Dynamical Systems and Ergodic Theory

Jiaqi Cui
East China Normal University

October 3, 2025

Abstract

This is my notes of the course "Dynamical Systems and Ergodic Theory" given by Manfred Einsiedler. https://video.ethz.ch/lectures/d-math/2024/spring/401-2374-24L.html

Contents

0 Examples

Let X be a set and $T\colon X \to X$ be a map.

Definition 0.1. *fixed point, periodic point, period, orbit...*

Definition 0.2. *Assume X has a topology. The ω-limit of $x \in X$ is*

$$\omega^{\pm}(x) := \left\{ \lim_{k\to\infty} T^{n_k}x : n_k \nearrow \pm\infty \right\}.$$

We could also ask about the "distribution" of $x, Tx, T^2x, \cdots, T^n x$ inside X as $n \to \infty$.
More generally, a dynamical system can be defined as a group action.

Example 0.3. $X = \mathbb{R}$, $Tx = x + 1$. The ω-limits are empty set. Thus we will restrict to compact metric spaces.

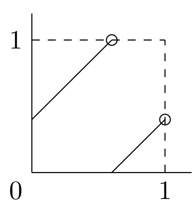
Example 0.4. $X = \mathbb{R} \cup \{\infty\}$ the one-point compactification of $\mathbb{R}$. $T(x) = \begin{cases} x+1, \ x \in \mathbb{R}; \\ \infty, \ x = \infty \end{cases}$. Then the ω-limits are all $\{\infty\}$.

Example 0.5. $X = \mathbb{R} \cup \{\pm\infty\}$. $T(x) = \begin{cases} x+1, \ x \in \mathbb{R}; \\ +\infty, \ x = +\infty; \\ -\infty, \ x = -\infty \end{cases}$. Then $\omega^+(x) = \begin{cases} +\infty, \ x \in \mathbb{R} \cup \{+\infty\}; \\ -\infty, \ x = -\infty \end{cases}$.

Example 0.6. North-South Dynamics

Example 0.7. $X = \mathbb{T} = \mathbb{R}/\mathbb{Z} \cong \mathbb{S}^1$ with the metric $d(x+\mathbb{Z}, y+\mathbb{Z}) = \min_{k\in\mathbb{Z}} |x - y + k|$. $R(x+\mathbb{Z}) := x+\alpha+\mathbb{Z}$ for a fixed $\alpha \in \mathbb{T}$. $R\colon \mathbb{T} \to \mathbb{T}$ is an isometry.

- If $\alpha = \frac{p}{q}$ is rational, then $R^q(x + \mathbb{Z}) = x + q\alpha + \mathbb{Z} = x + p + \mathbb{Z} = x + \mathbb{Z}$. Every point is periodic with period q.

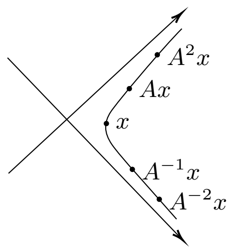


- If $\alpha \notin \mathbb{Q}/\mathbb{Z}$, then no point is periodic: say $R^n(x+\mathbb{Z}) = x+\mathbb{Z}$, then $n\alpha \in \mathbb{Z}$. Actually, all orbits are dense in this case.

Example 0.8. $X = \mathbb{T} = \mathbb{R}/\mathbb{Z}$. Fix $p \geqslant 2 \in \mathbb{N}$. $T(x) := px$. This map links to the base-p expansion of $x \in [0,1)$. Suppose $x = \sum_{k=1}^{\infty} \theta_k p^{-k}$ where $\theta_k\{0,\cdots,p-1\}$. Then $Tx = p\sum_{k=1}^{\infty} \theta_k p^{-k} + \mathbb{Z} = \sum_{k=1}^{\infty} \theta_{k+1} p^{-k} + \mathbb{Z}$.

Claim.
- There exist lots of periodic points– they are dense.
- There exist pre-periodic points that are not periodic, where x is pre-periodic if its orbit $|\mathcal{O}^+(x)| < \infty$.
- Many periods exist.
- The set of pre-periodic points is countable.
- There exist $x \in \mathbb{T}$ with $\omega^+(x) = \mathbb{T}$.
- There exist $x \in \mathbb{T}$ with $\omega^+(x)$ uncountable but not $\mathbb{T}$.
- There exist $x \in \mathbb{T}$ with $\omega^+(x)$ countable but not finite.

Example 0.9. $X = \mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. This is called a hyperbolic toral automorphism, because the orbit of any $x \neq 0 \in X$ is on a hyperbola.



