3. Dynamical systems

A *topological dynamical system* is an action of a semigroup G on a topological space $\mathcal{X}$ by continuous maps. Similarly, dynamical systems in other categories (e.g., on spaces with a measure, or on a manifold) are defined.

Example 3.1. If G is the additive semigroup $\mathbb{N}$, then the dynamical system is determined by the action of the generator. Thus such dynamical systems study iterations of a single continuous map $f : \mathcal{X} \longrightarrow \mathcal{X}$.

Example 3.2. If G is the infinite cyclic group $\mathbb{Z}$, then the dynamical system studies iterations (both forward and backward) of a homeomorphism $f : \mathcal{X} \longrightarrow \mathcal{X}$.

Actions of $(\mathbb{R}, +)$ are called *flows*. If it is a smooth action on a manifold, then it is uniquely determined by the corresponding vector field of velocities, and therefore study of such flows is equivalent to the study of (autonomous) ODE.

3.1. Full shift and subshifts.

Let G be a countable semigroup, and let A be a finite set, $|A| \geq 2$. Consider the direct product space A^G of all maps $\phi : G \longrightarrow A$. Its topology is given by the basis of open sets of the form

$$U_f = \{\phi : G \longrightarrow A \ : \ \phi|_D = f\},$$

where $f : D \longrightarrow A$ is a map defined on a finite subset $D \subset G$. Note that U_f are also closed. The space A^G is homeomorphic to the Cantor set.

For example, if $G = \mathbb{N}$, then A^G is the space of all sequences $a_1 a_2 \ldots$ of elements of A.

The semigroup G acts on A^G by the rule:

$$(g \cdot \phi)(x) = \phi(xg).$$

(For the right action the rule is $(\phi \cdot g)(x) = \phi(gx)$.)

Example 3.3. For $A^{\mathbb{Z}}$ the action of the generator s of $\mathbb{Z}$ is given by

$$s(a_1 a_2 \ldots) = a_2 a_3 \ldots .$$

Definition 3.1. The described dynamical system is called the *full shift* on the semigroup G over the alphabet A. Its closed G-invariant subsets are called *subshifts*.

Any closed subset of A^G is given by representing its complement as a union of the sets of the form U_f. In other words, it is given by a list of "prohibited" restrictions $\{f_i : D_i \longrightarrow G\}_{i \in I}$, so that a function ϕ belongs to the set if and only if $\phi|_{D_i} \neq f_i$ for every $i \in I$.

Therefore, a closed G-invariant subset $\mathcal{S}$ of A^G can be defined by a list of prohibited restrictions $\{f_i : D_i \longrightarrow G\}_{i \in I}$ so that $\phi \in \mathcal{S}$ if and only if $\phi|_{D_i \cdot g} \neq g \cdot f_i$ for all $i \in I$ and $g \in G$. We call the set $\{f_i\}_{i \in I}$ the set of *forbidden patterns* defining the subshift $\mathcal{S}$.

For example, a $\mathbb{Z}$-invariant closed subset $\mathcal{S}$ of $A^{\mathbb{Z}}$ is defined by a set P of finite words such that an infinite sequence w belongs to $\mathcal{S}$ if and only if no subword of w belongs to P.

Definition 3.2. A *shift of finite type* is a subshift of A^G defined by a finite set of forbidden patterns.

PROBLEM 3.1. Consider the shift of finite type $\mathcal{F} \subset \{0,1\}^{\mathbb{N}}$ of all sequences $a_1a_2\ldots$ such that $a_ia_{i+1} \neq 11$ for every i. Let u_n be the number of possible beginnings of length n of elements of $\mathcal{F}$. Find $\lim_{n\to\infty} \sqrt[n]{u_n}$.

3.2. Subshifts of $A^{\mathbb{Z}}$. The points of the space $A^{\mathbb{Z}}$ are usually represented by bi-infinite sequences $\ldots w(-1)w(0)w(1)\ldots$. The generator of the $\mathbb{Z}$-shift acts on it by the rule

$$s(w)(n) = w(n+1),$$

i.e., by shifting the sequence to the left.

A sub-shift is, by definition, a closed $\mathbb{Z}$-invariant subset of $A^{\mathbb{Z}}$. Every sub-shift can be defined by a set of *prohibited words* $P \subset A^*$, where A^* denotes the set of all finite words over the alphabet A.

Namely, the subshift defined by P is the set $\mathcal{X}_P$ of sequences $w \in A^{\mathbb{Z}}$ such that $w(i)w(i+1)\ldots w(i+k) \notin P$ for all $i \in \mathbb{Z}$ and $k \geq 0$.

