Bounds on Lyapunov exponents in non-Anosov systems

Patrick James Wright

Submitted in accordance with the requirements of the degree of:

Doctor of Philosophy

The University of Leeds

Department of Mathematics

United Kingdom

August 2018

Intellectual Property and Publication Statements

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

2018

University of Leeds

Patrick James Wright

Acknowledgements

First and foremost, I would like to thank my supervisors, Dr. Robert Sturman and Dr. Jitse Niesen, for their guidance, support and instruction throughout the entirety of the past four years. It goes without saying that their contribution to the thesis, and my knowledge of the subject area, has been invaluable, and this document could not have been completed without them.

In addition, I would like to thank Professor Grant Lythe for his assistance and advice for early versions of some of the chapters, as well as helping to steer me onto this course in the first place! I would also like to thank Dr. Jonathan Ward, in particular for pointers on writing and presentation given early on in the course, which I have (or hope to have!) taken on board in my writing subsequently.

I would like to thank my grandfather, Dr. John Rivière, who, along with my late grandmother, Sheila, provided the funds which have allowed me to undertake this course.

I would like to thank my father, Peter, who has helped ferry me across Leeds to different residences several times and has always been willing to provide help whenever I needed it. Finally, I would like to thank my mother, Judy, and her husband, Roger, for their love and support through trying times, and for providing a place I could flee to for a refreshing change of scenery!

Abstract

In this thesis we study a number of systems with varying degrees of hyperbolicity, including uniform and non-uniform hyperbolicity, and discuss the calculation of Lyapunov exponents in these cases. We derive and construct explicit, elementary bounds on the Lyapunov exponents of a collection of systems, which are collectively formed via the composition of shear mappings, upon the 2-torus. These bounds utilise the existence of invariant cones in tangent space to restrict the range of vectors considered in the calculations.

The bounds, with appropriate modifications, are (primarily) used to bound the Lyapunov exponents of two types of system in which their explicit calculation is not possible: a random dynamical system formed by choosing at random a hyperbolic toral automorphism, formed via shear composition, at each iterate, and the linked twist map, a deterministic system which has been used to model various physical phenomena in fluid mixing.

Following the derivation of the bounds, we discuss ways in which their accuracy can be improved. These improvements largely focus on finding a way to narrow the invariant cones used in the bounds, by considering possible preceding matrices within the orbit. We also investigate the practicality of the bounds, and how they compare to other bounds and methods of estimation for Lyapunov exponents.

Contents

1 Introduction

This thesis concerns the study of the *Lyapunov exponents* of a dynamical system, quantities which describe the rate of separation between infinitesimally close trajectories. These exponents categorise the tangent space of a system into regions within which vectors will expand (or contract) exponentially quickly, allowing many of the dynamical characteristics of a system to be studied in a tractable way. The presence of trajectories which separate from each other upon iteration is a necessary condition for chaotic behaviour, known as *sensitive dependence on initial conditions*, and as such the study of Lyapunov exponents typically concerns systems of a chaotic nature.

The first instance of the use of Lyapunov exponents is found in the doctoral thesis of their namesake, Aleksandr Mikhailovich Lyapunov [33], concerning the stability of motion. Specifically, given a function $x(t)$ with $x(t)e^{\lambda t}$ unbounded for $\lambda > \lambda_0$ and $x(t)e^{\lambda t} \to 0$ as $t \to \infty$ for $\lambda < \lambda_0$, then the critical value λ_0 is called the *characteristic number* of $x(t)$. These characteristic numbers are in fact the negative of the *largest* (or *maximal*) Lyapunov exponent as is typically defined in the literature. Positive characteristic numbers indicate that the function $x(t)$ diminishes at an exponential rate, while negative numbers indicate exponential growth; this intuition is reversed in the case of Lyapunov exponents, but the concept is identical. If, for example, the function $x(t)$ yielded the length of a vector v in tangent space after time t, then a negative (positive) characteristic number would tell us that v is being *stretched* (*contracted*) as t increases; the notions of stretching and contracting of tangent vectors will be key throughout this thesis. Following Lyapunov, further work into the theory of Lyapunov exponents can be attributed to Perron [38] in the context of linear differential equations, and Bylov *et al.* [13] regarding problems of stability.

It should be noted that Lyapunov exponents are not *a priori* guaranteed to exist, as they require the convergence of an infinite limit. Conditions for the existence of

Lyapunov exponents are given by Oseledec's Multiplicative Ergodic Theorem [35] (see Theorem 2.1), which guarantees the existence of a collection of Lyapunov exponents (known as the *Lyapunov spectrum*) in the presence of an invariant measure for the system. The theorem also guarantees that the number of distinct Lyapunov exponents cannot exceed the dimension of the system. In other words, the characterisation of the system described above into expanding, contracting and neutral regions will not return an infinite partitioning of tangent space, but rather a concise summary of the aggregate behaviour (at least in terms of growth rates) of vectors oriented in certain directions.

Another key concept we will discuss in this thesis is that of hyperbolicity. A wide variety of systems are encompassed by the term hyperbolic, from simple expanding and contracting maps, which are not chaotic, to more complex constructions, such as the linked twist map which we study in Chapters 3 and 5, which can display more interesting behaviours. In this thesis we will encounter different forms of hyperbolicity, with varying degrees of strictness.

The first, and most strict, form of hyperbolicity is known as uniform hyperbolicity; systems with this property are often referred to as Anosov diffeomorphisms, owing to the work of Anosov [3]. The archetypal example of such a system is the Arnold's Cat Map [6]. These systems fulfil a rigorous set of conditions (see Definition 2.7), including uniform bounds on minimum expansion/contraction rates and the existence of invariant subspaces in tangent space, and their behaviour has been well studied and understood; however, proving whether a system is uniformly hyperbolic can be tricky, and examples of such systems in practical applications are rare. In particular, Anosov diffeomorphisms have only been shown to exist on certain types of manifolds, and the question of which manifolds can admit them is still an open one; for example, it has been shown that no Anosov diffeomorphisms can exist upon a sphere [48].

By loosening the restrictions on the expansion and contraction rates for uniform hyperbolicity, one arrives at a more general property: non-uniform hyperbolicity. Due to the relaxation of these stringent conditions, non-uniform hyperbolicity encompasses a wider variety of systems, and subsequently a wider diversity of behaviours. The conditions for non-uniform hyperbolicity, at first glance, appear to be just as complex (and nearly as restrictive) as those of an Anosov diffeomorphism (see Section 3.2); in particular, one is still required to find invariant expanding and contracting subspaces within tangent space, and show that the expansion and contraction rates within those subspaces are non-zero. However, due to work by Pesin [39], referred to as Pesin theory, it is possible to prove that a system is non-uniformly hyperbolic by means of its Lyapunov exponents; specifically, one shows that the Lyapunov exponents of the system are non-vanishing. Non-uniformly hyperbolic systems can exist on a wider variety of manifolds than Anosov diffeomorphisms, and so their applications are more wide-ranging in practice; the key example of a non-uniformly hyperbolic system we will study in this thesis is the toral linked twist map (see Section 3.2); this map is a generalization of systems such as Arnold's Cat Map, allowing for regions with linear growth rates.

The requirement of finding systems with non-zero Lyapunov exponents seems simple in principle; however, Lyapunov exponents are notoriously difficult quantities to calculate in practice. Explicit calculation is only possible in rare cases, examples of which include Anosov diffeomorphisms such as Arnold's Cat Map (see Section 2.4). Typically an approximation is made instead, often involving *finite-time Lyapunov exponents*; finite truncations of the infinite limit one uses to calculate the Lyapunov exponents. Such approximations can vary in reliability and accuracy from system to system, as Oseledec's theorem does not say anything about the rate of convergence of finite-time Lyapunov exponents; we demonstrate this distinction for two systems in Chapter 5, and characteristic distributions of finite-time Lyapunov exponents

have been studied in a more general context [43].

An alternative approach to approximating the Lyapunov exponents is to instead find a way to bound them. If one can find a quantity which bounds the Lyapunov exponent, and this bound is non-zero (in the relevant direction), then it follows that the Lyapunov exponent itself is non-vanishing, and thus the results of Pesin theory apply. This thesis describes a method, building on the method proposed by Sturman and Thiffeault [54] for random products of shear matrices, which can provide both an upper and lower bound on systems formed via the composition of shears; such systems include the aforementioned linked twist map. Note that the Lyapunov exponents of the linked twist map have been shown to be non-vanishing (see [12]), however the bounds we obtain narrow the range of values much further than this.

Calculation (or estimation) of the (maximal) Lyapunov exponent of a system is important in many fields, as it is a solid indicator of the existence of chaotic behaviour and gives an idea of the time-scale upon which such behaviour takes place. In order to model real-world systems a random element is often utilised to modify otherwise deterministic systems, possibly through the addition of stochastic noise, or random choices of initial conditions. Similarly, one may study a system which is itself random, and indeed bounds on the Lyapunov exponents have been investigated substantially in the case of random products of matrices; see, for example, [42], [30]. Applications for the Lyapunov exponent of random products of matrices exist in areas such as physics (e.g. Schrödinger operators [11]), biology and ecology (e.g. population invasion models [24]), among many others; for further examples, see [15]. On the deterministic side, systems such as Arnold's Cat Map are viewed as simple, analytically-extractable, models of fluid mixing; in particular where two flows, or streamlines, of fluid cross one another. An example of such an application is given by the linked twist map, a generalization of such systems, which has been

used to model the Aref blinking vortex flow [4]. We discuss a method for bounding the Lyapunov exponents of random products of hyperbolic matrices in Chapter 4, and the linked twist map (and other similar deterministic systems) in Chapter 5.

In addition to establishing a close link between Lyapunov exponents and hyperbolicity in measure-preserving systems, Pesin theory also bridges the gap between hyperbolicity and measure theoretic properties such as ergodicity and mixing; this theory has been extended to the case of systems with singularities by Katok and Strelcyn [29]. The systems we study in this thesis are all Lebesgue measure-preserving, thus providing the link between the work of Pesin and the theorem of Oseledec in these cases.

Ergodicity is the notion of indecomposability of a dynamical system; the entire domain is involved in the dynamics, leaving no (non-trivial) invariant islands which fail to interact with other areas. Mixing encapsulates the idea of a deterministic system becoming independent of its initial state over time. Central to both of these concepts is the notion of recurrence, a phenomenon that exists within measure-preserving systems with a finite domain. Specifically, the Poincaré recurrence theorem [40] states that in such systems almost every orbit must return arbitrarily close to its initial condition (or indeed any other point on the orbit) infinitely many times; one way to think of this is to divide phase space into finitely-many regions of non-zero measure. Once you have visited each distinct region, you must return to a region you have previously visited.

The existence of ergodicity and mixing in the systems we study is important, as it tells us that the choice of initial condition does not matter in the evaluation of the Lyapunov exponent; by this we mean that a randomly chosen initial condition will return the same Lyapunov exponent as another with probability one. Additionally, the notion of recurrence is incorporated into the bounds we obtain for deterministic systems (i.e. the linked twist map), via the calculation of a *return time distribution*

- a construction which tells us the frequency with which points return to a particular reference region along their orbits.

Theory tells us that the Lyapunov exponents of an ergodic system are independent of the choice of initial condition. Seeing this from experimental/simulated data is not always easy however, as this independence can occur on long time scales; this can be particularly evident when the growth rate of vectors varies significantly from iterate to iterate, such as in the linked twist map. A quantity which gives an estimate of the time required to observe chaotic behaviour in a system is the *Lyapunov time*, given by the inverse of the maximal Lyapunov exponent; however, the Lyapunov exponent is a global property of the system, and so the Lyapunov time is better seen as an average time for which chaotic behaviour can occur for a typical orbit.

Alternatively, one may wish to measure this transition to independence via the time it takes for a system to achieve a certain level of homogeneity of observable functions (e.g. temperature, pressure); this is given by the *rate of decay of correlations*, which can be thought of as the rate of mixing of a system (see Section 2.9). This decay rate has been studied, using similar techniques and concepts to those we use in Chapters 4 and 5, by Ayyer and Stenlund [7], who found an upper bound on the decay rate depending upon a lower bound on the maximal Lyapunov exponent; note that the lower bounds we will obtain are an improvement upon the lower bound used by Ayyer and Stenlund. It should also be noted that while the Lyapunov exponent(s) and the rate of decay of correlations are seemingly related quantities, they are not intrinsically tied to each other; specifically, the choice of observable function can cause arbitrarily small correlation decay, despite the existence of a positive maximal Lyapunov exponent [50].

The format of the thesis is as follows: we begin in Chapter 2 by giving the background and definitions we will require for the methods later in the thesis. This

includes the (explicit) introduction of the concepts of uniform hyperbolicity, Lyapunov exponents, and invariant cones. In Chapter 3 we study a select few systems of interest, and define the notions of non-uniform hyperbolicity, with the motivating example of the linked twist map, and pseudo-Anosov maps, an alternative generalization of uniform hyperbolicity. In Chapter 4, we study and describe the formulation of bounds for random products of the composition of shear matrices, resulting in four of the main results of this thesis, Theorems 4.2 - 4.5. In Chapter 5, we apply this method to linked twist maps, with minor alterations, as well as an alternative map introduced in Chapter 3; this chapter contains the remaining three novel theorems of this thesis, Theorems 5.1 - 5.3, which concern bounds on Lyapunov exponents in linked twist maps. Finally, in Chapter 6, we summarise our findings and consider possible extensions to the theory.

2 Fundamental Properties of Measure-Preserving Dynamical Systems

The aim of this thesis is to study the quantities known as the Lyapunov exponents of a dynamical system - the growth rate of the separation between infinitesimally close trajectories - and to find methods to bound them in certain systems where explicit calculation is difficult or impossible. The Lyapunov exponents of a system are indicators of chaotic behaviour; positive Lyapunov exponents indicate expanding directions in tangent space, while negative Lyapunov exponents indicate contracting directions. We begin by giving definitions of the various mathematical objects we will require throughout the thesis in order to clarify the notations we will use.

2.1 Basic definitions

We will be concerning ourselves exclusively with systems which are measure-preserving for a measure μ which is defined on a (Borel) σ-algebra over a metric space M.

Definition 2.1. *A* **σ-algebra**, *σ, is a collection of subsets of M such that:*

(i) $M \in \sigma$,

(ii) $M \backslash A \in \sigma$ *for* $A \in \sigma$,

(iii) $\bigcup_{n\geq 0} A_n \in \sigma$ *for all* $A_n \in \sigma$ *forming a finite or infinite sequence* A_n *of subsets of* M.

We call σ the **Borel σ-algebra** *if it is the smallest σ-algebra containing all open subsets of M (i.e. if σ is* **generated** *by the collection of all open subsets of M).*

Note that we refer to a set in the Borel σ-algebra as a *Borel set*.

Definition 2.2. *A map $H : M \to M$ is said to be μ-preserving (or μ is an H invariant measure) if, for all $A \in \sigma$,*

$$\mu(H^{-1}(A)) = \mu(A). \tag{1}$$

If $\mu(M) = 1$, then μ is a *(Borel) probability measure.* If in addition (1) holds, then μ is an *invariant (Borel) probability measure.* In the case of the maps we will study, the preserved measure μ is the Lebesgue measure. Note that μ is assumed to be the Lebesgue measure unless otherwise specified; in all cases, the measure used will be an invariant Borel probability measure.

Most of the systems we study will be invertible (with the exception of random dynamical systems, see Chapter 3); in these cases, we can rewrite (1) as

$$\mu(H(A)) = \mu(A).$$

We now clarify the notation we will use for a measure-preserving dynamical system.

Definition 2.3. *A* **dynamical system** *is a quadruple (M, σ, H, μ) where M is a metric space, σ is a Borel σ-algebra over M, $H : M \to M$ is a transformation (or map), and μ is an H-invariant measure.*

As mentioned earlier, Lyapunov exponents are a quantity with close ties to chaos. Various definitions of chaotic behaviour in dynamical systems exist, and so we specify the definition of chaos as it will be understood in this thesis - that is, the definition given by Devaney.

Definition 2.4 (**[20]**)**.** *A continuous map $H : M \to M$ is said to be* **chaotic** *if:*

(i) H displays sensitive dependence on initial conditions.

(ii) H is topologically transitive.

(iii) the periodic points of H are dense in M.

In particular, positivity of the Lyapunov exponents of a system is indicative of the first of these conditions, and as such is a necessary, but not sufficient, component for chaos.

2.2 Maps formed from shear composition

The systems we study over the course of this thesis will all exhibit chaotic behaviour, and the mechanism we will use to give rise to this behaviour is the composition of shear maps on the 2-torus[1], given by

$$\mathbb{T}^2 = [0,1) \times [0,1) \mod 1.$$

Note that from this point forward if a map is defined upon $\mathbb{T}^2$ it will be considered to be mod 1, and the notation for this will be dropped. In general, let $F : \mathbb{T}^2 \to \mathbb{T}^2$ be given by

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & \alpha_i \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \text{ if } (x,y) \in P_i, \tag{2}$$

where $i \in [1,...,n]$ and $P_i = [0,1) \times [p_{i-1}, p_i)$ are horizontal annuli with $p_0 = 0$ and $p_n = 1$, and let $G : \mathbb{T}^2 \to \mathbb{T}^2$ be given by

$$G\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \beta_j & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \text{ if } (x,y) \in Q_j, \tag{3}$$

where $j \in [1,...,m]$ and $Q_j = [q_{j-1}, q_j) \times [0,1)$ are vertical annuli with $q_0 = 0$ and $q_m = 1$, and the α_i and β_j are constants in $\mathbb{R}$. To ensure continuity, we require that

$$\alpha_i(p_i - p_{i-1}), \beta_j(q_j - q_{j-1}) \in \mathbb{Z}. \tag{4}$$

We will study the dynamical system $(\mathbb{T}^2, \sigma, H, \mu)$, where $H = G \circ F : \mathbb{T}^2 \to \mathbb{T}^2$ and μ is either the Lebesgue measure or a probability measure derived from the Lebesgue

[1] In some cases we also consider subsets of the 2-torus.

measure. By convention we will always apply the horizontal shear first, however the behaviour and properties of the system are qualitatively identical with the order of application switched. We refer to the *Jacobian matrix* of the map H at the point $x \in \mathbb{T}^2$ as DH_x - when the Jacobian is identical for any choice of x, we will instead use DH. We now prove that H is measure-preserving.

Lemma 2.1. *F, G and H are Lebesgue measure-preserving.*

Proof. By the Kolmogorov Extension Theorem [31], we need only check that F and G preserve the Lebesgue measure of $A = (x_1, x_2) \times (y_1, y_2) \subset \mathbb{T}^2$, since the set of all possible A generates σ; that is, we can form any open set $A' \in \sigma$ via a possibly infinite number of unions/intersections/relative complements of sets of the form A. We have that

$$\mu(A) = (x_2 - x_1) \cdot (y_2 - y_1).$$

$F^{-1}(A)$ is a collection of parallelograms, continuously stacked upon each other, with the bottom corners at $(x_1 - \alpha_a y_1, y_1)$ and $(x_2 - \alpha_a y_1, y_1)$, and top corners at $(x_1 - \alpha_b y_2, y_2)$ and $(x_2 - \alpha_b y_2, y_2)$, where $a, b \in [1, ..., m]$, and without loss of generality $a \leq b$. Then

$$\mu(F^{-1}(A)) = (x_2 - \alpha_a y_1 - (x_1 - \alpha_a y_1)) \cdot (y_2 - y_1) = \mu(A).$$

Hence F is Lebesgue measure-preserving. The proof for G is identical, with x and y reversed. Hence

$$\mu(H^{-1}(A)) = \mu(F^{-1}(G^{-1}(A))) = \mu(G^{-1}(A)) = \mu(A),$$

and thus H is Lebesgue measure-preserving. □

The fact that the systems we study are measure-preserving means that their Lyapunov exponents will sum to zero - if a vector is expanded by some factor in one

direction, it must be contracted by this factor in another direction in order for measures to be preserved. Furthermore, many important results for measure-preserving dynamical systems are useful in the calculation and bounding of Lyapunov exponents. We will cover these results in detail over the next few sections.

2.3 Arnold's Cat Map

The archetypal example of a system formed by shear composition is Arnold's Cat Map [6]. Note that Arnold demonstrated some of the properties of this system (e.g. Poincaré recurrence) by iterating the polygons of a picture of a cat forwards, leading to the system obtaining its oft-used moniker. Let $F : \mathbb{T}^2 \to \mathbb{T}^2$ and $G : \mathbb{T}^2 \to \mathbb{T}^2$ be shears given by

$$F \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{5}$$

and

$$G \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{6}$$

then Arnold's Cat Map, $H = G \circ F$, is given by

$$H \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{7}$$

We will use the Cat Map to demonstrate the fundamental properties of measure-preserving dynamical systems. The action of the Cat Map on the unit square is shown in Figure 1. Note that the Cat Map is similarly defined by the composition of (2) and (3), with $i = j = \alpha_1 = \beta_1 = 1$.

2.4 Lyapunov exponents

We now discuss the *Lyapunov exponents*, which for $x, v \in \mathbb{R}^n$ are given by

$$\lambda_{\pm}(x, v) = \lim_{n \to \pm\infty} \frac{1}{n} \log ||DH_x^n v||, \tag{8}$$

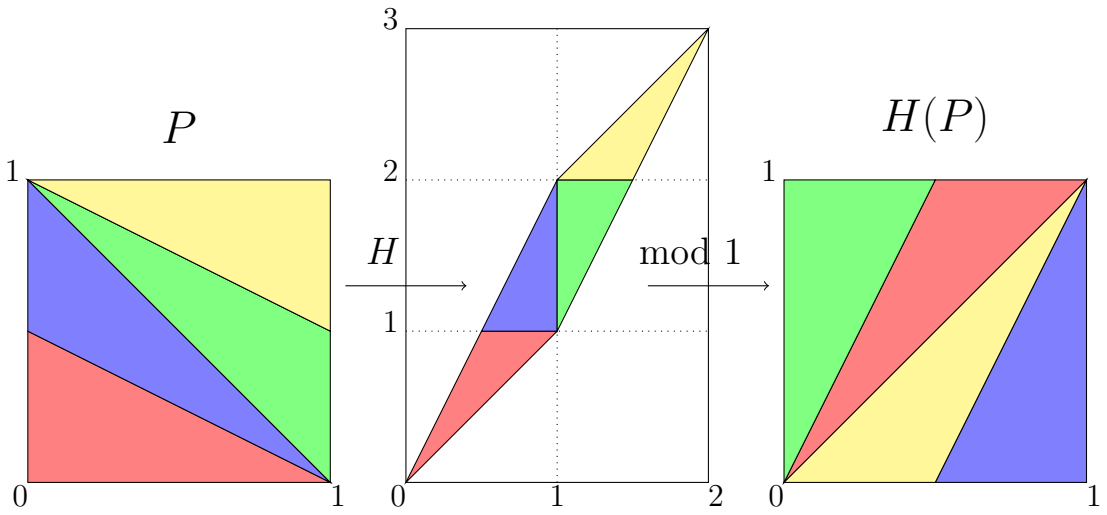


Figure 1: The action of Arnold's Cat Map on the unit square. The partition P is mapped to $H(P)$. $H(P)$ is shown on $\mathbb{R}^2$ before being reduced modulo 1.

provided the limit exists. Note that, in general, this formula yields a number of distinct Lyapunov exponents, known as the *Lyapunov spectrum*, corresponding to different choices of $x, v \in \mathbb{R}^n$. In the two dimensional measure-preserving case, there are two distinct Lyapunov exponents, $\lambda_{1,2}$, where $\lambda_1 = -\lambda_2$.

In two dimensions, the matrix DH_x is the Jacobian of the map H at the point $x \in \mathbb{R}^2$, however in the case of the Cat Map $DH_x = DH$ for all x, and so the Lyapunov exponents do not depend on x. In general this may not be true, however under certain conditions (see Theorem 2.1) $\lambda_\pm$ can be equal for μ-almost all x. Note that the Lyapunov exponents are independent of the choice of norm $||.||$; we will show this explicitly for the Cat Map.

The Lyapunov exponents quantify the rate at which infinitesimally close trajectories separate under application of the map H. In other words, Lyapunov exponents are the average rate at which the space (or vector) between two points expands or contracts as we iterate them forwards in the case of λ_1 or backwards in the case of

λ_2. A positive Lyapunov exponent in the direction v tells us that two trajectories, initially separated by the vector v, will separate exponentially quickly as we iterate H. On the other hand, a negative Lyapunov exponent tells us that the two trajectories will begin to converge upon one another at an exponential rate as we iterate H. Finally, Lyapunov exponents equal to zero indicate regions of tangent space where less than exponential (or no) growth of vectors occur.

A necessary (though not sufficient) condition for chaotic behaviour is *sensitive dependence on initial conditions*; no matter how close two trajectories are, given enough iterates, they will be separated by some minimum distance. A system which possesses only negative Lyapunov exponents cannot exhibit chaotic behaviour, since in such a system all trajectories will ultimately converge upon one another. On the other hand, a system with only positive Lyapunov exponents - known as an *expansive* system - will exhibit sensitive dependence on initial conditions; all orbits will separate from each other under iteration by H. Note that expansive systems can fail other conditions necessary for chaos, such as the existence of dense orbits (*topological transitivity*).

Let us consider the Lyapunov exponents $\lambda_{1,2}$ of the Cat Map H. In order to calculate $\lambda_{1,2}$ we require DH^n, which we obtain by diagonalising DH. The eigenvalues of DH are

$$\lambda_{u,s} = \frac{3 \pm \sqrt{5}}{2}, \tag{9}$$

with corresponding eigenvectors

$$v_{u,s} = \begin{pmatrix} 2 \\ 1 \pm \sqrt{5} \end{pmatrix}. \tag{10}$$

Let

$$P = (v_u, v_s) = \begin{pmatrix} 2 & 2 \\ 1+\sqrt{5} & 1-\sqrt{5} \end{pmatrix},$$

then

$$DH = P\Lambda P^{-1} = \frac{-1}{4\sqrt{5}} \begin{pmatrix} 2 & 2 \\ 1+\sqrt{5} & 1-\sqrt{5} \end{pmatrix} \cdot \begin{pmatrix} \lambda_u & 0 \\ 0 & \lambda_s \end{pmatrix} \cdot \begin{pmatrix} 1-\sqrt{5} & -2 \\ -1-\sqrt{5} & 1 \end{pmatrix},$$

and

$$DH^n = P\Lambda^n P^{-1}.$$

We can now rewrite (8) with $v = (x, y)$ as

$$\begin{aligned} \lambda_\pm(v) &= \lim_{n\to\pm\infty} \frac{1}{n} \log ||P\Lambda^n P^{-1} v||, \\ &= \lim_{n\to\pm\infty} \frac{1}{n} \log \left\| \begin{pmatrix} (2(1-\sqrt{5})\lambda_u^n - 2(1+\sqrt{5})\lambda_s^n)x + (-4\lambda_u^n + 4\lambda_s^n)y \\ (-4\lambda_u^n + 4\lambda_s^n)x + (-2(1+\sqrt{5})\lambda_u^n + 2(1-\sqrt{5})\lambda_s^n)y \end{pmatrix} \right\|, \\ &= \lim_{n\to\pm\infty} \frac{1}{n} \log \left\| \begin{pmatrix} (2(1-\sqrt{5})x - 4y)\lambda_u^n - (2(1+\sqrt{5})x + 4y)\lambda_s^n \\ -(4x + 2(1+\sqrt{5})y)\lambda_u^n + (4x + 2(1-\sqrt{5})y)\lambda_s^n \end{pmatrix} \right\|. \end{aligned}$$

Let $n \to +\infty$, then $\lambda_s^n \to 0$ as $n \to +\infty$, and for $v \neq v_s$ we have

$$\begin{aligned} \lambda_+ &= \log \lambda_u + \lim_{n\to\infty} \frac{1}{n} \log \left\| \begin{pmatrix} 2(1-\sqrt{5})x - 4y - (2(1+\sqrt{5})x + 4y)\left(\frac{\lambda_s}{\lambda_u}\right)^n \\ -(4x + 2(1+\sqrt{5})y) + (4x + 2(1-\sqrt{5})y)\left(\frac{\lambda_s}{\lambda_u}\right)^n \end{pmatrix} \right\|, \\ &= \log \lambda_u \approx 0.962, \end{aligned} \tag{11}$$

and for $v = v_s$ we have

$$\begin{aligned} \lambda_+ &= \log \lambda_s + \lim_{n\to\infty} \frac{1}{n} \log \left\| \begin{pmatrix} -8\sqrt{5} \\ 20 - 4\sqrt{5} \end{pmatrix} \right\|, \\ &= \log \lambda_s \approx -0.962, \end{aligned} \tag{12}$$

Similar calculations for $n \to -\infty$ yield $\lambda_- = \log \lambda^s$ for $v \neq v_s$ and $\lambda_- = \log \lambda^u$ for $v = v_s$. We have found a total of two Lyapunov exponents for the Cat Map, with the exponent we obtain depending on the choice of initial vector v, but not the choice of x. This indicates that this map possesses both expanding and contracting directions,

although we could have noted this from the eigenvector equations $DHv_{u,s} = \lambda_{u,s}v_{u,s}$ - clearly v_u is an expanding direction and v_s is a contracting direction.

We now know that Lyapunov exponents exist in the case of the Cat Map; however, for an arbitrary choice of dynamical system there is no guarantee a priori that $\lambda_\pm$ will converge at all, nor that $\lambda_+ = \lambda_-$ for any x. The following theorem by Oseledec [35] guarantees that these limits exist for almost all x, provided there exists an invariant measure.

Theorem 2.1 (**Oseledec Multiplicative Ergodic Theorem [35]**). *Let M be a compact manifold of dimension m, σ be the Borel σ-algebra on M, and $H : M \to M$ be a C^2 diffeomorphism, and let $\mathcal{M}(H)$ be the set of all invariant probability measures for H. Then there exists an invariant set $\sigma_H \in \sigma$ of full measure for every $\mu \in \mathcal{M}(H)$ such that the Lyapunov exponents exist for all points $x \in \sigma_H$. Specifically, the following properties are true.*

(a) The set σ_H is invariant, $H(\sigma_H) = \sigma_H$, and of full measure, $\mu(\sigma_H) = 1$ for all $\mu \in \mathcal{M}(H)$.

(b) For each $x \in \sigma_H$, the tangent space at x can be written as an increasing set of subspaces

$$\{0\} = V_x^0 \subset V_x^1 \subset \ldots \subset V_x^{s(x)} = T_xM$$

such that for $v \in V_x^j \backslash V_x^{j-1}$ the limit defining $\lambda(x, v)$ exists and $\lambda_j(x) = \lambda(x, v)$ is the same value for all such v, and the bundle of subspaces

$$\{V_x^j : x \in \sigma_H \text{ and } s(x) \geq j\}$$

are invariant in the sense that $DH_x V_x^j = V_{H(x)}^j$ for all $1 \leq j \leq s(x)$.

(c) The function $s : \sigma_H \to \{1, \ldots, m\}$ is a measurable function and invariant, $s \circ H = s$.

(d) If $x \in \sigma_H$, the Lyapunov exponents satisfy

$$-\infty \leq \lambda_1(x) < \lambda_2(x) < \ldots < \lambda_{s(x)}(x).$$

For $1 \leq j \leq m$, the function $\lambda_j(.)$ is defined and measurable on the set

$$\{x \in \sigma_H : s(x) \geq j\},$$

and is invariant, i.e. $\lambda_j \circ H = \lambda_j$.

This theorem guarantees that for an m-dimensional dynamical system there are at most m distinct Lyapunov exponents. In the case of the two-dimensional Cat Map, the Lebesgue measure is invariant, and we found two distinct Lyapunov exponents, $\lambda_1 = \log \lambda_u$ and $\lambda_2 = \log \lambda_s$, thus adhering to the theorem.

Calculation of Lyapunov exponents is, in general, not as simple as was the case for the Cat Map, where we obtained explicit expressions for $\lambda_\pm$; in fact, for systems where the Jacobian matrix is spatially dependent, obtaining an explicit expression may be impossible, outside of certain special cases (see Section 3.3). When explicit calculation is not possible, Lyapunov exponents are estimated numerically; however, problems arise when trying to evaluate these quantities by using (8) directly.

One such problem with a 'brute force' calculation is that the matrices DH_x^n can become *ill-conditioned*. When n is large[2], the columns of the matrix DH_x^n will tend to align themselves with the unstable eigenvector. This can lead to a small relative error in the calculation of the largest eigenvalue of DH_x^n having a large affect on the calculation of the smaller eigenvalues.

A widely used method for estimating the entire spectrum of Lyapunov exponents involves periodic *Gram-Schmidt orthonormalization* of vectors [36]; essentially, after each iterate, the algorithm 'resets' the orientation and length (norm) of the vectors, which removes the issue of the matrices becoming ill-conditioned.

[2]The number of iterates which constitutes large may vary from system to system.

The algorithm for this method in 2-dimensions is as follows (in this case, we orthonormalize after each iterate):

(i) Begin with the matrix $U = \mathbb{I}_2 = (u_1, u_2)$, a vector $v \in T_x M$, and $\Sigma = 0$, where $\mathbb{I}_2$ is the 2×2 identity matrix, and u_1 and u_2 are column vectors.

(ii) Obtain $DH \cdot U$ and set $v = DHv$.

(iii) Let $V = DH \cdot U = (v_1, v_2)$, where v_1 and v_2 are column vectors.

(iv) Set $v_2 = v_2 - (v_2 \cdot u_1)u_1$.

(v) Set $U = (\frac{v_1}{||v_1||}, \frac{v_2}{||v_2||})$.

(vi) Set $\Sigma = \Sigma + \log ||v_1|| + \log ||v_2||$.

(vii) Repeat steps (ii) through (vi) n times, where n is your desired number of iterates of H you wish to take for your estimate.

(viii) The Lyapunov exponent estimate is $\lambda = \Sigma/n$. Note that this step can be performed after each repetition of (ii) through (vi) to obtain a plot of the estimate for increasing n.

Typically the approximation improves in accuracy as n increases, with the trade-off being increased computing time. In essence this method calculates *finite-time Lyapunov exponents*, which for $x \in M$ and v in the i^{th} Lyapunov space are given by

$$\lambda_{i,n}(x, v) = \frac{1}{n} \log ||D_x H^n v||. \tag{13}$$

Finite-time Lyapunov exponents (FTLEs) are local properties of a system, whereas the Lyapunov exponents are global properties - that is, the FTLE's can depend on the choice of initial condition x. Clearly $\lambda_{i,n}$ tends to the true Lyapunov exponent λ_i as $n \to \infty$, however the rate at which this convergence occurs can vary from

system to system. In spatially dependent systems, different initial conditions can lead to varied FTLEs, even if the true Lyapunov exponents of the trajectories are in fact equal. These varying convergence rates are, in part, the motivation for finding bounds for the Lyapunov exponents of linked twist maps, which we discuss in Chapter 5.

Non-zero Lyapunov exponents are indicators of expanding or contracting regions in tangent space. As such, they are closely related to the idea of *hyperbolicity* and the study of hyperbolic dynamics. Hyperbolic measure-preserving systems exhibit both expanding and contracting behaviour, and are one of the most well understood ways of obtaining chaotic dynamics. In particular, due to the work of Pesin, non-zero Lyapunov exponents can imply various types of hyperbolicity under certain conditions. The strongest type of hyperbolicity, *uniform hyperbolicity*, is a property possessed by the Cat Map, and is the subject of the next section.

2.5 Uniform Hyperbolicity

We begin by defining what it means for a point x to be a *hyperbolic point*.

Definition 2.5. *A point $x \in M$ is a* **hyperbolic point** *of $H : M \to M$ if none of the eigenvalues of DH_x lie on the unit circle.*

This leads to the definition of what it means for a linear map to be hyperbolic.

Definition 2.6. *A linear map $H : M \to M$ is* **hyperbolic** *if none of the eigenvalues of DH lie on the unit circle.*

That is, a linear map is hyperbolic if every point in its domain is hyperbolic. Clearly the Cat Map is a hyperbolic map - its Jacobian DH has eigenvalues $\lambda_{\pm} = \frac{3\pm\sqrt{5}}{2}$ for any $x \in \mathbb{T}^2$, neither of which lie on the unit circle. The Cat Map exhibits expanding and contracting behaviour - the phase portrait of each point in the domain

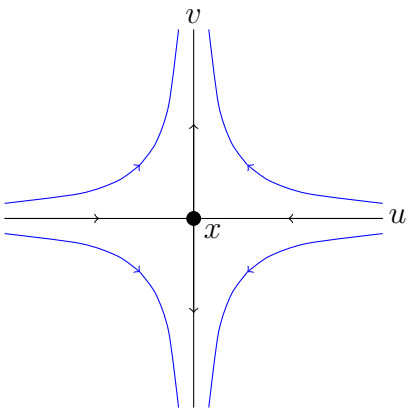


Figure 2: The phase portrait of a hyperbolic point x. The term *hyperbolic* arises due to the paths of nearby trajectories being hyperbolas relative to the orbit of x.

looks like that of a saddle point - however this is not necessarily indicative of a hyperbolic map; a hyperbolic map which is not measure-preserving may only exhibit one of these behaviours. For example, consider the expansive map $\mathcal{H} : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$\mathcal{H}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}. \tag{14}$$

$D\mathcal{H}$ has repeated eigenvalues, $\lambda = 2$, and since neither of these is on the unit circle, $\mathcal{H}$ is by definition hyperbolic. However, we see no contracting behaviour whatsoever in this map; instead, all orbits diverge from each other as we iterate. In a measure-preserving system, such as the Cat Map, if expansion occurs in one direction, contraction must occur in another to compensate. In particular, we see this behaviour at all points in the domain, not just at fixed/periodic points. Figure 2 shows an example of a local phase portrait of a point in a hyperbolic system such as the Cat Map.

In a map which possesses both expanding and contracting behaviour, one might wish to describe the expanding and contracting directions in a rigorous way. For

simplicity, let us consider a linear map $\mathcal{F} : \mathbb{R}^n \to \mathbb{R}^n$ (not necessarily hyperbolic), with eigenvalues λ_i for $i = 1, \ldots, n$, each with a corresponding eigenvector v_i (or generalised eigenvector in the case of repeated eigenvalues). There are 3 options for each i: $|\lambda_i| > 1$, $|\lambda_i| = 1$ or $|\lambda_i| < 1$.

We define a subspace E^s of $\mathbb{R}^n$ (since the map is linear, we can drop the x dependence of the tangent space) given by the span of all eigenvectors v_i with a corresponding eigenvalue $|\lambda_i| < 1$. E^s is known as the *contracting* or *stable* subspace; it contains all vectors which contract under the action of $\mathcal{F}$. We define the *expanding* (*unstable*) subspace E^u in a similar way, except it contains all vectors which expand under $\mathcal{F}$, and is given by the span of all eigenvectors with corresponding eigenvalue $|\lambda_i| > 1$. The remaining vectors are grouped into E^0; this *neutral* subspace contains all of the vectors which undergo expansion or contraction at a rate which is less than exponential and is given by the span of the eigenvectors corresponding to $|\lambda_i| = 1$.

The three subspaces E^s, E^u and E^0 contain all possible vectors, that is

$$E^s \oplus E^u \oplus E^0 = \mathbb{R}^n,$$

and are known as a *splitting* of tangent space. If $\mathcal{F}$ is a hyperbolic map, then $E^0 = \emptyset$ and all vectors undergo either exponential expansion or contraction upon iteration. In the case of the Cat Map, $E^s = \text{span}(v_-)$ and $E^u = \text{span}(v_+)$, and v_+ and v_- form a basis for $\mathbb{R}^n$; hence $E^s \oplus E^u = \mathbb{R}^n$ as desired. The Cat Map is hyperbolic, and so we do not obtain a neutral subspace E^0.

The above construction was for a linear map, which allowed for the splitting of tangent space to be defined globally (i.e. independently of the point $x \in \mathbb{R}^n$). In general this is not possible; if the Jacobian matrix $D\mathcal{F}_x$ of our map does depend on x, we instead need to find a splitting of tangent space at each point x. This idea leads to the definition of *uniform hyperbolicity*, which is the strongest form of hyperbolicity; uniformly hyperbolic systems are also commonly referred to as *Anosov diffeomorphisms*.

Definition 2.7 ([27]). *Let $\lambda < \mu$. A sequence of invertible linear maps $H_m : \mathbb{R}^n \to \mathbb{R}^n$, $m \in \mathbb{Z}$, is said to admit a (λ, μ)-***splitting** *if there exist decompositions $\mathbb{R}^n = E_m^+ \oplus E_m^-$ such that $H_m E_m^{\pm} = E_{m+1}^{\pm}$ and*

$$||H_{m|_{E_m^-}}|| \leq \lambda, \ ||H_m^{-1}{}_{|_{E_{m+1}^+}}|| \leq \mu^{-1}.$$

We say that $\{H_m\}_{m\in\mathbb{Z}}$ admits an **exponential splitting** *if it admits a (λ, μ)-splitting for some λ, μ and $\lambda < 1$, dim $E_m^- \geq 1$ or $\mu > 1$, dim $E_m^+ \geq 1$. We call $\{H_m\}_{m\in\mathbb{Z}}$* **uniformly hyperbolic** *if it admits a (λ, μ)-splitting for some $\lambda < 1 < \mu$.*

In other words, a map is uniformly hyperbolic if its Jacobian at each point admits a splitting of tangent space, upon which the growth and contraction rates of vectors are uniformly bound away from 1. The sequence of linear maps in the definition can be thought of as the linearisation at each point x on a trajectory of a possibly non-linear map H. In the case where H is linear, the H_m's are equal for all m. We now define what it means for a set of points (typically thought of as a trajectory) to be hyperbolic.

Definition 2.8 ([27]). *Let M be a smooth manifold, $U \subset M$ an open subset, $H : U \to M$ a C^1 diffeomorphism, and $\Lambda \subset U$ a compact H-invariant set. Λ is a* **hyperbolic set** *for H if there exists a Riemannian metric, known as a* **Lyapunov metric**, *in an open neighbourhood U of Λ and $\lambda < 1 < \mu$ such that for any point $x \in \Lambda$ the sequence of differentials*

$$(DH)_{H^n(x)} : T_{H^n(x)}M \to T_{H^{n+1}(x)}M,$$

$n \in \mathbb{Z}$, admits a (λ, μ)-splitting.

In other words, a set of points is hyperbolic if it is a trajectory of our map H, and the Jacobian at each point on the trajectory admits a splitting of tangent space upon which expansion and contraction rates are bound away from one.

Definition 2.9 ([**27**])**.** *A transitive, C^1 diffeomorphism $H : M \to M$ of a compact manifold M is called an* **Anosov diffeomorphism** *if M is a hyperbolic set for H.*

Hence a dynamical system (or map) is uniformly hyperbolic (or *Anosov*) if we can obtain a splitting of tangent space at every point x in the domain upon which expansion and contraction rates are bound away from one, and we can find constants which uniformly bound these rates away from one for all x. The conditions for a system to be uniformly hyperbolic are very strict - in practical applications one may find some hyperbolic sets in a system, but rarely is the entire domain so nicely behaved. As such, other less stringent forms of hyperbolicity exist, such as non-uniform hyperbolicity and pseudo-Anosov systems, where one or more of the strict conditions required for the system to be Anosov is relaxed. We will discuss these properties, and some examples of systems which possess them, in Chapter 3. Note that the transitivity of H, while not always included in the definition of an Anosov diffeomorphism, is required in order to construct Markov partitions, which we discuss in Section 2.8.

The Cat Map is an Anosov diffeomorphism - we already know we can obtain a splitting of tangent space for each x, and the constants λ and μ are simply the eigenvalues λ_u and λ_s respectively. It is a special example of a family of Anosov diffeomorphisms, which can be defined as follows: let $F : \mathbb{T}^2 \to \mathbb{T}^2$ and $G : \mathbb{T}^2 \to \mathbb{T}^2$ be shears given by

$$F \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{15}$$

and

$$G \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \beta & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{16}$$

then our family of maps, $H = G \circ F$, is given by

$$H\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & \alpha \\ \beta & 1+\alpha\beta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \tag{17}$$

where $\alpha, \beta \in \mathbb{Z}$ such that $\alpha\beta > 0$ or $\alpha\beta < -4$. We require these conditions on α and β in order to ensure hyperbolicity of the map and the ability to construct a suitable splitting of tangent space. To see this, consider the eigenvalues of H, which are given by

$$\lambda_{u,s} = \frac{2+\alpha\beta \pm \sqrt{\alpha\beta(4+\alpha\beta)}}{2}. \tag{18}$$

If $-4 < \alpha\beta < 0$, then $\alpha\beta(4+\alpha\beta) < 0$ and we therefore have complex eigenvalues, and no real eigenvectors. For example, if we let $\alpha = 1$ and $\beta = -1$ then we find that $H^6 = \mathbb{I}_2$; all points are periodic with period 6, hence no chaotic behaviour is observed. Note that λ_u and λ_s denote the unstable and stable eigenvalues respectively, and the sign in front of the square root can switch in certain cases.

Provided that α and β adhere to the aforementioned conditions, the behaviour of these maps is qualitatively identical to that of the Cat Map; we find expanding and contracting subspaces which correspond to the eigenvectors v_u and v_s, and we can find constants which uniformly bound the expansion and contraction rates away from one (again these are the eigenvalues $\lambda_{u,s}$). We also find that the Lyapunov exponents of H are $\lambda_{1,2} = \log \lambda_{u,s}$ via a similar calculation to the Lyapunov exponents of the Cat Map.

When considering expanding and contracting directions, one may wish to know which points approach or move away from each other (i.e. the vectors between them contract or expand) for all iterates. This gives rise to the idea of *local* and *global stable* and *unstable manifolds*. We first define what we mean by a local stable (or unstable) manifold of a fixed point x_*. Note that in the following definition the use of the inverse H^{-1} is justified since the map H is assumed to be a diffeomorphism.

Definition 2.10. *The* **local stable manifold** *of a fixed point x_* of a diffeomorphism $H : M \to M$ in a neighbourhood $B(x_*, r)$ of x_* is given by*

$$\gamma_s(x_*) = \{x \in B(x_*, r) : d(H^n(x), x_*) \to 0 \text{ as } n \to \infty\},$$

and its **local unstable manifold** *by*

$$\gamma_u(x_*) = \{x \in B(x_*, r) : d(H^{-n}(x), x_*) \to 0 \text{ as } n \to \infty\},$$

where $B(x_*, r)$ is the open ball of radius r about x_*.

The following theorem guarantees the existence of such manifolds provided that x_* is a hyperbolic fixed point.

Theorem 2.2 (**[27]**). *Let x_* be a hyperbolic fixed point of a local C^r diffeomorphism $H : U \to M$, $r \geq 1$, for some Borel subset U of M. Then there exist C^r embedded discs $\gamma^u_{x_*}, \gamma^s_{x_*} \subset U$ such that:*

(i) $T_{x_*}\gamma^u_{x_*} = E^u(DH_{x_*})$ and $T_{x_*}\gamma^s_{x_*} = E^s(DH_{x_*})$ (the unstable and stable subspaces under the respective Jacobian),

(ii) $H^{-1}(\gamma^u_{x_*}) \subset \gamma^u_{x_*}$ and $H(\gamma^s_{x_*}) \subset \gamma^s_{x_*}$.

Also, there exists $C(\delta)$ such that for any $y \in \gamma^s_{x_*}$, $z \in \gamma^u_{x_*}$, $m \geq 0$,

(iii) $d(H^m(y), x_*) < C(\delta)(\lambda(DH_{x_*}) + \delta)^m d(y, x_*)$,

(iv) $d(H^{-m}(z), x_*) < C(\delta)(\mu^{-1}(DH_{x_*}) + \delta)^m d(z, x_*)$,

Furthermore, there exists $\delta_0 > 0$ such that

(a) if $d(H^m(y), x_*) \leq \delta_0$ for $m \geq 0$, then $y \in \gamma^s_{x_*}$,

(b) if $d(H^m(z), x_*) \leq \delta_0$ for $m \leq 0$, then $z \in \gamma^u_{x_*}$.

The discs $\gamma^s_{x_*}$ and $\gamma^u_{x_*}$ are local stable and unstable manifolds of the fixed point, and may not be unique. Note that, by definition, $H^{-1}(\gamma^s_{x_*})$ and $H(\gamma^u_{x_*})$ are also local stable and unstable manifolds. We use this fact to define the global stable and unstable manifolds of a hyperbolic fixed point x_*.

Definition 2.11. *The manifolds*

$$\Gamma^s_{x_*} = \bigcup_{m \leq 0} H^m(\gamma^s_{x_*})$$

and

$$\Gamma^u_{x_*} = \bigcup_{m \geq 0} H^m(\gamma^u_{x_*})$$

are the **global stable** *and* **unstable manifolds** *of H at the point x_* respectively.*

Equivalently, the global stable (unstable) manifold contains all points which tend to x_* for all forward (backward) iterates of H, that is

$$\Gamma^s_{x_*} = \{x \in U : d(H^m(x), x_*) \to 0 \text{ as } m \to \infty\}, \tag{19}$$

and

$$\Gamma^u_{x_*} = \{x \in U : d(H^{-m}(x), x_*) \to 0 \text{ as } m \to \infty\}. \tag{20}$$

In the case of the fixed point $(0, 0)$ of the Cat Map (and indeed any map given by (17)), the global stable and unstable manifolds are simply the eigenvectors v_- and v_+ stretched to an infinite length; subsequently, the global stable and unstable manifolds are dense in $\mathbb{T}^2$ in these cases, each consisting of a vector wrapping around $\mathbb{T}^2$ an infinite number of times. In general, however, $\Gamma^s_{x_*}$ and $\Gamma^u_{x_*}$ are much more complicated, often forming structures called *homoclinic tangles* in non-linear maps ([41], see also [27]).

We noted earlier that the Cat Map possesses stable and unstable subspaces not only at its fixed point(s), but at every point in its domain. As such, we may consider whether stable and unstable manifolds exist for any point in the domain.

The extension of these ideas to non-fixed points is justified by the Hadamard-Perron theorem, which gives us the local stable and unstable manifolds for hyperbolic maps, as well as more general C^r diffeomorphisms under the specified conditions.

Theorem 2.3 (**Hadamard-Perron Theorem [27]**). *Let $\lambda < \mu$, $r \geq 1$ and for each $m \in \mathbb{Z}$ let $H_m : \mathbb{R}^n \to \mathbb{R}^n$ be a (surjective) C^r diffeomorphism such that for $(x, y) \in \mathbb{R}^k \bigoplus \mathbb{R}^{n-k}$*

$$H_m(x, y) = (A_m(x) + \alpha_m(x, y), B_m(y) + \beta_m(x, y))$$

for some linear maps $A_m : \mathbb{R}^k \to \mathbb{R}^k$ and $B_m : \mathbb{R}^{n-k} \to \mathbb{R}^{n-k}$ with $||A_m^{-1}|| \leq \mu^{-1}$, $||B_m|| \leq \lambda$, $\alpha_m(0) = 0$ and $\beta_m(0) = 0$.

Then for $0 < \omega < \min(1, \sqrt{\mu/\lambda} - 1)$ and

$$0 < \delta < \min\Big(\frac{\mu - \lambda}{\omega + 2 + 1/\omega}, \frac{\mu - (1+\omega)^2\lambda}{(1+\omega)(\omega^2 + 2\omega + 2)}\Big)$$

we have: If $||\alpha_m||_{C^1} < \delta$ and $||\beta_m||_{C^1} < \delta$ for all $m \in \mathbb{Z}$ then there is

(a) a unique family $\{\gamma_m^u\}_{m\in\mathbb{Z}}$ of k-dimensional C^1 manifolds

$$\gamma_m^u = \{(x, \phi_m^u(x)) : x \in \mathbb{R}^k\},$$

and

(b) a unique family $\{\gamma_m^s\}_{m\in\mathbb{Z}}$ of $(n-k)$-dimensional C^1 manifolds

$$\gamma_m^s = \{(x, \phi_m^s(x)) : x \in \mathbb{R}^{n-k}\},$$

where $\phi_m^u : \mathbb{R}^k \to \mathbb{R}^k$, $\phi_m^s : \mathbb{R}^{n-k} \to \mathbb{R}^{n-k}$, $\sup_{m\in\mathbb{Z}} ||D\phi_m^{u,s}|| < \omega$, and the following properties hold:

(i) $H_m(\gamma_m^s) = \gamma_{m+1}^s$ and $H_m(\gamma_m^u) = \gamma_{m+1}^u$.

(ii) $||H_m(z)|| < \lambda'||z||$ for $z \in \gamma_m^s$, and $||H_{m-1}^{-1}(z)|| < (\mu')^{-1}||z||$ for $z \in \gamma_m^u$, where

$$\lambda' := (1+\omega)(\lambda + \delta(1+\omega)) < \frac{\mu}{1+\omega} - \delta := \mu'.$$

(iii) *Let* $\lambda' < \nu < \mu'$. *If* $||H_{m+L-1} \circ \ldots \circ H_m(z)|| < C\nu^L||z||$ *for all* $L \geq 0$ *and some* $C > 0$ *then* $z \in \gamma_m^s$.
Similarly, if $||H_{m-L}^{-1} \circ \ldots \circ H_{m-1}^{-1}(z)|| \leq C\nu^{-L}||z||$ *for all* $L \geq 0$ *and some* $C > 0$ *then* $z \in \gamma_m^u$.

This theorem is analogous to Theorem 2.2, except we have to take care that the points which we aim to analyse are no longer fixed, and so we need to calculate their motion relative to each other. In the case of hyperbolic maps, we have $\lambda < 1 < \mu$, and the families of manifolds instead consist of C^r manifolds - in other words, they are at least as smooth as the map itself. The proof of the theorem (see [27]) involves the construction of families of invariant cones - these are an important concept in hyperbolic dynamics, and we will ultimately use invariant cones in the construction of rigorous bounds upon the Lyapunov exponent of specific families of maps in Chapters 4 and 5. Furthermore, due to results by Alekseev [2], the existence of invariant cones under certain conditions can imply the (uniform) hyperbolicity of a given map.

2.6 Invariant Cones

Determining whether a given map is (uniformly) hyperbolic can be difficult. In the case of the Cat Map and similar systems, the stable and unstable subspaces are relatively easy to obtain; however, in the case of non-linear maps, these subspaces can be much more difficult to find. A way of sidestepping this problem is provided by the *Alekseev Cone Criterion*, which states that the subspaces are given by the intersection of a family of *invariant cones* - if we can find these cones, we can subsequently calculate the subspaces. First, we define what it means for a bundle of tangent vectors to be called a *cone*.