Example 0.10. $X = (0,1) - \mathbb{Q}$, $Tx = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$. This relates to continued fraction expansion

$$x = \cfrac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots}}}$$

where $a_1, a_2, a_3, \cdots \in \mathbb{N}$. Note that

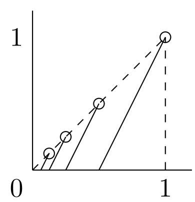
$$Tx = \cfrac{1}{a_2 + \frac{1}{a_3 + \frac{1}{a_4 + \cdots}}}.$$

Example 0.11 (Benford‘s law for powers of 2)**.** Given $j \in \{1,\cdots,9\}$, the limits

$$d_j := \lim_{N\to\infty} \frac{1}{N} \sharp \{2^n : 1 \leqslant n \leqslant N,\ 2^n \text{ starts in digital expension with } j\}$$

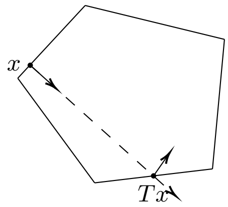
satisfy $d_1 > d_2 > \cdots > d_9 > 0$. In fact $d_1 = \log_{10} 2$.

Example 0.12. $X = [0,1]$, $T(x) = \begin{cases} 0, & x = 0,1; \\ nx - 1, & x \in \left[\frac{1}{n}, \frac{1}{n-1}\right) \end{cases}$.



We claim that $\lim_{n\to\infty} T^n x = 0$, and if $x \in \mathbb{Q}$, there exists n with $T^n x = 0$ and if $x \notin \mathbb{Q}$, $T^n x > 0$ for all n. For $x = e$ this can be used to show that $e \notin \mathbb{Q}$.

Example 0.13 (Billiards)**.** X is the set of boundary points with a vector and T is the movement of a boundary point along its vector to the next boundary point with a reflected vector.



Given a triangle, are there always periodic points? When the triangle is acute, right or obtuse with rational angles, the answer is yes. For the other cases, people don't know.

Example 0.14 (Geodesic flow)**.** Given a nice manifold M and its unit tangent bundle. There exists a way of following any vector in the tangent. When M is a sphere, the orbits are great circles. When M is a torus, whether an orbit is closed depending on whether the initial vector is rational. When M is a orientable surface of genus greater than 2, the orbits will be more complicated and more sensitive to initial values.

1 Topological Dynamics

Assume X is a compact metric space and $T\colon X \to X$ is continuous or even a homeomorphism.

Definition 1.1. *A homeomorphism $T\colon X \to X$ is called (topological) transitive if there exists a point $x_0 \in X$ for which the orbit is dense, i.e. $\overline{\mathcal{O}(x_0)} = X$.*

Definition 1.2. *A comtinuous map $T\colon X \to X$ is called forward transitive if there exists $x_0 \in X$ with $\overline{\mathcal{O}^+(x_0)} = X$.*

Example 1.3. $T_p\colon \mathbb{T} \to \mathbb{T}$ for $p \geqslant 2$ an integer which maps x to px is forward transitive. We will construct x_0 using base-p-expansion. We first list all finite sequences in the symbols $0, 1, \cdots, p-1$, and consider the result as one sequence of digits $a_1, a_2, \cdots \in \{0, 1, \cdots, p-1\}$. Then we define $x_0 := \sum_{k=1}^{\infty} a_k p^{-k}$. For any $\left[\frac{j}{p^i}, \frac{j+1}{p^i}\right)$, we can find an l such that $T^l x_0$ is in this interval. Thus $\mathcal{O}^+(x_0)$ is dense in $\mathbb{T}$. For example, for $p = 2$, we write

$$0, 1, 00, 01, 10, 000, \cdots, 111, 0000, \cdots, 1111, \cdots$$

Then we write it as a number sequence

$$0, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0, \cdots, 1, 1, 1, 0, 0, 0, 0, \cdots, 1, 1, 1, 1, \cdots$$

and then

$$x_0 = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} + 0 \cdot \frac{1}{8} + 0 \cdot \frac{1}{16} + 0 \cdot \frac{1}{32} + 1 \cdot \frac{1}{32} + 1 \cdot \frac{1}{64} + \cdots$$

When we apply T on x_0 for n times, the first n numbers of the number sequence will become 0. Then for any $\frac{j+1}{2^i}$, we can find an n such that the base-2-expansion of $\frac{j+1}{2^i}$ will at the start of the number sequence. This means $T^n x_0 \in \left[\frac{j}{2^i}, \frac{j+1}{2^i}\right)$.

Example 1.4. $R_\alpha\colon \mathbb{T} \to \mathbb{T}$ maps x to $x + \alpha$.

- If $\alpha \in \mathbb{Q}/\mathbb{Z}$, R_α only has periodic orbits and so is not transitive.
- If $\alpha \notin \mathbb{Q}/\mathbb{Z}$, R_α is topological transitive. See later.