The *language* of a subshift $\mathcal{X}$ is the set $\mathcal{L}_{\mathcal{X}} \subset A^*$ of all finite subwords of elements of $\mathcal{X}$. *Complexity* of the subshift is the function

$$p_{\mathcal{X}}(n) = |\{v \in A^n \ : \ v \in \mathcal{L}_{\mathcal{X}}\}.$$

PROPOSITION 3.1. *If $p_{\mathcal{X}}(n) = p_{\mathcal{X}}(n+1)$, then $p_{\mathcal{X}}(m) = p_{\mathcal{X}}(n)$ for all $m \geq n$.*

PROOF. If $p_{\mathcal{X}}(n) = p_{\mathcal{X}}(n+1)$, then for every word $v \in \mathcal{L}_{\mathcal{X}}$ of length n there exists only one letter $x_1 \in A$ such that $vx_1 \in \mathcal{L}_{\mathcal{X}}$. Considering the suffix of length n of vx_1, we conclude that there exists only one letter x_2 such that $vx_1x_2 \in \mathcal{L}_{\mathcal{X}}$. Repeating this argument $m-n$ times we get that v can be extended to a word of length m belonging to $\mathcal{L}_{\mathcal{X}}$ in a unique way, which implies that $p_{\mathcal{X}}(m) = p_{\mathcal{X}}(n)$. □

COROLLARY 3.2. *If $\mathcal{X}$ is infinite, then $p_{\mathcal{X}}(n) \geq n+1$ for all n.*

PROOF. If $p_{\mathcal{X}}(n)$ is bounded, then it is eventually constant (as $p_{\mathcal{X}}(n)$ is always non-decreasing). Then sufficiently long finite words belonging to $\mathcal{L}_{\mathcal{X}}$ will have unique finite extensions to both sides, hence an infinite sequence $w \in \mathcal{X}$ can be uniquely determined by a sufficiently long finite subword. This implies that $\mathcal{X}$ is finite if $p_{\mathcal{X}}(n)$ is bounded.

If $p_{\mathcal{X}}(n)$ is unbounded, then by the last proposition $p_{\mathcal{X}}(n+1) > p(n)$ for every n. We have $p_{\mathcal{X}}(1) \geq 2$, since otherwise $\mathcal{X}$ has only one possible element. Consequently, $p_{\mathcal{X}}(n) \geq n+1$ for all n. □

3.3. Example: Sturmian subshifts. Let us describe a class of subshifst with the lowest possible complexity $p(n) = n+1$.

Let $\theta \in (0,1)$ be an irrational number. Consider the corresponding rotation $R_\theta : x \mapsto x + \theta \pmod 1$ of the circle $\mathbb{R}/\mathbb{Z}$. Divide the circle into two arcs: $[0,\theta]$ and $[\theta, 1]$. For $x \in \mathbb{R}/\mathbb{Z}$ not belonging to the orbit $\{n\theta\}_{n\in\mathbb{Z}}$ of 0, define its *itinerary* as the sequence $I_\theta(x) \in \{0,1\}^{\mathbb{Z}}$ given by

$$I_\theta(x)(n) = \begin{cases} 0 & \text{if } R_\theta^n(x) \in [0,\theta]; \\ 1 & \text{if } R_\theta^n(x) \in [\theta, 1]. \end{cases}$$

If x belongs to the R_θ-orbit of 0, then there exists n such that $x + n\theta = 0$, and we have an ambiguity for the two coordinates $I_\theta(x)(n)$ and $I_\theta(x)(n+1)$ of $I_\theta(x)$. We define therefore two itineraries of x: $I_\theta(x_{+0})$ with $I_\theta(x)(n)I_\theta(x)(n+1) = 01$ and $I_\theta(x_{-0})$ with $I_\theta(x)(n)I_\theta(x)(n+1) = 10$. The remaining coordinates of $I_\theta(x_{+0})$ and $I_\theta(x_{-0})$ are defined by the original rule, since there is no ambiguity for them.

Denote by $\mathcal{X}_\theta$ the set of all itineraries $I_\theta(x)$ and $I_\theta(x_{\pm 0})$ for a given θ.

Let S_θ be the set $[0,1]$ in which every element $x \in (0,1)$ belonging to the R_θ-orbit of 0 is replaced by two copies: x_{+0} and x_{-0}. We identify 0_{+0} and 0_{-0} with the endpoints 0 and 1, respectively. The set S_θ has a natural ordering coming from the usual order of real numbers and the natural agreement $x_{-0} < x_{+0}$. Consider the order topology on S_θ, i.e., the topology with the basis of open sets equal to the set of all open intervals $(a,b) = \{x \in S_\theta \ : \ a < x < b\}$. Note that for $x, y \in \{R_\theta^n(0)\}_{n\in\mathbb{Z}}$, $x < y$, the open interval (x_{-0}, y_{+0}) is equal to the interval $[x_{+0}, y_{-0}]$, and since the set $\{R_\theta^n(0)\}_{n\in\mathbb{Z}}$ is dense in $\mathbb{R}/\mathbb{Z}$, the set of intervals of the form $[x_{+0}, y_{-0}]$ is a basis of topology. Such intervals are also closed, since their complements are equal to $[0, x_{+0}) \cup (y_{-0}, 1]$.

