal angles $0 \le \phi_1 \le \ldots \le \phi_s \le \frac{\pi}{2}$ and associated principal vectors $v_j \in V$, $w_j \in W$ are defined recursively for $j = 1, \ldots, s$ by*

$$\cos(\phi_j) = \max_{\substack{v \in V, \|v\|=1 \\ v^\top v_\ell = 0, \ell=1,\ldots,j-1}} \max_{\substack{w \in W, \|w\|=1 \\ w^\top w_\ell = 0, \ell=1,\ldots,j-1}} v^\top w = v_j^\top w_j. \tag{2.1}$$

The right-hand side of (2.1) lies in $[0, 1]$, so that $\phi_j \in [0, \frac{\pi}{2}]$ is uniquely defined by (2.1). While principal angles are unique, principal vectors are not, in general. Let us note that principal angles and principal vectors are also defined for subspaces of different dimension (see [17, Ch.6.4.3]), but this feature will not be used due to our assumption of invertibility. We further write $\phi_j = \phi_j(V, W)$ to indicate the dependence on the subspaces, and for the largest angle we introduce the notation

$$\phi_s(V, W) = \measuredangle(V, W).$$

If the subspaces V and W are one-dimensional we may write

$$\measuredangle(v, w) = \measuredangle(\mathrm{span}(v), \mathrm{span}(w)), \quad v, w \in \mathbb{R}^d, v, w \neq 0.$$

Let us also note that the usage of the angle between subspaces varies in the literature. For example, in [1, (3.3.13)], [2, p.216] this notion is used for $\sin(\phi_1)$ where ϕ_1 is the smallest angle. Then (2.1) turns into a min-min characterization, and the angle becomes zero if both subspaces share a common direction.

Principal values and vectors can be computed from a singular value decomposition (SVD) as follows.

Proposition 2.2. *([17, Algorithm 6.4.3]) Let $P, Q \in \mathbb{R}^{d,s}$ be two matrices with orthonormal columns and consider the SVD*

$$P^\top Q = Y \Sigma Z^\top, \quad Y, Z, \Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_s) \in \mathbb{R}^{s,s}, \ Y^\top Y = I_s = Z^\top Z. \tag{2.2}$$

Then the principal angles ϕ_j of $V = \mathcal{R}(P)$ and $W = \mathcal{R}(Q)$ satisfy

$$\sigma_j = \cos(\phi_j), j = 1, \ldots, s,$$

and principal vectors are given by

$$PY = \begin{pmatrix} v_1 & \cdots & v_s \end{pmatrix}, \quad QZ = \begin{pmatrix} w_1 & \cdots & w_s \end{pmatrix}. \tag{2.3}$$

Since the singular values of $P^\top Q$ and $Q^\top P$ agree, principal angles are symmetric with respect to V and W. In particular, the maximum angle satisfies

$$\measuredangle(V, W) = \measuredangle(W, V).$$

In Definition 2.1 the angles between two subspaces of equal dimension are defined recursively. For the computation of the j-th principal angle, the max-max characterization (2.1) requires knowledge of the principal vectors from index 1 to $j - 1$. In the following proposition we state a complementary min-max characterization. It begins with ϕ_s and computes ϕ_j via the known principal vectors for indices s to $j+1$. The result is motivated by the Hausdorff semi-distance between unit balls and proves to be better suited for the key estimates below. The proof will be given in the Supplementary materials I.

Proposition 2.3. *Let $V, W \subseteq \mathbb{R}^d$ be two s-dimensional subspaces. Then the principal angles and principal vectors satisfy for $j = s, \ldots, 1$*

$$\cos(\phi_j) = \min_{\substack{v \in V, \|v\|=1 \\ v^\top v_\ell = 0, \ell = j+1, \ldots, s}} \max_{\substack{w \in W, \|w\|=1 \\ w^\top w_\ell = 0, \ell = j+1, \ldots, s}} v^\top w = v_j^\top w_j. \tag{2.4}$$

In particular, the following relation holds

$$\measuredangle(V, W) = \phi_s(V, W) = \max_{\substack{v \in V \\ v \neq 0}} \min_{\substack{w \in W \\ w \neq 0}} \measuredangle(v, w) = \arccos\Big(\min_{\substack{v \in V \\ \|v\|=1}} \max_{\substack{w \in W \\ \|w\|=1}} v^\top w \Big). \tag{2.5}$$

Remark 2.4. *A related variational characterization appears in [31, Theorem 3]*

$$\cos(\phi_j) = \min_{\substack{U \subseteq V \\ \dim U = j-1}} \max_{\substack{x \in U^\perp \cap V, \|x\|=1 \\ y \in W, \|y\|=1}} |\langle x, y \rangle|.$$

If $j = s$ then $\dim U = s - 1$ and $x \in U^\perp \cap V$ runs through V with $\|x\| = 1$. Therefore, the formula implies (2.4) in the case $j = s$, but for $j < s$ the formulas differ.

Next we recall some well-known properties of the Grassmannian,

$$\mathcal{G}(s, d) = \{V \subseteq \mathbb{R}^d \text{ is a subspace of dimension } s\},$$

which may be found in [17, Ch.6.4.3], [21], for example.

Proposition 2.5. *The Grassmannian $\mathcal{G}(s, d)$ is a compact smooth manifold of dimension $s(d - s)$ and a metric space with respect to*

$$d(V, W) = \|P_V - P_W\|,$$

where P_V, P_W are the orthogonal projections onto V and W, respectively, and the formula

$$d(V, W) = \sin(\measuredangle(V, W)), \quad V, W \in \mathcal{G}(s, d)$$

holds. Furthermore, $\measuredangle(V, W)$ defines an equivalent metric on $\mathcal{G}(s, d)$ satisfying

$$\frac{2}{\pi} \measuredangle(V, W) \le d(V, W) \le \measuredangle(V, W).$$

Some useful geometric estimates for angles of vectors and subspaces are the following:

Lemma 2.6. *(Angle estimates)*

(i) For any two vectors $v, w \in \mathbb{R}^d$ with $\|v\| < \|w\|$ the following holds

$$\begin{aligned} \tan^2 \measuredangle(v + w, w) &\le \frac{\|v\|^2}{\|w\|^2 - \|v\|^2}, \\ \cos^2 \measuredangle(v + w, w) &\ge \frac{\|w\|^2 - \|v\|^2}{\|w\|^2}. \end{aligned} \tag{2.6}$$

(ii) Let $V \in \mathcal{G}(s,d)$ and $P \in \mathbb{R}^{d,d}$ be such that for some $0 \le q < 1$

$$\|(I-P)v\| \le q\|Pv\| \quad \forall\, v \in V. \tag{2.7}$$

Then $\dim(V) = \dim(PV)$ and the following estimate holds

$$\measuredangle(V, PV) \le \frac{q}{(1-q^2)^{1/2}}. \tag{2.8}$$

Proof. The first inequality in (2.6) follows from the second via the relation $\tan^2\alpha = \frac{1}{\cos^2\alpha} - 1$. The second inequality in (2.6) can be rewritten as

$$\big((v+w)^\top w\big)^2 - \|v+w\|^2(\|w\|^2 - \|v\|^2) \ge 0.$$

A short computation shows that the left-hand side agrees with $\big(v^\top w + \|v\|^2\big)^2$ which proves our assertion. The estimate (2.7) shows that $Pv = 0, v \in V$ implies $v = 0$, hence $\dim(V) = \dim(PV)$. Inequality (2.8) follows from (2.6) and the characterization (2.5)

$$\begin{aligned}\measuredangle(V, PV) &= \max_{\substack{v\in V\\ v\neq 0}} \min_{\substack{w\in PV\\ w\neq 0}} \measuredangle(v,w) \le \max_{\substack{v\in V\\ v\neq 0}} \measuredangle(v, Pv) \le \max_{\substack{v\in V\\ v\neq 0}} |\tan\measuredangle(v,Pv)| \\ &\le \max_{v\in V, v\neq 0} \frac{\|(I-P)v\|}{\|Pv\|}\Big(1 - \frac{\|(I-P)v\|^2}{\|Pv\|^2}\Big)^{-1/2} \le \frac{q}{(1-q^2)^{1/2}}.\end{aligned}$$

■

Remark 2.7. *The proof shows that the inequalities in (2.6) are strict for $v \neq 0$.*

Lemma 2.6 will be important for proving the Blocking Lemma 5.5 in the autonomous case. The next auxiliary result provides an angle-bound for an invertible matrix; it will be used in Proposition 5.4 for treating real eigenvalues in the autonomous case.

Lemma 2.8. *Let $S \in \mathrm{GL}(\mathbb{R}^d)$ and $\kappa = \|S^{-1}\|\|S\|$ be its condition number. Then the following estimate holds*

$$\measuredangle(SV, SW) \le \pi\kappa(1+\kappa)\measuredangle(V,W) \quad \forall\, V, W \in \mathcal{G}(s,d), \quad 1 \le s \le d. \tag{2.9}$$

Proof. Let us first prove (2.9) for $s = 1$. Then we can assume $V = \mathrm{span}(v)$, $W = \mathrm{span}(w)$ with $\|v\| = \|w\| = 1$ and $v^\top w \ge 0$. From Proposition 2.5 we have

$$\begin{aligned}\frac{1}{\pi}\measuredangle(v,w) &\le \frac{1}{2}d(V,W) = \frac{1}{2}\|vv^\top - ww^\top\| = \frac{1}{2}\|(v-w)v^\top + w(v-w)^\top\| \\ &\le \|v-w\| = (2(1-\cos(\measuredangle(v,w))))^{1/2} = 2\sin(\tfrac{1}{2}\measuredangle(v,w)) \le \measuredangle(v,w).\end{aligned} \tag{2.10}$$

We apply the first inequality in (2.10) to the image spaces and obtain

$$\begin{aligned}\measuredangle(Sv, Sw) &= \measuredangle(\|Sv\|^{-1}Sv, \|Sw\|^{-1}Sw) \le \pi\|S\left(\|Sv\|^{-1}v - \|Sw\|^{-1}w\right)\| \\ &\le \pi\|S\|\left(|\|Sv\|^{-1} - \|Sw\|^{-1}| + \|Sw\|^{-1}\|v-w\|\right) \\ &\le \pi\|S\|\|Sw\|^{-1}\left(\|Sv\|^{-1}\|S(w-v)\| + \|v-w\|\right).\end{aligned}$$

Now $\|Sw\|^{-1}, \|Sv\|^{-1} \le \|S^{-1}\|$ and the last inequality from (2.10) lead to

$$\measuredangle(Sv, Sw) \le \pi\kappa(1+\kappa)\|v-w\| \le \pi\kappa(1+\kappa)\measuredangle(v,w).$$

For the general case $s \ge 1$ we use (2.9) for all vectors $v \in V, v \neq 0$, $w \in W, w \neq 0$ and then apply the max-min characterization (2.5) from Proposition 2.3. ■

3. Basic theory of angular values. For an invertible nonautonomous linear system (1.1) we define the solution operator Φ_A by

$$\Phi_A(n,m) = \begin{cases} A_{n-1}\cdot \ldots \cdot A_m, & \text{for } n > m, \\ I, & \text{for } n = m, \\ A_n^{-1}\cdot \ldots \cdot A_{m-1}^{-1}, & \text{for } n < m. \end{cases}$$

Usually we suppress the dependence on the matrix sequence $A_n, n \in \mathbb{N}_0$ and simply write $\Phi = \Phi_A$. However, in Section 4 we consider matrix families generated by a linear random dynamical system for which the dependence on the family is essential.

3.1. Definitions and elementary properties. In the following we consider various ways of defining the average angular rotation that the system (1.1) exerts on subspaces of a fixed dimension. For this we use the notion of angles of subspaces from Section 2.

We reconsider a rigid rotation (1.5) as a simple motivating example, but now we allow $0 \le \varphi \le \pi$. For $v \in \mathbb{R}^2, v \neq 0$ and $j \in \mathbb{N}$ one obtains with Proposition 2.2 that

$$\measuredangle(v, T_\varphi v) = \measuredangle(T_\varphi^{j-1}v, T_\varphi^j v) = \arccos(|\cos(\varphi)|) = \min(\varphi, \pi-\varphi).$$

Hence we obtain for $n \in \mathbb{N}$ the arithmetic mean

$$\sup_{v\in\mathbb{R}^2} \frac{1}{n}\sum_{j=1}^{n} \measuredangle(T_\varphi^{j-1}v, T_\varphi^j v) = \sup_{V\in\mathcal{G}(1,2)} \frac{1}{n}\sum_{j=1}^{n} \measuredangle(\Phi(j-1,0)V, \Phi(j,0)V) = \min(\varphi, \pi-\varphi)$$

and the same value for both types of limits $\sup_{V\in\mathcal{G}(1,2)} \lim_{n\to\infty}$ and $\lim_{n\to\infty} \sup_{V\in\mathcal{G}(1,2)}$.

For general systems however, it turns out that the limit does not necessarily commute with the supremum, and sometimes the limit does not even exist. Therefore, we introduce several different types of angular values.

Definition 3.1. *Let the invertible nonautonomous system (1.1) be given. For every $s \in \{1,\ldots,d\}$ define the quantities*

$$a_{k+1,k+n}(V) = \sum_{j=k+1}^{k+n} \measuredangle(\Phi(j-1,0)V, \Phi(j,0)V) \quad n\in\mathbb{N},\ k\in\mathbb{N}_0,\ V\in\mathcal{G}(s,d). \tag{3.1}$$

i) The upper resp. lower s-th **inner angular value** *is defined by*

$$\bar{\theta}_s = \limsup_{n\to\infty} \frac{1}{n} \sup_{V\in\mathcal{G}(s,d)} a_{1,n}(V), \quad \underline{\theta}_s = \liminf_{n\to\infty} \frac{1}{n} \sup_{V\in\mathcal{G}(s,d)} a_{1,n}(V). \tag{3.2}$$

ii) The upper resp. lower s-th **outer angular value** *is defined by*

$$\hat{\theta}_s = \sup_{V\in\mathcal{G}(s,d)} \limsup_{n\to\infty} \frac{1}{n} a_{1,n}(V), \quad \underset{\hat{}}{\theta}_s = \sup_{V\in\mathcal{G}(s,d)} \liminf_{n\to\infty} \frac{1}{n} a_{1,n}(V). \tag{3.3}$$

iii) The upper resp. lower s-th **uniform inner angular value** *is defined by*

$$\begin{aligned} \bar{\theta}_{[s]} &= \limsup_{n\to\infty} \frac{1}{n} \sup_{V\in\mathcal{G}(s,d)} \sup_{k\in\mathbb{N}_0} a_{k+1,k+n}(V), \\ \underline{\theta}_{[s]} &= \liminf_{n\to\infty} \frac{1}{n} \sup_{V\in\mathcal{G}(s,d)} \inf_{k\in\mathbb{N}_0} a_{k+1,k+n}(V). \end{aligned} \tag{3.4}$$

iv) The upper resp. lower s-th **uniform outer angular value** *is defined by*

$$
\begin{aligned}
\hat{\theta}_{[s]} &= \sup_{V\in\mathcal{G}(s,d)} \lim_{n\to\infty} \frac{1}{n} \sup_{k\in\mathbb{N}_0} a_{k+1,k+n}(V),\\
\underset{\wedge}{\theta}_{[s]} &= \sup_{V\in\mathcal{G}(s,d)} \lim_{n\to\infty} \frac{1}{n} \inf_{k\in\mathbb{N}_0} a_{k+1,k+n}(V).
\end{aligned} \tag{3.5}
$$

Remark 3.2. *In the case $s = d$, all angular values are zero since the invertible system keeps the space $V = \mathbb{R}^d$ fixed and since $\measuredangle(\mathbb{R}^d, \mathbb{R}^d) = 0$.*

Our guiding principle in forming these quantities is to seek the subspace V which maximizes an angular value. The notions of 'upper' and 'lower' are motivated by the possible gap between $\limsup$ and $\liminf$ while 'outer' and 'inner' result from the noncommuting $\lim$ and $\sup$. The corresponding uniform angular values (and their 'lower' and 'upper' variants) become relevant when passing from autonomous to nonautonomous systems; see Sections 3.2 and 5.

As shorthand, we use an up/down bar for upper/lower inner angular values and an up/down hat for upper/lower outer angular values, while their uniform equivalents are indicated by the bracketed index $[s]$.

Clearly, the $\limsup$ and $\liminf$ in (3.2), (3.3), (3.4) are finite due to the boundedness of the angles. In Section 4 we prove that the $\limsup$ and $\liminf$ in (3.2) actually become limits in the setting of random dynamical systems. Let us further mention that the supremum for both quantities in (3.2) can be replaced by a maximum since $a_{1,n}(V)$ depends continuously on V in the compact space $\mathcal{G}(s, d)$.

In the following lemma we show that the limits in (3.5) always exist and that the $\limsup$ in the definition (3.4) of $\bar{\theta}_{[s]}$ is in fact a limit. Further, we collect some easy relations between the various angular values.

Lemma 3.3. *The limits in the definition (3.5) of the uniform outer angular values exist in $[0, \frac{\pi}{2}]$ and the* $\limsup$ *in the definition of $\bar{\theta}_{[s]}$ is a limit. Moreover, the relations of Diagram 3.1 hold for all $s = 1, \dots, d$.*

$$
\begin{array}{ccccccc}
\underset{\wedge}{\theta}_{[s]} & \leq & \underset{\wedge}{\theta}_{s} & \leq & \hat{\theta}_s & \leq & \hat{\theta}_{[s]}\\
\rotatebox{-90}{$\leq$} & & \rotatebox{-90}{$\leq$} & & \rotatebox{-90}{$\leq$} & & \rotatebox{-90}{$\leq$}\\
\underline{\theta}_{[s]} & \leq & \underline{\theta}_{s} & \leq & \bar{\theta}_s & \leq & \bar{\theta}_{[s]}
\end{array}
$$

Diagram 3.1: Comparison of angular values.

For the smallest and the largest value in this diagram we have the estimate

$$
\sup_{V\in\mathcal{G}(s,d)} \inf_{k\in\mathbb{N}_0} \measuredangle(\Phi(k,0)V, A_k\Phi(k,0)V) \leq \underset{\wedge}{\theta}_{[s]} \leq \bar{\theta}_{[s]} \leq \sup_{V\in\mathcal{G}(s,d)} \sup_{k\in\mathbb{N}_0} \measuredangle(V, A_kV). \tag{3.6}
$$

Proof. For every $V \in \mathcal{G}(s, d)$, the sequence $a_n(V) = \sup_{k\in\mathbb{N}_0} a_{k+1,k+n}(V)$ lies in $[0, \frac{n\pi}{2}]$ and is subadditive

$$
\begin{aligned}
a_{n+m}(V) &= \sup_{k\in\mathbb{N}_0} \left(a_{k+1,k+n}(V) + a_{k+n+1,k+n+m}(V)\right)\\
&\leq \sup_{k\in\mathbb{N}_0} a_{k+1,k+n}(V) + \sup_{\kappa\geq n} a_{\kappa+1,\kappa+m}(V) \leq a_n(V) + a_m(V).
\end{aligned}
$$

By Fekete's subadditive lemma [14, Lemma 4.2.7] this ensures

$$\lim_{n\to\infty}\frac{1}{n}a_n(V)=\inf_{n\in\mathbb{N}}\frac{1}{n}a_n(V)\in[0,\frac{\pi}{2}].$$

In a similar way, the sequence $a_n=\sup_{V\in\mathcal{G}(s,d)}a_n(V)$ turns out to be subadditive, which shows that $\limsup=\lim$ for the first quantity in (3.4). Further, the sequence $\alpha_n(V)=\inf_{k\in\mathbb{N}_0}a_{k+1,k+n}(V)$ turns out to be superadditive, i.e. $\alpha_{n+m}(V)\geq\alpha_n(V)+\alpha_m(V)$ for $n,m\in\mathbb{N}$, and thus

$$\lim_{n\to\infty}\frac{1}{n}\alpha_n(V)=\sup_{n\in\mathbb{N}}\frac{1}{n}\alpha_n(V)\in[0,\frac{\pi}{2}].$$

Next we prove the inequalities $\underset{\wedge}{\theta}_s\leq\hat{\theta}_s\leq\bar{\theta}_s\leq\bar{\theta}_{[s]}$,

$$\begin{aligned}\underset{\wedge}{\theta}_s&=\sup_{V\in\mathcal{G}(s,d)}\liminf_{n\to\infty}\frac{1}{n}a_{1,n}(V)\leq\sup_{V\in\mathcal{G}(s,d)}\limsup_{n\to\infty}\frac{1}{n}a_{1,n}(V)=\hat{\theta}_s\\&\leq\limsup_{n\to\infty}\frac{1}{n}\sup_{V\in\mathcal{G}(s,d)}a_{1,n}(V)=\bar{\theta}_s\leq\limsup_{n\to\infty}\frac{1}{n}\sup_{V\in\mathcal{G}(s,d)}\sup_{k\in\mathbb{N}_0}a_{k+1,k+n}(V)=\bar{\theta}_{[s]}.\end{aligned}$$

The remaining assertions in Diagram 3.1 follow in a similar way. Finally, note that Fekete's lemma leads to the representations

$$\sup_{V\in\mathcal{G}(s,d)}\sup_{n\in\mathbb{N}}\frac{1}{n}\inf_{k\in\mathbb{N}_0}a_{k+1,k+n}(V)=\underset{\wedge}{\theta}_{[s]}\leq\bar{\theta}_{[s]}=\inf_{n\in\mathbb{N}}\frac{1}{n}\sup_{V\in\mathcal{G}(s,d)}\sup_{k\in\mathbb{N}_0}a_{k+1,k+n}(V).$$

The inequalities (3.6) then follow by setting $n=1$ in $\sup_n$ and $\inf_n$. ■

We extend the motivating example (1.5) and analyze in detail the outer angular values of the 3-dimensional system defined by

$$A_n=A=\begin{pmatrix}\cos(\varphi)&-\sin(\varphi)&0\\\sin(\varphi)&\cos(\varphi)&0\\0&0&2\end{pmatrix},\quad n\in\mathbb{N}_0,\ 0<\varphi\leq\frac{\pi}{2}.\tag{3.7}$$

Denote by e_j the j-th unit vector in $\mathbb{R}^3$. For $v\in\operatorname{span}(e_1,e_2)$, we get, cf. (1.5), that $\measuredangle(A^{i-1}v,A^iv)=\varphi$ for all $i\in\mathbb{N}$. For $v\in\operatorname{span}(e_3)$ one has $\measuredangle(A^{i-1}v,A^iv)=0$, $i\in\mathbb{N}$. Next, we take a vector with components in both relevant subspaces. This vector is pushed under iteration with A towards the most unstable direction e_3. Thus, we expect that the angle between two subsequent iterates converges to 0. The following estimate proves that this convergence is indeed geometric. Consider $v=\binom{z}{1}$ with $0\neq z\in\mathbb{R}^2$. From the triangle inequality and the estimate (2.6) in Lemma 2.6 we find a constant $C>0$ such that for all $i\in\mathbb{N}$

$$\begin{aligned}\measuredangle(A^{i-1}v,A^iv)&=\measuredangle\left(\begin{pmatrix}T_\varphi^{i-1}z\\2^{i-1}\end{pmatrix},\begin{pmatrix}T_\varphi^iz\\2^i\end{pmatrix}\right)=\measuredangle\left(\begin{pmatrix}2^{1-i}T_\varphi^{i-1}z\\1\end{pmatrix},\begin{pmatrix}2^{-i}T_\varphi^iz\\1\end{pmatrix}\right)\\&\leq\measuredangle\left(\begin{pmatrix}2^{1-i}T_\varphi^{i-1}z\\1\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right)+\measuredangle\left(\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}2^{-i}T_\varphi^iz\\1\end{pmatrix}\right)\\&\leq\tan\measuredangle\left(\begin{pmatrix}2^{1-i}T_\varphi^{i-1}z\\1\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right)+\tan\measuredangle\left(\begin{pmatrix}2^{-i}T_\varphi^iz\\1\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right)\\&\leq C\cdot2^{-i}.\end{aligned}$$

Thus
$$\frac{1}{n}\sum_{i=1}^{n}\measuredangle(A^{i-1}v, A^i v) \le \frac{1}{n}\sum_{i=1}^{\infty} C\cdot 2^{-i} = \frac{2C}{n} \to 0 \text{ as } n\to\infty.$$
As a consequence, all first outer angular values from Definition 3.1 coincide and have the value φ, see Figure 3.1 and Theorem 5.7 for the inner angular values.