Proposition 1.5. *Let $T\colon X \to X$ be a homeomorphism. The followings are equivalent:*

1. *T is topological transitive;*
2. *if $U \subset X$ is open and $TU = U$, then either $U = \varnothing$ or $\overline{U} = X$;*
3. *if $U, V \subset X$ are non-empty and open, then there exists $n \in \mathbb{Z}$ so that $T^n U \cap V \neq \varnothing$;*
4. *the set $\left\{x_0 \in X : \overline{\mathcal{O}(x_0)} = X\right\}$ is a dense G_δ-set.*

Definition 1.6. *A set G is called a G_δ-set if it is a countable intersection of open sets.*

Theorem 1.7 (Baire Category Theorem)**.** *Let X be a complete metric space. Let $O_n \subset X$ be a sequence of dense open sets. Then $\bigcap_{n=1}^{\infty} O_n$ is a dense G_δ-set.*

Proof. We only prove that $\bigcup_{n=1}^{\infty} O_n$ is dense. For any open set U, we want to find a point in $U \cap \bigcup_{n=1}^{\infty} O_n$. First, $U \cap O_1$ is non-empty and open because O_1 is open and dense. Then we can find a open ball $B_{\varepsilon_1}(x_1) \subset U \cap O_1$. Repeat this process. We find a open ball $B_{\varepsilon_2}(x_2) \subset B_{\varepsilon_1}(x_1) \cap O_2$... WLOG we can suppose $\varepsilon_n \leqslant \frac{1}{n}$. Thus we can claim that $\{x_n\}$ is a Cauchy sequence and $x := \lim_{n\to\infty} x_n \in U \cap \bigcap_{n=1}^{\infty} O_n$: by construction, $x_m \in B_{\varepsilon_{m-1}}(x_{m-1}) \subset \cdots \subset B_{\varepsilon_n}(x_n)$ and then $d(x_m, x_n) < \varepsilon_n \leqslant \frac{1}{n}$. Note that X is complete, the limit $x = \lim_{n\to\infty} x_n \in X$ exists. For all n, taking the limit of m, we obtain $x = \lim_{m\to\infty} x_m \in \overline{B_{\varepsilon_n}(x_n)} \subset O_n$. Moreover, $x \in \overline{B_{\varepsilon_1}(x_1)} \subset U$. □

Corollary 1.8. *A countable intersection of dense G_δ-sets is a dense G_δ-set.*

Proof of Proposition 1.5. (1)⇒(2): Let $x_0 \in X$ with $\mathcal{O}(x_0)$ dense in X. Because U is open, $\mathcal{O}(x_0) \cap U \neq \varnothing$. Then there exists $n \in \mathbb{Z}$ such that $T^n x_0 \in U$. Note that $TU = U$, we have $x_0 \in T^{-n}U = U = T^{-m}U$ for any $m \in \mathbb{Z}$. This shows $\mathcal{O}(x_0) \subset U$ and then U is dense in X.

(2)⇒(3): Define $\widetilde{U} = \bigcup_{n\in\mathbb{Z}} T^n U$. We note that $T\widetilde{U} = \widetilde{U}$ is non-empty and open. Then it is dense. Because V is open, $\widetilde{U} \cap V \neq \varnothing$ and then there exists $n \in \mathbb{Z}$ such that $T^n U \cap V \neq \varnothing$.

(3)⇒(4): For any $n \in \mathbb{N}$, $\bigcup_{x\in X} B_{\frac{1}{n}}(x)$ is a open cover. Because X is compact, there exists $k(n) \in \mathbb{N}$ such that $\bigcup_{i=1}^{k(n)} B_{\frac{1}{n}}(x_i)$ covers X. Denote $B_{\frac{1}{n}}(x_i)$, $i = 1, \cdots, k(n)$, $n \in \mathbb{N}$ by $O_1, O_2, \cdots$. To show density of a set, it suffices to show that the set intersects any of those open sets $O_1, O_2, \cdots$. For any $j \in \mathbb{N}$, we define $\widetilde{O_j} = \bigcup_{n\in\mathbb{Z}} T^n O_j$. By assumption, $\widetilde{O_j}$ intersects any other open set. This means $\widetilde{O_j}$ is dense. By Baire Category Theorem, $G := \bigcap_{j=1}^{\infty} \widetilde{O_j}$ is a dense G_δ-set and consists precisely of all points $x_o \in X$ with dense orbit. To see that, if x_0 has dense orbit, $\mathcal{O}(x_0)$ must intersect all open set $O_1, O_2, \cdots$. Then for any O_i, there is $n \in \mathbb{Z}$ such that $T^n x_0 \in O_i$. Thus $x_0 \in \widetilde{O_i}$ and then $x_0 \in G$. Conversely, for any $x_0 \in G$, $\mathcal{O}(x_0)$ intersects open balls with any small radius. Thus it is dense.

