Markov partition and symbolic dynamics of hyperbolic toral automorphisms
Abstract

Hyperbolic toral automorphisms are naturally defined on the torus $\mathbb{T}^2$. These maps are chaotic on the entire torus $\mathbb{T}^2$ but they can be completely described. They are induced by linear maps in $\mathbb{R}^2$ but they have a much more complex and rich dynamical structure due to their chaotic property. We study hyperbolic toral automorphisms using a Markov partition. A Markov partition allows us to apply symbolic dynamics to a dynamical system that is topologically equivalent to the original system.

Keywords. Markov partition, symbolic dynamics, diffeomorphism, Anosov systems, Arnold’s cat map.
## Contents

1. Introduction ................................................. 4

2. Linear maps ................................................ 9

3. Hyperbolic toral automorphisms ....................... 13
   3.1 New properties arising only on the torus ........ 13

4. Markov partition ............................................ 15
   4.1 Construction ............................................ 16

5. Symbolic dynamics ......................................... 22

6. Conclusion .................................................. 27

7. Acknowledgments ........................................... 28

8. Bibliography ................................................. 29

9. Appendix ..................................................... 30
1 Introduction

In this thesis, we look at chaotic systems which can be completely described. Hence though they appear random, they possess an underlying structure. This structure is the main point of this thesis. The chaotic system that we will look at are hyperbolic toral automorphisms. Hyperbolic toral automorphisms were first introduced by D. V. Anosov, hence why they are also called Anosov systems. Several scholars like Robert L. Devaney and Clark Robinson have already investigated and written about this particular topic.

Hyperbolic toral automorphisms, which I will refer to as maps from now on, are induced by linear maps on the Euclidean plane. We will use this property to our advantage to describe the dynamics of these maps to some extent since the dynamics on the Euclidean plane are simple. To do this, we make use of the fact that we can build a torus from the $\mathbb{R}^2$ plane. First, let $[\alpha, \beta]$ be the set of all points in the $\mathbb{R}^2$ plane with an integer difference with $(\alpha, \beta)$, i.e.,

$$(x, y) \in [\alpha, \beta] \iff (x, y) - (\alpha, \beta) = (N, M), \text{ where } N, M \in \mathbb{Z}.$$  

Equivalently,

$$(x, y) \in [\alpha, \beta] \iff (x, y) \equiv (X \mod 1, Y \mod 1), \text{ where } (X, Y) \in \mathbb{R}^2.$$  

We say that all points $(x, y) \in [\alpha, \beta]$ make an equivalence class of $(\alpha, \beta)$. By doing this, we obtain all the equivalence classes. By starting in the unit square $0 \leq x, y \leq 1$, we only need to identify the boundaries of the unit square according to the above equivalence relation to obtain all equivalence classes. The line $y = 1$ is equivalent to the line $y = 0$, hence we can identify them as one line to obtain a cylinder from the square. The line $x = 1$ is equivalent to the line $x = 0$, hence by identifying them as one line we obtain a torus from the cylinder. Hence the torus is the set of all these equivalence classes. This construction gives us a natural projection

$$\pi : \mathbb{R}^2 \to \mathbb{T}^2$$

$$: (x, y) \to [x, y].$$

When we operate on $\mathbb{R}^2$ we use a linear map $L$, and when we are on $\mathbb{T}^2$ we use the induced hyperbolic toral automorphism $L_A$. This can be given by the following commutative diagram:
We will elaborate on this in the next section. We will also look at the dynamics that arise on the torus but not on the Euclidean space. Then we can finally construct a Markov partition on the torus, or rather on the unit square which can be made a torus through the equivalence relation. A Markov partition is a collection of rectangles, constructed in such a way that

- two sides of each rectangle lie in the stable set and the other two sides lie in the unstable set,
- if the forward iteration of one rectangle meets the interior of another rectangle, the image cuts completely across the latter rectangle in the unstable direction. Therefore, forward images of rectangles yield unstable rectangles,
- if the backward iteration of one rectangle meets the interior of another rectangle, the image cuts completely across the latter rectangle in the stable direction. Hence, backward images of rectangles yield stable rectangles.

This will enable us to apply symbolic dynamics in order to describe the trajectory of the points on the torus. By doing this, we obtain the full understanding of Hyperbolic toral automorphisms.

Example 1. (“Arnold’s cat map”)
To obtain an intuition of what this thesis is about, we show Arnold’s cat map. Arnold’s Cat Map is named after the Russian mathematician Vladimir I. Arnold, who first discovered it using a cat image. Arnold’s cat map $\Gamma : \mathbb{T}^2 \rightarrow \mathbb{T}^2$ is given by:

$$
\Gamma(\mathbf{x}) = A\mathbf{x} \mod 1
$$

where $\mathbf{x} \in \mathbb{R}^2$ and

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

$A$ is also called a representation matrix. Any image, in the unit square in $\mathbb{R}^2$, will shift one unit up in the $y$-direction and two units to right in the $x$-direction, this
corresponds to the image after the first red arrow in Figure 1. Taking “mod 1” corresponds to shifting the pieces of the image outside of the unit square back into the unit square, see the image after the long red arrow in Figure 1. Here is the first iteration of a cat image:

In Figure 2 we show how the step of taking mod 1 happens:
Since the determinant of $A$ is 1, any image in $\mathbb{R}^2$ that has been transformed by this matrix will retain its original area. Therefore, the area of second image in Figure 2(a) is the same as the area of the first image. The effect of taking mod 1 is to cut up the second image and reassemble it into the unit square. This is the completed mapping of Arnold’s Cat Map performed once on any image in the $\mathbb{R}^2$ plane.

Example 2. ("Discrete cat map")

In this example we will show Arnold’s cat map applied on an image with pixels. The pixels represent some color value which, put together, give an image. Hence an image can be represented by a matrix where each entry of the matrix will give a numeric value of the corresponding pixel (color) in the image at hand. The midpoints of the pixels have rational coordinates, hence they are periodic. Therefore, when we apply Arnold’s cat map on any image with pixels in the $\mathbb{R}^2$ plane, the original image will reappear after a finite number of iterations. Here follows the full cycle of a cat image of $150 \times 150$ pixels:
After 300 iterations, the original image reappears. Hence the cat image at hand has a period of 300. Though the map $\Gamma : T^2 \to T^2$ is chaotic, we are fully able to follow its dynamics.