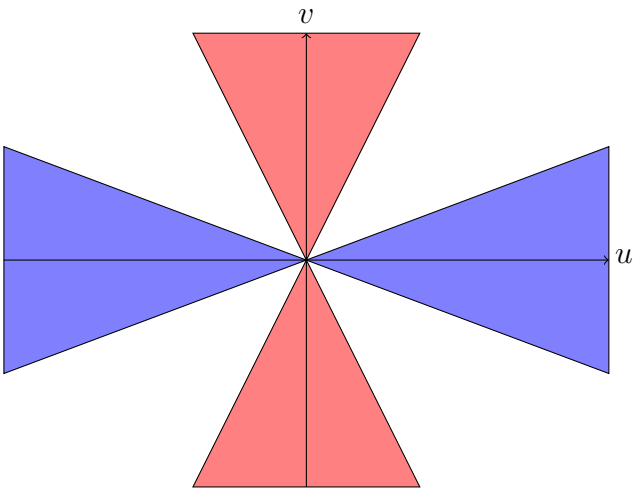


Figure 3: An example of a horizontal (blue) and vertical (red) cone.

Definition 2.12 ([27]). *The* **standard horizontal ω-cone** *at $x \in \mathbb{R}^n$ is defined by*

$$\alpha_x^\omega = \{(u, v) \in \mathbb{R}^n : ||v|| < \omega||u||\}.$$

The **standard vertical ω-cone** *at $x \in \mathbb{R}^n$ is defined by*

$$\beta_x^\omega = \{(u, v) \in \mathbb{R}^n : ||v|| < \omega||u||\}.$$

In general, a **cone** *$C \in \mathbb{R}^n$ is the image of one of these standard cones under an invertible linear map.*

All of the cones we study in this thesis will be two-dimensional, and hence will look similar to those depicted in Figure 3. We define a *cone field* to be a map that associates each $x \in \mathbb{R}^n$ with a cone $C_x \subset \mathbb{R}^n$, and a *cone family* to be a sequence of cone fields K. A sequence of diffeomorphisms $H = \{H_m\}_{m\in\mathbb{Z}}$ can act on a cone family by

$$(H(K))_{x,m} = (DH_{m-1})_{H_{m-1}^{-1}(x)}(K_{H_{m-1}^{-1}(x),m-1}).$$

Definition 2.13. *A cone family K is called* **(strictly) invariant** *if*

$$(H(K))_{x,m} \subset \mathrm{Int}(K_{x,m}) \cup \{0\}.$$

In other words, a cone is invariant if it is mapped within itself by the map H. We say that a cone C is a *minimal invariant cone* for the map H if

$$H(C) = \text{ Int}(C) \cup \{0\}. \tag{21}$$

That is, a minimal invariant cone is the smallest cone which is left invariant by H. Note that this cone can be obtained by intersecting all other invariant cones for H. With these definitions, we are now ready to state the Alekseev Cone Criterion [2]. We state the specific version given in [27].

Proposition 2.1 (**Alekseev Cone Criterion [27]**). *Let $\lambda' < \mu'$ and $\omega_m, \omega'_m > 0$ for $m \in \mathbb{Z}$. Let $L_m : \mathbb{R}^k \times \mathbb{R}^{n-k} \to \mathbb{R}^k \times \mathbb{R}^{n-k}$ be a sequence of invertible maps such that*

(i) $L_m \alpha^{\omega_m} \subset \text{Int}(\alpha^{\omega_{m+1}})$;

(ii) $L_m^{-1} \beta^{\omega'_{m+1}} \subset \text{Int}(\beta^{\omega'_m})$;

(iii) $||L_m(u,v)|| > \mu' ||(u,v)||$ *for* $(u,v) \in \alpha^{\omega_m}$;

(iv) $||L_m(u,v)|| < \lambda' ||(u,v)||$ *for* $(u,v) \in L_m^{-1} \beta^{\omega'_{m+1}}$.

Then

$$E^u_m := \bigcap_{i=0}^{\infty} L_{m-1} \circ L_{m-2} \circ \ldots \circ L_{m-i} \alpha^{\omega_{m-i}}$$

is a k-dimensional (unstable) subspace inside α^{ω_m} and

$$E^s_m := \bigcap_{i=0}^{\infty} L_m^{-1} \circ L_{m+1}^{-1} \circ \ldots \circ L_{m+i}^{-1} \beta^{\omega'_{m+i+1}}$$

is an $(n-k)$-dimensional (stable) subspace inside $\beta^{\omega'_m}$. Furthermore, if $\lambda' < 1 < \mu'$, then $\{L_m\}$ is a hyperbolic family of linear maps which admits a $(\lambda'\mu')$-splitting.

In other words, the cone criterion tells us that if we can find invariant cones for a map H and H^{-1} (conditions (i) and (ii)), one containing vectors which contract by

a factor of at least λ' (the *contracting cone*) and the other containing vectors which expand by a factor of at least μ' (the *expanding cone*) (conditions (iii) and (iv)), then the stable and unstable subspaces can be found by simply finding the images of these cones under repeated applications of the map H upon the expanding cone for the unstable subspace, and H^{-1} upon the contracting cone for the stable subspace.

In the case of the Cat Map (or any map of the form (17)), repeated applications of DH to a vector $v \neq v_-$ will begin to align the vector with v_+. Hence an invariant cone C^u for H is given by any cone which contains v_+ but does not contain v_-. Similarly, repeated applications of DH^{-1} align the vector $v \neq v_+$ with v_-, and so any cone C^s which contains v_- but not v_+ is an invariant cone for H^{-1}. The intersection of all such possible cones (of each particular case) yields the minimal invariant cones, which are the subspaces given by $E^u = \text{span}(v_+)$ and $E^s = \text{span}(v_-)$. Note that $E^u \bigoplus E^s = \mathbb{R}^2$, as required. In fact, an invariant cone (and by the intersection of all such cones, the minimal invariant cone) for any 2×2 diagonalizable matrix (with positive, distinct eigenvalues) is given by the following theorem by Rodman *et al.*

Theorem 2.4 ([**46**])**.** *Let a 2×2 matrix A be diagonalizable, with eigenvalues $\lambda_1 > \lambda_2 \geq 0$. Then a proper cone $C \subset \mathbb{R}^2$ is A-invariant if and only if it contains an eigenvector of A corresponding to λ_1 and its interior does not intersect the eigenline of A corresponding to λ_2.*

Thus the eigenvectors v_+ and v_- are the minimal invariant expansion and contraction cones respectively for any map H given by (17). We will discuss the invariant cones for the map given by the composition of (2) and (3) in Chapter 4.

We will use invariant cones to aid us in bounding the Lyapunov exponents of maps which, while not Anosov themselves, possess orbits that can be split up into chains of Jacobian matrices which yield Anosov maps, and therefore possess invariant cones. The difficulty in these calculations comes first in finding the invariant cone itself, and second in finding the frequency with which certain chains of matrices

occur. We will look at two specific cases for these bounds: Anosov maps of the form (17) chosen at random on each iterate in Chapter 4, and linked twist maps under specific parameter values in Chapter 5.

In the case of the linked twist map, the issue of calculating the frequency of certain orbits requires the partitioning of the domain in a special way in order to guarantee uniformly hyperbolic chains of matrices - a *return time partition/distribution*. This idea is made possible due to important results for general measure-preserving systems, as well as the fact that the linked twist map possesses a quality called *ergodicity*. Essentially, in an ergodic system, almost every orbit must visit every set of positive measure within the domain, and must do so an infinite number of times. In particular, a result called the *Poincaré recurrence theorem* [40] guarantees that, in a measure-preserving system, almost every point must return to a neighbourhood of itself, however small it may be, infinitely often; this guarantees that, in the case of the linked twist map, an (appropriately chosen) orbit will return to a reference set in the domain in a finite number of iterates, which allows for the construction of a return time partition. We discuss the topics of recurrence in measure-preserving systems and ergodicity in the next section.

2.7 Recurrence and ergodicity

An important notion in the context of measure-preserving systems is that of recurrence: the idea that a typical point will return arbitrarily close to where it started under iteration. An important result that tells us this fact is the Poincaré recurrence theorem, which states that not only must (almost every) point in a measure-preserving system return within a neighbourhood of itself in a finite number of iterates, but it must do so infinitely many times on its orbit. A proof of this theorem is given in [27].

Theorem 2.5 (**Poincaré recurrence theorem [40]**). *Let H be a measure-preserving transformation of a probability space (M, μ) and let $A \subset M$ be a measurable set ($\mu(A) > 0$). Then for any $N \in \mathbb{N}$*

$$\mu(\{x \in A : \{H^n(x)\}_{n \geq N} \subset X \backslash A\}) = 0.$$

Another important result for measure-preserving systems is the *Birkhoff ergodic theorem* [10]. In applications, one may be interested in measuring a quantity - known as an *observable* - at each particular point on an orbit; examples of observables include temperature, pressure, concentration or a characteristic function of a subset of the domain. From a theoretical standpoint, these quantities are space-dependent functions $\phi(x)$, which may be measurable, integrable, differentiable and/or continuous. The Birkhoff ergodic theorem tells us that in a measure-preserving system, for a typical (μ-almost every) initial condition, we can obtain a *time average*, $\phi^+(x)$, and the spatial average of this time average is equal to the spatial average of ϕ.

Theorem 2.6 (**Birkhoff ergodic theorem [10]**). *Let $H : (M, \mu) \to (M, \mu)$ be a measure-preserving transformation of a probability space and $\phi \in L^1(M, \mu)$ be an observable function. Then for μ-almost every $x \in M$ the following time average exists:*

$$\phi^+(x) := \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \phi(H^i(x)). \tag{22}$$

Moreover,

$$\int_M \phi^+(x) d\mu = \int_M \phi(x) d\mu.$$

The theorem can be restated to find that the backward time average

$$\phi^-(x) := \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \phi(H^{-i}(x)) \tag{23}$$

also exists for μ-almost every $x \in M$, and

$$\int_M \phi^-(x) d\mu = \int_M \phi(x) d\mu.$$

An immediate corollary of Theorem 2.6 is that forward time averages equal backward time averages for μ-almost every $x \in M$, since

$$\int_M \phi^+(x)d\mu = \int_M \phi(x)d\mu = \int_M \phi^-(x)d\mu,$$

via two applications of Theorem 2.6, or equivalently by finding that the measure of the set $A = \{x \in M : \phi^+(x) > \phi^-(x)\}$ is zero. Note that this theorem *does not* say that spatial averages equal time averages - in other words, that the average of ϕ over a single orbit is the same as the average over the entire domain - for this, we require *ergodicity*.

Definition 2.14. *A measure-preserving transformation $H : M \to M$ of a probability space (M, σ, μ) is* **ergodic** *if for any $A \in \sigma$,*

$$H^{-1}(A) = A \implies \mu(A) = 0 \text{ or } \mu(A) = 1. \tag{24}$$

Equivalently, one may also say that μ is an **ergodic measure** for H. The definition tells us that the only invariant sets under the action of an ergodic map are of measure zero or one; any set A with $0 < \mu(A) < 1$ is not invariant under an ergodic map. The following are alternative, equivalent definitions of ergodicity, which demonstrate various properties ergodic systems must possess.

Proposition 2.2 (**[21]**). *Under the conditions of Definition 2.14, the following statements are equivalent:*

(i) H is ergodic.

(ii) For any $A \in \sigma$, $\mu(H^{-1}(A)\Delta A) = 0$ implies that $\mu(A) = 0$ or $\mu(A) = 1$.[3]

(iii) For $A \in \sigma$, $\mu(A) > 0$ implies that $\mu(\bigcup_{n=1}^{\infty} H^{-n}(A)) = 1$.

[3] $A\Delta B = (A\backslash B) \cup (B\backslash A)$ denotes the *symmetric difference* of the sets A and B.

(iv) *For $A, B \in \sigma$, $\mu(A)\mu(B) > 0$ implies that there exists $n \geq 1$ with $\mu(H^{-n}(A) \cap B) > 0$.*

(v) *For measurable $f : M \to \mathbb{C}$, $f \circ H = f$ almost everywhere implies that f is equal to a constant almost everywhere.*

For a proof of the above proposition, see [21]. Statement (iii) highlights the idea of *indecomposability* of ergodic systems; it states that, if we look back far enough, the pre-image of A will intersect any set of positive measure in M. In other words, eventually the orbit of almost every point in the system will enter A, so we cannot split the domain into two or more sub-domains which are self-invariant. If a system is not ergodic, it may possess what is know as an *ergodic decomposition* - that is, a partitioning of the domain into subregions M_i, where the restriction of the map H to each M_i is an ergodic map.

Example 2.1. *Theorem 2.5 tells us that in a measure-preserving system (of finite measure), an orbit must return to near where it started. It does not, however, tell us that an orbit must visit any set of positive measure - ergodicity. To illustrate this, consider the dynamical system (M, σ, H, μ) where $M = M_1 \cup M_2$, M_1 and M_2 are separate, distinct copies of the 2-torus, and the restriction of H to M_i, $i = 1, 2$, is the Cat Map; clearly, H is measure-preserving. The orbit of a point $x \in M_1$ can never enter M_2, and as such cannot visit any region of positive measure in M_2.*

In addition to sensitive dependence on initial conditions, another required condition for chaotic behaviour is *topological transitivity.*

Definition 2.15. *A dynamical system $H : M \to M$ is* **topologically transitive** *if for every pair of non-empty, open sets $A, B \subset M$, there exists an integer n such that $H^n(A) \cap B \neq \emptyset$.*

In the case of homeomorphisms on a compact metric space (which encompasses all systems we study in this thesis), this property is equivalent to the existence of a

dense orbit (this is the *Birkhoff Transitivity Theorem*, see [45]). Ergodicity implies topological transitivity, provided $\mu(A) > 0$ for all non-empty, open $A \subset M$. If, in addition, such ergodic systems have positive Lyapunov exponents, then they possess (at least) two of the three conditions required for chaotic behaviour. We now show that the Cat Map (and its generalization (17)) are ergodic systems.

Lemma 2.2. *Let $H : \mathbb{T}^2 \to \mathbb{T}^2$ be given by (17). Then H is ergodic w.r.t. Lebesgue measure.*

Proof. (See, for example, [27]) We use Condition (v) of Proposition 2.2, and show that all H-invariant functions f are constant almost everywhere. Let us consider a function $f \in L^2$ with $f \circ H = H$, i.e. H is invariant under f. Since $f \in L^2$, it has a Fourier series given by

$$f(x,y) = \sum_{m,n\in\mathbb{Z}} c_{m,n} e^{2\pi imx} e^{2\pi iny},$$

where $c_{m,n} \in \mathbb{C}$. We have that

$$H\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + \alpha y \\ \beta x + (1 + \alpha\beta)y \end{pmatrix},$$

hence

$$f \circ H(x,y) = \sum_{m,n\in\mathbb{Z}} c_{m,n} e^{2\pi i(m+\beta n)x} e^{2\pi i(\alpha m+(1+\alpha\beta)n)y}.$$

By assumption we must have that $f \circ H$ and f are equal, hence

$$\sum_{m,n\in\mathbb{Z}} c_{m,n} e^{2\pi imx} e^{2\pi iny} = \sum_{m,n\in\mathbb{Z}} c_{m,n} e^{2\pi i(m+\beta n)x} e^{2\pi i(\alpha m+(1+\alpha\beta)n)y},$$

which can only happen for all $(x,y) \in \mathbb{T}^2$ if

$$c_{m,n} = c_{m+\beta n,\alpha m+(1+\alpha\beta)n},$$

for all $m, n \in \mathbb{Z}$, or in other words if

$$c_v = c_{Hv}$$

for all $v \in \mathbb{Z}^2$. By subsequent iteration, we therefore also have $c_v = c_{H^n v}$ for $n \in \mathbb{Z}$. H is hyperbolic, so we know that for $v \neq v_-$ (the stable eigenvector of DH) and $v \neq 0$, we have $\lim_{n\to\infty} ||H^n v|| \to \infty$. We know further that $v_- \neq v \in \mathbb{Z}^2$ since v_- has an irrational slope. We thus have either $v = 0$ or $\lim_{n\to\infty} ||H^n v|| \to \infty$.

For $f \in L^2$, $c_v \to 0$ as $||v|| \to \infty$ by the Riemann-Lebesgue lemma, and so we have

$$c_v = c_{H^n v} \to 0.$$

Hence c_v is only non-zero when $v = 0$, and the Fourier series of f is given by

$$f(x, y) = c_0.$$

Hence f is a constant function almost everywhere, and therefore H is ergodic. □

Example 2.2. *Consider again the system described in Example 2.1. We now know that the Cat Map is ergodic, and so the restriction of H to each M_i is ergodic. Hence, the ergodic decomposition of this map is given by $\{M_1, M_2\}$.*

The following corollary states that in ergodic systems, time averages and spatial averages of observable functions are equal.

Corollary 2.1. *Let $H : M \to M$ be an ergodic μ-preserving transformation with $\mu(M) = 1$, and let $\phi \in L^1(M, \mu)$. Then for every x outside of a set of measure zero,*

$$\phi_H(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \phi(H^i(x)) = \int_M \phi \, \mathrm{d}\mu.$$

Proof. H is ergodic and ϕ_H is H-invariant, hence, by Proposition 2.2(v), ϕ_H is constant almost everywhere. By Birkhoff's Ergodic Theorem, that constant must be $\int_M \phi \, \mathrm{d}\mu$. □

We now know that the Cat map and its generalization are ergodic; a typical orbit in these systems will visit any region of positive measure in the domain. A useful

tool for studying orbits in dynamical systems is *symbolic dynamics*, which involves partitioning the domain into various (finitely-many) regions (partition elements), assigning to each region a number, and then encoding orbits as a sequence of numbers corresponding to the regions that they visit. However, the choice of partition elements is important - a 'poor' choice of partition can lead to a single point being coded in multiple ways, or many points being coded by the same sequence. In an ergodic system, we would expect typical orbits to visit each of these regions in their own unique way, and as such we would want our coding to reflect that.

In the case of Anosov diffeomorphisms, such as the Cat Map, a partitioning of the domain exists which allows us to encode μ-almost all points uniquely - a *Markov partition* [49]. Furthermore, this partitioning of phase space possesses the *Markov property*; the future coding of a particular orbit depends only upon the region it finds itself in currently, and not what has come before. Note that Markov partitions can exist for non-Anosov systems, although finding them can be difficult. We discuss Markov partitions in the next section.

2.8 Markov partitions

In this section we will define what we mean by a Markov partition, and find a Markov partition for the Cat Map. We begin by describing the structure of the partition elements.

Definition 2.16 (**[45], [1]**). *Let $H : M \to M$ be a diffeomorphism which possesses a hyperbolic invariant set Λ with a local product structure (in the case of an Anosov diffeomorphism, $\Lambda = M$). A non-empty set $R \subset \Lambda$ is a* **(proper) rectangle** *if*

(i) $R = \mathrm{Cl}(\mathrm{Int}(R))$, *and*

(ii) $x, y \in R$ *implies that* $(\gamma^s_x \cap \gamma^u_y) \cap R$ *is exactly one point.*

In other words, a set R is a rectangle if it is closed, and the stable manifold for any point in R intersects exactly once, within R, with the unstable manifold of any other point in R (in the case of non-Anosov systems, these may instead be stable and unstable *foliations*, see Section 3.3). We say a point $x \in R$ is a *boundary point* of R if there exists $y \in B(x, r)$ such that $y \notin R$ for any $r > 0$. We label the boundary of R, the set of all boundary points $x \in R$, as ∂R. Note that $\mathrm{Int}(R) = R \backslash \partial R$. We now define what it means for a collection of rectangles to be a Markov partition.

Definition 2.17 (**[45]**). *Let $H : M \to M$ be a diffeomorphism which possesses a hyperbolic invariant set Λ with a local product structure. A* **Markov partition** *for H is a finite collection of rectangles, $\mathcal{R} = \{R_i\}_{i=1}^m$, that satisfies the following conditions.*

(i) $\mathcal{R}$ covers Λ, that is, $\Lambda = \bigcup_{i=1}^m R_i$.

(ii) If $i \neq j$, then $\mathrm{Int}(R_i) \cap \mathrm{Int}(R_j) = \emptyset$.

(iii) If $x \in \mathrm{Int}(R_i)$ and $H(x) \in \mathrm{Int}(R_j)$, then

$$H(\gamma_x^u \cap R_i) \supset \gamma_{H(x)}^u \cap R_j, \text{ and}$$

$$H(\gamma_x^s \cap R_i) \subset \gamma_{H(x)}^s \cap R_j.$$

(iv) (Optional) If $x \in \mathrm{Int}(R_i) \cap H^{-1}(\mathrm{Int}(R_j))$, then

$$\mathrm{Int}(R_j) \cap H(\gamma_x^u \cap \mathrm{Int}(R_i)) = \gamma_{H(x)}^u \cap \mathrm{Int}(R_j), \text{ and}$$

$$\mathrm{Int}(R_i) \cap H^{-1}(\gamma_{H(x)}^s \cap \mathrm{Int}(R_j)) = \gamma_x^s \cap \mathrm{Int}(R_i).$$

Condition (iii) ensures that if the image of a rectangle R_i intersects the interior of another rectangle R_j, then it must stretch all the way across R_j in the unstable direction and be a subset in the stable direction. In other words, applying H stretches the partition elements in the unstable direction, mapping unstable boundaries onto unstable boundaries, and contracts the partition elements in the stable direction. Conversely, applying H^{-1} flips the roles of the unstable and stable directions, causing the images to stretch across fully in the stable direction, and be a subset in the unstable direction.

Condition (iv) ensures that the transition matrix (see Definition 2.19 below) is well defined, by preventing the images of rectangles from stretching over the same rectangle more than once - that is, to prevent two points which may undergo the same sequence of rectangles to be indistinguishable from each other. We can still obtain a Markov partition without this condition - however, we must instead consider an *adjacency matrix* which takes into account the number of times various partition elements wrap around each other upon iteration.

Note that Markov partitions may not be unique; in fact, different Markov partitions for a particular system may contain different numbers of elements. We refer to a partition with the minimum possible number of elements as a *minimal Markov partition*. The symbolic dynamics upon the Markov partition are that of a *subshift of finite type*.

Definition 2.18. *For each natural number $N \geq 2$, let*

$$\Omega_N = \{\omega = (\dots, \omega_{-1}, \omega_0, \omega_1, \dots) : \omega_i \in \{0, 1, \dots, N-1\} \text{ for } i \in \mathbb{Z}\}$$

be the space of two-sided sequences of N symbols, and let $\sigma_N : \Omega_N \to \Omega_N$ be the left shift in Ω_N, that is

$$\sigma_N(\omega) = \omega' = (\dots, \omega'_0, \omega'_1, \dots),$$

where $\omega'_n = \omega_{n+1}$.

Let $A = (a_{ij})_{i,j=0}^{N-1}$ *be an* $N \times N$ *matrix, where* $a_{ij} \in \{0, 1\}$ *for all* i, j*, and let*

$$\Omega_A = \{\omega \in \Omega_N : a_{\omega_n \omega_{n+1}} = 1 \text{ for } n \in \mathbb{Z}\}.$$

Then the restriction

$$\sigma_{N|_{\Omega_A}} =: \sigma_A$$

is called a **subshift of finite type**.

In other words, the matrix A tells us all admissible transitions between the symbols $0, 1, \ldots, N-1$, and the subshift of finite type is the shift transformation resulting from these transitions. In the case of a Markov partition, the N symbols each represent a partition element, and the matrix A is a *transition matrix* which tells which elements the image of any particular element intersects under iteration of the map H.

Definition 2.19. *Let* $\mathcal{R} = \{R_i\}_{i=1}^m$ *be a Markov partition of the map* H*, then its* **transition matrix** $A = (a_{ij})$ *is given by*

$$a_{ij} = \begin{cases} 1 & \text{if } \operatorname{Int}(H(R_i)) \cap \operatorname{Int}(R_j) \neq \emptyset, \\ 0 & \text{if } \operatorname{Int}(H(R_i)) \cap \operatorname{Int}(R_j) = \emptyset. \end{cases}$$

A *is an* **adjacency matrix** *if* (a_{ij}) *instead takes the value of the number of times* $H(R_i)$ *wraps around (stretches fully across)* R_j *in the unstable direction.*

If none of the eigenvalues of the transition (or adjacency) matrix are one, then the Markov partition is minimal. If not, then it may be possible to merge two or more of the partition elements or find an entirely new Markov partition which contains fewer elements.

We will now discuss how to construct a Markov partition for the Cat Map (or, in principle, any Anosov diffeomorphism on $\mathbb{T}^2$). We begin at a fixed point of H, $(0, 0)$ say, and from here we extend v_u until it intersects with v_s extended from $(1, 1)$, which

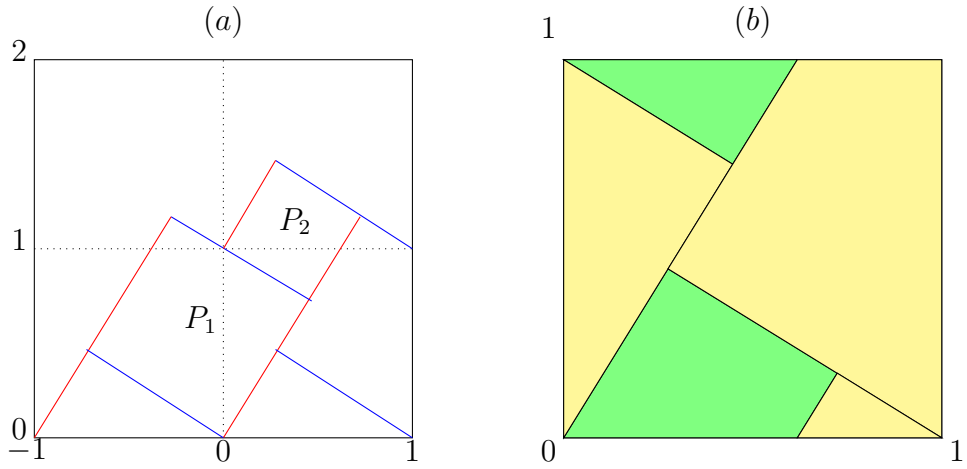


Figure 4: (a) Extending the eigenvectors of the Cat Map, v_u (red) and v_s (blue), to find appropriate Markov partition elements. (b) A two-element Markov partition for the Cat Map. The partition element P_1 is yellow, and P_2 is green.

we extend until it intersects with v_u extending from $(0, 1)$. We then extend v_s from both $(0, 1)$ and $(1, 0)$ until they intersect with v_u extended from $(0, 0)$, and v_s from $(0, 0)$ and $(0, 1)$ until they intersect with v_u extended from $(-1, 0)$. These give us all the intersections of v_u and v_s we need to form a two-element Markov partition for the Cat map. The intersections are shown in Figure $4(a)$, and the Markov partition is shown on $\mathbb{T}^2$ in Figure $4(b)$. We will refer to the Markov partition as P, and the partition elements as P_1 and P_2. Note that the boundaries of these partition elements, ∂P_1 and ∂P_2, are composed of two stable and two unstable manifolds each.

Now that we have found a Markov partition for H, we can find the adjacency matrix of this partition. To do this we need to calculate the images P_1 and P_2, and see how these intersect with P_1 and P_2. We can see the images of these partition elements in Figure $5(a)$. For this particular partition, the adjacency matrix is

$$P_{\text{adj}} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix},$$

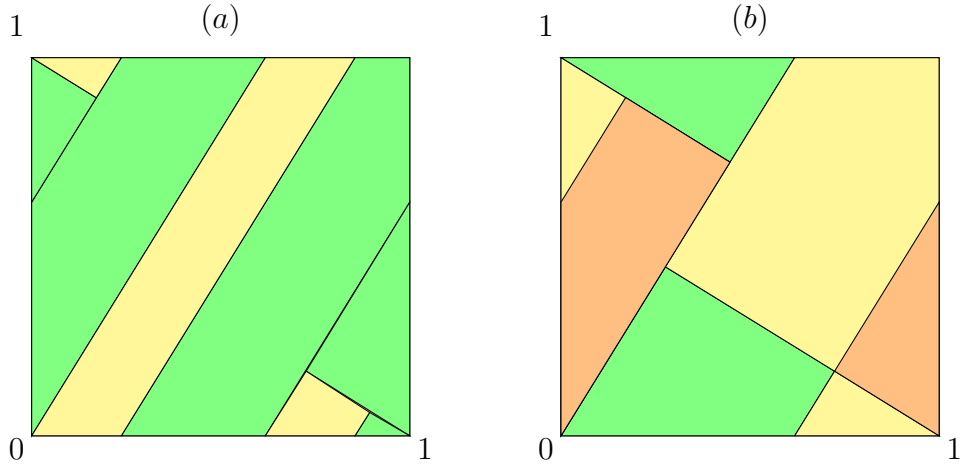


Figure 5: (a) The image of the Markov partition P of the Cat Map, where HP_1 (yellow) is the image of P_1 and HP_2 (green) is the image of P_2. (b) A Markov partition of the Cat Map with three partition elements.

since both images intersect P_1 and P_2, and HP_1 wraps around P_1 twice. Since $P_{\text{adj}} = DH_x^T$, its eigenvalues are $\lambda_{u,s}$. Neither of these are equal to one, which tells us that this Markov partition is minimal. An example of a Markov partition for the Cat Map with three elements is shown in Figure 5(b). We will discuss the Markov partition found by Mackay [34] for a particular system, originally studied by Cerbelli and Giona [14], which belongs to the class of maps known as *pseudo-Anosov* in the next chapter.

In the previous section, we discussed ergodicity, which encapsulates the idea of indecomposability of a dynamical system - that is, no non-trivial 'islands' (invariant sets of positive measure) exist in the domain which remain self-contained. A stronger notion than this is *mixing*; in a mixing system, the long term (i.e. infinite limit) state of the system is independent of its initial state. We discuss mixing in the next section.

2.9 Mixing and Decay of Correlations

To motivate the notion of mixing, we consider stirring milk into a cup of coffee. Ideally, as we stir, we would like the milk and coffee to distribute themselves evenly throughout the cup. This would ensure that any sip we were to take would contain the same proportions of milk and coffee as any other, no matter where we took the sip from - for example, we would obtain the same 'sip' whether we sipped the coffee from the top of the cup, or used a straw to sip from the bottom. Furthermore, we would want this distribution to depend only upon the proportions of milk and coffee found within the cup, and not on the original distribution before (or during) the stirring process; that is, no matter where we pour the milk into the cup of coffee, the stirring process should lead us to a state where the liquids in the cup are evenly distributed.

The notion of the limiting distribution being independent of the initial state of the system is clearly linked to the idea of ergodicity; if any pockets of liquid within the cup remain self-contained under the stirring process, then the sips taken from these pockets may differ from other regions. Consider our cup before we pour in the milk, and a stirring process which leaves a region within the cup invariant. If we do not pour any milk into this region, then due to invariance, none will ever enter it, and any sip we were to take from this region would contain only coffee and no milk.

Let us try to define these ideas in the context of a dynamical system (M, σ, H, μ). M is the region within the cup collectively encompassed by both the milk and the coffee, and H is the stirring process. Let the milk in the cup be given by the set A, then the proportion of milk in the cup is given by $\mu(A)/\mu(M)$. The amount of milk contained in a region B, the 'sip', after n applications of the stirring process is given by

$$\mu(H^n(A) \cap B),$$

and the proportion of milk in B is given by

$$\frac{\mu(H^n(A) \cap B)}{\mu(B)}.$$

Given enough applications of the stirring process, we would like for each sip to contain the same proportion of milk: the proportion of milk that is contained within the cup. Hence, we would like

$$\lim_{n\to\infty} \frac{\mu(H^n(A) \cap B)}{\mu(B)} - \frac{\mu(A)}{\mu(M)} = 0.$$

If we assume that μ is a probability measure ($\mu(M) = 1$), this leads us to the definition of what it means for a measure-preserving dynamical system to be mixing.

Definition 2.20. *A measure-preserving dynamical system is called* **mixing** *if for any two measurable sets A, B*

$$\lim_{n\to\infty} \mu(H^{-n}(A) \cap B) = \mu(A)\mu(B).$$

If H is invertible, then we can replace H^{-1} with H in the definition. Mixing implies ergodicity; let $H(A) = A$ and replace B with $M\backslash A$, then

$$\mu(H^{-n}(A) \cap (M\backslash A)) = 0$$

for every n, hence Definition 2.20 implies

$$\mu(A)\mu(M\backslash A) = 0.$$

Thus we have either $\mu(A) = 0$, or $\mu(M\backslash A) = 0$ (and thus $\mu(A) = 1$), and hence all H-invariant sets are measure zero or one — H is ergodic.

A map H of the form (17) is mixing. In fact, maps of this form possess an even stronger property, known as the *Bernoulli property*; specifically, this means they are isomorphic to a Bernoulli shift on the set of all sequences of symbols (in our case, the symbols could refer to the different elements of the Markov partition

of H). In practice, a system which possesses the Bernoulli property is statistically indistinguishable from a sequence of random coin or dice tosses.

Heuristically, any measurable $A \subset M$ gets stretched in the direction of the unstable subspace E^u and contracted in the direction of the stable subspace E^s under iteration by H. Repeated iteration yields the images $H^n(A)$, which on $\mathbb{R}^2$ are increasingly longer and thinner strips, with any original details of the set A (e.g. the structure of its boundary set ∂A) becoming less and less significant with each iterate. On $\mathbb{T}^2$, we see these strips wrapping around the domain in bands, with the number of wrappings increasing and the width of the bands decreasing as we iterate.

Formally, one must appeal to Pesin theory [39] to find a link between the stretching and contracting of the manifolds of H with ergodicity, and subsequently mixing. Loosely, Pesin theory tells us that the set of points with non-zero Lyapunov exponents form a finite or countable partitioning of ergodic components — regions upon which the map H, when restricted to a single component, is ergodic. We refer to a particular theorem to demonstrate the Bernoulli property (and therefore mixing) in H, but first we define what it means for a map H to be topologically mixing - a form of mixing which does not require the definition of a measure.

Definition 2.21. *An invertible, continuous map $H : M \to M$ is said to be* **topologically mixing** *if, given any pair of non-empty, open sets $A, B \subset M$, there exists an integer N, such that, for all $n > N$,*

$$H^n(A) \cap B \neq \emptyset.$$

Note that mixing and topological mixing are separate properties, and do not imply one another. We can see that H is topologically mixing by simply considering a line, from one side of A to another, in the direction of the unstable eigenvector. As H is iterated, this line will be stretched without altering its direction. Since the slope of the eigenvector is irrational, its images $H^n(A)$ will wrap around the torus

more and more with each iterate, never overlapping, and slowly filling the entire torus. Given enough iterates, $H^n(A)$ will eventually intersect B. We now refer to the following theorem to demonstrate the presence of the Bernoulli property for H.

Theorem 2.7 ([3]). *Let $H : M \to M$ be a measure-preserving Anosov diffeomorphism on a connected compact Riemannian manifold, and let H be topologically mixing. Then H has the Bernoulli property.*

After determining that a system is mixing, one may wish to quantify the rate at which mixing occurs. A typical way of quantifying the rate of mixing is by studying the rate of *decay of correlations of a scalar field*; this is done via the study of the *correlation function* $C_n(\phi, \psi)$, where ϕ and ψ are observable functions. We motivate this correlation function by first rewriting Definition 2.20 (assuming an invertible H) as

$$\lim_{n\to\infty} \int \chi_{H^n(A)\cap B} \mathrm{d}\mu = \int \chi_A \mathrm{d}\mu \cdot \int \chi_B \mathrm{d}\mu, \tag{25}$$

where χ_X is the characteristic function of the set $X \subseteq M$. A point x is in $H^n(A) \cap B$ if it is in both $H^n(A)$ and B, hence

$$\chi_{H^n(A)\cap B} = \chi_B \cdot \chi_{H^n(A)}.$$

Furthermore, since $\chi_{H^n(A)}$ consists of all points which map to A under H^{-n}, we have

$$\chi_{H^n(A)} = \chi_A \circ H^{-n}.$$

We can now rewrite (25) as

$$\lim_{n\to\infty} \int \chi_B(\chi_A \circ H^{-n}) \mathrm{d}\mu = \int \chi_A \mathrm{d}\mu \cdot \int \chi_B \mathrm{d}\mu. \tag{26}$$

In general, we can replace the characteristic functions with observable functions for the system, which yields the following definition.

Definition 2.22. *For observable functions $\phi, \psi \in L^2_M$, the* **correlation function** *is given by*

$$C_n(\psi, \phi) = \Big| \int \phi(\psi \circ H^{-n}) \mathrm{d}\mu - \int \phi \mathrm{d}\mu \cdot \int \psi \mathrm{d}\mu \Big|.$$

In applications one will often study $C_n(\phi, \phi)$, i.e. $\phi = \psi$, in order to study the rate of decay of correlations of a particular observable (scalar field) of interest - for example, temperature or pressure. If H is mixing, $C_n(\phi, \phi) \to 0$ as $n \to \infty$, and the rate of decay of C_n as n increases gives us an idea of how quickly the observable is being mixed. Specifically, one studies the exponential decay rate

$$\alpha_{\phi,\psi} = \sup\{a : \lim_{n\to\infty} \sup |e^{an} C_n(\phi, \psi)| < \infty\}. \tag{27}$$

The link between Lyapunov exponents and (exponential) decay of correlations has been studied for decades. Crawford and Cary [16] found that the decay rate $\alpha_{\phi,\psi}$ depends on the choice of observables for the Cat Map; for various observables, they found algebraic, exponential and faster than exponential decay rates. Slipantschuk *et al.* [50] showed that arbitrarily slow decay rates can be obtained by choosing non-regular observables. Ayyer and Stenlund [7] found explicit upper bounds upon the rate of decay of correlations in random products of hyperbolic toral automorphisms, which depend upon the maximal Lyapunov exponent. We will discuss a method for finding upper and lower bounds of such matrix products, and subsequently provide bounds on the maximal Lyapunov exponent which improve upon those given by Ayyer and Stenlund, in Chapter 4.

2.10 Summary

In this chapter, we have given the basic definitions we will need throughout this thesis, including that of the Lyapunov exponent of a dynamical system. We have discussed some fundamental properties of measure-preserving systems, including

uniform hyperbolicity, the existence of invariant cones and Markov partitions, ergodicity, and mixing. We have demonstrated that the Cat Map (and maps of the form (17)) possess these properties, and have shown a method for constructing Markov partitions for these maps. Furthermore, we have discussed a theorem which yields the invariant cones for diagonalizable matrices with positive eigenvalues, which we will use to find the invariant cones needed for our bounds upon the Lyapunov exponents of random products of matrices of this form.

3 Examples of non-uniform and non-Anosov systems

In this chapter we discuss several examples of non-uniform, non-Anosov systems upon the torus, with the intent of demonstrating why each system fails the conditions needed to be Anosov and/or uniformly hyperbolic. In addition to this, we will describe some examples of the different forms of hyperbolicity which arise when the strict conditions required for uniform hyperbolicity are loosened. We will also discuss random dynamical systems, focusing on those formed via compositions of shears, and the differences which arise between random and deterministic systems.

In the previous chapter we discussed and defined the notion of uniform hyperbolicity, with examples given by Arnold's Cat Map and more general Anosov diffeomorphisms upon the torus. These systems all have the property of being uniform upon the torus; that is, their Jacobian matrix DH has no spatial dependence. The deterministic systems we study in this chapter are all non-uniform, and so the Jacobian varies across the torus; furthermore, the systems will not be smooth, and thus their Jacobian matrix is not continuous across the entire domain.

Note that, to avoid confusion, the words uniform and non-uniform, when stated alone, are used to indicate the spatial dependence of the Jacobian matrix, while statements meant to regard the hyperbolicity of a system will always specify as such. For example, in Section 3.1, we will study a family of maps which are non-uniform, but uniformly hyperbolic; the systems themselves are not smooth (and hence are non-Anosov), but they do fulfil the requirements for Definition 2.7.

In Section 3.1, we study a generalisation of the Anosov diffeomorphisms studied in the previous chapter, and show that, in general, these maps are uniformly hyperbolic, but non-Anosov. In Section 3.2, we discuss the idea of non-uniform hyperbolicity, a looser form of hyperbolicity than uniform hyperbolicity, and a paradigmatic

example of such a system, the linked twist map. In Section 3.3, we discuss a generalization of Anosov diffeomorphisms, known as pseudo-Anosov maps, as well as discussing an interesting example of a (non-Anosov) pseudo-Anosov map, originally studied by Cerbelli and Giona [14]. Finally, in Section 3.6, we discuss random dynamical systems, and explain how (explicit) calculation of the Lyapunov exponent is complicated by randomness; in particular, we provide an example in the form of random products of toral Anosov diffeomorphisms, as studied in Chapter 2.

3.1 Non-uniform compositions of shears

In Sections 2.3 and 2.5 we discussed Arnold's Cat map and the family of Anosov systems to which it belongs. These maps are uniform upon $\mathbb{T}^2$; that is, the Jacobian matrix $D_x H$ of each map is identical for all choices of $x \in \mathbb{T}^2$. We now consider a generalisation of these systems, where uniformity across $\mathbb{T}^2$ is lost.

Let us first consider a simple example of such a system. Consider the system $(\mathbb{T}^2, \sigma, H, \mu)$ where σ is the Borel σ-algebra for $\mathbb{T}^2$, μ is Lebesgue measure, and $H : \mathbb{T}^2 \to \mathbb{T}^2$, given by $H = G \circ F$, where $F : \mathbb{T}^2 \to \mathbb{T}^2$ is given by

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = F_1 \begin{pmatrix} x \\ y \end{pmatrix}, & \text{for } y \leq \frac{1}{2}, \\ \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = F_2 \begin{pmatrix} x \\ y \end{pmatrix}, & \text{for } y > \frac{1}{2}, \end{cases}$$

and $G : \mathbb{T}^2 \to \mathbb{T}^2$ is given by

$$G\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

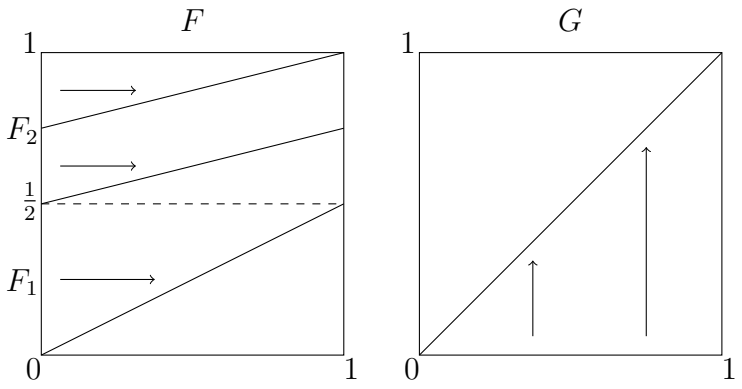


Figure 6: The shear maps F and G, which are composed to form the map H in (28).

Hence,

$$H\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = H_1\begin{pmatrix} x \\ y \end{pmatrix}, & \text{for } y \leq \frac{1}{2}, \\ \begin{pmatrix} 1 & 4 \\ 1 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = H_2\begin{pmatrix} x \\ y \end{pmatrix}, & \text{for } y > \frac{1}{2}. \end{cases} \tag{28}$$

The maps F and G are shown in Figure 6, and the action of H is shown in Figure 7. Note that this system is continuous, as $F_1 = F_2 \mod 1$ along $y = \frac{1}{2}$; specifically, the slopes of F_1 and F_2 are a multiple of $\frac{1}{y} = 2$.

The Jacobian matrix D_xH of H is dependent on the choice of $x \in \mathbb{T}^2$, and therefore so are the eigenvalues and eigenvectors of D_xH. When $y \leq \frac{1}{2}$, we have

$$\lambda_{H_1\pm} = 2 \pm \sqrt{3} \text{ and } v_{H_1\pm} = \begin{pmatrix} 2 \\ 1 \pm \sqrt{3} \end{pmatrix},$$

and when $y > \frac{1}{2}$, we have

$$\lambda_{H_2\pm} = 3 \pm 2\sqrt{2} \text{ and } v_{H_2\pm} = \begin{pmatrix} 1 \\ 1 \pm \sqrt{2} \end{pmatrix}.$$

We wish to determine whether or not H is uniformly hyperbolic. Intuitively, uniform hyperbolicity seems like a natural consequence of the systems H_1 and H_2

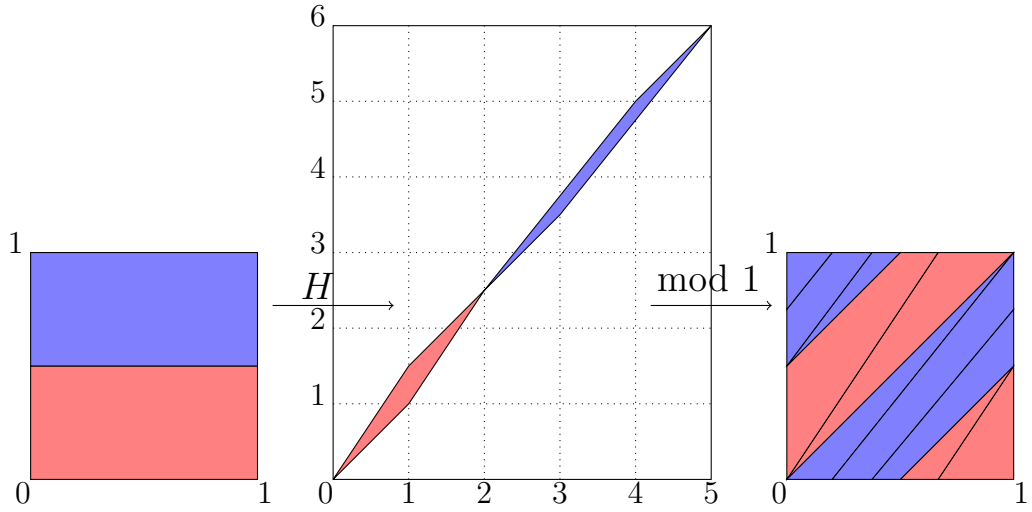


Figure 7: The action of the map H in (28). The region acted on by H_1 and its image are in red, while the region affected by H_2 and its image are in blue.

each being Anosov diffeomorphisms when extended to the entirety of $\mathbb{T}^2$; in particular, given this knowledge, it would seem (and in this case is true) that we could find uniform constants to bound the expansion and contractions rates away from one, perhaps by taking the minimum rates of expansion/contraction found in either maps unstable/stable subspace. However, the issue here is not these rates, but instead the subspaces themselves.

In order to determine if H is uniformly hyperbolic, we first need to find appropriate subspaces E^s_x and E^u_x such that

$$D_x H E^s_x = E^s_{H(x)},$$

and

$$D_x H E^u_x = E^u_{H(x)},$$

for any choice of $x \in \mathbb{T}^2$. It is worth noting that H is not a diffeomorphism, as clearly it is not differentiable on the lines $y = 0$ and $y = \frac{1}{2}$. However, this set of points

has measure 0, and so has little effect on the overall dynamics of the system [29]. As such, though H is not an Anosov diffeomorphism, we are considering whether it possesses all the qualities of such a system (i.e. uniform hyperbolicity), with the exception of differentiability for all $x \in \mathbb{T}^2$.

Consider the subspace ${E_1}^u_x$, the unstable subspace for the map $H_1 : \mathbb{T}^2 \to \mathbb{T}^2$, given by the span of the unstable eigenvector v_{H_1+} of $D_x H_1$. By the of definition an eigenvector we have that

$$D_x H_1 \cdot {E_1}^u_x = {E_1}^u_{H_1(x)}.$$

Now instead consider H. In this case, we instead require

$$D_x H_1 \cdot {E_1}^u_x = {E_1}^u_{H(x)},$$

for any choice of $x \in \mathbb{T}^2$. In particular, this means that if $H_1(x)$ lies in the region $y > \frac{1}{2}$, then we require

$$D_x H_1 \cdot {E_1}^u_x = {E_1}^u_{H_2(x)},$$

since $H = H_2$ for $y > \frac{1}{2}$. This is only possible if $H_1 = H_2$, and the problem is reduced to an Anosov diffeomorphism of the form studied in Section 2.5. However, if $H_1 \neq H_2$, we can find neither an unstable subspace nor, through a similar argument, a stable subspace of this form, which will satisfy the definition of uniform hyperbolicity for H. A sketch of this argument is shown in Figure 8.

A further result of the above argument is that constructing a Markov partition for H is not possible, since it is not possible to map boundaries of partition elements onto boundaries of other partition elements. This can make the calculation of a return-time distribution much trickier for maps of this form; we will see this in Chapter 5, where the calculation of such a distribution will be needed to evaluate upper and lower bounds for the Lyapunov exponents of H.

However, despite the above, it is possible to find subspaces which do produce a (λ, μ) splitting of $\mathbb{R}^2$ for H (for some λ, μ). To do this, we use invariant cones,

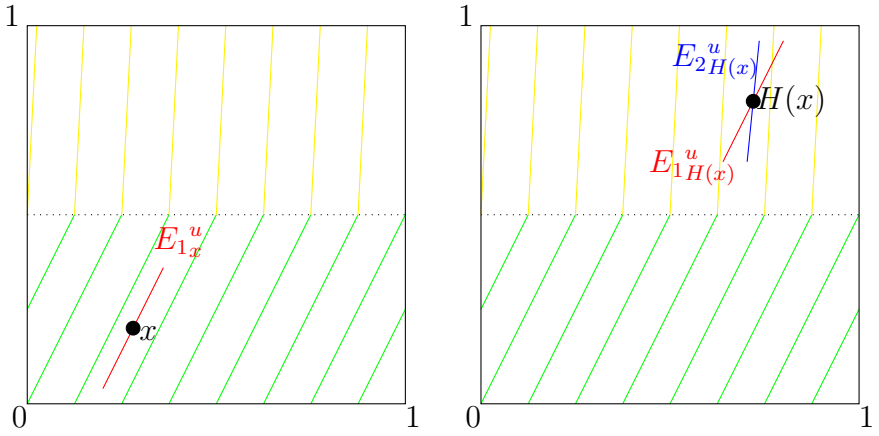


Figure 8: This sketch shows how we cannot use eigenvectors alone to satisfy the definition of uniform hyperbolicity for a map such as H in (28). The green lines represent $E_{1_x}^u$ and the yellow lines represent $E_{2_x}^u$ for other choices of x in the respective regions.

which we discussed in Section 2.6; specifically, we find two cones, an expansion and a contraction cone, which are mutually invariant under H_1 and H_2. We study such cones in detail in Chapter 4, where we discuss their form and for which systems they exist.

For H, an expansion cone, corresponding to the unstable subspace E_x^u, is given by

$$C_E = \{(x, y) \in \mathbb{R}^2 : \frac{1}{1+\sqrt{2}} \leq \frac{x}{y} \leq \frac{2}{1+\sqrt{3}}\},$$

and a contraction cone, corresponding to the stable subspace E_x^s, is given by

$$C_C = \{(x, y) \in \mathbb{R}^2 : \frac{2}{1-\sqrt{3}} \leq \frac{x}{y} \leq \frac{1}{1-\sqrt{2}}\}.$$

Note that the cone C_E has boundaries given by the unstable eigenvectors of H_1 and H_2, whilst the boundaries of C_C are given by their stable eigenvectors. Furthermore, for almost every orbit, all vectors $v \in \mathbb{R}^2$ satisfy

$$H^N v \in C_E \text{ for all } N > n \text{ for some } n \in \mathbb{N}, \tag{29}$$

since the stable eigenvectors v_{H_1-} and v_{H_2-} are not invariant under H. The remaining orbits consist of fixed and periodic points; in particular, the Jacobian of H is undefined at the fixed points, and so (29) cannot apply in these cases.

In order to verify uniform hyperbolicity of H we require a (λ, μ)-splitting, for which we need to find constants λ and μ such that

$$||H_{|_{C_C}}|| \leq \lambda < 1, \ ||{H^{-1}}_{|_{C_E}}|| \leq \mu^{-1} < 1.$$

In this case, using the spectral norm [52] given by

$$\|H\| = \frac{\|Hv\|_2}{\|v\|_2},$$

suitable constants λ and μ are given by

$$\lambda = \max_{v \in C_C} \left\{ \frac{\|H_1 v\|_2}{\|v\|_2}, \frac{\|H_2 v\|_2}{\|v\|_2} \right\} = 0.8931,$$

$$\mu = \min_{v \in C_E} \left\{ \frac{\|H_1 v\|_2}{\|v\|_2}, \frac{\|H_2 v\|_2}{\|v\|_2} \right\} = 3.6655.$$

We therefore have that $\lambda < 1 < \mu$, and so H is uniformly hyperbolic. Note that when calculating μ we are able to consider forward iterates of the map as opposed to pre-images as H is invertible.

We can generalise systems such as these to allow for multiple shears in both the horizontal and vertical directions. Consider partitioning the interval $[0, 1)$ into the intervals $P_i = [y_{i-1}, y_i)$ for $i = 1, \ . \ . \ . \ , n$, where $y_0 = 0$ and $y_n = 1$. The length - or height for reasons that will soon become clear - of the interval P_i is therefore given by

$$h_i = y_i - y_{i-1}.$$

We assign to each P_i a wrapping number α_i, which determines the number of times our shear map on P_i will wrap around $\mathbb{T}^2$. We require that $\frac{\alpha_i}{h_i} \in \mathbb{Z}$ so that our Jacobian matrices contain only integer values, and to ensure continuity. One can define continuous systems similar to these in which the Jacobian matrices do not

contain integer values; in this case, the torus is partitioned into annuli, and the maps F and G take the form of Dehn twists ([18], see also [32]) on these annuli (an explicit example of this is studied in Section 5.6).

We define $F_i : [0,1) \times P_i \to [0,1) \times P_i$ by

$$F_i \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & \frac{\alpha_i}{h_i} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{30}$$

and $F : \mathbb{T}^2 \to \mathbb{T}^2$ as the collection of all of the F_i. In other words, $F = F_i$ when F is restricted to $[0,1) \times P_i$. We define $G : \mathbb{T}^2 \to \mathbb{T}^2$ in a similar way, instead partitioning the interval $(0,1]$ into the intervals $Q_j = [x_{j-1}, x_j)$, for $j = 1, \ldots, m$, where $x_0 = 0$ and $x_m = 1$. Each Q_j has length (or width)

$$w_j = x_j - x_{j-1},$$

and is assigned a wrapping number β_j, where $\frac{\beta_j}{w_j} \in \mathbb{Z}$. We define $G_j : Q_j \times [0,1) \to Q_j \times [0,1)$ as

$$G_j \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{\beta_j}{w_j} & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{31}$$

$G : \mathbb{T}^2 \to \mathbb{T}^2$ is then defined as the collection of all the G_j, so $G = G_j$ where G is restricted to $Q_j \times [0,1)$. Essentially, F is a collection of n horizontal shears of varying slopes, laid on top of each other in such a way that they are continuous. G is similar, however we instead have m vertical shears placed side by side. Figure 9 shows a sketch of the maps F and G.