(4)⇒(1): Dense set can not be empty. □

Definition 1.9. *A homeomorphism $T\colon X \to X$ is called topological mixing if for any two non-empty open sets $U, V \subset X$, there exists N such that $T^n U \cap V \neq \varnothing$ for all $n \in \mathbb{Z}$ with $|n| > N$.*

Definition 1.10. *A homeomorphism $T\colon X \to X$ is called minimal if every orbit is dense.*

Proposition 1.11. *Let $T\colon X \to X$ be a homeomorphism of a compact metric space X. The followings are equivalent:*

1. *T is minimal;*
2. *if $E = TE \subset X$ is closed, then either $E = \varnothing$ or $E = X$;*
3. *if $U \subset X$ is open and non-empty, then $\bigcup_{n\in\mathbb{Z}} T^n U = X$.*

Proof. (1)⇒(2): Suppose $x \in E$. We have $X = \overline{\mathcal{O}(x)} \subset \overline{E} = E$.

(2)⇒(3): Denote $\widetilde{U} = \bigcup_{n\in\mathbb{Z}} T^n U$. $\widetilde{U}$ is open and $T\widetilde{U} = \widetilde{U}$. Hence $E := X - \widetilde{U}$ is closed and $TE = E$. Then we have $E = \varnothing$.

(3)⇒(1): Let $x_0 \in X$ and $U \neq \varnothing \subset X$ be open. Then there is $n \in \mathbb{Z}$ such that $x_0 \in T^n U$. This shows that $\mathcal{O}(x_0)$ intersects any non-empty open subset of X. Hence it is dense. □

Theorem 1.12. *Let X be a compact metric space and $T\colon X \to X$ be a homeomorphism. Then there exists a closed non-empty subset $Y \subset X$ so that $Y = TY$ and $T|_Y\colon Y \to Y$ is minimal.*

Proof. Define
$$\mathcal{E} := \{Y \subset X : Y \neq \varnothing, Y \text{ closed}, TY = Y\} \ni X.$$
We want to use Zorn's Lemma under $\subset$. We need to show that any chain in $\mathcal{E}$ has a lower bound in $\mathcal{E}$. Let $(Y_\alpha : \alpha \in \mathcal{I})$ be a chain. Define $Y = \bigcap_{\alpha\in\mathcal{I}} Y_\alpha$. Clearly, $TY = Y$ and Y is closed. Note that any intersection of finite Y_α is non-empty, by compactness of X, $Y \neq \varnothing$. These show that Y is the lower bound of chain $(Y_\alpha : \alpha \in \mathcal{I})$.

By Zorn's Lemma, there is a minimal element $Y \in \mathcal{E}$. Then $T|_Y\colon Y \to Y$ is minimal by Proposition 1.11. □

Corollary 1.13. *There exists $x_0 \in X$ so that $T^{n_k} x_0 \to x_0$ as $k \to \infty$ for some $n_k \nearrow \infty$.*

Proof. Let $Y \subset X$ be as in Theorem 1.12 and $x_0 \in Y$ be arbitrary. We have $\omega^+(x_0) \subset Y$ and $\omega^+(x_0)$ is T-invariant closed set. Note that $T|_Y : Y \to Y$ is minimal, by Proposition 1.11, $\omega^+(x_0) = \varnothing$ or Y. Note again that Y is a closed subset of the compact space X, it is also compact. Hence $\omega^+(x_0) \neq \varnothing$. This shows that $x_0 \in Y = \omega^+(x_0)$. □

Definition 1.14. *$x_0 \in X$ is called recurrent for $T\colon X \to X$ if $x_0 \in \omega_T^+(x_0)$.*

Example 1.15. Let $\alpha \notin \mathbb{Q}$ and $R\colon \mathbb{T} \to \mathbb{T}$ which maps x to $x + \alpha$. Then R is minimal: by 1.13, there is $x_0 \in \mathbb{T}$ which is recurrent. Let $\varepsilon > 0$. There is $n \in \mathbb{N}$ such that $\varepsilon > d(R^n x_0, x_0) = d(x_0 + n\alpha, x_0) = d(n\alpha, 0)$. Note that $n\alpha \neq 0$ in $\mathbb{T}$. For any $x \in \mathbb{T}$, $x, x \pm n\alpha, x \pm 2n\alpha, \cdots, x \pm \left\lfloor \frac{1}{d(n\alpha,0)} \right\rfloor n\alpha \in \mathcal{O}(x)$ is ε-dense in $\mathbb{T}$. This shows every orbit is dense.

Definition 1.16. *Let $T\colon X \to X$ be a homeomorphism. We say T is expansive if there exists (the expansive constant) $\delta > 0$ so that for any $x \neq y \in X$, there exists $n \in \mathbb{Z}$ so that $d(T^n x, T^n y) \geqslant \delta$.*

Example 1.17. $T_A\colon \mathbb{T}^2 \to \mathbb{T}^2$ given by matrix multiplication defined by $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ is expansive, topological transitive, mixing and not minimal.