Though we study chaotic dynamical systems which we are able to describe and understand completely, there are still many chaotic dynamical systems that we are not able to describe. Nonetheless, chaos gives us the tools needed to approach mathematical and physical phenomena with competence where before we were at a loss. Here are some examples of chaos in nature:

- the beating of a heart,
- the eccentricity of the planet Pluto’s orbit,
- the Great Red Spot of Jupiter,
• the turbulence of fluids,
• the capricious weather,
• the swinging of a pendulum,
• the formation of a snowflake.

Chaos brings us a step closer to understanding the universe we live in.

2 Linear maps

Definition 1. A map \( L : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is linear if

\[
L(\alpha v + \beta w) = \alpha L(v) + \beta L(w), \quad \text{where } \alpha, \beta \in \mathbb{R} \text{ and } v, w \in \mathbb{R}^n.
\]

Using standard basis vectors \( e_1, e_2, \ldots, e_n \) of \( \mathbb{R}^n \), one can find a matrix representation of the linear map as follows. Let \( L(e_i) = v_i \) for \( i = 1, \ldots, n \) and decompose the vector \( v \in \mathbb{R}^n \) into standard basis vectors \( e_i \) for \( i = 1, \ldots, n \):

\[
v = \begin{bmatrix}
\alpha_1 \\
\alpha_2 \\
\vdots \\
\alpha_n
\end{bmatrix}
= \alpha_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \ldots + \alpha_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}
= \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n.
\]

Then

\[
L(v) = L(\sum_{i=1}^{n} \alpha_i e_i) = \sum_{i=1}^{n} \alpha_i L(e_i) = \sum_{i=1}^{n} \alpha_i v_i.
\]

Using the \( v_i \)'s with \( i = 1, \ldots, n \), we construct a matrix \( A \) as follows:

\[
A = \begin{bmatrix} v_1 & v_2 & \ldots & v_n \end{bmatrix}
\]

Then for some \( x \in \mathbb{R}^n \), it follows that

\[
L(x) = Ax.
\]
A is called the matrix representation of the linear map \( L \). In the rest of the thesis we only deal with invertible linear maps and we fix \( n = 2 \) because we mainly focus on 2-dimensional dynamical systems. The results we obtain can be generalized to higher dimensional dynamical systems.

For the study of hyperbolic toral automorphisms, the matrix \( A \) has to satisfy the following conditions:

1. All entries of \( A \) are integers.
2. \( \det(A) \) is \( \pm 1 \).
3. \( A \) is hyperbolic, i.e., none of the eigenvalues of \( A \) have absolute value one. Moreover, all the eigenvalues must be real.

The first condition ensures that \( L \) takes the integer lattice \( \mathbb{Z}^2 \) into itself. The second condition ensures that \( A^{-1} \) is also a matrix with only integers as entries. Consequently, \( L^{-1} \) also takes the integer lattice \( \mathbb{Z}^2 \) into itself. The third condition ensures that the toral automorphisms are hyperbolic. Moreover, we know a bit more about the eigenvalues of \( A \), i.e.,

**Theorem 1.** If \( A \) is hyperbolic, then the eigenvalues of \( A \) are irrational.

**Proof.** Suppose \( A \) is hyperbolic and let \( t = \text{tr}(A) \) be the trace of \( A \), i.e., the sum of the diagonal elements of \( A \). We can write the eigenvalues of \( A \) as

\[
\frac{t \pm \sqrt{t^2 - 4\det(A)}}{2}.
\]

This equation is only rational if \( t^2 - 4\det(A) = s^2 \) for some \( s \in \mathbb{Z} \). This occurs only when \( t \in \{0,2\} \), but then \( A \) has \( \pm 1 \) as its eigenvalues. This is a contradiction since \( A \) is hyperbolic. Therefore, the eigenvalues of \( A \) must be irrational. \( \square \)

Since we are working with 2-dimensional dynamical systems, we also know that \( A \) has two eigenvalues. Since \( A \) is hyperbolic and \( \det(A) = \pm 1 \), one of the eigenvalues is in the interval \((-1,1)\) and the absolute value of the second eigenvalue is strictly bigger than 1. This has the following implications:

- if \( \lambda \in (-1,1) \), then the system will be contracted in the direction of the corresponding eigenvector. All forward iterations will converge to a fixed point, which is the origin in this case:

\[
L^n(x) \to 0 \text{ as } n \to \infty \text{ for all } x \in \text{span}\{v\},
\]

where \( v \) is the eigenvector corresponding to \( \lambda \in (-1,1) \).
• if $|\lambda| > 1$, then the system will be expanded in the direction of the corresponding eigenvector. Here, the backward iterations converge to a fixed point, which is also the origin:

$$L^{-n}(x) \to 0 \text{ as } n \to \infty \text{ for all } x \in \text{span}\{w\},$$

where $w$ is the eigenvector corresponding to $|\lambda| > 1$.

Based on these implications, it is useful to define the concepts of stable and unstable subspace.

**Definition 2.**

1. The line $W^s$ through the origin is called the stable subspace of $L$ if for every $x \in W^s$ it holds that $L(x) \in W^s$ and $L^n(x) \to 0$ as $n \to \infty$.

2. The line $W^u$ through the origin is called the unstable subspace of $L$ if for every $x \in W^u$ it holds that $L(x) \in W^u$ and $L^{-n}(x) \to 0$ as $n \to \infty$.

3. If $x \notin W^s \cup W^u$ then $|L^n(x)| \to \infty$ as $n \to \pm \infty$.

**Example 3.** For computational purposes, we will use the following representation matrix throughout the thesis:

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

Its eigenvalues are:

$$\lambda_1 = \frac{1 + \sqrt{5}}{2}$$

$$\lambda_2 = \frac{1 - \sqrt{5}}{2}$$

and the corresponding eigenvectors are:

$$v_1 = \begin{pmatrix} 1 \\ -1 + \sqrt{5} \end{pmatrix}$$

$$v_2 = \begin{pmatrix} 1 \\ -1 - \sqrt{5} \end{pmatrix}$$

Since $|\lambda_1| > 1$ then $\lambda_1 = \lambda_u$ is the unstable eigenvalue. Similarly, $\lambda_2 \in (-1, 1)$ so $\lambda_2 = \lambda_s$ is the stable eigenvalue. It follows that $v_1 = v_u$ and $v_2 = v_s$ are the unstable and stable eigenvectors respectively.
Here is the plot of the stable and unstable subsets of $A$ through the origin,

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{plot.png}
\caption{Lines spanned by the eigenvectors of $A$}
\end{figure}

Later on, we will see how to use these subsets of $A$ to construct a Markov partition. We proved that the eigenvalues of $A$ are irrational if $A$ is hyperbolic, a direct consequence of that is the following:

**Proposition 1.** The slopes of the lines in the direction of the eigenvectors of $A$ are irrational.

**Proof.** We prove this by contradiction. A line $l$ through the origin has a rational slope if and only if $l \cap \mathbb{Z}^2 \neq \{(0,0)\}$. This means, if $l = \{(x,y) | y = \frac{c}{d}x\}$ then $(c,d) \in l$. Vice versa, if $(m,n) \in l$ and $l$ is given by the equation $y = \alpha x$, then $n = \alpha m$. Hence $\alpha = \frac{n}{m}$. Now assume that $l \cap \mathbb{Z}^2 = \{(0,0)\}$, and take $\lambda \in (-1,1)$. Since $A(l) = l$ and $A(\mathbb{Z}^2) = \mathbb{Z}^2$ then $A(l \cap \mathbb{Z}^2) = l \cap \mathbb{Z}^2$. Let $(m,n) \in l \cap \mathbb{Z}^2$, where $m$ and $n$ are not simultaneously zero, be a point closest to the origin. Then $A(m,n) \in l \cap \mathbb{Z}^2$ and $A(m,n) \neq (0,0)$ since $A$ is invertible. It follows that $d(A(m,n),(0,0)) = |\lambda|d((m,n),(0,0)) < d((m,n),(0,0))$, where $d$ is the Euclidean distance. This gives us a contradiction. Hence $l \cap \mathbb{Z}^2 = \{(0,0)\}$ and $l$ has an irrational slope. The case of $|\lambda| > 1$ is done by similar reasoning. \qed
3 Hyperbolic toral automorphisms

First things first, let us define what the words hyperbolic and automorphism mean.

**Definition 3.** An invertible linear map is hyperbolic if none of its eigenvalues, i.e., the eigenvalues of its representation matrix, have absolute value one.

**Definition 4.** An automorphism $f$ is an isomorphism from a smooth manifold $M$ to itself.

When $M$ is defined to be a torus $\mathbb{T}^2$ and $f$ is hyperbolic, then one speaks of a hyperbolic toral automorphism. The linear map $L : \mathbb{R}^2 \to \mathbb{R}^2$ induces a well defined automorphism $L_A$ on the torus $\mathbb{T}^2$.

**Note.** The projection map $\pi$ is not bijective since it is not injective. This is because points in the same equivalence class will be mapped to one point on the torus $\mathbb{T}^2$. The consequence of this way of mapping though, is that the map $L_A : \mathbb{T}^2 \to \mathbb{T}^2$ is by construction an automorphism.

Since $L$ is a diffeomorphism on $\mathbb{R}^2$ it follows that $L_A$ is also a diffeomorphism on $\mathbb{T}^2$. Hence all properties that $L$ has, will be inherited by $L_A$. But there are some properties that only arise on $\mathbb{T}^2$ but not on $\mathbb{R}^2$.

3.1 New properties arising only on the torus

First, a reminder of the words periodic and dense since that we will use them quite often.

**Definition 5.** A point $[p] \in \mathbb{T}^2$ is periodic if $L^n_A[p] = [p]$ for some $n \in \mathbb{Z}$.

**Definition 6.** A subset $S$ of the torus $\mathbb{T}^2$ is dense in $\mathbb{T}^2$ if the closure of $S$ is the torus $\mathbb{T}^2$ itself, i.e., $\overline{S} = \mathbb{T}^2$.

Using definition 5 and 6, we can prove the following property:

**Proposition 2.** Periodic points are dense in $\mathbb{T}^2$.

**Proof.** Let $[p]$ be any rational point in $\mathbb{T}^2$. We can always find a common denominator such that $[p]$ is of the form $[\alpha \frac{a}{k}, \beta \frac{b}{k}]$, where $\alpha, \beta, k \in \mathbb{Z}$. Note that there are $k^2$ such points in $\mathbb{T}^2$ and by choosing $k$ arbitrary large, such points are dense in $\mathbb{T}^2$.

We want to prove that $[p]$ is periodic with a period less or equal to $k^2$. Note that the iterations of points of the form $[\alpha \frac{a}{k}, \beta \frac{b}{k}]$ with $0 \leq \alpha, \beta < k$ by $L_A$ are also of the same form as $[p]$ since $A$ is an integer matrix. In other words, $L_A$ only permutes such points. Hence there exist $j, i \in \mathbb{Z}$ such that $L^j_A([p]) = L^i_A([p])$ and $|i - j| \leq k^2$. By applying $L^{-j}_A$ to both sides of the equation we obtain $[p] = L^{-j}_A L^i_A([p]) = L^{-j+i}_A([p])$.

This shows that $[p]$ is periodic with a period less or equal to $k^2$. Therefore periodic points are dense in $\mathbb{T}^2$. 

\[\square\]
For us to be able to fully analyse the dynamics on the torus, we desperately need the dynamics to always stay on the torus. Hence the non-wandering property comes in handy.

**Definition 7.** A point \([p]\) is called non-wandering for \(L_A\) if, for any open neighbourhood \(U\) containing \([p]\), there exists an integer \(n > 0\) such that \(L_A^n(U) \cap U \neq \emptyset\). Thus, there exist a point \([q]\) \(\in U\) with \(L_A^n[q] \in U\).

**Note.** We do not require that \([p]\) returns to \(U\).

From proposition 2 and definition 7 we can prove the following:

**Proposition 3.** There are infinitely many periodic points and the non-wandering set of \(L_A\) is all of \(T^2\).

**Proof.** Since \(k \in \mathbb{Z}\) in the proof of proposition 2 was arbitrary and there are infinitely many integers to choose from, it follows that there are infinitely many periodic points in \(T^2\). It also follows that the non-wandering set of \(L_A\) is all of \(T^2\) because for all neighbourhood \(U\) around any \([p]\) \(\in T^2\), there is a rational point \([q]\) \(\in U\) with \(L_A^n[q] \in U\) where \(n > 1\). Take \(n\) equal to the period of \([q]\), then we have found our non-wandering point in \(U\). Since periodic points are dense in \(T^2\), then the non-wandering set of \(L_A\) is all \(T^2\). □

One can not speak of periodic points and forget homoclinic points.

**Definition 8.** Let \([p]\) \(\in T^2\) be a periodic point for \(L_A\). A homoclinic point to \([p]\) is a point \([q] \neq [p]\) which lies in \(W^s[p] \cap W^u[p]\). In the case of toral automorphisms, \(W^s[p]\) and \(W^u[p]\) meet at a non-zero angle. In this case the homoclinic points are then called transverse.