For analyzing the second outer angular values, we first note that for all $V \in \mathcal{G}(2,3)$ there exists a $u \in \operatorname{span}(e_1, e_2)$ such that $V = \operatorname{span}(u, v)$ with $v \in \mathbb{R}^3$. Without loss of generality, we assume that $u = e_1$. We observe for $v \in \operatorname{span}(e_1, e_2)$ that $a_{1,n}(V) = 0$ and for $v \in \operatorname{span}(e_3)$, we obtain $a_{1,n}(V) = \varphi$. Next, we consider the mixed case $v = \begin{pmatrix} z_1 & z_2 & 1\end{pmatrix}^\top$, with $0 \ne z \in \mathbb{R}^2$. Let $W = \operatorname{span}(e_1, e_3)$ then we get for $i \in \mathbb{N}$
$$\measuredangle(A^{i-1}V, A^i V) \le \measuredangle(A^{i-1}V, A^{i-1}W) + \measuredangle(A^{i-1}W, A^i W) + \measuredangle(A^i W, A^i V).$$
The second term is equal to φ for all $i \in \mathbb{N}$. We conclude that all second outer angular values coincide with φ by showing that the first and third term converge to zero with a geometric rate. Note that for $i \in \mathbb{N}_0$ we have
$$\begin{aligned} A^i V &= \operatorname{span}(A^i e_1, A^i e_3 + A^i \begin{pmatrix} z_1 & z_2 & 0\end{pmatrix}^\top) \\ &= \operatorname{span}(A^i e_1, e_3 + 2^{-i} A^i \begin{pmatrix} z_1 & z_2 & 0\end{pmatrix}^\top), \\ A^i W &= \operatorname{span}(A^i e_1, e_3). \end{aligned}$$
With $P_i = I + Q_i$, $Q_i = 2^{-i}A^i \begin{pmatrix} 0 & 0 & z_1 \\ 0 & 0 & z_2 \\ 0 & 0 & 0 \end{pmatrix}$ it follows that $A^i V = P_i A^i W$. Furthermore, we find an i-independent constant $C > 0$ such that $\|(I - P_i)v\| = \|Q_i v\| \le 2^{-i}C\|P_i v\|$ for all $v \in \mathbb{R}^3$. Thus, Lemma 2.6, (ii) applies for sufficiently large $i \in \mathbb{N}$ and provides the estimate
$$\measuredangle(A^i V, A^i W) \le 2^{-i}\frac{C}{(1 - 2^{-i}C)^{\frac{1}{2}}}$$
which completes the proof.

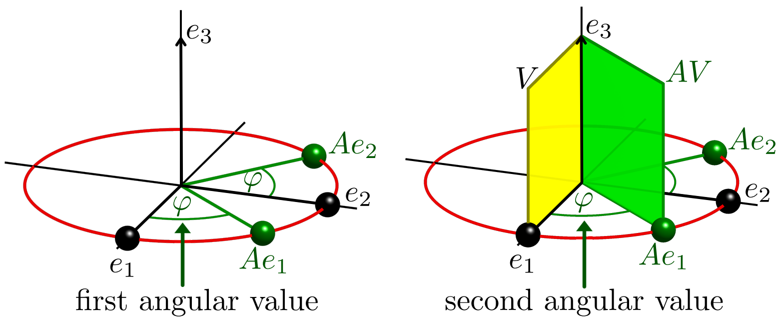

Figure 3.1: First and second angular values for the motivating three-dimensional system (3.7).

In general, equality does not hold in Diagram 3.1. This phenomenon is illustrated in Section 3.2 by Examples 3.10 and 3.11. However, angular values do agree when the angles of iterates or their averages occurring in Definition 3.1 have some uniformity properties. For this purpose let us introduce for $n \in \mathbb{N}$ the functions

$$b_n : \mathcal{G}(s,d) \to \mathbb{R}, \quad b_n(V) = \measuredangle(\Phi(n-1,0)V, \Phi(n,0)V), \tag{3.8}$$

and recall $a_{1,n}(V) = \sum_{j=1}^{n} b_j(V)$ from Definition 3.1. Let us also recall the notion of uniform almost periodicity for a sequence of functions.

Definition 3.4. *Given a set $\mathcal{V}$ and a Banach space $(\mathcal{W}, \|\cdot\|)$. A sequence of mappings $b_n : \mathcal{V} \to \mathcal{W}$, $n \in \mathbb{N}$ is called uniformly almost periodic if*

$$\begin{aligned} &\forall \varepsilon > 0\ \exists P \in \mathbb{N} : \forall V \in \mathcal{V}\ \forall \ell \in \mathbb{N}\ \exists p \in \{\ell, \dots, \ell + P\} : \\ &\forall n \in \mathbb{N} : \|b_n(V) - b_{n+p}(V)\| \le \varepsilon. \end{aligned}$$

Remark 3.5. *Our definition is slightly weaker than the standard notion ([28, Ch.4.1]) which requires for each $\varepsilon > 0$ the existence of a relatively dense set $\mathcal{P} \subset \mathbb{N}$ such that*

$$\forall n \in \mathbb{N}\ \forall p \in \mathcal{P}\ \forall V \in \mathcal{V} : \|b_n(V) - b_{n+p}(V)\| \le \varepsilon.$$

This is more restrictive, since the choice of $p \in \{\ell, \dots, \ell + P\} \cap \mathcal{P}$ is uniform in V.

The following Proposition 3.7 will be used repeatedly when determining angular values for the two-dimensional case; see Proposition 5.2. First, we state a crucial observation, which is proven in the Supplementary materials II.

Lemma 3.6. *Let $b_n : \mathcal{V} \to \mathcal{W}$, $n \in \mathbb{N}$ be a sequence of uniformly almost periodic and uniformly bounded functions. Then for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n \ge m \ge N$, $k \in \mathbb{N}$, $V \in \mathcal{V}$*

$$\Big\| \frac{1}{n} \sum_{j=1}^{n} b_j(V) - \frac{1}{m} \sum_{j=1}^{m} b_{j+k}(V) \Big\| \le \varepsilon.$$

Proposition 3.7. *The following statements hold for all $s \in \{1, \dots, d\}$.*

(a) If the functions $\frac{1}{n} a_{1,n} : \mathcal{G}(s,d) \to \mathbb{R}$ converge uniformly to the constant function $\varphi \in [0, \frac{\pi}{2}]$ as $n \to \infty$, then all nonuniform angular values coincide, i.e. $\bar{\theta}_s = \hat{\theta}_s = \underline{\theta}_s = \underset{\wedge}{\theta}_s = \varphi$.

(b) If the functions b_n, $n \in \mathbb{N}$ from (3.8) are uniformly almost periodic, then all angular values coincide,

$$\underset{\wedge}{\theta}_{[s]} = \underset{\wedge}{\theta}_s = \hat{\theta}_s = \hat{\theta}_{[s]} = \underline{\theta}_{[s]} = \underline{\theta}_s = \bar{\theta}_s = \bar{\theta}_{[s]}.$$

Proof. The claim in (a) is clear since lim sup and lim inf in (3.3) are limits and the supremum is continuous w.r.t. uniform convergence.

Lemma 3.3 shows that it suffices for (b) to prove $\bar{\theta}_{[s]} \le \underset{\wedge}{\theta}_{[s]}$. By the definition (3.4) we find for every $\varepsilon > 0$ a number $N_1 \in \mathbb{N}$ and for all $n \ge N_1$ elements $V_n \in \mathcal{G}(s,d)$, $k_n \in \mathbb{N}$ such that

$$\Big| \bar{\theta}_{[s]} - \frac{1}{n} \sum_{j=1}^{n} b_{j+k_n}(V_n) \Big| \le \frac{\varepsilon}{2}.$$

From Lemma 3.6 we obtain $N = N(\varepsilon) \in \mathbb{N}$, $N \ge N_1$ such that for all $n \ge N$, $h \in \mathbb{N}_0$

$$\Big| \frac{1}{n} \sum_{j=1}^{n} b_j(V_N) - \frac{1}{N} \sum_{j=1}^{N} b_{j+k_N}(V_N) \Big| + \Big| \frac{1}{n} \sum_{j=1}^{n} b_j(V_N) - \frac{1}{n} \sum_{j=1}^{n} b_{j+h}(V_N) \Big| \le \frac{\varepsilon}{2}.$$

Combining these results yields for all $h \in \mathbb{N}_0$, $n \geq N$:

$$\begin{aligned}\bar{\theta}_{[s]} &\leq \frac{1}{n}\sum_{j=1}^{n} b_{j+h}(V_N) + \Big|\frac{1}{n}\sum_{j=1}^{n} b_j(V_N) - \frac{1}{n}\sum_{j=1}^{n} b_{j+h}(V_N)\Big| \\ &+ \Big|\frac{1}{N}\sum_{j=1}^{N} b_{j+k_N}(V_N) - \frac{1}{n}\sum_{j=1}^{n} b_j(V_N)\Big| + \Big|\bar{\theta}_{[s]} - \frac{1}{N}\sum_{j=1}^{N} b_{j+k_N}(V_N)\Big| \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} + \frac{1}{n}\sum_{j=1}^{n} b_{j+h}(V_N).\end{aligned}$$

Taking the infimum over h and the limit $n \to \infty$ (see Lemma 3.3 for its existence) shows that there exists $V(\varepsilon) := V_{N(\varepsilon)} \in \mathcal{G}(s,d)$ satisfying

$$\bar{\theta}_{[s]} - \varepsilon \leq \lim_{n\to\infty} \inf_{h\in\mathbb{N}_0} \frac{1}{n}\sum_{j=1}^{n} b_{j+h}(V(\varepsilon)).$$

Thus we deduce

$$\bar{\theta}_{[s]} - \varepsilon \leq \sup_{V\in\mathcal{G}(s,d)} \lim_{n\to\infty} \inf_{h\in\mathbb{N}_0} \frac{1}{n}\sum_{j=1}^{n} b_{j+h}(V) = \underset{\frown}{\theta}_{[s]} \leq \bar{\theta}_{[s]}.$$

■

Next, we apply a kinematic similarity, induced by a transformation $\tilde{u}_n = Q_n u_n$ with $Q_n \in \mathrm{GL}(\mathbb{R}^d)$ to (1.1), i.e. we consider

$$\tilde{u}_{n+1} = \tilde{A}_n \tilde{u}_n, \quad \tilde{A}_n = Q_{n+1} A_n Q_n^{-1}, \tag{3.9}$$

and ask when angular values remain unchanged.

Proposition 3.8. *(Invariance of angular values)*

(i) Assume $Q_n = r_n Q$, $n \in \mathbb{N}$ with $r_n \in \mathbb{R}$, $r \neq 0$ and $Q \in \mathbb{R}^{d,d}$ orthogonal. Then the angular values of (1.1) and (3.9) agree.

(ii) Assume constant transformation matrices $Q_n = Q, n \in \mathbb{N}_0$ with Q invertible. If any of the values $\theta_s \in \{\underset{\frown}{\theta}_{[s]}, \underset{\frown}{\theta}_s, \hat{\theta}_s, \hat{\theta}_{[s]}, \underline{\theta}_{[s]}, \underline{\theta}_s, \bar{\theta}_s, \bar{\theta}_{[s]}\}$ in Definition 3.1 vanishes for the system (1.1) then the same angular value vanishes for the transformed system (3.9).

Proof. First note that the solution operators Φ_A, $\Phi_{\tilde{A}}$ of (1.1), (3.9) are related by

$$\Phi_{\tilde{A}}(n,m)Q_m = Q_n\Phi_A(n,m), \quad n,m \in \mathbb{N}_0. \tag{3.10}$$

The result of (i) follows from (3.10) and the invariance of angles under scalings and orthogonal transformations (cf. Proposition 2.2)

$$\begin{aligned}\measuredangle(\Phi_{\tilde{A}}(j-1,0)QV, \Phi_{\tilde{A}}(j,0)QV) &= \measuredangle(\frac{r_{j-1}}{r_0}Q\Phi_A(j-1,0)V, \frac{r_j}{r_0}Q\Phi_A(j,0)V) \\ &= \measuredangle(\Phi_A(j-1,0)V, \Phi_A(j,0)V).\end{aligned}$$

For case (ii) the relation (3.10) reads $\Phi_{\tilde{A}}(j,0)Q = Q\Phi_A(j,0)$ and the assertion follows from the angle-boundedness (2.9) of the matrix $S = Q$. ■

One can strengthen Proposition 3.8 as follows. For assertion (i) it is sufficient if $Q_n = r_n P_n$ where P_n converges to some orthogonal matrix, and for assertion (ii) it is sufficient if Q_n converges to some invertible matrix. However, we do not expect substantially more general transformations to leave all angular values invariant. For example, if a kinematic similarity preserves the single terms $\measuredangle(u_n, A_n u_n) = \measuredangle(Q_n u_n, Q_{n+1} A_n u_n)$, for all $n \in \mathbb{N}_0$ and if a condition is desired which does not depend on the particular choice of A_n, one is led to the property $\measuredangle(Q_n u, Q_{n+1} v) = \measuredangle(u, v)$ for all $u, v \in \mathbb{R}^d$, $n \in \mathbb{N}_0$. The latter condition implies that all matrices Q_n, $n \in \mathbb{N}_0$ are multiples of a common orthogonal matrix.

Finally, we discuss an invariance property of maximizers which occur with the outer angular values. Starting with the difference equation (1.1), we define for $\eta, n \in \mathbb{N}_0$ the matrices $A_n(\eta) := A_{n+\eta}$. Denote by Φ^+_η the solution operator of the shifted difference equation

$$u_{n+1} = A_{n+\eta} u_n, \quad n \in \mathbb{N}_0 \tag{3.11}$$

and observe that for all $n, m, \eta \in \mathbb{N}_0$

$$\Phi^+_\eta(n, m) = \Phi(n + \eta, m + \eta).$$

Let $\hat{\theta}_s(\eta)$ be the s-th upper outer angular value for (3.11). The corresponding maximizers that occur with the outer values are given by

$$\hat{\mathcal{V}}_s(\eta) = \Big\{ V \in \mathcal{G}(s, d) : \hat{\theta}_s(\eta) = \limsup_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \measuredangle(\Phi^+_\eta(j-1, 0)V, \Phi^+_\eta(j, 0)V) \Big\}.$$

Note that this set may be empty. We obtain the following invariance.

Proposition 3.9. *Let $A_n \in \mathbb{R}^{d,d}$, $n \in \mathbb{N}_0$ be invertible matrices. Then the following relation holds for all $\eta \in \mathbb{N}$,*

$$A_\eta \hat{\mathcal{V}}_s(\eta) = \hat{\mathcal{V}}_s(\eta + 1). \tag{3.12}$$

Proof. Fix $\eta \in \mathbb{N}$ and let $V \in \mathcal{G}(s, d)$. Then we get

$$\begin{aligned}
&\frac{1}{n} \sum_{j=1}^{n} \measuredangle(\Phi^+_{\eta+1}(j-1, 0)A_\eta V, \Phi^+_{\eta+1}(j, 0)A_\eta V) \\
&= \frac{1}{n} \sum_{j=1}^{n} \measuredangle(\Phi(j + \eta, \eta + 1)A_\eta V, \Phi(j + \eta + 1, \eta + 1)A_\eta V) \\
&= \frac{1}{n} \sum_{j=1}^{n} \measuredangle(\Phi(j + \eta, \eta)V, \Phi(j + \eta + 1, \eta)V) \\
&= \frac{n+1}{n} \frac{1}{n+1} \left(\sum_{j=1}^{n+1} \measuredangle(\Phi^+_\eta(j-1, 0)V, \Phi^+_\eta(j, 0)V) - \measuredangle(V, \Phi^+_\eta(1, 0)V) \right).
\end{aligned}$$

Taking $\limsup$ as $n \to \infty$ we have

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \measuredangle(\Phi^+_{\eta+1}(j-1, 0)A_\eta V, \Phi^+_{\eta+1}(j, 0)A_\eta V) \\
&= \limsup_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \measuredangle(\Phi^+_\eta(j-1, 0)V, \Phi^+_\eta(j, 0)V).
\end{aligned}$$

In the case $\hat{\mathcal{V}}_s(\eta) = \emptyset$ then $\hat{\mathcal{V}}_s(\eta+1) = \emptyset$ and (3.12) is trivial. Otherwise, the invertibility of A_η yields

$$V \in \hat{\mathcal{V}}_s(\eta) \Leftrightarrow A_\eta V \in \hat{\mathcal{V}}_s(\eta+1)$$

which proves (3.12). ■

A corresponding result also holds for lower outer angular values as well as for uniform outer angular values.

3.2. Some nonautonomous key examples. Upper, lower, uniform respectively non-uniform outer and inner angular values do not coincide in general. The following examples illustrate this fact.

First, we construct an example which possesses different upper, lower and uniform angular values. A related example in continuous time can be found in [12, Example 2.2]. There, the authors illustrate that the Lyapunov spectrum may be a proper subset of the Sacker-Sell spectrum and that generally both spectra do not consist of isolated points only.

Example 3.10. *Fix $0 \le \varphi_0 < \varphi_1 \le \frac{\pi}{2}$ and let $T_\varphi := \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$. For $n \in \mathbb{N}_0$, we define*

$$A_n = \begin{cases} T_{\varphi_0}, & \text{for } n = 0 \vee n \in \bigcup_{\ell=1}^{\infty} [2^{2\ell-1}, 2^{2\ell}-1] \cap \mathbb{N}, \\ T_{\varphi_1}, & \text{otherwise.} \end{cases}$$

Table 3.2 illustrates this construction.

n	0	1	2	3	4	5	6	7	8	...	15	16
A_n	T_{φ_0}	T_{φ_1}	T_{φ_0}	T_{φ_0}	T_{φ_1}	T_{φ_1}	T_{φ_1}	T_{φ_1}	T_{φ_0}	...	T_{φ_0}	T_{φ_1}

Table 3.2: Construction of $(A_n)_{n\in\mathbb{N}_0}$.

Inner and outer angular values coincide for the nonautonomous difference equation

$$u_{n+1} = A_n u_n, \quad n \in \mathbb{N}_0,$$

since all one-dimensional subspaces rotate through the same angle.