We consider the system $(\mathbb{T}^2, \sigma, H, \mu)$, where $H = G \circ F$. In other words, we have

$$H_{i,j} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & \frac{\alpha_i}{h_i} \\ \frac{\beta_j}{w_j} & 1 + \frac{\alpha_i \beta_j}{h_i w_j} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \text{ for } y \in P_i \text{ and } x + \frac{\alpha_i}{h_i} y \in Q_j. \tag{32}$$

Note that there are now $m + n$ lines upon which H is not differentiable; however, these still amount to a set of measure zero, and so again do not affect the dynamics of the system as a whole.

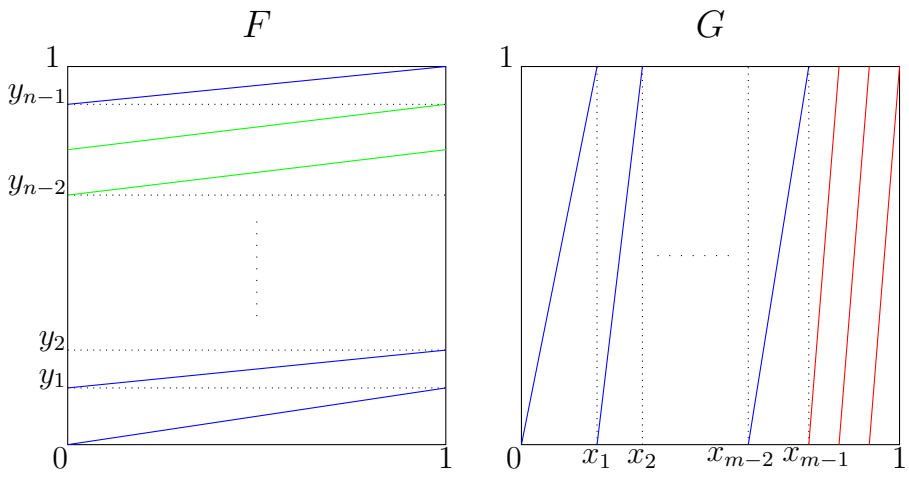


Figure 9: An example of the maps F and G from equations (30) and (31) respectively, with positive wrapping numbers α_i and β_j for all i and j. P_i is the region between $y = y_{i-1}$ and $y = y_i$, and Q_j is the region between $x = x_{j-1}$ and $x = x_j$.

We will consider the cases where all the α_i are of the same sign, and all the β_j are of the same sign. When this is not the case, the analysis is much harder. A particular example where much progress has been made is the case where $n = 2$, $m = 1$, $\alpha_1 = 2$, $\alpha_2 = -2$ and $\beta_1 = 1$, which was studied by Cerbelli and Giona [14], and later Mackay [34]; we will discuss this example in detail in Section 3.3.

Assuming all α_i share the same sign and all β_j share the same sign, we have two main cases to consider:

1. $\alpha_i, \beta_j > 0$ for all i, j.

2. $\alpha_i < 0, \beta_j > 0$, and $\alpha_i \beta_j < -4$ for all i, j.

Note that the cases where we have $\alpha_i, \beta_j < 0$ and $\alpha_i > 0, \beta_j < 0$ are equivalent to the above cases up to rotation. The requirement of $\alpha_i \beta_j < -4$ ensures hyperbolicity (though not necessarily uniform hyperbolicity) of H in the second case.

As earlier, we require both stable and unstable subspaces, with uniform bounds

upon the expansion and contraction rates of vectors therein, in order to infer uniform hyperbolicity of H. We again turn to mutually invariant cones in an attempt to yield such subspaces, and we construct these cones in the same way as we did before, by using the eigenvectors of the various Jacobian matrices of H as boundaries for the cones. We require such a cone to be invariant for any Jacobian we may choose, and so to obtain an invariant cone for H, we pick the widest possible cone that can be obtained from these eigenvectors.

However, an issue arises for general systems of the form given by (32), in that mutually invariant cones need not always exist; we will study this problem in more detail in Chapter 4, where we provide an explicit example of a random dynamical system for which invariant cones cannot be constructed (see Figure 22b). The cases for which invariant cones do exist are those in which none of the stable eigenvectors of H_{ij} lie amongst the unstable eigenvectors of H_{ij}, for all i, j. In essence, we require a 'neat' partitioning of tangent space, where all the stable eigenvectors are located in one region, all of the unstable eigenvectors are located in another region, and there is no overlap between the two.

With the above in mind, we find that the cases where $\alpha_i, \beta_j > 0$ for all i, j (and similarly, $\alpha_i, \beta_j < 0$ for all i, j) do possess invariant cones, as all unstable (stable) eigenvectors are located in the region $\{(x, y) \in \mathbb{R}^2 : \frac{x}{y} > 0\}$, and all stable (unstable) eigenvectors are located in the region $\{(x, y) \in \mathbb{R}^2 : \frac{x}{y} < 0\}$.

The cases where $\alpha_i \beta_j < -4$ are not as simple, and typically need to be studied on a case by case basis to determine whether invariant cones exist, and thus if the system is uniformly hyperbolic. An example of such a system where invariant cones do exist occurs when $\alpha_i = \alpha < 0$ for all i, with β_j such that $\alpha\beta_j < -4$ for all j. This results in a partitioning of the quadrant $\{(x, y) \in \mathbb{R}^2 : \frac{x}{y} > 0\}$, with all unstable eigenvectors located between the vectors $(0, 1)$ and $(-\alpha, 2)$, and all stable eigenvectors located between $(-\alpha, 2)$ and $(1, 0)$.

It is fairly simply to establish numerically whether invariant cones exist for a particular choice of matrices. This particular scheme only checks whether an invariant cone exists for a pair of matrices, but could be easily expanded upon to allow for more matrices. Once invariant cones have been obtained, it is a matter of checking the constants λ and μ to determine if the system is uniformly hyperbolic. Again labelling the cones which yield our stable and unstable subspaces as C_C and C_E respectively, and noting that H is invertible, we check if

$$\lambda = \max_{v \in C_C} \left\{ \frac{\|H_{1,1}v\|_2}{\|v\|_2}, \ldots, \frac{\|H_{n,m}v\|_2}{\|v\|_2} \right\} < 1,$$

and if

$$\lambda = \min_{v \in C_E} \left\{ \frac{\|H_{1,1}v\|_2}{\|v\|_2}, \ldots, \frac{\|H_{n,m}v\|_2}{\|v\|_2} \right\} > 1.$$

If these conditions hold, then H is uniformly hyperbolic. If not, H may possess a less strict form of hyperbolicity, known as non-uniform hyperbolicity, which we discuss in the next section.

3.2 Linked Twist Maps and non-uniform hyperbolicity

We now discuss toral linked twist maps (LTMs), which are a realization of the idea of intersecting flows in various models of fluid dynamics; in particular, linked twist maps are connected to the notion of streamlines crossing, examples of which include the Aref blinking vortex flow [4], which has been used to model tidal advection along a headland (see [55] and references therein). We will see that these systems differ from those studied in the previous section, as their Jacobian matrix is not hyperbolic in certain regions of the domain, and orbits can become trapped in non-hyperbolic regions for long periods of time. As in the previous section, we begin by studying a specific, simple example of a LTM, before moving on to the general case.

Consider the map $H : P \cup Q \to P \cup Q$ given by $H = G \circ F$, where $F : P \to P$

is given by

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \tag{33}$$

and $G : Q \to Q$ is given by

$$G\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \tag{34}$$

where $P = \{(x, y) \in \mathbb{T}^2 : y \leq \frac{1}{2}\}$ and $Q = \{(x, y) \in \mathbb{T}^2 : x \leq \frac{1}{2}\}$ are annuli. Let

$$A = \big\{(x, y) \in \mathbb{T}^2 : y \leq \frac{1}{2}, x + 2y \leq \frac{1}{2}\big\},$$
$$B = \big\{(x, y) \in \mathbb{T}^2 : y \leq \frac{1}{2}, x + 2y \geq \frac{1}{2}\big\},$$
$$C = \big\{(x, y) \in \mathbb{T}^2 : y \geq \frac{1}{2}, x \leq \frac{1}{2}\big\},$$

then we can write H as

$$H\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} & \text{for } (x, y) \in A, \\ \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} & \text{for } (x, y) \in B, \\ \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} & \text{for } (x, y) \in C. \end{cases} \tag{35}$$

In other words, we first apply a horizontal shear with a slope of 2 onto the annulus P, and then a vertical shear with a slope of 2 onto the annulus Q. The map H is a linked twist map. Note that while we discuss this map, instead of the usual Lebesgue measure μ_L, we will instead consider a normalized measure over the annuli P and Q, given by

$$\mu_H(Z) = \frac{4}{3}\int_{y\in Z}\int_{x\in Z} H(x, y)dxdy = \frac{4}{3}\mu_L(Z). \tag{36}$$

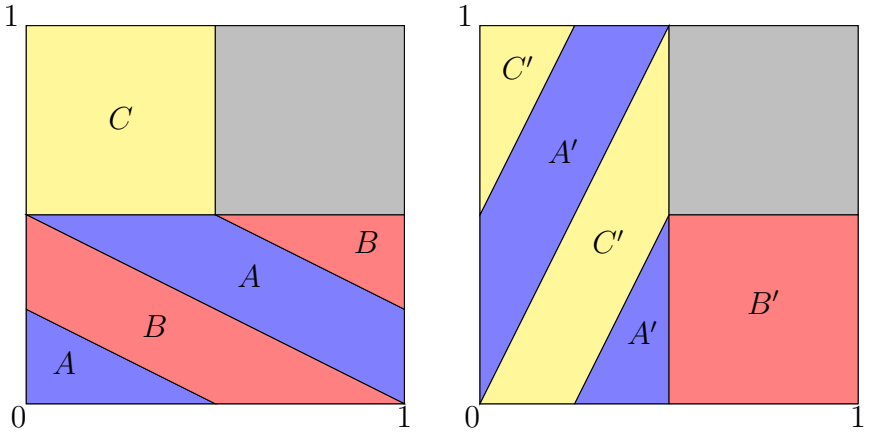


Figure 10: The regions A, B and C and their images under H: A', B' and C' respectively. Note that B' intersects B, and C' intersects both B and C. This tells us that it is possible to have multiple iterates in a row without landing in the hyperbolic region A.

From the definition of H, we can immediately see that its Jacobian matrix D_xH is not hyperbolic when $(x, y) \in B$ or C, as in both cases H is a shear map. However, D_xH is hyperbolic for $(x, y) \in A$, and the iterates for which this is the case are the iterates when vectors in tangent space will undergo exponential expansion. In other words, the positivity of the Lyapunov exponent of this map will derive entirely from the points on the orbit which lie in A.

A, B, and C, as well as their images under H, are shown in Figure 10. We can see from the figure that B' intersects B and C' intersects both B and C. This means it is possible for orbits to spend consecutive iterates in the non-hyperbolic region of the map; when this happens, vectors in tangent space will undergo sub-exponential expansion for two consecutive iterates. In fact, if we sketch $H^n(B)$ and $H^n(C)$ for a given n, we find that they both intersect the non-hyperbolic region of the map, and the intersection has positive measure.

In order to find the set of points which remain in the shear regions for at least 2 iterates, we need to find

1. $H^{-1}(H(B) \cap B)$.

2. $H^{-1}(H(B) \cap C)$.

3. $H^{-1}(H(C) \cap B)$.

4. $H^{-1}(H(C) \cap C)$.

The union of all of these sets is the set of points which do not undergo exponential expansion for at least two iterates. In this case, we know that $\mu_H(H^{-1}(H(B) \cap C)) = 0$ since $\mu_H(H(B) \cap C) = 0$, and H is measure-preserving. If we find $H(H(B) \cap B)$, $H(H(B) \cap C)$, $H(H(C) \cap B)$ and $H(H(C) \cap C)$ and see how they intersect B and C, then by calculating the 2nd pre-image H^{-2} of these intersections we can find the set of points which do not enter the hyperbolic region A for at least three iterates of H. Repeating this process for n steps requires calculating the n^{th} pre-image of 2^{n+1} sets. The measure of the union of the pre-images of these sets is the proportion of points which do not enter the hyperbolic region A for at least n iterates. The sets we find are defined inductively by

$$B_n = H(B_{n-1}) \cap \Gamma_n, \tag{37}$$

for $n \in \mathbb{N}$, $\Gamma_i = B$ or C, and $B_0 = \{B\}$, and

$$C_n = H(C_{n-1}) \cap \Gamma_n, \tag{38}$$

for $n \in \mathbb{N}$, $\Gamma_i = B$ or C, and $C_0 = \{C\}$. Note that both (37) and (38) describe 2^n different sets. Let

$$W_n = H^{-n}(B_n) \cup H^{-n}(C_n). \tag{39}$$

Then $\mu_H(W_n)$ is the proportion of points which do not enter A for at least n iterates. We can find a W_n of positive measure for any $n \in \mathbb{N}$, though $\mu_H(W_n)$ will decrease as n increases. Figure 11 shows the sets W_1 and W_2.

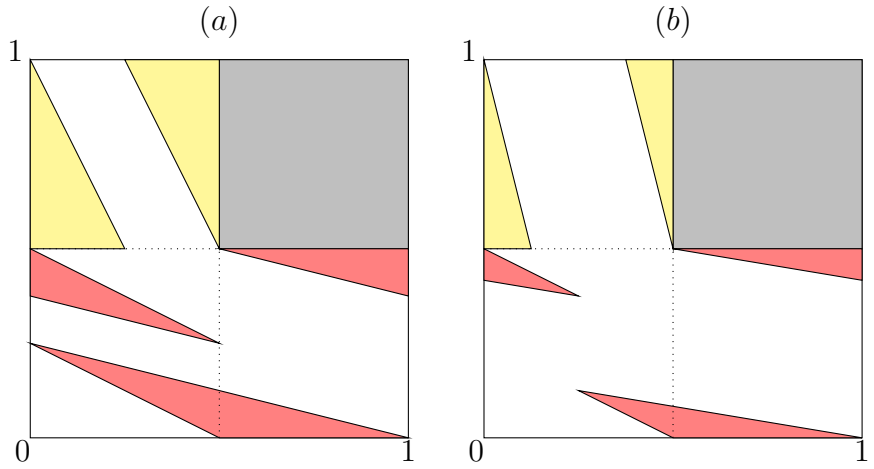


Figure 11: (*a*) W_1 and (*b*) W_2, the sets of points which spend at least 1 and 2 consecutive iterates in non-hyperbolic regions respectively, for the map H given by (32). The red region is $H^{-i}(B_i)$ and the yellow region is $H^{-i}(C_i)$, for $i = 1, 2$.

The fact that W_n has positive measure for any n affects the calculation of Lyapunov exponents of this map; specifically, finite-time Lyapunov exponents (FTLEs) can vary significantly depending on the choice of initial condition. If, by chance, we chose a point in W_{10^5}, then the iterates would remain entirely in non-hyperbolic regions[4] for 10^5 iterates; a FTLE calculated over 10^5 iterates of H would be zero, as tangent vectors would have undergone sub-exponential expansion on each iterate. Similarly, choosing a point which begins or whose orbit intersects a W_m, where m is a significant fraction of the 10^5 iterates, will also affect the FTLE. In contrast, if we performed a similar calculation for Arnold's Cat Map, we would find that our choice of initial condition $x_0 \in \mathbb{T}^2$ does not affect our estimate at all, since the Jacobian matrix is uniform on $\mathbb{T}^2$.

The spatial dependence of FTLEs in LTMs is a problem, since seemingly sim-

[4]This assumes our computer can calculate to a sufficient degree of accuracy such that any error in calculation does not cause the orbit to intersect a hyperbolic region.

ilar, nearby, randomly chosen initial conditions can return significantly different estimations for λ, with possible estimates of zero in rare cases; however, due to the arguments put forth by Burton and Easton [12], we know that LTMs are ergodic and have non-zero Lyapunov exponents almost everywhere, precluding the usefulness of such an estimate.

From the above argument we can conclude that H fails the conditions required for uniform hyperbolicity; we cannot find constants $\lambda < 1$ and $\mu < 1$ for which H is uniformly hyperbolic by Definition 2.7. Instead, suitable bounds upon the expansion and contraction rates will depend upon x; in the case of a LTM, for the iterates spent in the shear regions, $\lambda = \mu = 1$, whereas for those spent in the hyperbolic regions, $\lambda < 1 < \mu$, as in the uniformly hyperbolic case. Systems with this property, which is a relaxation of the conditions required for uniform hyperbolicity, are called non-uniformly hyperbolic.

Definition 3.1 ([8])**.** *Let $H : M \to M$ be a diffeomorphism of a compact smooth Riemannian manifold M. An H-invariant Borel subset $\mathcal{R} \subset M$ is said to be* **non-uniformly hyperbolic** *if there exist:*

(a) numbers λ and μ such that $0 < \lambda < 1 < \mu$;

(b) a number ϵ and Borel functions C, $K : \mathcal{R} \to (0, \infty)$;

(c) subspaces $E^s(x)$ and $E^u(x)$ for each $x \in \mathcal{R}$,

which satisfy the following conditions:

1. *the subspaces $E^s(x)$ and $E^u(x)$ depend measurably on x and form an invariant splitting of the tangent space, i.e.,*

$$T_xM = E^s(x) \oplus E^u(x),$$

$$D_xHE^s(x) = E^s(H(x)),\ D_xHE^u(x) = E^u(H(x));$$

2. *for* $v \in E^s(x)$ *and* $n > 0$,

$$\|D_x H^n v\| \leq C(x)\lambda^n e^{\epsilon n}\|v\|;$$

3. *for* $v \in E^u(x)$ *and* $n < 0$,

$$\|D_x H^n v\| \leq C(x)\mu^n e^{\epsilon|n|}\|v\|;$$

4. $\angle(E^s(x), E^u(x)) \geq K(x)$;

5. *for* $n \in \mathbb{Z}$,

$$C(H^n(x)) \leq C(x)e^{\epsilon|n|},\ K(H^n(x)) \geq K(x)e^{-\epsilon|n|}.$$

Non-uniformly hyperbolic systems differ from uniformly hyperbolic systems (and Anosov diffeomorphisms) due to the inability to guarantee exponential expansion (or contraction) of vectors within a subspace of tangent space. In the case of the linked twist map H given by (35), almost every orbit will eventually fall into the shear regions B and C of the map, and during these iterates exponential expansion will not occur. This is due to the fact that H is ergodic, and so a typical orbit will visit every region (of positive measure) within the domain; furthermore, all LTMs of the general form we discuss next are also ergodic (see [12], [44]). In fact, LTMs possess even stronger properties than this, such as mixing and the K-property (see [59]).

Note that we can still obtain invariant expanding and contracting subspaces for H, again by finding invariant cones, although we cannot guarantee exponential expansion or contraction of vectors within those subspaces on *all* iterates. The subspaces are required to be invariant under the shear maps F and G, since it is possible for H to take either as its Jacobian on a single iterate. We study the cones which are invariant for LTMs such as H in Chapter 5; note that these cones are also used by Sturman and Thiffeault [54] when finding bounds upon the Lyapunov exponents for random products of shear matrices.

In order to define a LTM in general, we first define annuli P_i and Q_j, where $i \in \{1, \ldots, n\}$, and $j \in \{1, \ldots, m\}$, given by

$$P_i = \{(x, y) \in \mathbb{T}^2 : y_{2(i-1)} \leq y \leq y_{2i-1}\},$$

$$Q_j = \{(x, y) \in \mathbb{T}^2 : x_{2(j-1)} \leq x \leq x_{2j-1}\},$$

where $0 \leq x_0 < \ldots < x_{2m-1} \leq 1$, and $0 \leq y_0 < \ldots < y_{2n-1} \leq 1$. In other words, the sets P_i are mutually disjoint horizontal annuli, ordered so that P_1 is below P_2, which is below P_3, and so on. Similarly, the sets Q_j are mutually disjoint vertical annuli, ordered so that Q_1 is the furthest to the left, and Q_m the furthest to the right, of $\mathbb{T}^2$.

On each P_i we define a horizontal shear with wrapping number α_i and, similarly, on each Q_j we define a vertical shear with wrapping number β_j. Specifically, letting $h_i = y_{2i-1} - y_{2(i-1)}$ be the height of P_i and $w_j = x_{2i-1} - x_{2(i-1)}$ be the width of Q_j, we define horizontal shears $F_i : P_i \to P_i$ as

$$F_i \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & \frac{\alpha_i}{h_i} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ y_{2(i-1)} \end{pmatrix},$$

and vertical shears $G_j : Q_j \to Q_j$ as

$$G_j \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{\beta_j}{w_j} & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x_{2(j-1)} \\ 0 \end{pmatrix}.$$

We define the maps $F : \mathbb{T}^2 \to \mathbb{T}^2$ and $G : \mathbb{T}^2 \to \mathbb{T}^2$ as follows:

$$F \begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} F_i \begin{pmatrix} x \\ y \end{pmatrix} & \text{for } (x, y) \in P_i, \\ I_2 \begin{pmatrix} x \\ y \end{pmatrix} & \text{otherwise,} \end{cases} \tag{40}$$

$$G\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} G_j \begin{pmatrix} x \\ y \end{pmatrix} & \text{for } (x,y) \in Q_j, \\ I_2 \begin{pmatrix} x \\ y \end{pmatrix} & \text{otherwise,} \end{cases} \tag{41}$$

where I_2 is the 2×2 identity matrix. The maps F and G are shown in Figure 12. The restriction of the map $H : \mathbb{T}^2 \to \mathbb{T}^2$ to $(\bigcup_i P_i) \cup (\bigcup_j Q_j)$, given by $H = G \circ F$, is a linked twist map; the example we studied earlier, given by (35), is the case where $n = m = 1$, $x_0 = y_0 = 0$, $x_1 = y_1 = \frac{1}{2}$ and $\alpha_1 = \beta_1 = 1$. We can write H as

$$H\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} H_{i,j} \begin{pmatrix} x \\ y \end{pmatrix} & \text{for } (x,y) \in P_i, (x + \frac{\alpha_i}{h_i} y, y + y_{2(i-1)}) \in Q_j, \\ F_i \begin{pmatrix} x \\ y \end{pmatrix} & \text{for } (x,y) \in P_i, (x + \frac{\alpha_i}{h_i} y, y + y_{2(i-1)}) \notin Q_j, \\ G_j \begin{pmatrix} x \\ y \end{pmatrix} & \text{for } (x,y) \notin P_i, (x,y) \in Q_j, \\ I_2 \begin{pmatrix} x \\ y \end{pmatrix} & \text{otherwise,} \end{cases} \tag{42}$$

where

$$H_{i,j}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & \frac{\alpha_i}{h_i} \\ \frac{\beta_j}{w_j} & 1 + \frac{\alpha_i \beta_j}{h_i w_j} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x_{2(j-1)} \\ y_{2(i-1)} \end{pmatrix}.$$

Note that in order to guarantee hyperbolicity in the regions analogous to the region A for the map given by (35), we require $\alpha_i \beta_j > 0$ or $\alpha_i \beta_j < -4$ for all i, j.

The family of maps described by (42) shares many of the same properties as the earlier example, such as ergodicity and non-uniform hyperbolicity. Similarly to the system given by (32), obtaining an invariant cone for H involves taking the widest cone for all possible combinations of α_i, β_j; note that in the case where $\alpha_i, \beta_j > 0$ for

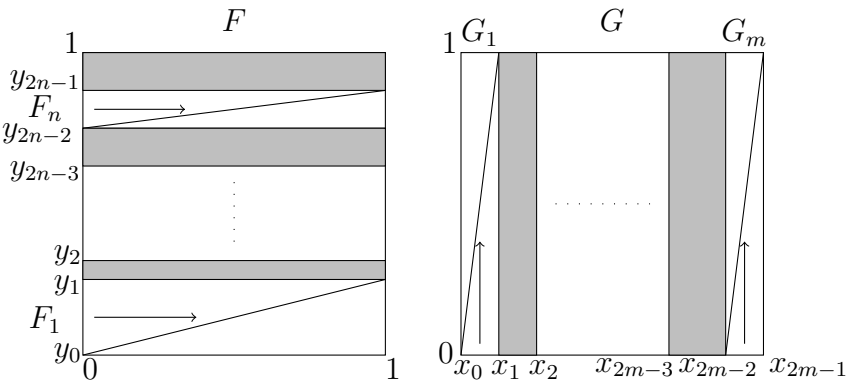


Figure 12: Sketches of the shear maps F and G from (40) and (41), which are composed to form the general LTM H. The grey shaded areas are the regions upon which F and G are the identity.

all i, j, the region $\{(x, y) \in \mathbb{R}^2 : x/y \geq 0\}$ is an invariant cone for H. In Chapter 5 we consider alternative cones to this which, while not invariant under H, are invariant under the return map to the region where the annuli overlap. Doing so allows us to find elementary bounds on the Lyapunov exponents of a sub-family of the systems given by (42); specifically, we study the systems obtained by taking $x_0 = y_0 = 0$, $x_1 = y_2 = \alpha_1 = \alpha_2 = \beta_1 = 1$, and $y_1 = p$ for some $p \in [0, 1]$. The method we describe is also applicable (in principle) to (42) as well.

3.3 Cerbelli and Giona's 'archetype for non-uniform chaos' and pseudo-Anosov maps

In this section we discuss a class of maps known as pseudo-Anosov [57], and focus on a specific example, formed via shear composition, which has been extensively studied by Cerbelli and Giona [14] and later Mackay [34]. The system is referred to by the authors as an 'archetype for non-uniform chaos' - specifically, it captures aspects of the dynamics of two important classes of map: hyperbolic toral automorphisms such as the Cat Map, which are represented by expanding and contracting direc-

tions in tangent space, and generalized baker transformations, represented by the discontinuity within the system, which can be used to model deformations caused by non-linear dynamics.

Pseudo-Anosov maps loosen the strict conditions required for a system H to be uniformly hyperbolic by replacing the need for smooth contracting and expanding subspaces in tangent space, E^s_x and E^u_x, with stable and unstable foliations upon the domain, F^s and F^u, which are differentiable everywhere except for a finite number of singularities. These singularities typically occur at critical points in the domain where foliation leaves of varying slopes meet. In addition, we require that vectors upon these foliations are stretched (on F^u) or contracted (on F^s) by some 'stretch factor'.

In order to define a pseudo-Anosov map explicitly, we first define what we mean by a measured foliation upon M.

Definition 3.2 ([57], see also [22])**.** *A measured foliation F on a closed surface M is a geometric structure on M which consists of a singular foliation and a measure in the transverse direction. In some neighbourhood U of a regular point of F, there is a 'flow box' $\phi : U \to \mathbb{R}^2$ which sends the leaves of F to the horizontal lines in $\mathbb{R}^2$. If two such neighbourhoods U_i and U_j overlap then there is a transition function ϕ_{ij} defined on $\phi_j(U_j)$, with the standard property*

$$\phi_{ij} \circ \phi_j = \phi_i$$

which must have the form

$$\phi(x, y) = (H(x, y), c \pm y)$$

for some constant c. A finite number of singularities of F of the type of a 'p-pronged saddle', $p \geq 3$, are allowed (see Figure 15).

In other words, the foliation leaves stretch across the entire domain, but need not be smooth at all points; singularities will arise where boundaries of 'non-differentiability' meet. In the system we will study, these boundaries are the lines along which the foliations undergo a sudden change in gradient. If two of the neighbourhoods described above overlap, there must be a one to one mapping between the foliation leaves which lie in both neighbourhoods; that is, leaves do not 'disappear' or end in regions away from singularities.

We now define what it means for a map to be pseudo-Anosov.

Definition 3.3 ([37],[57])**.** *A mapping class $\mathcal{H}$ on a surface M of negative Euler characteristic is said to be* **pseudo-Anosov** *if there is a pair F^s and F^u of transverse arational (i.e. no closed leaves) measured foliations on M and a representative H of $\mathcal{H}$ so that $H(F^u) = \lambda F^u$ and $H(F^s) = \frac{1}{\lambda}F^s$ with $\lambda > 1$. λ is called the* **dilatation** *(or* **stretch factor***) of $\mathcal{H}$.*

From the definition, it is fairly simple to see that the Cat Map is pseudo-Anosov; taking the eigenvectors for the foliations, we see that the dilatation factor corresponds to the largest eigenvalue. In fact, pseudo-Anosov maps are a generalization of Anosov diffeomorphisms, and as such any Anosov (or uniformly hyperbolic) system is also pseudo-Anosov. Note that Anosov diffeomorphisms possess Markov partitions, and this is also possible (though not necessary) for pseudo-Anosov maps. We now study a specific example of a pseudo-Anosov map, formed via shear composition, which has been shown to possess a Markov partition.

Consider the map $H : \mathbb{T}^2 \to \mathbb{T}^2$ given by

$$H\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + f(y) \\ x + f(y) + y \end{pmatrix}, \tag{43}$$

where

$$f(y) = \begin{cases} 2y & \text{when } y \leq 1/2, \\ 2 - 2y & \text{when } y \geq 1/2. \end{cases}$$

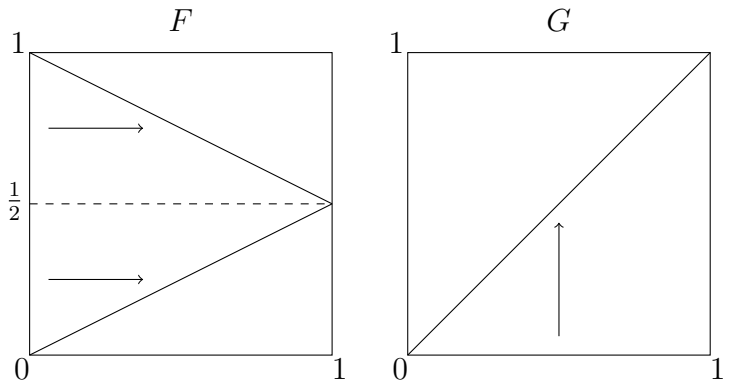


Figure 13: The shears which when composed form H. F is the tent map turned on its side, while G is a shear of slope one.

Figure 13 shows the shears F and G which when composed form H. This map was studied extensively by Cerbelli and Giona in [14], and has been shown to have some interesting properties. It is not uniformly hyperbolic, but Mackay showed in [34] that it does possess a Markov partition and is pseudo-Anosov.

The Jacobian matrix DH of H is given by

$$DH = \begin{cases} \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} & \text{if } y \leq 1/2, \\ \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix} & \text{if } y \geq 1/2. \end{cases}$$

When $y \leq \frac{1}{2}$, DH is a hyperbolic matrix and vectors in tangent space undergo uniform expansion and contraction. This is the 'hyperbolic part' of the map, where all of the expansion takes place. However, when $y \geq \frac{1}{2}$, DH is not a hyperbolic matrix. In fact,

$$DH^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I_2,$$

and thus no expansion occurs when $y \geq \frac{1}{2}$. However, after 2 iterates spent in the

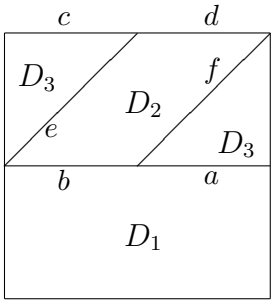


Figure 14: A splitting of the torus into various subsets which are the building blocks for the subsets A, B and C.

region $y \geq \frac{1}{2}$, the next iterate must enter the region $y \leq \frac{1}{2}$, since $DH^2 = -I_2$. This means that μ-a.e. orbit of the system spends at least one third of its iterates in the hyperbolic part of the map.

In [14], Cerbelli and Giona split the torus up into 3 disjoint subsets A, B and C, whose 'building blocks' are shown in Figure 14, in order to discuss point set dynamics. They are given by

$$A = D_1 \cup a \cup c,$$

$$B = D_2 \cup b \cup d,$$

$$C = D_3 \cup e \cup f,$$

and satisfy the relations

$$H(A) \subset A \cup B,$$

$$H(B) = C,$$

$$H(C) \subset A.$$

These relations tell us that points in A either remain there or are mapped into B after 1 iterate. Once mapped into B, they are then mapped into C on the next

iteration, and back into A on the subsequent iteration. This structure helps us to define the stable and unstable foliations, and hence to find a Markov partition for this map. Let:

$$DH_1 = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \text{ and } DH_2 = \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix}.$$

Their eigenvalues are $\lambda_{1\pm} = 2 \pm \sqrt{3}$ and $\lambda_2^2 = -1$ respectively. The eigenvectors of DH_1 corresponding to λ_{1+} and λ_{1-} respectively are:

$$u_1 = \begin{pmatrix} 2 \\ 1+\sqrt{3} \end{pmatrix} \text{ and } s_1 = \begin{pmatrix} 2 \\ 1-\sqrt{3} \end{pmatrix}.$$

These eigenvectors form the expanding and contracting directions for $x \in A$.

Consider the expanding and contracting directions for $x \in B$. For two iterates these manifolds will undergo no expansion/contraction as DH_2 is a non-hyperbolic matrix. However, on the following iterate they will again be 'hit' by DH_1 as the orbit returns to A. Hence the expanding and contracting directions will be the same as those for $x \in A$, but with two applications of DH_2. This means that the expanding and contracting directions for $x \in B$ will be:

$$DH_2^2 u_1 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} u_1 = u_1,$$

$$DH_2^2 s_1 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} s_1 = s_1,$$

and are thus the same as those for $x \in A$. Using a similar argument, the expanding and contracting directions for $x \in C$ are:

$$u_2 = DH_2 u_1 = \begin{pmatrix} 2\sqrt{3} \\ \sqrt{3}-1 \end{pmatrix},$$

$$s_2 = DH_2 s_1 = \begin{pmatrix} 2\sqrt{3} \\ \sqrt{3}+1 \end{pmatrix}.$$

Using these facts, Cerbelli and Giona defined two fields of direction, ε^u and ε^s, as the expanding and contracting fibre bundles respectively, in the following way:

$$\varepsilon^u = (E^u_x)_{x\in\mathbb{T}^2}, \tag{44a}$$

$$\varepsilon^s = (E^s_x)_{x\in\mathbb{T}^2}, \tag{44b}$$

where:

$$E^u_x = \begin{cases} \text{span}(u_1), & \text{for } x \in A \cup B, \\ \text{span}(u_2), & \text{for } x \in C, \end{cases}, \tag{45a}$$

$$E^s_x = \begin{cases} \text{span}(s_1), & \text{for } x \in A \cup B, \\ \text{span}(s_2), & \text{for } x \in C. \end{cases} \tag{45b}$$

In [14] Cerbelli and Giona proved that these fibres are dense in $\mathbb{T}^2$ and that vectors parallel to unstable fibres undergo exponential growth while vectors parallel to stable fibres undergo exponential decay when iterated forwards.

We will now use these fibre bundles to construct a Markov partition for H. Mackay [34] first constructed a 6-element Markov partition for this map and noticed that it should be possible to construct a 4-element partition, due to the appearance of 2 unitary eigenvalues associated with his adjacency matrix. Demers and Wojtkowski [19] then described a general method to construct finite Markov partitions for a family of pseudo-Anosov maps, of which H is a member. They also explicitly constructed a 4-element Markov partition for a case very similar to H, although in a different coordinate system to that of Cerbelli and Giona's map. We will use Wojtkowski and Demers method to construct a 4-element Markov partition for H in the coordinate system used by Cerbelli and Giona.

The first part of the method involves defining two invariant piecewise linear foliations, one stable and one unstable, for our map H. In our case, these foliations will be the expanding and contracting fibre bundles ε^s and ε^u defined in (44). We must then identify the singularity points of H and how many prongs they have.

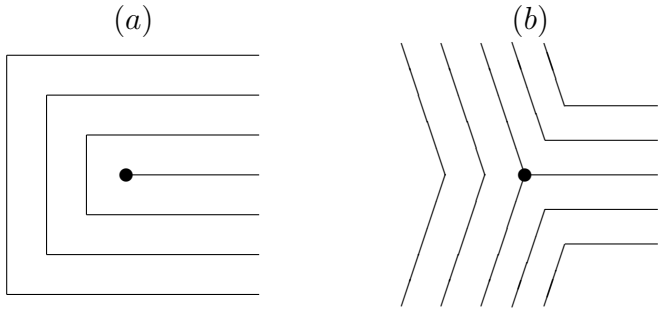


Figure 15: (a) A 1-prong singularity. (b) A 3-prong singularity. See also [56].

Definition 3.4 ([19])**.** *A singularity p of H is called an n-prong singularity ($n \in \mathbb{N}, n \neq 2$) if the local stable and unstable leaves are homeomorphic to the curves $Re(z^{n/2}) = constant$ and $Im(z^{n/2}) = constant$ respectively near to $z = 0$ in $\mathbb{C}$.*

The singularities of H are the fixed points $(0, 0)$ and $(0, 1/2)$, which are 1-prong singularities, and the period 2 points $(1/2, 0)$ and $(1/2, 1/2)$, which are 3-prong singularities. In order to be considered a 1- or 3-prong singularity, the structure of the local stable and unstable leaves of our foliations around the singularity must be qualitatively the same as that shown in Figure 15. With these facts in hand, we are able to start constructing the Markov partition.

Let the elements of our Markov partition be denoted by P_1, P_2, P_3 and P_4. We extend the stable leaves of $(1/2, 0)$ and $(1/2, 1/2)$ outside of D_3 until they intersect with the unstable leaves of $(0, 0)$ and $(0, 1/2)$. We then define P_3 to be the parallelogram bounded by these intersections (see Figure 16). We then define the remaining partition elements as follows:

$$P_2 = H^{-1}(P_3), P_4 = H(P_3), P_1 = T^2/(P_2 \cup P_3 \cup P_4).$$

Each of these elements has boundaries that are composed of a finite union of stable and unstable leaves. The stable leaves emanate from one of the 3-prong singularities,

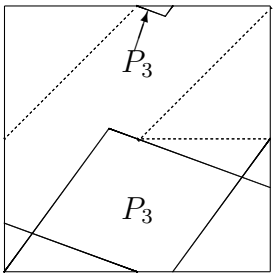


Figure 16: The partition element P_3. The dotted area is D_3.

whereas the unstable leaves emanate from one of the 1-prong singularities. The Markov partition for H is shown in Figure 17.

Now that we have a Markov partition for H, we can calculate its image, and use this to write down a connectivity matrix for the Markov partition [1]. The calculation of the image is somewhat simplified by noting that $H(P_2) = P_3$ and $H(P_3) = P_4$. With this in mind, we obtain the matrix

$$M = \begin{pmatrix} 2 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 2 & 0 & 0 \end{pmatrix}, \tag{46}$$

where the i, jth entry corresponds to the number of different connected components of the image of P_i that cross over P_j.

Note that the matrix M is an adjacency (or connectivity) matrix as opposed to a transition matrix, since the Markov partition we have found is not generating i.e. there is a set of points of positive measure which share the same orbit (are coded in the same way) from the perspective of the symbolic dynamics of our Markov partition. An adjacency matrix is similar to a transition matrix, except its entries may be in $\mathbb{N}$ rather than just being equal to one or zero, with these numbers cor-

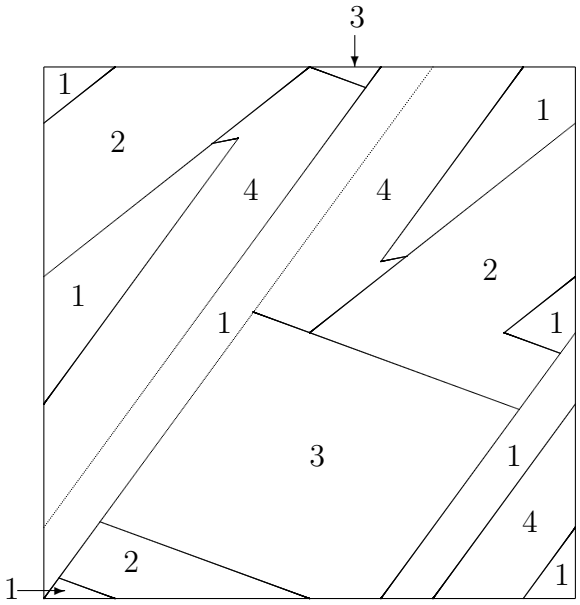


Figure 17: The Markov partition of H. The number n, $n = 1, ..., 4$, corresponds to the partition element P_n.

responding to the number of different connected components which cross over the specified region; in other words, to find M, we determine how many times the image of each partition element crosses over each other element.

The characteristic equation for M is

$$\lambda^4 - 2\lambda^3 - 2\lambda + 1 = 0,$$

which factorises to

$$(\lambda^2 - (1+\sqrt{3})\lambda + 1)(\lambda^2 - (1-\sqrt{3})\lambda + 1) = 0;$$

the eigenvalues of M are then given by

$$\lambda_{1,2} = \frac{1+\sqrt{3} \pm \sqrt{2\sqrt{3}}}{2} \text{ and } \lambda_{3,4} = \frac{1-\sqrt{3} \pm i\sqrt{2\sqrt{3}}}{2}.$$

The largest of the real eigenvalues, λ_1, is the dilatation factor of H. This is the factor by which the measure of the unstable fibre bundles is stretched (and the measure of the stable fibre bundles is contracted) on iteration by H.

The characteristic equation and eigenvalues found here agree with those found by Mackay in [34], minus the two eigenvalues equal to 1, indicating that this is the 4-element partition he suggested should exist. It is worth noting that the element P_2 in this partition is identical to element number 4 in Mackay's partition. This element seems to be somewhat key to building a Markov partition, since it covers most of D_3, the area where the fibre bundles point in the direction of u_2 and s_2.

This map has some important properties which allow the construction of a Markov partition, such as the relationships between the sets A, B, and C and having $DH_2^2 = -I$. These properties prevent the orbits from being trapped in non-hyperbolic areas for extended periods, and thus allow expansion to occur sooner than, for example, in toral linked twist maps, where orbits can become trapped in non-hyperbolic regions for a large number of iterates. As such, this map is a rather special case within the context of compositions of shears; in particular, no Markov partition exists for the systems we studied in Sections 3.1 and 3.2. Note that it has been shown that pseudo-Anosov diffeomorphisms can often be obtained from Dehn twists [23].

3.4 Random dynamical systems

In this section we discuss random dynamical systems. There are multiple ways in which to introduce randomness to a system including, for example, incorporating noise and/or choosing an initial condition at random; we will implement randomness via choosing a Jacobian matrix from a predetermined selection, at random, on each iterate. In particular, we will study the systems formed by selecting two Anosov diffeomorphisms of the form given by (17), and choosing each with a certain probability on every iterate.

In order to define a random dynamical system, we first specify what we mean by a cocycle over a dynamical system.

Definition 3.5 ([28]). *Let $H : M \to M$ be an invertible measure-preserving transformation of a Lebesgue space (M, σ, μ) (with $\mu(M) = 1$) and let $GL(n, \mathbb{R})$ denote the group of invertible linear transformations of $\mathbb{R}^n$. Then for any measurable function $A : M \to GL(n, \mathbb{R})$ and $x \in M$, if we set*

$$\begin{aligned} \mathcal{A}(x, i) &= A(H^{i-1}(x)) \cdots A(x) \text{ for } i > 0, \\ \mathcal{A}(x, i) &= A(H^{-i}(x))^{-1} \cdots A(H^{-1}(x))^{-1} \text{ for } i < 0, \end{aligned} \tag{47}$$

and

$$\mathcal{A}(x, 0) = \ id_M,$$

then it follows that

$$\mathcal{A}(x, i + k) = \mathcal{A}(H^k(x), i)\mathcal{A}(x, k). \tag{48}$$

We call any measurable function $\mathcal{A} : M \times \mathbb{Z} \to GL(n, \mathbb{R})$ satisfying (48) a **measurable linear cocycle over** *H (or simply a cocycle).*

In other words, A assigns to each point x on M the function $A(x)$; similarly, for an orbit $x, H(x), \ldots, H^i(x)$, A assigns the functions $A(x), A(H(x)), \ldots, A(H^i(x))$. The cocycle $\mathcal{A}(x, i)$ forms an ordered product of the functions assigned to each point on the orbit.

We now define what we mean by a random dynamical system. Note that for the systems we study in this section (and Chapter 4), and for any (invertible) discrete time system, we have time $T = \mathbb{Z}$.

Definition 3.6 ([5]). *A* **measurable random dynamical system** *on the measurable space (M, σ) over a metric dynamical system $(\Omega, H, \mathbb{P}, (\theta(t))_{t \in T})$ with time T is a mapping*

$$\phi : T \times \Omega \times M \to M, \ (t, \omega, x) \mapsto \phi(t, \omega, x),$$

with the following properties:

1. *ϕ is $\sigma(T) \otimes H \otimes \sigma$, σ - measurable.*

2. *The mappings* $\phi(t,\omega) := \phi(t,\omega,\cdot) : M \to M$ *form a cocycle over* $\theta(\cdot)$*, i.e. they satisfy*

$$\phi(0,\omega) = id_M \text{ for all } \omega \in \Omega \text{ (if } 0 \in T), \tag{49}$$

$$\phi(t+s,\omega) = \phi(t,\theta(s)\omega) \circ \phi(s,\omega) \text{ for all } s,t \in T,\ \omega \in \Omega. \tag{50}$$

Note that, in this definition, Ω gives us the possible choices of Jacobian matrix we can take at each iterate. $\mathbb{P}$ is a probability distribution, which assigns a probability to each Jacobian matrix in Ω. We then apply at each iterate one of the elements of Ω, at a rate given by the corresponding probability.

We now discuss a specific example of a random dynamical system; this is a specific case of a family of random dynamical systems for which we find bounds on the Lyapunov exponents in Chapter 4. Consider the system $(\mathbb{T}^2, \sigma, A, \mu)$, where $A : \mathbb{T}^2 \to \mathbb{T}^2$ is given by

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \tag{51}$$

is an Anosov diffeomorphism. Similarly, consider the system $(\mathbb{T}^2, \sigma, B, \mu)$, where $B : \mathbb{T}^2 \to \mathbb{T}^2$ is given by

$$B\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}. \tag{52}$$

We will study the random dynamical system on the measurable space $(\mathbb{T}^2, \mathbb{R}^2)$ over the metric dynamical system $(\{A,B\}, \mathbb{T}^2, \{\frac{1}{2}, \frac{1}{2}\}, (t)_{t\in\mathbb{Z}})$ given by $H : \mathbb{Z} \times \{A,B\} \times \mathbb{T}^2 \to \mathbb{T}^2$; that is, a map $H : \mathbb{T}^2 \to \mathbb{T}^2$ which at each iterate chooses A with probability $\frac{1}{2}$, and B with probability $\frac{1}{2}$.

Given our earlier discussion of the systems in Section 3.1, we can infer that the cones

$$C_E = \{(x,y) \in \mathbb{R}^2 : \frac{2}{1+\sqrt{3}} \leq \frac{x}{y} \leq \frac{2}{1+\sqrt{5}}\},$$

and

$$C_C = \{(x,y) \in \mathbb{R}^2 : \frac{2}{1-\sqrt{5}} \leq \frac{x}{y} \leq \frac{2}{1-\sqrt{3}}\},$$

are invariant under H, as they are mutually invariant under A and B; consequently, they form the expanding subspace E_x^u and contracting subspace E_x^s respectively.

In the infinite limit, all possible orbits except for those consisting entirely of As and Bs (which occur with probability zero) will fall into C_E upon forward iteration of H for all vectors $v \in \mathbb{R}^2$; furthermore, in the case of the aforementioned orbits, only the stable eigenvectors v_A^s and v_B^s respectively do not enter C_E upon iteration. We can find lower bounds upon the spectral norms of $\frac{\|Av\|}{\|v\|}$ and $\frac{\|Bv\|}{\|v\|}$ for $v \in C_E$, which yield the minimum growth rates (w.r.t. the l_2 norm) within the cone. We obtain

$$\min_{v \in C_E} \left\{ \frac{\|Av\|_2}{\|v\|_2} \right\} = 2.6102, \text{ and } \min_{v \in C_E} \left\{ \frac{\|Bv\|_2}{\|v\|_2} \right\} = 3.7321,$$

the former of which is the minimum possible growth rate within C_E for H. This tells us that the (maximal) Lyapunov exponent λ can be (almost everywhere) uniformly bound from below by $\log(2.6102) = 0.9594$.

Both this system and similar systems we will study in Chapter 4 possess many similarities to the uniformly hyperbolic deterministic systems discussed in Section 3.1. The differences between the random and deterministic cases stem mainly from the frequency with which specific orbits occur. Each orbit (of equal length) occurs with equal probability in the random case (or, for a more general probability distribution $\mathbb{P}$, with an implicitly specified probability), whereas the measures of the sets of points undergoing certain orbits need not be equal in the deterministic case. This is important, since calculations of the Lyapunov exponent (and similarly bounds on the Lyapunov exponent) need to take all possible orbits into account.

We can introduce randomness into the deterministic cases by choosing an initial condition at random; this can have a large effect on FTLEs in the case of an intermittently hyperbolic map such as the linked twist map, but in the case of the maps in Section 3.1, the influence of the choice of initial condition diminishes quickly with the number of iterates. In the random case, the choice of initial condition has no

effect on the dynamics; variation in the FTLE also diminishes as the number of iterates increases, however, it is always possible to have an unusually large number of either As or Bs at any point on the orbit, which can yield a FTLE significantly lower or higher than expected.

3.5 Summary

In this chapter we have studied various non-uniform systems formed via the composition of shear maps, as well as random dynamical systems. We have seen that constructing maps in this way can lead to various properties and systems emerging, including uniform and non-uniform hyperbolicity, and pseudo-Anosov maps.

We studied a generalization of the uniform Anosov diffeomorphisms defined in Chapter 2, and found that they satisfied the conditions required for uniform hyperbolicity, but are not Anosov. We formed the subspaces required for uniform hyperbolicity from invariant cones for the map, as opposed to the eigenvectors which were used in the uniform case. We find bounds upon the Lyapunov exponent of a specific example of this type of system in Section 5.6.

We then discussed non-uniform hyperbolicity, which loosens the requirement for uniform bounds upon the expansion and contraction rates. The example we discussed was the linked twist map; we studied the sets of points which repeatedly stay in the shear regions, and showed that orbits exist which remain in the shear regions for any finite number of iterates. This result meant that we could not bound the expansion rate of vectors in tangent space from below. Chapter 5 discusses a method for bounding Lyapunov exponents for linked twist maps.

In Section 3.3 we discussed an example of a system studied extensively by Cerbelli and Giona [14], and Mackay [34]. The map was found to be pseudo-Anosov, which is a generalization of Anosov diffeomorphisms. We also constructed a Markov partition for this map, using the method outlined by Demers and Wojtkowski [19].

Finally, we defined random dynamical systems, and discussed how they differ from deterministic systems. We gave a specific example of a system which chooses an Anosov diffeomorphism at random on each iterate, and saw that its expansion rate can be uniformly bound from below. We discuss a method for finding rigorous elementary bounds upon the Lyapunov exponent for similar maps in Chapter 4.

4 Bounding Lyapunov exponents of random products of shear compositions

This chapter will discuss a method for obtaining an upper and a lower bound on the largest Lyapunov exponent, λ, for random products of hyperbolic matrices; in particular, those formed by the composition of shear matrices. The method for calculating bounds on λ is a continuation of the work done by Sturman and Thiffeault [54], who obtained explicit and elementary bounds for random products of shear matrices.

The method utilises cones which are invariant under both of the randomly chosen matrices (mutually invariant cones), the existence of which allow us to restrict our search for the maximum and minimum possible growth rates for vectors (upon iteration) to just the vectors within these cones. We will also perform the relevant calculations for three different norms in order to improve the bounds further.

To obtain these bounds, we will split the sequence of iterates up into subsequences of smaller matrices, $M_{a,b}$, as in [54], in order to allow for diagonalisation of the matrix products. Following this, we will improve upon the bounds further via a method also discussed in [54], which involves considering the possible preceding matrices within the larger product. This will allow us to narrow the invariant cones we use, and thus narrow the range of vectors we consider within each cone.

In Section 4.1 we define the systems we will be studying in this chapter. In Section 4.2 we discuss the Lyapunov exponents of these systems and ways in which we might try to bound them. In Section 4.3 we find cones which are invariant under both matrices in our system and determine when they exist. In Section 4.4 we calculate the simplest version of the bounds using these cones, which consider only one iterate of the system at a time. In Section 4.5 we improve upon these bounds by considering multiple iterates of the system at once. In Section 4.6, we attempt

to improve further upon the bounds in the previous section, particularly in cases where the mutually invariant cone is wide, by finding a way to narrow the cone in specific situations. Finally, in Section 4.7 we summarise what we find.

4.1 Systems formed by random products of shear compositions

In this section we define the family of random dynamical systems for which we will obtain rigorous bounds upon the Lyapunov exponent(s). Consider the system $(\mathbb{T}^2, \sigma, A, \mu)$ where σ is the Borel σ-algebra for $\mathbb{T}^2$, μ is Lebesgue measure, and $A : \mathbb{T}^2 \to \mathbb{T}^2$, given by

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & \alpha \\ \beta & 1 + \alpha\beta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{53}$$

is the composition of a horizontal shear of slope α and a vertical shear of slope β, where $\alpha, \beta \in \mathbb{R}$. Similarly, consider the system $(\mathbb{T}^2, \sigma, B, \mu)$, where $B : \mathbb{T}^2 \to \mathbb{T}^2$ is given by

$$B \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & \gamma \\ \delta & 1 + \gamma\delta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{54}$$

We will study the random dynamical system on the measurable space $(\mathbb{T}^2, \mathbb{T}^2)$ over the metric dynamical system $(\{A, B\}, \mathbb{T}^2, \{\frac{1}{2}, \frac{1}{2}\}, (\theta(t))_{t\in\mathbb{Z}})$ given by $H : \mathbb{Z}\times\{A, B\}\times \mathbb{T}^2 \to \mathbb{T}^2$; that is, a map $H : \mathbb{T}^2 \to \mathbb{T}^2$ which at each iterate behaves as A with probability $\frac{1}{2}$, and as B with probability $\frac{1}{2}$. In order to simplify notation we will refer to both the systems and the maps described above as A and B respectively, where no confusion arises.