**Proposition 4.** Transverse homoclinic points are dense in \(T^2\).

**Proof.** We have proved that the slopes of the lines in the direction of the eigenvectors of \(A\) are irrational, hence \(W^s[p]\) and \(W^u[p]\) are dense in \(T^2\) for any \([p]\) \(\in T^2\). Their intersection is by definition also dense in \(T^2\). Hence Transverse homoclinic points are dense in \(T^2\). □

Finally, let us define what topologically transitive and sensitive dependence on initial conditions mean for hyperbolic toral automorphisms.

**Definition 9.** The toral automorphism \(L_A\) is topologically transitive, i.e., for any non-empty open subsets \(U\) and \(V\) of \(T^2\) and for some point \([p]\) \(\in U\), there exists a positive integer \(k\) such that \(L^k_A[p] \in V\). That means every pair of two non-empty open subsets of \(T^2\) shares a periodic orbit, equivalently \(L^k_A(U) \cap V \neq \emptyset\).
Proposition 5. The toral automorphism $L_A$ has sensitive dependence on initial conditions.

Proof. For every point $[p] \in \mathbb{T}^2$ we know its stable and unstable subset in $\mathbb{T}^2$. Then for every $[r] \in W^u[p]$, $d([p], [r]) < d(L^n_A[p], L^n_A[r])$ for some $n \in \mathbb{Z}$ and $n > 1$. Similarly, for every $[s] \in W^s[p]$, $d([p], [s]) < d(L^{-m}_A[p], L^{-m}_A[s])$ for some $m \in \mathbb{Z}$ and $m > 1$. Therefore, $L_A$ indeed has sensitive dependence on initial conditions. □

Definition 9 and proposition 2 and 5 combined give us the most important property for the hyperbolic toral automorphism $L_A$.

Definition 10. $L_A$ is chaotic on the entire torus $\mathbb{T}^2$ if and only if

1. periodic points under the action of $L_A$ are dense in $\mathbb{T}^2$,
2. $L_A$ is topologically transitive, and
3. $L_A$ has sensitive dependence on initial conditions.

Now we can study the dynamics on $\mathbb{T}^2$ using symbolic dynamics. To do so, we first need to construct a Markov partition on $\mathbb{T}^2$.

4 Markov partition

Symbolic dynamics represent trajectories by infinite length sequences using a finite number of symbols. To represent the state space of a dynamical system with a finite number of symbols, we must partition the space into a finite number of rectangles and assign a symbol to each one. The rectangles must also satisfy a few conditions to be a Markov partition:

Definition 11. A Markov partition for $L_A : \mathbb{T}^2 \to \mathbb{T}^2$ is a finite collection of rectangles $\mathcal{R} = \{R_i\}_{i=1}^n$ that satisfies the following conditions:

1. The collection of rectangles must cover $\mathbb{T}^2$, i.e., $\mathbb{T}^2 = \bigcup_{i=1}^n R_i$.
2. If $i \neq j$, then $\text{int}(R_i) \cap \text{int}(R_j) = \emptyset$ and $\text{int}(R_i) \cap R_j = \emptyset$.
3. If $z \in \text{int}(R_i)$ and $L_A(z) \in \text{int}(R_j)$, then
   
   $L_A(W^u(z, R_i)) \supset W^u(L_A(z), R_j)$ and
   $L_A(W^s(z, R_i)) \subset W^s(L_A(z), R_j)$

The third condition basically means that
• if $L_A(R_i) \cap \text{int}(R_j) \neq \emptyset$, then the image of $R_i$ will cut across $R_j$ in the unstable direction.

• if $L_A^{-1}(R_i) \cap \text{int}(R_j) \neq \emptyset$, then the image of $R_i$ will cut across $R_j$ in the stable direction.

**Note.** As the image of $R_i$ crosses $R_j$ either in the stable or unstable direction, it will stop when it hits the next boundary line.

According to these properties, two sides of each $R_i \in \mathcal{R}$ are in the unstable direction and the other two sides of $R_i \in \mathcal{R}$ are in the stable direction.

**Note.** Any finite set of rectangles that satisfies the three conditions in definition 11 gives a Markov partition of the torus. Hence a Markov partition for a torus is not unique.

### 4.1 Construction

The construction will be done on the unit square in $\mathbb{R}^2$. Using the points with coordinates: $(0, 0), (1, 0), (0, 1)$ and $(1, 1)$, we draw $W^u$ and $W^s$ of each of these points inside the unit square in the following manner:

1. From $(0, 0)$ draw a line $l_{(0,0)}$ in the unstable direction and stops when it hits the boundary of the unit square.

2. Then from $(0, 1)$ and $(1, 0)$ draw the lines $l_{(0,1)}$ and $l_{(1,0)}$ in the stable direction. These lines stop when they meet $l_{(0,0)}$. The intersection of $l_{(0,0)}$ with $l_{(0,1)}$ is marked by a red dot, and the intersection of $l_{(0,0)}$ with $l_{(1,0)}$ is marked by a blue dot.

3. From $(1, 1)$ draw a line $l_{(1,1)}$ in the unstable direction and this line stops when it meets $l_{(0,1)}$. The intersection of $l_{(1,1)}$ with $l_{(0,1)}$ is marked by a yellow dot.

4. Using symmetries, draw an extension of line $l_{(0,0)}$ to complete the rectangles. The intersection of this extension with $l_{(0,1)}$, is marked by a green dot.

Each of these four steps is shown in Figure 5 below:
These rectangles satisfy the first two conditions in definition 11. The third condition is also satisfied, but to see this we have to apply $L_A$ on the Markov partition at least once. We first label the rectangles, as depicted in Figure 6 below, so that we can clearly see how the rectangles transform after we apply $L_A$. Secondly, we show the image of the intersection points after we apply $L_A$:
We find the image of the rectangles by computing the image of random points in each rectangle. We immediately see whether $L_A(R_i) \cap R_j$ for $i \neq j$ and $i, j = 1, 2, 3$:

- For $R_1$:
  
  (0.1, 0.8) $\xrightarrow{L_A} (0.9, 0.1)$
  
  (0.45, 0.3) $\xrightarrow{L_A} (0.75, 0.45)$
  
  (0.9, 0.9) $\xrightarrow{L_A} (0.8, 0.9)$
  
  (0.9, 0.6) $\xrightarrow{L_A} (0.5, 0.9)$

We obtain the following plots:

- For $R_2$:
\[(0.75, 0.3) \xrightarrow{L_A} (0.05, 0.75)\]
\[(0.1, 0.95) \xrightarrow{L_A} (0.05, 0.1)\]
\[(0.3, 0.6) \xrightarrow{L_A} (0.9, 0.3)\]