The rotation R_θ is naturally lifted to a homeomorphism of S_θ by the rule $R_\theta(x_{+0}) = R_\theta(x)_{+0}$ and $R_\theta(x_{-0}) = R_\theta(x)_{-0}$. Then the itinerary $I_\theta(x)$ for $x \in S_\theta$ is non-ambiguously given by the rule

$$I_\theta(x)(n) = \begin{cases} 0 & \text{if } R_\theta^n(x) \in [0, \theta_{-0}]; \\ 1 & \text{if } R_\theta^n(x) \in [\theta_{+0}, 1]. \end{cases}$$

It follows directly from the definitions that the set $\mathcal{X}_\theta$ of all itineraries $I_\theta(x)$, $x \in S_\theta$ is shift-invariant, since $s(I_\theta(x)) = I_\theta(R_\theta(x))$.

For every $n \in \mathbb{Z}$ the points $n\theta$ and $(n+1)\theta$ subdivide the circle $\mathbb{R}/\mathbb{Z}$ into two arcs with ends $n\theta_{\pm 0}$ and $(n+1)\theta_{\pm 0}$. If we know to which of these arcs the point x belongs, we know the nth letter $I_\theta(x)(n)$ of the itinerary. It follows that in order to know the subword $I_\theta(x)(0)I_\theta(x)(1)\ldots I_\theta(x)(n-1)$ of the itinerary, it is enough to know which of the arcs into which the points $0, \theta, 2\theta, \ldots, n\theta$ divide the circle x belongs to. This implies that complexity $p_{\mathcal{X}_\theta}(n)$ is equal to the number of these arcs, i.e., to $n+1$.

Proposition 3.3. *The map $I_\theta : S_\theta \longrightarrow \mathcal{X}_\theta$ is a homeomorphism.*

Proof. The arguments of the paragraph preceding the proposition show that for every cylindrical set $U_v = \{w \in \mathcal{X}_\theta \ : \ w(0)w(1)\ldots w(n-1) = v\}$ the set $I_\theta^{-1}(U_v)$ is either empty or one of the $n+1$ arcs into which the points $0, \theta, 2\theta, \ldots, n\theta$ subdivide the circle, i.e., it is of the form $[i\theta_{+0}, j\theta_{-0}]$, where $i, j \in \{0, 1, \ldots, n\}$ are such that there is no $k \in \{0, 1, \ldots, n\} \setminus \{i, j\}$ such that $k\theta_{\pm 0}$ are between $i\theta_{+0}$ and $j\theta_{-0}$. This proves that I_θ is continuous. It easily follows from the fact that the orbit of 0 under R_θ is dense that I_θ is one-to-one. Since the set of arcs $I_\theta^{-1}(U_v)$ is a basis of topology of S_θ, the inverse map $I_\theta^{-1} : \mathcal{X}_\theta \longrightarrow S_\theta$ is also continuous. $\square$

In particular, $\mathcal{X}_\theta$ is compact, hence closed in $\{0,1\}^{\mathbb{Z}}$.

3.4. Markovian shifts. If a shift of finite type $\mathcal{X} \subset A^{\mathbb{Z}}$ is defined by a set of prohibited words P such that $P \subset A^2$, then it is called a *topological Markov shift*. A topological Markov shift can be defined by the complement $A^2 \setminus P$ of the set of prohibited words, which we will call the *set of allowed transitions*.

Every shift of finite type can be realized as a topological Markov shift in the following way.

Definition 3.3. Let $f : \mathcal{X} \longrightarrow \mathcal{X}$ and $g : \mathcal{Y} \longrightarrow \mathcal{Y}$ be continuous maps (seen as dynamical systems). A *semiconjugacy* $\phi : \mathcal{X} \longrightarrow \mathcal{Y}$ is a continuous map such that the diagram

$$\begin{array}{ccc} \mathcal{X} & \xrightarrow{f} & \mathcal{X} \\ \downarrow{\scriptstyle\phi} & & \downarrow{\scriptstyle\phi} \\ \mathcal{Y} & \xrightarrow{g} & \mathcal{Y} \end{array}$$

is commutative. In general, if $\mathcal{X}$ and $\mathcal{Y}$ are topological spaces with actions of a semigroup G, then $\phi : \mathcal{X} \longrightarrow \mathcal{Y}$ is a semi-conjugacy if for all $g \in G$ and $x \in \mathcal{X}$ we have $\phi(g \cdot x) = g \cdot \phi(x)$.

If a semiconjugacy ϕ is a homeomorphism, then it is called a *conjugacy*. Two dynamical systems are said to be *topologically conjugate* if there exists a conjugacy between them.