Denote by p_ℓ the number of occurrences of T_{φ_1} in $(A_n)_{0\le n\le \ell}$. One observes for $n \in \mathbb{N}$ that

$$p_{2^{2n-1}-1} = \frac{1}{3}(4^n - 1) = p_{2^{2n}-1}$$

and

$$\lim_{n\to\infty} \frac{1}{2^{2n-1}-1} p_{2^{2n-1}-1} = \frac{2}{3}, \quad \frac{1}{2^{2n}-1} p_{2^{2n}-1} = \frac{1}{3}.$$

Thus, we obtain

$$\underline{\theta}_1 = \underset{\wedge}{\theta}_1 = \frac{2}{3}\varphi_0 + \frac{1}{3}\varphi_1, \quad \bar{\theta}_1 = \hat{\theta}_1 = \frac{1}{3}\varphi_0 + \frac{2}{3}\varphi_1.$$

For each $n \in \mathbb{N}$, we find infinitely many indices $\nu \in \mathbb{N}$ such that $A_{\nu+\ell} = T_{\varphi_0}$ (resp. $A_{\nu+\ell} = T_{\varphi_1}$) for all $\ell = 0, \dots, n-1$. As a consequence, the Diagram 3.1 has the explicit form in Diagram 3.3.

Although inner and outer angular values coincide for Example 3.10, this coincidence is in general not true. We discuss the following example.

$$\begin{array}{ccccccc}
\varphi_0 = \underset{\wedge}{\theta}_{[1]} & < & \frac{2}{3}\varphi_0 + \frac{1}{3}\varphi_1 = \underset{\wedge}{\theta}_1 & < & \frac{1}{3}\varphi_0 + \frac{2}{3}\varphi_1 = \hat{\theta}_1 & < & \hat{\theta}_{[1]} = \varphi_1 \\
\| & & \| & & \| & & \| \\
\varphi_0 = \underline{\theta}_{[1]} & < & \frac{2}{3}\varphi_0 + \frac{1}{3}\varphi_1 = \underline{\theta}_1 & < & \frac{1}{3}\varphi_0 + \frac{2}{3}\varphi_1 = \bar{\theta}_1 & < & \bar{\theta}_{[1]} = \varphi_1
\end{array}$$

Diagram 3.3: Angular values of Example 3.10.

Example 3.11. *Let*

$$C := \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \quad R := \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In the case of the reflection R, we observe for $v = \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}$, $\phi \in [0, \frac{\pi}{2}]$ that

$$\measuredangle(v, Rv) = \begin{cases} 2\phi, & \text{for } 0 \le \phi \le \frac{\pi}{4}, \\ \pi - 2\phi, & \text{for } \frac{\pi}{4} < \phi \le \frac{\pi}{2} \end{cases}$$

and the maximal angle is achieved at $v \in \operatorname{span}\{\binom{1}{1}\}$.

For $n \in \mathbb{N}_0$, we define

$$A_n := \begin{cases} R, & \text{for } n \in \bigcup_{\ell=1}^{\infty} [2 \cdot 2^\ell - 4, 3 \cdot 2^\ell - 5], \\ C, & \text{otherwise.} \end{cases}$$

Table 3.4 illustrates this construction.

n	0	1	2	3	4	5	6	7	8	9	10	11	12	…	19	20
A_n	R	R	C	C	R	R	R	R	C	C	C	C	R	…	R	C

Table 3.4: Construction of $(A_n)_{n \in \mathbb{N}_0}$.

We prove that inner and outer angular values of the nonautonomous difference equation

$$u_{n+1} = A_n u_n, \quad n \in \mathbb{N}_0$$

do not coincide. First we show that $\hat{\theta}_1 = 0$. Let

$$V_\phi := \operatorname{span}\begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}, \quad b_j(\phi) = \measuredangle(\Phi(j-1, 0)V_\phi, \Phi(j, 0)V_\phi).$$

For $\phi \in \{0, \frac{\pi}{2}\}$ we get $b_j(\phi) = 0$ for all $j \in \mathbb{N}$. In the case $\phi \in (0, \frac{\pi}{2})$ we observe that $\Phi(j, 0)V_\phi \to V_0$ as $j \to \infty$. Thus for each $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that $b_j(\phi) \le \varepsilon$ for all $j \ge N$. As a consequence we get for n sufficiently large

$$\frac{1}{n}\sum_{j=1}^{n} b_j(\phi) = \frac{1}{n}\left(\sum_{j=1}^{N-1} b_j(\phi) + \sum_{j=N}^{n} b_j(\phi)\right) \le \frac{1}{n}\Big((N-1)\frac{\pi}{2} + (n+1-N)\varepsilon\Big) \le 2\varepsilon$$

and this shows

$$\hat{\theta}_1 = \sup_{V\in\mathcal{G}(1,2)} \limsup_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \measuredangle(\Phi(j-1,0)V, \Phi(j,0)V) = 0.$$

Similarly, all outer angular values are zero.

Next, we determine an estimate for the upper inner angular value. We claim that $\bar{\theta}_1 \geq \frac{\pi}{6}$.

For $\ell \in \mathbb{N}$ *let* $p_\ell := 3\cdot 2^\ell - 5$. *Note that the matrix* C *appears* $2^\ell - 2$ *times in* $(A_n)_{0\leq n\leq p_\ell}$ *and* R *appears* 2^ℓ *times in* $(A_n)_{p_{\ell-1}<n\leq p_\ell}$.

Let $V(\ell) := \operatorname{span}\{C^{-2^\ell+2}\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)\}$. *We obtain*

$$\begin{aligned}\bar{\theta}_1 &= \limsup_{n\to\infty} \sup_{v\in\mathcal{G}(1,2)} \frac{1}{n}\sum_{j=1}^{n} \measuredangle(\Phi(j-1,0)V, \Phi(j,0)V)\\ &\geq \limsup_{\ell\to\infty} \frac{1}{p_\ell+1}\sum_{j=1}^{p_\ell+1} \measuredangle(\Phi(j-1,0)V(\ell), \Phi(j,0)V(\ell))\\ &\geq \limsup_{\ell\to\infty} \frac{1}{p_\ell+1} 2^\ell \frac{\pi}{2} = \lim_{\ell\to\infty} \frac{2^\ell}{3\cdot 2^\ell - 4}\cdot\frac{\pi}{2} = \frac{\pi}{6}.\end{aligned}$$

We obtain estimates for $\underline{\theta}_1$ *by analyzing the subsequence* $n = p_\ell - 2^\ell + 1$ *for* $\ell \in \mathbb{N}$. *These indices detect the end of each block of* C*s. In particular, we observe that* $\frac{\pi}{12} \leq \underline{\theta}_1 < \frac{\pi}{6}$ *and present results for all angular values in Diagram 3.5.*

$$\begin{array}{ccccccc} 0 = \underset{\hat{}}{\theta}_{[1]} & = & 0 = \underset{\hat{}}{\theta}_1 & = & 0 = \hat{\theta}_1 & = & \hat{\theta}_{[1]} = 0 \\ \| & & \wedge & & \wedge & & \wedge \\ 0 = \underline{\theta}_{[1]} & < & \frac{\pi}{12} \leq \underline{\theta}_1 & < & \frac{\pi}{6} \leq \bar{\theta}_1 & < & \bar{\theta}_{[1]} = \frac{\pi}{2} \end{array}$$

Diagram 3.5: Angular values of Example 3.11.

4. Angular values of random linear cocycles. Following [2, Ch.3.3.1] we consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $T : \Omega \circlearrowleft$ be a measurable, $\mathbb{P}$-preserving, ergodic transformation. Let $A : \Omega \to \mathrm{GL}(\mathbb{R}^d)$ and set $A_\omega^{(n)} = A(T^{n-1}\omega)\cdots A(T\omega)A(\omega)$. Note that $A_\omega^{(n)}$ corresponds to a random solution operator $\Phi(n, 0, \omega)$ in the setting of Section 3; cf. [2, (3.3.2)]. In analogy to the right-hand side of (1.3), for $n \geq 1$, define for $s \in \{1, \ldots, d\}$ and $V \in \mathcal{G}(s, d)$

$$a_n(\omega, V) = \sum_{j=0}^{n-1} \measuredangle(A_\omega^{(j)}V, A_\omega^{(j+1)}V).$$

Define a skew product $\tau : \Omega\times\mathcal{G}(s,d) \circlearrowleft$ by $\tau(\omega, V) = (T\omega, A(\omega)V)$ and $f : \Omega\times\mathcal{G}(s,d) \to \mathbb{R}$ by $f(\omega, V) = \measuredangle(V, A(\omega)V)$. One has the Birkhoff sum representation:

$$a_n(\omega, V) = \sum_{j=0}^{n-1} f(\tau^j(\omega, V)).$$

The following result provides general conditions for the angular value limits to be independent of the initial condition ω and reference subspace V. We use the notation ϑ_s for angular values in this setting, to distinguish them from the angular values θ_s in Section 3.

Theorem 4.1. *Suppose τ preserves an ergodic probability measure μ on $\Omega \times \mathcal{G}(s,d)$, where μ has marginal $\mathbb{P}$ on Ω; that is, $\mu(\cdot, \mathcal{G}(s,d)) = \mathbb{P}$.*
Then there is a $\vartheta_s \in [0, \frac{\pi}{2}]$ satisfying

$$\vartheta_s = \lim_{n\to\infty} \frac{1}{n} a_n(\omega, V) = \int_{\Omega\times\mathcal{G}(s,d)} \measuredangle(V, A(\omega)V)\ d\mu(\omega, V),$$

for μ-almost every $(\omega, V) \in \Omega \times \mathcal{G}(s,d)$.

Proof. This follows immediately from Birkhoff's ergodic theorem and ergodicity of τ. ■

The existence of an ergodic invariant measure μ for τ is connected with a certain irreducibility condition on the action on subspaces, leading to an independence of the angular values with respect to V. We would like to treat general linear cocycles and so in analogy to Section 3.1, we consider angular values of a random linear cocycle w.r.t. $\omega \in \Omega$ and ask for extreme values w.r.t. $V \in \mathcal{G}(s,d)$. To keep the analogy with Definition 3.1 we use an up/down bar for upper/lower inner angular values and an up/down hat for upper/lower outer angular values, while their uniform equivalents are denoted by the bracketed index $[s]$.

Theorem 4.2. *Let $T : \Omega \circlearrowleft$ be a measurable, $\mathbb{P}$-preserving, ergodic transformation and let $A : \Omega \to \mathrm{GL}(\mathbb{R}^d)$. Then the following assertions hold.*

1. *There is a number $\bar{\vartheta}_s$ such that for $\mathbb{P}$-a.e. ω,*

$$\bar{\vartheta}_s = \lim_{n\to\infty} \max_{V\in\mathcal{G}(s,d)} \frac{a_n(\omega, V)}{n} = \inf_{n\in\mathbb{N}} \frac{1}{n} \int_\Omega \max_{V\in\mathcal{G}(s,d)} a_n(\omega, V)\ d\mathbb{P}(\omega). \tag{4.1}$$

In particular, one has

$$\bar{\vartheta}_s \le \int_\Omega \max_{V\in\mathcal{G}(s,d)} \measuredangle(V, A(\omega)V)\ d\mathbb{P}(\omega).$$

2. *There is a number $\hat{\vartheta}_s$ such that for $\mathbb{P}$-a.e. ω,*

$$\hat{\vartheta}_s = \sup_{V\in\mathcal{G}(s,d)} \limsup_{n\to\infty} \frac{a_n(\omega, V)}{n}.$$

Furthermore, if for $\mathbb{P}$-a.e. ω the supremum over V is achieved by at most $K < \infty$ subspaces $V_1(\omega), \dots, V_K(\omega)$, then one may create K equivariant collections $\{V_k(\omega)\}_{1\le k\le K, \omega\in\Omega}$, satisfying $V_k(T\omega) = A(\omega)V_k(\omega)$, $k = 1, \dots, K$.

3. *There is a number $\underset{\wedge}{\vartheta}_s$ such that for $\mathbb{P}$-a.e. ω,*

$$\underset{\wedge}{\vartheta}_s = \sup_{V\in\mathcal{G}(s,d)} \liminf_{n\to\infty} \frac{a_n(\omega, V)}{n}.$$

Furthermore, if for $\mathbb{P}$-a.e. ω the supremum over V is achieved by at most $K < \infty$ subspaces $V_1(\omega), \dots, V_K(\omega)$, then one may create K equivariant collections $\{V_k(\omega)\}_{1\le k\le K, \omega\in\Omega}$, satisfying $V_k(T\omega) = A(\omega)V_k(\omega)$, $k = 1, \dots, K$.

4. *There is a number $\bar{\vartheta}_{[s]}$ such that*

$$\bar{\vartheta}_{[s]} = \lim_{n\to\infty} \sup_{V\in\mathcal{G}(s,d)} \operatorname*{ess\,sup}_{\omega\in\Omega} \frac{a_n(\omega,V)}{n} = \inf_{n\to\infty} \operatorname*{ess\,sup}_{\omega\in\Omega} \max_{V\in\mathcal{G}(s,d)} \frac{a_n(\omega,V)}{n}.$$

In particular, one has

$$\bar{\vartheta}_{[s]} \leq \operatorname*{ess\,sup}_{\omega\in\Omega} \max_{V\in\mathcal{G}(s,d)} \measuredangle(V, A(\omega)V).$$

5. *There is a number $\underset{\wedge}{\vartheta}_{[s]}$ such that*

$$\underset{\wedge}{\vartheta}_{[s]} = \sup_{V\in\mathcal{G}(s,d)} \lim_{n\to\infty} \operatorname*{ess\,inf}_{\omega\in\Omega} \frac{a_n(\omega,V)}{n} = \sup_{V\in\mathcal{G}(s,d)} \sup_{n\geq 1} \operatorname*{ess\,inf}_{\omega\in\Omega} \frac{a_n(\omega,V)}{n}.$$

In particular, one has

$$\sup_{V\in\mathcal{G}(s,d)} \operatorname*{ess\,inf}_{\omega\in\Omega} \measuredangle(V, A(\omega)V) \leq \underset{\wedge}{\vartheta}_{[s]}.$$

6. *There is a number $\hat{\vartheta}_{[s]}$ such that*

$$\hat{\vartheta}_{[s]} = \sup_{V\in\mathcal{G}(s,d)} \lim_{n\to\infty} \operatorname*{ess\,sup}_{\omega\in\Omega} \frac{a_n(\omega,V)}{n} = \sup_{V\in\mathcal{G}(s,d)} \inf_{n\geq 1} \operatorname*{ess\,sup}_{\omega\in\Omega} \frac{a_n(\omega,V)}{n}.$$

In particular, one has

$$\hat{\vartheta}_{[s]} \leq \sup_{V\in\mathcal{G}(s,d)} \operatorname*{ess\,sup}_{\omega\in\Omega} \measuredangle(V, A(\omega)V).$$

Remark 4.3. *In Part 1 of Theorem 4.2, if Ω is a metric space, $\omega \mapsto A(\omega)$ is continuous, and $T : \Omega \circlearrowleft$ is uniquely ergodic, then using the fact that $\omega \mapsto \max_V \measuredangle(V, A(\omega)V)$ is continuous, we obtain the following stronger conclusion. By Theorem 1.5 [33], the limit in (4.1) converges in a semi-uniform way (the limit exists for* all *$\omega \in \Omega$ and converges uniformly in ω): given $\epsilon > 0$, there exists n_0 such that for all $n \geq n_0$,*

$$\bar{\vartheta}_s \leq \max_{V\in\mathcal{G}(s,d)} \frac{a_n(\omega,V)}{n} \leq \bar{\vartheta}_s + \epsilon \qquad \textit{for all } \omega \in \Omega.$$

A simple example of such a system is an irrational rotation on the unit circle $\Omega = S^1$ and $T\omega = \omega + \varphi$, where $\varphi \notin \mathbb{Q}$. The Lebesgue measure on S^1 is the unique invariant probability measure. One may choose any *continuous matrix-valued function A. More generally, one may consider rationally independent translations on higher-dimensional tori.*

Proof of Theorem 4.2.

1. We note that the invertibility of the matrices $A(\omega)$ implies $A(\omega)\mathcal{G}(s,d) = \mathcal{G}(s,d)$ for $\mathbb{P}$ a.e. ω. As in the proof of Lemma 3.3 one can easily show that

$$\max_{V\in\mathcal{G}(s,d)} a_{n+m}(\omega,V) \leq \max_{V\in\mathcal{G}(s,d)} a_n(\omega,V) + \max_{V\in\mathcal{G}(s,d)} a_m(T^n\omega,V)$$

for every $n,m \geq 0$, and therefore $g_n(\omega) := \max_{V\in\mathcal{G}(s,d)} a_n(\omega,V)$ is a subadditive sequence of functions. Recall that $0 \leq a_n(\omega,V) \leq \frac{\pi}{2}$ for all ω, V, implying $0 \leq g_n \leq \frac{\pi}{2}$. The results are now immediate by the subadditive ergodic theorem applied to g_n, using ergodicity of $\mathbb{P}$.

2. By direct computation, one verifies that

$$a_n(T\omega, A(\omega)V) = a_n(\omega, V) - \measuredangle(V, A(\omega)V) + \measuredangle(A_\omega^{(n)}V, A_\omega^{(n+1)}V).$$

Thus, because values of angles are bounded, one has

$$\limsup_{n\to\infty} \frac{1}{n} a_n(T\omega, A(\omega)V) = \limsup_{n\to\infty} \frac{1}{n} a_n(\omega, V) =: g(\omega, V), \tag{4.2}$$

where g is measurable in ω and V. By the invertibility of $A(\omega)$ for a.e. ω, we see that

$$\begin{aligned} h(\omega) &:= \sup_{V\in\mathcal{G}(s,d)} g(\omega, V) = \sup_{V\in\mathcal{G}(s,d)} \limsup_{n\to\infty} \frac{1}{n} a_n(T\omega, A(\omega)V) \\ &= \sup_{V\in\mathcal{G}(s,d)} \limsup_{n\to\infty} \frac{1}{n} a_n(T\omega, V) = h(T\omega). \end{aligned}$$

By ergodicity, the T-invariant function h is constant a.e. The expression (4.2) demonstrates equivariance of a particular maximizing V (if it exists). By hypothesis for $\mathbb{P}$-a.e. ω there are at most K distinct subspaces $\tilde{V}$ satisfying $\sup_{V\in\mathcal{G}(s,d)} g(\omega, V) = g(\omega, \tilde{V})$. By ergodicity, the invertibility of the $A(\omega)$, and the equivariance property, the number of solutions must be independent of ω on a full $\mathbb{P}$-measure set; let us call this number K. The equivariance property allows us to "match" the K pointwise solutions to create K families of maximizing subspaces $V_1(\omega), \dots, V_K(\omega)$ obeying equivariance.
3. This proof is analogous to Part 2.
4. Since $V \mapsto \frac{1}{n} a_n(\omega, V)$ is continuous for each $n \in \mathbb{N}$ and $\mathbb{P}$-a.e. ω, and $\mathcal{G}(s,d)$ is compact, we may replace the $\max_{V\in\mathcal{G}(s,d)}$ with $\sup_{V\in\mathcal{G}(s,d)}$ in all statements of Part 4. One may now interchange the operations $\operatorname{ess\,sup}_\omega$ and $\sup_{V\in\mathcal{G}(s,d)}$. Similarly to the proof of Part 1, one shows that $\sup_{V\in\mathcal{G}(s,d)} \operatorname{ess\,sup}_\omega a_n(\omega, V)$ is a subadditive sequence. Then the results follow immediately from Fekete's subadditive lemma.
5. We note that for fixed V we obtain a superadditive sequence of numbers $g_n(V) := \operatorname{ess\,inf}_\omega a_n(\omega, V)$. By Fekete's superadditive lemma one has $\lim_{n\to\infty} g_n(V)$ exists and equals $\sup_{n\in\mathbb{N}} g_n(V)$. This proves all statements concerning $\underset{\wedge}{\vartheta}_{[s]}$.
6. The results for $\hat{\vartheta}_{[s]}$ follow similarly, replacing superadditivity with subadditivity. ■

In the following we compare the various angular values as in Diagram 3.1. First note that we have a limit in equation (4.1). Therefore, it is unnecessary to distinguish between upper and lower angular values $\bar{\vartheta}_s$, $\underline{\vartheta}_s$ as in Diagram 3.1 for the nonautonomous case. To complete the following diagram, we introduce the lower uniform inner angular value

$$\underline{\vartheta}_{[s]} = \liminf_{n\to\infty} \sup_{V\in\mathcal{G}(s,d)} \operatorname*{ess\,inf}_{\omega\in\Omega} \frac{a_n(\omega, V)}{n},$$

which does not appear in Theorem 4.2.

Lemma 4.4. *Let the assumptions of Theorem 4.2 hold. Then the angular values defined above are related by Diagram 4.1.*

Proof. The proof of the inequalities in Diagram 4.1 is similar to the proof of Lemma 3.3. ■

$$
\begin{array}{ccccccc}
\underaccent{\hat}{\vartheta}_{[s]} & \leq & \underaccent{\hat}{\vartheta}_{s} & \leq & \hat{\vartheta}_{s} & \leq & \hat{\vartheta}_{[s]} \\
\rotatebox[origin=c]{90}{$\geq$} & & \rotatebox[origin=c]{90}{$\geq$} & & \rotatebox[origin=c]{90}{$\geq$} & & \rotatebox[origin=c]{90}{$\geq$} \\
\underline{\vartheta}_{[s]} & \leq & \bar{\vartheta}_{s} & = & \bar{\vartheta}_{s} & \leq & \bar{\vartheta}_{[s]}
\end{array}
$$

Diagram 4.1: Comparison of angular values for random dynamical systems.