We obtain the following plots:

(a) 3 points in \(\mathbb{R}^2\)

(b) The images of the 3 points

Figure 8: First iteration in \(\mathbb{R}^2\)

- For \(\mathbb{R}^3\):

\[(0.2, 0.6) \xrightarrow{L_A} (0.8, 0.2)\]
\[(0.4, 0.3) \xrightarrow{L_A} (0.7, 0.4)\]
\[(0.8, 0.4) \xrightarrow{L_A} (0.2, 0.8)\]

We obtain the following plots:

(a) 3 points in \(\mathbb{R}^3\)

(b) The images of the 3 points

Figure 9: First iteration in \(\mathbb{R}^3\)
Using the images of the intersection points and the images of the points of each rectangle as guide, we can plot the images of $R_1, R_2$ and $R_3$:

Figure 10: $L_A$ applied to the rectangles

Comparing part (a) to part (b) of Figure 7, 8 and 9, we note that:

- $L_A(R_1) \cap \text{int}(R_2) \neq \emptyset$ and $L_A(R_1) \cap \text{int}(R_3) \neq \emptyset$.
- $L_A(R_2) \cap \text{int}(R_1) \neq \emptyset$ and $L_A(R_2) \cap \text{int}(R_3) \neq \emptyset$.
- $L_A(R_3) \cap \text{int}(R_2) \neq \emptyset$.

Using the third condition in definition 11, $L_A(R_i)$ cut across $R_j$ in the unstable direction for $i, j = 1, 2, 3$ to obtain part(b) of Figure 10. Therefore, the third condition of definition 11 is also satisfied. We finally have the first forward iteration of the Markov partition by $L_A$:

Figure 11: The first iteration of the Markov partition

20
Remark 1. The construction of the pre-image (the backward iteration) of the Markov partition follows the same techniques and reasoning as the construction of the image (the forward iteration) of the Markov partition. The only difference is that we use $L^{-1}_A$ during the computation.

From the above non-empty intersections $L(R_i) \cap \text{int}(R_j)$ for $i, j = 1, 2, 3$, we can write the transition matrix for our system:

$$B = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Note. The transition matrix $B$ does not admit fixed points, but we do have a fixed point in $T^2$, namely $[0]$. Moreover, there are some overlaps at the boundaries of the rectangles that we ignored, namely:

- $L_A(R_1)$ and $R_1$ overlap along the boundaries of $R_1$ and $R_2$,
- $L_A(R_2)$ and $R_2$ overlap along the boundaries of $R_1$, $R_2$ and $R_3$,
- $L_A(R_3)$ and $R_3$ overlap along the boundaries of $R_1$, $R_2$ and $R_3$,
- $L_A(R_3)$ and $R_1$ also overlap along the boundaries of $R_1$, $R_2$ and $R_3$.

We ignored these overlaps since the points on the stable and unstable sets will stay on the stable and unstable sets. Moreover, we will show how to get the symbolic dynamics of the points in the overlaps. As we will see, there is not just one unique way to write sequences for such points. Hence this will give rise to some identification of the sequences of the points in the overlaps. Also, there are points that reach the stable and unstable sets after some iterations. After reaching the stable and unstable sets, they will be treated in a similar way as the points in the overlaps. But first, let us study symbolic dynamics.
5 Symbolic dynamics

In our case the Markov partition is a collection of three rectangles, hence our the symbol space is given by \{1, 2, 3\}. We then define a map on the symbol space. The map will give the itinerary of every point on the torus as the time evolves, i.e., the map memorizes the label of the rectangle in which the point is for each iteration.

**Definition 12.** The itinerary of \(x\) is a sequence \(S(x) = (\ldots s_{-2}s_{-1}s_0s_1s_2\ldots)\) where \(s_j = i\) if \(L^j_A(x) \in R_i\) for \(i = 1, 2, 3\).

The negative indices tell us about the backward iterations and the positive indices tell us about the forward iterations. Therefore, we can completely track each point. Hence the name "Symbolic dynamics".

Let \(\Sigma_3 = \{s = (\ldots s_{-1}s_0s_1\ldots) | s_j = 1, 2, 3\}\) be the sequence space. We can define a metric on this sequence space as follows:

**Theorem 2.** Let \(s, t \in \Sigma_3\) where \(s = (\ldots s_{-1}s_0s_1\ldots)\), \(t = (\ldots t_{-1}t_0t_1\ldots)\). Then,

\[
d[(s), (t)] = \sum_{i=-\infty}^{\infty} \frac{\delta(s_i, t_i)}{2^{|i|}} \quad \text{where} \quad \delta(l, j) = \begin{cases} 0 & \text{if } l = j \\ 1 & \text{if } l \neq j \end{cases}
\]

defines a metric on \(\Sigma_3\).

**Proof.** Let \(s, t, r \in \Sigma_3\), then

1. For any \(s, t \in \Sigma_3\), \(d[(s), (t)] \geq 0\) because \(\delta\) is never negative and \(d[(s), (t)] = 0\) if and only if \(s_i = t_i\) for all \(i \in \mathbb{Z}\).

2. \(d[(s), (t)] = d[(t), (s)]\) by definition of \(\delta\).

3. Because \(\delta(r_i, s_i) + \delta(s_i, t_i) \geq \delta(r_i, t_i)\) for any \(r, s, t \in \Sigma_3\), it follows that \(d[(r), (s)] + d[(s), (t)] \geq d[(r), (t)]\)

Hence \(d[(s), (t)] = \sum_{i=-\infty}^{\infty} \frac{\delta(s_i, t_i)}{2^{|i|}}\) is a metric on \(\Sigma_3\).

On \(\Sigma_3\) we also want to define a Shift map \(\sigma : \Sigma_3 \to \Sigma_3\) as follows:

\[
\sigma(s) = t \quad \text{where} \quad t_k = s_{k+1}.
\]

\(\sigma\) is invertible, i.e, \(\sigma^{-1}(t) = s\) where \(s_k = t_{k-1}\).

**Proposition 6.** \(\sigma\) is continuous.
Before we can prove the continuity of $\sigma$, we need the following theorem:

**Theorem 3.** *(The proximity theorem)*

1. If $d((s), (t)) < \frac{1}{2^m}$, then $s_i = t_i$ for $i = 0, 1, \ldots, m$.
2. If $s_i = t_i$ for $i = 0, 1, \ldots, m$, then $d((s), (t)) < \frac{1}{2^m}$.