Let $\mathcal{X} \subset A_1^{\mathbb{Z}}$ and $\mathcal{Y} \subset A_2^{\mathbb{Z}}$ be subshifts, and suppose that $\phi : \mathcal{X} \longrightarrow \mathcal{Y}$ is a semi-conjugacy.

For every $x \in A_2$ the set $\phi^{-1}(\{w \in \mathcal{Y} \ : \ w(0) = x\})$ is open in $\mathcal{X}$, hence there exists $k > 0$ and a finite set $B \subset A_1^{2k+1}$ such that

$$\phi^{-1}(\{w \in \mathcal{Y} \ : \ w(0) = x\}) = \bigcup_{v \in B} \{w \in \mathcal{X} \ : \ w(-k)w(-k+1)\ldots w(k-1)w(k) = v\}.$$

It follows that there exists k and a map $\psi : A_1^{2k+1} \longrightarrow A_2$ such that for every $w \in \mathcal{X}$ we have

$$\phi(w)(0) = \psi(w(-k)w(-k+1)\ldots w(k-1)w(k)).$$

Since ϕ is a semi-conjugacy, this implies that for every $n \in \mathbb{Z}$

$$\phi(w)(n) = \psi(w(n-k)w(n-k+1)\ldots w(n+k-1)w(n+k)).$$

In other words, in order to know the nth coordinate of $\phi(w)$, one has to look through the window $[n-k, \ldots, n+k]$ at w and apply the same rule for all n. Such maps are called *block codes*.

As a consequence of the above arguments, we get the following relation between complexities of semi-conjugated subshifts.

LEMMA 3.4. *Suppose that $\phi : \mathcal{X} \longrightarrow \mathcal{Y}$ is a surjective semi-conjugacy of subshifts. Then there exists k such that $p_{\mathcal{Y}}(n) \le p_{\mathcal{X}}(n+k)$ for all n.*

As an example, consider the "identity" block code:

$$\kappa_k : A^{\mathbb{Z}} \longrightarrow (A^{2k+1})^{\mathbb{Z}}$$

given by

$$\kappa_k(w)(n) = w(n-k)w(n-k+1)\ldots w(n+k+1)w(n+k).$$

The block code $\kappa_k^* : (A^{2k+1})^{\mathbb{Z}} \longrightarrow A^{\mathbb{Z}}$ given by

$$\kappa_k^*(w)(n) = a_n, \text{ if } w(n) = a_{n-k}a_{n-k+1}\ldots a_{n+k-1}a_{n+k}$$

obviously satisfies $\kappa_k^* \circ \kappa_k = Id$. It follows that κ_k is a homeomorphism from $A^{\mathbb{Z}}$ to $\kappa_k(A^{\mathbb{Z}})$.

It is easy to see that $\kappa_k(A^{\mathbb{Z}})$ is a Markov subshift of $(A^{2k+1})^{\mathbb{Z}}$ defined by the set of allowed transitions $\{(xv)(vy) \ : \ v \in A^{2k}, x, y \in A\}$.

The following is straightforward.

Proposition 3.5. *Let* $\mathcal{X} \subset A^{\mathbb{Z}}$ *be a shift of finite type defined by a finite set* P *of prohibited words. If* k *is such that* $2k+1$ *is larger than the length of every element of* P*, then* $\kappa_k(\mathcal{X})$ *is a topological Markov shift. In particular, every shift of finite type is topologically conjugate to a topological Markov shift.*

Let $\mathcal{X} \subset A^{\mathbb{Z}}$ be a topological Markov shift. Consider the graph Γ with the set of vertices A, in which we have an arrow from x to y if and only if xy belongs to the set of allowed transitions. Then $\mathcal{X}$ coincides with the set of all bi-infinite oriented paths in Γ (seen as sequences of vertices).

Proposition 3.6. *Let* M *be the adjacency matrix of* Γ*. Then* $p_{\mathcal{X}}(n)$ *is equal to the sum of entries of* M^n*. In particular,* $p_{\mathcal{X}}(n)$ *is a linear combination of exponential functions multiplied by polynomials, and has either exponential or polynomial growth rate.*

If $p_{\mathcal{X}}(n)$ has polynomial growth, then all simple cycles (i.e., cycles without self-intersections) in Γ are disjoint. The cycles of Γ can be connected in a "cascade", i.e., we have a partial order on the cycles, where a cycle C_1 is "below" a cycle C_2 if there is an oriented path from a vertex of C_2 to a vertex of C_1. Define *depth* of Γ as the largest size of a chain in this partially ordered set of cycles.