In the special case where Ω consists of a single point, we are in the autonomous setting with a single matrix A. We may apply the results of Theorem 4.2 Parts 1–3 to obtain the following corollary.

Corollary 4.5. *Let* $a_n(V) = \sum_{j=0}^{n-1} \measuredangle(A^j V, A^{j+1} V)$ *with* $A \in \mathrm{GL}(\mathbb{R}^d)$ *and* $V \in \mathcal{G}(s,d)$. *Then the following holds:*

1. *The limit*
$$\bar{\vartheta}_s := \lim_{n\to\infty} \max_{V\in\mathcal{G}(s,d)} \frac{a_n(V)}{n} \text{ exists and equals } \inf_{n\in\mathbb{N}} \max_{V\in\mathcal{G}(s,d)} \frac{a_n(V)}{n}.$$
In particular, one has
$$\bar{\vartheta}_s \leq \max_{V\in\mathcal{G}(s,d)} \measuredangle(V, AV).$$
2. *There is a number* $\hat{\vartheta}_s$ *such that*
$$\hat{\vartheta}_s = \sup_{V\in\mathcal{G}(s,d)} \limsup_{n\to\infty} \frac{a_n(V)}{n}.$$
Furthermore, if the supremum over V *is achieved by a subspace* V *then the supremum is also achieved by* $A^j V$ *for all* $j \in \mathbb{Z}$.
3. *There is a number* $\underaccent{\hat}{\vartheta}_s$ *such that*
$$\underaccent{\hat}{\vartheta}_s = \sup_{V\in\mathcal{G}(s,d)} \liminf_{n\to\infty} \frac{a_n(V)}{n}.$$
Furthermore, if the supremum over V *is achieved by a subspace* V *then the supremum is also achieved by* $A^j V$ *for all* $j \in \mathbb{Z}$.

Finally, to contrast the random setting with the nonautonomous setting, let us reexamine the nonautonomous Examples 3.10 and 3.11. There we found $\underline{\theta}_1 < \bar{\theta}_1$ for the nonautonomous inner angular values in Diagrams 3.3 and 3.5. However, such a distinction is unnecessary in the random setting, see Diagram 4.1. Therefore, the $0, 1$ sequences underlying the choice of matrices in Tables 3.2 and 3.4 cannot occur for a set of full measure with an ergodic measure-preserving map.

5. Angular values for the autonomous case. A linear dynamical system, generated by a single matrix $A \in \mathrm{GL}(\mathbb{R}^d)$, fits into both – the nonautonomous setting of Section 3 and the random setting of Section 4. Therefore, the various angular values θ_s from Section 3 and ϑ_s from Section 4 coincide. Even for this case the computation of angular values turns out to be nontrivial. Since we will vary the matrix A, we write $\theta_s(A)$ to indicate the dependence of the angular values on the matrix.

The following Corollary collects some equalities in Diagram 3.1 for autonomous systems.

Corollary 5.1. *For an invertible autonomous system the following equalities hold for the values from Definition 3.1*

$$\underline{\theta}_s(A) = \bar{\theta}_s(A) = \bar{\theta}_{[s]}(A) = \lim_{n\to\infty} \frac{1}{n} \sup_{V\in\mathcal{G}(s,d)} \sum_{j=1}^{n} \measuredangle(A^{j-1}V, A^j V). \tag{5.1}$$

Proof. The existence of the limit on the RHS of (5.1) is due to Part 1 of Corollary 4.5. Its existence can also be derived directly from Fekete's subadditive lemma. The fact that $\bar{\theta}_s(A)$ is equal to the limit on the RHS of (5.1) is by definition in Part 1 of Corollary 4.5. The equality $\underline{\theta}_s(A) = \bar{\theta}_s(A)$ is trivial because we are discussing limits rather than limsup or liminf. The equality $\bar{\theta}_s(A) = \bar{\theta}_{[s]}(A)$ follows from Part 4 of Theorem 4.2, since Ω consists of a single point. It also follows from Definition 3.1 and the identity

$$a_{k+1,k+n}(V) = \sum_{j=k+1}^{k+n} \measuredangle(A^{j-1}V, A^j V) = \sum_{\nu=1}^{n} \measuredangle(A^{\nu-1}A^k V, A^\nu A^k V) = a_{1,n}(A^k V). \qquad \blacksquare$$

Next, we determine some explicit formulas for angular values in the autonomous case. Proposition 3.8(i) shows that we can assume A to be in real Schur form (cf. [20, Theorem 2.3.4]), i.e. A is quasi-upper triangular

$$A = \begin{pmatrix} \Lambda_1 & A_{12} & \cdots & A_{1k} \\ 0 & \Lambda_2 & \cdots & A_{2k} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \Lambda_k \end{pmatrix}, \quad \Lambda_i \in \mathbb{R}^{d_i,d_i}, A_{ij} \in \mathbb{R}^{d_i,d_j}, \tag{5.2}$$

where either $d_i = 1$ and $\Lambda_i = \lambda_i \in \mathbb{R}$ is a real eigenvalue or $d_i = 2$ and

$$\Lambda_i = \begin{pmatrix} \mathrm{Re}(\lambda_i) & -\frac{1}{\rho_i}\mathrm{Im}(\lambda_i) \\ \rho_i \mathrm{Im}(\lambda_i) & \mathrm{Re}(\lambda_i) \end{pmatrix}, \quad 0 < \rho_i \le 1, \tag{5.3}$$

for a complex eigenvalue $\lambda_i \in \mathbb{C} \setminus \mathbb{R}$.

5.1. The two-dimensional case. Later on we use (5.2) to reduce the computation of angular values to those of diagonal blocks. Therefore, we look at 2×2-matrices first and compute $\bar{\theta}_1(A)$ in terms of the spectrum $\sigma(A)$. This is already a nontrivial task. Consider $A \in \mathbb{R}^{2,2}$ with complex conjugate eigenvalues $\lambda, \bar{\lambda}$, $\mathrm{Im}(\lambda) > 0$ and set $\varphi = \arg(\lambda)$ where $\lambda = |\lambda|\exp(i\varphi), 0 < \varphi < \pi$. By orthogonal similarity transformations and a scaling with $|\lambda|^{-1}$ one can put A into the normal form (see (5.3))

$$A(\rho,\varphi) = \begin{pmatrix} \cos(\varphi) & -\rho^{-1}\sin(\varphi) \\ \rho\sin(\varphi) & \cos(\varphi) \end{pmatrix}, \quad 0 < \rho \le 1, \quad 0 < \varphi < \pi. \tag{5.4}$$

According to Proposition 3.8(i) these are the transformations which leave all angular values invariant. Further, the matrix $A(\rho,\varphi)$ leaves the ellipse $x^2 + \rho^{-2}y^2 = 1$ invariant, so that $\rho \le 1$ can be achieved by a permutation.

Finally, we introduce the skewness of a matrix $A \in \mathrm{GL}(\mathbb{R}^d)$ by

$$\mathrm{skew}(A) = \frac{1}{2r(A)}\|A - A^\top\|, \quad r(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$$

and note that this quantity is also invariant under scalings and orthogonal similarity transformations. For the matrix (5.4) we have $\mathrm{skew}(A(\rho,\varphi)) = \frac{1}{2}(\rho + \rho^{-1})|\sin(\varphi)|$.

Proposition 5.2. *For a matrix $A \in \mathrm{GL}(\mathbb{R}^2)$ all first angular values $\theta_1(A)$ with $\theta_1 \in \{\underset{\wedge}{\theta}_{[1]}, \underset{\wedge}{\theta}_1, \hat{\theta}_1, \hat{\theta}_{[1]}, \underline{\theta}_{[1]}, \underline{\theta}_1, \bar{\theta}_1, \bar{\theta}_{[1]}\}$ coincide. Moreover, the following holds:*

(a) If A has only real eigenvalues then

$$\theta_1(A) = \begin{cases} \frac{\pi}{2}, & \text{if } \sigma(A) = \{-\lambda, \lambda\} \subset \mathbb{R}, \lambda > 0, \\ 0, & \text{otherwise.} \end{cases}$$

(b) If $\sigma(A) = \{\lambda, \bar{\lambda}\}, \mathrm{Im}(\lambda) \neq 0$ then

$$\theta_1(A) \le \min(|\arg(\lambda)|, \pi - |\arg(\lambda)|). \tag{5.5}$$

If additionally, $\mathrm{skew}(A) \le 1$ *then we have equality, i.e.*

$$\theta_1(A) = \min(|\arg(\lambda)|, \pi - |\arg(\lambda)|). \tag{5.6}$$

Proof. By Proposition 3.8(i) we can assume A to be in Schur form and scale A such that the largest eigenvalue has (absolute) value 1. Further we mention that $\bar{\theta}_1(A) = 0$ causes all other angular values to vanish by Corollary 5.1 and Lemma 3.3.

For $A = I$ the result is trivial and we are left with the cases

$$A = \begin{cases} \begin{pmatrix} 1 & \eta \\ 0 & \lambda \end{pmatrix}, & \begin{array}{ll} 0 < |\lambda| < 1, \lambda \in \mathbb{R}, \eta \ge 0, & \text{case (i)} \\ \lambda = 1, \eta > 0, & \text{case (ii)} \\ \lambda = -1, \eta \ge 0, & \text{case (iii)} \end{array} \\ A(\rho, \varphi), & \begin{array}{ll} 0 < \rho \le 1, 0 < \varphi < \pi, & \text{case (iv).} \end{array} \end{cases} \tag{5.7}$$

It suffices to consider spaces $V = \mathrm{span}(v_0)$ where $v_0 = \begin{pmatrix} \cos(\theta_0) \\ \sin(\theta_0) \end{pmatrix}$ and $|\theta_0| \le \frac{\pi}{2}$. We write the iterates in polar coordinates

$$v_j = A^j v_0, \quad v_j = r_j \begin{pmatrix} \cos(\theta_j) \\ \sin(\theta_j) \end{pmatrix}, \tag{5.8}$$

where $r_j = \|v_j\|$ and the angles $\theta_j \in \mathbb{R}$ will be determined appropriately. If $|\theta_j - \theta_{j-1}| \le \pi$ one finds that the angle between successive spaces is

$$\measuredangle(\mathrm{span}(v_{j-1}), \mathrm{span}(v_j)) = \chi(\theta_j - \theta_{j-1}), \quad \chi(x) := \min(|x|, \pi - |x|). \tag{5.9}$$

In the following we study the matrices from (5.7) case by case.

(i) Since $|\lambda| < 1$ the Blocking Lemma 5.5 below applies and reduces the formula to the one-dimensional case, i.e. $\bar{\theta}_1(A) = \max(\bar{\theta}_1(1), \bar{\theta}_1(\lambda)) = 0$ and similarly for $\underline{\theta}_1, \hat{\theta}_1, \underset{\wedge}{\theta}_1$. Nevertheless, for later use and for the purpose of illustration we discuss the simple subcase $\eta = 0 < \lambda$ explicitly. In this case we obtain $|\theta_j| \le \frac{\pi}{2}$ for all $j \in \mathbb{N}$ and the following formula

$$\theta_j = \Psi_\lambda(\theta_{j-1}), \quad \Psi_\lambda(\theta) = \begin{cases} \arctan(\lambda \tan(\theta)), & |\theta| < \frac{\pi}{2}, \\ \theta, & |\theta| = \frac{\pi}{2}, \end{cases} \tag{5.10}$$

cf. Figure 5.1. For $\lambda > 0$ we have $\Psi'_\lambda(\theta) > 0$ for all $|\theta| \le \frac{\pi}{2}$, $0 < \Psi_\lambda(\theta) < \theta$ for $\theta \in (0, \frac{\pi}{2})$, and $0 > \Psi_\lambda(\theta) > \theta$ for $\theta \in (-\frac{\pi}{2}, 0)$. The values θ_j are monotone decreasing resp. increasing if $\theta_0 > 0$ resp. $\theta_0 < 0$, and therefore

$$a_{1,n} = \sum_{j=1}^{n} \measuredangle(\mathrm{span}(v_{j-1}), \mathrm{span}(v_j)) = |\sum_{j=1}^{n}(\theta_{j-1} - \theta_j)| = |\theta_0 - \theta_n| \le \frac{\pi}{2}. \tag{5.11}$$

The assertion then follows from Proposition 3.7 (a) with $\varphi = 0$.

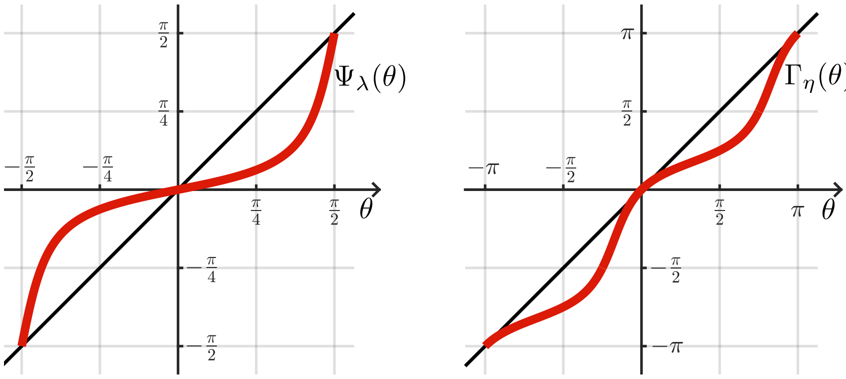

Figure 5.1: Graphs of Ψ_λ for $\lambda = 0.2$ and of Γ_η for $\eta = 1$.

(ii) For the matrix in (5.7) (ii) we obtain

$$\theta_j = \Gamma_\eta(\theta_{j-1}), \quad \Gamma_\eta(\theta) = \begin{cases} \mathrm{arccot}(\eta + \cot(\theta)), & 0 < \theta < \pi, \\ \mathrm{arccot}_{-1}(\eta + \cot(\theta)), & -\pi < \theta < 0, \\ \theta, & \theta = -\pi, 0, \pi, \end{cases} \tag{5.12}$$

where arccot_{-1} is the first negative branch of arccot, see Figure 5.1. The function Γ_η is strictly monotone increasing and satisfies $\Gamma_\eta(\theta) < \theta$ for $0 < |\theta| < \pi$. Therefore, the sequence θ_j is monotone decreasing and converges to 0 if $0 \le \theta_0 < \pi$ and to $-\pi$ if $\theta_0 < 0$. Thus the minimum in (5.9) is achieved at $|\theta_j - \theta_{j-1}|$ and we obtain as in (5.11)

$$a_{1,n} = \sum_{j=1}^{n} \measuredangle(\mathrm{span}(v_{j-1}), \mathrm{span}(v_j)) \le |\sum_{j=1}^{n}(\theta_{j-1} - \theta_j)| = |\theta_0 - \theta_n| \le \pi.$$

(iii) The third case describes a reflection which satisfies $A^2 = I$. Moreover, we find

$$v_0^\top A v_0 = \cos(2\theta_0) - \eta \sin(2\theta_0)$$

which vanishes for $\theta_0 = \frac{\pi}{4}$ if $\eta = 0$, and otherwise for

$$\theta_0 = \frac{1}{2}\arctan(\eta^{-1}) \in \big(0, \frac{\pi}{4}\big).$$

Then we have $\measuredangle(v_0, Av_0) = \frac{\pi}{2} = \measuredangle(A^j v_0, A^{j-1} v_0)$ for all $j \ge 1$. Since $\frac{\pi}{2}$ is the maximum possible angular value our assertion is proved. A reflection turns out to have the same angular value as a rotation by $\frac{\pi}{2}$.

(iv) In (5.7) we can assume $\varphi \le \frac{\pi}{2}$ since $A(\rho,\varphi)$ is orthogonally similar to $-A(\rho,\pi-\varphi)$. For this rotational case we use ergodic theory and employ almost periodicity; see [28, Ch.4.1, Remarks 1.3-1.7]. We extend the function Ψ_ρ defined in (5.10) from $[-\frac{\pi}{2},\frac{\pi}{2}]$ to $\mathbb{R}$ by setting

$$\Psi_\rho(\theta+n\pi)=\Psi_\rho(\theta)+n\pi, \quad |\theta|\le\frac{\pi}{2}, n\in\mathbb{Z}\setminus\{0\}. \tag{5.13}$$

For this extended function there exists a constant $C_\rho>0$ such that

$$|\Psi_\rho(x)-x|, |\Psi_\rho'(x)|, |\Psi_\rho''(x)|\le C_\rho \quad \text{for all} \quad x\in\mathbb{R}. \tag{5.14}$$

The factorization

$$\begin{pmatrix}\cos(\varphi) & -\rho^{-1}\sin(\varphi)\\ \rho\sin(\varphi) & \cos(\varphi)\end{pmatrix}=\begin{pmatrix}1&0\\0&\rho\end{pmatrix}\begin{pmatrix}\cos(\varphi) & -\sin(\varphi)\\ \sin(\varphi) & \cos(\varphi)\end{pmatrix}\begin{pmatrix}1&0\\0&\rho^{-1}\end{pmatrix}$$

shows that the angles $\theta_j\in\mathbb{R}$ in (5.8) are accumulated according to

$$\theta_j=F_{\rho,\varphi}(\theta_{j-1}), j\in\mathbb{N}, \quad F_{\rho,\varphi}(\theta):=\Psi_\rho(\varphi+\Psi_{\rho^{-1}}(\theta)). \tag{5.15}$$

The new variables $\varphi_j=\varphi+\Psi_{\rho^{-1}}(\theta_j)$ then satisfy the recursion

$$\varphi_j=\varphi+\Psi_{\rho^{-1}}(\Psi_\rho(\varphi+\Psi_{\rho^{-1}}(\theta_{j-1})))=\varphi+\varphi_{j-1},$$

hence $\varphi_j=\varphi_0+j\varphi=(j+1)\varphi+\Psi_{\rho^{-1}}(\theta_0)$ and

$$\theta_j=\Psi_\rho(j\varphi+\Psi_{\rho^{-1}}(\theta_0)). \tag{5.16}$$

In particular, the values θ_j are monotone increasing. From (5.16) and (5.14) we infer

$$\begin{aligned}\frac{1}{n}a_{1,n}&=\frac{1}{n}\sum_{j=1}^{n}\chi(\theta_{j-1}-\theta_j)\le\frac{1}{n}\sum_{j=1}^{n}(\theta_j-\theta_{j-1})=\frac{1}{n}(\theta_n-\theta_0)\\ &=\varphi+\frac{1}{n}\big(\Psi_\rho(\varphi_n-\varphi)-(\varphi_n-\varphi)+\Psi_{\rho^{-1}}(\theta_0)-\theta_0\big)\\ &\le\varphi+\frac{C_\rho+C_{\rho^{-1}}}{n}.\end{aligned} \tag{5.17}$$

This will lead to the estimate (5.5) as $n\to\infty$ provided we have shown the equality of all angular values. For this purpose we apply Proposition 3.7(b) where we identify $V\in\mathcal{G}(1,2)$ with $\theta+2\pi\mathbb{Z}\in S^{2\pi}=\mathbb{R}/(2\pi\mathbb{Z})$ via $V=\operatorname{span}\begin{pmatrix}\cos(\theta)\\ \sin(\theta)\end{pmatrix}$. The function Ψ_ρ is a lift of the circle map $\psi_\rho: S^{2\pi}\to S^{2\pi}$ defined by $\psi_\rho(\theta+2\pi\mathbb{Z})=\Psi_\rho(\theta)+2\pi\mathbb{Z}$. Further, the iteration (5.15) may be written by means of a circle map $T_{\rho,\varphi}:S^{2\pi}\to S^{2\pi}$ as follows

$$\theta_j+2\pi\mathbb{Z}=T_{\rho,\varphi}(\theta_{j-1}+2\pi\mathbb{Z}), \quad T_{\rho,\varphi}=\psi_\rho\circ\tau_\varphi\circ\psi_\rho^{-1}, \tag{5.18}$$

where the shift $\tau_\varphi: S^{2\pi}\to S^{2\pi}$ is defined by $\tau_\varphi(\theta+2\pi\mathbb{Z})=\theta+\varphi+2\pi\mathbb{Z}$. The map $F_{\rho,\varphi}$ in (5.15) is then a lift of $T_{\rho,\varphi}$. It is well known (see [5, Ch.2.6.2]) that τ_φ is an ergodic isometry of $S^{2\pi}$ with respect to Lebesgue measure μ_1 and the standard metric

$$d_1(\theta_1+2\pi\mathbb{Z},\theta_2+2\pi\mathbb{Z})=\min_{z\in\mathbb{Z}}|\theta_1-\theta_2+2\pi z|$$

if and only if $\frac{\varphi}{\pi} \notin \mathbb{Q}$. In this case the conjugacy (5.18) implies that $T_{\rho,\varphi}$ is an ergodic isometry of $S^{2\pi}$ with respect to the image measure $\mu_\rho = \mu_1 \circ \psi_\rho^{-1}$ and the image metric $d_\rho(\cdot,\cdot) = d_1(\psi_\rho^{-1}\cdot, \psi_\rho^{-1}\cdot)$. We conclude from [28, Remark 1.3] that the map $T_{\rho,\varphi}$ is uniformly almost periodic, i.e. for every $\varepsilon > 0$ there exists a relatively dense set $\mathcal{P} \subseteq \mathbb{N}_0$ such that $d_\rho(x, T^p_{\rho,\varphi}x) \le \varepsilon$ for all $x \in S^{2\pi}$, $p \in \mathcal{P}$. Moreover, for any continuous function $g : S^{2\pi} \to \mathbb{R}$ the sequence of functions $b_n(x) = g(T^{n-1}_{\rho,\varphi}x)$, $x \in S^{2\pi}$, $n \in \mathbb{N}$ is uniformly almost periodic in the sense of Definition 3.4. To see this, let $\varepsilon_0 > 0$ be given and take $\varepsilon > 0$ such that $|g(x_1) - g(x_2)| \le \varepsilon_0$ whenever $d_\rho(x_1, x_2) \le \varepsilon$, $x_1, x_2 \in S^{2\pi}$. For the relatively dense set $\mathcal{P} \subset \mathbb{N}$ belonging to ε we then find

$$|b_n(x) - b_{n+p}(x)| = |g(T^{n-1}_{\rho,\varphi}x) - g(T^p_{\rho,\varphi}(T^{n-1}_{\rho,\varphi}x))| \le \varepsilon_0 \quad \forall n \in \mathbb{N}, p \in \mathcal{P}, x \in S^{2\pi}.$$