According to this theorem any two sequences that are close enough to each other must agree in their first few terms. Now we can prove the continuity of $\sigma$ using the metric $d$, i.e., if for all $\epsilon > 0$ there exists $\delta > 0$ such that $d((s), (t)) < \delta$, then $d(\sigma(s), \sigma(t)) < \epsilon$.

**Proof.** Let $\epsilon > 0$. If for any two sequences $s, t \in \Sigma_3$ it holds that $d(\sigma(s), \sigma(t)) < \epsilon$, we know by the proximity theorem that the first $m + 1$ terms of $\sigma(s)$ and $\sigma(t)$ agree only for $\frac{1}{2^m} < \epsilon$. This means that before we had applied the shift map $\sigma$, $s$ and $t$ had $m + 2$ terms that agreed. Choose $\delta < \frac{1}{2^{m+1}}$. This way we make sure that every two sequences in $\Sigma_3$ that are at $\delta$-distance apart from each other agree in their first $m + 2$ terms. Hence by applying $\sigma$ we get points that are at a distance less than $\frac{1}{2^m}$ apart. Finally we choose an $m$ such that $\frac{1}{2^m} < \epsilon$. \qed

Now we want to see why the shift map $\sigma$ defines a chaotic dynamical system. Therefore, we need to check that the shift map satisfies the three conditions in definition 10 on $\Sigma_3$ instead of $T^2$. So let us check that definition 10 applies to $\sigma$:

**Theorem 4.** $\sigma$ is chaotic on the entire $\Sigma_3$ if and only if

1. periodic points under the action of $\sigma$ are dense in $\Sigma_3$,
2. $\sigma$ is topologically transitive, and
3. $\sigma$ has sensitive dependence on initial conditions.

**Proof.** We use the proximity theorem to prove all three conditions in definition 10.

- First, we prove by construction that periodic points under the action of $\sigma$ give a dense subset $X$ of $\Sigma_3$: Take $\epsilon > 0$ and a periodic point $x = (x_0, \ldots, x_m, x_{m+1}, x_{m+2}, \ldots, x_l) \in \Sigma_3$. Choose $m$ such that $\frac{1}{2^m} < \epsilon$. We want to find another periodic point $y$ such that $d(\sigma(x), \sigma(y)) < \epsilon$. We need $y$ to agree with $x$ in the first $m + 1$ terms. So choose $y = (x_0, \ldots, x_{m+1})$. We have found a point $y$ within an $\epsilon$-distance from $x$. This holds for all periodic points in $\Sigma_3$, hence periodic points form a dense subset $X$ of $\Sigma_3$.  

23
• Secondly, we prove by construction that \( \sigma \) is transitive:
Take \( x, y \in \Sigma_3 \). We want to find \( z \in \Sigma_3 \) such that the orbit of \( z \) is at an \( \epsilon \)-distance from \( x \) and \( y \). Choose \( m \) such that \( \frac{1}{2^m} < \epsilon \). We construct \( z \) in such a way that all possible combinations of the 1, 2, 3 symbols representing the Markov partitions are present in \( z \), especially all combinations of 1, 2, 3 of length \( m + 1 \). This way we know there exists a \( k_1 \) and a \( k_2 \) such that the first \( m + 1 \) terms of \( \sigma^{k_1}(z) \) and \( \sigma^{k_2}(z) \) are the same as those of \( x \) and \( y \) respectively. Hence the orbit of \( z \) is within an \( \epsilon \)-distance of \( x \) and \( y \). Therefore, \( \sigma \) is transitive.

• Thirdly, we prove that \( \sigma \) on \( \Sigma_3 \) describes a dynamical system that has sensitive dependence on initial conditions:
Take \( \epsilon = 1 \). For two points \( x, y \in \Sigma_3 \) where \( x \neq y \) and \( d((x), (y)) < \frac{1}{2^m} \), we know that their first \( m + 1 \) terms are the same. Choose a \( k > m + 1 \) where we choose \( m \) such that \( \frac{1}{2^m} < \epsilon \). Then evaluate the distance between \( \sigma^k(x) \) and \( \sigma^k(y) \). We know \( x_k \neq y_k \) since \( k > m + 1 \) and \( x, y \) agree only in their first \( m + 1 \) terms.

\[
d[\sigma^k(x), \sigma^k(y)] = \sum_{i=-\infty}^{\infty} \frac{\delta(x_{i+k}, y_{i+k})}{2^{|i|}} \geq \frac{\delta(x_k, y_k)}{2^0} = \epsilon.
\]

Hence the orbits of \( x \) and \( y \) are at least at an \( \epsilon \)-distance from each other. Therefore, the dynamical system described by \( \sigma \) has sensitive dependence on initial conditions.

Indeed, the shift map \( \sigma \) defines a chaotic dynamical system which is a topologically equivalent system to the system we were originally studying. Therefore, we can define a conjugacy between \( L_A \) on \( \mathbb{T}^2 \) and \( \sigma \) on \( \Sigma_3 \):

**Definition 13.** Given two maps \( L_A : \mathbb{T}^2 \to \mathbb{T}^2 \) and \( \sigma : \Sigma_3 \to \Sigma_3 \), a homeomorphism \( \phi : \mathbb{T}^2 \to \Sigma_3 \) such that \( \phi \circ \sigma = L_A \circ \phi \) is called a topological conjugacy.

This is equivalent to the following commutative diagram:

\[
\begin{array}{ccc}
\Sigma_3 & \xrightarrow{\sigma} & \Sigma_3 \\
\phi \downarrow & & \phi \downarrow \\
\mathbb{T}^2 & \xrightarrow{L_A} & \mathbb{T}^2
\end{array}
\]
Because the torus is a closed set of equivalent classes, we know that \( \Sigma_3 \) is also a closed set through topological conjugacy. Even if \( \phi \) would be a semi-conjugacy, \( \Sigma_3 \) will still be a closed set because \( \Sigma_3 \) is restricted to the allowable sequences on the set \( \{1, 2, 3\} \). These allowable sequences are finite. Moreover,

**Proposition 7.** \( \Sigma_3 \) is invariant under the action of \( \sigma \).

**Proof.** Let \( s \in \Sigma_3 \) and \( t = \sigma(s) \). Because \( t_k = s_k + 1 \), then all the iterations of \( t \) are in fact the iterations of \( s \) omitting the very first entry at the beginning of the sequence of \( s \). Hence \( t \in \Sigma_3 \). The converse is also true, just run this proof backward. \( \square \)

Since \( \sigma \) leaves \( \Sigma_3 \) invariant and \( \Sigma_3 \) is closed and the allowable sequences are finite, We call this shift map \( \sigma \) a subshift of finite type.