Theorem 1.18 (Multiple Recurrence Theorem)**.** *Let X be a compact metric space and let $T_1, \cdots, T_d\colon X \to X$ be pairwise commutative homeomorphisms. Then there exists some $x_0 \in X$ and some $n_k \nearrow \infty$ as $k \to \infty$ so that $T_j^{n_k} x_0 \to x_0$ for $j = 1, \cdots, d$.*

Proof. We will use induction. For $d = 1$, Corollary 1.13 works. Now we assume the theorem holds for $d-1$. And we also assume X is d-minimal in the sense that $Y \subset X$ closed and $Y = T_1 Y = \cdots = T_d Y$ imply $Y = \varnothing$ or X. This can be done by reapplying the proof of Theorem 1.12.

Denote

$$S = T_1 \times \cdots \times T_d\colon X^d \to X^d,$$
$$\widehat{T_j} = T_j \times \cdots \times T_j\colon X^d \to X^d.$$

$S, \widehat{T_1}, \cdots, \widehat{T_d}$ are pairwise commutative. For $\underline{n} = (n_1, \cdots, n_d) \in \mathbb{Z}^d$, we define

$$T^{\underline{n}} = T_1^{n_1} \circ \cdots \circ T_d^{n_d}\colon X \to X,$$
$$\widehat{T}^{\underline{n}} = \widehat{T_1}^{n_1} \circ \cdots \circ \widehat{T_d}^{n_d}\colon X^d \to X^d.$$

Denote $\Delta(x) = (x, \cdots, x)$ the diagonal element in X^d and $\Delta_X = \{\Delta(x) : x \in X\}$. By commutativity, $\widehat{T}^{\underline{n}}(\Delta(x)) = \Delta(T^{\underline{n}x})$. We need to prove that there exist some $x_0 \in X$ and $n_k \nearrow \infty$ such that $S^{n_k}(\Delta(x_0)) \to \Delta(x_0)$ as $k \to \infty$.

Claim (A)**.** $\Delta_X \cap \omega_S^+(\Delta_X) \neq \varnothing$, where for the subset $\Delta_X \subset X$, we define $\omega_S^+(\Delta_X) = \{\lim_{k\to\infty} S^{n_k}(\Delta(y_k)) : \Delta(y_k) \in \Delta_X, n_k \nearrow \infty\}$.

Proof of Claim (A). Denote

$$R_1 = T_1 T_d^{-1}, \cdots, R_{d-1} = T_{d-1} T_d^{-1}\colon X \to X.$$

By inductive hypothesis, there exists $x_0 \in X$ and $n_k \nearrow \infty$ so that $R_j^{n_k} x_0 \to x_0$ as $k \to \infty$ for all $j = 1, \cdots, d-1$. Define $y_k = T_d^{n_k} x_0$. For $j < d$, $T_j^{n_k} y_k = R_j^{n_k} x_0 \to x_0$; for $j = d$, $T_d^{n_k} y_k = x_0$, as $k \to \infty$. This means $\Delta(x_0) = \lim_{k\to\infty} S^{n_k}(\Delta(y_k)) \in \Delta_X \cap \omega_S^+(\Delta_X) \neq \varnothing$. □

Claim (B)**.** $\Delta_X \subset \omega_S^+(\Delta_X)$.

Proof of Claim (B). This proof needs the minimality assumption. Denote $Y = \{x \in X : \Delta(x) \in \omega_S^+(\Delta_X)\}$. Because $[x \mapsto \Delta(x)]$ is continuous and $\omega_S^+(\Delta_X)$ is closed by diagonal principle, Y is closed. By Claim (A), $Y \neq \varnothing$. By Proposition 1.11, we only need to prove that $T_j^{\pm 1} Y \subset Y$ for $j = 1, \cdots, d$, then $Y = X$ and the claim follows.

Let $x \in Y$. Then there exists $n_k \nearrow \infty$ and $y_k \in X$ such that $\Delta(x) = \lim_{k\to\infty} S^{n_k}(\Delta(y_k))$. Hence we have

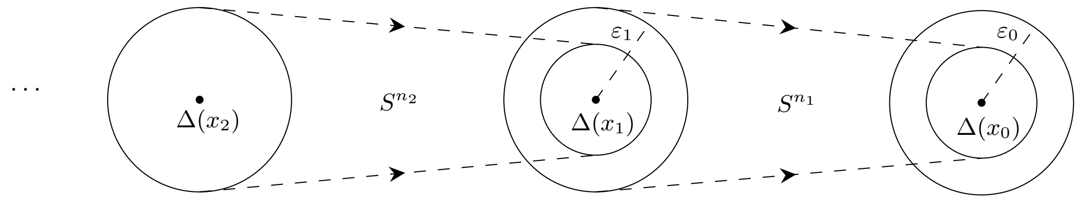
$$\Delta(T_j^{\pm 1} x) = \widehat{T_j}^{\pm 1}(\Delta(x)) = \lim_{k\to\infty} \widehat{T_j}^{\pm 1} S^{n_k}(\Delta(y_k)) = \lim_{k\to\infty} S^{n_k} \Delta(T_j^{\pm 1} y_k) \in \omega_S^+(\Delta_X).$$