Proposition 3.7. *If complexity of* $\mathcal{X}$ *is polynomial, then the degree of polynomial growth of* $p_{\mathcal{X}}$ *is equal to the depth of* Γ *minus one. The space* $\mathcal{X}$ *is in this case countable, and Cantor-Bendixson rank of* $\mathcal{X}$ *is equal to the depth of* Γ*.*

Here Cantor-Bendixson rank of $\mathcal{X}$ is defined in the following way. Denote $\mathcal{X}^{(0)} = \mathcal{X}$, and define inductively $\mathcal{X}^{(\alpha+1)}$ as the set of limit points (i.e., non-isolated points) of $\mathcal{X}^{(\alpha)}$. Define for limit ordinals λ the set $\mathcal{X}^{(\lambda)}$ as the intersection of all sets $\mathcal{X}^{(\alpha)}$ for $\alpha < \lambda$. Then *Cantor-Bendixson rank* of $\mathcal{X}$ is the smallest ordinal α such that $\mathcal{X}^{(\alpha+1)} = \mathcal{X}^{(\alpha)}$. If $\mathcal{X}$ is a countable Hausdorff compact space, then we have $\mathcal{X}^{(\alpha)} = \emptyset$ for α equal to the Cantor-Bendixson rank.

For example, if depth of Γ is 1, then there are no cycles connected by an oriented path, and $\mathcal{X}$ is a finite set.

We see, for example, that Sturmian shifts are not of finite type, since their complexity is $n+1$, but they are uncountable.

3.5. Angle doubling map. Consider the circle $\mathbb{R}/\mathbb{Z}$. Define $f : \mathbb{R}/\mathbb{Z} \longrightarrow \mathbb{R}/\mathbb{Z}$ by $f(x) = 2x$. We call it *angle doubling map*. It is a degree 2 covering map.

Definition 3.4. For a dynamical system $F : \mathcal{X} \longrightarrow \mathcal{X}$ a point $x \in \mathcal{X}$ is called *periodic* if there exists n such that $F^n(x) = x$. It is called *pre-periodic* if it is not periodic, but there exists k such that $F^k(x)$ is periodic.

A point has finite (forward) orbit if and only if it is either periodic or pre-periodic. If F is one-to-one, then there are no pre-periodic points.

The following is straightforward.

Proposition 3.8. *A point* $x \in \mathbb{R}/\mathbb{Z}$ *has finite orbit for the angle doubling map if and only if it is rational. If it has odd denominator (when written as a fraction in lowest terms), then it is periodic. Otherwise it is pre-periodic. In particular the sets of periodic and of pre-periodic points are dense in* $\mathbb{R}/\mathbb{Z}$*.*

Consider the map $\phi : \{0,1\}^{\mathbb{N}} \longrightarrow \mathbb{R}/\mathbb{Z}$ given by

$$\phi(a_1a_2\ldots) = \sum_{n\geq 1}^{\infty} \frac{a_n}{2^n}.$$

PROPOSITION 3.9. *The map ϕ is surjective. For $x \in \mathbb{R}/\mathbb{Z}$ the set $\phi^{-1}(x)$ consists either of one or two elements. In the latter case x is of the form $\frac{m}{2^k}$ for $m, k \in \{0,1,2,\ldots\}$, and $\phi^{-1}(x)$ consists of sequences of the form $a_1a_2\ldots a_{k-1}01111\ldots$ and $a_1a_2\ldots a_{k-1}10000\ldots$, or of the form $0000\ldots$ and $1111\ldots$ (if $x = 0$).*

DEFINITION 3.5. Let $\phi : \mathcal{X} \longrightarrow \mathcal{Y}$ be a semi-conjugacy between two spaces on which a semigroup G acts. Its *kernel* is the space $\mathcal{E}_\phi = \{(x_1, x_2) \in \mathcal{X}^2 \ : \ \phi(x_1) = \phi(x_2)\}$ together with the action $g \cdot (x_1, x_2) = (g \cdot x_1, g \cdot x_2)$.

It follows directly from the definition of a semi-conjugacy that the space $\mathcal{E}_\phi$ is G-invariant. Note also that it is closed. In particular, if the space $\mathcal{X}$ is compact, then $\mathcal{E}_\phi$ is also compact.

We have a natural identification (a conjugacy) of the direct product $A^G \times A^G$ with the shift $(A^2)^G$. Namely, any pair of functions $(f_1, f_2) \in A^G \times A^G$ can be written as one function $(f_1, f_2) : G \longrightarrow A^2$ by the rule $(f_1, f_2)(g) = (f_1(g), f_2(g))$. It follows that if $\phi : \mathcal{X} \longrightarrow \mathcal{Y}$ is a semi-conjugacy such that $\mathcal{X}$ is a subshift of A^G, then $\mathcal{E}_\phi$ is a subshift of $(A^2)^G$.

DEFINITION 3.6. A topological dynamical system $(\mathcal{Y}, G)$ is *finitely presented* if there exists a shift of finite type $\mathcal{X} \subset A^G$ and a surjective sem-conjugacy $\phi : \mathcal{X} \longrightarrow \mathcal{Y}$ such that $E_\phi \subset (A^2)^G$ is a shift of finite type.