The eigenvalues of the Jacobian matrix of A, DA (which we shall, again for notational convenience, refer to as A where no confusion arises), are given by

$$\lambda_{A\pm} = \frac{2 + \alpha\beta \pm \sqrt{\alpha\beta(4 + \alpha\beta)}}{2}, \tag{55}$$

and the respective eigenvectors by

$$v_{A\pm} = \begin{pmatrix} 2\alpha \\ \alpha\beta \pm \sqrt{\alpha\beta(4+\alpha\beta)} \end{pmatrix}. \tag{56}$$

Note that $\lambda_{B\pm}$ and $v_{B\pm}$ are obtained by swapping γ for α and δ for β in (55) and (56). The systems A and B are both hyperbolic automorphisms provided the constants α, β adhere to either

(i) $\alpha\beta > 0$, or

(ii) $\alpha\beta < -4$,

and similarly for γ, δ. The method we discuss will make use of invariant cones shared by the matrices A and B, and as such we require these matrices to be hyperbolic in order for invariant cones to exist for each matrix individually.

We will frequently refer to the ratio of the x and y components of vectors in tangent space throughout this chapter for the purpose of comparison; in particular, the ratios corresponding to the eigenvectors $v_{A\pm}$ and $v_{B\pm}$ are given by

$$r_{A\pm} = \frac{2\alpha}{\alpha\beta \pm \sqrt{\alpha\beta(4+\alpha\beta)}}, \tag{57}$$

and

$$r_{B\pm} = \frac{2\gamma}{\gamma\delta \pm \sqrt{\gamma\delta(4+\gamma\delta)}} \tag{58}$$

respectively.

Note that λ_{A+} is the unstable eigenvalue when $\alpha\beta > 0$ and the stable eigenvalue when $\alpha\beta < -4$, and vice versa for λ_{A-}; the same properties hold for $\lambda_{B\pm}$, for the corresponding γ and δ. We will label the stable and unstable eigenvectors of A as v_A^s and v_A^u respectively, to avoid confusion in cases such as $\alpha\beta > 0$ and $\gamma\delta < -4$, where v_{A+} and v_{B-} are the unstable eigenvectors of each matrix. Similarly, we will refer to the corresponding eigenvalues and ratios of components of the eigenvectors as $\lambda_A^{s,u}$

and $r_A^{s,u}$ respectively. The possible orientations of v_A^s and v_A^u for the different cases we will consider are shown in Figure 18. Knowing the positions of these vectors is useful as, later in this chapter, we will use them to form the boundaries of invariant cones for the map H.

4.2 Lyapunov exponents

Recall the Lyapunov exponents $\lambda_{1,2}$ of a deterministic system G:

$$\lambda_{1,2}(x,v) = \lim_{n\to\pm\infty} \frac{1}{n} \log ||DG_x^n v||.$$

In the case of the random system H, the Jacobian matrix depends on our choice of A or B at each iterate, but not on x. Furthermore, since our system is probabilistic, we require the expected value of this quantity. The Lyapunov exponents of H are therefore given by

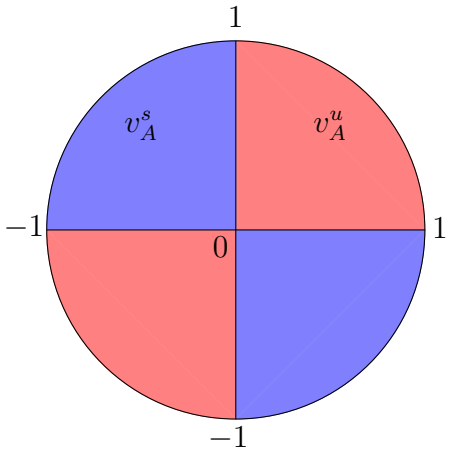
$$\lambda_{1,2}(v) = \lim_{n\to\pm\infty} \frac{1}{n} \mathbb{E}(\log ||H^n v||). \tag{59}$$

Note that this quantity was shown to converge almost everywhere by Furstenberg and Kesten [25]. In this chapter we will study bounds upon the largest Lyapunov exponent of H, $\lambda_1 = \lambda$; note that since both A and B are measure-preserving, the method will also obtain bounds upon the smallest Lyapunov exponent, $\lambda_2 = -\lambda$.

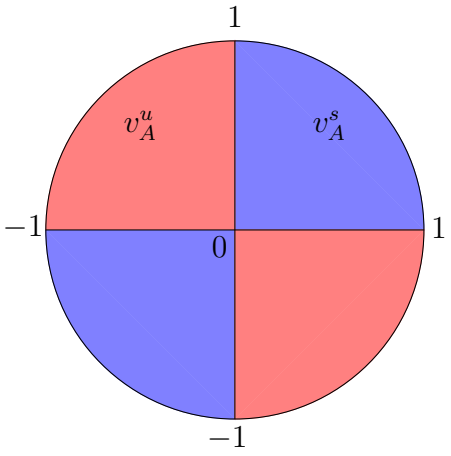
The simplest rigorous upper bound for λ is obtained by assuming H picks the matrix with the largest unstable eigenvalue (in modulus) at each iterate. In this case we obtain

$$\lambda \leq \max\{\lambda_{A_1}, \lambda_{B_1}\},$$

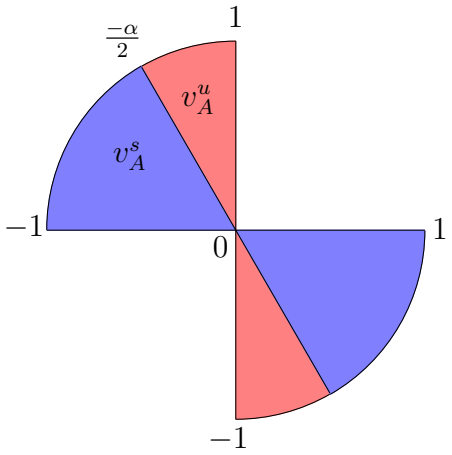
where λ_{A_1,B_1} are the maximal Lyapunov exponents for A and B respectively. Any upper bound we calculate must therefore be less than the maximum of λ_{A_1} and λ_{B_1} to be of any use. We cannot obtain a lower bound in a similar way (i.e. by taking the lesser of λ_{A_1} and λ_{B_1}). To see this, consider the case where $\alpha, \beta = 1$ and $\gamma, \delta = -1$. The eigenvalues of A and B are given by $\lambda_{A\pm} = \lambda_{B\pm} = \frac{3\pm\sqrt{5}}{2}$, and so taking the lower



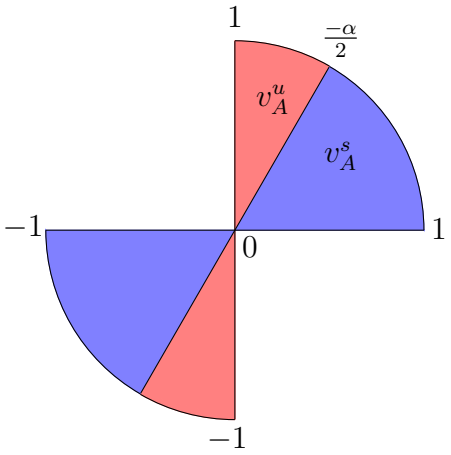
(a) $\alpha, \beta > 0$.



(b) $\alpha, \beta < 0$.



(c) $\alpha\beta < -4, \alpha > 0, \beta < 0$, for fixed α.



(d) $\alpha\beta < -4, \alpha < 0, \beta > 0$, for fixed α.

Figure 18: The regions which the stable and unstable eigenvectors v_A^s and v_A^u can inhabit for various choices of α and β. In cases (c) and (d), the vectors v_A^u and v_A^s are found on either side of the vector with slope $\frac{-\alpha}{2}$, with their exact orientation depending on the choice of β; varying β while keeping α fixed allows the eigenvectors to attain the ranges shown in the figures.

bound to be the maximal Lyapunov exponent of the system corresponding to the smaller of these would yield $\lambda = \log|\frac{3+\sqrt{5}}{2}| \approxeq 0.962$. However, numerical calculation of λ using Gram-Schmidt orthonormalization yields $\lambda \approxeq 0.663$, and thus our lower bound must be invalid. The reason for this is that the vectors we consider when calculating λ are constantly being shifted around, and so do not converge to v_{A+} or v_{B+} for an infinite number of iterates, as is the case when calculating the Lyapunov exponent of A or B individually. We therefore regularly observe expansion rates less than the smaller of the two Lyapunov exponents of A and B in this case.

A common upper bound for λ is given by submultiplicativity of matrix norms [30]. Let M_{2^k} be the set of matrices formed from all possible sequences of A and B of length 2^k. For example,

$$M_{2^0} = \{A, B\},$$

$$M_{2^1} = \{AA, AB, BA, BB\},$$

$$M_{2^2} = \{AAAA, AAAB, \ldots, BBBA, BBBB\}.$$

An upper bound is then given by

$$E_k = 2^{-k}\mathbb{E}\{\log||M_{2^k}||\}, \tag{60}$$

where 2^{-k} arises due to each matrix product in the set M_{2^k} corresponding to 2^k iterates of H. Note that, by submultiplicativity of matrix norms,

$$\begin{aligned}\log||H^n|| &= \log||AABBABAA\ldots AABAAA||,\\ &\leq \Big(c_1\log||A\ldots A|| + c_2\log||AA\ldots B|| + \ldots + c_{2^k}\log||B\ldots B||\Big),\end{aligned}$$

where c_i is the number of times the i^{th} matrix of the set M_{2^k} appears in the sequence H^n. Thus

$$\lim_{n\to\infty}\mathbb{E}\{\log||H^n||\} \leq 2^{-2^k}\Big(\log||A\ldots A|| + \ldots + \log||B\ldots B||\Big) = \mathbb{E}\{\log||M_{2^k}||\},$$

since $\frac{c_i}{n} \to 2^{-2^k}$ as $n \to \infty$.

The E_k's will decrease monotonically to λ as $k \to \infty$, for any choice of matrix norm $||\cdot||$ and for any two matrices A and B that we choose. This is due to the fact that as we increase k, the bound accounts for more of the underlying distribution. For example, the sequences

$$AABBAABBAABB\ldots \text{ and } ABABABABABAB\ldots$$

return the same frequencies for matrices in M_{2^0}, but different frequencies for matrices in M_{2^1}. A and B need not be hyperbolic, nor formed by composition of shears, for this bound to hold. However, in the case we are interested in, A and B are hyperbolic, and we can use this fact to obtain a more practical upper bound than E_k, and to also obtain a lower bound.

Let us consider $H^n v_0$, i.e. applying the map H n times to a vector v_0 in the tangent space $\mathbb{R}^2$, where x_0 is our initial condition. We obtain a random sequence of A's and B's applied to our vector v_0:

$$H^n v_0 = AABABB\ldots BBABv_0.$$

Now consider $||H^n v_0||$. Let $H_k = A$ if on the k^{th} iterate we apply an A, and $H_k = B$ if on the k^{th} iterate we apply a B. We have

$$||H^n v_0|| = ||v_n|| = ||H_n v_{n-1}|| = \frac{||H_n v_{n-1}||}{||v_{n-1}||} \cdot ||v_{n-1}||,$$

where $v_{n-1} = H^{n-1} v_0$. Continuing this process we obtain

$$||H^n v_0|| = \frac{||H_n v_{n-1}||}{||v_{n-1}||} \cdot \ \ldots \ \cdot \frac{||H_1 v_0||}{||v_0||} \cdot ||v_0||. \tag{61}$$

Thus, to obtain an upper and lower bound on λ, we need to find a way to bound

$$\frac{||H_i v_{i-1}||}{||v_{i-1}||} \tag{62}$$

for each $i \in \{1, ..., n\}$. To do this, we utilise the existence of a mutually invariant cone for A and B; such a cone will be invariant under all applications of H. Since vectors

are trapped within the cone for all iterates, finding the maximum and minimum values of (62) for vectors within this cone will yield upper and lower bounds upon this quantity for all iterates of H. To obtain bounds upon λ, we need to obtain bounds for $||H^n v||$ for any vector v in such a cone.

Explicitly evaluating $||H^n v||$ (and subsequently λ) when the Jacobian matrix H is subject to change can be difficult or impossible, since the growth of the norm is dependent on the orientation of v; however in some special cases, such as Cerbelli and Giona's archetype for non-uniform chaos, the Lyapunov exponent can be formulated explicitly. In the case where H is constant for all x, such as the Cat Map, we find that repeated application of H aligns almost all vectors v with the unstable eigenvector (see Lemma 4.1); when matrices are chosen at random it is possible to be pulled away from this eigenvector by application of a different Jacobian, altering the rate of expansion for subsequent iterates. Thus, in order to evaluate $||H^n v||$, we need to know how the orientation of v will change under application by A or B. To do this, we will first define what we mean by alignment of vectors in 2 dimensions.

Definition 4.1. *Let $u = (u_1, u_2)$ and $v = (v_1, v_2)$ be vectors in $\mathbb{R}^2$. We assign to each of these vectors an angle $\theta \in [0, \pi]$ between themselves and the vector $(-1, 0)$, that is*

$$\theta(v) = \frac{\pi}{2} + \arctan \frac{v_1}{v_2},$$

taking the appropriate limits for the ratios $\pm\infty$. Then the smallest angle between the vectors is given by

$$\theta_{u,v} = \min \big\{ |\theta(v) - \theta(u)|, |\pi - \theta(v) + \theta(u)| \big\}.$$

We say a vector u is more aligned with v than with w if

$$\theta_{u,v} < \theta_{u,w}.$$

Figure 19 shows the angle $\theta_{u,v}$. The following lemma describes the alignment of a vector after application by A.

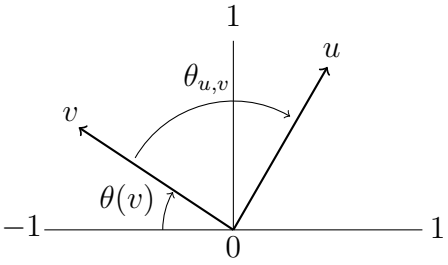


Figure 19: $\theta_{u,v}$.

Lemma 4.1. *Let v be a vector in the tangent space $\mathbb{R}^2$ of a point $x \in \mathbb{T}^2$, and let*

$$A = \begin{pmatrix} 1 & \alpha \\ \beta & 1+\alpha\beta \end{pmatrix}.$$

Let λ_A^u and λ_A^s be the unstable and stable eigenvalues of A, and v_A^u and v_A^s be their corresponding eigenvectors. Then

1. *if $v \notin \{v_A^u, v_A^s\}$, we have $\theta_{Av,v_A^u} < \theta_{v,v_A^u}$.*

2. *if $v \in \{v_A^u, v_A^s\}$, $Av = \lambda v$, for $\lambda \in \{\lambda_A^u, \lambda_A^s\}$ respectively.*

Proof. Part 2 is simply the equation for the eigenvectors of A, which v satisfies by definition in this case. For Part 1, let $v \in \mathbb{R}^2$, then

$$v = c_1 v_A^u + c_2 v_A^s,$$

for some $c_1, c_2 \in \mathbb{R}$. We have that

$$Av = c_1 A v_A^u + c_2 A v_A^s = c_1 \lambda_A^u v_A^u + c_2 \lambda_A^s v_A^s,$$

$$\theta_{v,v_A^u} = \min\Big\{ \left| \arctan \frac{v_{A1}^u}{v_{A2}^u} - \arctan \frac{c_1 v_{A1}^u + c_2 v_{A1}^s}{c_1 v_{A2}^u + c_2 v_{A2}^s} \right|, \\ \left| \pi - \arctan \frac{v_{A1}^u}{v_{A2}^u} + \arctan \frac{c_1 v_{A1}^u + c_2 v_{A1}^s}{c_1 v_{A2}^u + c_2 v_{A2}^s} \right| \Big\},$$

and

$$\theta_{Av,v_A^u} = \min\Big\{ \left|\arctan\frac{v_{A1}^u}{v_{A2}^u} - \arctan\frac{c_1\lambda_A^u v_{A1}^u + c_2\lambda_A^s v_{A1}^s}{c_1\lambda_A^u v_{A2}^u + c_2\lambda_A^s v_{A2}^s}\right|, \\ \left|\pi - \arctan\frac{v_{A1}^u}{v_{A2}^u} + \arctan\frac{c_1\lambda_A^u v_{A1}^u + c_2\lambda_A^s v_{A1}^s}{c_1\lambda_A^u v_{A2}^u + c_2\lambda_A^s v_{A2}^s}\right| \Big\}.$$

Noting that $0 < |\lambda_A^s| < 1 < |\lambda_A^u|$, we have either

$$\left|\frac{v_{A1}^u}{v_{A2}^u}\right| < \left|\frac{c_1\lambda_A^u v_{A1}^u + c_2\lambda_A^s v_{A1}^s}{c_1\lambda_A^u v_{A2}^u + c_2\lambda_A^s v_{A2}^s}\right| < \left|\frac{c_1 v_{A1}^u + c_2 v_{A1}^s}{c_1 v_{A2}^u + c_2 v_{A2}^s}\right|,$$

or

$$\left|\frac{c_1 v_{A1}^u + c_2 v_{A1}^s}{c_1 v_{A2}^u + c_2 v_{A2}^s}\right| < \left|\frac{c_1\lambda_A^u v_{A1}^u + c_2\lambda_A^s v_{A1}^s}{c_1\lambda_A^u v_{A2}^u + c_2\lambda_A^s v_{A2}^s}\right| < \left|\frac{v_{A1}^u}{v_{A2}^u}\right|.$$

Using this, we can conclude that

$$\left|\arctan\frac{v_{A1}^u}{v_{A2}^u} - \arctan\frac{c_1\lambda_A^u v_{A1}^u + c_2\lambda_A^s v_{A1}^s}{c_1\lambda_A^u v_{A2}^u + c_2\lambda_A^s v_{A2}^s}\right| < \left|\arctan\frac{v_{A1}^u}{v_{A2}^u} - \arctan\frac{c_1 v_{A1}^u + c_2 v_{A1}^s}{c_1 v_{A2}^u + c_2 v_{A2}^s}\right|,$$

and hence

$$\theta_{Av,v_A^u} < \theta_{v,v_A^u},$$

as desired. □

Lemma 4.1 tells us that applications of A or B will align our tangent vectors with their respective unstable eigenvectors. If, for example, we always chose A's for our Jacobian, then $A^n v$ would begin to align with the unstable eigenvector of A, v_A^u, as $n \to \infty$ [5]. Bounding the norm is then a matter of observing the growth rate of the vector v_A^u, as for all $v \neq v_A^s$, we have

$$\lim_{n\to\infty}\frac{1}{n}\log||A^n v|| = \lim_{n\to\infty}\frac{1}{n}\log||A^n v_A^u||.$$

However, H chooses between 2 different matrices for its Jacobian, so this orientation can vary significantly. For example, if we imagine H choosing a long string of A's, then $H^n v$ will begin to align with v_A^u. If H then chooses a B, our vector will begin

[5] Provided $v \neq v_A^s$, the stable eigenvector of A, in which case its orientation will not change.

to align with v_B^u instead, and this will have an impact on the growth rate of the vector. We will use this information to help us in the construction of an invariant cone for H.

Furthermore, Lemma 4.1 does not tell us *how* we approach the unstable eigenvectors, merely that the angle between the image of our vector and the unstable eigenvector shrinks after each iterate. This is important, since the manner of this convergence will determine our choice of an invariant cone. If a vector approaches the eigenvector from one side only, then we need only consider the vectors on one side of the eigenvector to obtain an invariant cone. If the vector instead approaches by flipping from one side of the eigenvector to the other with each iterate, then our cone will need to contain vectors on both sides of the eigenvector in order to account for this. We discuss the construction of these invariant cones in the next section.

4.3 Mutually invariant cones

In order to obtain bounds upon the quantity given by (62), we will restrict the range of vectors we consider from the entire tangent space to a smaller bundle of tangent vectors: an invariant cone. In order for a cone to be invariant under applications of H, it must be invariant under both A and B. Recall Theorem 2.4, which states that an invariant cone for a 2×2 diagonalizable matrix A with non-negative eigenvalues is given by any cone which contains v_A^u and whose interior does not contain v_A^s. Hence we can find an invariant cone of this form for the matrix A in the cases when $\alpha\beta > 0$.

When $\alpha\beta < -4$, Lemma 4.1 tells us that applications of A to a vector $v \neq v_A^s$ will cause v to become more aligned with v_A^u than v_A^u is with v_A^s, and repeated applications of A will cause the ratio of the x and y components of v, r_v, to tend to r_A^u. Thus the same cone given by Theorem 2.4 will work in this case as well; the difference is that the orientation of the vector will flip after each iterate in the

cases where the eigenvalues are negative. In other words, a vector pointing into the positive quadrant, $x, y \geq 0$, will flip its orientation into the negative quadrant, $x, y \leq 0$, on the following application of A, and vice versa. The invariant cone for any matrix A of the form (53) is given by the following Lemma.

Lemma 4.2. *Let A be given by (53) and let C_A be the cone with boundaries given by the vectors $u = (u_1, u_2)$ and $v = (v_1, v_2)$, with the following conditions:*

1. *Any of the following hold:*
   - (a) $r_A^s \leq r_A^u$, *and* $r_u = \frac{u_1}{u_2} \in [r_A^s, r_A^u]$,
   - (b) $r_A^s \geq r_A^u$, *and* $r_u \in [r_A^u, r_A^s]$,
   - (c) $0 < r_A^s < \infty$, $-\infty < r_A^u < 0$ *and* $r_u \in [r_A^s, \infty] \cup [-\infty, r_A^u]$,[6]
   - (d) $0 < r_A^u < \infty$, $-\infty < r_A^s < 0$ *and* $r_u \in [r_A^u, \infty] \cup [-\infty, r_A^s]$,
2. $\theta_{u,v_A^u} \leq \theta_{v_A^s,v_A^u}$,
3. $\theta_{v,v_A^u} \leq \pi - \theta_{v_A^s,v_A^u}$.

Then C_A is an invariant cone for A.

A sketch of C_A is shown in Figure 20.

Proof. From Lemma 4.1, we know that both Au and Av are more aligned with v_A^u than u and v respectively. Therefore, by continuity of A, all vectors w between u and v_A^u must have that Aw is more aligned with v_A^u than Au, and similarly for all vectors w between v and v_A^u. We therefore have $AC_A \subset C_A$, and thus C_A is invariant under A. □

[6]The closed boundaries on $\pm\infty$ here are to indicate the inclusion of the vectors $(1, 0)$ and $(-1, 0)$ as possibilities for the vector u, and similarly in part 1d.

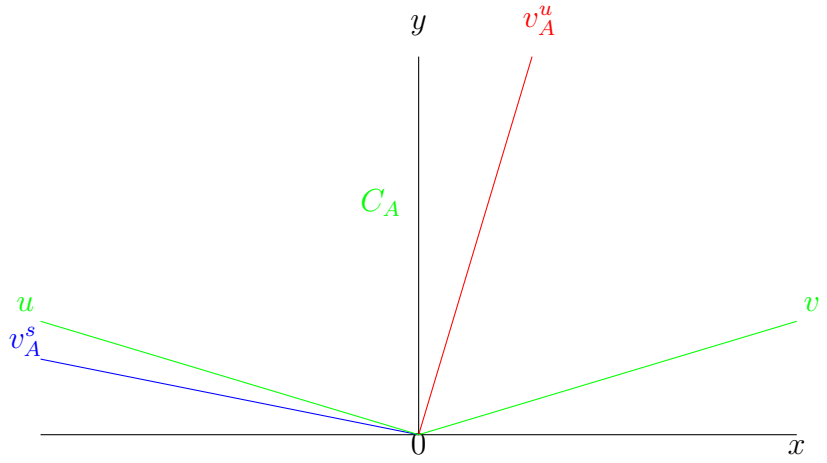


Figure 20: An example of a possible cone C_A in tangent space, for the case where $\alpha, \beta > 0$. The boundaries of C_A are given by the vectors u and v (green lines). Note that C_A is invariant under A for u up to and including v_A^s.

The most important factor in the construction of mutually invariant cones for A and B is the invariance of the eigenvectors under their respective maps. The idea will be to create a cone which traps the tangent vector between two eigenvectors - one boundary of the cone is impassable under applications by A, and the other boundary by B. By Lemma 4.2, this will only be possible if:

1. both of v_A^u and v_B^u are contained within the cone, and

2. neither of v_A^s nor v_B^s are contained in the interior of the cone.

The following theorem tells us the conditions required for such a cone to exist.

Theorem 4.1. *Let A and B be given by (53) and (54) respectively and, without loss of generality, assume that*

$$r_A^u \leq r_B^u.$$

Then a cone which is invariant under applications of both A and B exists if and only if either

(i) $r_A^s \leq r_A^u$ *and* $r_B^s \geq r_B^u$, *or*

(ii) $r_A^s \geq r_B^u$ *and* $r_B^s \leq r_A^u$, *or*

(iii) $r_A^s, r_B^s \geq r_A^u$, *and* $r_A^s, r_B^s \leq r_B^u$.

Furthermore, the minimal such cone for cases (i) and (ii), is given by

$$C_{AB} = \left\{(x, y) \in T_x\mathbb{T}^2 : r_A^u \leq \frac{x}{y} \leq r_B^u\right\},$$

and for case (iii) by

$$\hat{C}_{AB} = {C_{AB}}^C \cup \{v_A^u, v_B^u\}.$$

Proof. By Lemma 4.2, a cone is invariant under A or B if it contains the respective unstable eigenvector and its interior does not contain the respective stable eigenvector. Hence, in order for a single cone to be invariant under both A and B, we must meet all four of these conditions at once. In other words, we must be able to find an invariant cone for A which contains v_B^u but whose interior does not contain v_B^s. This can only occur if either

$$r_A^s \leq r_A^u \text{ and } r_B^s \geq r_B^u,$$

or

$$r_A^s \geq r_B^u \text{ and } r_B^s \leq r_A^u,$$

or

$$r_A^s, r_B^s \geq r_A^u, \text{ and } r_A^s, r_B^s \leq r_B^u,$$

and thus this proves the first statement.

To prove the second statement, we note that any cone which contains both v_A^u and v_B^u is by definition a (possibly non-strict) superset of either C_{AB} or $\hat{C}_{AB}$, and so by taking the intersection of all such invariant cones under A and B we obtain the minimal cone, C_{AB}, in cases (i) and (ii), or $\hat{C}_{AB}$ in case (iii). □

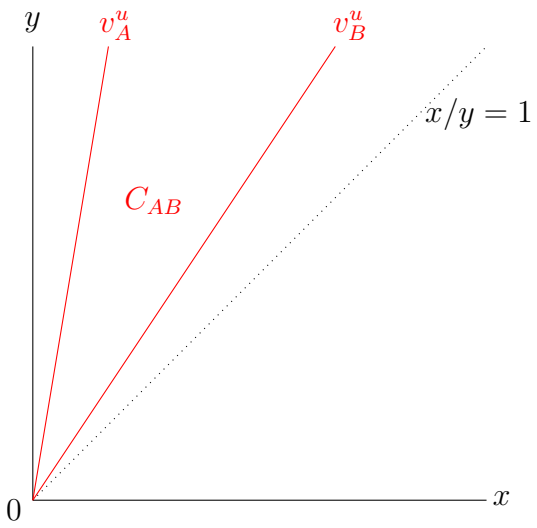
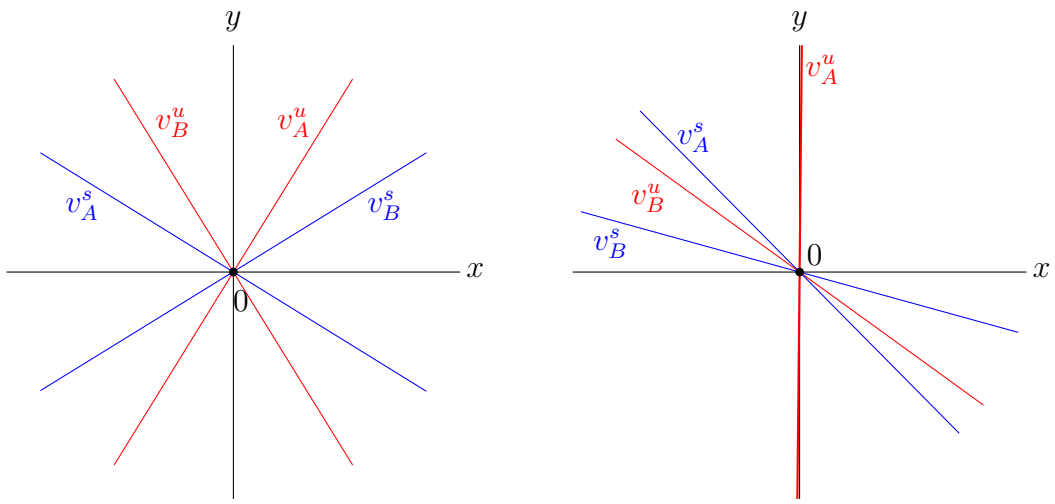


Figure 21: The cone C_{AB} in tangent space, for the case where α, β, γ, and δ are positive integers.

Note that the cones C_{AB} and $\hat{C}_{AB}$ given by Theorem 4.1 belong to the set of expansion cones discussed by Ayyer and Stenlund in [7] and are the minimal choice for such cones. The cones C_{AB} are analogous to the cones C_+ and C_- studied by Sturman and Thiffeault in the case of random products of shear matrices; C_{AB} has the capacity to be both narrower and wider than these cones, depending on the choice of $\alpha, \beta, \gamma, \delta$.

An example of C_{AB} for the case when α, β, γ, and δ are positive integers is shown in Figure 21. Note that the unstable eigenvectors partition tangent space into two distinct cones, C_{AB} and ${C_{AB}}^C$. Theorem 4.1 tells us that in order to discover if a particular pair of matrices (of the appropriate form) possess a mutually invariant cone, we must calculate their eigenvectors, and then determine whether the stable eigenvectors are contained within the same partition element (see Figure 22). If they are, then a mutually invariant cone for A and B exists, and if not, then no such cone exists. From this point forward, we will refer to the minimal, mutually



(a) $\alpha = 1, \beta = 1, \gamma = -1, \delta = -1$. (b) $\alpha = 1, \beta = 100, \gamma = 5, \delta = -1$.

Figure 22: Examples of cases where a mutually invariant cone for the matrices A and B either (a) can or (b) cannot exist.

invariant cone for the matrices A and B as C_{AB}, regardless of the case of Theorem 4.1 we are studying, in order to simplify notation. Note that C_{AB} is an invariant cone for any product of A and B, and is therefore also an invariant cone for the system H.

Note that not all A and B containing only integer entries satisfy Theorem 4.1. An example of a case which does not is $\alpha = 1, \beta = 100, \gamma = 5, \delta = -1$. In this case (and others like it) the stable eigenvectors are located on either side of v_B^u, and so the cone C_{AB} is not invariant (see Figure 22b).

4.4 Calculating the bounds

Now that we have found an invariant cone for H, we can find the maximum and minimum values of $\frac{||H_i v_{i-1}||}{||v_{i-1}||}$ for a vector in the cone for a given norm. For a vector

$v = (x, y) \in \mathbb{R}^2$, we will consider the l_2 norm, given by

$$||v||_2 = \sqrt{x^2 + y^2}, \tag{63}$$

the l_1 norm, given by

$$||v||_1 = |x| + |y|, \tag{64}$$

and the l_∞ norm, given by

$$||v||_\infty = \max\big\{|x|, |y|\big\}. \tag{65}$$

For each of these, we will calculate a strict upper and lower bound upon the value of both $\frac{||Av||}{||v||}$ and $\frac{||Bv||}{||v||}$ for $v \in C_{AB}$. Our upper and lower bounds on λ, which we label Φ_0 and Ψ_0 respectively, are given by inputting the bounds on $\frac{||Av||}{||v||}$ and $\frac{||Bv||}{||v||}$ into (59); this is summarised by the following theorem.

Theorem 4.2. *Let H be the random dynamical system on the measurable space $(\mathbb{T}^2, \mathbb{T}^2)$ over the metric dynamical system $(\{A, B\}, \mathbb{T}^2, \{\frac{1}{2}, \frac{1}{2}\}, (\theta(t))_{t\in\mathbb{Z}})$ given by $H : \mathbb{Z} \times \{A, B\} \times \mathbb{T}^2 \to \mathbb{T}^2$, where A and B are given by (53) and (54) respectively. Then an upper bound for the maximal Lyapunov exponent λ is*

$$\begin{aligned}\Phi_0(A, B) &= \lim_{n\to\infty} \frac{1}{n} \log\Big(\max_{v\in C_{AB}} \big\{\frac{||Av||}{||v||}\big\}\Big)^a \cdot \Big(\max_{v\in C_{AB}} \big\{\frac{||Bv||}{||v||}\big\}\Big)^b, \\ &= \frac{1}{2}\log\Big(\max_{v\in C_{AB}} \big\{\frac{||Av||}{||v||}\big\}\Big) + \frac{1}{2}\log\Big(\max_{v\in C_{AB}} \big\{\frac{||Bv||}{||v||}\big\}\Big),\end{aligned} \tag{66}$$

where a and b are the number of occurrences of A and B respectively; note that $\frac{a}{n}$ and $\frac{b}{n} \to \frac{1}{2}$ as $n \to \infty$. Similarly, we obtain the lower bound

$$\Psi_0(A, B) = \frac{1}{2}\log\Big(\min_{v\in C_{AB}} \big\{\frac{||Av||}{||v||}\big\}\Big) + \frac{1}{2}\log\Big(\min_{v\in C_{AB}} \big\{\frac{||Bv||}{||v||}\big\}\Big). \tag{67}$$

It is worth noting that both Φ_0 and Ψ_0 reduce to (8) when $A = B$. In this case the invariant cone can be taken to be the line v_A^u, the maximum and minimum growth rates are equal, and we obtain the Lyapunov exponent of A.

In the case of the l_2 norm, the quantity $\frac{||Av||_2}{||v||_2}$ is the *spectral norm* of the matrix A, usually denoted $||A||$ (see, for example, [52]). We have that

$$\max_{v\neq 0} \frac{||Av||_2^2}{||v||_2^2} = \max_{v\neq 0} \frac{v^T A^T A v}{||v||_2^2} = \lambda^u_{A^T A},$$

and similarly

$$\min_{v\neq 0} \frac{||Av||_2^2}{||v||_2^2} = \min_{v\neq 0} \frac{v^T A^T A v}{||v||_2^2} = \lambda^s_{A^T A},$$

where $\lambda^{u,s}_{A^T A}$ are the unstable and stable eigenvalues of $A^T A$ respectively. Note that $\lambda^{u,s}_{A^T A}$ are always positive. These give us the maximum and minimum growth rates respectively:

$$\max_{v\neq 0} ||A|| = \sqrt{\lambda^u_{A^T A}} \text{ and } \min_{v\neq 0} ||A|| = \sqrt{\lambda^s_{A^T A}}.$$

These growth rates are obtained at the corresponding eigenvectors of $A^T A$, and the spectral norm is monotonic between these eigenvectors. This raises an issue, as it is possible for these vectors to be within the interior of C_{AB}, and thus our maximum/minimum is not always obtained on the boundaries of the cone. In other words, if we were to write

$$||A|| = f(z),$$

where $z = x/y$, we would find that this function is not always monotonic over C_{AB}. The following lemma takes into account the cases where $v^u_{A^T A}$ or $v^u_{B^T B}$ are either in the interior or exterior C_{AB}. In the lemma and proof, $v^{u,s}_{A^T A}$ are the unstable and stable eigenvectors of $A^T A$ respectively.

Lemma 4.3. *Let A and B of the form given by (53) and (54). Then we obtain the following upper bounds for $\frac{||Av||_2}{||v||_2}$ for $v \in C_{AB}$:*

1. *If $v^u_{A^T A} \in C_{AB}$ then,*

$$\frac{||Av||_2}{||v||_2} \leq \sqrt{\lambda^u_{A^T A}},$$

2. *If $v^u_{A^T A} \notin C_{AB}$ then,*

$$\frac{||Av||_2}{||v||_2} \leq \max\left\{ \frac{||Av^u_A||_2}{||v^u_A||_2}, \frac{||Av^u_B||_2}{||v^u_B||_2} \right\}.$$

We also obtain the following lower bounds for $\frac{||Av||_2}{||v||_2}$ *for* $v \in C_{AB}$*:*

1. *If* $v^s_{A^TA} \in C_{AB}$ *then,*
$$\frac{||Av||_2}{||v||_2} \geq \sqrt{\lambda^s_{A^TA}},$$

2. *If* $v^s_{A^TA} \notin C_{AB}$ *then,*
$$\frac{||Av||_2}{||v||_2} \geq \min\Big\{\frac{||Av^u_A||_2}{||v^u_A||_2}, \frac{||Av^u_B||_2}{||v^u_B||_2}\Big\}.$$

Note that we can restate this lemma and proof to give us the upper and lower bounds of $\frac{||Bv||}{||v||}$ for $v \in C_{AB}$, by substituting $v^{u,s}_{B^TB}$ for $v^{u,s}_{A^TA}$ and $\lambda^{u,s}_{B^TB}$ for $\lambda^{u,s}_{A^TA}$ in the formulas.

Proof. First, let us consider the upper bounds. If $v^u_{A^TA} \in C_{AB}$, then we attain our upper bound for $||A||$ when $v = v^u_{A^TA} \in C_{AB}$, giving us $\max\{\frac{||Av||}{||v||}\} = \sqrt{\lambda^u_{A^TA}}$. Otherwise, $v^u_{A^TA} \notin C_{AB}$, and hence the spectral norm is monotonic on C_{AB}; we therefore attain our upper bound when v is a vector on the boundary of C_{AB}, which will be either v^u_A or v^u_B, depending on our choices of $\alpha, \beta, \gamma, \delta$.

Now, let us consider the lower bounds. If $v^s_{A^TA} \in C_{AB}$, then we attain our lower bound for $||A||$ when $v = v^s_{A^TA} \in C_{AB}$, giving us $\min\{\frac{||Av||}{||v||}\} = \sqrt{\lambda^s_{A^TA}}$. Otherwise, $v^s_{A^TA} \notin C_{AB}$, and the spectral norm is monotonic over C_{AB}; we therefore attain our lower bound when v is a vector on the boundary of C_{AB}, which will be either v^u_A or v^u_B, depending on our choices of $\alpha, \beta, \gamma, \delta$. □

We now consider an explicit example of calculating Φ_0 and Ψ_0 using only the l_2 norm.

Example 4.1. *Consider the case where we choose*

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

and

$$B\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix},$$

each with probability $\frac{1}{2}$*, at each iterate of our system* H*. The eigenvectors of* A *and* B *are*

$$v_A^{u,s} = \begin{pmatrix} 2 \\ 1 \pm \sqrt{5} \end{pmatrix},\ v_B^{u,s} = \begin{pmatrix} 2 \\ 2 \pm 2\sqrt{3} \end{pmatrix},$$

respectively, yielding the minimal, mutually invariant cone

$$C_{AB} = \Big\{(x,y) \in \mathbb{R}^2 : \frac{2}{2+2\sqrt{3}} \leq \frac{x}{y} \leq \frac{2}{1+\sqrt{5}}\Big\},$$

which is shown in Figure 23a. The eigenvectors of the matrices A^TA *and* B^TB *are*

$$v_{A^TA}^{u,s} = \begin{pmatrix} 2 \\ 1 \pm \sqrt{5} \end{pmatrix} = v_A^{u,s},$$

since A *is symmetric, and*

$$v_{B^TB}^{u,s} = \begin{pmatrix} 14 \\ 5 \pm \sqrt{221} \end{pmatrix}$$

respectively. Hence we obtain the following upper bounds:

$$\max_{v \in C_{AB}} \Big\{\frac{||Av||_2}{||v||_2}\Big\} = \sqrt{\lambda_{A^TA}^u} \approx 2.618,$$

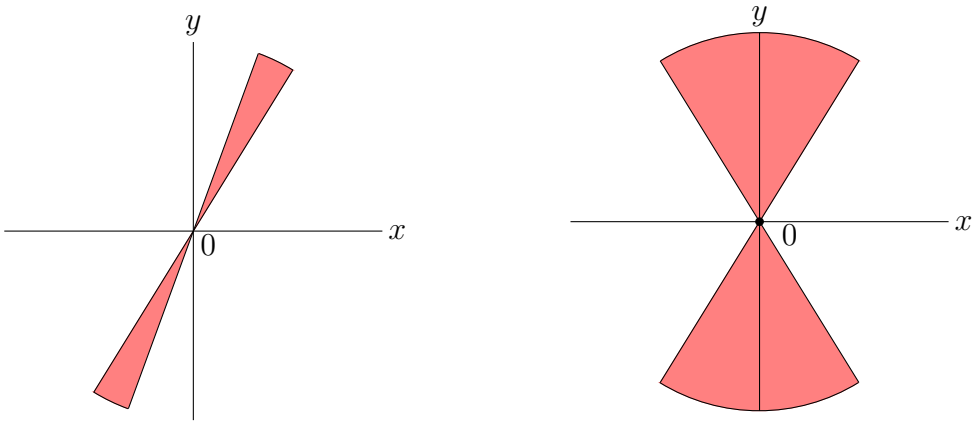
since $v_{A^TA}^u \in C_{AB}$*, and*

$$\max_{v \in C_{AB}} \Big\{\frac{||Bv||_2}{||v||_2}\Big\} = \max\Big\{\frac{||Bv_B^u||_2}{||v_B^u||_2}, \frac{||Bv_A^u||_2}{||v_A^u||_2}\Big\} = \frac{||Bv_A^u||_2}{||v_A^u||_2} \approx 3.857.$$

Similarly, we also obtain the following lower bounds:

$$\min_{v \in C_{AB}} \Big\{\frac{||Av||_2}{||v||_2}\Big\} = \frac{||Av_B^u||_2}{||v_B^u||_2} \approx 2.566,$$

$$\min_{v \in C_{AB}} \Big\{\frac{||Bv||_2}{||v||_2}\Big\} = \frac{||Bv_B^u||_2}{||v_B^u||_2} = \lambda_B^u = 2 + \sqrt{3} \approx 3.732.$$



(a) Example 4.1. (b) Example 4.2.

Figure 23: The cones C_{AB} from Examples 4.1 and 4.2. In Example 4.2 the cone is much wider, and this leads to the gap between the bounds Ψ_0 and Φ_0 being significantly larger than in Example 4.1.

Substituting these values into (66) and (67) yields

$$\Psi_0 = 1.130 \leq \lambda \leq 1.156 = \Phi_0 \text{ } \textit{(to 3.d.p.)}.$$

Using Gram-Schmidt orthonormalization to calculate λ over 500000 iterates of H yields

$$\lambda \approx 1.144.$$

The gap between the bounds in Example 4.1 is fairly small since the cone C_{AB} is fairly narrow. Specifically, using the estimate of λ as a baseline, we see a percentage error of -1.22% for Ψ_0 and $+1.05\%$ for Φ_0. The simplest upper bound is given by $\max\{\lambda_A, \lambda_B\} = \log(2+\sqrt{3}) = 1.317$, which is a percentage error of $+15.12\%$; we therefore see a significant improvement between Φ_0 and the simplest bound, by a factor of 14.4. Example 4.2 considers a case where the cone is wider.

Example 4.2. *Consider instead the case where we choose*

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

and

$$B\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

In Example 4.1 we calculated $v_A^{u,s}$ *and* $v_{A^TA}^{u,s}$*. Since* B *is also symmetric, we obtain*

$$v_B^{u,s} = v_{B^TB}^{u,s} = \begin{pmatrix} -2 \\ 1 \pm \sqrt{5} \end{pmatrix}.$$

The minimal, mutually invariant cone is

$$C_{AB} = \Big\{(x,y) \in \mathbb{R}^2 : \frac{-2}{1+\sqrt{5}} \leq \frac{x}{y} \leq \frac{2}{1+\sqrt{5}}\Big\},$$

which in this case contains both $v_{A^TA}^u$ *and* $v_{B^TB}^u$ *(see Figure 23b). Note further that this cone is much wider (i.e. contains a wider range of ratios of vectors) than the cone in Example 4.1. We obtain the upper bounds*

$$\max_{v \in C_{AB}} \Big\{\frac{||Av||_2}{||v||_2}\Big\} = \frac{||Av_A^u||_2}{||v_A^u||_2} = \lambda_A^u = \frac{3+\sqrt{5}}{2} \approx 2.618,$$

$$\max_{v \in C_{AB}} \Big\{\frac{||Bv||_2}{||v||_2}\Big\} = \frac{||Bv_B^u||_2}{||v_B^u||_2} = \lambda_B^u = \frac{3+\sqrt{5}}{2} \approx 2.618,$$

and the lower bounds

$$\min_{v \in C_{AB}} \Big\{\frac{||Av||_2}{||v||_2}\Big\} = \frac{||Av_B^u||_2}{||v_B^u||_2} \approx 1.220,$$

$$\min_{v \in C_{AB}} \Big\{\frac{||Bv||_2}{||v||_2}\Big\} = \frac{||Bv_A^u||_2}{||v_A^u||_2} \approx 1.220.$$

We therefore obtain the following bounds on λ*:*

$$\Psi_0 = 0.199 \leq \lambda \leq 0.962 = \Phi_0 \text{ (to 3.d.p.).}$$

Using Gram-Schmidt orthonormalization to calculate λ over 500000 iterates of H yields

$$\lambda \approx 0.662.$$

The percentage errors in this case, again using the estimate for λ as a baseline, are -69.94% for Ψ_0 and $+45.32\%$ for Φ_0, which are drastically larger than the errors we saw for Example 4.1; furthermore, there is no improvement over the simplest upper bound, given by $\max\{\lambda_A, \lambda_B\}$, in this particular case (this is due to the matrices A and B sharing the same eigenvalues).

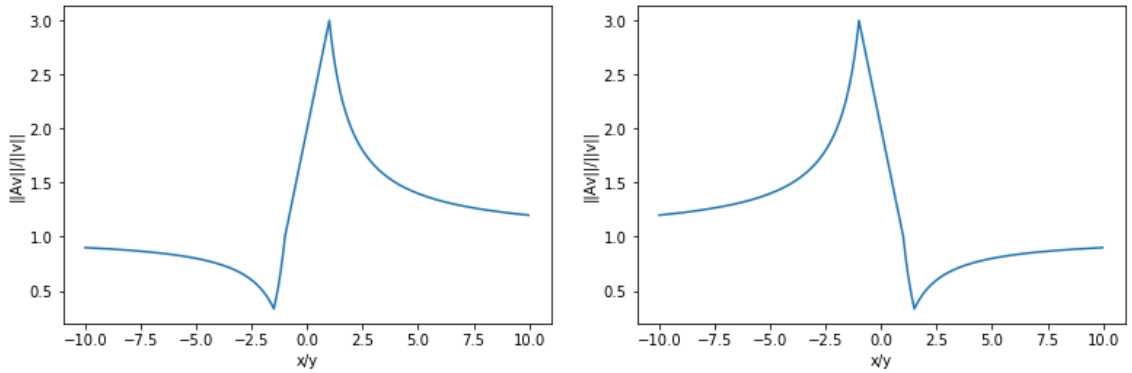
These examples demonstrate that the accuracy of the bounds Ψ_0 and Φ_0 depends on our choices of the matrices A and B, and in particular on the width of their mutually invariant cone.[7] The narrower the cone, the closer the bounds will be to the true Lyapunov exponent. The accuracy of the bounds does also depend on the number of iterates of H we consider at once - so far we have only considered finding bounds upon one iterate of H at a time, however we will later improve upon this by considering matrices of the form $M_{a,b} = A^a B^b$, which will yield more accurate bounds. However, we will still find that these improved bounds are limited by the cone C_{AB}, and we will discuss ways to try to improve upon this.

Let us now consider the l_∞ norm. We have

$$\frac{||Av||_\infty}{||v||_\infty} = \frac{\max\left\{|x + \alpha y|, |\beta x + (1 + \alpha\beta)y|\right\}}{\max\left\{|x|, |y|\right\}}. \tag{68}$$

Graphs of $\frac{||Av||_\infty}{||v||_\infty}$ against $\frac{x}{y}$ for particular choices of α and β are shown in Figure 24. The l_∞ norm is monotonic between its maximum and minimum values for all $v \in \mathbb{R}^2$, so in order to determine the maximum and minimum values within C_{AB}, we need to know whether the vectors which yield these values are in C_{AB}.

[7] So far we have only demonstrated this in the case of the l_2 norm, however this problem arises with the l_1 and l_∞ norms as well, and also if we pick the best bounds from any of these norms.



(a) $\alpha = \beta = 1$. (b) $\alpha = \beta = -1$.

Figure 24: Plots of the quantity $\frac{||Av||_\infty}{||v||_\infty}$ for vectors of the form $(x/y, 1)$.

Consider the unit ball centred at $(0, 0)$ for the l_∞ norm, B_{l_∞} (see Figure 25a). This ball contains all vectors v for which $||v||_\infty = 1$. In order to find the vectors which maximise and minimise $\frac{||Av||_\infty}{||v||_\infty}$, we need to find the vectors $v \in B_{l_\infty}$ for which $||Av||_\infty$ is maximised or minimised, since $||v||_\infty = 1$. The quantity $\frac{||Av||_\infty}{||v||_\infty}$ is then maximised or minimised along any scalar multiple of these vectors.

The image of B_{l_∞} has corners as follows:

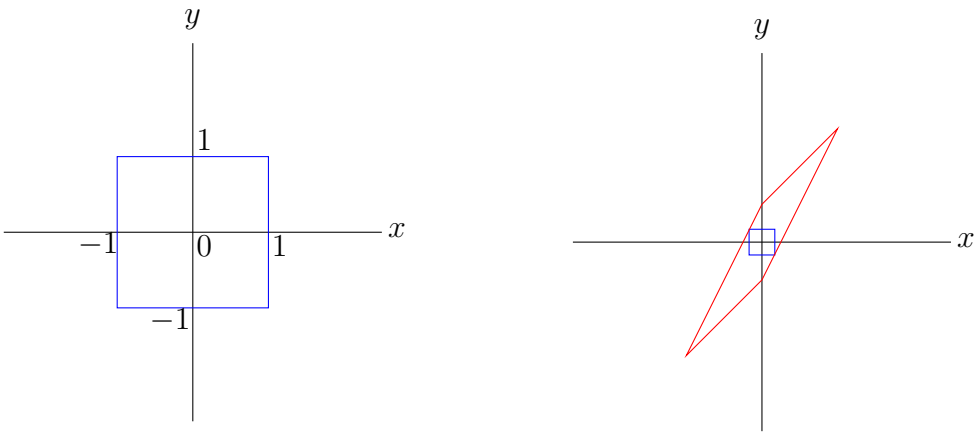
$$A\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}1 & \alpha\\ \beta & 1+\alpha\beta\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}1+\alpha\\ \beta+1+\alpha\beta\end{pmatrix},$$

$$A\begin{pmatrix}-1\\1\end{pmatrix} = \begin{pmatrix}\alpha-1\\ 1+\alpha\beta-\beta\end{pmatrix},$$

$$A\begin{pmatrix}-1\\-1\end{pmatrix} = \begin{pmatrix}-1-\alpha\\ -\beta-1-\alpha\beta\end{pmatrix},$$

$$A\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}1-\alpha\\ \beta-1-\alpha\beta\end{pmatrix}.$$

The following lemma uses the above to tell us the vectors v for which $\frac{||Av||_\infty}{||v||_\infty}$ is maximised and minimised for all permitted choices of α and β.



(a) The l_∞ norm unit ball, B_{l_∞}. (b) Extending B_{l_∞} when $\alpha = \beta = 1$.

Figure 25: In order to find the vector which minimises $||Av||_\infty$, we extend B_{l_∞} from the origin until it intersects AB_{l_∞}.

Lemma 4.4. *Let A be given by (53). Let $v_{\max}, v_{\min} \in \mathbb{R}^2$ be the vectors for which the quantity $\frac{||Av||_\infty}{||v||_\infty}$ is maximised and minimised for all $v \in \mathbb{R}^2$ respectively. Then:*

(i) if $\alpha, \beta > 0$,

$$v_{\max} = (1, 1), v_{\min} = (-\alpha\beta - \alpha - 1, \beta + 1).$$

(ii) if $\alpha, \beta < 0$,

$$v_{\max} = (-1, 1), v_{\min} = (\alpha\beta + \alpha + 1, \beta + 1).$$

(iii) if $\alpha < 0$, $\beta > 0$ and $\alpha\beta < -4$,

$$v_{\max} = (-1, 1), v_{\min} = (-\alpha\beta - \alpha - 1, \beta + 1).$$

(iv) if $\alpha > 0$, $\beta < 0$ and $\alpha\beta < -4$,

$$v_{\max} = (1, 1), v_{\min} = (\alpha\beta + \alpha + 1, \beta + 1).$$

Proof. In all cases, the maximum x and y coordinates (in modulus) of AB_{l_∞} are attained at its corners, or in other words, at the image of one of the corners of

B_{l_∞}. This means our maximum value for $||Av||_\infty$ for $v \in B_{l_\infty}$ is always attained by either the vector $(1,1)$ (and $(-1,-1)$), or $(-1,1)$ (and $(1,-1)$). In the cases where $\alpha > 0$, we have that the largest x coordinate in modulus is $1+\alpha$, and the largest y coordinate in modulus is $(\beta + 1 + \alpha\beta)$, yielding $v_{\max} = (1,1)$ (or $(-1,-1)$). When $\alpha < 0$, we instead have that the largest x coordinate in modulus is $1-\alpha$, and the largest y coordinate in modulus is $(1+\alpha\beta-\beta)$, yielding $v_{\max} = (-1,1)$ (or $(1,-1)$).

The minimum in each of these cases is attained by expanding an l_∞ ball from the origin until it meets with the image of AB_{l_∞} (see Figure 25b). The point at which the two meet yields the vector v whose pre-image is $A^{-1}v = v_{\min}$. The first point of the expanding ball to intersect with AB_{l_∞} will be a corner[8], and so will be the pre-image of either $(1,1)$ or $(-1,1)$. In the cases where $\beta > 0$, we find $v_{\min} = A^{-1}(-1,1)^T$, and in the cases where $\beta < 0$, we find $v_{\min} = A^{-1}(1,1)^T$. □

In order to determine what the maximum and minimum values for $\frac{||Av||_\infty}{||v||_\infty}$ are for $v \in C_{AB}$, we first check if either (or both) of $v_{\max}$ or $v_{\min}$ is contained in C_{AB}. If so, then the respective maximum or minimum will occur at that vector. We can then take our other bound (if needed) to be the minimum or maximum respectively of the boundary vectors to C_{AB}, since the l_∞ norm is monotonic between $v_{\max}$ and $v_{\min}$. This idea is summarised in the following lemma.