Now, let us focus on the overlaps that we had ignored. We can establish the identification of the sequences of the points on the boundaries of the rectangles. Note that the boundaries of the rectangles lie on the stable and unstable sets.

**Example 4.** Choose a point \([p]\) that lies on the stable boundary of \( R_2 \cap R_3 \). Let \( S([p]) = (\ldots s_{-1}s_0s_1 \ldots) \) be its natural sequence. Therefore, \( s_0([p]) \) could be 2 or 3. Hence we have a choice to make, but we first have to look at the behavior of \([p]\) as we take forward iterations. The first forward iteration of \([p]\) by \( L_A \) lies in \( R_1 \cap R_2 \). As we continue to take forward iterations of \([p]\) we notice that the images alternatively lie in \( R_2 \cap R_3 \) and \( R_1 \cap R_2 \). Hence, if we choose \( s_0([p]) = 2 \), then \( s_1([p]) = 1 \) because our transition matrix \( B \) does not admit fixed points, i.e., points like \((\ldots s_{-1}2222 \ldots)\) are not allowed. We have \( S([p]) = (\ldots s_{-1}12121 \ldots) \) in this case. On the other hand, if we choose \( s_0([p]) = 3 \), then \( s_1([p]) \) is 2 or 1. We obtain \( S([p]) = (\ldots s_{-1}3232 \ldots) \) or \( S([p]) = (\ldots s_{-1}3131 \ldots) \). Therefore, these three representations of \([p]\) must be identified. For the points that reach the stable set after some iterations will be treated in a similar manner, i.e., points with the following sequences

\[
(\ldots s_{k-1}s_ks_{k+1}2121 \ldots) \\
(\ldots s_{k-1}s_ks_{k+1}3232 \ldots) \\
(\ldots s_{k-1}s_ks_{k+1}3131 \ldots)
\]

will also be identified. Moreover, since

\[
\sigma(\ldots 2121 \ldots) = (\ldots 1212 \ldots) \\
\sigma(\ldots 3232 \ldots) = (\ldots 2323 \ldots) \\
\sigma(\ldots 3131 \ldots) = (\ldots 1313 \ldots)
\]

it follows that the points with these sequences represent the fixed point \([0]\) and hence should be identified and regarded as identical.
Using the same reasoning we can show how to identify the sequences of points on the unstable set as well as for the points which finally reach the unstable set after some iterations. The quotient of $\Sigma_3$ obtained by making all the identifications is denote by $\tilde{\Sigma}_3$. Therefore, $\phi$ will again become a topological conjugacy between $L_A : T^2 \to T^2$ and $\sigma : \tilde{\Sigma}_3 \to \tilde{\Sigma}_3$. Hence we obtain the following commutative diagram:

\[
\begin{array}{ccc}
T^2 & \xrightarrow{L_A} & T^2 \\
\downarrow{\phi} & & \downarrow{\phi} \\
\tilde{\Sigma}_3 & \xrightarrow{\sigma} & \tilde{\Sigma}_3 \\
\end{array}
\]

On the complement of the stable and unstable sets $\phi$ describes the dynamics there completely and is well behaved. By construction, $\phi$ gives a well defined bijective map from the torus $T^2$ to $\tilde{\Sigma}_3$.  

26
6 Conclusion

During the study of hyperbolic toral automorphisms we have seen that we rather study the dynamics in $\mathbb{R}^2$, because the dynamics on $\mathbb{R}^2$ are much simpler to understand. Since the end results have to be on the torus, we established equivalent classes of all points in $\mathbb{R}^2$ which gave us a unit square. By construction, we identified the boundaries of the unit square to obtain the torus $\mathbb{T}^2$. For the dynamics on $\mathbb{R}^2$ to be transported to the torus, we then used the natural projection $\pi : \mathbb{R}^2 \to \mathbb{T}^2$ to obtain the induced hyperbolic toral automorphisms.

Once on the torus $\mathbb{T}^2$, we observed that there are some special properties that arise. These properties were initially not present on the $\mathbb{R}^2$ plane. Periodicity and the non-wandering properties are some of the main properties. We also saw that the hyperbolic toral automorphisms are chaotic on the entire torus $\mathbb{T}^2$. This means that:

- periodic points are dense in $\mathbb{T}^2$ under the action of hyperbolic automorphisms,
- hyperbolic toral automorphisms are topologically transitive, and
- hyperbolic toral automorphisms have sensitive dependence on initial conditions.

These properties make it difficult to study the dynamics on the torus $\mathbb{T}^2$. Therefore, we opted to study a topologically equivalent system using symbolic dynamics. To do this we constructed a Markov partition on the torus $\mathbb{T}^2$. This is a finite collection of rectangles that have to satisfy a few conditions:

- The collection of rectangles must cover the torus $\mathbb{T}^2$,
- The interiors of the rectangles in the collection are disjoint,
- If a point is in one rectangle and its forward (or backward) iteration is in another rectangle, the image will cut across the other rectangle in the unstable (or stable) direction.

Using these conditions, we constructed a Markov partition and its first forward iteration.

We then proceeded with symbolic dynamics. Since we had 3 rectangles in the Markov partition, we defined a map on the symbol space $\{1, 2, 3\}$ that gives the itinerary of every point on the torus as the time evolves. Using the proximity theorem, we proved that the topologically equivalent system to our original system is
also chaotic. All that was left was to couple the topologically equivalent system to the original system. Here we ran into a problem because the transition matrix that we defined does not admit fixed points, yet we knew that there is a fixed point on the torus $\mathbb{T}^2$, namely $[0]$. Moreover, we had ignored the overlaps on the boundaries of the rectangles that arose when we made iterations. To remedy this problem we made a new kind of identification so that we would not have two or many sequences that represent one and the same point. Doing this enabled us to find a way to represent the fixed point $[0]$ and the points in the overlaps. Finally, we could establish a topological conjugacy $\phi$ which couples the original system to the topologically equivalent system.