PROPOSITION 3.10. *Let $\phi : \{0,1\}^{\mathbb{N}} \longrightarrow \mathbb{R}/\mathbb{Z}$ be the semi-conjugacy $\phi(a_1a_2\ldots) = \sum_{n=1}^{\infty} \frac{a_n}{2^n}$ of the shift with the angle doubling map. Then $\mathcal{E}_\phi$ is a shift of finite type.*

PROOF. It is easy to check that $\mathcal{E}_\phi$ is defined by the following set of 12 prohibited sub-words. Here we write elements of $\{0,1\}^2$ as columns.

$$\begin{array}{cc}
\begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}x\\x\end{pmatrix}, & \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}x\\x\end{pmatrix}, \\
\begin{pmatrix}x\\x\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}, & \begin{pmatrix}x\\x\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}, \\
\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}, & \begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}, \\
\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}, & \begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix},
\end{array}$$

where $x = 0, 1$. □

3.6. Smale's solenoid. Let $f : \mathcal{X} \longrightarrow \mathcal{X}$ be a continuous map. Consider the sequence

$$\mathcal{X} \xleftarrow{f} \mathcal{X} \xleftarrow{f} \mathcal{X} \xleftarrow{f} \cdots$$

and let $\widehat{\mathcal{X}}$ be its inverse limit, i.e., the subspace of $\mathcal{X}^{\mathbb{Z}}$ consisting of sequences $(x_1, x_2, \ldots)$ such that $f(x_{n+1}) = x_n$. Then the map $\hat{f} : \widehat{\mathcal{X}} \longrightarrow \widehat{\mathcal{X}}$ given by $\hat{f}(x_1, x_2, \ldots) = (f(x_1), f(x_2), \ldots)$ is a homeomorphism. Its inverse is the map $(x_1, x_2, \ldots) \mapsto (x_2, x_3, \ldots)$. The dynamical system $(\widehat{\mathcal{X}}, \hat{f})$ is called the *natural extension* of the dynamical system $(\mathcal{X}, f)$.

The natural extension of the angle doubling map is called the *Smale's solenoid*.

If $(x_1, x_2, \dots) \in (\mathbb{R}/\mathbb{Z})^{\mathbb{N}}$ is a point of the solenoid, then the binary representations of x_i are of the form

$$(.a_1a_2\dots, .a_0a_1a_2\dots, .a_{-1}a_0a_1a_2\dots, \dots).$$

It is natural, therefore, to represent the point $(x_1, x_2, \dots)$ by the bi-infinite binary number

$$\dots a_{-1}a_0.a_1a_2\dots.$$

The natural extension of the map $x \mapsto 2x$ is then represented by the shift

$$\dots a_{-1}a_0.a_1a_2\dots \mapsto \dots a_0a_1.a_2a_3\dots.$$

It also follows directly from the properties of the binary numeration system on $\mathbb{R}/\mathbb{Z}$ that two different sequences represent the same point of the solenoid if and only if they are of the form $\dots a_{n-1}a_n011111\dots$ and $\dots a_{n-1}a_n100000\dots$, or they are $\dots 0000000\dots$ and $\dots 1111111\dots$. We see that the set of pairs of sequences representing the same point of the solenoid is a two-sided shift of finite type (over the alphabet $\{0,1\}^2$) given by the same set of prohibited words as the one-sided shift in Proposition 3.10.

Recall that the *ring of dyadic integers* $\mathbb{Z}_2$ is the completion of $\mathbb{Z}$ with respect to the norm $\|n\| = 2^k$, where k is the non-negative integer such that n is divisible by 2^k. It is the ring of formal infinite series of the form

$$\sum_{k=0}^{\infty} a_k 2^k, \qquad a_k \in \{0,1\}$$

with natural operations (coming from the arithmetic rules of adding and multiplying integers in the binary numeration system). The map $\sum_{k=0}^{\infty} a_k 2^k \mapsto (a_0, a_1, \dots)$ is a homeomorphism of $\mathbb{Z}_2$ with the Cantor set $\{0,1\}^{\mathbb{N}}$.

It follows from the described representation of the points of solenoid by binary sequences that the solenoid is naturally homeomorphic with the quotient of $\mathbb{Z}_2 \times \mathbb{R}$ by the action $(x, y) \mapsto (x - n, y + n)$ of $\mathbb{Z}$. In other words, it is the quotient of the group $\mathbb{Z}_2 \oplus \mathbb{R}$ by the subgroup of all elements of the form $(-n, n)$ for $n \in \mathbb{Z}$. Note that this shows that the solenoid is an abelian group, which also follows from the fact that it is the inverse limit of the groups $\mathbb{R}/\mathbb{Z}$ with respect to the group homomorphism $x \mapsto 2x$.