In the case $\frac{\varphi}{\pi} \in \mathbb{Q}$ we have the same result since then every point $x \in S^{2\pi}$ has the same period q where $\frac{\varphi}{\pi} = \frac{2p}{q}$.
Let us apply this to the continuous function

$$g(x) = \min(d_1(x, T_{\rho,\varphi}x), d_1(\tau_\pi x, T_{\rho,\varphi}x)), \quad x \in S^{2\pi}. \tag{5.19}$$

Setting $x = \theta_0 + 2\pi\mathbb{Z}$ we obtain

$$T^{j-1}_{\rho,\varphi}x = \theta_{j-1} + 2\pi\mathbb{Z}, \quad j \in \mathbb{N}.$$

Using $\theta_{j-1} < \theta_j \le \theta_{j-1} + \pi$ and (5.9) for $j \in \mathbb{N}$ then leads to

$$\begin{aligned} b_j(x) = g(T^{j-1}_{\rho,\varphi}x) &= \min(\theta_j - \theta_{j-1}, \theta_{j-1} + \pi - \theta_j) \\ &= \chi(\theta_j - \theta_{j-1}) = \measuredangle(\operatorname{span}(v_{j-1}), \operatorname{span}(v_j)). \end{aligned} \tag{5.20}$$

Therefore, all angular values agree by Proposition 3.7 (b).
Next we show that the assumption $\operatorname{skew}(A) = \frac{1}{2}(\rho+\rho^{-1})|\sin(\varphi)| \le 1$ implies $\theta_j - \theta_{j-1} \le \frac{\pi}{2}$. Then the minimum in (5.9) is always achieved with the first term and the first inequality in (5.17) becomes an equality. Thus we find

$$|\frac{1}{n}a_{1,n} - \varphi| \le \frac{C_\rho + C_{\rho^{-1}}}{n},$$

and Proposition 3.7(a) implies the assertion. It remains to analyze the inequality

$$F_{\rho,\varphi}(\theta) = \Psi_\rho(\varphi + \Psi_{\rho^{-1}}(\theta)) \le \theta + \frac{\pi}{2}, \quad \theta \in \mathbb{R}. \tag{5.21}$$

For later purposes we perform a rather explicit calculation. First note that it is enough to consider $0 < |\theta| < \frac{\pi}{2}$ since $F_{\rho,\varphi}(\theta + n\pi) = F_{\rho,\varphi}(\theta) + n\pi$ holds by (5.13) and since (5.21) is obvious for $\theta = 0, \pm\frac{\pi}{2}$. By the monotonicity of $\Psi_{\rho^{-1}}$ and the sum formula[1] for arctan we obtain that $F_{\rho,\varphi}(\theta) \le \theta + \frac{\pi}{2}$ holds for $0 < |\theta| < \frac{\pi}{2}$ if and only if

$$\varphi \le \Psi_{\rho^{-1}}(\theta + \frac{\pi}{2}) - \Psi_{\rho^{-1}}(\theta) = \begin{cases} r(\theta,\rho), & 0 < \theta < \frac{\pi}{2}, \\ \pi + r(\theta,\rho), & -\frac{\pi}{2} < \theta < 0, \end{cases}$$

$$r(\theta,\rho) := \arctan\Big(\frac{\tan(\theta) + \frac{1}{\tan(\theta)}}{\rho^{-1} - \rho}\Big) = \arctan\Big(\frac{2}{(\rho^{-1} - \rho)\sin(2\theta)}\Big).$$

[1] $\arctan(x) + \arctan(y) = \operatorname{sgn}(x)\pi - \arctan(\frac{x+y}{xy-1})$ for $x \neq 0, xy > 1$.

In the case $\theta < 0$ this inequality always holds since $\varphi \le \frac{\pi}{2}$, while for $\theta > 0$ it is equivalent to

$$\sin(2\theta) \le \frac{2}{\tan(\varphi)(\rho^{-1} - \rho)} =: \beta(\rho, \varphi). \tag{5.22}$$

Expressing $\tan(\varphi)$ in terms of $\sin(\varphi) = \frac{2\,\mathrm{skew}(A)}{\rho^{-1}+\rho}$ leads to

$$\beta(\rho, \varphi) = \Big(1 + \frac{4(1 - \mathrm{skew}(A)^2)}{(\rho^{-1} - \rho)^2}\Big)^{1/2}. \tag{5.23}$$

Hence condition (5.22) holds for all $\theta \in \mathbb{R}$ if $\mathrm{skew}(A) \le 1$. ■

Remark 5.3. *Let us relate the result of Proposition 5.2 to the theory of rotation numbers; see [24, Ch.11]. First, note that this theory uses $[0, 1)$ instead of $[0, 2\pi)$ as the interval of periodicity. Every matrix $A \in \mathrm{GL}(\mathbb{R}^2)$ induces a homeomorphism $f : S^1 \to S^1$ of $S^1 = \mathbb{R}/\mathbb{Z}$ via the relation (one step of the iteration (5.8))*

$$v = \|v\| \begin{pmatrix} \cos(2\pi x) \\ \sin(2\pi x) \end{pmatrix} \mapsto Av = \|Av\| \begin{pmatrix} \cos(2\pi f(x)) \\ \sin(2\pi f(x)) \end{pmatrix}, \quad x \in S^1. \tag{5.24}$$

The homeomorphism is orientation-preserving if and only if $\det(A) > 0$. For such a homeomorphism the rotation number $\tau(f) \in [0, 1)$ is well defined. Iterating (5.24) and comparing with (5.8) then shows the equality $2\pi\tau(f) = \hat{\theta}_1(A)$, provided no vector rotates by more than $\frac{\pi}{2}$. For the matrices in (5.7) these conditions hold in case (i) if $\lambda > 0$, in case (ii), and in case (iv) if $\mathrm{skew}(A) \le 1$ (see (5.15)). The corresponding f-maps are $2\pi f(x) = \Psi_\lambda(2\pi x)$ (see case (i), $\lambda > 0$, $\eta = 0$, equation (5.10)), $2\pi f(x) = \Gamma_\eta(2\pi x)$ (case (ii), equation (5.12)), and $2\pi f(x) = \Psi_\rho(\varphi + \Psi_{\rho^{-1}}(2\pi x))$ (case (iv)). Determining the exact first angular value $\theta_1(A)$ in case $\mathrm{skew}(A) > 1$ of (a) is more involved. In Theorem 6.1 we will show that the inequality (5.5) is generally strict except for some resonant values of $\varphi = \arg(\lambda)$.

5.2. Systems of higher dimension. As a first step we consider a matrix with a single eigenvalue which generalizes the second case in (5.7). Its proof is stated in the Supplementary materials III.

Proposition 5.4. *Assume that the spectrum of $A \in \mathbb{R}^{d,d}$ consists of one eigenvalue $\lambda \in \mathbb{R}$, $\lambda \neq 0$. Then all first angular values vanish, i.e.*

$$\theta_1(A) = 0 \quad \textit{for} \quad \theta_1 \in \{\underset{\wedge}{\theta}_{[1]}, \underset{\wedge}{\theta}_1, \hat{\theta}_1, \hat{\theta}_{[1]}, \underline{\theta}_{[1]}, \underline{\theta}_1, \bar{\theta}_1, \bar{\theta}_{[1]}\}.$$

To proceed further, we require the following lemma.

Lemma 5.5. *(Blocking Lemma) Let $\mathbb{R}^d = X_s \oplus X_u$ be a decomposition into invariant subspaces of $A \in \mathbb{R}^{d,d}$ such that $A_s = A_{|X_s}$ and $A_u = A_{|X_u}$ satisfy*

$$|\sigma(A_s)| < |\sigma(A_u)|. \tag{5.25}$$

Then the following holds for all types of angular values $\theta_1 \in \{\bar{\theta}_1, \underline{\theta}_1, \hat{\theta}_1, \underset{\wedge}{\theta}_1\}$

$$\theta_1(A) = \max(\theta_1(A_s), \theta_1(A_u)). \tag{5.26}$$

We refer to the Supplementary materials IV for a proof of the Blocking Lemma.

Remark 5.6. *By Corollary 5.1 it is clear that formula (5.26) also holds for the uniform first angular value $\bar{\theta}_{[1]}(A)$. We did not succeed in proving this for the remaining three uniform first angular values. However, we will be able to treat these three values in the subsequent main Theorem 5.7 under a special assumption.*

The following Theorem combines the results of Propositions 5.2, 5.4 and Lemma 5.5.

Theorem 5.7. *Let the spectrum of $A \in \mathrm{GL}(\mathbb{R}^d)$ satisfy*

$$\lambda \in \sigma(A), \lambda \notin \mathbb{R} \Longrightarrow \lambda \text{ is simple and } |\eta| \neq |\lambda| \quad \forall \eta \in \sigma(A) \setminus \{\lambda, \bar{\lambda}\}. \tag{5.27}$$

Then all 8 types of angular values $\theta_1(A)$ with $\theta_1 \in \{\underset{\wedge}{\theta}_{[1]}, \underset{\wedge}{\theta}_1, \hat{\theta}_1, \hat{\theta}_{[1]}, \underline{\theta}_{[1]}, \underline{\theta}_1, \bar{\theta}_1, \bar{\theta}_{[1]}\}$ coincide. Let $\mathbb{R}^d = \bigoplus_{i=1}^k \mathcal{R}(Q_i)$, $Q_i \in \mathbb{R}^{d,d_i}$, $Q_i^\top Q_i = I_{d_i}$ be a decomposition of $\mathbb{R}^d$ into invariant subspaces of A corresponding to eigenvalues of equal modulus, i.e.

$$AQ_i = Q_i A_i, \quad A_i \in \mathbb{R}^{d_i,d_i}, \quad |\sigma(A_1)|, \dots, |\sigma(A_k)| \text{ pairwise different.} \tag{5.28}$$

Then the following equality holds

$$\theta_1(A) = \max_{i=1,\dots,k} \theta_1(A_i). \tag{5.29}$$

If there exist two real eigenvalues of opposite sign in $\sigma(A)$ then $\theta_1(A) = \frac{\pi}{2}$. Otherwise, the following estimate holds

$$\theta_1(A) \le \max_{\lambda \in \sigma(A)} \min(|\arg(\lambda)|, \pi - |\arg(\lambda)|). \tag{5.30}$$

Equality holds if the maximum on the right-hand side of (5.30) is zero or if it is achieved for an eigenvalue $\lambda_{i_0} \in \sigma(A_{i_0})$, $i_0 \in \{1, \dots, k\}$ with $\mathrm{Im}(\lambda_{i_0}) \neq 0$ and $\mathrm{skew}(A_{i_0}) \le 1$; i.e.

$$\theta_1(A) = \min(|\arg(\lambda_{i_0})|, \pi - |\arg(\lambda_{i_0})|). \tag{5.31}$$

Remark 5.8. *For the formulas (5.29) and (5.31) it is essential to choose orthonormal bases for the invariant subspaces. Other bases will preserve the spectra of the matrices A_i but neither the values $\mathrm{skew}(A_i)$ nor the angular values $\theta_1(A_i)$, see Proposition 5.2 (b). Except for the first block Λ_1, the angular values of Λ_i in the Schur form (5.2) generally do not agree with $\theta_1(A_i)$, see Algorithm 6.2 and the example in Section 6.3.2.*

Proof. Let us first prove (5.29) for all 4 nonuniform types of angular value. Note that a decomposition $\mathbb{R}^d = \bigoplus_{i=1}^k \mathcal{R}(Q_i)$ of the desired type always exists since we can decompose $\sigma(A)$ into subsets of equal modulus and then select an orthogonal basis for each of the corresponding invariant subspaces. In this way we transform A into block-diagonal form in a specific way (see (5.28)),

$$A \begin{pmatrix} Q_1 & \cdots & Q_k \end{pmatrix} = \begin{pmatrix} Q_1 & \cdots & Q_k \end{pmatrix} \mathrm{diag}(A_1, \dots, A_k). \tag{5.32}$$

If one does not insist on orthonormal bases for the subspaces then one can keep the diagonal blocks Λ_i from the Schur form; see [17, Thm 7.1.6]. From $Q_i^\top Q_i = I_{d_i}$ we obtain for every $v_i \in \mathbb{R}^{d_i}, v_i \neq 0$, $i = 1, \dots, k$, $j \in \mathbb{N}$,

$$\measuredangle(A^{j-1} Q_i v_i, A^j Q_i v_i) = \measuredangle(Q_i A_i^{j-1} v_i, Q_i A_i^j v_i) = \measuredangle(A_i^{j-1} v_i, A_i^j v_i),$$

so that all first angular values of A_i and of the restriction $A_{|\mathcal{R}(Q_i)}$ coincide. Hence Lemma 5.5 shows (5.29). Note that (5.29) also holds for the $\bar{\theta}_{[1]}$-values by Corollary 5.1.

Now assume that there exist two real eigenvalues $\lambda, -\lambda \in \sigma(A)$ and w.l.o.g. assume $\lambda, -\lambda \in \sigma(A_1)$. Then there exists an orthogonal $S \in \mathbb{R}^{d_1,d_1}$ and some $\eta \in \mathbb{R}$ such that

$$A_1 S = SM, \quad M = \begin{pmatrix} M_{11} & M_{12} \\ 0 & M_{22} \end{pmatrix}, M_{11} = \begin{pmatrix} \lambda & \eta \\ 0 & -\lambda \end{pmatrix}.$$

The first two columns of S form an orthonormal basis of the span of eigenvectors which belong to λ and $-\lambda$. Choosing initial vectors $v_1 = (v_0, 0, \dots, 0)^\top \in \mathbb{R}^{d_1}, v_0 \in \mathbb{R}^2$ we find for all $j \in \mathbb{N}$

$$\measuredangle(A_1^{j-1} S v_1, A_1^j S v_1) = \measuredangle(S M^{j-1} v_1, S M^j v_1) = \measuredangle(M^{j-1} v_1, M^j v_1) = \measuredangle(M_{11}^{j-1} v_0, M_{11}^j v_0).$$

The proof of Proposition 5.2(a) (see case (iii) in (5.7)) shows that there exists $v_0 \in \mathbb{R}^2$, $v_0 \neq 0$ such that for all $j \in \mathbb{N}$

$$\frac{\pi}{2} = \measuredangle(M_{11}^{j-1} v_0, M_{11}^j v_0) = \measuredangle(A_1^{j-1} S v_1, A_1^j S v_1) = \measuredangle(A^{j-1} Q_1 S v_1, A^j Q_1 S v_1).$$

Since $\frac{\pi}{2}$ is the maximum of all angular values, Definition 3.1 implies that all 8 types of angular values are equal to $\frac{\pi}{2}$.

If such a pair of real eigenvalues does not exist then assumption (5.27) shows that the matrices A_i either are two-dimensional as in Proposition 5.2 (see case (iv) in (5.7)) or have a single real eigenvalue as in Proposition 5.4. In both cases the propositions guarantee all first angular values of the matrices A_i to coincide. Thus the four nonuniform angular values of the given matrix A are equal by (5.29). For $\bar{\theta}_{[1]}(A)$ the result then follows from Corollary 5.1. Moreover, Lemma 3.3 yields formula (5.29) and the coincidence of the $\hat{\theta}_{[1]}$-values:

$$\bar{\theta}_1(B) = \hat{\theta}_1(B) \leq \hat{\theta}_{[1]}(B) \leq \bar{\theta}_{[1]}(B) = \bar{\theta}_1(B), \quad B \in \{A, A_i (i = 1, \dots, k)\}.$$

Next we show $\underset{\wedge}{\theta}_{[1]}(A) = \underset{\wedge}{\theta}_1(A)$. From (5.29) we find an index $\ell \in \{1, \dots, k\}$ for which $\underset{\wedge}{\theta}_1(A) = \underset{\wedge}{\theta}_1(A_\ell)$ holds. Then we use Lemma 3.3 and the equality of angular values from Propositions 5.2 and 5.4,

$$\begin{aligned}
\underset{\wedge}{\theta}_1(A) = \underset{\wedge}{\theta}_{[1]}(A_\ell) &= \sup_{V_\ell \in \mathcal{G}(1,d_\ell)} \liminf_{n\to\infty} \frac{1}{n} \inf_{k \in \mathbb{N}_0} \sum_{j=k+1}^{k+n} \measuredangle(Q_\ell A_\ell^{j-1} V_\ell, Q_\ell A_\ell^j V_\ell) \\
&= \sup_{V_\ell \in \mathcal{G}(1,d_\ell)} \liminf_{n\to\infty} \frac{1}{n} \inf_{k \in \mathbb{N}_0} \sum_{j=k+1}^{k+n} \measuredangle(A^{j-1} Q_\ell V_\ell, A^j Q_\ell V_\ell) \\
&\leq \sup_{V \in \mathcal{G}(1,d)} \liminf_{n\to\infty} \frac{1}{n} \inf_{k \in \mathbb{N}_0} \sum_{j=k+1}^{k+n} \measuredangle(A^{j-1} V, A^j V) = \underset{\wedge}{\theta}_{[1]}(A) \leq \underset{\wedge}{\theta}_1(A).
\end{aligned}$$

Using Lemma 3.3 we obtain the result for the last angular value $\underline{\theta}_{[1]}(A)$:

$$\underset{\wedge}{\theta}_1(B) = \underset{\wedge}{\theta}_{[1]}(B) \leq \underline{\theta}_{[1]}(B) \leq \underline{\theta}_1(B) = \underset{\wedge}{\theta}_1(B), \quad B \in \{A, A_i (i = 1, \dots, k)\}.$$

Finally, the estimate (5.30) follows from (5.5) and Proposition 5.4. If the maximum value on the right of (5.30) is zero then the assertion (5.31) is obvious. Otherwise, it follows from (5.29) and (5.6) in Proposition 5.2 when applied to the 2×2 matrix A_{i_0}. ∎

Note that condition (5.27) excludes a complex eigenvalue of multiplicity ≥ 2 and another eigenvalue of the same modulus. Let us consider such an exceptional case, namely a block diagonal matrix with two rotations

$$A = \begin{pmatrix} T_{\varphi_1} & 0 \\ 0 & T_{\varphi_2} \end{pmatrix}, \quad 0 \leq \varphi_1, \varphi_2 \leq \frac{\pi}{2}. \tag{5.33}$$

We claim that every type of angular value is given by

$$\theta_1(A) = \max(\varphi_1, \varphi_2).$$

Let $v = \begin{pmatrix} v_1 & v_2 \end{pmatrix}^\top$, $v_1, v_2 \in \mathbb{R}^2$ and $|v_1|^2 + |v_2|^2 = 1$. Since A is orthogonal we obtain

$$\begin{aligned} \cos(\measuredangle(Av, v)) &= |\langle Av, v\rangle| = |\langle T_{\varphi_1} v_1, v_1\rangle + \langle T_{\varphi_2} v_2, v_2\rangle| \\ &= \cos(\varphi_1)|v_1|^2 + \cos(\varphi_2)|v_2|^2 = |v_1|^2(\cos(\varphi_1) - \cos(\varphi_2)) + \cos(\varphi_2). \end{aligned}$$

By the orthogonal invariance of the angle (see Proposition 3.8) this leads to

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \measuredangle(A^{j-1}v, A^j v) = \arccos\left(|v_1|^2(\cos(\varphi_1) - \cos(\varphi_2)) + \cos(\varphi_2)\right). \tag{5.34}$$

Suppose w.l.o.g. that $\varphi_2 \geq \varphi_1$ so that $\cos(\varphi_1) - \cos(\varphi_2) \geq 0$. Then the maximum w.r.t. v in (5.34) occurs for $|v_1| = 0$, hence $\theta_1(A) = \varphi_2$. The same argument applies to a block diagonal matrix with k blocks $T_{\varphi_i}, i = 1, \ldots, k$ on the diagonal, leading to $\theta_1(A) = \max_{i=1,\ldots,k} \varphi_i$. However, we did not find a formula for $\theta_1(A)$ in cases which violate (5.27) but which are more general than (5.33).

6. Numerical algorithms and results. In this section, our main goal is to discuss algorithms for the computation of the first outer angular value $\hat{\theta}_1(A)$ of an autonomous system generated by a matrix $A \in \mathbb{R}^{d,d}$. First we investigate the two-dimensional case where our focus is on matrices with $\mathrm{skew}(A) > 1$. We extend the theory underlying Proposition 5.2 and compare with numerical computations.