Lemma 4.5. *Let A and B be given by (53) and (54) respectively, and let $v_{\max}$ and $v_{\min}$, given by Lemma 4.4, be the vectors which maximise and minimise $\frac{||Av||_\infty}{||v||_\infty}$ respectively for all $v \in \mathbb{R}^2$. Then we obtain the following upper bounds on $\frac{||Av||_\infty}{||v||_\infty}$ for $v \in C_{AB}$:*

U1. *If $v_{\max} \in C_{AB}$, then*

$$\frac{||Av||_\infty}{||v||_\infty} \leq \frac{||Av_{\max}||_\infty}{||v_{\max}||_\infty}.$$

[8]Unless the slope of the edges between subsequent corners of AB_{l_∞} is either 0 or ∞, neither of which are possible given our restrictions on α and β.

U2. *If* $v_{\max} \notin C_{AB}$*, then*

$$\frac{||Av||_\infty}{||v||_\infty} \leq \max\Big\{\frac{||Av_A^u||_\infty}{||v_A^u||_\infty}, \frac{||Av_B^u||_\infty}{||v_B^u||_\infty}\Big\}.$$

We also obtain the following lower bounds on $\frac{||Av||_\infty}{||v||_\infty}$ *for* $v \in C_{AB}$*:*

L1. *If* $v_{\min} \in C_{AB}$*, then*

$$\frac{||Av||_\infty}{||v||_\infty} \geq \frac{||Av_{\min}||_\infty}{||v_{\min}||_\infty}.$$

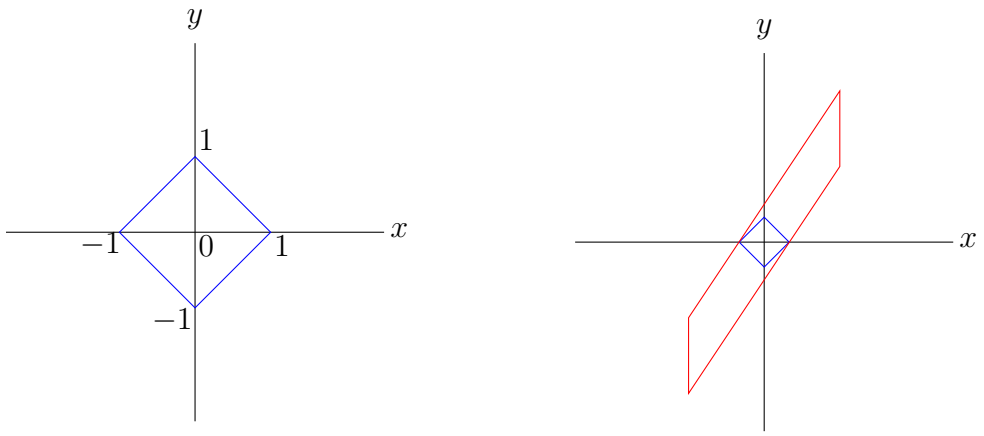
L2. *If* $v_{\min} \notin C_{AB}$*, then*

$$\frac{||Av||_\infty}{||v||_\infty} \geq \min\Big\{\frac{||Av_A^u||_\infty}{||v_A^u||_\infty}, \frac{||Av_B^u||_\infty}{||v_B^u||_\infty}\Big\}.$$

Proof. Clearly if the condition for **U1.** is met then, by Lemma 4.4, $\frac{||Av||_\infty}{||v||_\infty}$ is maximised at $v_{\max}$, and similarly if the condition for **L1.** is met, then $\frac{||Av||_\infty}{||v||_\infty}$ is minimised at $v_{\min}$. The l_∞ norm is monotonic between these vectors, and so if one (or both) of these vectors is not located within C_{AB}, we find the maximum (or minimum) for $v \in C_{AB}$ by taking the maximum (or minimum) of the value of $\frac{||Av||_\infty}{||v||_\infty}$ for v on the boundaries of C_{AB}, or in other words for $v \in \{v_A^u, v_B^u\}$. □

We now undergo a similar process to obtain upper and lower bounds on $\frac{||Av||_1}{||v||_1}$, that is, using the l_1 norm instead of the l_∞ norm. We have

$$\frac{||Av||_1}{||v||_1} = \frac{|x + \alpha y| + |\beta x + (1 + \alpha\beta)y|}{|x| + |y|}. \tag{69}$$



(a) The l_1 norm unit ball, B_{l_1}.

(b) Extending B_{l_1} when $\alpha = \beta = 1$.

Figure 26: In order to find the vector which minimises $||Av||_1$, we extend B_{l_1} from the origin until it intersects AB_{l_1}.

We again consider the unit ball centred around $(0, 0)$, now for the l_1 norm, which we label B_{l_1} (see Figure 26a). The corners of the image, AB_{l_1}, are as follows:

$$A\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\\beta\end{pmatrix},$$

$$A\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}\alpha\\1+\alpha\beta\end{pmatrix},$$

$$A\begin{pmatrix}-1\\0\end{pmatrix} = \begin{pmatrix}-1\\-\beta\end{pmatrix},$$

$$A\begin{pmatrix}0\\-1\end{pmatrix} = \begin{pmatrix}-\alpha\\-1-\alpha\beta\end{pmatrix}.$$

The following lemma uses the above to tell us the vectors v for which $\frac{||Av||_1}{||v||_1}$ is maximised and minimised for all permitted choices of α and β.

Lemma 4.6. *Let A be given by (53). Let $v_{\max}, v_{\min} \in \mathbb{R}^2$ be the vectors upon which the quantity $\frac{||Av||_1}{||v||_1}$ is maximised and minimised for all $v \in \mathbb{R}^2$ respectively. Then:*

(i) if $\alpha\beta > 0$ *and* $|\alpha| \geq \frac{|\beta|}{|\beta|+1}$,

$$v_{\max} = (0, 1), v_{\min} = (\alpha\beta + 1, -\beta).$$

(ii) if $\alpha\beta > 0$ *and* $|\alpha| \leq \frac{|\beta|}{|\beta|+1}$,

$$v_{\max} = (1, 0), v_{\min} = (\alpha\beta + 1, -\beta).$$

(iii) if $\alpha\beta < -4$ *and* $|\alpha| \geq \frac{|\beta|+2}{|\beta|+1}$,

$$v_{\max} = (0, 1), v_{\min} = (\alpha\beta + 1, -\beta).$$

(iv) if $\alpha\beta < -4$ *and* $|\alpha| \leq \frac{|\beta|+2}{|\beta|+1}$,

$$v_{\max} = (1, 0), v_{\min} = (\alpha\beta + 1, -\beta).$$

Proof. When calculating where we obtain $v_{\max}$ for the various cases, we note that this maximum must be attained at one of the corners of AB_{l_1}. In the case where the slope of the edge between subsequent corners (for example, the corners $(\alpha, 1 + \alpha\beta)$ and $(1, \beta)$ when $\alpha, \beta > 0$) is equal to one, we would still attain a maximum at these corners, but also at any point on the edge connecting them - however this case is not possible given our restrictions on α and β.

When $\alpha\beta > 0$, we have that

$$1 + |\beta| \geq |\alpha| + |1 + \alpha\beta|,$$

provided

$$|\alpha| \leq \frac{|\beta|}{|\beta| + 1}.$$

Thus the maximum is attained at the pre-image of $(1, \beta)$ - $(1, 0)$ - in this case. Otherwise, we obtain our maximum at the pre-image of $(\alpha, 1+\alpha\beta)$ - $(0, 1)$. Similarly, when $\alpha\beta < -4$, we have that

$$1 + |\beta| \geq | - \alpha| + | - 1 - \alpha\beta|,$$

provided

$$|\alpha| \leq \frac{|\beta|+2}{|\beta|+1},$$

and the maximum is attained at $(1,0)$ in this case, and $(0,1)$ otherwise.

To obtain the minimums, we note that the slopes of the edges of AB_{l_1} which cross the y-axis are always greater than one in modulus, and thus the first vector of an l_1 ball to intersect them will be $(1,0)$. Hence, in all cases, $v_{\min} = A^{-1}(1,0) = (\alpha\beta+1, -\beta)$. □

We now state the analogous result to Lemma 4.5 for the l_1 norm.

Lemma 4.7. *Let A and B be given by (53) and (54) respectively, and let $v_{\max}$ and $v_{\min}$, given by Lemma 4.6, be the vectors which maximise and minimise $\frac{||Av||_1}{||v||_1}$ respectively for all $v \in \mathbb{R}^2$. Then we obtain the following upper bounds on $\frac{||Av||_1}{||v||_1}$ for $v \in C_{AB}$:*

U1 *If $v_{\max} \in C_{AB}$, then*

$$\frac{||Av||_1}{||v||_1} \leq \frac{||Av_{\max}||_1}{||v_{\max}||_1}.$$

U2 *If $v_{\max} \notin C_{AB}$, then*

$$\frac{||Av||_1}{||v||_1} \leq \max\Big\{\frac{||Av_A^u||_1}{||v_A^u||_1}, \frac{||Av_B^u||_1}{||v_B^u||_1}\Big\}.$$

We also obtain the following lower bounds on $\frac{||Av||_1}{||v||_1}$ for $v \in C_{AB}$:

L1 *If $v_{\min} \in C_{AB}$, then*

$$\frac{||Av||_1}{||v||_1} \geq \frac{||Av_{\min}||_1}{||v_{\min}||_1}.$$

L2 *If* $v_{\min} \notin C_{AB}$*, then*

$$\frac{||Av||_1}{||v||_1} \geq \min\Big\{\frac{||Av_A^u||_1}{||v_A^u||_1}, \frac{||Av_B^u||_1}{||v_B^u||_1}\Big\}.$$

Proof. Noting that the l_1 norm, like the l_∞ norm, is monotonic between its maximum and minimum, the proof of this lemma is identical to that of Lemma 4.5. □

We can calculate the bounds Φ_0 and Ψ_0 for any of these norms and obtain rigorous upper and lower bounds upon λ. Thus, by choosing the smallest Φ_0 and largest Ψ_0, we obtain the best bounds upon λ using these three norms. In the following example we return to the systems we considered in Examples 4.1 and 4.2, calculate Φ_0 and Ψ_0 for the l_∞ and l_1 norms using the above lemmas, and compare these to what we obtained when using the l_2 norm to see if any improvement is made. Note that we use the notation $\Phi_0^{(k)}$ and $\Psi_0^{(k)}$ for $k \in \{1, 2, \infty\}$ to indicate Φ_0 and Ψ_0 evaluated using the l_k norm.

Example 4.3. *First, consider the system studied in Example 4.1. We have* $\alpha, \beta > 0$*, thus for A,* $v_{A\max_\infty} = (1, 1)$ *and* $v_{A\min_\infty} = (-3, 2)$*. Furthermore, we have*

$$|\alpha| = 1 \geq \frac{|\beta|}{|\beta| + 1} = \frac{1}{2},$$

so $v_{A\max_1} = (0, 1)$ *and* $v_{A\min_1} = (2, -1)$*. These conditions also hold for B, and we obtain* $v_{B\max_\infty} = (1, 1)$*,* $v_{B\min_\infty} = (-4, 3)$*,* $v_{B\max_1} = (0, 1)$*, and* $v_{B\min_1} = (3, -2)$*. We note that none of these vectors are contained within the cone* C_{AB}*, so our maximum and minimum are attained upon the boundaries of the cone for both* $\frac{||Av||}{||v||}$ *and* $\frac{||Bv||}{||v||}$ *using both norms. Taking into account whether the norms are increasing or decreasing between* v_B^u *and* v_A^u*, the* l_∞ *norm yields the upper bound*

$$\begin{aligned}\Phi_0^{(\infty)} &= \frac{1}{2}\log\frac{||Av_A^u||_\infty}{||v_A^u||_\infty} + \frac{1}{2}\log\frac{||Bv_A^u||_\infty}{||v_A^u||_\infty} \\ &= \frac{1}{2}\log(\frac{3+\sqrt{5}}{2}) + \frac{1}{2}\log(2+\sqrt{5}) \cong 1.203,\end{aligned}$$

and the lower bound

$$\begin{aligned}\Psi_0^{(\infty)} &= \frac{1}{2}\log\frac{||Av_B^u||_\infty}{||v_B^u||_\infty} + \frac{1}{2}\log\frac{||Bv_B^u||_\infty}{||v_B^u||_\infty} \\ &= \frac{1}{2}\log(\frac{3+\sqrt{3}}{2}) + \frac{1}{2}\log(2+\sqrt{3}) \approxeq 1.089,\end{aligned}$$

while the l_1 norm yields the upper bound

$$\begin{aligned}\Phi_0^{(1)} &= \frac{1}{2}\log\frac{||Av_B^u||_1}{||v_B^u||_1} + \frac{1}{2}\log\frac{||Bv_B^u||_1}{||v_B^u||_1} \\ &= \frac{1}{2}\log(1+\sqrt{3}) + \frac{1}{2}\log(2+\sqrt{3}) \approxeq 1.161,\end{aligned}$$

and the lower bound

$$\begin{aligned}\Psi_0^{(1)} &= \frac{1}{2}\log\frac{||Av_A^u||_1}{||v_A^u||_1} + \frac{1}{2}\log\frac{||Bv_A^u||_1}{||v_A^u||_1} \\ &= \frac{1}{2}\log(\frac{3+\sqrt{5}}{2}) + \frac{1}{2}\log(\frac{5+\sqrt{5}}{2}) \approxeq 1.124.\end{aligned}$$

Thus, in this case, the bounds obtained via the l_1 and l_∞ norms are worse than those obtained via the l_2 norm.

Now consider the system studied in Example 4.2. We now have that $\gamma, \delta < 0$, so $v_{B\max_\infty} = (-1, 1)$ and $v_{B\min_\infty} = (1, 0)$. We still have

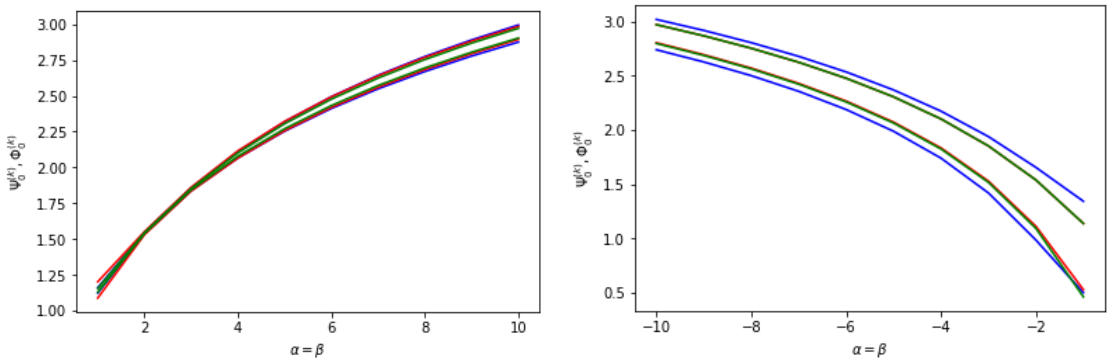
$$|\gamma| = 1 \geq \frac{|\delta|}{|\delta|+1} = \frac{1}{2},$$

so $v_{B\max_1} = (0, 1)$ and $v_{B\min_1} = (2, 1)$. Note that both of $v_{A\max_1}$ and $v_{B\max_1}$ are contained within C_{AB} in this case. The l_∞ norm yields the upper bound

$$\begin{aligned}\Phi_0^{(\infty)} &= \frac{1}{2}\log\frac{||Av_A^u||_\infty}{||v_A^u||_\infty} + \frac{1}{2}\log\frac{||Bv_B^u||_\infty}{||v_B^u||_\infty} \\ &= \log(\frac{3+\sqrt{5}}{2}) \approxeq 0.962,\end{aligned}$$

and the lower bound

$$\begin{aligned}\Psi_0^{(\infty)} &= \frac{1}{2}\log\frac{||Av_B^u||_\infty}{||v_B^u||_\infty} + \frac{1}{2}\log\frac{||Bv_A^u||_\infty}{||v_A^u||_\infty} \\ &= \log(\frac{5-\sqrt{5}}{2}) \approxeq 0.324,\end{aligned}$$



(a) $\alpha = \beta \in \{1, \ldots, 10\}$. (b) $\alpha = \beta \in \{-10, \ldots, -1\}$.

Figure 27: $\Phi_0^{(k)}$ and $\Psi_0^{(k)}$, $k \in \{1, 2, \infty\}$, for $\gamma = 1$, $\delta = 2$, for various $\alpha = \beta$.

while the l_1 norm yields the upper bound

$$\begin{aligned}\Phi_0^{(1)} &= \frac{1}{2}\log\frac{||Av_{A\max_1}||_1}{||v_{A\max_1}||_1} + \frac{1}{2}\log\frac{||Bv_{B\max_1}||_1}{||v_{B\max_1}||_1} \\ &= \log(3) \approxeq 1.099,\end{aligned}$$

and the lower bound

$$\begin{aligned}\Psi_0^{(1)} &= \frac{1}{2}\log\frac{||Av_B^u||_1}{||v_B^u||_1} + \frac{1}{2}\log\frac{||Bv_B^u||_1}{||v_B^u||_1} \\ &= \frac{1}{2}\log(\frac{-9+5\sqrt{5}}{2}) + \frac{1}{2}\log(\frac{3+\sqrt{5}}{2}) \approxeq 0.524.\end{aligned}$$

Thus we obtain no improvement upon Φ_0, however the l_1 norm provides the best Ψ_0 in this case.

Figure 27 shows Φ_0 and Ψ_0 for each of the l_2, l_1 and l_∞ norms, fixing $\gamma = 1$, $\delta = 2$, for a range of values of $\alpha = \beta$. The cones C_{AB} are much wider for the cases shown in Figure 27b, and we see that the gap between the bounds is much wider as a result.

In the upcoming sections we will improve upon the bounds Φ_0 and Ψ_0 using two methods. In Section 4.5 we will consider looking at longer chains of matrices, of the form $M_{a,b} = A^a B^b$, in order to take into account the differing growths rates

that may occur under these longer chains. Furthermore, this method will allow us to capture more of the qualities of the underlying distribution than simply looking at one iterate at a time, as we did with Φ_0 and Ψ_0. In Section 4.6 we will try to eliminate the large gap that we observe between the bounds as the width of the cone C_{AB} widens. To do this, we will discuss a way to look at narrower cones - subsets of C_{AB} - for each of the terms in our bounds, by considering the matrices which precede each $M_{a,b}$.

4.5 Improving the bounds

In this section we will improve upon the accuracy of the bounds Φ_0 and Ψ_0 by considering multiple iterates of our infinite sequence of A's and B's at once. To do this, we will split up the sequence of matrices in a particular way, as suggested in [54]. Recall equation (61). If, instead of splitting $H^n v_0$ into n individual iterates, we split it into all sequences of the form

$$M_{a,b} = A^a B^b \tag{70}$$

we obtain

$$\begin{aligned} H^n v_0 &= AABABB \ldots AABBABv_0, \\ &= (AAB) \cdot (ABB) \cdot \ldots \cdot (AABB) \cdot (AB) v_0, \\ &= M_{a_N,b_N} M_{a_{N-1},b_{N-1}} \ldots M_{a_1,b_1} v_0. \end{aligned} \tag{71}$$

and thus

$$||H^n v_0|| = \frac{||M_{a_N,b_N} v_{n-1}||}{||v_{n-1}||} \cdot \ \ldots \ \cdot \frac{||M_{a_1,b_1} v_0||}{||v_0||} \cdot ||v_0||, \tag{72}$$

where $v_i = M_{a_i,b_i} v_{i-1}$ for $i \in \{0, \ldots, n\}$. Thus in order to obtain bounds upon $||H^n v_0||$, we need to bound $\frac{||M_{a_i,b_i} v_{i-1}||}{||v_{i-1}||}$ for each i. We are trying to obtain bounds upon the Lyapunov exponent, and thus are concerned with what happens as $n \to \infty$. In the infinite limit the sequence H^n will contain every possible sequence $M_{a,b}$ for

$a, b \geq 1$, each with an associated frequency of occurrence. We therefore need to obtain bounds upon $\frac{||M_{a,b}v||}{||v||}$ for each a and b, for $v \in C_{AB}$, and find the frequency with which each of the sequence occurs within the infinite sequence. Note that in [54], matrices of the form $M_{a,b}$ were required in order to ensure hyperbolicity; in our case, we do this to ensure that the diagonalisation of the matrices can be written in a general form.

It is worth noting that, in (72), $N \neq n$. To find a relationship between n and N, we look at the average length of the matrices $M_{a,b}$. In this case, length means the number of matrices multiplied together to create $M_{a,b}$, and is equal to $a+b$. We consider a sequence m of A's and B's which starts with a B. b is the number of B's we obtain before we get an A, and at each step we obtain either A or B with probability $1/2$, hence

$$\mathbb{P}(b = 1) = \mathbb{P}(m = \ldots AB | m = \ldots B) = 1/2,$$

$$\mathbb{P}(b = 2) = \mathbb{P}(m = \ldots ABB | m = \ldots B) = 1/4,$$

$$\mathbb{P}(b = 3) = \mathbb{P}(m = \ldots ABBB | m = \ldots B) = 1/8,$$

$$\vdots$$

$$\mathbb{P}(b = N) = \mathbb{P}(m = \ldots AB^N | m = \ldots B) = 1/2^N.$$

Thus,

$$\mathbb{E}(b) = \sum_{i=1}^{\infty} \mathbb{P}(b = i) \cdot i = \frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \ \ldots \ = \sum_{i=1}^{\infty} \frac{i}{2^i} = 2.$$

Via a similar calculation, this time starting m with an A instead of a B, we obtain $\mathbb{E}(a) = 2$, and hence

$$\mathbb{E}(\text{Length of } M_{a,b}) = \mathbb{E}(a + b) = 4.$$

From this we deduce that, as $n \to \infty$,

$$N = \frac{n}{4},$$

and thus

$$\lambda = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}(\log ||H^n v_0||) = \lim_{N\to\infty} \frac{1}{4N}\mathbb{E}\Big(\log \big(\prod_{i=1}^{N} \frac{||M_{a_i,b_i} v_{i-1}||}{||v_{i-1}||}\big)\Big). \tag{73}$$

We choose A or B at each iterate by means of a simple Bernoulli distribution, and so the expected frequency with which each $M_{a,b}$ occurs is trivial to calculate: 2^{-a-b}. Using this we can simplify (73) to obtain our bounds on λ; this is summarised by the following theorem.

Theorem 4.3. *Let H be the random dynamical system on the measurable space $(\mathbb{T}^2, \mathbb{T}^2)$ over the metric dynamical system $(\{A, B\}, \mathbb{T}^2, \{\frac{1}{2}, \frac{1}{2}\}, (\theta(t))_{t\in\mathbb{Z}})$ given by $H : \mathbb{Z} \times \{A, B\} \times \mathbb{T}^2 \to \mathbb{T}^2$, where A and B are given by (53) and (54) respectively. Let $M_{a,b} = A^a B^b$. For $v \in \mathbb{R}^2$ the maximal Lyapunov exponent λ is given by*

$$\lambda = \frac{1}{4} \sum_{a,b=1}^{\infty} 2^{-a-b} \log \big(\frac{||M_{a,b} v||}{||v||}\big). \tag{74}$$

Let C_{AB} be the minimal mutually invariant cone for A and B, given by Theorem 4.1. Then an upper bound on λ is given by

$$\Phi_1 = \frac{1}{4} \sum_{a,b=1}^{\infty} 2^{-a-b} \log \big(\max_{v\in C_{AB}} \Big\{ \frac{||M_{a,b} v||}{||v||} \Big\}\big), \tag{75}$$

and a lower bound on λ is given by

$$\Psi_1 = \frac{1}{4} \sum_{a,b=1}^{\infty} 2^{-a-b} \log \big(\min_{v\in C_{AB}} \Big\{ \frac{||M_{a,b} v||}{||v||} \Big\}\big). \tag{76}$$

In the above theorem, (75) and (76) are obtained by observing that in systems of this form, all vectors are eventually orientated within C_{AB}; this includes the stable eigenvectors of DA and DB, as neither are invariant under H. We can therefore restrict the range of vectors considered in the calculation of λ from the entirety of $\mathbb{R}^2$ to just C_{AB}, as the time taken for a particular vector to enter C_{AB} will be finite, and thus will not contribute to λ; we then obtain bounds by taking the maximum and minimum growth rates within this cone.

It is worth noting that, when computing these bounds, we truncate the infinite sum, and so the *numerical* result we would obtain for Φ_1 is not a rigorous upper bound on λ, but more of an estimate of the infinite sum. Specifically, the truncation

$$\Phi_{1_n} = \frac{1}{4}\sum_{a,b=1}^{n} \frac{2^{-a-b}}{\log}\Big(\max_{v\in C_{AB}}\Big\{\frac{||M_{a,b}v||}{||v||}\Big\}\Big)$$

converges to Φ_1 exponentially quickly as $n \to \infty$. On the other hand, the truncation of Ψ_1 *is* a rigorous lower bound. For 'large' a and b, the contribution from individual terms becomes smaller - we see exponential growth rates in the term $\frac{||M_{a,b}v||}{||v||}$ as we increase a and b, however, due to taking the log of this term, 2^{-a-b} is the dominating term for large a and b. Thus, if we calculate all terms in the sum to 'large enough' a and b, then the error between the truncation and Φ_1 will be small. We will then replace the remaining terms in the sequence with a 'remainder' term, which we discuss later (see (77)). This remainder term will ensure that the numerical result we obtain for Φ_1 is a rigorous upper bound.

In order to find bounds upon $\frac{||M_{a,b}v||}{||v||}$, we need to calculate $M_{a,b}$ for each a and b. To do this, we need to find both A^a and B^b. These are obtained by diagonalizing A and B so that

$$A = P\Lambda_A P^{-1} \text{ and } B = Q\Lambda_B Q^{-1},$$

where

$$P = \begin{pmatrix} 2\alpha & 2\alpha \\ \alpha\beta + \sqrt{\alpha\beta(4+\alpha\beta)} & \alpha\beta - \sqrt{\alpha\beta(4+\alpha\beta)} \end{pmatrix},$$

is a matrix consisting of the eigenvectors of A, and

$$\Lambda_A = \begin{pmatrix} \lambda_{A+} & 0 \\ 0 & \lambda_{A-} \end{pmatrix}.$$

The definitions for Q and Λ_B are similar; we simply substitute B for A, γ for α and

δ for β in the formulas. Raising these matrices to the appropriate powers gives us

$$A^a = \frac{1}{\alpha(\lambda_{A-} - \lambda_{A+})} \begin{pmatrix} \alpha & \alpha \\ \lambda_{A+} - 1 & \lambda_{A-} - 1 \end{pmatrix} \begin{pmatrix} \lambda_{A+}^a & 0 \\ 0 & \lambda_{A-}^a \end{pmatrix} \begin{pmatrix} \lambda_{A-} - 1 & -\alpha \\ 1 - \lambda_{A+} & \alpha \end{pmatrix},$$

and

$$B^b = \frac{1}{\gamma(\lambda_{B-} - \lambda_{B+})} \begin{pmatrix} \gamma & \gamma \\ \lambda_{B+} - 1 & \lambda_{B-} - 1 \end{pmatrix} \begin{pmatrix} \lambda_{B+}^b & 0 \\ 0 & \lambda_{B-}^b \end{pmatrix} \begin{pmatrix} \lambda_{B-} - 1 & -\gamma \\ 1 - \lambda_{B+} & \gamma \end{pmatrix}.$$

Multiplying A^a and B^b together yields $M_{a,b}$. Due to the length of the expressions for each entry of this matrix, we list it entry by entry:

$$\begin{aligned}
\Lambda \cdot M_{a,b(1,1)} &= \Big(\lambda_{A+}^a(\lambda_{A-} - 1) + \lambda_{A-}^a(1 - \lambda_{A+})\Big) \\
&\quad \cdot \Big(\lambda_{B+}^b(\lambda_{B-} - 1) + \lambda_{B-}^b(1 - \lambda_{B+})\Big) + \alpha\delta(\lambda_{A-}^a - \lambda_{A+}^a)(\lambda_{B-}^b - \lambda_{B+}^b), \\
\Lambda \cdot M_{a,b(1,2)} &= \gamma(\lambda_{B-}^b - \lambda_{B+}^b)\Big(\lambda_{A+}^a(\lambda_{A-} - 1) + \lambda_{A-}^a(1 - \lambda_{A+})\Big) \\
&\quad + \alpha(\lambda_{A-}^a - \lambda_{A+}^a)\Big(\lambda_{B-}^b(\lambda_{B-} - 1) + \lambda_{B+}^b(1 - \lambda_{B+})\Big), \\
\Lambda \cdot M_{a,b(2,1)} &= \delta(\lambda_{B-}^b - \lambda_{B+}^b)\Big(\lambda_{A-}^a(\lambda_{A-} - 1) + \lambda_{A+}^a(1 - \lambda_{A+})\Big) \\
&\quad + \beta(\lambda_{A-}^a - \lambda_{A+}^a)\Big(\lambda_{B+}^b(\lambda_{B-} - 1) + \lambda_{B-}^b(1 - \lambda_{B+})\Big), \\
\Lambda \cdot M_{a,b(2,2)} &= \Big(\lambda_{A-}^a(\lambda_{A-} - 1) + \lambda_{A+}^a(1 - \lambda_{A+})\Big) \\
&\quad \cdot \Big(\lambda_{B-}^b(\lambda_{B-} - 1) + \lambda_{B+}^b(1 - \lambda_{B+})\Big) + \beta\gamma(\lambda_{A-}^a - \lambda_{A+}^a)(\lambda_{B-}^b - \lambda_{B+}^b),
\end{aligned}$$

where $\Lambda = (\lambda_{A-} - \lambda_{A+})(\lambda_{B-} - \lambda_{B+})$. Note that $M_{a,b(1,1)}$ has the smallest modulus of these quantities, while $M_{a,b(2,2)}$ has the largest modulus (for large a and b). These formulas apply to all cases of $\alpha, \beta, \gamma, \delta$ we consider; however, note that the stabilities of the eigenvalues differ for different cases.

We already know an invariant cone for $M_{a,b}$, or indeed any product of A and B: C_{AB}. However, it is possible to narrow this cone further in the case of the matrices $M_{a,b}$. This will help to narrow the gap we observed between the upper and lower bounds somewhat, by narrowing all cones C_{AB} slightly; however, a gap will still

remain, albeit smaller. In Section 4.6, we will consider a method which causes this gap to tend to zero as we take more terms in our bounds.

Lemma 4.8. *Fix $a \in \mathbb{N}$. The cone bounded by v_A^u and $A^a v_B^u$ is an invariant cone for $M_{a,b}$ for all $b \in \mathbb{N}$, and is the minimal cone which is mutually invariant for all choices of b. In particular, the cone bounded by v_A^u and Av_B^u, $\tilde{C}_{AB}$, is the global, minimal, invariant cone for all matrices $M_{a,b}$; that is, for all $a, b \in \mathbb{N}$.*

The cone $\tilde{C}_{AB}$ is shown (qualitatively) in Figure 28 for the case when $\alpha, \beta, \gamma, \delta$ are integers greater than zero.

Proof. Consider $M_{a,b}v = A^a B^b v$ for a vector $v \in C_{AB}$. Lemma 4.1 tells us that applications of B to a vector v will cause it to align with v_B^u. If we let $b \to \infty$, then $B^b v$ can be as close to v_B^u as we like. In order to apply $M_{a,b}$, we then need to apply A^a, and thus our vector will begin to align with v_A^u. The furthest away our vector can be from v_A^u after a applications of A is $A^a v_B^u$. Thus the minimal cone we can take is the cone whose boundaries are v_A^u and $A^a v_B^u$, which we shall label as C_a. We have that

$$v_A^u \in \ \ldots \ \subset C_4 \subset C_3 \subset C_2 \subset C_1 = \tilde{C}_{AB},$$

so if we wish our cone to be invariant for any choice of a, we must choose $\tilde{C}_{AB}$, which by definition is bounded by the vectors v_A^u and Av_B^u. □

We now consider the upper and lower bounds upon $\frac{||M_{a,b}v||}{||v||}$ for the l_2, l_1 and l_∞ norms, within the cone $\tilde{C}_{AB}$. Note that Sturman and Thiffeault [54] provided similar results to these in the case of random products of shear matrices. First, we consider the l_2 norm. In the following lemma, $\lambda_{M_{a,b}^T M_{a,b}^{u,s}}$ and $v_{M_{a,b}^T M_{a,b}^{u,s}}$ refer to the unstable and stable eigenvalues and eigenvectors of $M_{a,b}^T M_{a,b}$ respectively.

Lemma 4.9. *Let A and B be matrices of the form given by (53) and (54), and let $\tilde{C}_{AB}$ be the cone given by Lemma 4.8. Then we obtain the following upper bounds for $\frac{||M_{a,b}v||_2}{||v||_2}$ for $v \in \tilde{C}_{AB}$:*

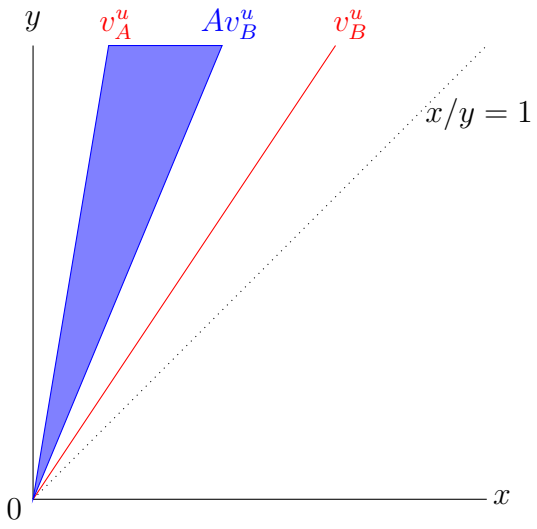


Figure 28: The cones C_{AB} and $\tilde{C}_{AB}$ in tangent space, for the case when $\alpha, \beta, \gamma, \delta$ are integers greater than zero; C_{AB} is the region between v_A^u and v_B^u, and $\tilde{C}_{AB}$ is the region between v_A^u and Av_B^u.

1. *If* $v_{M_{a,b}^T M_{a,b}^u} \in \tilde{C}_{AB}$ *then,*

$$\frac{||M_{a,b}v||_2}{||v||_2} \leq \sqrt{\lambda^u_{M_{a,b}^T M_{a,b}}}.$$

2. *If* $v_{M_{a,b}^T M_{a,b}^u} \notin \tilde{C}_{AB}$ *then,*

$$\frac{||M_{a,b}v||_2}{||v||_2} \leq \max\left\{\frac{||M_{a,b}v_A^u||_2}{||v_A^u||_2}, \frac{||M_{a,b}Av_B^u||_2}{||Av_{B+}||_2}\right\}.$$

We also obtain the following lower bounds for $\frac{||Av||_2}{||v||_2}$ *for* $v \in \tilde{C}_{AB}$*:*

1. *If* $v_{M_{a,b}^T M_{a,b}^s} \in \tilde{C}_{AB}$ *then,*

$$\frac{||M_{a,b}v||_2}{||v||_2} \geq \sqrt{\lambda^s_{M_{a,b}^T M_{a,b}}}.$$

2. *If* $v_{M_{a,b}^T M_{a,b}^s} \notin \tilde{C}_{AB}$ *then,*

$$\frac{||M_{a,b}v||_2}{||v||_2} \geq \min\left\{\frac{||M_{a,b}v_A^u||_2}{||v_A^u||_2}, \frac{||M_{a,b}Av_B^u||_2}{||Av_{B+}||_2}\right\}.$$

Proof. The proof of this lemma is entirely analogous to that of Lemma 4.3; we simply substitute $\tilde{C}_{AB}$ for C_{AB} and $M_{a,b}$ for A. □

We now consider the l_∞ norm. The images of the corners of the l_∞ unit ball under $M_{a,b}$ are

$$M_{a,b}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}M_{a,b(1,1)} + M_{a,b(1,2)}\\M_{a,b(2,1)} + M_{a,b(2,2)}\end{pmatrix},$$

$$M_{a,b}\begin{pmatrix}-1\\1\end{pmatrix} = \begin{pmatrix}-M_{a,b(1,1)} + M_{a,b(1,2)}\\-M_{a,b(2,1)} + M_{a,b(2,2)}\end{pmatrix},$$

$$M_{a,b}\begin{pmatrix}-1\\-1\end{pmatrix} = \begin{pmatrix}-M_{a,b(1,1)} - M_{a,b(1,2)}\\-M_{a,b(2,1)} - M_{a,b(2,2)}\end{pmatrix},$$

$$M_{a,b}\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}M_{a,b(1,1)} - M_{a,b(1,2)}\\M_{a,b(2,1)} - M_{a,b(2,2)}\end{pmatrix}.$$

Lemma 4.10. *Let A and B be matrices of the form given by (53) and (54), and let $v_{\max}$ and $v_{\min}$ be the vectors for which the quantity $\frac{||M_{a,b}v||_\infty}{||v||_\infty}$ is maximised and minimised respectively for $v \in T_x\mathbb{T}^2$. Then:*

(i) if $\left\|M_{a,b}\begin{pmatrix}1\\1\end{pmatrix}\right\|_\infty \geq \left\|M_{a,b}\begin{pmatrix}-1\\1\end{pmatrix}\right\|_\infty$, $v_{\max} = \begin{pmatrix}1\\1\end{pmatrix}$.

(ii) if $\left\|M_{a,b}\begin{pmatrix}1\\1\end{pmatrix}\right\|_\infty < \left\|M_{a,b}\begin{pmatrix}-1\\1\end{pmatrix}\right\|_\infty$, $v_{\max} = \begin{pmatrix}-1\\1\end{pmatrix}$.

(iii) if $\left\|M_{a,b}^{-1}\begin{pmatrix}1\\1\end{pmatrix}\right\|_\infty \geq \left\|M_{a,b}^{-1}\begin{pmatrix}-1\\1\end{pmatrix}\right\|_\infty$, $v_{\min} = \begin{pmatrix}-M_{a,b(2,2)} - M_{a,b(1,2)}\\M_{a,b(2,1)} + M_{a,b(1,1)}\end{pmatrix}$.

(iv) if $\left\|M_{a,b}^{-1}\begin{pmatrix}1\\1\end{pmatrix}\right\|_\infty < \left\|M_{a,b}^{-1}\begin{pmatrix}-1\\1\end{pmatrix}\right\|_\infty$, $v_{\min} = \begin{pmatrix}M_{a,b(2,2)} - M_{a,b(1,2)}\\-M_{a,b(2,1)} + M_{a,b(1,1)}\end{pmatrix}$.

Proof. The largest x and y coordinates, and thus largest values for $||M_{a,b}v||_\infty$ for $v \in B_{l_\infty}$, will be attained at the corner points of $M_{a,b}B_{l_\infty}$. Thus $v_{\max}$ must be a vector corresponding to a corner point of B_{l_∞}; either $(1,1)$ or $(-1,1)$.

We find $v_{\min}$ by extending an l_∞ ball out from the origin until its first intersection with $M_{a,b}B_{l_\infty}$, which will occur at a corner point of the ball. $v_{\min}$ is then the pre-image of the vector corresponding to this point of intersection - either $M_{a,b}^{-1}(1,1)$ or $M_{a,b}^{-1}(-1,1)$. □

Lemma 4.11. *Let A and B be matrices of the form given by (53) and (54), let $\tilde{C}_{AB}$ be the cone given by Lemma 4.8 and let $v_{\max}$ and $v_{\min}$ be given by Lemma 4.10. Then we obtain the following upper bounds for $\frac{||M_{a,b}v||_\infty}{||v||_\infty}$ for $v \in \tilde{C}_{AB}$:*

U1 *If $v_{\max} \in \tilde{C}_{AB}$, then*

$$\frac{||M_{a,b}v||_\infty}{||v||_\infty} \leq \frac{||M_{a,b}v_{\max}||_\infty}{||v_{\max}||_\infty}.$$

U2 *If $v_{\max} \notin \tilde{C}_{AB}$, then*

$$\frac{||M_{a,b}v||_\infty}{||v||_\infty} \leq \max\left\{\frac{||M_{a,b}v_A^u||_\infty}{||v_A^u||_\infty}, \frac{||M_{a,b}Av_B^u||_\infty}{||Av_B^u||_\infty}\right\}.$$

We also obtain the following lower bounds upon $\frac{||M_{a,b}v||_\infty}{||v||_\infty}$ for $v \in \tilde{C}_{AB}$:

L1 *If $v_{\min} \in \tilde{C}_{AB}$, then*

$$\frac{||M_{a,b}v||_\infty}{||v||_\infty} \geq \frac{||M_{a,b}v_{\min}||_\infty}{||v_{\min}||_\infty}.$$

L2 *If $v_{\min} \notin \tilde{C}_{AB}$, then*

$$\frac{||M_{a,b}v||_\infty}{||v||_\infty} \geq \min\left\{\frac{||M_{a,b}v_A^u||_\infty}{||v_A^u||_\infty}, \frac{||M_{a,b}Av_B^u||_\infty}{||Av_B^u||_\infty}\right\}.$$

Proof. This lemma follows from Lemma 4.10 and monotonicity of the l_∞ norm between its maximum and minimum. □

We now undergo the same process for the l_1 norm. The images of the corners of the l_1 unit ball under $M_{a,b}$ are

$$M_{a,b}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}M_{a,b(1,1)}\\M_{a,b(2,1)}\end{pmatrix},$$

$$M_{a,b}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}M_{a,b(1,2)}\\M_{a,b(2,2)}\end{pmatrix},$$

$$M_{a,b}\begin{pmatrix}-1\\0\end{pmatrix} = \begin{pmatrix}-M_{a,b(1,1)}\\-M_{a,b(2,1)}\end{pmatrix},$$

$$M_{a,b}\begin{pmatrix}0\\-1\end{pmatrix} = \begin{pmatrix}-M_{a,b(1,2)}\\-M_{a,b(2,2)}\end{pmatrix}.$$

Lemma 4.12. *Let A and B be matrices of the form given by (53) and (54), and let $v_{\max}$ and $v_{\min}$ be the vectors for which the quantity $\frac{||M_{a,b}v||_1}{||v||_1}$ is maximised and minimised respectively for $v \in T_x\mathbb{T}^2$. Then:*

(i) if $\left\|M_{a,b}\begin{pmatrix}1\\0\end{pmatrix}\right\|_1 \geq \left\|M_{a,b}\begin{pmatrix}0\\1\end{pmatrix}\right\|_1$, $v_{\max} = \begin{pmatrix}1\\0\end{pmatrix}$.

(ii) if $\left\|M_{a,b}\begin{pmatrix}1\\0\end{pmatrix}\right\|_1 < \left\|M_{a,b}\begin{pmatrix}0\\1\end{pmatrix}\right\|_1$, $v_{\max} = \begin{pmatrix}0\\1\end{pmatrix}$.

(iii) if $\left\|M_{a,b}^{-1}\begin{pmatrix}1\\0\end{pmatrix}\right\|_1 \geq \left\|M_{a,b}^{-1}\begin{pmatrix}0\\1\end{pmatrix}\right\|_1$, $v_{\min} = \begin{pmatrix}-M_{a,b(2,1)}\\M_{a,b(1,1)}\end{pmatrix}$.

(iv) if $\left\|M_{a,b}^{-1}\begin{pmatrix}1\\0\end{pmatrix}\right\|_1 < \left\|M_{a,b}^{-1}\begin{pmatrix}0\\1\end{pmatrix}\right\|_1$, $v_{\min} = \begin{pmatrix}M_{a,b(2,2)}\\-M_{a,b(1,2)}\end{pmatrix}$.

Proof. The greatest value of $|x| + |y|$ for $v = (x, y) \in M_{a,b}B_{l_1}$ must be attained at the image of a corner point, and thus $v_{\max}$ must correspond to one of these corner points; either $(1, 0)$ or $(0, 1)$.

The vector $v_{\min}$ is found by extending an l_1 ball out from the origin until its first intersection with $M_{a,b}B_{l_1}$, which will occur at a corner point of the ball. $v_{\min}$ is then the pre-image of the vector corresponding to this point of intersection - either $M_{a,b}^{-1}(1, 0)$ or $M_{a,b}^{-1}(0, 1)$. □

Lemma 4.13. *Let A and B be matrices of the form given by (53) and (54), let $\tilde{C}_{AB}$ be the cone given by Lemma 4.8 and let $v_{\max}$ and $v_{\min}$ be given by Lemma 4.12. Then we obtain the following upper bounds for $\frac{||M_{a,b}v||_1}{||v||_1}$ for $v \in \tilde{C}_{AB}$:*

U1 *If $v_{\max} \in \tilde{C}_{AB}$, then*

$$\frac{||M_{a,b}v||_1}{||v||_1} \leq \frac{||M_{a,b}v_{\max}||_1}{||v_{\max}||_1}.$$

U2 *If $v_{\max} \notin \tilde{C}_{AB}$, then*

$$\frac{||M_{a,b}v||_1}{||v||_1} \leq \max\left\{\frac{||M_{a,b}v_A^u||_1}{||v_A^u||_1}, \frac{||M_{a,b}Av_B^u||_1}{||Av_B^u||_1}\right\}.$$

We also obtain the following lower bounds upon $\frac{||M_{a,b}v||_1}{||v||_1}$ for $v \in \tilde{C}_{AB}$:

L1 *If $v_{\min} \in \tilde{C}_{AB}$, then*

$$\frac{||M_{a,b}v||_1}{||v||_1} \geq \frac{||M_{a,b}v_{\min}||_1}{||v_{\min}||_1}.$$

L2 *If $v_{\min} \notin \tilde{C}_{AB}$, then*

$$\frac{||M_{a,b}v||_1}{||v||_1} \geq \min\left\{\frac{||M_{a,b}v_A^u||_1}{||v_A^u||_1}, \frac{||M_{a,b}Av_B^u||_1}{||Av_B^u||_1}\right\}.$$

Proof. This lemma follows from Lemma 4.12 and monotonicity of the l_1 norm between its maximum and minimum. □

An additional improvement we can make to the bounds comes from noticing that our choice of splitting up the sequence H^n into the matrices $M_{a,b} = A^a B^b$ is arbitrary. The Lyapunov exponents are a property of the system H as a whole, and as such any splitting of the sequence into smaller chunks should yield the same value for λ. Specifically, consider a sequence

$$H^n = \dots A^{a_2} B^{b_2} A^{a_1} B^{b_1} = \dots M_{a_2,b_2} M_{a_1,b_1}.$$

Let $N_{b,a} = B^b A^a$, then we can rewrite the above as

$$H^n = \dots N_{b_3,a_2} N_{b_2,a_1} B^{b_1}.$$

Since b_1 is finite, the contribution that B^{b_1} makes to the Lyapunov exponent is zero. We can therefore also obtain bounds on λ by calculating the maximum and minimum values of $\frac{||N_{b,a}v||}{||v||}$ for all $a, b \geq 1$ and v in the cone $\tilde{C}_{BA}$, which has boundary vectors of v_B^u and Bv_A^u. This will tend to yield an improvement upon one of Φ_1 or Ψ_1 (depending on the choices of A and B), but not both. Note that in some cases doing this does not yield an improvement; an example of this is the case where $\alpha, \beta = -1$ and $\gamma, \delta = 1$. Note that the bounds shown by the figures in this section and section 4.6 take the information in this paragraph into account; that is, to produce the figures we have calculated upper and lower bounds when splitting the sequence H^n into subsequences of both $M_{a,b}$ and $N_{b,a}$, then plotted the smaller of the two upper bounds and the larger of the two lower bounds.

Taking a truncation of the infinite sum up to any finite a, b in Theorem 4.3 will yield a rigorous lower bound; however, the upper bound is not rigorous until the entire sum is completed. In order to ensure that the upcoming numerical results return a rigorous upper bound as well, we have added a 'remainder' term. Specifically,

we calculate

$$\Phi_1 = \frac{1}{4}\Bigg(\sum_{a,b=1}^{i} 2^{-a-b} \log\big(\max_{v\in\tilde{C}_{AB}} \Big\{ \frac{||M_{a,b}v||}{||v||} \Big\}\big) + \big(1 - \sum_{a,b=1}^{i} 2^{-a-b}\big)\Big(\max_{v\in\tilde{C}_{AB}} \Big\{ \frac{||Av||}{||v||}, \frac{||Bv||}{||v||} \Big\}\Big)\Bigg). \tag{77}$$

The expression on the second line assumes that the terms cut off by the truncation are 'as expansive as possible' - that is, they consist entirely of A's or B's, and always yield the *largest possible* growth for vectors within the cone. By doing this, the remainder term is overestimating the terms in the infinite sum we did not consider, and thus Φ_1 is a rigorous upper bound.

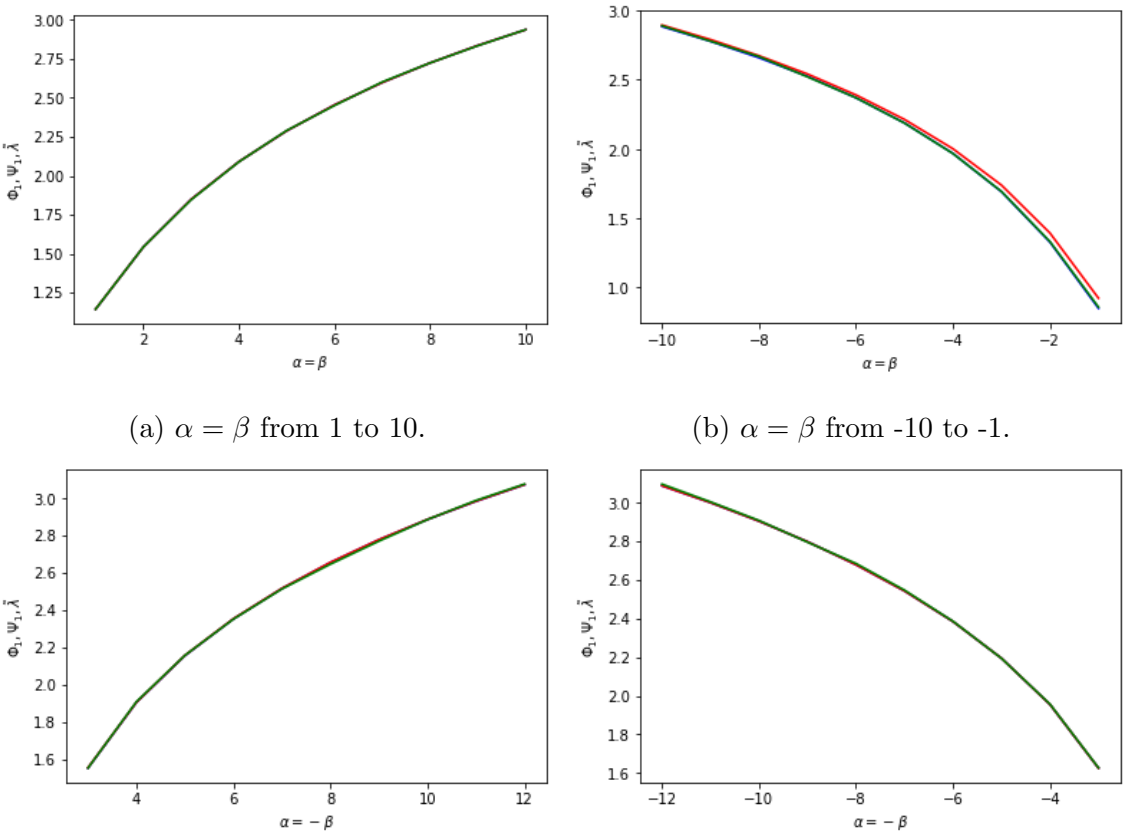
Applying the results from Lemmas 4.9 - 4.13 to Theorem 4.3 yields rigorous, explicit bounds on λ for any of the norms we have studied. We now show the numerical results obtained when calculating these bounds in Figures 29 and 30, and compare them to an estimate for λ, $\tilde{\lambda}$, obtained using Gram Schmidt orthonormalization. Figure 31 shows the range of values which the bounds encompass in the cases studied in the previous figures (i.e. the width of the gap between the bounds).

Figure 30 shows the bounds Ψ_1 and Φ_1, as well as $\tilde{\lambda}$, for $\gamma = 1$, $\delta = 2$, and varying α and β. As mentioned earlier, to obtain these graphs (and the table in Figure 29), we calculated the truncation of the infinite sums, in this case up to and including the terms where $a = b = 40$. On the given scale, we see little difference between $\tilde{\lambda}$ and either of Ψ_1 or Φ_1 for all cases except Figure 30b, where $\alpha = \beta < 0$. The wider gap between the bounds in this case is explained by noting that when $\alpha = \beta < 0$, the width of the cone is at its widest of all those studied in Figure 30.

Figure 31 shows that Ψ_1 and Φ_1 are useful for obtaining λ between 2 s.f., for small α, β in modulus for case (b), to 6+ s.f., for large α, β in modulus for case (d). In general, we see that increasing the modulus of α and β tends to narrow the gap between the bounds. This is due to the fact that a matrix (in this case A) with

U, Ψ_1, Φ_1, $\tilde{\lambda}$, and $\sigma_{\tilde{\lambda}}$ when $\gamma = 1$, $\delta = 2$.					
α, β	U	Ψ_1	Φ_1	$\tilde{\lambda}$	$\sigma_{\tilde{\lambda}}$
1,1	1.31696	1.142968	1.143659	1.143119	0.001902
5,5	3.29446	2.287968	2.287986	2.293997	0.010411
10,10	4.64488	2.936932	2.936934	2.941293	0.011061
-1,-1	1.31696	0.845339	0.918686	0.852869	0.002807
-5,-5	3.29446	2.186900	2.212729	2.184764	0.006954
-10,-10	4.64488	2.885053	2.896129	2.884233	0.017851
3,-3	1.92485	1.553993	1.555999	1.555323	0.002917
7,-7	3.84969	2.517056	2.517991	2.514943	0.009602
12,-12	4.95578	3.073275	3.073727	3.074438	0.016346
-3,3	1.92485	1.624771	1.624787	1.623924	0.002715
-7,7	3.84969	2.542714	2.542723	2.545598	0.013857
-12,12	4.95578	3.087957	3.087958	3.088964	0.014868

Figure 29: Table showing the values of the naive upper bound $U = \max\{\lambda_A, \lambda_B\}$, Φ_1, Ψ_1, an estimate of λ, $\tilde{\lambda}$, and the standard deviation relating to this estimate, $\sigma_{\tilde{\lambda}}$, for various choices of α and β, with $\gamma = 1$, $\delta = 2$. The best lower and upper bound have been selected from $\Psi_1^{(k)}$ and $\Phi_1^{(k)}$, $k \in \{1, 2, \infty\}$, and $\tilde{\lambda}$ is obtained by taking the average of 20 'trials' of 10^4 iterates of Gram-Schmidt orthonormalization; $\sigma_{\tilde{\lambda}}$ is the standard deviation related to this random sample. Note that it takes approximately 2 seconds to calculate both Ψ_1 and Φ_1, and approximately 8.8 seconds to calculate $\tilde{\lambda}$; however, the code required to obtain $\tilde{\lambda}$ is significantly shorter.