Consider $D = \mathbb{Z}^2 \times [0,1] \subset \mathbb{Z}^2 \times \mathbb{R}$. This set is a fundamental region for the $\mathbb{Z}$-action: the images of D under the action of $\mathbb{Z}$ cover the whole space $\mathbb{Z}^2 \times \mathbb{R}$, and have disjoint interiors. It follows that the solenoid is obtained from $\mathbb{Z}^2 \times [0,1]$ by making identifications

$$(x, 1) \mapsto (x + 1, 0).$$

We have proved the following description of the solenoid.

Proposition 3.11. *The solenoid is homeomorphic to the mapping torus of the map $x \mapsto x + 1$ on $\mathbb{Z}_2$. It is connected, but has uncountably many path connected components. Each path connected component is dense.*

The statement about path connected components follows from the fact that the map $x \mapsto x + 1$ on $\mathbb{Z}_2$ is a minimal dynamical system (which in turn follows just from the fact that $\mathbb{Z}$ is dense in $\mathbb{Z}_2$). Connectivity of the solenoid follows from density of its path connected components, or from the fact that it is inverse limit of connected spaces.

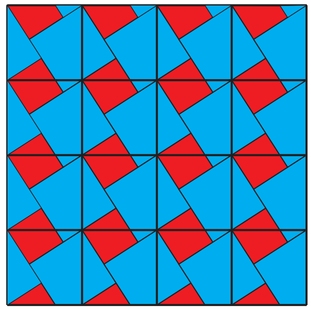

FIGURE 5. A Markov partition

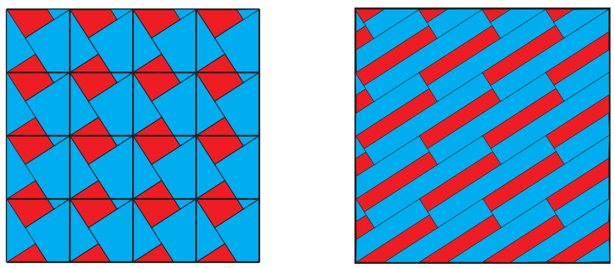

FIGURE 6. Image of the partition

The original Smale's construction of the solenoid realizes it as an attractor of a diffeomorphism of an open subset of $\mathbb{R}^3$, as follows.

Consider a torus in $\mathbb{R}^3$ (say, obtained by rotating the circle $(x-2)^2 + y^2 = 1$ around the z-axis). Let U be the open region inside the torus. Consider a "skinny" torus U_1 inside U winding twice, as it is shown on Figure... Let f be a diffeomorphism $f : U_1 \longrightarrow U$ which is expanding in the longitudinal direction of the torus, and contracting in the planes of the rotated circle. Denote $U_n = f^n(U)$. Then $U_{n+1} \subset U_n$ and $S = \bigcap_{n\geq 1} U_n$ is homeomorphic to the solenoid. Moreover, the action of f on S is topologically conjugate to the natural extension of the angle doubling map.

3.7. Arnold's Cat Map. Consider the map $f : \mathbb{R}^2/\mathbb{Z}^2 \longrightarrow \mathbb{R}^2/\mathbb{Z}^2$ induced by the linear transformation of $\mathbb{R}^2$ with the matrix $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$. Since $\det A = 1$, the map f is a homeomorphism (and an automorphism of the group $\mathbb{R}^2/\mathbb{Z}^2$).

The eigenvalues of A are satisfy $\lambda^2 - 3\lambda + 1 = 0$, hence they are equal to $\lambda = \frac{3+\sqrt{5}}{2}$ and $\lambda^{-1} = \frac{3-\sqrt{5}}{2}$. Note that $\lambda > 1$ and $0 < \lambda^{-1} < 1$. The eigenvectors of A are $\begin{pmatrix} 1 \\ \frac{-1+\sqrt{5}}{2} \end{pmatrix}$ and $\begin{pmatrix} \frac{1-\sqrt{5}}{2} \\ 1 \end{pmatrix}$. Note that they are orthogonal (since A is symmetric).

Consider the following partition (see Figure 5) of the torus $\mathbb{R}^2/\mathbb{Z}^2$ into two squares (red and blue) with sides parallel to the eigenvectors of A:

Their images under the action of A are shown on Figures 6 and 7.