Then we use the results from Lemma 5.5 and Theorem 5.7 to develop an algorithm for matrices of arbitrary dimension. Let us emphasize that the whole calculation aims at first outer angular values. In the autonomous case we know the coincidence with inner angular values by Theorem 5.7. However, for general nonautonomous systems the computation of inner angular values turns out to be quite involved since one has to solve an optimization problem in every time step.

Let us also note that simple algorithms based on subspace iterations tend to fail. The forward iteration of a generic one-dimensional subspace converges to the most unstable direction. However, we must consider also non-generic directions, i.e. all invariant subspaces, in order to compute $\hat{\theta}_1(A)$.

6.1. Two dimensional autonomous examples. Consider the normal form (5.4) with increasing skewness and recall from Proposition 5.2 that all angular values coincide,

$$A(\rho, \varphi) = \begin{pmatrix} \cos(\varphi) & -\rho^{-1}\sin(\varphi) \\ \rho\sin(\varphi) & \cos(\varphi) \end{pmatrix}, \quad 0 < \rho \leq 1, \quad 0 < \varphi \leq \frac{\pi}{2}. \tag{6.1}$$

In the following Table 6.1 we compare for three cases the value of φ from its normal form with the numerical value $\hat{\theta}_{1,\mathrm{num}}$ obtained by solving the optimization problem

$$\hat{\theta}_{1,\mathrm{num}} = \max_{v\in\mathbb{R}^2,\|v\|=1} \frac{1}{N}\sum_{j=1}^{N} \measuredangle(A^{j-1}v, A^j v), \quad N = 1000 \tag{6.2}$$

with the MATLAB-routine `fminbnd`. Note that in this case computations using forward iteration are not spoilt by a dominating direction since A has two eigenvalues of equal modulus.

A	skew(A)	eigenvalues	φ	$\hat{\theta}_{1,\mathrm{num}}$	$\varphi - \hat{\theta}_{1,\mathrm{num}}$
$\begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix}$	$\frac{1}{\sqrt{7}} < 1$	$\frac{5}{2} \pm \frac{\sqrt{3}}{2} i$	$\arctan(\frac{\sqrt{3}}{5})$	0.33347	$6 \cdot 10^{-17}$
$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$	$\frac{1}{\sqrt{2}} < 1$	$1 \pm i$	$\frac{\pi}{4}$	0.78540	$1 \cdot 10^{-16}$
$\begin{pmatrix} 2 & 1 \\ -49 & 3 \end{pmatrix}$	$\frac{5\sqrt{5}}{\sqrt{11}} > 1$	$\frac{5}{2} \pm \frac{\sqrt{195}}{2} i$	$\arctan(\sqrt{\frac{39}{5}})$	0.52709	0.6999

Table 6.1: First angular values for autonomous examples with increasing skewness.

The first and second example in Table 6.1 have skewness ≤ 1. Then the numerical angular value $\hat{\theta}_{1,\mathrm{num}}$ agrees with $\min(|\arg(\lambda)|, \pi - |\arg(\lambda)|)$ to machine accuracy, as predicted by Proposition 5.2. However, the third example belongs to the values $\varphi = \arctan(\sqrt{\frac{39}{5}})$, $\rho = \frac{10\sqrt{5}-\sqrt{461}}{\sqrt{39}} \approx \frac{1}{7}$ and skew$(A) \approx 3.37$, so that Proposition 5.2 provides no explicit expression for the first angular value. The solution of (6.2) yields a substantially smaller value $\hat{\theta}_{1,\mathrm{num}} < \varphi$ in this case. Indeed, the first angular value exhibits a rather subtle dependence on the matrix entries for skew$(A) > 1$. Figure 6.1 (left panel) shows the result of an extensive computation of the angular value for the matrix (6.1) with $\rho = \frac{1}{7}$ and for 25210 equidistant points $\varphi \in [0, \frac{\pi}{2}]$. The vertical red line on the left marks the critical value $\varphi_c = \arcsin(\frac{2}{\rho+\rho^{-1}})$ below which we have skew$(A(\rho,\varphi)) \leq 1$ and Proposition 5.2 guarantees φ as the first angular value. The value φ_c seems to be sharp, and for values $\varphi > \varphi_c$ we observe resonances occurring at rational multiples of π.

The following theorem gives an explicit formula for irrational multiples of π and reduces the computation of the angular value to a finite optimization problem for rational multiples of π. For comparison we show in Figure 6.1(right panel) the diagram of angular values when evaluated directly from the result of Theorem 6.1. Continuing this evaluation for several values of ρ yields the three-dimensional diagram in Figure 6.2.

Theorem 6.1. *For $0 < \rho \leq 1$ and $0 < \varphi \leq \frac{\pi}{2}$ the first angular value $\hat{\theta}_1(A(\rho,\varphi))$ of the matrix from (6.1) is given by*

$$\hat{\theta}_1(A(\rho,\varphi)) = \begin{cases} \varphi, & \mathrm{skew}(A(\rho,\varphi)) \leq 1, & \text{case(i)} \\ \varphi + \frac{1}{\pi}\int_{\{\delta<0\}} \delta(\theta)\mathrm{d}\theta, & \mathrm{skew}(A(\rho,\varphi)) > 1, \frac{\varphi}{\pi} \notin \mathbb{Q}, & \text{case(ii)} \\ \varphi, & \mathrm{skew}(A(\rho,\varphi)) > 1, \frac{\varphi}{\pi} = \frac{1}{q}, q \geq 2, & \text{case(iii)} \\ \frac{1}{q}\max_{0\leq\theta\leq\frac{\pi}{2}} \sum_{j=1}^{q} g_j(\theta), & \mathrm{skew}(A(\rho,\varphi)) > 1, \frac{\varphi}{\pi} = \frac{p}{q}, q \notin p\mathbb{N}, & \text{case(iv).} \end{cases} \tag{6.3}$$

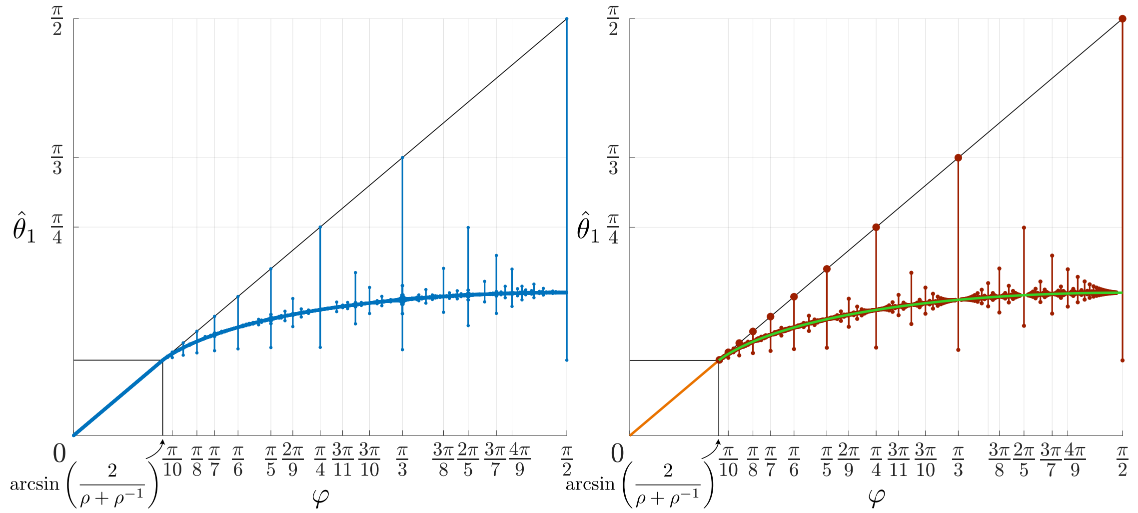

Figure 6.1: Angular value $\hat{\theta}_1$ for (6.1) with $\rho = \frac{1}{7}$. Left panel: For 25210 equidistant points $\varphi \in [0, \frac{\pi}{2}]$ the minimal and maximal first angular value are computed by solving an optimization problem; minima and maxima are connected with lines. Right panel: Computation of first angular value via formula (6.3) in Theorem 6.1. Results for case 1 (orange), case 2 (green), case 3 (big points on the diagonal), case 4 (small points above the green curve). In case 4, minima are also shown (small points below the green curve) and connected with corresponding maxima.

Here the functions $g_j, \delta : [0, \pi] \to \mathbb{R}$, $j \in \mathbb{N}$ are defined as follows:

$$\delta(\theta) = 2\Psi_\rho(\theta) - 2\Psi_\rho(\theta + \varphi) + \pi, \quad \textit{with } \Psi_\rho \textit{ from } (5.10),$$
$$g_j(\theta) = \min(\theta_j - \theta_{j-1}, \theta_{j-1} + \pi - \theta_j), \quad \theta_{j-1} = \Psi_\rho((j-1)\varphi + \Psi_{\rho^{-1}}(\theta)).$$

If $\frac{\pi}{2} < \varphi < \pi$ then $\hat{\theta}_1(A(\rho, \varphi)) = \hat{\theta}_1(A(\rho, \pi - \varphi))$.

Proof. The proof is done sequentially for cases (i), (ii), (iv), and (iii).

(i) This case follows from (5.6) in Proposition 5.2 since $A(\rho, \varphi)$ has eigenvalues $e^{\pm i\varphi}$.

(ii) For nonresonant values $\frac{\varphi}{\pi} \notin \mathbb{Q}$ we return to the proof of (5.7) case (iv) in Proposition 5.2. Let us apply Birkhoff's ergodic theorem to the ergodic isometry $T_{\rho,\varphi}$ of $(S^{2\pi}, d_\rho, \mu_\rho)$ (see (5.18)) and to the continuous map g from (5.19),

(6.4)
$$\hat{\theta}_1(A(\rho, \varphi)) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} g(T_{\rho,\varphi}^{j-1}\xi) = \frac{1}{2\pi} \int_{S^{2\pi}} g(y)\mathrm{d}\mu_\rho(y) = \frac{1}{2\pi} \int_{S^{2\pi}} g(\psi_\rho(x))\mathrm{d}\mu_1(x).$$

The last equality is due to the transformation formula. Also note that the convergence is uniform in $\xi \in S^{2\pi}$, see [28, Ch.4.1, Remark 1.5]. We evaluate the integrand for $x = \theta + 2\pi\mathbb{Z}$, $\theta \in [0, 2\pi)$,

$$\begin{aligned} g(\psi_\rho(x)) &= \min(d_1(\psi_\rho(x), T_{\rho,\varphi} \circ \psi_\rho(x)), d_1(\tau_\pi \circ \psi_\rho(x), T_{\rho,\varphi} \circ \psi_\rho(x))) \\ &= \min(d_1(\psi_\rho(x), \psi_\rho \circ \tau_\varphi(x)), d_1(\psi_\rho \circ \tau_\pi(x), \psi_\rho \circ \tau_\varphi(x))) \\ &= \min(\Psi_\rho(\theta + \varphi) - \Psi_\rho(\theta), \Psi_\rho(\theta + \pi) - \Psi_\rho(\theta + \varphi)), \end{aligned}$$

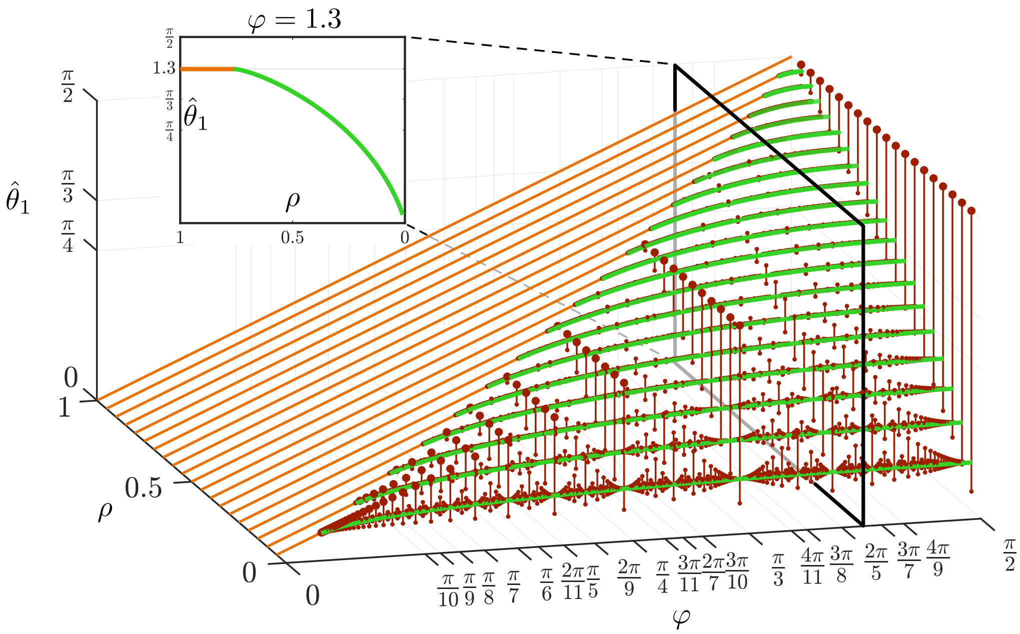

Figure 6.2: Angular value $\hat\theta_1$ for (6.1) with $\varphi \in [0, \frac{\pi}{2}]$ and $\rho = 0.05, 0.1, \ldots, 1$. Computation via formula (6.3) in Theorem 6.1.

where the last equality follows from $\Psi_\rho(\theta) \le \Psi_\rho(\theta+\varphi) < \Psi_\rho(\theta+\pi) = \Psi_\rho(\theta)+\pi$. Combining this with (6.4) and using (5.13) leads to

$$\begin{aligned}\hat\theta_1(A(\rho,\varphi)) &= \frac{1}{2\pi}\int_0^{2\pi} \min(\Psi_\rho(\theta+\varphi)-\Psi_\rho(\theta), \Psi_\rho(\theta+\pi)-\Psi_\rho(\theta+\varphi))\mathrm{d}\theta \\ &= \frac{1}{\pi}\int_0^{\pi} \min(\Psi_\rho(\theta+\varphi)-\Psi_\rho(\theta), \Psi_\rho(\theta+\pi)-\Psi_\rho(\theta+\varphi))\mathrm{d}\theta.\end{aligned}$$

We investigate the minimum by looking at the sign of the difference

$$\Psi_\rho(\theta+\pi)-\Psi_\rho(\theta+\varphi)-(\Psi_\rho(\theta+\varphi)-\Psi_\rho(\theta)) = \delta(\theta).$$

For $\mathrm{skew}(A(\rho,\varphi)) > 1$ the equivalence of (5.21) and (5.22) yields

$$\delta(\theta)\begin{cases} \ge 0, & \theta \in [0,\theta_-] \cup [\theta_+,\pi], \\ < 0, & \theta \in (\theta_-,\theta_+), \end{cases} \tag{6.5}$$

where the values $\theta_\pm$ are given as follows

$$\begin{aligned}\theta_\pm &= \Psi_{\rho^{-1}}(\theta'_\pm), \\ \theta'_- &= \frac{1}{2}\arcsin\left(\frac{2}{\tan(\varphi)(\rho^{-1}-\rho)}\right) \in (0,\frac{\pi}{4}),\ \theta'_+ = \frac{\pi}{2}-\theta'_- \in (\frac{\pi}{4},\frac{\pi}{2}).\end{aligned}$$

Using (5.13) the following computation completes the proof of assertion (ii)

$$\begin{aligned}
\hat{\theta}_1(A(\rho,\varphi)) &= \frac{1}{\pi}\Big\{\Big(\int_0^{\theta_-} + \int_{\theta_+}^{\pi}\Big)\Psi_\rho(\theta+\varphi) - \Psi_\rho(\theta)\mathrm{d}\theta + \int_{\theta_-}^{\theta_+} \Psi_\rho(\theta+\pi) - \Psi_\rho(\theta+\varphi)\mathrm{d}\theta\Big\} \\
&= \frac{1}{\pi}\Big\{\int_0^{\pi} \Psi_\rho(\theta+\varphi) - \Psi_\rho(\theta)\mathrm{d}\theta + \int_{\theta_-}^{\theta_+} \delta(\theta)\mathrm{d}\theta\Big\} \\
&= \frac{1}{\pi}\left\{\int_{\varphi}^{\pi+\varphi} - \int_0^{\pi}\right\} \Psi_\rho(\theta)\mathrm{d}\theta + \frac{1}{\pi}\int_{\theta_-}^{\theta_+} \delta(\theta)\mathrm{d}\theta \\
&= \frac{1}{\pi}\int_0^{\varphi} \Psi_\rho(\theta) + \pi - \Psi_\rho(\theta)\mathrm{d}\theta + \frac{1}{\pi}\int_{\theta_-}^{\theta_+} \delta(\theta)\mathrm{d}\theta \\
&= \varphi + \frac{1}{\pi}\int_{\{\delta<0\}} \delta(\theta)\mathrm{d}\theta.
\end{aligned}$$

Recall that the values $\theta_\pm = \theta_\pm(\rho,\varphi)$, depend on ρ, φ and satisfy

$$0 < \theta_-(\rho,\varphi) < \theta_+(\rho,\varphi) < \tfrac{\pi}{2}, \quad \text{if } \operatorname{skew}(A(\rho,\varphi)) > 1 \tag{6.6}$$

by the strict monotonicity of $\Psi_{\rho^{-1}}$. Therefore, (6.5) implies $\hat{\theta}_1(A(\rho,\varphi)) < \varphi$.

(iv) Next we consider $\frac{\varphi}{\pi} = \frac{p}{q}$ for some natural numbers $0 < p \le q$. From the definition (5.18) of $T_{\rho,\varphi}$ we obtain

$$T_{\rho,\varphi}^q = \psi_\rho \circ \tau_\varphi^q \circ \psi_\rho^{-1} = \psi_\rho \circ \tau_{p\pi} \circ \psi_\rho^{-1} = \tau_{p\pi},$$

where we used $\psi_\rho \circ \tau_{p\pi} = \tau_{p\pi} \circ \psi_\rho$ due to (5.13). Moreover, translation invariance of the metric d_1 yields that the function g in (5.19) is π-periodic:

$$\begin{aligned}
g(\tau_\pi x) &= \min(d_1(\tau_\pi x, T_{\rho,\varphi}(\tau_\pi x)), d_1(\tau_{2\pi} x, T_{\rho,\varphi}(\tau_\pi x))) \\
&= \min(d_1(\tau_\pi x, \tau_\pi(T_{\rho,\varphi} x)), d_1(x, \tau_\pi(T_{\rho,\varphi} x))) \\
&= \min(d_1(x, T_{\rho,\varphi} x), d_1(\tau_\pi x, \tau_{2\pi}(T_{\rho,\varphi} x))) \\
&= \min(d_1(x, T_{\rho,\varphi} x), d_1(\tau_\pi x, T_{\rho,\varphi} x)) = g(x).
\end{aligned}$$

Therefore, decomposing $n = kq + r$ with $k \ge 0, 1 \le r \le q$ leads to

$$\begin{aligned}
\frac{1}{n}\sum_{j=1}^{n} g(T_{\rho,\varphi}^{j-1} x) &= \frac{1}{n}\Big(\sum_{\nu=0}^{k-1}\sum_{\ell=1}^{q} g(T_{\rho,\varphi}^{\nu q+\ell-1} x) + \sum_{\ell=1}^{r} g(T_{\rho,\varphi}^{kq+\ell-1} x)\Big) \\
&= \frac{1}{n}\Big(\sum_{\nu=0}^{k-1}\sum_{\ell=1}^{q} g(\tau_{\nu p\pi}(T_{\rho,\varphi}^{\ell-1} x)) + \sum_{\ell=1}^{r} g(\tau_{kp\pi}(T_{\rho,\varphi}^{\ell-1} x))\Big) \\
&= \frac{k}{kq+r}\sum_{\ell=1}^{q} g(T_{\rho,\varphi}^{\ell-1} x) + \frac{1}{n}\sum_{\ell=1}^{r} g(T_{\rho,\varphi}^{\ell-1} x) \xrightarrow{n\to\infty} \frac{1}{q}\sum_{\ell=1}^{q} g(T_{\rho,\varphi}^{\ell-1} x).
\end{aligned}$$

Maximizing over $x = \theta_0 + 2\pi\mathbb{Z}$ and using (5.20) then proves case (iv) of formula (6.3).

(iii) It remains to show that the maximum in case $p = 1$ is given by $\varphi = \frac{\pi}{q}$. In this case we have $T_{\rho,\varphi}^q = \tau_\pi$ and thus equation (5.20) yields for all $\theta_0 \in [0, \frac{\pi}{2}]$

$$\frac{1}{q}\sum_{\ell=1}^{q} g(T_{\rho,\varphi}^{\ell-1} x) = \frac{1}{q}\sum_{\ell=1}^{q} \chi(\theta_j - \theta_{j-1}) \le \frac{1}{q}\sum_{j=1}^{q}(\theta_j - \theta_{j-1}) = \frac{1}{q}(\theta_q - \theta_0) = \frac{\pi}{q} = \varphi. \tag{6.7}$$

We set $\theta_0 = \theta_-$ and recall that θ_- has been chosen such that $\theta_1 - \theta_0 = \frac{\pi}{2}$. Since $\theta_j - \theta_{j-1} \geq 0$ sum up to π there is no index $j > 1$ with $\theta_j - \theta_{j-1} > \frac{\pi}{2}$, hence equality holds in (6.7). ■

Theorem 6.1 and Figures 6.1, 6.2 show that angular values can be quite sensitive to parametric perturbations. For example, approximate a rational multiple $\varphi_0 = \frac{\pi}{q}, q \geq 2$ by irrational multiples φ of π for some value ρ with $\mathrm{skew}(A(\rho, \varphi_0)) > 1$. Then the formula (6.3) and the relations (6.5), (6.6) imply

$$\liminf_{\varphi \to \varphi_0} \hat{\theta}_1(A(\rho, \varphi)) = \varphi_0 + \int_{\theta_-(\rho,\varphi_0)}^{\theta_+(\rho,\varphi_0)} \delta(\theta) d\theta < \varphi_0 = \hat{\theta}_1(A(\rho, \varphi_0)). \tag{6.8}$$

Hence the angular value $\hat{\theta}_1$ is not lower semi-continuous. However, angular values may still be upper semi-continuous.