This shows $T_j^{\pm 1} x \in Y$. □

Claim (C)**.** For every $\varepsilon > 0$, there exists a point $x \in X$ and some $n \geqslant 1$ so that $d(S^n(\Delta(x)), \Delta(x)) < \varepsilon$.

Proof of Claim (C). Let $x_0 \in X$ be arbitrary and $\varepsilon_0 \in \left(0, \frac{\varepsilon}{2}\right)$. By Claim (B), there exists $x_1 \in X$, $n_1 \geqslant 1$ so that $d(S^{n_1}(\Delta(x_1)), \Delta(x_0)) < \varepsilon_0$. By continuity of S^{n_1} there exists $\varepsilon_1 \in (0, \varepsilon_0)$ so that if $y = (y_1, \cdots, y_d) \in B_{\varepsilon_1}(\Delta(x_1))$, then $S^{n_1}(y) \in B_{\varepsilon_0}(\Delta(x_0))$.

Continuing inductively we find new points $x_k \in X$ and $n_k \geqslant 1$ so that $d(S^{n_k}(\Delta(x_k)), \Delta(x_{k-1})) \leqslant \varepsilon_{k-1}$, where $\varepsilon_k \in (0, \varepsilon_{k-1})$ so that if $y = (y_1, \cdots, y_d) \in B_{\varepsilon_k}(\Delta(x_k))$, then $S^{n_k}(y) \in B_{\varepsilon_{k-1}}(\Delta(x_{k-1}))$.



By compactness of X, we can find $0 \leqslant k < l$ so that $d(\Delta(x_k), \Delta(x_l)) < \frac{\varepsilon}{2}$. Applying S^{n_l} to $\Delta(x_l)$, we obtain a point in $B_{\varepsilon_{l-1}}(\Delta(x_{l-1}))$. By construction, note that $S^{n_l}(\Delta(x_l)) \in B_{\varepsilon_{l-1}}(\Delta(x_{l-1}))$, we obtain $S^{n_{l-1}+n_l}(\Delta(x_l)) \in B_{\varepsilon_{l-2}}(\Delta(x_{l-2}))$. Continuing inductively, we obtain $S^{n_{k+1}+\cdots+n_l}(\Delta(x_l)) \in B_{\varepsilon_k}(\Delta(x_k))$. Then $n = n_{k+1} + \cdots + n_l$ and $x = x_l$ are as desired:

$$d(S^{n_{k+1}+\cdots,n_l}(\Delta(x_l)), \Delta(x_l)) \leqslant d(S^{n_{k+1}+\cdots+n_l}(\Delta(x_l)), \Delta(x_k)) + d(\Delta(x_k), \Delta(x_l)) < \varepsilon_k + \frac{\varepsilon}{2} < \varepsilon.$$

$\square$

We define

$$F(x) = \inf_{n \in \mathbb{N}} d(S^n(\Delta(x)), \Delta(x)).$$

Claim (D)**.** If there exists $x_0 \in X$ such that $F(x_0) = 0$, then the theorem also holds for d maps. Thus the proof will be completed.

Proof of Claim (D). If there exists n so that $d(S^n(\Delta(x)), \Delta(x)) = 0$, then for $n_k = nk$, we have $\Delta^{n_k}(\Delta(x_0)) = \Delta(x_0) \to \Delta(x_0)$. If not, there exists n_k with $S^{n_k}(\Delta(x_0)) \to \Delta(x_0)$. $\square$

To prove the existence of such $x_0 \in X$, we need a lemma:

Lemma 1.19. *Let X be a compact metric space. Let $f_n \colon X \to [0, \infty)$ be a sequence of continuous. We define $F(x) = \inf_{n \in \mathbb{N}} f_n(x)$ for any $x \in X$.*

1. *F is upper semi-continuous: for any $x \in X$ and $\varepsilon > 0$, there exists some $\delta > 0$ so that for any $y \in B_\delta(x)$, we have $F(y) < F(x) + \varepsilon$.*

2. *The sets*
$$\mathcal{J}_\varepsilon := \left\{ x \in X : \forall \eta > 0, \sup_{y,z \in B_\eta(x)} |F(y) - F(z)| > \varepsilon \right\}$$
are closed with empty interior for all $\varepsilon > 0$.