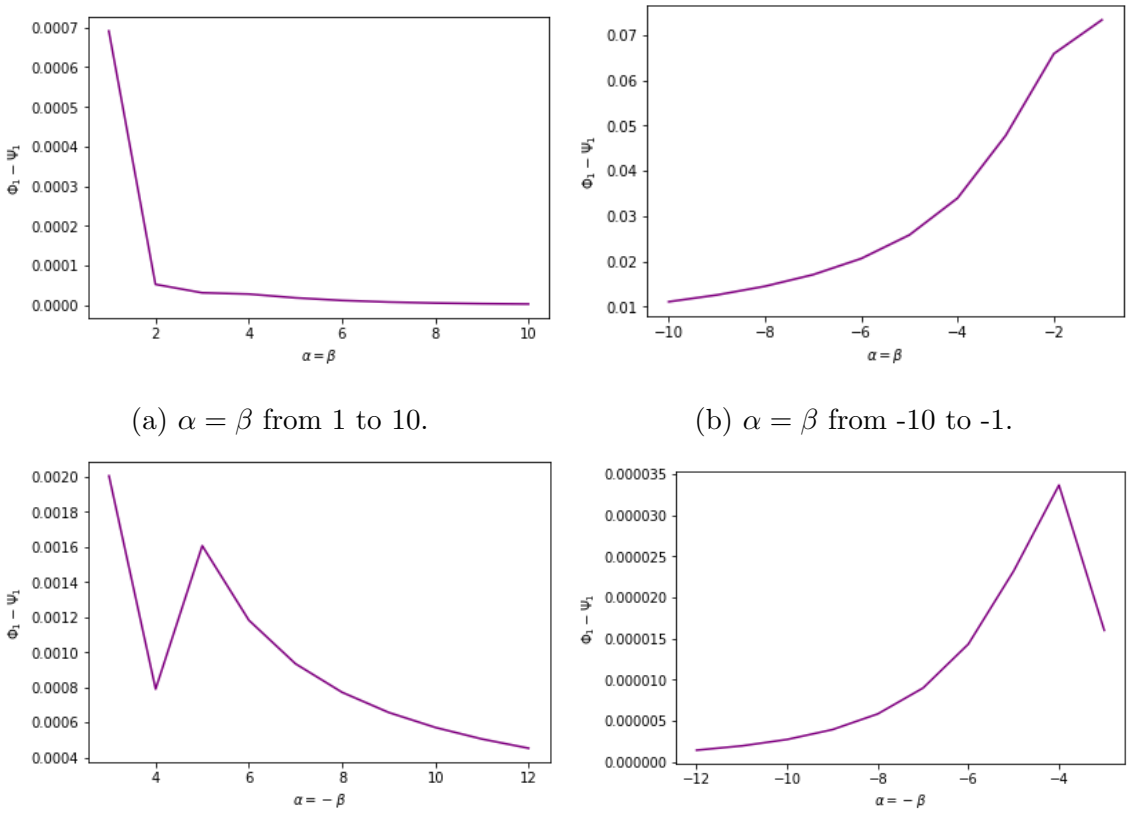
(a) $\alpha = \beta$ from 1 to 10.

(b) $\alpha = \beta$ from -10 to -1.

(c) $\alpha = -\beta$ from 3 to 12.

(d) $\alpha = -\beta$ from -12 to -3.

Figure 30: Φ_1 and Ψ_1 under varying choices of α and β, with $\gamma = 1$, $\delta = 2$, and an estimate of λ, $\tilde{\lambda}$, obtained by taking the average of 10 'trials' of 10^4 iterates of Gram-Schmidt orthonormalization. Note that, on this scale, noticeable gaps between the bounds occur in case (b) only - this is the case (of those studied here) where the eigenvector v_A^u is located the furthest away from v_B^u. Note that it takes approximately 1.91 seconds to calculate both Ψ_1 and Φ_1.



(a) $\alpha = \beta$ from 1 to 10.

(b) $\alpha = \beta$ from -10 to -1.

(c) $\alpha = -\beta$ from 3 to 12.

(d) $\alpha = -\beta$ from -12 to -3.

Figure 31: The range of values the bounds encompass (i.e. $\Phi_1 - \Psi_1$) under varying choices of α and β, with $\gamma = 1$, $\delta = 2$.

larger eigenvalues will exert a larger 'pull' on vectors within the cone, aligning them with v_A^u more quickly than a matrix with smaller eigenvalues. This is of particular relevance for the cone $\tilde{C}_{AB}$; its boundary vector Av_B^u will be more aligned with its other boundary vector v_A^u for larger eigenvalues, and thus the cone itself will be narrower in these cases. This is a similar result to that of Sturman and Thiffeault, who found that increasing the slope of the shears in the random product would cause their bounds to become more accurate.

It is more likely for the vectors $v_{\max}$ and $v_{\min}$ for some $M_{a,b}$ and any of the l_1,

l_2 or l_∞ norms to be contained within a wider cone. Ideally we would like to avoid calculating upper and lower bounds using these vectors wherever possible, as they represent the extreme cases of growth rates of vectors within the cone. Section 4.6 discusses a way to remove some of these extreme cases in order to improve the accuracy of the bounds, via restriction of the calculations to narrower cones.

4.6 Improving the cones

In this section we discuss a way in which we can narrow the cone $\tilde{C}_{AB}$ to a collection of smaller subcones, with the intention of narrowing the gap which remains between the bounds Φ_1 and Ψ_1. This is achieved by considering the matrices which precede the matrices $M_{a,b}$ we used to calculate the bounds. The collection of narrower subcones of $\tilde{C}_{AB}$ will contain some cones which do not include the maximum and minimum growth rates we find within the cone $\tilde{C}_{AB}$. In particular, if $\tilde{C}_{AB}$ contains the vectors $v_{\max}$ or $v_{\min}$ for any of the l_2, l_1 or l_∞ norms, then considering 'large enough' preceding matrices will yield subcones which do not contain these vectors, and thus allow us to improve the contribution to the bounds we obtain from these cones. This idea was also studied in [54] in the case of random products of shears matrices, where a significant improvement was found over the bounds involving wider cones.

Theorem 4.3 gave us the formulae for Φ_1 and Ψ_1. To calculate these bounds, we find the maximum and minimum possible contributions to the growth rate we can obtain for each $M_{a,b}$ and for vectors in the cone $\tilde{C}_{AB}$. This was justified by noting that we could split the infinite sequence of matrices we obtain by iterating H into the matrices $M_{a,b}$. Consider now the matrices which precedes each of these $M_{a,b}$, $M_{m,n}$ say. The frequency with which each $M_{m,n}$ precedes the matrix $M_{a,b}$ is identical to the frequency with which each matrix appears: 2^{-m-n}. Thus the frequency with

which we find a particular matrix $M_{a,b}$ preceded by another matrix $M_{m,n}$ is

$$\mathbb{P}(M_{a,b}M_{m,n}) = 2^{-a-b-m-n}. \tag{78}$$

Consider now the location of a vector $v \in \tilde{C}_{AB}$ following application by $M_{m,n}$; the vector must be contained within the cone $M_{m,n}\tilde{C}_{AB}$, with boundary vectors given by $M_{m,n}v_A^u$ and $M_{m,n}Av_B^u$. Thus, with frequency 2^{-m-n}, we can choose vectors $v \in M_{m,n}\tilde{C}_{AB}$ instead of $\tilde{C}_{AB}$ when calculating $\frac{||M_{a,b}v||}{||v||}$. We can therefore rewrite the $M_{a,b}$ term (for this particular a and b) in Theorem 4.3 as

$$2^{-a-b-m-n}\log\big(\max_{v\in M_{m,n}\tilde{C}_{AB}}\big\{\frac{||M_{a,b}v||}{||v||}\big\}\big) + (1-2^{-m-n})\cdot 2^{-a-b}\log\big(\max_{v\in\tilde{C}_{AB}}\big\{\frac{||M_{a,b}v||}{||v||}\big\}\big) \tag{79}$$

and

$$2^{-a-b-m-n}\log\big(\min_{v\in M_{m,n}\tilde{C}_{AB}}\big\{\frac{||M_{a,b}v||}{||v||}\big\}\big) + (1-2^{-m-n})\cdot 2^{-a-b}\log\big(\min_{v\in\tilde{C}_{AB}}\big\{\frac{||M_{a,b}v||}{||v||}\big\}\big) \tag{80}$$

respectively. The second term in each expression essentially says that for all $M_{m,n}$ for which we do not explicitly calculate the cone $M_{m,n}\tilde{C}_{AB}$, we instead take the cone $\tilde{C}_{AB}$, which we know is invariant under any $M_{m,n}$ by Lemma 4.8.

We can of course consider more than just one preceding matrix $M_{m,n}$. If we proceed with the process outlined above, adding in additional terms with the appropriate frequencies for each preceding $M_{m,n}$, we instead obtain, for each $M_{a,b}$, the upper bound

$$U_{a,b} = 2^{-a-b}\cdot\sum_{m,n=1}^{\infty}2^{-m-n}\log\big(\max_{v\in M_{m,n}\tilde{C}_{AB}}\big\{\frac{||M_{a,b}v||}{||v||}\big\}\big) \tag{81}$$

and the lower bound

$$L_{a,b} = 2^{-a-b}\cdot\sum_{m,n=1}^{\infty}2^{-m-n}\log\big(\min_{v\in M_{m,n}\tilde{C}_{AB}}\big\{\frac{||M_{a,b}v||}{||v||}\big\}\big) \tag{82}$$

The term 2^{-m-n} becomes very small as m and n increase, and so in a similar fashion to Φ_1 and Ψ_1, taking 'large enough' m and n is sufficient to obtain the terms $U_{a,b}$

and $L_{a,b}$ to reasonable degree of accuracy; note that in the upcoming figures we take $a = b = 40$ and $m = n = 10$. Summing over all a, b, we obtain the upper bound

$$\lambda \leq \Phi_2 = \frac{1}{4} \sum_{a,b=1}^{\infty} U_{a,b}, \tag{83}$$

and the lower bound

$$\lambda \geq \Psi_2 = \frac{1}{4} \sum_{a,b=1}^{\infty} L_{a,b}. \tag{84}$$

We label these bounds as Φ_2 and Ψ_2, where the 2 indicates that we are considering a total of 2 matrices ($M_{a,b}$ and $M_{m,n}$) in each term. This idea is summarised in the following theorem.

Theorem 4.4. *Under the same conditions as Theorem 4.3, let*

$$U_{m,n,a,b} = 2^{-m-n-a-b} \log\big(\max_{v \in M_{m,n}\tilde{C}_{AB}} \{ \frac{||M_{a,b}v||}{||v||} \} \big) \tag{85}$$

and

$$L_{m,n,a,b} = 2^{-m-n-a-b} \log\big(\min_{v \in M_{m,n}\tilde{C}_{AB}} \{ \frac{||M_{a,b}v||}{||v||} \} \big). \tag{86}$$

Then the maximal Lyapunov exponent λ has the following upper and lower bounds:

$$\lambda \leq \Phi_2 = \frac{1}{4} \sum_{a,b=1}^{\infty} \sum_{m,n=1}^{\infty} U_{m,n,a,b}, \tag{87}$$

$$\lambda \geq \Psi_2 = \frac{1}{4} \sum_{a,b=1}^{\infty} \sum_{m,n=1}^{\infty} L_{m,n,a,b}. \tag{88}$$

Note that we can continue this process indefinitely, looking at more and more preceding matrices to obtain much narrower cones. Let us relabel our first matrix as M_{a_1,b_1}, the first preceding matrix as M_{a_2,b_2}, the matrix which precedes this as M_{a_3,b_3}, and so on. The following theorem defines bounds Φ_k and Ψ_k which take into account all of the matrices M_{a_κ,b_κ} where $\kappa \leq k$.

Theorem 4.5. *Under the same conditions as Theorem 4.3, with $a = a_1$ and $b = b_1$, let*

$$\hat{C}_{AB}(a_k, b_k, \ldots, a_2, b_2) = M_{a_k,b_k} \cdot \ldots \cdot M_{a_2,b_2} \tilde{C}_{AB}, \tag{89}$$

$$U_{a_k,b_k,\ldots,a_1,b_1} = 2^{-a_k-b_k-\ldots-a_1-b_1} \log\Big(\max_{v \in \hat{C}_{AB}(a_k,b_k,\ldots,a_2,b_2)} \Big\{ \frac{||M_{a_1,b_1} v||}{||v||} \Big\} \Big) \tag{90}$$

and

$$L_{a_k,b_k,\ldots,a_1,b_1} = 2^{-a_k-b_k-\ldots-a_1-b_1} \log\Big(\min_{v \in \hat{C}_{AB}(a_k,b_k,\ldots,a_2,b_2)} \Big\{ \frac{||M_{a_1,b_1} v||}{||v||} \Big\} \Big). \tag{91}$$

Then the maximal Lyapunov exponent λ has the following upper and lower bounds:

$$\lambda \leq \Phi_k = \frac{1}{4} \sum_{a_1=1,b_1=1}^{\infty} \ldots \sum_{a_k=1,b_k=1}^{\infty} U_{a_k,b_k,\ldots,a_1,b_1}, \tag{92}$$

$$\lambda \geq \Psi_k = \frac{1}{4} \sum_{a_1=1,b_1=1}^{\infty} \ldots \sum_{a_k=1,b_k=1}^{\infty} L_{a_k,b_k,\ldots,a_1,b_1}. \tag{93}$$

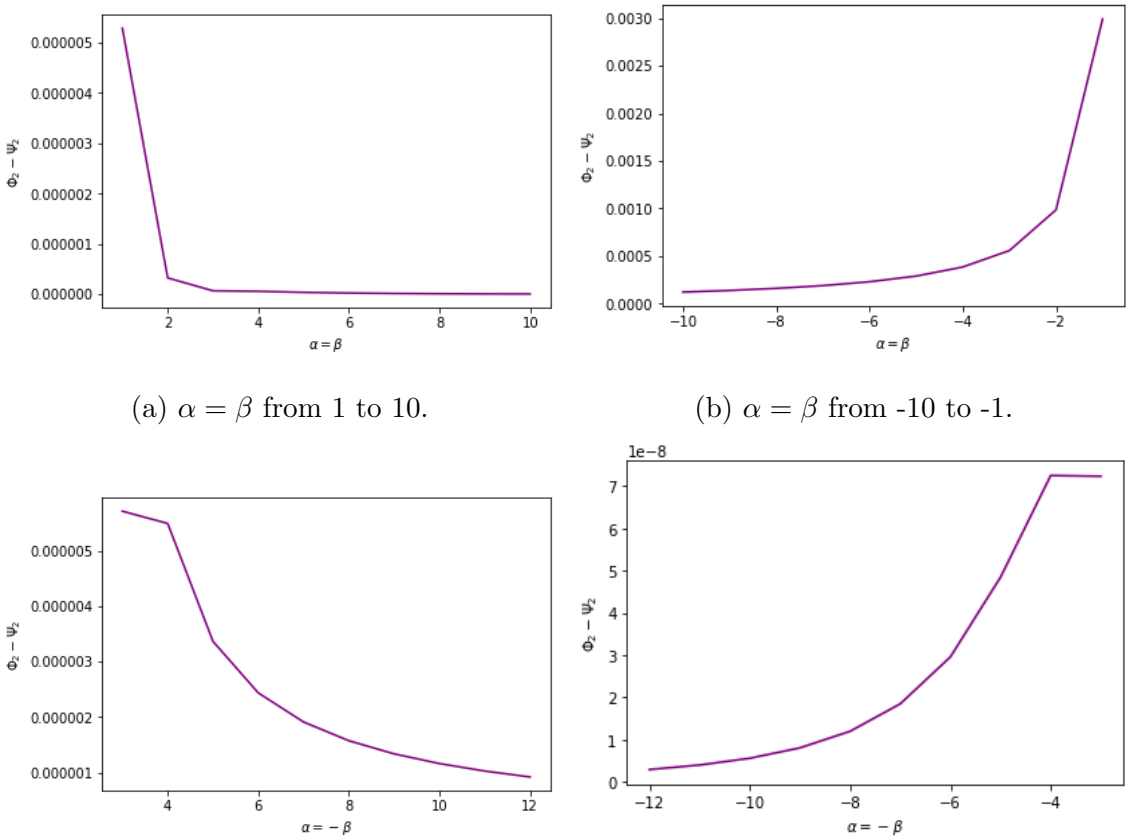
Figures 32 and 33 show the numerical results for Ψ_2 and Φ_2 for the same cases we studied in Figures 29 - 31 in Section 4.5. Note that we again calculate Φ_2 using a remainder term as in (77), but in addition to this, we also require a remainder term for the preceding matrices for which we do not calculate the cone $M_{m,n}\tilde{C}_{AB}$ (i.e. those outside of the finite truncation). Specifically, we calculate

$$\begin{aligned} \Phi_2 = \frac{1}{4} \Bigg(& \sum_{a,b=1}^{i} \Big(\sum_{m,n=1}^{j} 2^{-m-n-a-b} \log\big(\max_{v \in M_{m,n}\tilde{C}_{AB}} \{ \frac{||M_{a,b} v||}{||v||} \} \big) \\ & + 2^{-a-b}(1 - \sum_{m,n=1}^{j} 2^{-m-n}) \log\big(\max_{v \in \tilde{C}_{AB}} \{ \frac{||M_{a,b} v||}{||v||} \} \big) \Big) \\ & + \big(1 - \sum_{a,b=1}^{i} 2^{-a-b}\big) \Big(\max_{v \in \tilde{C}_{AB}} \Big\{ \frac{||Av||}{||v||}, \frac{||Bv||}{||v||} \Big\} \Big) \Bigg). \end{aligned} \tag{94}$$

The term on the second line says that for the preceding matrices outside of the finite truncation, we simply take $\tilde{C}_{AB}$ as the invariant cone. We thus make no improvement over Φ_1 for these particular cases; however, this does ensure that $\Phi_2 \leq \Phi_1$. If we

U, Ψ_2, Φ_2, $\tilde{\lambda}$ and $\sigma_{\tilde{\lambda}}$ when $\gamma = 1$, $\delta = 2$.					
α, β	U	Ψ_2	Φ_2	$\tilde{\lambda}$	$\sigma_{\tilde{\lambda}}$
1,1	1.31696	1.1433078679	1.1433131537	1.1432098	0.0005483
5,5	3.29446	2.2879800912	2.2879801286	2.2871329	0.0029977
10,10	4.62488	2.9369335256	2.9369335306	2.9355497	0.0041252
-1,-1	1.31696	0.8526194278	0.8556089639	0.8526771	0.0008515
-5,-5	3.29446	2.1869793464	2.1872709486	2.1867099	0.0033443
-10,-10	4.62488	2.8850566596	2.8851807964	2.8853677	0.0059216
3,-3	1.92485	1.5554080772	1.5554137897	1.5554526	0.0008910
7,-7	3.84969	2.5170765398	2.5170784508	2.5173760	0.0032072
12,-12	4.95578	3.0732764275	3.0732773469	3.0724321	0.0083176
-3,3	1.92485	1.6247782255	1.6247782978	1.6245508	0.0009367
-7,7	3.84969	2.5427198562	2.5427198747	2.5431770	0.0030175
-12,12	4.95578	3.0879578906	3.0879578935	3.0862290	0.0067713

Figure 32: Table showing the values of the naive upper bound $U = \max\{\lambda_A, \lambda_B\}$, Φ_2, Ψ_2, an estimate of λ, $\tilde{\lambda}$, and the standard deviation relating to this estimate, $\sigma_{\tilde{\lambda}}$, for various choices of α and β, with $\gamma = 1$, $\delta = 2$. The best lower and upper bound have been selected from $\Psi_2^{(k)}$ and $\Phi_2^{(k)}$, $k \in \{1, 2, \infty\}$, and $\tilde{\lambda}$ is obtained by taking the average of 20 'trials' of 10^5 iterates of Gram-Schmidt orthonormalization; $\sigma_{\tilde{\lambda}}$ is the standard deviation related to this random sample. The calculation time for the bounds is approximately 90 seconds (to obtain both), and for $\tilde{\lambda}$ is approximately 85 seconds. The bounds $\Psi_2^{(k)}$ and $\Phi_2^{(k)}$ take approximately 50 seconds to obtain individually for each k.



(a) $\alpha = \beta$ from 1 to 10.

(b) $\alpha = \beta$ from -10 to -1.

(c) $\alpha = -\beta$ from 3 to 12.

(d) $\alpha = -\beta$ from -12 to -3.

Figure 33: The range of values the bounds encompass (i.e. $\Phi_2 - \Psi_2$) under varying choices of α and β, with $\gamma = 1$, $\delta = 2$ (includes the data from Figure 32).

wished to calculate numerically a rigorous upper bound Φ_k for some k, we would have to undergo a similar process to this for each set of preceding matrices.

Figure 32 shows the value of a naive upper bound $U = \max\{\lambda_A, \lambda_B\}$, Ψ_2, and Φ_2 for various choices of α and β, with $\gamma = 1$, $\delta = 2$. We see that, with the exception of the case $\alpha = \beta < 0$, Ψ_2 and Φ_2 agree to 6-8 s.f.. The improvements in accuracy are 2-3 s.f. (a factor of 100-1000) for all cases over Ψ_1 and Φ_1, while the calculation time has increased by a factor of roughly 45.

We also compare these values to the average of a sample of 20 finite time Lyapunov exponents of 10^5 iterates, $\tilde{\lambda}$, as well as the standard deviation, $\sigma_{\tilde{\lambda}}$, of the sample. Note that this sample size was chosen to provide a similar computation time to the bounds, so as to compare the results one may achieve using the two methods for a similar time period. Note that in each case except $\alpha, \beta = -1$, we have that the interval $[\Psi_2, \Phi_2] \subset [\tilde{\lambda} - \sigma_{\tilde{\lambda}}, \tilde{\lambda} - \sigma_{\tilde{\lambda}}]$, indicating that Ψ_2 and Φ_2 provide a tighter range of values for λ than $\tilde{\lambda}$. In addition to this, we see that as we increase α and β in modulus the tightness of Ψ_2 and Φ_2 improves, whereas $\sigma_{\tilde{\lambda}}$ increases, indicating that the reliability of $\tilde{\lambda}$ decreases. It is worth noting that the code required to obtain $\tilde{\lambda}$ is significantly shorter and widely known, however.

Figure 33 shows the range of values the bounds Ψ_2 and Φ_2 encompass for various choices of α and β, with $\gamma = 1$, $\delta = 2$. We also see that, similar to what we found for Ψ_1 and Φ_1, the gap between the bounds tends to decrease as α and β increase in modulus for all cases.

It should be noted that it is possible to 'pick and choose' which matrices $M_{m,n}$ we consider in our truncation, as we can simply add those we do not consider into the remainder term in (94). It therefore may be possible to only choose the matrices which provide the biggest improvement in accuracy in order to reduce the computing time.

4.7 Summary

We have studied bounds on maps H consisting of random products of matrices formed via the composition of shear matrices on the 2-torus, which utilise the existence of mutually invariant cones to bound the term $||H^n v||$ in the definition of the Lyapunov exponent λ. We have the considered ways in which we can improve these bounds, by finding ways to narrow the cones we consider in the calculations.

This work has been an extension to that of Sturman and Thiffeault [54], who studied random products of shear matrices; in particular, the method of considering a sequence of preceding matrices to narrow the cones we consider in the bounds is novel, and in principle, could be applied to the bounds in [54] as well, with appropriate modifications to the cones being considered. Similarly, the work on generalised Lyapunov exponents they discuss could be applied to systems of this form as well, using the cones discussed in this chapter.

A possible application of the lower bound Ψ_i, $i \in \mathbb{N}$, is found in Ayyer and Stenlund [7], who obtain an upper bound upon the correlation decay of the systems we have studied in this chapter. The upper bound depends upon their lower bound upon the maximal Lyapunov exponent, which is given by

$$\lambda \geq \log \max\{\lambda_C^{-1}, \lambda_\epsilon\},$$

where λ_C is the minimum contraction rate of vectors within a contraction cone C (whose image remains within C), and λ_ϵ is the minimum expansion rate of vectors in an expansion cone E. Note that, in the context of this chapter, $C_{AB} \subseteq E$ is the minimal such expansion cone, and the cones we consider in subsequent iterations of the bounds ($M_{m,n}\tilde{C}_{AB}$) are significantly narrower than any E; specifically, the cones narrow exponentially with m. The lower bound λ given above is simply the smallest possible growth rate of a vector within the cone E, and is therefore significantly improved upon by the bounds defined in this chapter, as these are the vectors we

try to avoid by narrowing our cones.

Pollicott studies the systems in this chapter where A and B have only positive entries in [42], where he derives an efficient method for calculating λ in these cases. This method involves the uses of complex 'determinant' functions; these are used to express the Lyapunov exponent explicitly via a series involving their partial derivatives. The coefficients of this series can then be calculated in turn to yield more accurate estimations of λ, labelled λ_i, $i \in \mathbb{N}$. He gives an explicit calculation for the case equivalent to $\alpha, \beta = 1$, $\gamma = 1$, and $\delta = 2$, the 9th step of which yields the approximation

$$\lambda_9 = 1.14331103510294924584325185365558829940 25\ldots$$

which appears accurate to 32 d.p., and subsequent steps converge super exponentially. For the purposes of accurate computation in these cases, this method clearly is superior to the bounds Ψ_i and Φ_i, which increase exponentially in computing time (assuming $a_i = c$ for $i, c \in \mathbb{N}$). However, the code for the bounds Ψ_i and Φ_i are relatively simple, and they are not restricted to positive entries.

5 Bounds on Lyapunov exponents in Linked Twist Maps

The Lyapunov exponents of a dynamical system describe the average rate of separation of infinitesimally close trajectories in phase space, or in other words, the rate of stretching of the vectors between these trajectories. In measure-preserving systems these exponents converge (almost-everywhere) to a limit independent of the choice of initial condition [35]. However, this convergence need not be quick; in practical applications, one must study finite-time Lyapunov exponents (FTLEs), which are spatially dependent. Some regions of the phase space may undergo little to no expansion for large numbers of iterates, resulting in abnormally small FTLEs for some initial conditions.

A good example of a family of maps which encapsulates this behaviour is (linear) toral linked twist maps, which we studied in Chapter 3. Recall that these consist of regions which behave similarly to an Anosov diffeomorphism, with exponential expansion and contraction rates, while other regions behave like shears, with sub-exponential expansion rates. Orbits in such maps can become trapped in the shear regions for any finite number of iterates, and thus contribute nothing to their Lyapunov exponent for arbitrarily large numbers of iterates along their orbits; in particular, this can occur when an orbit lands very close to a boundary [53].

In this chapter we discuss a method, similar to that of the previous chapter, for finding explicit upper and lower bounds for the Lyapunov exponents of a parametrised family of linked twist maps; however, in principle, this method would work for any LTM, or indeed any map for which we can obtain invariant cones upon orbits returning to a particular region - an example of such a map is given in Section 5.6. We again consider invariant cones for the maps to obtain one part of the bound, and the distributions of points undergoing particular orbits for the other. The main

difference between the bounds in the previous chapter (and those in Sturman and Thiffeault [54]) is that the systems discussed therein consisted of randomly chosen matrices, and so the realisation of points undergoing iteration by certain sequences of matrices was trivial. Linked twist maps are deterministic and so the analogous construction, known as the return time distribution, needs to be considered more carefully.

In Section 5.1, we define the family of linked twist maps we will be studying and briefly revisit some of the theory discussed in Chapter 4, as well as stating the theorem central to this chapter, which gives explicit elementary bounds on the Lyapunov exponents of these systems. In Section 5.2 we study the return time distribution of points beginning in a particular region of the domain. More specifically, we study the frequency with which points experience particular orbits in order to determine that the distribution can be written inductively. In Section 5.3 we study the invariant cones of the family of linked twist maps, and calculate upper and lower bounds for various choices of norms under specific sequences of shears. In Section 5.4 we combine the results from the previous two sections to obtain our upper and lower bounds for the Lyapunov exponent, and thus complete the proof of the theorem stated in Section 5.1. In Section 5.5 we discuss a way to narrow the cones used in the calculation of these bounds in order to improve their accuracy, as well as discussing the general practicality of the bounds and ways in which it may be improved upon. Finally, in Section 5.6 we discuss other examples of deterministic maps for which these bounds can be obtained.

5.1 The parametrised family of linked twist maps

In this section we describe the specific parametrised family of linked twist maps we will study in this chapter, outline the concepts used in the calculation of bounds upon the Lyapunov exponents of these maps, and state a theorem which yields

rigorous upper and lower bounds for their Lyapunov exponents, λ_1, λ_2. Note that since these maps are measure-preserving, they satisfy $\lambda_1 + \lambda_2 = 0$, thus any bound on the maximal Lyapunov exponent λ_1 will also give a bound on the minimal Lyapunov exponent λ_2; we henceforth refer to the Lyapunov exponent as λ.

The maps we study act (non-trivially) upon two annuli, P and Q, which are parametrised by $p \in (0, 1)$. We will initially focus on the range of parameters $0.682 < p < 1$ for simplicity; we will see that this is the range of parameter values for which the return time distribution is at its simplest (see Section 5.2). However, note that return time distributions of such maps can be studied for any parameter $0 < p < 1$ (see, for example, the similar constructions in [51]), and, in principle, for non-linear systems too.

We define two maps: $F : \mathbb{T}^2 \to \mathbb{T}^2$ given by

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} F_1\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{p} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} & \text{if } (x, y) \in P, \\ F_2\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} & \text{otherwise,} \end{cases} \tag{95}$$

and $G : \mathbb{T}^2 \to \mathbb{T}^2$, given by

$$G\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} G_1\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{1}{p} & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} & \text{if } (x, y) \in Q, \\ G_2\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} & \text{otherwise,} \end{cases} \tag{96}$$

with annuli P and Q defined by

$$P = \{(x, y) \in \mathbb{T}^2 : 0 \leq y \leq p\},$$

$$Q = \{(x, y) \in \mathbb{T}^2 : 0 \leq x \leq p\}.$$

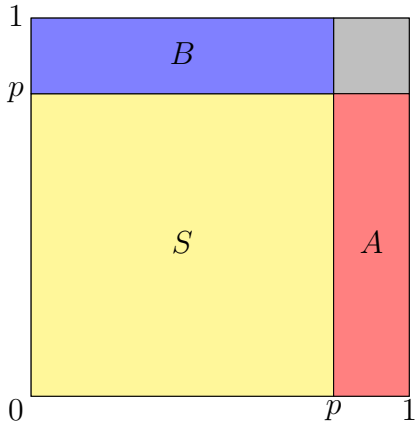


Figure 34: The regions of the linked twist map $H = F \circ G$ defined on the annuli $P = S \cup A$ and $Q = S \cup B$. We will study the return time distribution of the points in S.

We define the map $H = G \circ F : P \cup Q \to P \cup Q$ as the linked twist map given by the composition of the restrictions of F and G to $P \cup Q$, and we define the regions $S = \{(x, y) \in \mathbb{T}^2 : x, y \leq p\}$, $A = \{(x, y) \in \mathbb{T}^2 : y \leq p, x > p\}$ and $B = \{(x, y) \in \mathbb{T}^2 : x \leq p, y > p\}$, which are shown in Figure 34.

We note here that, in addition to the cases stated earlier, we see no theoretical barrier to the calculation of such bounds in the case where

$$P = \{(x, y) \in \mathbb{T}^2 : 0 \leq y \leq p\},$$

$$Q = \{(x, y) \in \mathbb{T}^2 : 0 \leq x \leq q\},$$

for $p \neq q$, that is, where the annuli P and Q have independent widths, nor in the case where we have multiple disjoint horizontal and/or vertical annuli, $P_1, \ldots, P_n$ and $Q_1, \ldots, Q_m$ respectively. In practice, however, the calculation of the corresponding return time distributions is not a simple task.

Each iterate of H is composed of an application of F followed by an application of G; however, since F_2 and G_2 are both the identity map, this can be simplified further to just applications of F_1 and G_1. In a similar method to Chapter 4, and for

reasons justified in Section 5.2, we will split this sequence up in the following way:

$$\begin{aligned} H^N v_0 &= GFGF \cdot \ \ldots \ \cdot GFGF v_0, \\ &= G_1F_2G_1F_1 \cdot \ \ldots \ \cdot G_1F_2G_2F_1 v_0, \\ &= G_1G_1F_1 \cdot \ \ldots \ \cdot G_1F_1 v_0, \\ &= G_1^{b_n}F_1^{a_n} \cdot \ \ldots \ \cdot G_1^{b_1}F_1^{a_1} v_0, \\ &= M_{a_n,b_n} \cdot \ \ldots \ \cdot M_{a_1,b_1} v_0, \end{aligned}$$

where

$$M_{a_i,b_i} = G_1^{b_i}F_1^{a_i} = \begin{pmatrix} 1 & \frac{a_i}{p} \\ \frac{b_i}{p} & 1 + \frac{a_i b_i}{p^2} \end{pmatrix}, \tag{97}$$

so that, assuming $||v_0|| = 1$,

$$||H^N v_0|| = \prod_{i=1}^{n} \frac{||M_{a_i,b_i} v_{i-1}||}{||v_{i-1}||}, \tag{98}$$

where $v_i = M_{a_i,b_i} v_{i-1}$.

To obtain an upper (lower) bound on $||H^N v||$, and consequently λ, we need to find an upper (lower) bound for $\frac{||M_{a,b}v||}{||v||}$ for each *possible* combination of a and b. This is different to the random cases we have studied before - it is possible that not all sequences of these matrices are achievable under these maps. We discuss how to bound these quantities in Section 5.3.

Let us assume we have found an upper bound $\phi(a,b)$ and a lower bound $\psi(a,b)$ for each possible choice of a and b. We will also need to know the distribution of these sequences - the frequency with which they occur along an orbit - in order to calculate their overall contribution to the Lyapunov exponent. Note that ergodicity of the LTM on $P \cup Q$ ([12], [44]) ensures that these frequencies are equal for μ a.a. orbits, where μ is the renormalised Lebesgue measure. We will discuss this return time distribution in detail in Section 5.2, but for now let $R(a,b)$ be the measure of points which return to our reference region S under application of the matrix $M_{a,b}$,

and let n_S be the average return time, in terms of the number of iterates of H, of all points in this region. Then we have

$$\lambda \leq \frac{1}{n_S} \sum_{a,b=1}^{\infty} R(a,b) \log \phi(a,b) = \Phi_H, \tag{99}$$

and

$$\lambda \geq \frac{1}{n_S} \sum_{a,b=1}^{\infty} R(a,b) \log \psi(a,b) = \Psi_H. \tag{100}$$

Note that the $R(a,b)$ (taken over all a,b) yields the return time distribution; this is analogous to the realisation of the matrices in the random case, which was simply 2^{-a-b} for each $M_{a,b}$. Similarly, the average return time n_S is analogous to the expected length of the matrices $M_{a,b}$ in the random case.

In order for these bounds to converge, the frequency with which the $M_{a,b}$'s occur as a and b increase must decrease at a faster rate than the bounds $\phi(a,b)$ and $\psi(a,b)$ increase; in other words, longer matrices $M_{a,b}$ must be sparser within an orbit than shorter ones. Note that, since the bounds $\phi(a,b)$ and $\psi(a,b)$ will increase at most exponentially, any polynomial $R(a,b)$ is sufficient to meet this condition, and we will find that this is indeed the case for the family of maps H. Furthermore, we will see that only a select few of these $M_{a,b}$ are possible, and the frequency with which they occur within an orbit can be written inductively for 'large' a or b. The results are summarised in the following theorem.

Theorem 5.1. *Let $H : P \cup Q \to P \cup Q$ be the linked twist map defined by the restriction of the composition of (95) and (96) to $P \cup Q$. Let p^* be the real root of $p^3 + p - 1 = 0$ ($p^* \approx 0.682$), and let $p^* \leq p \leq 1$. Then the Lyapunov exponents $\lambda_{1,2}$ satisfy*

$$\Psi_{H_k} = \sum_{a,b=1}^{\infty} R(a,b) \log \psi_k(a,b) \leq n_S |\lambda_{1,2}| \leq \sum_{a,b=1}^{\infty} R(a,b) \log \phi_k(a,b) = \Phi_{H_k}, \tag{101}$$

for $k \in \{1, 2, \infty\}$, where all non-zero $R(i,j)$ are given by Lemmas 5.2 - 5.5, $\psi_k(a,b)$ and $\phi_k(a,b)$ are given by Lemma 5.7, and $n_S = \frac{2}{p} - 1$. We therefore have $\Psi_H \leq$

$|\lambda_{1,2}| \leq \Phi_H$, *where*

$$\Psi_H = \frac{p}{2-p} \cdot \sup_{i=1,2,\infty} \{\Psi_{H_i}\} \text{ and } \Phi_H = \frac{p}{2-p} \cdot \inf_{i=1,2,\infty} \{\Phi_{H_i}\}. \tag{102}$$

As noted earlier, the statement given by (101), with appropriate adjustments to the terms $R(a,b)$, $\psi_k(a,b)$, $\phi_k(a,b)$, and n_S, could be obtained in the case where $P = \{(x,y) \in \mathbb{T}^2 : y \leq p\}$, $Q = \{(x,y) \in \mathbb{T}^2 : x \leq q\}$, and $p \neq q$. In fact, bounds such as these can be obtained in principle for any map whose domain possesses a region to which the return map is hyperbolic, and the possible variants of the Jacobian matrices of the return map share an invariant cone. We discuss examples of such maps in Section 5.6.

5.2 Finding the return time distribution

We will now find a return time distribution for points in a specific region of our domain (the region S), which will serve as the distribution element of our bounds. Note that return (or recurrence) times have been studied in detail by Young [61] for systems with similar properties to linked twist maps - systems which are hyperbolic on large parts of their domain, though not necessarily its entirety. The tail of the return time distribution was of particular interest in this case, as a sufficiently fast decay rate (polynomial or greater) implied the existence of a Central Limit Theorem for all observables ϕ with $\int \phi d\mu = 0$, in addition to an equivalently fast mixing rate (decay of correlations). In our case, the observable we are interested in is the first return time; the number of iterates it takes for a point in S to return to S for the first time.

Throughout this section we will use F and G to refer to the non-identity parts of the maps, F_1 and G_1. For example, if we say a point x returns to S after experiencing $GGFF$, we really mean that it returned following application of the sequence $G_1F_2G_1F_1G_2F_1$. We will also use the following probability measure for the

points in S, derived from the Lebesgue measure μ_L:

$$\mu_S(A) = \frac{\mu_L(A)}{p^2}. \tag{103}$$

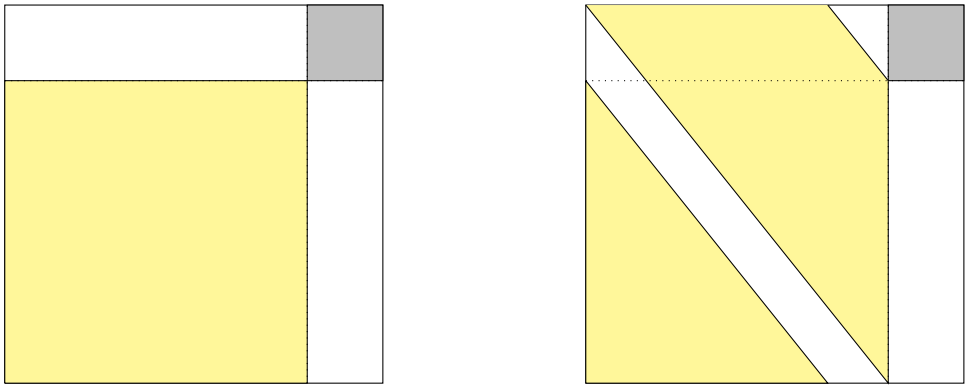
We require a matrix to be hyperbolic in order to obtain its invariant cone. In the case of the map H, an orbit experiences a hyperbolic matrix provided it undergoes application of both F and G, or in other words, provided the orbit is not just a sequence of either Fs or Gs. The following lemma guarantees that orbits beginning in S must experience a hyperbolic matrix before returning to S, or equivalently, the return map to S is itself hyperbolic.

Lemma 5.1. *Let $X = (x, y) \in S$. The return map to S of X under H, R_X, is of the form*

$$R_X \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & \frac{k}{p} \\ \frac{n-k+1}{p} & 1 + \frac{k(n-k+1)}{p^2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{104}$$

where n is the number of iterates of H required for X to return to S, and $1 \leq k \leq n$. R_X is hyperbolic. The only possible sequences through which X can return to S under H are those of the form $G^{n-k+1}F^k$.

Proof. We have two possibilities for $H(X)$ after one iterate: F or GF. If $H = F$, then after one iterate of H we find ourselves in A. We then must undergo application by only Fs until we re-enter S, at which point we undergo application by at least one G, since we apply G after F for a full iterate of H. If this application of G moves us into B, then we undergo application by only Gs until we next re-enter S, at which point we have returned. Hence the sequences $G^{n-k+1}F^k$ are possible for $k \geq 2$, where n is the number of iterates of H. If $H = GF$, then we either return immediately, or find ourselves in B. While in B we can only undergo application by Gs until we return to S, hence the sequences G^nF are possible. In other words, the only possible sequences for points in S is a string of Fs followed by a string of Gs, of the form $G^{n-k+1}F^k$ for $k \geq 1$.



(a) The set S is shown shaded yellow.

(b) The yellow set shows $G^{-1}(S)$.

Figure 35: The sets S and $G^{-1}(S)$. We apply F^{-1} to $G^{-1}(S)$ to find $H^{-1}(S)$.

Finally,

$$F^k = \begin{pmatrix} 1 & \frac{k}{p} \\ 0 & 1 \end{pmatrix},$$

$$G^{n-k+1} = \begin{pmatrix} 1 & 0 \\ \frac{n-k+1}{p} & 1 \end{pmatrix},$$

and the composition of these gives us the required result for R_X, which has eigenvalues

$$\lambda_\pm = \frac{1 + \frac{k(n-k+1)}{p^2} \pm \sqrt{(1 + \frac{k(n-k+1)}{p^2})^2 - 4}}{2}.$$

Since $n \geq k \geq 1$, we have that $\lambda_+ > 1$ and $0 < \lambda_- < 1$, and hence R_X is hyperbolic.

□

To find the return time distribution, we calculate the pre-images of the set S under H, and see where the pre-images intersect with the original set S. To do this, we first calculate $G^{-1}(S)$, which is shown in Figure 35(b). We then find which parts of $G^{-1}(S)$ are in P, and calculate the image of these under F^{-1} to obtain $H^{-1}(S)$, which is shown in Figure 36.

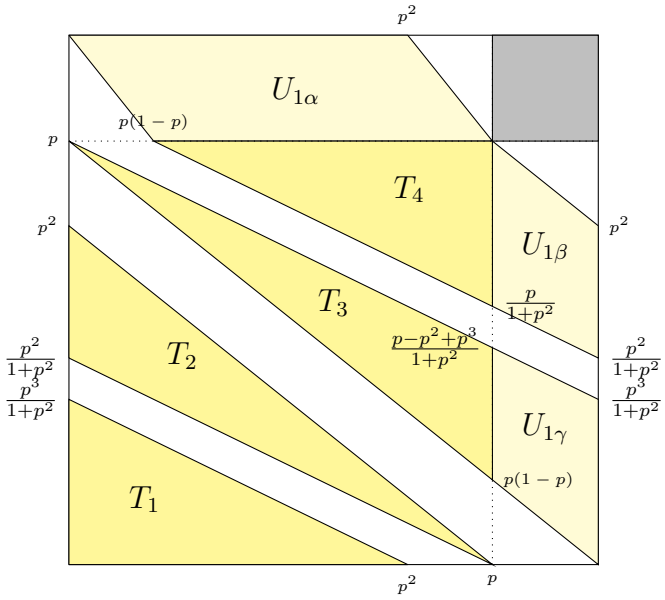


Figure 36: $H^{-1}(S)$. The four bright yellow triangles form the set $H^{-1}(S) \cap S = R_1 = R_{FG}$. The three pale yellow quadrilaterals form the set $U_1 = H^{-1}(S) \backslash S$.

$H^{-1}(S)$ consists of four triangles which intersect with S, and three quadrilaterals which are outside of S. The quadrilateral in B is a parallelogram, while the two in A are not. The three quadrilaterals are the sets we will continue to iterate backwards in order to find points which return on iterates later than the first. The four triangles consist of all the points which return after the first iterate of H, under the sequence GF. From this point onwards, we will refer to these four triangles as R_{GF} - the points which have returned under the sequence GF - and follow a similar naming convention for points which return after experiencing other sequences. We will also refer to the points which have not returned by the first iterate (the three quadrilaterals) as U_1, with the parallelogram referred to as $U_{1\alpha}$, and the remaining quadrilaterals as $U_{1\beta}$ and $U_{1\gamma}$ respectively (see Figure 36).

Lemma 5.2. *The measure of the set R_{GF} is given by $\mu_S(R_{GF}) = \frac{2p^3}{1+p^2}$.*

Proof. We need to calculate the measure of each of the four triangles which intersect

with S, and sum these in order to find $\mu_S(R_{GF})$. Let us label the triangles from bottom (corner at $(0,0)$) to top (corner at (p,p)) as T_1 through T_4. First we note that T_1 and T_4 are identical triangles, as are T_2 and T_3, so we need only calculate $\mu_S(T_1)$ and $\mu_S(T_2)$.

T_1 has corners at $(0,0)$, $(p^2,0)$ and $(0,\frac{p^3}{1+p^2})$. Hence

$$\mu_L(T_1) = \frac{p^5}{2(1+p^2)}.$$

Similarly, T_2 has corners $(p,0)$, $(0,\frac{p^2}{1+p^2})$ and $(0,p^2)$. Hence

$$\mu_L(T_2) = \frac{p^5}{2(1+p^2)}.$$

Hence $\mu_L(R_{GF}) = \frac{2p^5}{1+p^2}$, and therefore $\mu_S(R_{GF}) = \frac{2p^3}{1+p^2}$. □

Lemma 5.2 tells us that the measure of the points yet to return after the first iterate (U_1) is

$$\mu_S(U_1) = 1 - \frac{2p^3}{1+p^2}.$$

From now on we will only iterate the points in U_1 and see how these intersect with S. Let us begin by considering $U_{1\alpha}$. We first find its image under G^{-1}. The corners of $U_{1\alpha}$ are mapped under G^{-1} as follows:

$$G^{-1}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}1 & 0\\ \frac{-1}{p} & 1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix},$$

$$G^{-1}\begin{pmatrix}p(1-p)\\p\end{pmatrix} = \begin{pmatrix}p(1-p)\\2p-1\end{pmatrix},$$

$$G^{-1}\begin{pmatrix}p\\p\end{pmatrix} = \begin{pmatrix}p\\p-1\end{pmatrix},$$

$$G^{-1}\begin{pmatrix}p^2\\1\end{pmatrix} = \begin{pmatrix}p^2\\1-p\end{pmatrix}.$$

$G^{-1}(U_{1\alpha})$ is shown in Figure 37(a). We must now apply F^{-1} to the dark red region, which we shall label Δ, in order to find $H^{-1}(U_{1\alpha})$. The corners of Δ are:

$$\left(\frac{p(1-p)}{2}, p\right),\ \Big(p(1-p), p\Big),\ \Big(p^2, 1-p\Big),\ \left(\frac{p(p+1)}{2}, 0\right),\ \Big(p^2, 0\Big),\ \Big(p(1-p), 2p-1\Big),$$

which are mapped under F^{-1} as follows:

$$F^{-1}\begin{pmatrix}\frac{p(1-p)}{2}\\ p\end{pmatrix} = \begin{pmatrix}1 & \frac{-1}{p}\\ 0 & 1\end{pmatrix}\begin{pmatrix}\frac{p(1-p)}{2}\\ p\end{pmatrix} = \begin{pmatrix}\frac{p(1-p)}{2}-1\\ p\end{pmatrix},$$

$$F^{-1}\begin{pmatrix}p(1-p)\\ p\end{pmatrix} = \begin{pmatrix}1 & \frac{-1}{p}\\ 0 & 1\end{pmatrix}\begin{pmatrix}p(1-p)\\ p\end{pmatrix} = \begin{pmatrix}p(1-p)-1\\ p\end{pmatrix},$$

$$F^{-1}\begin{pmatrix}p^2\\ 1-p\end{pmatrix} = \begin{pmatrix}1 & \frac{-1}{p}\\ 0 & 1\end{pmatrix}\begin{pmatrix}p^2\\ 1-p\end{pmatrix} = \begin{pmatrix}\frac{p^3+p-1}{p}\\ 1-p\end{pmatrix},$$

$$F^{-1}\begin{pmatrix}\frac{p(p+1)}{2}\\ 0\end{pmatrix} = \begin{pmatrix}1 & \frac{-1}{p}\\ 0 & 1\end{pmatrix}\begin{pmatrix}\frac{p(p+1)}{2}\\ 0\end{pmatrix} = \begin{pmatrix}\frac{p(p+1)}{2}\\ 0\end{pmatrix},$$

$$F^{-1}\begin{pmatrix}p^2\\ 0\end{pmatrix} = \begin{pmatrix}1 & \frac{-1}{p}\\ 0 & 1\end{pmatrix}\begin{pmatrix}p^2\\ 0\end{pmatrix} = \begin{pmatrix}p^2\\ 0\end{pmatrix},$$

$$F^{-1}\begin{pmatrix}p(1-p)\\ 2p-1\end{pmatrix} = \begin{pmatrix}1 & \frac{-1}{p}\\ 0 & 1\end{pmatrix}\begin{pmatrix}p(1-p)\\ 2p-1\end{pmatrix} = \begin{pmatrix}\frac{-p^3+p^2-2p+1}{p}\\ 2p-1\end{pmatrix}.$$

$H^{-1}(U_{1\alpha})$ is shown in Figure 37(b). Note that in order for this picture to be accurate, we require the following condition to be met:

$$p^3 + p - 1 \geq 0 \text{ (or } p \geq 0.682...). \tag{105}$$

If we do not meet this condition, then the change in gradient which occurs at the points $(\frac{p^3+p-1}{p}, 1-p)$ and $(\frac{-p^3+p^2-2p+1}{p}, 2p-1)$ will instead be contained within the unreturned regions, and will continue into later iterates, making the return time

distribution more complicated than in the following construction. From this point forward, we will assume that condition (105) is met.

In order to calculate the measure of R_{GGF}, we first calculate the measure of the two triangles in B and the parallelogram in A. The triangles have corners at $(0,1)$, $(\frac{p(1-p)}{2}, p)$, $(p(1-p), p)$ and $(p^2, 1)$, $(\frac{p(1+p)}{2}, 1)$, (p, p) respectively, and so are identical to each other. Let us refer to these triangles as τ_{2a} and τ_{2b}. Note that the 2 refers to the number of iterates of H^{-1} we have taken to obtain these triangles, and we will follow this labelling convention for the future unreturned parts of the pre-images of these triangles. They both have Lebesgue measure given by

$$\mu_L(\tau_{2a}) = \frac{p(1-p)^2}{4},$$

and thus we have

$$\mu_S(\tau_{2a} \cup \tau_{2b}) = \frac{(1-p)^2}{2p}. \tag{106}$$

The parallelogram, which we label ρ, has corners at $(p, \frac{p}{1+p^2})$, $(p, \frac{p-p^2+p^3}{1+p^2})$, $(1, \frac{p^2}{1+p^2})$ and $(1, \frac{p^3}{1+p^2})$, and thus has Lebesgue measure

$$\mu_L(\rho) = \frac{p^2(1-p)^2}{1+p^2},$$

and thus

$$\mu_S(\rho) = \frac{(1-p)^2}{1+p^2}. \tag{107}$$

We now consider $U_{1\beta}$ and $U_{1\gamma}$, neither of which undergo application by G^{-1}. The corners of $U_{1\beta}$ are mapped under F^{-1} as follows:

$$F^{-1}\begin{pmatrix} p \\ p \end{pmatrix} = \begin{pmatrix} 1 & \frac{-1}{p} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} p \\ p \end{pmatrix} = \begin{pmatrix} p-1 \\ p \end{pmatrix},$$

$$F^{-1}\begin{pmatrix} 1 \\ p^2 \end{pmatrix} = \begin{pmatrix} 1 & \frac{-1}{p} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ p^2 \end{pmatrix} = \begin{pmatrix} 1-p \\ p^2 \end{pmatrix},$$

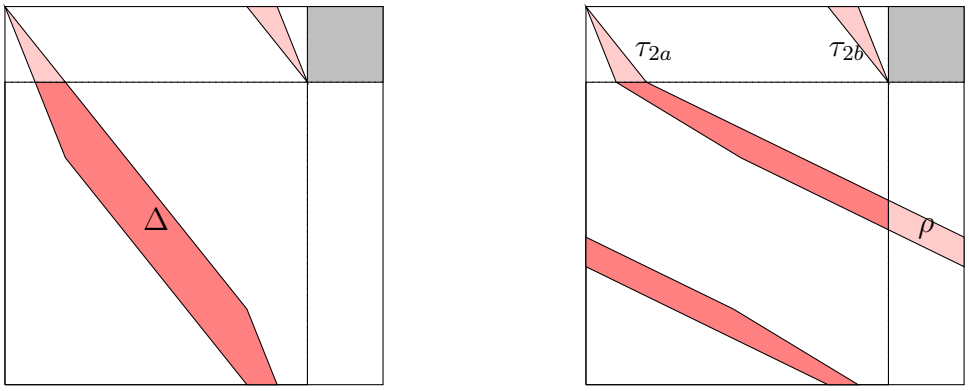


Figure 37: The pre-image of $U_{1\alpha}$ under G, and then under H.