6.2. An algorithm for computing first angular values. Based on Theorem 6.1 and on the results from Section 5, we propose the following numerical scheme for autonomous systems; see Algorithm 6.2. In case $A \in \mathbb{R}^{d,d}$ is invertible and satisfies the assumption (5.27) our numerical approach is justified by Theorem 5.7.

Algorithm 6.2 Computation of $\hat{\theta}_1(A)$

---

(1) Compute a real Schur decomposition of A

$$A = QSQ^\top, \quad S = \begin{pmatrix} \Lambda_1 & & \star \\ & \ddots & \\ 0 & & \Lambda_k \end{pmatrix}, \quad Q \in \mathbb{R}^{d,d} \text{ orthogonal, cf. (5.2),}$$

such that the diagonal blocks $\Lambda_1, \ldots, \Lambda_\ell$ are two-dimensional and $\Lambda_{\ell+1}, \ldots \Lambda_k$ are reals (such a Schur decomposition always exists, see [20, Theorem 2.3.4]). Let $\lambda_i, \bar{\lambda}_i$ be the eigenvalues of Λ_i, $i = 1, \ldots \ell$ and let $A_i = \lambda_i = \Lambda_i$ for $i = \ell + 1, \ldots, k$.

(2) Compute $\hat{\theta}_1(A)$ as follows

if $\exists i \neq j \in \{\ell + 1, \ldots, k\} : \lambda_i = -\lambda_j$ **then**
  $\hat{\theta}_1(A) = \frac{\pi}{2}$
else
  for $i = 1, \ldots, \ell$ **do**
    if $i = 1$ **then**
      $A_1 = \Lambda_1$
    else
      Compute a reordered Schur decomposition of A using `ordschur`,
      such that the upper left 2×2-block has the eigenvalue λ_i.
      Denote this upper left 2×2-block by A_i.
    end if
    Determine φ_i, ρ_i such that $A_i = |\lambda_i| A(\rho_i, \varphi_i)$.
    Compute $\hat{\theta}_1(A_i) = \theta_1(A(\rho_i, \varphi_i))$ using Theorem 6.1.
  end for
  $\hat{\theta}_1(A) = \max\{0, \hat{\theta}_1(A_i), i = 1, \ldots, \ell\}$.
end if

As explained in Remark 5.8, the Schur decomposition of A is reordered several times to obtain the diagonal blocks A_i, $i = 2, \dots, \ell$. We apply the MATLAB command `ordschur` for this task.[2] The value $\theta_1(A(\rho_i, \varphi_i))$ is calculated for $i = 1, \dots, \ell$ with Theorem 6.1. Note that the fourth case in (6.3) results in a one-dimensional optimization problem which we solve with a derivative-free method implemented in the MATLAB-routine `fminbnd`.

6.3. Numerical experiments. Let us apply Algorithm 6.2 to autonomous models with increasing complexity and dimension. The example in Section 6.3.2 particularly illustrates the need for reordering the Schur decomposition in Algorithm 6.2.

6.3.1. Block-diagonal examples. We begin with autonomous examples which have a block-diagonal structure. Due to the invariance of corresponding coordinate spaces, one can read off first angular values without the need for reordering Schur decompositions. Furthermore, the numerical calculation can be done with high accuracy and even exactly when these expressions are evaluated symbolically. Therefore, approximation errors are not discussed in Table 6.2.

A	A_1	$\hat{\theta}_1(A_1)$	A_2	$\hat{\theta}_1(A_2)$	$\hat{\theta}_1(A)$
$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$	2	0	3	0	0
$\begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$	2	0	-2	0	$\frac{\pi}{2}$
$\begin{pmatrix} c & s \\ s & -c \end{pmatrix}$	1	0	-1	0	$\frac{\pi}{2}$
$\begin{pmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & -2 \end{pmatrix}$	$\begin{pmatrix} c & -s \\ s & c \end{pmatrix}$	φ	-2	0	φ

Table 6.2: Angular values of block-diagonal examples. We abbreviate $c = \cos(\varphi)$, $s = \sin(\varphi)$, $0 < \varphi < \frac{\pi}{2}$.

6.3.2. An illustrative four-dimensional example. Using the normal form (6.1) we consider a 4×4-matrix which has already Schur form

$$A = \begin{pmatrix} A(1, \frac{1}{2}) & I_2 \\ 0 & \eta A(\frac{1}{2}, 1.4) \end{pmatrix} = \begin{pmatrix} \Lambda_1 & I_2 \\ 0 & \Lambda_2 \end{pmatrix}$$

with $\eta = 1.2$. For this matrix we have $\theta_1(\Lambda_1) = \frac{1}{2}$ and $\theta_1(\Lambda_2) = 1.128$. The algorithm sets $A_1 = \Lambda_1$ and reorders the Schur form so that the eigenvalues of Λ_2 appear in the first 2×2-block:

$$Q^T A Q = \begin{pmatrix} \eta A(0.7493, 1.4) & \star \\ 0 & A(0.6142, \frac{1}{2}) \end{pmatrix}, \quad \text{whereby } A_2 = \eta A(0.7493, 1.4).$$

From Theorem 6.1 the algorithm then finds $\theta_1(A_2) = 1.355$ and thus we have

$$\max(\theta_1(\Lambda_1), \theta_1(\Lambda_2)) = 1.128 < 1.355 = \max(\theta_1(A_1), \theta_1(A_2)) = \theta_1(A).$$

This example illustrates that first angular values can generally not be computed from the diagonal blocks of a single Schur decomposition.

[2] We are not aware of a MATLAB procedure that computes the block decomposition (5.32) directly.

$\dim(A)$	$\hat{\theta}_1(A)$	number of 2×2 blocks	initial Schur	max reordering
10^2	1.5370	45	0.0045 sec	0.0001 sec
10^3	1.5643	488	0.43 sec	0.013 sec
10^4	1.5705	4958	105 sec	1.18 sec

Table 6.3: First angular values of three random matrices: number of 2×2 blocks, analyzed by Algorithm 6.2; computing time for initial Schur decomposition; maximal time for one reordering Schur step.

6.3.3. High dimensional examples. We illustrate the performance of our algorithm for three matrices of dimension 10^2, 10^3 and 10^4. Their entries are uniformly distributed in $(0, 1)$ and generated by the MATLAB random number generator initialized with `rng(1)`. Table 6.3 documents our numerical results. We measure the time for the initial Schur decomposition and the maximal time for one reordering with `ordschur`. It turns out that the computing time for one reordering step grows linearly with the position i of the block Λ_i in the Schur form. The numerical experiments are carried out on an Intel Xeon W-2140B CPU with MATLAB 2020a.

Applying Algorithm 6.2 to the 10^4-dimensional random matrix yields $\ell = 4958$ two-dimensional blocks for which we calculate the first angular value, using Theorem 6.1. Then there are 84 real eigenvalues of different modulus leading to a vanishing angular value. Summing up we obtain $k = 5042$ one- resp. two-dimensional blocks A_i. For the presentation in Figure 6.3, these blocks are rearranged, such that

$$\theta_1(A_i) \le \theta_1(A_{i+1}) \text{ for all } i = 1, \dots, k-1. \tag{6.9}$$

The left panel shows a plot of the pairs $(i, \theta_1(A_i))_{i=1,\dots,k}$. Except for an initial ramp due to the 84 real eigenvalues, the plot suggests an almost uniform distribution of angular values. This is also confirmed by the corresponding histogram shown in the right panel. As expected, further experiments show no correlation between the modulus $|\lambda_i|$ of the eigenvalue and the angular value $\theta_1(A_i)$ of the corresponding 2×2 matrix A_i.

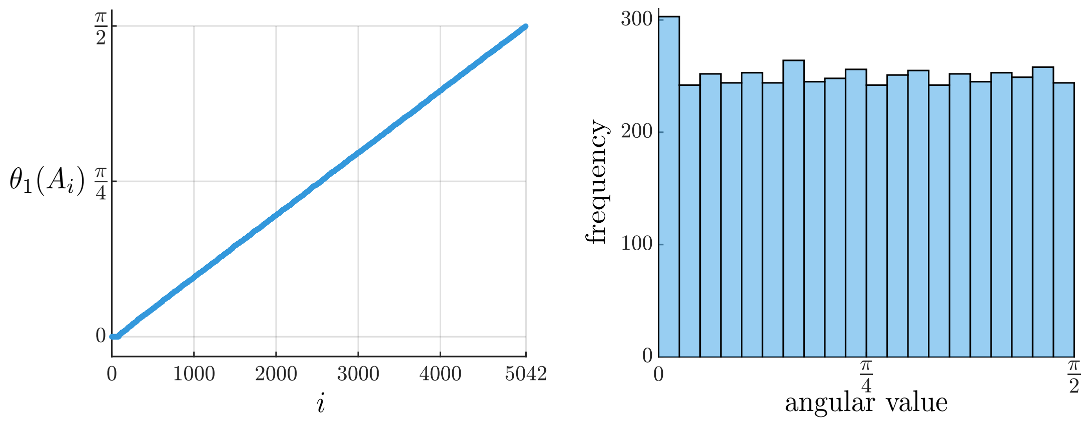

Figure 6.3: Left: sorted angular values $\theta_1(A_i)$, see (6.9), of a 10^4-dimensional random matrix; right: histogram of angular values.

Outlook. The approach of this article lends itself to several extensions and further problems which we discuss in the following.

Numerics for nonlinear systems. The content of Sections 3 and 4 applies to linear difference equations arising from variational equations of nonlinear dynamical systems of the form

$$u_{n+1} = DF(\xi_n)u_n, \quad n \in \mathbb{N}_0, \tag{6.10}$$

where $\xi_{n+1} = F(\xi_n), n \in \mathbb{N}_0$ is a bounded trajectory of a nonlinear diffeomorphism $F : \mathbb{R}^d \to \mathbb{R}^d$. To numerically estimate the angular values of these variational equations, one has to extend the numerical algorithm for the autonomous case (Section 6) to general nonautonomous systems (1.1) in dimension $d \geq 2$ and to angular values of arbitrary type $s \geq 1$. This is the topic of the forthcoming work [7], which uses a reduction procedure to generalize Theorem 5.7 from eigenvalues and eigenspaces to the dichotomy spectrum and stable and unstable fibers (see [32], [4], [30]). The algorithm is applied to variational equations. In fact, Figure 1.1 shows for the well-known Hénon map ([19]) the succession of those subspaces which lead to the outer angular value for the linearized equation (6.10).

Continuous-time systems. It is natural to set up a theory of angular values for continuous-time systems. Such an extension requires one to handle derivatives of principal angles between moving subspaces both theoretically and numerically. By Proposition 2.2 principal angles can be computed from singular values of matrices, which employ orthogonal bases of subspaces—obtained by a QR-decomposition, for example. Thus one is led to the well-known problem of computing smooth singular value and QR-decompositions which has been studied extensively in the literature; see [8], [10], [11]. One approach is to solve suitable differential equations for smooth decompositions ([10]), and this has turned out to be efficient with numerical methods for Lyapunov exponents; see [11, Section 4]. A corresponding analysis of the angle function from Section 3 and a resulting algorithm are currently under investigation.

Perturbation theory. As noted after Theorem 6.1 (see (6.8) and Figures 6.1, 6.2) angular values can be quite sensitive to parametric perturbations. In particular, without further assumptions they are not lower semi-continuous. It is an open question whether they are still upper semi-continuous in general. More specifically, it will be desirable to have criteria which ensure continuity of angular values. For Lyapunov exponents in continuous time such criteria are well known; see [1, Ch.IV,V], [11, Section 2].

Regularity theory. On the one hand, the examples in Section 3.2 demonstrate that $\liminf$ and $\limsup$ generally do not coincide for outer angular values. On the other hand, the $\liminf$ and $\limsup$ do coincide for the inner angular values and the uniform angular values in random dynamical systems and in all cases for autonomous dynamical systems; see Sections 4 and 5. It will be of interest to identify a larger class of systems for which the corresponding limits exist. This will provide a weak analogy to the class of regular continuous-time dynamical systems that have sharp Lyapunov exponents; see [1, Theorem 3.9.1].

Acknowledgments. WJB thanks the CRC 701 'Spectral Structures and Topological Methods in Mathematics' at Bielefeld University and the School of Mathematics and Statistics at the University of New South Wales for supporting his visit to GF in 2017, where this research was initiated. Further support by the CRC 1283 'Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications' is gratefully acknowledged. GF is partially supported by an ARC Discovery Project. TH thanks the Research Centre for Mathematical Modelling (RCM^2) and the Faculty of Mathematics at Bielefeld University for supporting this research project. The authors thank the referees for pointing them to various related topics and references from the literature.

REFERENCES

[1] L. Y. Adrianova. *Introduction to linear systems of differential equations*, volume 146 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1995.

[2] L. Arnold. *Random dynamical systems*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.

[3] L. Arnold and L. San Martin. A multiplicative ergodic theorem for rotation numbers. *J. Dynam. Differential Equations*, 1(1):95–119, 1989.

[4] B. Aulbach and J. Kalkbrenner. Exponential forward splitting for noninvertible difference equations. *Comput. Math. Appl.*, 42(3-5):743–754, 2001.

[5] L. Barreira. *Ergodic theory, hyperbolic dynamics and dimension theory*. Universitext. Springer, Heidelberg, 2012.

[6] L. Barreira. *Lyapunov exponents*. Birkhäuser/Springer, Cham, 2017.

[7] W.-J. Beyn and T. Hüls. Angular values of linear dynamical systems and their connection to the dichotomy spectrum. Technical report, Department of Mathematics, Bielefeld University, 2020. arXiv: 2012.11340.

[8] A. Bunse-Gerstner, R. Byers, V. Mehrmann, and N. K. Nichols. Numerical computation of an analytic singular value decomposition of a matrix valued function. *Numer. Math.*, 60(1):1–39, 1991.

[9] W. de Melo and S. van Strien. *One-dimensional dynamics*, volume 25 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3)*. Springer-Verlag, Berlin, 1993.

[10] L. Dieci and T. Eirola. On smooth decompositions of matrices. *SIAM J. Matrix Anal. Appl.*, 20(3):800–819, 1999.

[11] L. Dieci and C. Elia. The singular value decomposition to approximate spectra of dynamical systems. Theoretical aspects. *J. Differential Equations*, 230(2):502–531, 2006.

[12] L. Dieci and E. S. Van Vleck. Lyapunov and Sacker-Sell spectral intervals. *J. Dynam. Differential Equations*, 19(2):265–293, 2007.

[13] C. Elia and R. Fabbri. Rotation number and exponential dichotomy for linear Hamiltonian systems: from theoretical to numerical results. *J. Dynam. Differential Equations*, 25(1):95–120, 2013.

[14] T. Fisher and B. Hasselblatt. *Hyperbolic flows*. Zürich lectures in advanced mathematics. European Mathematical Society, Zürich, 2019.

[15] A. Galántai and C. J. Hegedűs. Jordan's principal angles in complex vector spaces. *Numer. Linear Algebra Appl.*, 13(7):589–598, 2006.

[16] I. Gohberg, M. A. Kaashoek, and J. Kos. Classification of linear time-varying difference equations under kinematic similarity. *Integral Equations Operator Theory*, 25(4):445–480, 1996.

[17] G. H. Golub and C. F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, fourth edition, 2013.

[18] K. Gustafson. *Antieigenvalue analysis*. World Scientific, Hackensack, NJ [u.a.], 2012.

[19] M. Hénon. A two-dimensional mapping with a strange attractor. *Comm. Math. Phys.*, 50(1):69–77, 1976.

[20] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, second edition, 2013.

[21] S. Jiang. Angles between Euclidean subspaces. *Geom. Dedicata*, 63(2):113–121, 1996.

[22] R. Johnson and M. Nerurkar. Exponential dichotomy and rotation number for linear Hamiltonian systems. *J. Differential Equations*, 108(1):201–216, 1994.

[23] R. A. Johnson. m-functions and Floquet exponents for linear differential systems. *Ann. Mat. Pura Appl. (4)*, 147:211–248, 1987.

[24] A. B. Katok and B. Hasselblatt. *Introduction to the modern theory of dynamical systems*. Ency-

clopedia of mathematics and its applications ; vol. 54. Cambridge University Press, Cambridge, 1995.

[25] C. Meyer. *Matrix analysis and applied linear algebra*. SIAM,Philadelphia, PA, 2000.

[26] M. Misiurewicz and K. Ziemian. Rotation sets for maps of tori. *J. London Math. Soc. (2)*, 40(3):490–506, 1989.

[27] Z. Nitecki. *Differentiable dynamics. An introduction to the orbit structure of diffeomorphisms*. The M.I.T. Press, Cambridge, Mass.-London, 1971.

[28] K. Petersen. *Ergodic theory*, volume 2 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1989.

[29] K. Polotzek, K. Padberg-Gehle, and T. Jäger. Set-oriented numerical computation of rotation sets. *J. Comput. Dyn.*, 4(1-2):119–141, 2017.

[30] C. Pötzsche. Dichotomy spectra of triangular equations. *Discrete Contin. Dyn. Syst.*, 36(1):423–450, 2016.

[31] V. Rakočević and H. K. Wimmer. A variational characterization of canonical angles between subspaces. *J. Geom.*, 78(1-2):122–124, 2003.

[32] R. J. Sacker and G. R. Sell. A spectral theory for linear differential systems. *J. Differential Equations*, 27(3):320–358, 1978.

[33] R. Sturman and J. Stark. Semi-uniform ergodic theorems and applications to forced systems. *Nonlinearity*, 13(1):113–143, 2000.

[34] P. Zhu and A. V. Knyazev. Angles between subspaces and their tangents. *J. Numer. Math.*, 21(4):325–340, 2013.

Supplementary materials.