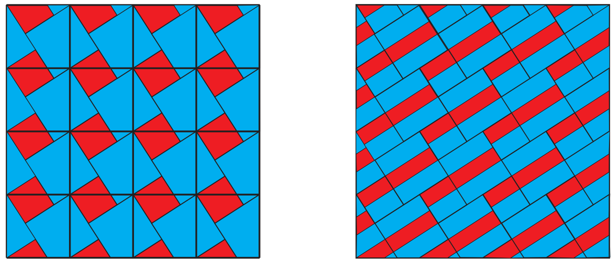

FIGURE 7. Image of the partition

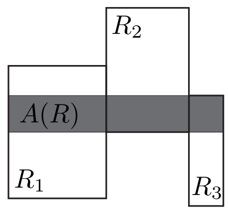

FIGURE 8. Markov partition

If $R \subset \mathbb{R}^2$ is a rectangle with sides parallel to the eigenspaces of A such that the quotient map $\mathbb{R}^2 \longrightarrow \mathbb{R}^2/\mathbb{Z}^2$ restricted to R is injective, then we also the image of R in the torus an A-rectangle. If R is an A-rectangle, and $x \in R$, then we denote by $W_s(x, R)$ the maximal segment inside R containing x and parallel to the contracting eigenspace (i.e., to the eigenspace of the eigenvalue $\lambda < 1$), and by $W_u(x, R)$ we denote the maximal segment inside R containing x and parallel to the expanding eigenspace.

DEFINITION 3.7. Let $\mathcal{R}$ be a finite set of A-rectangles $R \subset \mathbb{R}^2/\mathbb{Z}^2$ satisfying the following conditions:

(1) The interiors of the rectangles $R \in \mathcal{R}$ are disjoint, and union of their closures is the whole torus.
(2) If x belongs to the interior of $R_1 \in \mathcal{R}$ and $f(x)$ belongs to the interior of $R_2 \in \mathcal{R}$, then
$$f(W_s(x, R_1)) \subset W_s(f(x), R_2)$$
and
$$W_u(f(x), R_2) \subset f(W_u(x, R_1)).$$

Then the set $\mathcal{R}$ is called a *Markov partition*.

For example, the set $\mathcal{R}$ consisting of the two rectangles shown on Figure 5 is a Markov partition, which follows from Figure 7.

Let $\mathcal{R}$ be a Markov partition. Then for every $R \in \mathcal{R}$ the intersections of the form $f(R) \cap R_i$ for $R_i \in \mathcal{R}$ subdivide the rectangle $f(R)$ into a finite number of disjoint sub-rectangles, by cutting the expanding direction into pieces. For each of these sub-rectangles R' and $x \in R'$ we have $W_s(x, R) = W_s(x, R')$, see Figure 8. Note that an intersection $f(R) \cap R_i$ can consist of several sub-rectangles of $f(R)$.

Consider an oriented graph with the vertices identified with the elements of $\mathcal{R}$ and an edge from R to R_i for every rectangular piece of a non-empty intersection

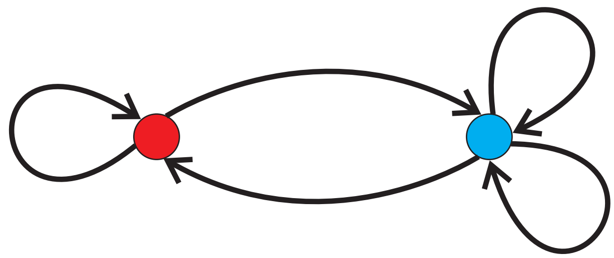

FIGURE 9. Structural graph

$f(R) \cap R_i$ (of open rectangles). Let us call this graph the *structural graph* of the Markov partition. For an edge e of the structural graph, denote by R_e the corresponding (closed) piece of an intersection $f(R_1) \cap R_2$ for $R_i \in \mathcal{R}$ (so that R_1 and R_2 correspond to the beginning and end of the edge e).

For example, the structural graph of the Markov partition from Figure 5 is shown on Figure 9. Note that its adjacency matrix is equal to $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$.

Consider the space $\mathcal{M}$ of all bi-infinite paths in the structural graph. It is obviously a topological Markov shift (in particular, it is a shift of finite type). It follows from the adjacency matrix of the graph that complexity of this shift is equal to a function of the form $C_1\lambda^n + C_2\lambda^{-n}$, hence *entropy* of the shift is $\lim_{n\to\infty} \frac{\log p_{\mathcal{M}}(n)}{n} = \log \lambda$.

PROPOSITION 3.12. *For every infinite path $w = \ldots e_{-1}e_0e_1\ldots$ in the structural graph of the Markov partition there exists exactly one point $\phi(w) \in \mathbb{R}^2/\mathbb{Z}^2$ such that $A^n(\phi(w)) \in R_{e_n}$ for all $n \in \mathbb{Z}$. The map $\phi : \mathcal{M} \longrightarrow \mathbb{R}^2/\mathbb{Z}^2$ is a semi-conjugacy of the shift with $(\mathbb{R}^2/\mathbb{Z}^2, A)$.*

We will prove later a general result implying that the constructed semi-conjugacy is a finite presentation (i.e., that its kernel $\mathcal{E}_\phi$ is a shift of finite type).