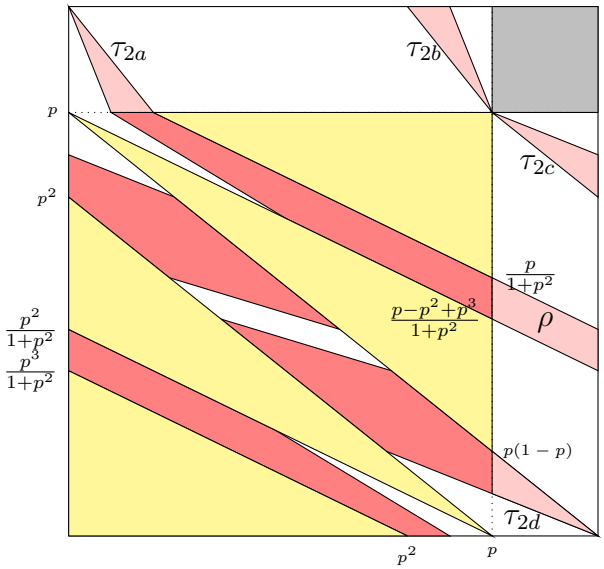


Figure 38: $H^{-1}(U_1)$ is shown in red, with dark red indicating the regions which have returned after two applications of H^{-1} and light red the regions yet to return (U_2).

$$F^{-1}\begin{pmatrix}1\\ \frac{p^2}{1+p^2}\end{pmatrix}=\begin{pmatrix}1 & \frac{-1}{p}\\ 0 & 1\end{pmatrix}\begin{pmatrix}1\\ \frac{p^2}{1+p^2}\end{pmatrix}=\begin{pmatrix}1-\frac{p}{1+p^2}\\ \frac{p^2}{1+p^2}\end{pmatrix},$$

$$F^{-1}\begin{pmatrix}p\\ \frac{p}{1+p^2}\end{pmatrix}=\begin{pmatrix}1 & \frac{-1}{p}\\ 0 & 1\end{pmatrix}\begin{pmatrix}p\\ \frac{p}{1+p^2}\end{pmatrix}=\begin{pmatrix}\frac{p^3+p-1}{1+p^2}\\ \frac{p}{1+p^2}\end{pmatrix}.$$

Similarly, for the corners of $U_{1\gamma}$:

$$F^{-1}\begin{pmatrix}1\\ 0\end{pmatrix}=\begin{pmatrix}1 & \frac{-1}{p}\\ 0 & 1\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix}=\begin{pmatrix}1\\ 0\end{pmatrix},$$

$$F^{-1}\begin{pmatrix}p\\ p(1-p)\end{pmatrix}=\begin{pmatrix}1 & \frac{-1}{p}\\ 0 & 1\end{pmatrix}\begin{pmatrix}p\\ p(1-p)\end{pmatrix}=\begin{pmatrix}2p-1\\ p(1-p)\end{pmatrix},$$

$$F^{-1}\begin{pmatrix}p\\ \frac{p^3-p^2+p}{1+p^2}\end{pmatrix}=\begin{pmatrix}1 & \frac{-1}{p}\\ 0 & 1\end{pmatrix}\begin{pmatrix}p\\ \frac{p^3-p^2+p}{1+p^2}\end{pmatrix}=\begin{pmatrix}\frac{p^3-p^2+2p-1}{1+p^2}\\ \frac{p^3-p^2+p}{1+p^2}\end{pmatrix},$$

$$F^{-1}\begin{pmatrix}1\\ \frac{p^3}{1+p^2}\end{pmatrix}=\begin{pmatrix}1 & \frac{-1}{p}\\ 0 & 1\end{pmatrix}\begin{pmatrix}1\\ \frac{p^3}{1+p^2}\end{pmatrix}=\begin{pmatrix}\frac{1}{1+p^2}\\ \frac{p^3}{1+p^2}\end{pmatrix}.$$

We end up with two quadrilaterals which stretch into S from A, each with a triangle remaining in A (see Figure 38). The two triangles, which we label τ_{2c} and τ_{2d} from top to bottom, have corners at (p,p), $(1,p^2)$, $(1,\frac{p(p+1)}{2})$ and $(1,0)$, $(p,p(1-p))$ and $(p,\frac{p(1-p)}{2})$ respectively, and thus have identical dimensions to τ_{2a} and τ_{2b}. We therefore have that

$$\mu_S(\tau_{2c}\cup\tau_{2d})=\frac{(1-p)^2}{2p}. \tag{108}$$

We now use this information to calculate the measure of the set of points which return under each of the possible sequences of matrices we can have in two iterates of H. These sets are R_{GGF}, which consists of the points belonging to $U_{1\alpha}$ which return on the 2nd iterate of H, and R_{GFF}, which consists of the points returning from $U_{1\beta}$ and $U_{1\gamma}$.

Lemma 5.3. *The measures of the sets of points which return to S after undergoing application by the sequence GGF, R_{GGF}, and the sequence GFF, R_{GFF}, respectively are*

$$\mu_S(R_{GGF}) = \mu_S(R_{GFF}) = \frac{-3p^4 + 2p^3 + 2p - 1}{2p(1+p^2)}.$$

Proof. First, consider the points which return from $U_{1\alpha}$. We have that

$$\mu_L(U_{1\alpha}) = p^2(1-p),$$

and thus

$$\mu_S(U_{1\alpha}) = (1-p).$$

Thus, using (106) and (107), we have that

$$\begin{aligned}\mu_S(R_{GGF}) &= \mu_S(U_{1\alpha}) - \mu_S(\tau_{2a} \cup \tau_{2b}) - \mu_S(\rho),\\ &= (1-p) - \frac{(1-p)^2}{2p} - \frac{(1-p)^2}{1+p^2},\\ &= \frac{-3p^4 + 2p^3 + 2p - 1}{2p(1+p^2)}.\end{aligned}$$

Now consider $U_{1\beta}$ and $U_{1\gamma}$. We can calculate $\mu_S(U_{1\beta})$ by considering three triangles: T_4, the triangle with corners at (p,p), $(1,p)$ and $(1,p^2)$, δ_1 say, and the triangle with corners at $(p-p^2,p)$, $(1,p)$ and $(1,\frac{p^2}{1+p^2})$, σ_1 say. We have that

$$\mu_S(\sigma_1) = \frac{(p^2-p+1)^2}{2p(1+p^2)},$$

$$\mu_S(\delta_1) = \frac{(1-p)^2}{2p},$$

$$\mu_S(T_4) = \frac{p^3}{2(1+p^2)},$$

and so

$$\begin{aligned}\mu_S(U_{1\beta}) &= \mu_S(\sigma_1) - \mu_S(\delta_1) - \mu_S(T_4),\\ &= \frac{p(1-p^2)}{2(1+p^2)}.\end{aligned}$$

Similarly, we find that

$$\mu_S(U_{1\gamma}) = \frac{p(1-p^2)}{2(1+p^2)}.$$

Hence, using (108), we have

$$\begin{aligned}\mu_S(R_{GFF}) &= \mu_S(U_{1\beta} \cup U_{1\gamma}) - \mu_S(\tau_{2c} \cup \tau_{2d}),\\ &= \frac{p(1-p^2)}{(1+p^2)} - \frac{(1-p)^2}{2p},\\ &= \frac{-3p^4+2p^3+2p-1}{2p(1+p^2)}.\end{aligned}$$

□

We now need to iterate U_2 backwards. Let us begin with the parallelogram ρ, which only undergoes application by F^{-1}. We know from previous calculations where the bottom edge of $U_{1\beta}$ and the top edge of $U_{1\gamma}$ are mapped to, and by continuity ρ must be mapped in between them. Thus, provided that condition (105) is met, ρ will entirely return on the 3rd iterate, filling the region between the $F^{-1}(U_{1\beta})$, $F^{-1}(U_{1\gamma})$, T_2 and T_3 (see Figure 39).

Now consider τ_{2c}, which also only undergoes application by F^{-1}. Its corners are mapped as follows:

$$F^{-1}\begin{pmatrix}p\\p\end{pmatrix} = \begin{pmatrix}1 & \frac{-1}{p}\\0 & 1\end{pmatrix}\begin{pmatrix}p\\p\end{pmatrix} = \begin{pmatrix}p-1\\p\end{pmatrix},$$

$$F^{-1}\begin{pmatrix}1\\p^2\end{pmatrix} = \begin{pmatrix}1 & \frac{-1}{p}\\0 & 1\end{pmatrix}\begin{pmatrix}1\\p^2\end{pmatrix} = \begin{pmatrix}1-p\\p^2\end{pmatrix},$$

$$F^{-1}\begin{pmatrix}1\\\frac{p(1+p)}{2}\end{pmatrix} = \begin{pmatrix}1 & \frac{-1}{p}\\0 & 1\end{pmatrix}\begin{pmatrix}1\\\frac{p(1+p)}{2}\end{pmatrix} = \begin{pmatrix}\frac{1-p}{2}\\\frac{p(1+p)}{2}\end{pmatrix}.$$

Thus $F^{-1}(\tau_{2c})$ is a triangle which stretches into the region between $F^{-1}(U_{1\beta})$ and T_3, meeting with the corner of $U_{1\beta}$ at $(1-p, p^2)$, as shown in Figure 39.

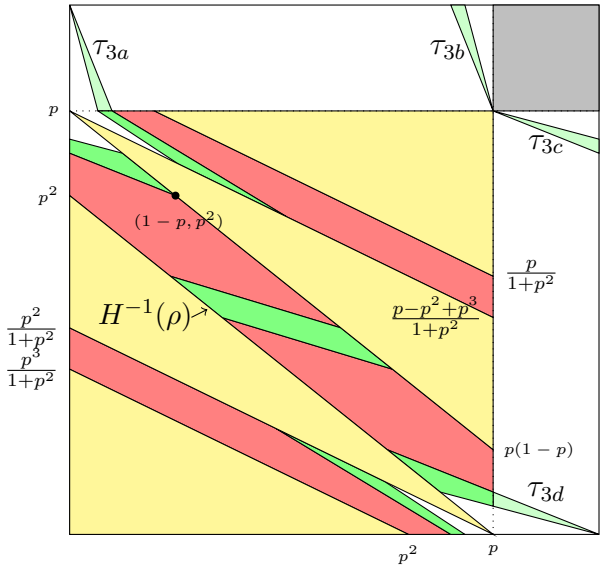


Figure 39: $H^{-1}(U_2)$ is shown in green, with bright green indicating the regions which have returned after three applications of H^{-1} and pale green the regions yet to return (U_3).

We now consider what will happen to this triangle on future iterates. We define τ_{3c} to be the points in τ_{2c} which did not return on the 3rd iterate, and note that τ_{3c} has corners at (p,p), $(1, \frac{p(1+p)}{2})$, $(1, \frac{p(2+p)}{3})$. Hence the bottom edge of τ_{3c} is the top edge of τ_{2c}. We therefore find that $F^{-1}(\tau_{3c})$ will lie on top of $F^{-1}(\tau_{2c})$, filling in more of the region between $F^{-1}(U_{1\beta})$ and T_3. We can repeat this process indefinitely to find the corners of the triangle τ_{nc} for any $n \geq 2$:

$$\begin{pmatrix} p \\ p \end{pmatrix}, \begin{pmatrix} 1 \\ \frac{p(n-2+p)}{n-1} \end{pmatrix}, \begin{pmatrix} 1 \\ \frac{p(n-1+p)}{n} \end{pmatrix}. \tag{109}$$

From this, we can see that the images of each triangle $\tau_{(n-1)c}$ will consist of τ_{nc} and a quadrilateral in S which lies on top of the quadrilateral in S belonging to the image of $\tau_{(n-2)c}$. The measure of τ_{nc} tells us the measure of points which do not return on iterate n, corresponding to this sequence of triangles. In order to find the measure of all unreturned points, we also have to consider the other three triangles: τ_{na}, τ_{nb},

and τ_{nd}.

$F^{-1}(\tau_{2d})$ behaves in a similar way to $F^{-1}(\tau_{2c})$, however it instead stretches into the region between $F^{-1}(U_{1\gamma})$ and T_2. The corners of τ_{3d} are $(1,0)$, $(p, \frac{p(1-p)}{2})$, $(p, \frac{p(1-p)}{3})$. In general, the corners for τ_{nd} are

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} p \\ \frac{p(1-p)}{n-1} \end{pmatrix}, \begin{pmatrix} p \\ \frac{p(1-p)}{n} \end{pmatrix}. \tag{110}$$

Both τ_{na} and τ_{nb} first undergo application by G^{-1}, then any regions in S undergo application by F^{-1}. The triangles τ_{na} and τ_{nb} are identical to τ_{nd} and τ_{nc} respectively, for any $n \geq 2$, except that their x and y coordinates have been swapped around. For the corners of τ_{na}, we have

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{p(1-p)}{n-1} \\ p \end{pmatrix}, \begin{pmatrix} \frac{p(1-p)}{n} \\ p \end{pmatrix}, \tag{111}$$

and for τ_{nb}, we have

$$\begin{pmatrix} p \\ p \end{pmatrix}, \begin{pmatrix} \frac{p(n-2+p)}{n-1} \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{p(n-1+p)}{n} \\ 1 \end{pmatrix}. \tag{112}$$

The remainders of the triangles $\tau_{(n-1)a}$ and $\tau_{(n-1)b}$ are mapped into the regions between $H^{-1}(U_{1\alpha})$ and T_3, and $H^{-1}(U_{1\alpha})$ and T_2 respectively (see Figure 39). For all of these triangles, we can calculate the measure of the set of returning points by subtracting the measure of τ_{ni} from that of $\tau_{(n-1)i}$ for $i \in \{a, b, c, d\}$. Since ρ fully returns on the third iterate, these four triangles comprise all of the unreturned points from the fourth iterate onwards. The following lemma tells us the measures of the sets of returning points on the third iterate.

Lemma 5.4. *The measures of the sets of points which return to S after undergoing application by the sequence $GGGF$, R_{GGGF}, and the sequence GFF, R_{GFFF}, respectively are*

$$\mu_S(R_{GGGF}) = \mu_S(R_{GFFF}) = \frac{(1-p)^2}{3p}.$$

The measure of the set of points which return to S after undergoing application by the sequence $GGFF$, R_{GGFF}, is

$$\mu_S(R_{GGFF}) = \frac{(1-p)^2}{1+p^2}.$$

Proof. First consider R_{GGFF}. The only points which experience the sequence $GGFF$ are those originally belonging to $U_{1\alpha}$, which then reside in A after the second iterate. The only points which fulfil this condition are those in ρ, which entirely returns on the third iterate, and so, using (107), we have

$$\mu_S(R_{GGFF}) = \mu_S(\rho) = \frac{(1-p)^2}{1+p^2}.$$

Now consider R_{GGGF}. The points which experience the sequence $GGGF$ are those which return from τ_{2a} and τ_{2b} on the third iterate. We therefore have

$$\begin{aligned}\mu_S(R_{GGGF}) &= \mu_S(\tau_{2a} - \tau_{3a}) + \mu_S(\tau_{2b} - \tau_{3b}),\\ &= 2 \cdot \frac{1}{2} \cdot (\frac{(1-p)^2}{2p} - \frac{(1-p)^2}{6p}),\\ &= \frac{(1-p)^2}{3p},\end{aligned}$$

since τ_{na} and τ_{nb} have the same measure for $n \geq 2$. The calculation for $\mu_S(R_{GFFF})$ is identical, since the measures of τ_{na} and τ_{nb} are identical to those of τ_{nd} and τ_{nc} respectively. □

From the fourth iterate onwards, the returning points on the n^{th} iterate feed in from the triangles $\tau_{(n-1)}$ exclusively. We can therefore calculate the measures of the sets of returning points for all remaining iterates, since we know the coordinates for the corners of each τ_n.

Lemma 5.5. *Let $n \geq 4$. The measures of the sets of points which return to S after experiencing the sequence G^nF, R_{G^nF}, and the sequence GF^n, R_{GF^n}, respectively are*

$$\mu_S(R_{G^nF}) = \mu_S(R_{GF^n}) = \frac{2(1-p)^2}{np(n-1)(n-2)}.$$

Proof. As in the proof for Lemma 5.4, we simply need to calculate the difference between the $(n-1)^{th}$ and n^{th} triangles. We have

$$\begin{aligned}\mu_S(R_{G^nF}) &= \mu_S(\tau_{(n-1)a} - \tau_{na}) + \mu_S(\tau_{(n-1)b} - \tau_{nb}),\\ &= 2 \cdot \frac{1}{2} \cdot \big(\frac{(1-p)^2}{p(n-2)(n-1)} - \frac{(1-p)^2}{np(n-1)}\big),\\ &= \frac{2(1-p)^2}{np(n-1)(n-2)},\end{aligned}$$

The calculation for $\mu_S(R_{GF^n})$ is identical for the same reason as in the proof of Lemma 5.4. □

In Section 5.4 we combine the results from Lemmas 5.2 - 5.5, which give an explicit formula for the frequency with which each possible $M_{a,b}$ occurs, with those of Lemma 5.7 in Section 5.3, which gives bounds upon the norms of each possible sequence of matrices, to obtain the bounds Ψ_H and Φ_H. A quick check to show the validity of this distribution is given by Kac's lemma [26], which states that in an ergodic, measure-preserving transformation, the expected first return time to a set S is $1/\mu(S)$, where μ is the normalized measure on the domain. In other words, if D is the domain, and $A \subset D$, then

$$\mu(A) = \frac{\mu_L(A)}{\mu_L(D)}.$$

In order to find the normalized measure in this case, we need to find the measure of S with respect to our entire domain, which in this case is $A \cup B \cup S$. We have

$$\mu_L(S) = p^2 \text{ and } \mu_L(A \cup B \cup S) = p(2-p),$$

so we have

$$\mu(S) = \frac{\mu_L(S)}{\mu_L(A \cup B \cup S)} = \frac{p}{2-p}.$$

Kac's Lemma then tells us that, since H is ergodic on $P \cup Q$ w.r.t. the normalized Lebesgue measure, the average first return time to S is

$$n_S = \frac{1}{\mu(S)} = \frac{2}{p} - 1. \tag{113}$$

We can now check whether our distribution agrees with Kac's Lemma by finding the expected return time to S. We obtain the following:

$$\begin{aligned} n_S &= R(1,1) + 2\cdot(R(1,2)+R(2,1)) \\ &\quad + 3\cdot(R(1,3)+R(2,2)+R(3,1)) + \sum_{n=4}^{\infty} n\cdot(R(1,n)+R(n,1)), \\ &= \frac{-4p^4+7p^3-6p^2+7p-2}{p(1+p^2)} + \frac{2(1-p)^2}{p}\big(1+2\sum_{r=2}^{\infty}\frac{1}{r(r+1)}\big), \\ &= \frac{-4p^4+7p^3-6p^2+7p-2}{p(1+p^2)} + \frac{4(1-p)^2}{p}, \ \big(\text{since } \sum_{r=2}^{\infty}\frac{1}{r(r+1)} = \frac{1}{2}\ \big) \\ &= \frac{2}{p} - 1. \end{aligned}$$

Hence our distribution gives the same result as Kac's Lemma. We will now use this distribution, in combination with the upper and lower bounds for the l_∞, l_1 and l_2 norms within a globally invariant cone calculated by Sturman and Thiffeault [54] in the case of random products of shear matrices, to calculate explicit upper and lower bounds for the Lyapunov exponents of this family of linked twist maps.

5.3 Invariant cones and bounds on norms

In this section we will explicitly calculate upper and lower bounds for the quantity $\frac{||M_{a,b}v||}{||v||}$ in a particular invariant cone, where the matrix $M_{a,b}$ is defined below. First we discuss the cones we will use to obtain these bounds.

Consider the matrices

$$M_{ab} = B^b A^a \tag{114}$$

where

$$A = \begin{pmatrix} 1 & \frac{1}{p} \\ 0 & 1 \end{pmatrix},$$

$$B = \begin{pmatrix} 1 & 0 \\ \frac{1}{p} & 1 \end{pmatrix},$$

and so

$$M_{ab} = \begin{pmatrix} 1 & \frac{a}{p} \\ \frac{b}{p} & 1 + \frac{ab}{p^2} \end{pmatrix}.$$

These matrices and their invariant cones were studied by Sturman and Thiffeault [54], in the context of bounding the Lyapunov exponents of random products of shear matrices by splitting up random sequences of As and Bs into blocks of the form M_{ab}. Linked twist maps can be treated in much the same way as random products of shear matrices - if we imagine iterating our map H discussed in Section 5.2, and taking the norm after N iterations, in a similar way to (61), we obtain (for some $v_0 \in \mathbb{R}^2$, $n \in \mathbb{N}$)

$$\begin{aligned} ||H^N v_0|| &= ||GFFGGF...FFGv_0||, \\ &= \frac{||M_{a_n,b_n} v_{n-1}||}{||v_{n-1}||} \cdot ... \cdot \frac{||M_{a_1,b_1} v_0||}{||v_0||} \cdot ||v_0||, \end{aligned}$$

where the maps F and G have taken the place of the matrices A and B mentioned above. Hence, in order to calculate $||H^N v_0||$, we need to bound each of these $\frac{||M_{a_i,b_i} v_{i-1}||}{||v_{i-1}||}$, and find out how n relates to N. Fortunately, due to Kac's Lemma, the latter of these problems is relatively simple. We know that the average return time to S is given by $n_S = \frac{2}{p} - 1$, and that each M_{a_i,b_i} represents $a_i + b_i - 1$ iterates of H. Thus, since each M_{a_i,b_i} corresponds to a return to S, we have

$$\mathbb{E}(a_i + b_i - 1) = \frac{2}{p} - 1,$$

and subsequently

$$n = \frac{N}{\mathbb{E}(a_i + b_i - 1)} = \frac{N}{\frac{2}{p} - 1}. \tag{115}$$

We will now consider invariant cones for the matrices M_{ab}. Let us assume that both α and β are positive, which guarantees that M_{ab} is a hyperbolic matrix. Note that the sequences of matrices we studied when finding the return time distribution for S in Section 5.2 were all of the form M_{ab}.

Recall Lemma 4.2, which gives us an invariant cone for any hyperbolic matrix formed by shear composition - we need a cone which contains the unstable eigenvector of M_{ab}, but not the stable eigenvector. It is possible for the unstable eigenvector of M_{ab} to be anywhere within the cone

$$C = \big\{(x, y) \in \mathbb{R}^2 : 0 < x/y < 1\big\}, \tag{116}$$

for any a,b, since $\alpha, \beta > 1$. So, in order for a cone to be invariant for any M_{ab}, it must at least contain C. Furthermore, all possible stable eigenvectors of M_{ab} lie outside of C, which makes C itself an invariant cone for M_{ab} by Lemma 4.2.

We now calculate upper and lower bounds for $\frac{||M_{a,b}v||}{||v||}$ for the l_∞, l_1 and l_2 norms for $v \in C$. To do this we make use of the existence of invariant cones [2] for the matrices $M_{a,b}$. Note that $M_{a,b}$ is diagonalizable with positive eigenvalues for any a, b, and thus any invariant cone for $M_{a,b}$ must contain the unstable eigenvector, and its interior must not contain the stable eigenvector [46]. We need to find a cone which is invariant under all possible $M_{a,b}$; we will refer to this cone as $C_{global}(p)$. The eigenvectors of $M_{a,b}$ are given by

$$v_\pm = \begin{pmatrix} 2 \\ \frac{b}{p} \pm \sqrt{\frac{4b}{a} + \frac{b^2}{p^2}} \end{pmatrix}.$$

A simple calculation then yields that the unstable eigenvector v_+ is always contained within the cone

$$C_{\text{global}}(p) = \big\{(x, y) \in \mathbb{R}^2 : 0 < \frac{x}{y} < p\big\}, \tag{117}$$

for any choice of a, b, and furthermore the stable eigenvector v_- lies outside this cone. Hence C_{global} is an invariant cone for every $M_{a,b}$. Note that C_{global} is *not* an invariant cone for H, as its Jacobian matrix D_xH is not always hyperbolic.

In order to bound $||H^N v||$, we find upper and lower bounds for the quantity $\frac{||M_{a,b}v||}{||v||}$, for $v \in C_{\text{global}}$ and each possible a, b, using the l_∞, l_1 and l_2 norms. We first

consider which vectors in tangent space maximise and minimise the growth rates of these norms following iteration by $M_{a,b}$.

Lemma 5.6. *Let $v^{(i)}_{\max}, v^{(i)}_{\min} \in \mathbb{R}^2$ be the vectors for which the quantity $\frac{||M_{a,b}v||_i}{||v||_i}$ is maximised and minimised for all $v \in \mathbb{R}^2$ and $i \in \{1, 2, \infty\}$, and let $v^{u,s}_{M^T_{a,b}M_{a,b}}$ be the unstable and stable eigenvectors of $M^T_{a,b}M_{a,b}$ respectively. Then:*

(i) $v^{(2)}_{\max} = v^{u}_{M^T_{a,b}M_{a,b}}$ *and* $v^{(2)}_{\min} = v^{s}_{M^T_{a,b}M_{a,b}}$.

(ii) $v^{(\infty)}_{\max} = (1, 1)$ *and* $v^{(\infty)}_{\min} = (-\frac{ab}{p^2} - \frac{a}{p} - 1, \frac{b}{p} + 1)$.

(iii) If $\frac{a}{p} \geq \frac{b}{b+p}$ *then* $v^{(1)}_{\max} = (0, 1)$. *Otherwise* $v^{(1)}_{\max} = (1, 0)$.

(iv) $v^{(1)}_{\min} = (\frac{ab}{p^2} + 1, -\frac{b}{p})$.

Proof. First we consider the l_2 norm. The quantity $\frac{||M_{a,b}v||_2}{||v||_2}$ is the spectral matrix norm which, since $M_{a,b}$ is non-singular, is by definition maximised or minimised when v is the unstable or stable eigenvector of the matrix $M^T_{a,b}M_{a,b}$ [52].

We now consider the l_∞ norm. We extend the unit l_∞ ball under $M_{a,b}$, as shown in Figure 25 in Chapter 4. We have that $\frac{1}{p}, a, b > 1$, hence by Lemma 4.4

$$v_{\max} = (1, 1), v_{\min} = (-\frac{ab}{p^2} - \frac{a}{p} - 1, \frac{b}{p} + 1),$$

where we have substituted $\alpha = \frac{a}{p}$ and $\beta = \frac{b}{p}$.

Similarly, consider the l_1 norm. The unit l_1 ball is shown in Figure 26 in Chapter 4. We have that $\frac{ab}{p^2} > 0$, and whether we have $\frac{a}{p} > \frac{b}{b+p}$ depends on the choice of a and b. Specifically, if the sequence $M_{a,b}$ we consider has large a (> 2), then $b = 1$. Similarly, if we have large b, $a = 1$. Hence, by Lemma 4.6, we have that if $\frac{a}{p} \geq \frac{b}{b+p}$,

$$v^{(1)}_{\max} = (0, 1)$$

and otherwise

$$v^{(1)}_{\max} = (1, 0).$$

In both cases, $v^{(1)}_{\min} = (\frac{ab}{p^2} + 1, -\frac{b}{p})$. □

Finding where $\frac{||M_{a,b}v||}{||v||}$ is maximised or minimised within C_{global} for the l_i norm involves checking if $v_{\max}^{(i)}$ and $v_{\min}^{(i)}$ are contained within the cone. If not, then the maximum and minimum are attained on the boundaries of the cone, as these norms are monotonic between their maximum and minimum. Lemma 5.7 gives us the bounds $\phi(a,b)$ and $\psi(a,b)$ for each a, b and for three different norms.

Lemma 5.7. *Let $v_{\max}^{(i)}, v_{\min}^{(i)}$ be given by Lemma 5.6 for $i \in \{1, 2, \infty\}$, and let $\lambda^{u,s}_{M_{a,b}^T M_{a,b}}$ be the unstable and stable eigenvalues of $M_{a,b}^T M_{a,b}$ respectively. We obtain the following upper bounds upon $\frac{||M_{a,b}v||_i}{||v||_i}$ for $v \in C_{\text{global}}$.*

1. *l_2 norm:*
$$\frac{||M_{a,b}v||_2}{||v||_2} \leq \sqrt{\lambda^u_{M_{a,b}^T M_{a,b}}}.$$

2. *l_1 norm:*
$$\frac{||M_{a,b}v||_1}{||v||_1} \leq 1 + \frac{a}{p} + \frac{ab}{p^2}.$$

3. *l_∞ norm:*
$$\frac{||M_{a,b}v||_\infty}{||v||_\infty} \leq 1 + b + \frac{ab}{p^2}.$$

We also obtain the following lower bounds upon $\frac{||M_{a,b}v||_i}{||v||_i}$ for $v \in C_{\text{global}}$.

1. *l_2 norm:*
$$\frac{||M_{a,b}v||_2}{||v||_2} \geq \min\left\{\sqrt{\frac{(p+\frac{a}{p})^2 + (1+b+\frac{ab}{p^2})^2}{1+p^2}}, \sqrt{(\frac{a}{p})^2 + (1+\frac{ab}{p^2})^2}\right\}.$$

2. *l_1 norm:*
$$\frac{||M_{a,b}v||_1}{||v||_1} \geq 1 + \frac{ab + pa + p^2 b}{p^2(1+p)}.$$

3. *l_∞ norm:*
$$\frac{||M_{a,b}v||_\infty}{||v||_\infty} \geq 1 + \frac{ab}{p^2}.$$

Proof. First we consider the l_2 norm. The quantity $\frac{||M_{a,b}v||_2}{||v||_2}$ is the spectral matrix norm which, since $M_{a,b}$ is non-singular, is by definition maximised or minimised when v is the unstable or stable eigenvector of the matrix $M_{a,b}^T M_{a,b}$. Its value varies monotonically between the maximum and minimum, thus we need to determine where these eigenvectors lie (i.e. in or out of C_{global}) in order to find where our maximum and minimum are obtained for $v \in C_{\text{global}}$. We have

$$M_{a,b}^T M_{a,b} = \begin{pmatrix} 1 + \frac{b^2}{p^2} & \frac{a}{p} + \frac{b}{p}(1 + \frac{ab}{p^2}) \\ \frac{a}{p} + \frac{b}{p}(1 + \frac{ab}{p^2}) & \frac{a^2}{p^2} + (1 + \frac{ab}{p^2})^2 \end{pmatrix},$$

which is a symmetric matrix, and thus its eigenvectors v_+ and v_- are orthogonal. Let

$$\Gamma_{a,b} = \frac{b^2}{p^2} + \frac{a^2}{p^2} + \frac{2ab}{p^2} + (\frac{ab}{p^2})^2,$$

then the eigenvalues of $M_{a,b}^T M_{a,b}$ are given by

$$\lambda_\pm = \frac{2 + \Gamma_{a,b} \pm \sqrt{\Gamma_{a,b}(4 + \Gamma_{a,b})}}{2},$$

and the corresponding eigenvectors are

$$v_\pm = \begin{pmatrix} 2(\frac{a}{p} + \frac{b}{p}(1 + \frac{ab}{p^2})) \\ \Gamma_{a,b} - \frac{2b^2}{p^2} \pm \sqrt{\Gamma_{a,b}(4 + \Gamma_{a,b})} \end{pmatrix} = \begin{pmatrix} r \\ s_\pm \end{pmatrix}.$$

We have that $r > 0$ since $\frac{1}{p} > 0$ and $s_+ > 0$ since $\Gamma_{a,b} > \frac{2b^2}{p^2}$. We can also infer that $s_- < 0$ due to orthogonality of eigenvectors. We therefore have that v_+ lies in the quadrant $\{(x, y) \in \mathbb{R}^2 : x, y \geq 0\}$, the same quadrant as C_{global}. In order for v_+ to lie in C_{global}, we require $\frac{r}{s_+} < p$. We have

$$s_+ > 2\Gamma_{a,b} - \frac{2b^2}{p^2} = \frac{4ab}{p^2} + \frac{2a^2}{p^2} + 2(\frac{ab}{p^2})^2 > \frac{2b}{p^2} + \frac{2a}{p^2} + \frac{2ab^2}{p^4} = \frac{r}{p},$$

and thus $s_+ > \frac{r}{p}$. Hence $v_+ \in C_{\text{global}}$, and therefore we attain our maximum at v_+. Since $s_- < 0$, $v_- \notin C_{\text{global}}$ and thus our minimum is attained at one of the

boundaries of the cone - either $v = (0, 1)$ or $v = (p, 1)$. We have, for $v = (u, v)$,

$$\frac{||M_{a,b}v||_2^2}{||v||_2^2} = \frac{(u + \frac{a}{p}v)^2 + (\frac{b}{p}u + (1 + \frac{ab}{p^2})v)^2}{u^2 + v^2},$$

so for our minimum, we choose the smaller of

$$\frac{||M_{a,b}(0,1)||_2^2}{||(0,1)||_2^2} = \frac{(p + \frac{a}{p})^2 + (1 + b + \frac{ab}{p^2})^2}{1 + p^2},$$

or

$$\frac{||M_{a,b}(1,1)||_2^2}{||(1,1)||_2^2} = (\frac{a}{p})^2 + (1 + \frac{ab}{p^2})^2.$$

Our maximum is given by the square root of the eigenvalue corresponding to the unstable eigenvector of $M_{a,b}^T M_{a,b}$, $\lambda^u_{M_{a,b}^T M_{a,b}}$. Specifically,

$$\frac{||M_{a,b}(r, s_+)||_2^2}{||(r, s_+)||_2^2} = \frac{(r + \frac{a}{p}s_+)^2 + (\frac{b}{p}r + (1 + \frac{ab}{p^2})s_+)^2}{r^2 + s_+^2}.$$

Now consider the l_∞ norm. For a vector $v = (x, y) \in C_{\text{global}}$, we have

$$||v||_\infty = |y|,$$

and since C_{global} is an invariant cone for M_{ab}, we have that

$$\frac{||M_{a,b}v||_\infty}{||v||_\infty} = \frac{|\frac{b}{p}x + (1 + \frac{ab}{p^2})y|}{|y|} = \frac{bx}{py} + 1 + \frac{ab}{p^2}.$$

Relabelling $\frac{x}{y} = z$, let

$$f(z) = 1 + \frac{b}{p}(z + \frac{a}{p}).$$

Then

$$f'(z) = \frac{b}{p} > 0$$

provided $b > 0$, and so f is a monotonically increasing function, and is maximised when z is at its largest. Hence we take the largest value of z possible for $v \in C_{\text{global}}$ as our upper bound, and the smallest for our lower bound; p and 0 respectively.

Finally we consider the l_1 norm. For $a, b, \frac{1}{p} > 0$, we have that

$$\frac{||M_{a,b}v||_1}{||v||_1} = \frac{|x + \frac{a}{p}y| + |\frac{b}{p}x + (1 + \frac{ab}{p^2})y|}{|x| + |y|} = \frac{(1 + \frac{b}{p})\frac{x}{y} + 1 + \frac{a}{p} + \frac{ab}{p^2}}{\frac{x}{y} + 1}.$$

Relabelling $\frac{x}{y} = z$, let

$$f(z) = \frac{(1 + \frac{b}{p})z + 1 + \frac{a}{p} + \frac{ab}{p^2}}{z + 1}.$$

Then

$$f'(z) = \frac{\frac{b}{p} - \frac{a}{p}(1 + \frac{b}{p})}{(z + 1)^2},$$

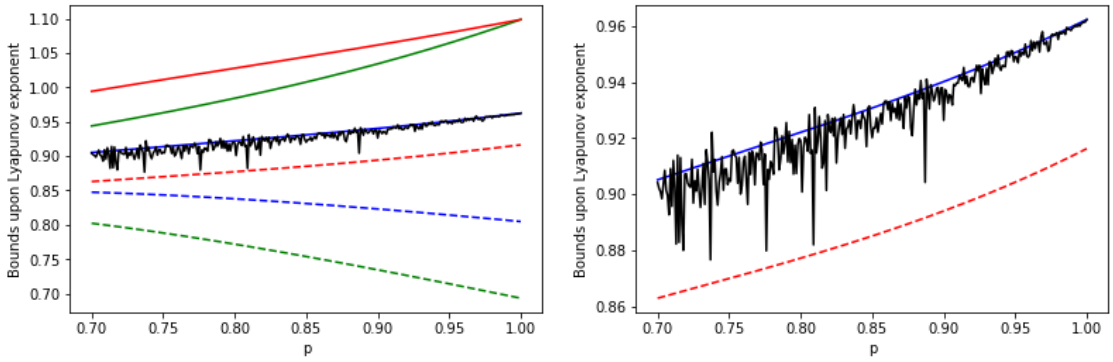
which is a monotonically decreasing function in z for $\alpha, \beta > 1$. Since $\alpha, \beta = \frac{1}{p}$ in our linked twist maps, this is true for the cases we are interested in. Hence $f(z)$ is maximised when z is at its smallest, and minimised when z is at its largest; 0 and p respectively. □

Note that the lemmas in this and the previous section, taken collectively, prove Theorem 5.1.

5.4 Calculation of the bounds

We now present the results given by Theorem 5.1, and compare these to the Lyapunov exponent calculated using a standard algorithm involving Gram Schmidt orthonormalization (see [9, 47]). Figure 40 shows Φ_{H_i}, Ψ_{H_i}, Φ_H, and Ψ_H, for $p \in [0.7, 1]$ and $i \in \{1, 2, \infty\}$.

The bounds Φ_H and Ψ_H require an infinite number of calculations, so a slight modification is needed in order to use them practically. Let us choose an integer $k \geq 4$. This is the maximum number of iterates of H for which we will calculate the truncation of the bounds on $\frac{||M_{a,b}v||}{||v||}$. We note that since each of the terms in the lower bound Ψ_H is positive, we need only calculate up to $i = k$ in the final summation to obtain a strict lower bound - any lower bound we obtain for larger k will be greater.



(a) Ψ_{H_i} and Φ_{H_i}.

(b) Ψ_H and Φ_H.

Figure 40: The bounds given by the l_1, l_2, and l_∞ norms for $p \in [0.7, 1]$ are shown in Figure 40a, where $\Psi_H^{(k)}$ is given by the dashed and $\Phi_H^{(k)}$ by the solid line, for $k \in \{1, 2, \infty\}$. Figure 40b shows Ψ_H and Φ_H, which in this case are given by Ψ_{H_1} and Φ_{H_2} respectively. The black line is a finite-time Lyapunov exponent, calculated using a standard algorithm involving Gram Schmidt orthonormalization ([9],[47]).

For the upper bound Φ_H, we cannot obtain a strict upper bound in a finite number of iterates in this way; however, the difference between the upper bound for the k^{th} and $k+1^{th}$ iterates decreases as k increases, and is quickly insignificant compared to the difference between the upper and lower bounds; for example, when $p = 0.9$, the difference between the upper bounds obtained when calculating up to the 4^{th} and 5^{th} iterates respectively is 1.49% of the gap between the upper and lower bounds obtained on the 5^{th} iterate, and this proportion decreases (at a sub-exponential rate) with each subsequent iterate.

We see that the best choices in these cases are the l_1 and l_2 norms for the lower and upper bounds respectively. Both bounds approximate λ to within a few decimal places; specifically, over a square lattice of initial conditions within S, the average error between the finite time Lyapunov exponent calculated using 10^4 iterates of a standard Gram Schmidt algorithm and Ψ_H is 3.95%, and with Φ_H is 0.97%. Note

that at $p = 1$, $\lambda = \Phi_{H_2}$; this is because the vector $v^2_{\max}$ given by Lemma 5.6 is equal to the eigenvector of H when $p = 1$, and this vector lies within C_{global}. The l_1 and l_∞ norms also coincide when $p = 1$; the formulas given in Lemma 5.7 are equal in this case.

These bounds, in their current state, are not especially useful in determining λ; the bounds are near instantaneous to calculate, and can provide an indication as to whether a given calculation of a FTLE has begun to converge, although only to within a few decimal places. Despite this, we can see from Figure 40 that the fluctuations observed in the calculated FTLEs do occasionally exceed the upper bound and approach the lower bound. The explicit calculation of the return time distribution does give us some insight into how the results of Young ([60], [61]) would tie into our system; we see that the proportion of points whose first return time is on iterate n (for $n \geq 4$) decays polynomially, as $\mathcal{O}(n^{-3})$, suggesting that the rate of mixing is at least polynomial.

The main limiting factor upon the accuracy of these bounds is the width of the cone C_{global}; the cone encompasses a 'wide' range of vectors, and in particular contains some of the $v_{\max}$ and $v_{\min}$ vectors given in Lemma 5.6. The bounds are calculated for either the vectors which undergo the largest or smallest expansion rates, or are taken at the boundaries of the cone. In reality however, over the course of an orbit, an initial vector v will orientate itself in various directions between the boundaries of the cone. If we can find a way to narrow the cones we consider for the terms $\psi(a, b)$ and $\phi(a, b)$ in Theorem 5.1, then we can obtain better bounds than Ψ_H and Φ_H. This is the topic of the next section.

5.5 Improving the Cone

In this section we discuss the circumstances under which we can consider narrower cones than C_{global}, in order to improve upon the bounds Φ_H and Ψ_H in Theorem

5.1. These modifications were simple in the case of bounding random products of matrices, in both [54] and Chapter 4, as the frequency with which the matrix M_{a_2,b_2} preceded the matrix M_{a_1,b_1} in an orbit was independent of the choice of a_1 and b_1. In the deterministic case, however, this independence does not hold.

We consider the sequences of matrices which can precede each $M_{a,b}$. We know that $M_{a,b}v \in C_{\text{global}}$ for $v \in C_{\text{global}}$ and any $a, b \geq 1$. In particular, $M_{a,b}v \in M_{a,b}C_{\text{global}}$. Thus, if we know that a particular sequence M_{a_1,b_1} is preceded by M_{a_2,b_2}, then we can restrict the calculations of $\phi(a_1, b_1)$ and $\psi(a_1, b_1)$ to the cone $M_{a_2,b_2}C_{\text{global}}$. This will happen with a certain frequency, which can be calculated by finding the Lebesgue measure of the intersection of $R_{G^{b_1}F^{a_1}}$ and $M_{a_2,b_2}R_{G^{b_2}F^{a_2}}$; in other words, we find the points which land in $R_{G^{b_1}F^{a_1}}$ following an application of M_{a_2,b_2}. We then divide this quantity by the Lebesgue measure of $R_{G^{b_1}F^{a_1}}$, and label the resulting frequency $P(a_2, b_2, a_1, b_1)$. Figure 41 shows this intersection in the case where $a_i = b_i = 1$ for $i = 1, 2$. This idea is summarised by the following theorem.

Theorem 5.2. *With the same conditions as Theorem 5.1, let*

$$P(a_2, b_2, a_1, b_1) = \frac{\mu_L(M_{a_2,b_2}R_{G^{b_2}F^{a_2}} \cap R_{G^{b_1}F^{a_1}})}{\mu_L(R_{G^{b_1}F^{a_1}})} \cdot \mu_S(R_{G^{b_1}F^{a_1}}),$$

$$\psi_k(a_2, b_2, a_1, b_1) = \min_{v \in M_{a_2,b_2}C_{\text{global}}} \left\{ \frac{||M_{a_1,b_1}v||_k}{||v||_k} \right\},$$

and

$$\phi_k(a_2, b_2, a_1, b_1) = \max_{v \in M_{a_2,b_2}C_{\text{global}}} \left\{ \frac{||M_{a_1,b_1}v||_k}{||v||_k} \right\}.$$

Then the Lyapunov exponents $\lambda_{1,2}$ satisfy

$$|\lambda_{1,2}| \geq \Psi^1_{H_k} = \frac{1}{n_S} \sum_{a_1,b_1=1}^{\infty} \sum_{a_2,b_2=1}^{\infty} P(a_2, b_2, a_1, b_1) \log \psi_k(a_2, b_2, a_1, b_1),$$

and

$$|\lambda_{1,2}| \leq \Phi^1_{H_k} = \frac{1}{n_S} \sum_{a_1,b_1=1}^{\infty} \sum_{a_2,b_2=1}^{\infty} P(a_2, b_2, a_1, b_1) \log \phi_k(a_2, b_2, a_1, b_1).$$

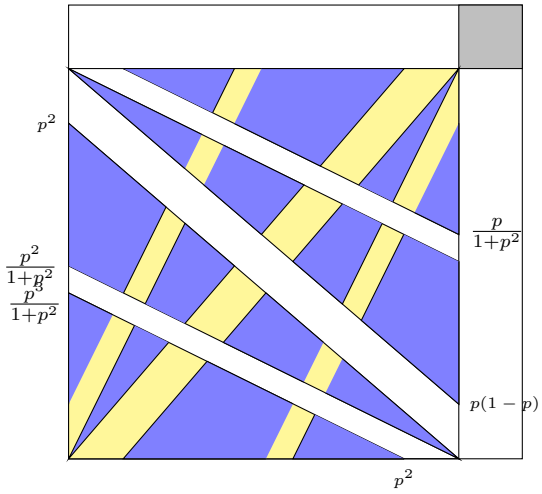


Figure 41: The regions corresponding to the frequency $P(1,1,1,1)$ are shaded in blue. The yellow regions are the remainder of the set R_{GF}, which has measure given by the sum of $P(1,1,i,j)$ for all i,j where $i,j \neq 1$.

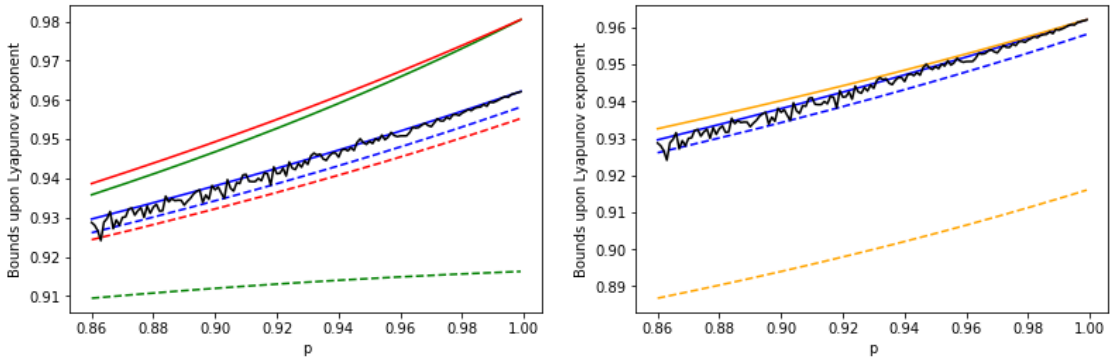
The distribution of preceding sequences of matrices is at its simplest when all $M_{a,b}$ with $a \geq 3$ or $b \geq 3$ always precede the matrix $M_{1,1}$ - that is, when any extended period spent in the shear regions is followed immediately by at least one hyperbolic iterate (i.e. an immediate return to S). This occurs when

$$p^5 + 3p^3 - 3p^2 + p - 1 > 0$$

i.e. $p \geq 0.8562$. This behaviour (i.e. remaining in S for a number of iterates following iterates spent in the shear region) is typical when the number of iterates spent within the shear region is large, even for p below this value; this was shown by Sturman and Springham [53].

A smaller improvement over the previous bounds can be obtained by simply finding the intersection shown in Figure 41, which is valid for any $p \in (0.682, 1)$. The distribution is simplified somewhat by noting that

$$P(a_1,b_1,a_2,b_2) = \frac{\mu_L(R_{G^{b_1}F^{a_1}})}{\mu_L(R_{G^{b_2}F^{a_2}})} P(a_2,b_2,a_1,b_1).$$



(a) $\Psi^1_{H_k}$ and $\Phi^1_{H_k}$.

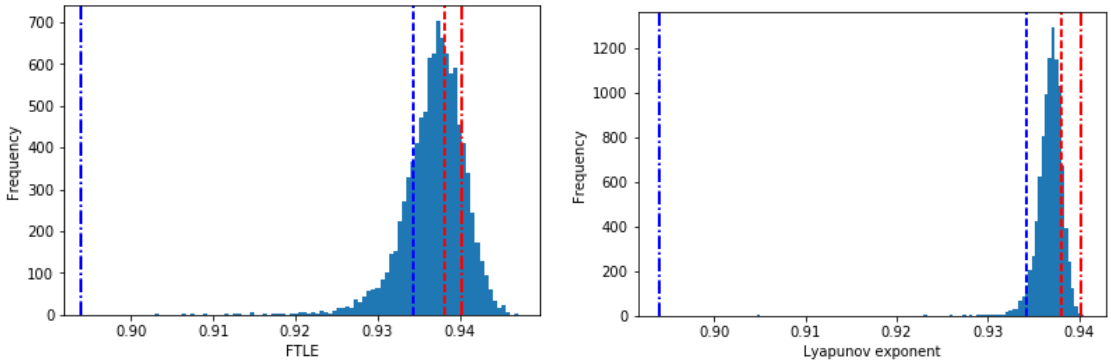
(b) Ψ^1_H and Φ^1_H.

Figure 42: (a) $\Psi^1_{H_k}$ and $\Phi^1_{H_k}$ for the l_1, l_2, and l_∞ norms, with the black line showing FTLEs calculated using 10^5 iterates of the Gram-Schmidt algorithm. (b) The best of these yield Ψ^1_H and Φ^1_H - in this case both are given by the l_2 norm. For comparison purposes we include Ψ_H and Φ_H in orange.

This is because the Lebesgue measure of the relevant intersections is the same in both cases due to the symmetry between the image and pre-image.

Figure 42 shows Ψ^1_H and Φ^1_H when $p \geq 0.86$ - the cases where the distribution of preceding sequences is at its simplest. Comparing this to Figure 40, we see there is a significant improvement, particularly on the accuracy of the lower bound. Note that the black line is the calculated FTLE for a particular initial condition (varying with p) within S. We also see that the l_2 norm yields the best upper and lower bounds, whereas before the best lower bound was given by the l_1 norm. We find that the average error between the bounds and the FTLE calculated for 10^5 iterates over a square lattice of initial conditions in S is 0.71% for Ψ^1_H and 0.82% for Φ^1_H. In the lower bound, this marks an improvement in the error by a factor of 5.56 on average.

Figure 43 shows the distributions of FTLEs for n iterates of the map when $p = 0.9$, and compares these to the bounds. We find that for $n = 10^4$, 84.2% of



(a) $n = 10000$. (b) $n = 100000$.

Figure 43: Histograms showing the distribution of FTLEs, $\lambda_n(x)$, for a square lattice of 10^4 points, $p = 0.9$, and n iterates of the map taken at each point. The dot-dash lines are Ψ_H and Φ_H, and the dotted lines are Ψ_H^1 and Φ_H^1.

the points upon the lattice yield exponents between Ψ_H and Φ_H, while 40.3% yield exponents between Ψ_H^1 and Φ_H^1. For $n = 10^5$, 99.9% of the points yield exponents between Ψ_H and Φ_H, while 81.3% yield exponents between Ψ_H^1 and Φ_H^1.

Note that Theorem 5.2 is the first step in a long sequence of improvements one can make to the bounds; specifically, it describes looking at the first preceding sequence of the form M_{a_2,b_2}. We could consider the sequences M_{a_3,b_3} which precede these M_{a_2,b_2}, and calculate the corresponding frequencies with which these occur; in principle, we could carry this process on indefinitely by finding the frequencies $P(a_i, b_i, \ldots a_1, b_1)$ for increasing i. The calculation of these frequencies is an intricate task in practice, however.

For the following theorem, we let $R_{P(a_i,b_i,\ldots,a_1,b_1)}$ be the set of points corresponding to the frequency $P(a_i, b_i, \ldots, a_1, b_1)$.

Theorem 5.3. *With the same conditions and definitions as Theorem 5.2, we define*

inductively

$$P(a_i, b_i, \ldots, a_1, b_1) = \frac{\mu_L(M_{a_i,b_i} R_{G^{b_i} F^{a_i}} \cap R_{P(a_{i-1},b_{i-1},\ldots,a_1,b_1)})}{\mu_L(R_{P(a_{i-1},b_{i-1},\ldots,a_1,b_1)})} \cdot \mu_S(R_{P(a_{i-1},b_{i-1},\ldots,a_1,b_1)}).$$

Also, let

$$\psi_k(a_i, b_i, \ldots, a_1, b_1) = \min_{v \in M_{a_i,b_i} \cdot \ldots \cdot M_{a_2,b_2} C_{\text{global}}} \left\{ \frac{||M_{a_1,b_1} v||_k}{||v||_k} \right\},$$

and

$$\phi_k(a_i, b_i, \ldots, a_1, b_1) = \max_{v \in M_{a_i,b_i} \cdot \ldots \cdot M_{a_2,b_2} C_{\text{global}}} \left\{ \frac{||M_{a_1,b_1} v||_k}{||v||_k} \right\}.$$

Then the Lyapunov exponents $\lambda_{1,2}$ *satisfy*

$$|\lambda_{1,2}| \geq \Psi^i_{H_k} = \frac{1}{n_S} \sum_{a_i,b_i=1}^{\infty} \ldots \sum_{a_1,b_1=1}^{\infty} P(a_i, b_i, \ldots, a_1, b_1) \log \psi_k(a_i, b_i, \ldots, a_1, b_1),$$

and

$$|\lambda_{1,2}| \leq \Phi^i_{H_k} = \frac{1}{n_S} \sum_{a_i,b_i=1}^{\infty} \ldots \sum_{a_1,b_1=1}^{\infty} P(a_i, b_i, \ldots, a_1, b_1) \log \phi_k(a_i, b_i, \ldots, a_1, b_1).$$

Considering an infinite number of preceding sequences of matrices would allow our bounds to converge upon the true Lyapunov exponent. In particular, when $p = 1$ (i.e. the Cat map), we can reduce the bounds $\Psi^{\infty}_{H_k}$ and $\Phi^{\infty}_{H_k}$ to the explicitly calculable Lyapunov exponent. In this case the only distribution elements which contribute to the sum are those where $a_i = b_i = 1$ for all i, and each occurs with frequency one. Let λ^u and v^u be the unstable eigenvalue and eigenvector of the Cat map respectively, then we obtain

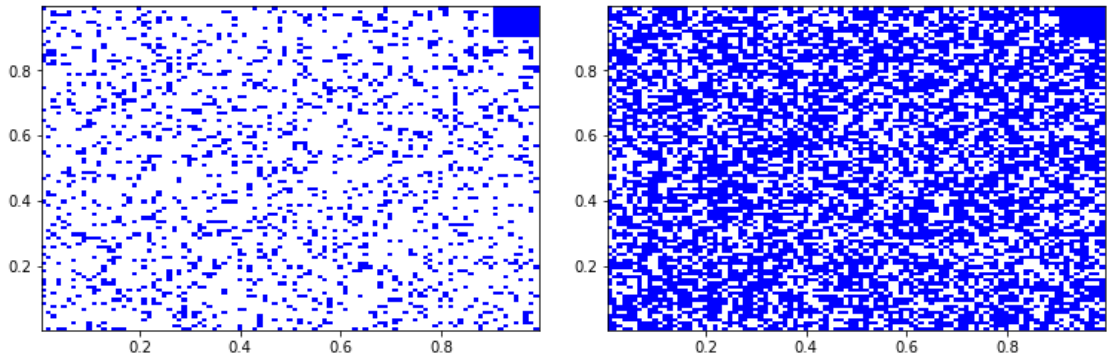
$$\begin{aligned} \Psi^{\infty}_{H_k} &= \frac{1}{n_S} P(1, 1, \ldots, 1, 1) \log \psi_k(1, 1, \ldots, 1, 1), \\ &= \log \max_{v \in GF \ldots GFC_{\text{global}}} \left\{ \frac{||GFv||_k}{||v||_k} \right\}, \\ &= \log \frac{||GFv^u||_k}{||v^u||_k} = \log \lambda^u, \end{aligned} \tag{118}$$

and similarly for $\Phi^{\infty}_{H_k}$. Alternatively, we obtain the same result by replacing C_{global} with v^u, which is the minimal invariant cone for the Cat map.