I. Variational characterization of maximum principal angle – Proof of Proposition 2.3.

Proof. We use the following elementary fact

$$\max_{y\in\mathbb{R}^j,\|y\|=1} v^\top y = \|v\| \quad \forall\, v\in\mathbb{R}^j, v\neq 0, \tag{I.1}$$

with the maximum achieved at $y=\frac{1}{\|v\|}v$ for $v\neq 0$. Consider $w\in W$ with $\|w\|=1$ and $w^\top w_\ell = 0$, $\ell=j+1,\dots,s$. By Proposition 2.2 there exists $b\in\mathbb{R}^s$ such that

$$w = Qb = QZZ^\top b = \begin{pmatrix} w_1 & \cdots & w_s\end{pmatrix} Z^\top b. \tag{I.2}$$

Since $\|w\|=1$ and $w^\top w_\ell = 0$ for $\ell = j+1,\dots,s$ we obtain the partitioning

$$Z^\top b = \begin{pmatrix} b^I \\ 0 \end{pmatrix}, b^I\in\mathbb{R}^j, \|b^I\|=1, \quad Z = \begin{pmatrix} Z^I & Z^{II}\end{pmatrix}, Z^I\in\mathbb{R}^{d,j}.$$

By (I.1) this implies for all $v\in\mathbb{R}^d$, $v\neq 0$

$$\max_{\substack{w\in W,\|w\|=1\\ w^\top w_\ell=0,\ell=j+1,\dots,s}} v^\top w = \max_{b^I\in\mathbb{R}^j,\|b^I\|=1} v^\top QZ^I b^I = \|Z^{I\top}Q^\top v\|. \tag{I.3}$$

In a similar way, for $v\in\mathbb{R}^d$ with $\|v\|=1$ and $v^\top v_\ell=$ for $\ell = j+1,\dots,d$ we find vectors $a\in\mathbb{R}^d$, $a^I\in\mathbb{R}^j$ such that

$$v = Pa = PYY^\top a = \begin{pmatrix} v_1 & \cdots & v_s\end{pmatrix} Y^\top a, \quad Y^\top a = \begin{pmatrix} a^I \\ 0\end{pmatrix}, \quad \|a^I\|=1.$$

Using this and (2.2) in (I.3) and setting $\Sigma^I = \operatorname{diag}(\sigma_1,\dots,\sigma_j)$ leads to

$$\begin{aligned}
&\min_{\substack{v\in V,\|v\|=1\\ v^\top v_\ell=0,\ell=j+1,\dots,s}} \max_{\substack{w\in W,\|w\|=1\\ w^\top w_\ell=0,\ell=j+1,\dots,s}} v^\top w\\
&= \min_{a^I\in\mathbb{R}^j,\|a^I\|=1} \|Z^{I\top}Q^\top PY\begin{pmatrix} a^I\\ 0\end{pmatrix}\| = \min_{a^I\in\mathbb{R}^j,\|a^I\|=1} \|Z^{I\top}Z\Sigma\begin{pmatrix} a^I\\ 0\end{pmatrix}\|\\
&= \min_{a^I\in\mathbb{R}^j,\|a^I\|=1} \|Z^{I\top}Z^I\Sigma^I a^I\| = \min_{a^I\in\mathbb{R}^j,\|a^I\|=1} \|\Sigma^I a^I\|.
\end{aligned}$$

Since $\sigma_1\geq\ldots\geq\sigma_j$ the last minimum is σ_j and it is achieved at the j-th unit vector $a^I = e_j^I\in\mathbb{R}^j$. With Proposition 2.2 this yields the minimizer $v = PYe_j = v_j$, where $e_j = \begin{pmatrix} e_j^I\\ 0\end{pmatrix}\in\mathbb{R}^d$. Returning to (I.3) we obtain the maximizer $b^I = \frac{1}{\sigma_j}Z^{I\top}Q^\top v_j$ where $\sigma_j = \|Z^{I\top}Q^\top v_j\|$ is the maximum value. By (I.2) and (2.3) this leads to the maximizer of the original problem

$$\begin{aligned}
w &= \frac{1}{\sigma_j}QZ\begin{pmatrix} Z^{I\top}Q^\top v_j\\ 0\end{pmatrix} = \frac{1}{\sigma_j}QZ\begin{pmatrix} Z^{I\top}Q^\top PYe_j\\ 0\end{pmatrix}\\
&= \frac{1}{\sigma_j}QZ\begin{pmatrix} Z^{I\top}Z\Sigma e_j\\ 0\end{pmatrix} = \frac{1}{\sigma_j}QZ\begin{pmatrix} \sigma_j e_j^I\\ 0\end{pmatrix} = w_j.
\end{aligned}$$

Finally, note that taking arccos reverses min and max in (2.5). ■

II. Uniform almost periodicity – Proof of Lemma 3.6.

Proof. Let $\varepsilon > 0$. By the uniform almost periodicity there exists a $P \in \mathbb{N}$ such that for every $V \in \mathcal{V}$ and each $k \in \mathbb{N}_0$ we find a $p_k \in \{k, \dots, P+k\}$ (which may depend on V) with

$$\|b_n(V) - b_{n+p_k}(V)\| \le \frac{\varepsilon}{8} \quad \forall n \in \mathbb{N}. \tag{II.1}$$

Let $b_\infty = \sup_{n,V} \|b_n(V)\|$ and $L = \lceil \frac{16}{\varepsilon} P b_\infty \rceil$. It follows for each $k \in \mathbb{N}_0$ that

$$\begin{aligned} \Big\| \sum_{j=1}^{L} b_j(V) - \sum_{j=1}^{L} b_{j+k}(V) \Big\| &\le \sum_{j=1}^{L} \|b_j(V) - b_{j+p_k}(V)\| + \Big\| \sum_{j=1}^{L} b_{j+p_k}(V) - \sum_{j=1}^{L} b_{j+k}(V) \Big\| \\ &\le \sum_{j=1}^{L} \frac{\varepsilon}{8} + 2Pb_\infty \le L\frac{\varepsilon}{4}. \end{aligned} \tag{II.2}$$

Let $N = \lceil \frac{8}{\varepsilon} L b_\infty \rceil$ and decompose $n \ge m \ge N$ modulo L, i.e.

$$m = \ell_m L + r_m,\ 0 \le r_m < L, \quad n = \ell_n L + r_n,\ 0 \le r_n < L.$$

For $c(V) := \sum_{j=1}^{L} b_j(V)$ we obtain from (II.1) and (II.2) for each $k \in \mathbb{N}_0$ the estimates

$$\Big\| \sum_{j=1}^{L\ell_n} b_j(V) - \ell_n c(V) \Big\| \le \sum_{i=1}^{\ell_n} \Big\| \sum_{j=1}^{L} b_{j+(i-1)L}(V) - c(V) \Big\| \le \ell_n L \frac{\varepsilon}{4},$$

$$\Big\| \frac{\ell_n}{n} c(V) - \frac{\ell_m}{m} c(V) \Big\| = \Big| \frac{\ell_n r_m - \ell_m r_n}{nm} \Big| \|c(V)\| \le \frac{\ell_n L}{nm} \|c(V)\| \le \frac{1}{m} \|c(V)\| \le \frac{L}{N} b_\infty \le \frac{\varepsilon}{4}.$$

Combining these estimates, we find for every $k \in \mathbb{N}_0$

$$\begin{aligned} &\Big\| \frac{1}{n} \sum_{j=1}^{n} b_j(V) - \frac{1}{m} \sum_{j=1}^{m} b_{j+k}(V) \Big\| \\ &\le \Big\| \frac{1}{n} \sum_{j=1}^{L\ell_n} b_j(V) - \frac{1}{m} \sum_{j=1}^{L\ell_m} b_{j+k}(V) \Big\| + \Big\| \frac{1}{n} \sum_{j=L\ell_n+1}^{L\ell_n+r_n} b_j(V) - \frac{1}{m} \sum_{j=L\ell_m+1}^{L\ell_m+r_m} b_{j+k}(V) \Big\| \\ &\le \Big\| \frac{1}{n} \sum_{j=1}^{L\ell_n} b_j(V) - \frac{\ell_n}{n} c(V) \Big\| + \Big\| \frac{\ell_n}{n} c(V) - \frac{\ell_m}{m} c(V) \Big\| \\ &+ \Big\| \frac{\ell_m}{m} c(V) - \frac{1}{m} \sum_{j=1}^{L\ell_m} b_{j+k}(V) \Big\| + \Big\| \frac{1}{n} \sum_{j=L\ell_n+1}^{L\ell_n+r_n} b_j(V) - \frac{1}{m} \sum_{j=L\ell_m+1}^{L\ell_m+r_m} b_{j+k}(V) \Big\| \\ &\le \frac{\ell_n L}{n} \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\ell_m L}{m} \frac{\varepsilon}{4} + \frac{2L}{N} b_\infty \le \varepsilon. \end{aligned}$$

■

III. A matrix with a single eigenvalue – Proof of Proposition 5.4.

Proof. By Lemma 3.3 and Corollary 5.1 it suffices to show that $\bar{\theta}_1(A) = 0$. Further, by Proposition 3.8 we can assume $\lambda = 1$ and A to be in (real) Jordan normal form

$$\begin{aligned} A = \operatorname{diag}(\Lambda_1, \dots, \Lambda_k), \quad \Lambda_\ell = I_{d_\ell} + E_\ell \in \mathbb{R}^{d_\ell, d_\ell}, \\ (E_\ell)_{ij} = \delta_{i+1,j}, 1 \le i, j \le d_\ell, \ell = 1, \dots, k. \end{aligned} \tag{III.1}$$

Consider first the case $k = 1$ and drop the index ℓ. For a vector $v \in \mathbb{R}^d, v \neq 0$ let $d_\star + 1 = \max\{j \in \{1,\dots,d\} : v_j \neq 0\}$ and assume w.l.o.g. $v_{d_\star+1} = 1$. Further, we define vectors $v^j, j \in \mathbb{N}_0$ and polynomials q_i of degree $d_\star + 1 - i$ for $i = 1,\dots,d_\star+1$ by

$$\text{(III.2)} \qquad v^j := A^j v = \sum_{\nu=0}^{d_\star} \binom{j}{\nu} E^\nu v, \quad q_i(j) = (v^j)_i = \sum_{\nu=0}^{d_\star+1-i} \binom{j}{\nu} v_{i+\nu}.$$

If $d_\star = 0$ then we have $v^j = v$ for all $j \in \mathbb{N}_0$, hence all angles $\measuredangle(v^j, v^{j+1}) = 0$ vanish and do not contribute to the supremum in (3.2). Therefore, we can assume $d_\star \geq 1$. Let $z_\nu \in \mathbb{C}$, $\nu = 1,\dots,d_\star$ denote the roots of q_1 (repeated according to multiplicity) and set $x_\nu = \mathrm{Re}(z_\nu)$. Our goal is to show that there exists a constant $C_\star > 0$ independent of v such that for all $j \in \mathbb{N}_0$

$$\text{(III.3)} \qquad \measuredangle(v^j, v^{j+1}) \leq \frac{C_\star}{\min_{\nu=1,\dots,d_\star} |j - x_\nu|}, \quad \text{if} \min_{\nu=1,\dots,d_\star} |j - x_\nu| \geq 1.$$

Suppose this has been shown, then the set $M = \bigcup_{\nu=1,\dots,d_\star}(x_\nu - 1, x_\nu + 1)$ contains at most $2d_\star$ natural numbers and (III.3) leads to the estimate

$$\begin{aligned}
\frac{1}{n}\sum_{j=0}^{n-1} \measuredangle(v^j, v^{j+1}) &\leq \frac{d_\star \pi}{n} + \frac{1}{n} \sum_{j \in \{0,\dots,n-1\}\setminus M} \frac{C_\star}{\min_{\nu=1,\dots,d_\star}|j - x_\nu|} \\
&\leq \frac{d\pi}{n} + \frac{C_\star}{n} \sum_{j \in \{0,\dots,n-1\}\setminus M} \sum_{\nu=1}^{d_\star} \frac{1}{|j - x_\nu|} \\
&\leq \frac{d\pi}{n} + \frac{C_\star}{n} 2d(\log(n) + 1).
\end{aligned}$$

In the last step we used the standard estimate of the harmonic sum. The right-hand side is independent of v, taking the supremum over v and letting $n \to \infty$ shows $\bar{\theta}_1(A) = 0$.

For the proof of (III.3) let us first notice the relation $v^{j+1} - v^j = (A - I_d)v^j = Ev^j$. By (III.2) this leads to the recursion (setting $q_{d_\star+2} \equiv 0$)

$$\text{(III.4)} \qquad q_i(j+1) - q_i(j) = q_{i+1}(j), \quad j \in \mathbb{N}_0,\ i = 1,\dots,d_\star+1$$

and to the expression

$$\text{(III.5)} \qquad \|v^{j+1} - v^j\|^2 = \sum_{i=1}^{d_\star+1} q_{i+1}(j)^2 = \|v^j\|^2 - q_1(j)^2 \leq \|v^j\|^2.$$

If $q_1(j) \neq 0$ then Lemma (2.6) (i) applies and yields

$$\text{(III.6)} \qquad \begin{aligned}
\measuredangle(v^j, v^{j+1}) \leq \tan \measuredangle(v^j, v^{j+1}) &\leq \left[\frac{\|v^j\|^2 - q_1(j)^2}{q_1(j)^2}\right]^{1/2} \\
&= \left[\sum_{i=2}^{d_\star+1} \frac{q_i(j)^2}{q_1(j)^2}\right]^{1/2} \leq \sqrt{d_\star} \max_{i=2,\dots,d_\star+1} \frac{|q_i(j)|}{|q_1(j)|}.
\end{aligned}$$

In view of the recursion (III.4) and (III.6) it is sufficient to prove for some constant C_2, independent of v, and for all $\tau = 0, \dots, d_\star$ the estimate

$$\left| \frac{q_2(j+\tau)}{q_1(j)} \right| \le \frac{C_2}{\min_{\nu=1,\dots,d_\star} |j - x_\nu|}, \quad \text{if} \min_{\nu=1,\dots,d_\star} |j - x_\nu| \ge 1. \tag{III.7}$$

From $q_1(j) = \prod_{\nu=1}^{d_\star}(j - z_\nu)$ and (III.4) we obtain by expanding products

$$\begin{aligned} \left| \frac{q_2(j+\tau)}{q_1(j)} \right| &= \prod_{\nu=1}^{d_\star} |j - z_\nu|^{-1} \left| \prod_{\nu=1}^{d_\star} (j - z_\nu + \tau + 1) - \prod_{\nu=1}^{d_\star} (j - z_\nu + \tau) \right| \\ &= \prod_{\nu=1}^{d_\star} |j - z_\nu|^{-1} \left| \sum_{\substack{J \subset \{1,\dots,d_\star\} \\ |J| < d_\star}} \prod_{\nu \in J} (j - z_\nu) \left[(\tau+1)^{d_\star - |J|} - \tau^{d_\star - |J|} \right] \right| \\ &\le \sum_{\substack{J \subset \{1,\dots,d_\star\} \\ |J| \ge 1}} \left[(\tau+1)^{|J|} - \tau^{|J|} \right] \prod_{\nu \in J} |j - z_\nu|^{-1}. \end{aligned}$$

Because of $|j - z_\nu| \ge |j - x_\nu|$ and $|J| \ge 1$ we have

$$\prod_{\nu \in J} |j - z_\nu|^{-1} \le \frac{1}{\min_{\nu=1,\dots,d_\star} |j - x_\nu|}, \quad \text{if} \min_{\nu=1,\dots,d_\star} |j - x_\nu| \ge 1,$$

which proves (III.7).

The proof is easily adapted to the general Jordan form (III.1). Assertion (III.3) remains the same, but now we have block vectors $v^j = (v_1^j, \dots, v_k^j)^\top$ and polynomials $q_{i,\ell}, i = 1, \dots, d_\ell$, $\ell = 1, \dots, k$. The formula (III.5) turns into

$$\|v^{j+1} - v^j\|^2 = \|v^j\|^2 - \sum_{\ell=1}^{k} q_{1,\ell}(j)^2 = \sum_{\ell=1}^{k} (\|v_\ell^j\|^2 - q_{1,\ell}(j)^2),$$

and the estimate (III.6) is modified by using

$$\frac{\|v^j\|^2 - \sum_{\ell=1}^{k} q_{1,\ell}^2(j)}{\sum_{\ell=1}^{k} q_{1,\ell}^2(j)} \le \sum_{\ell=1}^{k} \frac{\|v_\ell^j\|^2 - q_{1,\ell}(j)^2}{q_{1,\ell}(j)^2}.$$

The subsequent arguments remain unchanged. ■

IV. Proof of the Blocking Lemma 5.5.

Proof. By scaling A and (5.25) we can arrange that $|\sigma(A_s)|, |\sigma(A_u^{-1})| < q < 1$. Then there exists a constant $C_\star$ such that

$$\|A_u^{-j} v_u\| \le C_\star q^j \|v_u\|, \quad \|A_s^j v_s\| \le C_\star q^j \|v_s\|, \quad \forall\ v_u \in X_u, v_s \in X_s. \tag{IV.1}$$

Let us first consider outer angular values and decompose $v \in \mathbb{R}^d$ as $v = v_s + v_u$, $v_s \in X_s, v_u \in X_u$.

The cases $v_s = 0$ resp. $v_u = 0$ immediately show that $\theta_1(A) \geq \max(\theta_1(A_s), \theta_1(A_u))$ holds for $\theta_1 \in \{\hat{\theta}_1, \underset{\wedge}{\theta}_1\}$. To prove the converse, we assume $v_u \neq 0$ and obtain from the triangle inequality

$$\Big|\frac{1}{n}\sum_{j=1}^{n} \measuredangle(A^{j-1}v, A^j v) - \frac{1}{n}\sum_{j=1}^{n} \measuredangle(A_u^{j-1}v_u, A_u^j v_u)\Big| \leq \frac{2}{n}\sum_{j=0}^{n} \measuredangle(A^j v, A_u^j v_u). \tag{IV.2}$$

We show that the right-hand side converges to zero as $n \to \infty$ for every v. Then the $\liminf$ and the $\limsup$ of the first two sums in (IV.2) agree and our assertion follows by taking the supremum over v. With $C_\star$, q from (IV.1) there exist an index $j_\star = j_\star(v)$ such that

$$2C_\star^2 q^{2j}\|v_s\| \leq \sqrt{3}\|v_u\|, \quad \text{for all } j \geq j_\star. \tag{IV.3}$$

The estimate (2.6) in Lemma 2.6 then shows for $j \geq j_\star$

$$\begin{aligned} \measuredangle(A^j v, A_u^j v_u) &\leq \tan\measuredangle(A^j v, A_u^j v_u) \leq \frac{\|A_s^j v_s\|}{\left(\|A_u^j v_u\|^2 - \|A_s^j v_s\|^2\right)^{1/2}} \\ &\leq \frac{C_\star q^j \|v_s\|}{\left(C_\star^{-2} q^{-2j}\|v_u\|^2 - C_\star^2 q^{2j}\|v_s\|^2\right)^{1/2}} \leq 2C_\star^2 q^{2j}\frac{\|v_s\|}{\|v_u\|}. \end{aligned} \tag{IV.4}$$

Since the right-hand side is summable our conclusion follows.

Next we analyze the inner angular values. By Corollary 5.1 it suffices to consider $\theta_1 = \bar{\theta}_1$. As above, Definition 3.1 implies the estimate $\theta_1(A) \geq \max(\theta_1(A_s), \theta_1(A_u))$, and it remains to prove the converse. From (IV.3) and (IV.4) we infer that for each $v = v_s + v_u$ with $v_u \neq 0$ the following index exists

$$k_\star = k_\star(v) = \min\{j \in \mathbb{N} : \|A^j v_s\| \leq \|A^j v_u\|\}.$$

Further choose $j_\star$ such that $2C_\star^2 q^{2j_\star} \leq \sqrt{3}$. Then (IV.3) holds for $A_s^{k_\star}v_s, A_u^{k_\star}v_u$ instead of v_s, v_u and the estimate (IV.4) yields

$$\measuredangle(A^j v, A^j v_u) \leq 2C_\star^2 q^{2(j-k_\star)} \quad \text{for } j - k_\star \geq j_\star.$$

We use $\|A_u^{k_\star-1}v_u\| \leq \|A_s^{k_\star-1}v_s\|$, (IV.1) and Lemma 2.6 to derive a corresponding estimate of angles to the stable part for $j \leq k_\star - j_\star - 1$:

$$\begin{aligned} \measuredangle(A^j v, A^j v_s) &\leq \tan\measuredangle(A^j v, A^j v_s) \\ &\leq \frac{\|A_u^{j-k_\star+1}(A_u^{k_\star-1}v_u)\|}{\left(\|A_s^{j-k_\star+1}A_s^{k_\star-1}v_s\|^2 - \|A_u^{j-k_\star+1}(A_u^{k_\star-1}v_u)\|^2\right)^{1/2}} \\ &\leq \frac{C_\star q^{k_\star-j-1}\|A_u^{k_\star-1}v_u\|}{\left(C_\star^{-2}q^{-2(k_\star-j-1)}\|A_s^{k_\star-1}v_s\|^2 - C_\star^2 q^{2(k_\star-j-1)}\|A_s^{k_\star-1}v_s\|^2\right)^{1/2}} \\ &\leq \frac{C_\star^2 q^{2(k_\star-j-1)}\|A_u^{k_\star-1}v_u\|}{\left(1 - C_\star^4 q^{4(k_\star-j-1)}\right)^{1/2}\|A_s^{k_\star-1}v_s\|} \leq 2C_\star^2 q^{2(k_\star-j-1)}. \end{aligned}$$

With these preparations the triangle inequality leads to (recall $\sum_m^n = 0$ if $m > n$)

$$\frac{1}{n}\sum_{j=1}^{n} \measuredangle(A^{j-1}v, A^j v) \le \frac{1}{n}\Big[\Big(\sum_{j=1}^{k_\star-1-j_\star} + \sum_{j=k_\star+1+j_\star}^{n}\Big)\measuredangle(A^{j-1}v, A^j v) + (2j_\star+1)\frac{\pi}{2}\Big]$$
$$\le \frac{1}{n}\Big[2C_\star^2\Big(\sum_{j=1}^{k_\star-1-j_\star} q^{2(k_\star-j-1)} + \sum_{j=k_\star+1+j_\star}^{n} q^{2(j-k_\star)}\Big) + (j_\star+1)\pi$$
$$+ \sum_{j=1}^{\min(k_\star,n)} \measuredangle(A_s^{j-1}v_s, A_s^j v_s) + \sum_{j=k_\star+1}^{n} \measuredangle(A_u^{j-1}v_u, A_u^j v_u)\Big].$$

For any given $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $k \ge n_0$, $v_s \in X_s$, $v_s \neq 0$, $v_u \in X_u$, $v_u \neq 0$ the following holds

$$\sum_{j=1}^{k} \measuredangle(A_s^{j-1}v_s, A_s^j v_s) \le k(\bar\theta_1(A_s)+\varepsilon), \quad \sum_{j=1}^{k} \measuredangle(A_u^{j-1}v_u, A_u^j v_u) \le k(\bar\theta_1(A_u)+\varepsilon).$$

Thus we have for $n \ge n_0$

$$\sum_{j=1}^{\min(k_\star,n)} \measuredangle(A_s^{j-1}v_s, A_s^j v_s) \le \begin{cases} \min(k_\star,n)(\bar\theta_1(A_s)+\varepsilon), & k_\star \ge n_0, \\ n_0\frac{\pi}{2}, & k_\star \le n_0. \end{cases}$$

With a similar estimate for $\sum_{j=k_\star+1}^{n} \measuredangle(A_u^{j-1}v_u, A_u^j v_u)$ we obtain for $n \ge n_0$ and some constant C independent of v and n

$$\frac{1}{n}\sum_{j=1}^{n} \measuredangle(A^{j-1}v, A^j v) \le \frac{1}{n}\Big[C + n_0\frac{\pi}{2} + \min(k_\star,n)(\bar\theta_1(A_s)+\varepsilon)$$
$$+ n_0\frac{\pi}{2} + (n-\min(k_\star,n))(\bar\theta_1(A_u)+\varepsilon)\Big]$$
$$\le \max(\bar\theta_1(A_s), \bar\theta_1(A_u)) + \varepsilon + \frac{1}{n}(C + n_0\pi).$$

Now take the supremum over $v \in \mathbb{R}^d, v_u \neq 0$ and then n large so that the last summand is less than ε. This finishes the proof of (5.26). ■