7 Acknowledgments

In Writing this thesis, the following books were mostly used: An Introduction to Chaotic Dynamical Systems by Robert L. Devaney; and Dynamical systems: Stability, Symbolic Dynamics, and Chaos by Clark Robinson. There were some other sources used for literature and they can be found in the references. I would like to thank my first supervisor Dr. Alef E. Sterk for the frequent feedbacks and extra information as it was much needed. I also would like to acknowledge Prof. Dr. Harry L. Trentelman who helped assess my final thesis. I also thank my good friends Sirin Yildiz, Rishi Yildiz and Ben Vader for helping with the final details and embellishments of the thesis.
8 Bibliography


9 Appendix

Matlab Code for the Markov partition

```matlab
clf
axis([0 1 0 1]);
hold on;
%Draw the unit square:
x1=0;
x2=1;
y1=0;
y2=1;
x = [x1, x2, x2, x1, x1];
y = [y1, y1, y2, y2, y1];
plot(x, y, 'b-', 'LineWidth', 3);
hold on;
%Draw l_(0,0):
a=sqrt(5);
s=0:1;
t=([-1+a]/2)*s;
plot(s, t, 'r-');
hold on;
%By solving [(-1+a)/2]*x=[((-1-a)/2)*x]+1 we obtained the coordinates of the intersection point of l_(0,0) and l_(0,1). i1=(1/a,(-1+a)/(2*a)). Now we can draw l_(0,1):
i1=[1/a;(-1+a)/(2*a)];
l1=[x1, i1(1)];
k1=[y2, i1(2)];
%plot(l1, k1, 'r-');
%hold on;
%plot(i1(1), i1(2), 'r.', 'markersize', 18)
hold on;
x3 = 0.5;
y8 = 0.5;
text(x3, y8, 'R_1');

%By solving [(-1+a)/2]*x=[((-1-a)/2)*x-1] we obtained the coordinates of the intersection point of l_(0,0) and l_(1,0). i2=((1+a)/(2*a),1/a). Now we can draw l_(0,1):
i2=[(1+a)/(2*a),1/a];
l2=[x2, i2(1)];
```

\textbf{By solving} \[\left\lbrack (1-a)/2 \right\rbrack x + 1 = \left\lbrack \left( (1+a)/2 \right) x - 1 \right\rbrack + 1\] we obtained the coordinates of the intersection point of \(l_x(0,1)\) and \(l_y(1,1)\). \(i_3 = \left( (1+a)/(2a), (-1+a)/a \right)\). Now we can draw \(l_x(0,1)\):

\(i_3 = \left( (1+a)/(2a), (-1+a)/a \right)\); \(l_3 = [x_2, i_3(1)]; \ k_3 = [y_2, i_3(2)];\)

\%plot \((l_3, k_3, 'r-')\)
\%hold on
\%plot \((i_3(1), i_3(2), 'b.', 'markersize', 18)\)
\%hold on

\%x4 = 0.6;
\%y9 = 0.1;
\%txt1 = 'R_2';
%text (x4, y9, txt1)

We close off the rectangles by drawing the extension of the line \(l_x(0,0)\). By first projecting the end point of \(l_x(0,0)\) onto \((0,0)\) and then drawing a line in the direction of the unstable vector until we hit \(l_x(0,1)\). So we solve \[\left[ (1+a)/2 \right] x + (1+a)/2 = \left[ (1-a)/2 \right] x + 1\] to obtain the coordinates of the intersection. \(i_4 = \left( (3-a)/(2a), (1+a)/(2a) \right)\).

\(i_4 = \left[ (3-a)/(2a), (1+a)/(2a) \right]; \ l_4 = [x_1, i_4(1)]; \ k_4 = [(-1+a)/2, i_4(2)];\)

\%plot \((l_4, k_4, 'r-')\)
\%hold on
\%plot \((i_4(1), i_4(2), 'g.', 'markersize', 18)\)
\%hold on

%**************************************************************************
% we compute where the intersection points will go after applying L_A on them.
A=[1 1;1 0];
L1=mod(A*i1,1);
L2=mod(A*i2,1);
L3=mod(A*i3,1);
L4=mod(A*i4,1);
plot(L1(1),L1(2),'r.','markersize',18)
hold on
plot(L2(1),L2(2),'b.','markersize',18)
hold on
plot(L3(1),L3(2),'y.','markersize',18)
hold on
plot(L4(1),L4(2),'g.','markersize',18)
hold on
%
%take the following points from R1:
%plot(0.45,0.3,'k*','markersize',6)
%hold on
%plot(0.1,0.8,'k+','markersize',6)
%hold on
%plot(0.9,0.9,'kx','markersize',6)
%hold on
%plot(0.9,0.6,'k.','markersize',10)
%hold on
%apply L_A on them to obtain:
%plot(0.75,0.45,'k*','markersize',6)
%hold on
%plot(0.9,0.1,'k+','markersize',6)
%hold on
%plot(0.8,0.9,'kx','markersize',6)
%hold on
%plot(0.5,0.9,'k.','markersize',10)
%hold on
%take the following points from R2:
%plot (0.75, 0.3, 'm.', 'markersize', 10)
%hold on
%plot (0.1, 0.95, 'm*', 'markersize', 6)
%hold on
%plot (0.3, 0.6, 'mx', 'markersize', 6)
%hold on
%apply L_A on them to obtain:
%plot (0.05, 0.75, 'm.', 'markersize', 10)
%hold on
%plot (0.05, 0.1, 'm*', 'markersize', 6)
%hold on
%plot (0.9, 0.3, 'mx', 'markersize', 6)
%hold on

%take the following points from R3:
%plot (0.2, 0.6, 'c.', 'markersize', 10)
%hold on
%plot (0.4, 0.3, 'c*', 'markersize', 6)
%hold on
%plot (0.8, 0.4, 'cx', 'markersize', 6)
%hold on
%apply L_A on them to obtain:
%plot (0.8, 0.2, 'c.', 'markersize', 10)
%hold on
%plot (0.7, 0.4, 'c*', 'markersize', 6)
%hold on
%plot (0.2, 0.8, 'cx', 'markersize', 6)
%hold on

%Using the images of the intersection points and the images of the 3 points of each rectangle we can plot the first iteration of the rectangles:

n1=[x1,(3-a)/(-1+a)];
n2=[(-1+a)/2,y2];
plot (n1,n2, 'r-')
hold on;
n3=[L4(1),(3-a)/(-1+a)];
n4 = [L4(2), y1];
plot(n3, n4, 'r-')
hold on;
n5 = [L3(1), x2];
n6 = [L3(2), (3-a)/2];
plot(n5, n6, 'r-')
hold on;
n7 = [x1, x2];
n8 = [(3-a)/2, y2];
plot(n7, n8, 'r-')
hold on;
n9 = [x1, L2(1)];
n0 = [y2, L2(2)];
plot(n9, n0, 'r-')
hold on;
% here are the images of the rectangle:
m0 = 0.5;
m1 = 0.8;
txt1 = 'R_1';
text(m0, m1, txt1)
hold on;
x3 = 0.5;
y8 = 0.5;
txt1 = 'R_2';
text(x3, y8, txt1)
hold on;
x3 = 0.6;
y8 = 0.1;
txt1 = 'R_3';
text(x3, y8, txt1)