The distributions of FTLEs in LTMs do not appear Gaussian; they exhibit long (stretched-exponential) tails on the left hand side, corresponding to points which spend large portions of their orbits in the shear regions, similar to the distributions discussed in [43]. For a given set of points, these tails will begin to shrink as $n \to \infty$, as more of the points trapped in the shear regions escape into the hyperbolic regions, and the distribution appears to converge upon a Gaussian distribution; this is due to LTMs obeying a central limit theorem as described in [61]. Figure 43 shows that, even after 10^5 iterates, the probability of a randomly chosen initial condition having a FTLE outside of Ψ_H^1 and Φ_H^1 is still very significant.

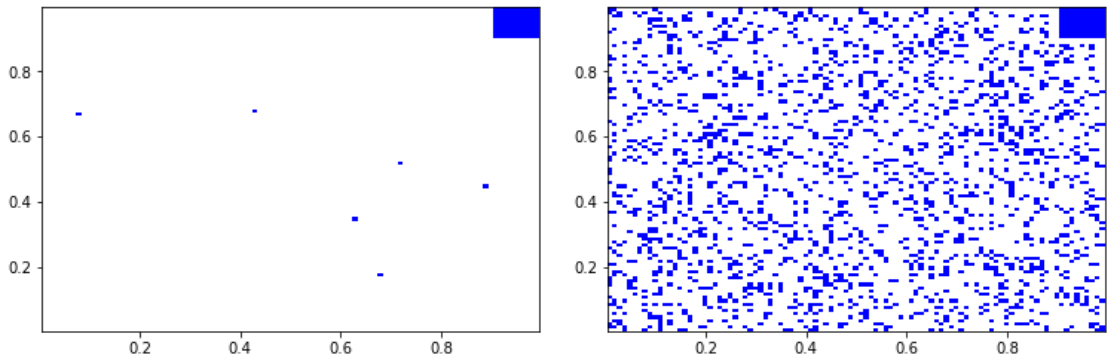
The issue of poorly converging FLTEs could be sidestepped if the initial conditions which yielded them were concentrated in particular regions within the domain. If this were the case, one could simply pick an initial condition outside of these regions and be confident that the calculated FTLE would converge to a reasonable degree. However, linked twist maps are ergodic, and so the points for which we obtain slower convergence are representative of a portion of every orbit; at some (and indeed, infinitely many) iterates upon every orbit, we will land within a shear region for any finite number of iterates, corresponding to initial conditions with FTLEs below the lower bounds, or, alternatively, return to the hyperbolic regions unusually often within a subsequence of iterates, corresponding to the initial conditions above the upper bounds.

Figure 44 shows in blue the location of initial conditions whose FLTEs lie outside of the bounds after n iterates of H, when $p = 0.9$. Note that the dataset and lattice used to create these figures is the same as those used in Figure 43. The figures show that there is no clear pattern to the location of initial conditions with slowly converging FTLEs; we find such initial conditions are spread fairly evenly throughout the domain, favouring neither the shear nor hyperbolic regions. This indicates that the choice of initial condition to obtain a 'sufficiently' converged FTLE



(a) Ψ_H and Φ_H, $n = 10000$ iterates.

(b) Ψ_H^1 and Φ_H^1, $n = 10000$ iterates.



(c) Ψ_H and Φ_H, $n = 100000$ iterates.

(d) Ψ_H^1 and Φ_H^1, $n = 100000$ iterates.

Figure 44: The points in blue are the initial conditions in a lattice of 10^4 points (as used in Figure 43) whose FTLEs are outside of the respective bounds after n iterates. The points in white are those with FTLEs within the bounds after n iterates. There is no obviously discernible pattern to the location of initial conditions with slowly converging FTLEs.

is not obvious. Also, following choosing an initial condition, the number of iterates needed to obtain sufficient convergence is not clear without further experimentation. Since the Lyapunov exponent is itself not explicitly calculable, these bounds provide a benchmark for the notion of sufficient convergence.

5.6 Alternative maps

The bounds we have discussed in this chapter were conceived with the trickiness of calculating Lyapunov exponents in linked twist maps in mind; however, the concepts used in their derivation are not limited to linked twist maps exclusively. In this section we discuss and construct similar bounds to Φ_H and Ψ_H for another map formed from shear compositions. Note that this is done primarily for demonstrative purposes, since we will see that the FTLEs we obtain for such a map are typically more reliable than was the case for the LTM; unlike in the LTM, orbits in this map do not undergo arbitrarily long sequences of sub-exponential (linear) growth.

For $p \in (0,1)$, let $H_\Omega : \mathbb{T}^2 \to \mathbb{T}^2$ be given by

$$H_\Omega \begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} H_{\Omega_2} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{p} \\ 1 & \frac{p+1}{p} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} & \text{if } y \leq p, \\ H_{\Omega_1} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{1-p} \\ 1 & \frac{2-p}{1-p} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \frac{-p}{1-p} \\ \frac{-p}{1-p} \end{pmatrix} & \text{if } y \geq p. \end{cases} \tag{119}$$

This map applies a horizontal shear which wraps once around either the annulus $P = \{(x,y) : y \leq p\}$ or $Q = \{(x,y) : y \geq p\}$, followed by a vertical shear which wraps once around $\mathbb{T}^2$ (see Figure 45). In fact, H_Ω is an example of the family of systems discussed in Section 3.1, where $n = 2$, $m = 1$, and $\alpha_{1,2} = \beta_1 = 1$; note that H_{Ω_1} is shifted in order to maintain continuity of the map. At $p = 1/2$ and $p = 1$, H_Ω reduces to an Anosov diffeomorphism, with an explicitly calculable Lyapunov

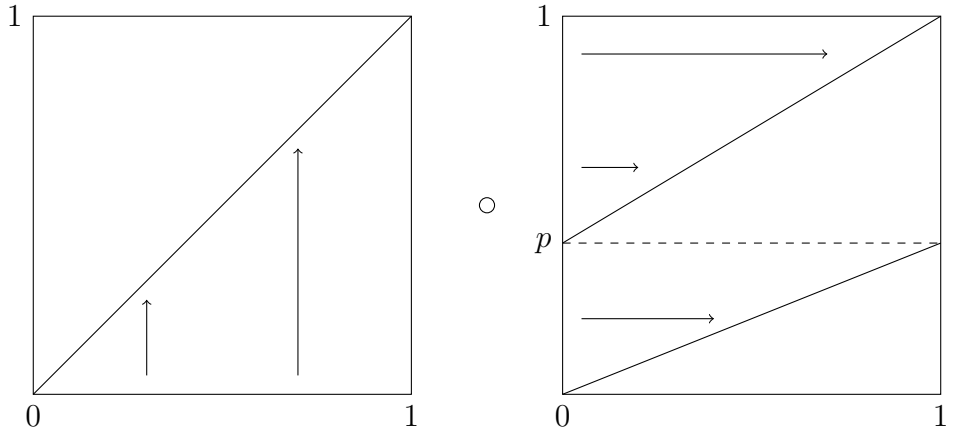


Figure 45: The shears which are composed to obtain the map H_Ω in (119).

exponent. The cases where $p \in (0, \frac{1}{2}]$ are qualitatively identical to those where $p \in [\frac{1}{2}, 1)$, and so we only focus on the former.

We again set out to obtain a return time distribution to a particular region, in this case the annulus Q. For $A \subseteq Q$, we concern ourselves with the conditional measure

$$\mu_Q(A) = \frac{\mu_L(A)}{1-p}. \tag{120}$$

A point which begins in Q will undergo application by H_{Ω_1} for one iterate, returning immediately or undergoing application by only H_{Ω_2} for all iterates until its return to Q. Thus all sequences $M_{a,b}$ that we concern ourselves with will be of the form

$$\begin{aligned} M_{1,b} &= H_{\Omega_2}^b \circ H_{\Omega_1}, \\ &= DH_{\Omega_2}^b DH_{\Omega_1} + DH_{\Omega_2}^b \mathbf{c}, \end{aligned} \tag{121}$$

where

$$\mathbf{c} = \begin{pmatrix} \frac{-p}{1-p} \\ \frac{-p}{1-p} \end{pmatrix},$$

and, for $i = 1, 2$, DH_{Ω_i} is the respective Jacobian matrix of H at each iterate. Note that these matrices share an invariant cone by Theorem 2.4, with boundaries given by their unstable eigenvectors. We label this cone C_{H_Ω}.

In a similar fashion to Section 5.2, we will refer to the set of points which first return to Q after application by the sequence $M_{1,b}$ as R_b, where $b \in \mathbb{N}_0$, and their corresponding frequencies in the return time distribution as $R(b)$. Unlike in the case of the linked twist map, we can allow $b = 0$ since both DH_{Ω_1} and DH_{Ω_2} are hyperbolic. We will also refer to the set of points yet to return on the i^{th} iterate of H as U_i. Note that the calculations we do throughout this section are for the images of these points rather than the sets themselves; since this map is measure-preserving, this does not affect the distribution.

We obtain the following for the measure of the set of points whose first return to Q is after one and two iterates of H (see the yellow and red regions in Figure 46 respectively):

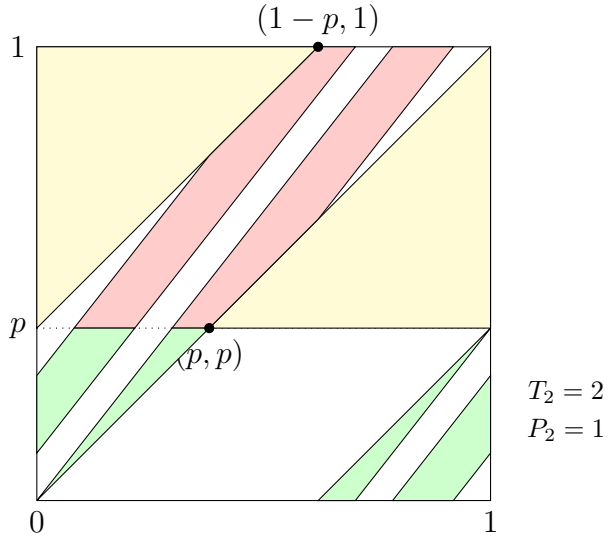
$$R(0) = \mu_Q(R_0) = 1 - p,$$

and

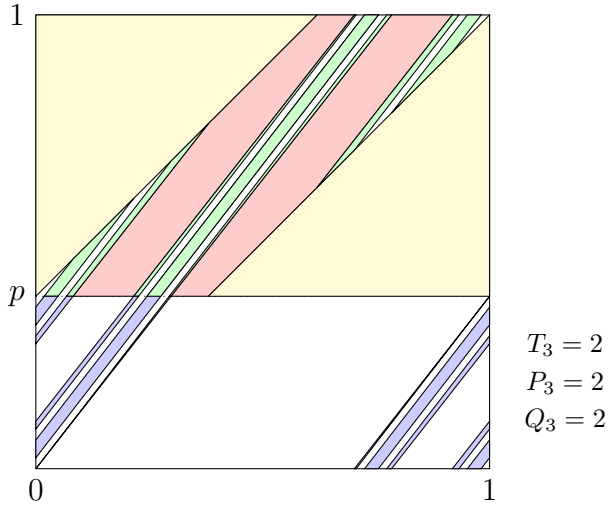
$$R(1) = \mu_Q(R_1) = \frac{p(1-2p^2)}{(2p+1)(1-p)}.$$

We note that U_2 consists of two triangles and a parallelogram. On the following iterate, the triangles are stretched between their fixed points at $(0, 0)$ and $(1, p)$ respectively into Q, wrapping around the entire torus once in the process. The result of this is that the triangles in U_2 admit a quadrilateral each in U_3, with the same slopes as the triangles in U_3. This process will repeat at each iterate, with these triangles admitting two more quadrilaterals; note that the quadrilaterals admitted by the triangles are not parallelograms. Similarly, the parallelogram in U_2 is wrapped once around the torus stretching into Q, and admits another parallelogram in U_3; this also repeats for later iterates. Finally, any existing quadrilaterals from earlier triangles are also stretched in a similar way, and so admit a new quadrilateral each on subsequent iterates.

Let T_i, P_i and Q_i be the number of triangles, parallelograms and other quadri-



(a) The sets $R_{H_{\Omega_1}}$ (yellow), $R_{H_{\Omega_2}H_{\Omega_1}}$ (red) and $U_{H_{\Omega_2}}$ (green).



(b) The sets $R_{H_{\Omega_1}}$ (yellow), $R_{H_{\Omega_2}H_{\Omega_1}}$ (red), $R_{H^2_{\Omega_2}H_{\Omega_1}}$ (green) and $U_{H_{\Omega_3}}$ (blue).

Figure 46: The points returning and yet to return after the second and third iterates of H. Note that the sets R_b and U_i are pre-images, and the above pictures show the images of the points returning on the respective iterate; for example, $R_{H_{\Omega_1}}$ shows the image of the points which remain in A for at least one iterate, that is, $H(R_0)$. For the purposes of calculating the return time distribution this distinction does not matter, as the measures of the corresponding sets are identical (e.g. $\mu(R_0) = \mu(R_{H_{\Omega_1}})$).

laterals which comprise U_i, then, for $i \geq 3$,

$$T_i = 2, P_i = 2^{i-2}, Q_i = 2 + 2Q_{i-1}, \tag{122}$$

where $Q_2 = 0$. Noting that the total number of elements of U_i doubles at each iterate, starting with 3 elements when $i = 2$, we have

$$Q_i = 3 \cdot 2^{i-2} - P_i - T_i = 2(2^{i-2} - 1).$$

The triangles have the same measure as each other for each i, as do all the parallelograms, however the other quadrilaterals do not. We need to find the width of the bases of each of these shapes in order to obtain their areas, and subsequently the return time distribution.

Let δ_j be the slope of the vector obtained by applying DH_{Ω_2} j times to the vector $(1, 1)$. Then, the width of the bases of the triangles, w_{T_i}, is given by

$$w_{T_i} = p\left(\frac{1}{\delta_{i-3}} - \frac{1}{\delta_{i-2}}\right),$$

For the parallelograms, the width $w_{P_{i+1}}$ for $i \geq 4$ is given inductively by

$$w_{P_{i+1}} = \frac{p \cdot w_{P_i}}{1 + p + p\delta_{i-3}},$$

where $w_{P_3} = \frac{p(1-p)}{2p+1}$. The widths of the horizontal sides of the quadrilaterals admitted by the triangles, $w^{1,2}_{Q_i}(\epsilon)$, are given by

$$w^1_{Q_i}(\epsilon) = \epsilon\left(\frac{1}{\delta_{i-3}} - \frac{1}{\delta_{i-2}}\right),$$

$$w^2_{Q_i}(\epsilon) = (\epsilon + p)\left(\frac{1}{\delta_{i-3}} - \frac{1}{\delta_{i-2}}\right),$$

where $\epsilon \in \mathbb{N}$ represents the region $\{(x, y) \in \mathbb{R}^2 : \epsilon \leq y \leq \epsilon + p\}$ which that particular quadrilateral is located in, $w^1_{Q_i}(\epsilon)$ is the width at $y = \epsilon$, and $w^2_{Q_i}(\epsilon)$ is the width at $y = \epsilon + p$. In this case we only concern ourselves with the quadrilaterals admitted by the triangle with a corner at $(0, 0)$, since those admitted by the other triangle

possess the same areas. Note that since each of the quadrilaterals for a given i is located in a different region $\{(x, y) \in \mathbb{R}^2 : \epsilon \leq y \leq \epsilon + p\}$ to the others, their areas will vary depending on ϵ. See Figure 47 for a sketch of this process.

The Lebesgue measures of these shapes are as follows:

$$\mu_L(T_i) = \frac{p \cdot w_{T_i}}{2},$$

$$\mu_L(P_i) = p \cdot w_{P_i},$$

$$\mu_L(Q_i(\epsilon)) = \frac{p \cdot (w^1_{Q_i}(\epsilon) + w^2_{Q_i}(\epsilon))}{2}.$$

Thus we have

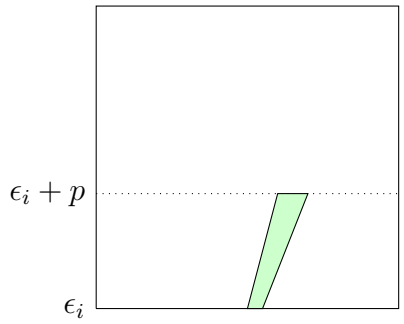
$$\mu_Q(U_i) = \frac{1}{(1-p)}(2 \cdot \mu_L(T_i) + 2^{i-2} \cdot \mu_L(P_i) + \sum_{\epsilon} 2n(\epsilon) \cdot \mu_L(Q_i(\epsilon))),$$

where $n(\epsilon)$ is the number of quadrilaterals in the region $\{(x, y) \in \mathbb{R}^2 : \epsilon \leq y \leq \epsilon + p\}$ - this is either 0 or 1. We then have, for $i \geq 3$,

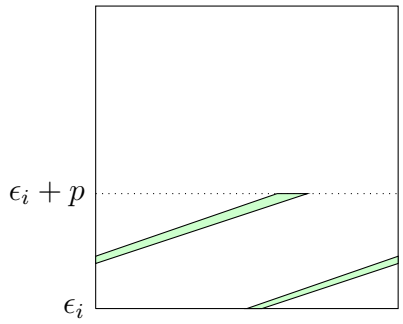
$$R(i-1) = \mu_Q(R_{i-1}) = \mu_Q(U_{i-1}) - \mu_Q(U_i).$$

To obtain $n(\epsilon)$ for a given U_i, we perform the following algorithm:

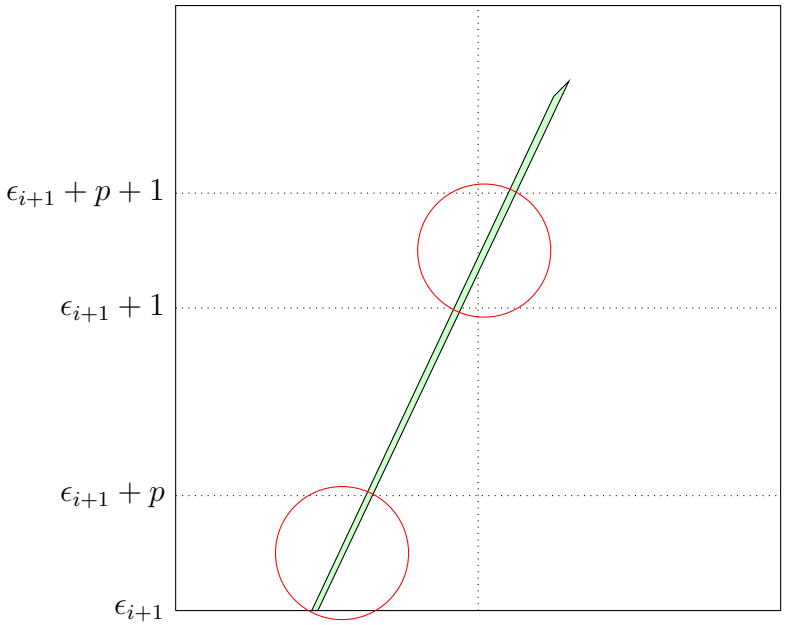
1. If $i \leq 3$, there are no quadrilaterals in U_i.
2. Start with the (list of) coordinates $(x, y) = (1, 1)$. Let $j = 4$.
3. If $j < i$, replace each entry (x, y) in the list with $(x + y, x + 2y)$ and $(x + y + 1, x + 2y + 1)$.
4. Add $(1, 1)$ to the list, and let $j = j + 1$.
5. Repeat steps 3 and 4 until $j = i$.
6. The list of ϵ for which $n(\epsilon) = 1$ is given by the y coordinate of each entry in the list.



(a) A quadrilateral in U_i.



(b) Applying a horizontal shear.



(c) Applying a vertical shear yields two quadrilaterals in U_{i+1}.

Figure 47: A sketch of the process undertaken by each quadrilateral in U_i to yield two quadrilaterals in U_{i+1} (highlighted by the red circles). Note that, for the purpose of clarity, none of these figures are to scale, as the images become narrow very quickly; this can be seen in Figure 46. Also note that ϵ_i and ϵ_{i+1} indicate the y-coordinate in the case of U_i and U_{i+1} respectively.

For $k \in \{1, 2, \infty\}$, let

$$\psi^k_{H_\Omega}(b) = \min_{v \in C_{H_\Omega}} \{\frac{||M_{H^b_{\Omega_2}, H_{\Omega_1}} v||_k}{||v||_k}\},$$

and

$$\phi^k_{H_\Omega}(b) = \max_{v \in C_{H_\Omega}} \{\frac{||M_{H^b_{\Omega_2}, H_{\Omega_1}} v||_k}{||v||_k}\},$$

then our bounds upon the Lyapunov exponent λ, Ψ_{H_Ω} and Φ_{H_Ω}, are given by

$$\lambda \geq \Psi^k_{H_\Omega} = (1-p) \sum_{b=0}^{\infty} R(b) \log \psi^k_{H_\Omega}(b), \tag{123}$$

and

$$\lambda \leq \Phi^k_{H_\Omega} = (1-p) \sum_{b=0}^{\infty} R(b) \log \phi^k_{H_\Omega}(b), \tag{124}$$

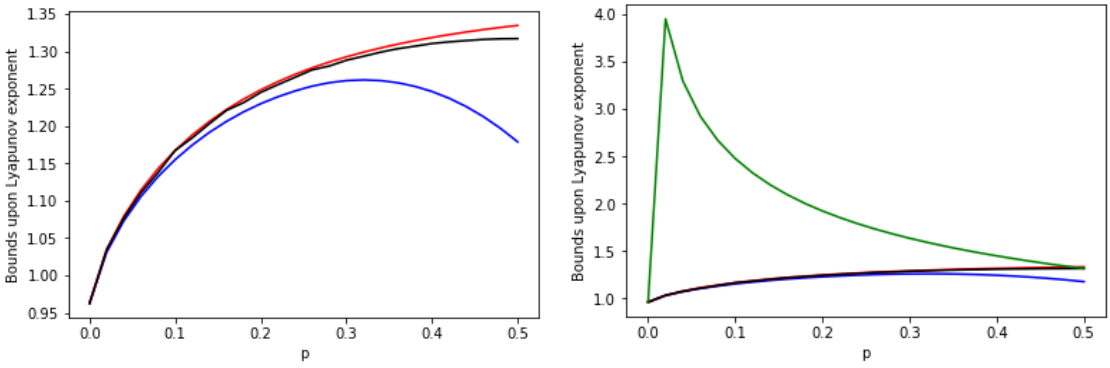
where the average return time to Q is $\frac{1}{1-p}$ by Kac's lemma [26]. We again take the best lower and upper bounds from the three norms to obtain Ψ_{H_Ω} and Φ_{H_Ω}.

The above is all exact and can be written down explicitly, similar to the bounds we obtained for the LTM, however an issue in practicality arises when trying to calculate the total area of the quadrilaterals, $\sum_\epsilon \mu_L(Q_i(\epsilon))$. Specifically, the number of area calculations required to obtained an exact value increases exponentially with i after $i = 4$. In order to prevent this, we can instead consider bounds upon these areas for $i \geq 5$; the lower bound is given by taking $\epsilon = 1$ for all quadrilaterals in U_i, whereas the upper bound is given by taking the largest value of ϵ - in this case, by applying the map

$$(x, y) \to (x + y + 1, x + 2y + 1)$$

$i - 4$ times to the coordinates $(1, 1)$, and taking the final y coordinate as our value for ϵ.

Figure 48a shows the bounds Ψ_{H_Ω} and Φ_{H_Ω} for $p \in [0, 0.5]$ - note that the bounds in this figure were calculated using the upper and lower bounds upon the areas of the quadrilaterals, as mentioned in the previous paragraph. We see that for small



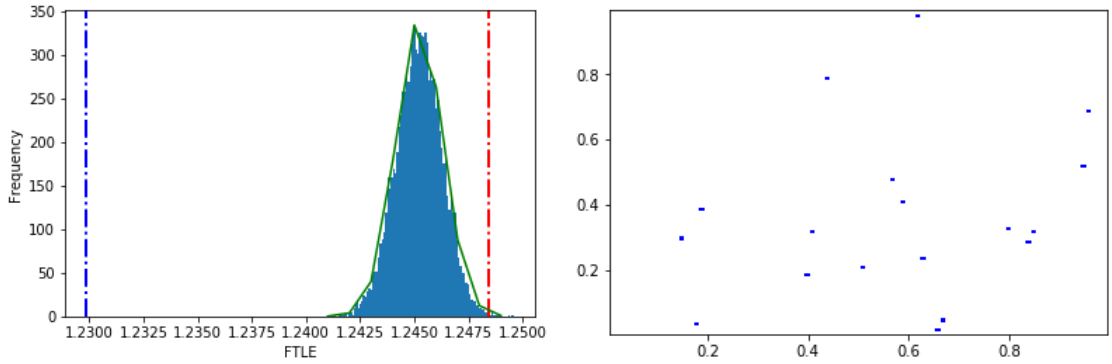
(a) Ψ_{H_Ω} and Φ_{H_Ω}. (b) Comparing with a 'naive' upper bound.

Figure 48: Ψ_{H_Ω} and Φ_{H_Ω} for $p \in [0, 0.5]$. The FTLE calculated over 100000 iterates for the initial condition $(\frac{1}{\pi}, 2 - \sqrt{2})$ is shown in black. In (b), the bounds are compared to the naive upper bound given by the larger Lyapunov exponent of the matrices DH_{Ω_1} and DH_{Ω_2}.

values of p the bounds are much closer, and the gap between them increases with p. The FTLE remains comparatively close to Φ_{H_Ω} for all p, whereas the accuracy of Ψ_{H_Ω} appears to diminish as p increases.

In Figure 48b, the bounds are compared to a naive upper bound given by the larger Lyapunov exponent of the matrices DH_{Ω_1} and DH_{Ω_2} - in this case, DH_{Ω_2}. We see that the improvement over this bound is quite large until p gets close to 0.5; note that at $p = 0.5$, the system is an Anosov diffeomorphism. The sudden rise in the naive bound from $p = 0$ to $p = 0.02$ is due to the horizontal shear in the map changing from a slope of 1 (H_Ω is the Cat map when $p = 0$) to two horizontal shears of slopes 0.02 and 0.98 respectively - the matrix corresponding to the shear with slope 0.02 provides the upper bound in this case, leading to the sudden jump.

Figure 49a shows a histogram of FTLEs of H_Ω after 10^5 iterates for a square lattice of 10^4 initial conditions when $p = 0.2$. The distribution of the FTLEs in this case is much narrower than in the case of the linked twist map, with the vast



(a) A histogram of the FTLEs compared to Ψ_{H_Ω} and Φ_{H_Ω}.

(b) The initial conditions which yield FTLEs outside of Ψ_{H_Ω} and Φ_{H_Ω}.

Figure 49: The FTLEs in the above were obtained using a lattice of 10^4 initial conditions upon $\mathbb{T}^2$, each iterated under H_Ω 10^5 times. In this case, all of the FTLEs outside of the bounds are larger than Φ_{H_Ω}.

majority of FTLEs concentrated close to Φ_{H_Ω}. In particular, the distribution itself appears to be at least approximately Gaussian - specifically, the green fitted curve is given by

$$N(x) = \frac{c \cdot e^{\frac{-(x-\lambda)^2}{2\sigma^2}}}{2\pi\sigma^2},$$

where λ is the Lyapunov exponent, σ^2 the variance of the distributions of FTLEs after 10^5 iterates, and c a scaling factor depending on the number of initial conditions and size of bins chosen; in the case of Figure 49a, $\sigma^2 = 1.1541 \times 10^{-6}$ and $c = 0.91966$. In this case, these quantities have been calculated directly from the dataset. We see that there is no stretched tail to this distribution, as there was in the case of the LTM; the FTLE obtained from a randomly chosen initial condition for H_Ω is typically a more reliable estimate than for the LTM.

Figure 49b shows the initial conditions within the lattice which yield FTLEs outside of the bounds; we see that very few do, although the number is comparable to that of the FTLEs we found outside of Ψ_H and Φ_H in the case of the linked twist

map in Figure 44c. The difference in this case, however, is that all of the initial conditions outside of the bounds yield FTLEs above the upper bound Φ_{H_Ω}, and are closer to the mean FTLE than was the case for the LTM.

The reason for Φ_{H_Ω} being the closer of the two bounds to the FTLEs is due to the choice of reference region Q. The largest proportion of points return to Q after one iterate of H_Ω, and so only the matrix DH_1 is considered when bounding their norms; this matrix possesses the smaller Lyapunov exponent of the two Jacobians in the system, and on average stretches tangent vectors less than DH_2 does. If the reference region were P instead, or equivalently, if we considered the case where $p \in [0.5, 1]$, then DH_1 would possess the larger Lyapunov exponent, and so both Ψ_{H_Ω} and Φ_{H_Ω} would be increased.

The Lyapunov exponents for the cases where $p \in [0, 0.5]$ are identical to those where $p \in [0.5, 1]$, where we equate the case p with $1-p$; the systems are dynamically identical and simply differ by a shift in coordinates. Thus, a way to improve upon these bounds (in particular, Ψ_{H_Ω}) would be to undergo the same process for the cases where $p \in [0.5, 1]$, then take the upper bound from the case p and the lower bound from the case $1-p$. Following this, one could, in principle, perform the same process we saw in the case of the LTM, considering preceding sequences of matrices in order to the narrow the cones used. However, the calculation of a preceding distribution would not be a simple task.

5.7 Summary

We have studied a method for obtaining bounds upon the Lyapunov exponents of a parametrized family of linked twist maps. The method involved the use of the existence of invariant cones for orbits returning to a reference region - in this case, the region where the two annuli overlap - as well as the ability to fully express a return time distribution for the points in this region for a range of parameter values.

We improved upon the bounds we obtained by considering the sequences of matrices which could precede these orbits, and explicitly calculating the preceding distribution for a smaller range of parameter values. Although the best bounds using this method are obtained by finding the entire distribution, a smaller improvement is possible by simply finding any preceding distribution element, and applying it in the proper way within the expression for the bounds. Furthermore, we stated a theorem which repeats this process, to obtain a sequence of bounds which should narrow upon the Lyapunov exponent.

We showed further that these bounds are applicable to systems other than LTMs by explicitly calculating them for the case of another system formed via shear composition. In this case the return time distribution was explicitly calculable, however we instead elected to bound the distribution for practicality reasons.

6 Summary and Outlook

In this final section, we summarise our findings throughout this thesis and discuss some possible extensions to the theory.

6.1 Summary

In this thesis we have studied the Lyapunov exponents of various systems in which their explicit calculation is not possible. These systems are characterised by the existence of some form of hyperbolicity within their dynamics. The strictest form of hyperbolicity, uniform hyperbolicity, requires uniform bounds on the expansion and contraction rates upon invariant regions of tangent space; the most typical systems which possess this quality, such as Arnold's Cat Map, have explicitly calculable Lyapunov exponents, however adding some form of non-uniformity to the system - for example by adding randomness as in Chapter 4, or by introducing spatial dependence of the Jacobian as in Chapter 5 - can remove explicit calculation as a possibility.

Where explicit calculation is not possible, one must resort to estimation, often via the numerical calculation of finite-time Lyapunov exponents, or to obtaining rigorous bounds on the Lyapunov exponents. This thesis has provided a method which yields the latter of these two options, by utilising the existence of hyperbolicity in the systems to find invariant cones in tangent space; these cones are used to reduce the range of vectors one must consider when calculating the maximum and minimum growth rates in tangent space.

We have derived and studied these bounds (primarily) for two families of maps. The first was in the case of random products of hyperbolic toral automorphisms formed via the composition of shear matrices, a method which built upon that of Sturman and Thiffeault [54] in the case of random products of shear matrices. The

second case concerned deterministic systems, with particular focus on the linked twist map, where one could explicitly obtain (or at least bound) a return time distribution to a reference region. Despite the linked twist map possessing regions upon which the Jacobian is not hyperbolic, these bounds could still be obtained due to the return map to the reference region being hyperbolic.

In the case of the family of random dynamical systems studied in Chapter 4, we obtained explicit, elementary bounds on the Lyapunov exponents, labelled Ψ_1 and Φ_1. We noted that the bounds provided reasonable estimates (up to 6 s.f.) for the Lyapunov exponents in some cases where the cone was particularly narrow due to the matrices chosen having eigenvectors which were more closely aligned, however in other cases (i.e. wider cones) the bounds were not as tight.

With the results of calculating Ψ_1 and Φ_1 in mind, we endeavoured to improve upon these bounds via narrowing the cones; this was made possible by considering the matrices which could precede other matrices within a realisation of the infinite orbit. This led to an improved version of the bounds, Ψ_2 and Φ_2, which improved accuracy over the previous bounds by a factor of between 100-1000 (2-3 s.f.) on a case by case basis. Furthermore, the process of narrowing the cones could in principle be repeated indefinitely, leading to a sequence of bounds Ψ_k and Φ_k, with the bounds increasing in tightness with each k. While this process would be relatively simple to code, the calculation time would increase exponentially with k, since every possible combination of matrix and preceding matrix must be considered.

Following the results of Chapter 4, we applied a similar method to obtain bounds on the Lyapunov exponent of a family of linked twist maps. In order to guarantee that the orbits we considered were hyperbolic, we calculated a return-time distribution to the overlap region of the two annuli on which the shears were defined. The theory involving the invariant cones in this case is similar to that of Sturman and Thiffeault [54], and so bounds upon growth rates of vectors in tangent space were

explicitly obtainable.

The return-time distribution for the linked twist map was explicitly obtainable, although for the purpose of clarity we focused on a range of parameter values in which the distribution was at its simplest. The distribution was obtained using elementary geometry and as such is somewhat general in its application in principle; to demonstrate this, we also calculated a return-time distribution (and subsequently bounds) for another map formed via shear composition which possessed similar properties in Section 5.7.

The first iteration of the bounds for the linked twist map provided an indication of the value of the Lyapunov exponent to within a few decimal places. The calculation time for this version was very quick, since the number of matrix products we considered was reduced significantly from the random case, where every product was possible. In order to improve the tightness of the bounds, we again considered narrowing the cones in the same way as in Section 4.6, by considering the matrices which can precede each matrix in an orbit. The calculation of this 'preceding distribution', while again involving only simple geometry, is intricate in practice but seems widely applicable in such systems due to its simplicity.

Following the above calculations, we obtained an improved version of the bounds upon the Lyapunov exponent of the linked twist map, particularly in the tightness of the lower bound. For the purposes of accurate estimation, these bounds did not initially appear as useful as those obtained for the random family of maps, however investigation into the distribution of finite-time Lyapunov exponents of the linked twist map appeared to indicate that, even after 10^5 iterates of a typical estimation process involving Gram-Schmidt orthonormalization, a significant fraction (18.7%) of the estimates lay outside of the improved bounds. We also noted that, in a similar fashion to the random case, we could in principle extend this process of considering the preceding matrices indefinitely; however, this process would be much more time

consuming that in the random case, due to the requirement of calculating more and more intricate preceding distributions.

6.2 Random bounds: three or more hyperbolic matrices

The bounds Ψ_k and Φ_k from Theorem 4.5 are stated specifically for the case where we choose the matrices A and B with probability $\frac{1}{2}$ at each iterate. The choice to present this theorem (as well as others in the chapter) for this simple case was made to attempt to emphasise the more interesting side of the bounds - the use of the mutually invariant cones for these matrices. The generalization of these theorems to cases where the chosen probability distributions are non-uniform is immediate and obvious, as the theory regarding the cones does not change at all; in fact, the only things that change are the frequency with which particular matrices $M_{a,b}$ occur, and the expected length of each $M_{a,b}$, as discussed in Section 4.5. For example, in Theorem 4.5, if we were to instead choose the matrix A with probability $\frac{2}{3}$ and B with probability $\frac{1}{3}$, then we would simply replace the term $2^{-a-b-m-n}$ with $2^{a+m} \cdot 3^{-a-b-m-n}$. We then find, via ratio tests, that the expected length of each $M_{a,b}$ is given by $\mathbb{E}(a+b) = \frac{11}{4}$, so we replace the term $\frac{1}{4}$ with $\frac{4}{11}$ and proceed as before. In general, these bounds can be obtained provided one can acquire a joint probability distribution for the matrices and those that precede them, and are not limited to solely geometric distributions.

Extending the theory to the case where we consider three or matrices at once is not as obvious. Let us consider a system where we choose from the matrices $A_1, A_2, \dots A_\Omega$ at random on each iterate, with respective probabilities $p_1, p_2, \dots p_\Omega$ (where $\sum_{i=1}^{\Omega} p_i = 1$). The first issue is the existence of a global, mutually invariant cone for the matrices, as such a cone would need to be invariant under any product $A_j^{a_j} A_i^{a_i}$ for $a_i, a_j \in \mathbb{N}$. Fortunately, determining whether such a cone exists is no more difficult than was the case for two matrices: one simply checks each possible

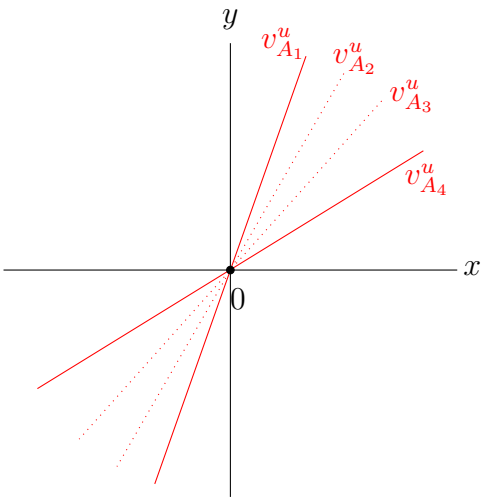


Figure 50: A sketch of the globally invariant cone in the case of a system which chooses from four matrices at random. The solid lines indicate the cones boundaries, in this case given by the unstable eigenvectors of the matrices A_1 and A_4, which yield the widest possible cone with boundaries given by unstable eigenvectors.

pair A_i and A_j to determine whether a mutually invariant cone exists for them. If such a cone exists for every pair of matrices, then a globally invariant cone is given by choosing the largest (widest) of these cones, as it will by definition encompass all of the other pairwise invariant cones (see Figure 50). The condition for such a cone existing remains similar: no stable eigenvectors for any of the matrices A_i can be located between the unstable eigenvectors for any pair A_j, A_k for $i, j, k \in \{1, \ldots, \Omega\}$.

Another problem arises in the choice of matrix product to consider; that is, the analogue to the matrices $M_{a,b} = B^b A^a$ which we used to rewrite the term $\|H^n v\|$ in the definition of the Lyapunov exponent. The most obvious choice appears to be the matrix

$$M_{a_1,a_2,\ldots,a_\Omega} = A_\Omega^{a_\Omega} \cdot \ldots \cdot A_2^{a_2} A_1^{a_1},$$

however, in order to account for sequences which skip one or more matrices in this sequence (e.g. $A_5^4 A_3^7$) or which occur in a different order (e.g. $A_2^8 A_7 A_4^3$) we must

allow for the numbers a_i to be zero. In principle this does not add any complications to the theory, as the resulting products could still be written down in general via diagonalization as in Chapter 4; in the bounds, the summations would begin from 0 instead of 1, and we would need to carefully consider the expected length of the matrices $M_{a_1,\ldots,a_\Omega}$.

In principle, it would appear that such bounds could be obtained for three or more matrices; they would be explicit and elementary as desired, and would require little modification to the theory already discussed. One concern may be that since the global cone is, in some sense, the 'worst-case scenario' out of the pairwise mutually invariant cones, the resulting bounds Ψ_k and Φ_k may not be especially accurate for low k. The cones should narrow significantly as k increases however, so it is possible this problem would not persist to high k values. The computing time for these bounds would increase significantly with k (at least exponentially), and so their practicality could also be called into question.

6.3 Random bounds: generalized Lyapunov exponents

Positivity of the maximal Lyapunov exponent of a system tells us that vectors will grow at exponential rates following sufficient iteration. While this is in itself useful to know due to reasons discussed previously, such as the indication of chaotic behaviour, one may wish to obtain further information about the convergence of the Lyapunov exponent. One method for achieving this is to observe the distribution obtained when calculating finite-time Lyapunov exponents, similar to those we saw in Chapter 5, and see how this evolves when taking increasing numbers of iterates. This can be a very lengthy procedure however, as one must calculate many finite-time Lyapunov exponents to obtain a sufficient sample size.

An alternative method is to study the *generalized Lyapunov exponents* [17], $\ell(q)$

for $q \in \mathbb{N}$, which are given by

$$\ell(q) = \lim_{n\to\infty} \log \mathbb{E}\|H^n v_0\|^q, \tag{125}$$

for $v_0 \in \mathbb{R}^2$. Each $\ell(q)$ gives the growth rate of the q^{th} moment of the norm of the matrix product H^n. In a similar way to how Lyapunov exponents characterise the expansion rates in tangent space, generalized Lyapunov exponents quantify the fluctuations of these growth rates. For a more detailed discussion of generalized Lyapunov exponents and their uses, as well as a method for estimating generalized Lyapunov exponents, see Vanneste [58] and the references therein.

In our case, and in a similar fashion to Sturman and Thiffeault [54], we note that

$$\begin{aligned}\mathbb{E}\|H^n v_0\|^q &= \mathbb{E}\left(\frac{\|M_{a_N,b_N} v_{N-1}\|}{\|v_{N-1}\|} \cdot \ldots \cdot \frac{\|M_{a_1,b_1} v_0\|}{\|v_0\|}\right)^q, \\ &= \mathbb{E}\left(\frac{\|M_{a_N,b_N} v_{N-1}\|}{\|v_{N-1}\|}\right)^q \cdot \ldots \cdot \mathbb{E}\left(\frac{\|M_{a_1,b_1} v_0\|}{\|v_0\|}\right)^q,\end{aligned}$$

since $\|v_i\|$ are independent of each other. Finally, since we choose the matrices at random, the a_i and b_i are independent and identically distributed random variables, and so we obtain the lower bound

$$\Psi_1^{(q)} = \frac{1}{4}\sum_{a,b=1}^{\infty} 2^{-a-b} \log\left(\min_{v\in C_{AB}}\left\{\frac{\|M_{a,b} v\|}{\|v\|}\right\}\right)^q \leq \ell(q), \tag{126}$$

and the upper bound

$$\Phi_1^{(q)} = \frac{1}{4}\sum_{a,b=1}^{\infty} 2^{-a-b} \log\left(\max_{v\in C_{AB}}\left\{\frac{\|M_{a,b} v\|}{\|v\|}\right\}\right)^q \geq \ell(q). \tag{127}$$

6.4 Deterministic bounds: general application

In Chapter 5 we derived and studied bounds for the Lyapunov exponents of a family of linked twist maps, as well as an additional map formed via shear composition which we introduced in Chapter 3. We note here that these bounds are, in principle, applicable to a wider range of systems than those presented in this thesis.

The most important condition in the acquisition of bounds of this form is the existence of invariant subspaces (cones) in tangent space. Due to this, the method naturally lends itself to finding bounds on Lyapunov exponents in systems with some form of hyperbolicity, as the existence of invariant subspaces is a necessity in such systems. An example of another system discussed in this thesis to which the method could be applied is the map originally studied by Cerbelli and Giona [14], which we discussed in Section 3.3, although it should be noted that the Lyapunov exponents of this map can be evaluated explicitly. In fact, the invariant cones one obtains for this map are simply the eigenvectors of the Jacobian corresponding to the hyperbolic region, and so our bounds reduce to the exact Lyapunov exponent; similarly, this happens in the case of systems such as Arnold's Cat Map, where the invariant subspaces also consist of a single vector.

The main limiting factor in the calculation of the bounds (assuming the pre-requisites, such as the existence of invariant cones, are met) in deterministic systems is the calculation of the return-time distribution, and subsequently the distributions characterising the matrices which can precede other matrices in an orbit. It should be noted that an exact expression of the distribution is not necessary to obtain bounds, as we showed in Section 5.6 where we found bounds upon certain elements of the distribution. If one could obtain an expression which could (non-trivially) bound the distribution elements at each subsequent iteration of the bounds, then in principle higher iterations of the bounds could become attainable. Note also that while the above statement will hold true in theory, the existence of a system where one could obtain useful bounds upon these distribution elements is not known to the author, and a way to bound the distribution elements of subsequent preceding distributions in the systems we have studied also appears unlikely.

References

[1] Roy L Adler and Benjamin Weiss. Entropy, a complete metric invariant for automorphisms of the torus. *Proceedings of the National Academy of Sciences*, 57(6):1573–1576, 1967.

[2] Vladimir Mikhailovich Alekseev. Quasi-random dynamical systems. *Mathematical Notes*, 6(4):749–753, 1969.

[3] Dmitrij V Anosov. *Geodesic flows on closed Riemann manifolds with negative curvature*. Number 90. American Mathematical Society, 1969.

[4] Hassan Aref. Stirring by chaotic advection. *Journal of fluid mechanics*, 143: 1–21, 1984.

[5] Ludwig Arnold. *Random dynamical systems*. Springer Science & Business Media, 2013.

[6] VI Arnold and A Avez. New results in qualitative dynamics.(Book reviews: Problemes ergodiques de la mecanique classique). *Science*, 159:1344, 1968.

[7] Arvind Ayyer and Mikko Stenlund. Exponential decay of correlations for randomly chosen hyperbolic toral automorphisms. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 17(4):043116, 2007.

[8] Luis Barreira and Ya B Pesin. *Lyapunov exponents and smooth ergodic theory*, volume 23. American Mathematical Soc., 2002.

[9] G Bennetin, L Galgani, A Giorgilli, and JM Strelcyn. Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems: A method for computing all of them. *Meccanica*, 15(9):27, 1980.

[10] George D Birkhoff. Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences*, 17(12):656–660, 1931.

[11] Philippe Bougerol et al. *Products of random matrices with applications to Schrödinger operators*, volume 8. Springer Science & Business Media, 2012.

[12] Robert Burton and Robert W Easton. Ergodicity of linked twist maps. In *Global Theory of Dynamical Systems*, pages 35–49. Springer, 1980.

[13] BF Bylov, RE Vinograd, DM Grobman, and VV Nemyckii. The theory of Lyapunov exponents and its applications to problems of stability. *Nauka, Moscow*, 1966.

[14] Stefano Cerbelli and Massimiliano Giona. A continuous archetype of nonuniform chaos in area-preserving dynamical systems. *Journal of Nonlinear Science*, 15(6):387–421, 2005.

[15] Joel E Cohen, Harry Kesten, and Charles Michael Newman. *Random matrices and their applications*, volume 50. American Mathematical Soc., 1986.

[16] John David Crawford and John R Cary. Decay of correlations in a chaotic measure-preserving transformation. *Physica D: Nonlinear Phenomena*, 6(2): 223–232, 1983.

[17] Andrea Crisanti, Giovanni Paladin, and Angelo Vulpiani. *Products of Random Matrices: in Statistical Physics*, volume 104. Springer Science & Business Media, 2012.

[18] Max Dehn. Die gruppe der abbildungsklassen. *Acta mathematica*, 69(1):135–206, 1938.

[19] Mark F Demers and Maciej P Wojtkowski. A family of pseudo-Anosov maps. *Nonlinearity*, 22(7):1743, 2009.

[20] Robert Devaney. *An introduction to chaotic dynamical systems*. Westview press, 2008.

[21] Manfred Einsiedler and Thomas Ward. *Ergodic theory*. Springer, 2013.

[22] Albert Fathi. Travaux de Thurston sur les surfaces. *Astérisque*, pages 66–67, 1979.

[23] Albert Fathi. Dehn twists and pseudo-Anosov diffeomorphisms. *Inventiones mathematicae*, 87(1):129–151, 1987.

[24] R Ferriere and M Gatto. Lyapunov exponents and the mathematics of invasion in oscillatory or chaotic populations. *Theoretical Population Biology*, 48(2): 126–171, 1995.

[25] Harry Furstenberg and Harry Kesten. Products of random matrices. *The Annals of Mathematical Statistics*, 31(2):457–469, 1960.

[26] Mark Kac. On the notion of recurrence in discrete stochastic processes. *Bulletin of the American Mathematical Society*, 53(10):1002–1010, 1947.

[27] Anatole Katok and Boris Hasselblatt. *Introduction to the modern theory of dynamical systems*, volume 54. Cambridge university press, 1997.

[28] Anatole Katok and Leonardo Mendoza. Dynamical systems with nonuniformly hyperbolic behavior. *Introduction to the Modern Theory of Dynamical Systems*, 1995.

[29] Anatole Katok and Jean-Marie Strelcyn. *Invariant manifolds, entropy and billiards. Smooth maps with singularities*, volume 1222. Springer, 2006.

[30] Eric S Key. Lower bounds for the maximal Lyapunov exponent. *Journal of Theoretical Probability*, 3(3):477–488, 1990.

[31] Andreĭ Nikolaevich Kolmogorov. *Foundations of the Theory of Probability: Second English Edition*. Courier Dover Publications, 2018.

[32] WB Raymond Lickorish. A representation of orientable combinatorial 3-manifolds. *Annals of Mathematics*, pages 531–540, 1962.

[33] A. M. Lyapunov. The general problem of motion stability. *Annals of Mathematics Studies*, 17, 1892.

[34] Robert S MacKay. Cerbelli and Giona's map is pseudo-Anosov and nine consequences. *Journal of Nonlinear Science*, 16(4):415–434, 2006.

[35] Valery Iustinovich Oseledec. A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc*, 19(2): 197–231, 1968.

[36] Thomas S Parker and Leon Chua. *Practical numerical algorithms for chaotic systems*. Springer Science & Business Media, 2012.

[37] Robert C Penner. A construction of pseudo-Anosov homeomorphisms. *Transactions of the American Mathematical Society*, 310(1):179–197, 1988.

[38] Oskar Perron. Die ordnungszahlen linearer differentialgleichungssysteme. *Mathematische Zeitschrift*, 31(1):748–766, 1930.

[39] Yakov Borisovich Pesin. Characteristic Lyapunov exponents and smooth ergodic theory. *Russian Mathematical Surveys*, 32(4):55–114, 1977.

[40] Henri Poincaré. Sur le probleme des trois corps et les équations de la dynamique. *Acta mathematica*, 13(1):A3–A270, 1890.

[41] Henri Poincaré and R Magini. Les méthodes nouvelles de la mécanique céleste. *Il Nuovo Cimento (1895-1900)*, 10(1):128–130, 1899.

[42] Mark Pollicott. Maximal Lyapunov exponents for random matrix products. *Inventiones mathematicae*, 181(1):209–226, 2010.

[43] Awadhesh Prasad and Ramakrishna Ramaswamy. Characteristic distributions of finite-time Lyapunov exponents. *Physical Review E*, 60(3):2761, 1999.

[44] Feliks Przytycki. Ergodicity of toral linked twist mappings. In *Annales scientifiques de l'Ecole normale supérieure*, volume 16, pages 345–354. Elsevier, 1983.

[45] Clark Robinson. *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos*. CRC Press, Inc, 1995.

[46] Leiba Rodman, Hakan Seyalioglu, and Ilya M Spitkovsky. On common invariant cones for families of matrices. *Linear Algebra and its Applications*, 432(4):911–926, 2010.

[47] Marco Sandri. Numerical calculation of Lyapunov exponents. *Mathematica Journal*, 6(3):78–84, 1996.

[48] Kenichi Shiraiwa. Manifolds which do not admit Anosov diffeomorphisms. *Nagoya Mathematical Journal*, 49:111–115, 1973.

[49] Ya G Sinai. Markov partitions and C-diffeomorphisms. *Functional Analysis and its applications*, 2(1):61–82, 1968.

[50] Julia Slipantschuk, Oscar F Bandtlow, and Wolfram Just. On the relation between Lyapunov exponents and exponential decay of correlations. *Journal of Physics A: Mathematical and Theoretical*, 46(7):075101, 2013.

[51] James Springham. Ergodic properties of linked-twist maps. *arXiv preprint arXiv:0812.0899*, 2008.

[52] Gilbert Strang. *Introduction to linear algebra*, volume 3. Wellesley-Cambridge Press Wellesley, MA, 1993.

[53] Rob Sturman and James Springham. Rate of chaotic mixing and boundary behavior. *Physical Review E*, 87(1):012906, 2013.

[54] Rob Sturman and Jean-Luc Thiffeault. Lyapunov exponents for the random product of two shears. *arXiv preprint arXiv:1706.03398*, 2017.

[55] Rob Sturman, Julio M Ottino, and Stephen Wiggins. *The mathematical foundations of mixing: the linked twist map as a paradigm in applications: micro to macro, fluids to solids*, volume 22. Cambridge University Press, 2006.

[56] Jean-Luc Thiffeault, Matthew D Finn, Emmanuelle Gouillart, and Toby Hall. Topology of chaotic mixing patterns. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 18(3):033123, 2008.

[57] William P Thurston et al. On the geometry and dynamics of diffeomorphisms of surfaces. *Bulletin (new series) of the American Mathematical Society*, 19(2): 417–431, 1988.

[58] J Vanneste. Estimating generalized Lyapunov exponents for products of random matrices. *Physical Review E*, 81(3):036701, 2010.

[59] Maciej Wojtkowski. Linked twist mappings have the K-property. *Annals of the New York Academy of Sciences*, 357(1):65–76, 1980.

[60] Lai-Sang Young. Statistical properties of dynamical systems with some hyperbolicity. *Annals of Mathematics*, 147(3):585–650, 1998.

[61] Lai-Sang Young. Recurrence times and rates of mixing. *Israel Journal of Mathematics*, 110(1):153–188, 1999.