3. *F is continuous on a dense G_δ-set in X.*

Proof of Lemma 1.19. (1): By definition, there is some n so that $f_n(x) < F(x) + \frac{\varepsilon}{2}$. Because f_n is continuous, there exists $\delta > 0$ so that for any $y \in B_\delta(x)$, $f_n(y) < f_n(x) + \frac{\varepsilon}{2}$. Now we have

$$F(y) \leqslant f_n(y) < f_n(x) + \frac{\varepsilon}{2} < F(x) + \varepsilon.$$

(2): Let $\overline{x} \in \overline{\mathcal{J}_\varepsilon}$. Let $\eta > 0$. Then there exists $x \in \mathcal{J}_\varepsilon \cap B_{\frac{\eta}{2}}(\overline{x})$. By triangle inequality, $B_{\frac{\eta}{2}}(x) \subset B_\eta(\overline{x})$. Thus by definition,

$$\sup_{y,z \in B_\eta(\overline{x})} |F(y) - F(z)| \geqslant \sup_{y,z \in B_{\frac{\eta}{2}}(x)} |F(y) - F(z)| > \varepsilon.$$

This implies $\overline{x} \in \mathcal{J}_\varepsilon$ and then $\mathcal{J}_\varepsilon$ is closed.

Suppose the there exists some $x_0 \in \mathcal{J}_\varepsilon^\circ$. By (1), there is a $\delta_0 > 0$ so that $F(y) < F(x_0) + \frac{\varepsilon}{2}$ for all $y \in B_{\delta_0}(x_0)$. We choose δ_0 small enough such that $B_{\delta_0}(x_0) \subset \mathcal{J}_\varepsilon$. We claim that we can find a point $x_1 \in B_{\delta_0}(x_0)$ such that $F(x_1) \leqslant F(x_0) - \frac{\varepsilon}{2}$. If not, for any $y \in B_{\delta_0}(x_0)$, $F(y) > F(x_0) - \frac{\varepsilon}{2}$. Associating our choice of δ_0, $|F(y) - F(x_0)| < \frac{\varepsilon}{2}$. Then for any $y, z \in B_{\delta_0}(x_0)$, $|F(y) - F(z)| \leqslant |F(y) - F(x_0)| + |F(x_0 - F(z))| < \varepsilon$. This gives a contradiction of the definition of $\mathcal{J}_\varepsilon$. Now we can repeat this to find $B_{\delta_1}(x_1) \subset B_{\delta_0} \subset \mathcal{J}_\varepsilon$ and

$x_2 \in B_{\delta_1}(x_1)$ such that $F(x_2) \leqslant F(x_1) - \frac{\varepsilon}{2} \leqslant F(x_0) - 2 \cdot \frac{\varepsilon}{2}$. Repeating this process, we get a sequence $x_n \in B_{\delta_0}(x_0)$ so that $F(x_n) \leqslant F(x_0) - n \cdot \frac{\varepsilon}{2}$. This contradicts to the fact that $F \geqslant 0$.

(3): By (2), $X - \mathcal{J}_{\frac{1}{n}}$ is open and dense. Thus by Baire Category Theorem, $G := \bigcap_{n \geqslant 0} \left(X - \mathcal{J}_{\frac{1}{n}} \right)$ is a dense G_δ-set. Fix $x \in G$. For any $\varepsilon > 0$, choose n such that $\frac{1}{n} < \varepsilon$. By construction, $x \notin \mathcal{J}_{\frac{1}{n}}$, i.e. there exists $\eta > 0$ so that $\sup_{y,z \in B_\eta(x)} |F(y) - F(z)| < \frac{1}{n} < \varepsilon$. In particular, $|F(y) - F(x)| < \varepsilon$ for all $y \in B_\eta(x)$. Hence F is continuous at x. □

Especially, we have

Claim (E)**.** There exists $x_0 \in X$ so that F is continuous at x_0.

Claim (F)**.** If F is continuous at x_0, then $F(x_0) = 0$.

Proof of Claim (F). Assume $F(x_0) > 0$. Then there is an open neighborhood U of x_0 and $\delta > 0$ such that $F(x) > \delta > 0$ in U. Note that $\widetilde{U} := \bigcup_{\underline{n} \in \mathbb{Z}^d} T^{-\underline{n}} U$ is non-empty and open, and $T_1 \widetilde{U} = \cdots = T_d \widetilde{U} = \widetilde{U}$. By minimality and Proposition 1.11, considering the closed set $Y = X - \widetilde{U}$, we obtain $\widetilde{U} = X$. By compactness, there is a finite set $F \subset \mathbb{Z}^d$ such that $X = \bigcup_{\underline{n} \in F} T^{-\underline{n}} U$. By continuity of $\widehat{T}^{\underline{n}}$ for $\underline{n} \in F$ and compactness of X, $\widehat{T}^{\underline{n}}$ is uniform continuous on X. And because F is finite, there exists some $\varepsilon > 0$ so that for all $x, y \in X^d$, $d(x, y) < \varepsilon$, we have $d(\widehat{T}^{\underline{n}} x, \widehat{T}^{\underline{n}} y) < \delta$, for any $\underline{n} \in F$.

By Claim (C), for this $\varepsilon > 0$, we can find an $x_\varepsilon \in X$ with $F(x) < \varepsilon$. Especially, we can also find an $m \geqslant 1$ so that $d(S^m(\Delta(x_\varepsilon), \Delta(x_\varepsilon)) < \varepsilon$. Besides, we can find an $\underline{n} \in F$ so that $x_\varepsilon \in T^{-\underline{n}} U$. Now by continuity of $\widehat{T}^{\underline{n}}$ and commutativity of the maps, we have

$$\delta > d(\widehat{T}^{\underline{n}}(S^m(\Delta(x_\varepsilon))), \widehat{T}^{\underline{n}}(\Delta(x_\varepsilon))) = d(S^m(\Delta(T^{\underline{n}}(x_\varepsilon))), \Delta(T^{\underline{n}}(x_\varepsilon)))$$

for $T^{\underline{n}}(x_\varepsilon) \in U$. Recall that $F > \delta$ on U. We get a contradiction. □

□

Remark 1.20. If T_j are not commutative, it fails. For example, consider the North-South dynamics and "East-West" dynamics on $\mathbb{S}^1$.

As corollary, we have:

Theorem 1.21 (van der Werden)**.** *Let $\mathbb{Z} = B_1 \sqcup B_2 \sqcup \cdots \sqcup B_k$ be a finite partition. Then there exists $B = B_j$ such that it contains arbitrarily long arithmetic progressions: for arbitrary N, there exist some $a, d \in \mathbb{Z}$ such that*

$$a, a + d, a + 2d, \cdots, a + Nd \in B$$

Proof. To prove this theorem, we need to construct a related dynamical system.

Let $X_{full} = \{1, \cdots, k\}^{\mathbb{Z}}$. The full shift is defined as $\sigma \colon X_{full} \to X_{full}$, $(\sigma(x))_n = x_{n+1}$.

We can define a metric on X_{full}:

$$d(x, y) := \begin{cases} 0, \ x = y; \\ \frac{1}{n+1}, \ n = \min\{|k| : x_k \neq y_k\} \end{cases}.$$

Claim. X_{full} is compact under this metric.

Proof of Claim. Note that d induces the standard product topology. Then Tychonoff's Theorem gives the claim. □

Now we turn to famous Furstenberg's correspondence. We define $z \in X_{full}$ by $z_n = j$ if $n \in B_j$ and then $X := \overline{\mathcal{O}_\sigma(z)} \subset X_{full}$. Consider the shift map $\sigma = \sigma|_X \colon X \to X$. Note that X is closed in a compact space, it's compact. Hence σ is a homeomorphism on a compact metric space X. Moreover, we define $T_1 = \sigma, \cdots, T_N = \sigma^N$.

By multiple recurrence theorem, there exists some $x \in X$ and some $n_l \nearrow \infty$ such that $T_1^{n_l} x \to x, \cdots, T_N^{n_l} x \to x$, i.e. $\sigma^{n_l} x \to x, \cdots, \sigma^{N n_l} x \to x$ as $l \to \infty$. Then for $\varepsilon = 1$, we can find an n_l such that $d(\sigma^{n_l} x, x) < 1, \cdots, d(\sigma^{N n_l} x, x) < 1$. Denote $d = n_l$. By definition, this happens if and only if $x_d = x_0, \cdots, x_{Nd} = x_0$.

Now we work with the facts that $x_0 = x_d = \cdots = x_{Nd}$ and $x \in X = \overline{\mathcal{O}_\sigma(z)}$. There exists some $a \in \mathbb{Z}$ such that $d(x, \sigma^a z) < \frac{1}{Nd+1}$. By definition, x and $\sigma^a z$ have the same symbols at coordinates $-Nd, \cdots, 0, \cdots, Nd$. Therefore,

$$\begin{tikzcd}
x_0 \arrow[r, equal] \arrow[d, equal] & x_d \arrow[r, equal] \arrow[d, equal] & \cdots \arrow[r, equal] \arrow[d, equal] & x_{Nd} \arrow[d, equal] \\
z_a & z_{a+d} & \cdots & z_{a+Nd}
\end{tikzcd}$$

This forces,

$$\begin{tikzcd}
x_0 \arrow[r, equal] \arrow[d, equal] & x_d \arrow[r, equal] \arrow[d, equal] & \cdots \arrow[r, equal] \arrow[d, equal] & x_{Nd} \arrow[d, equal] & \\
z_a \arrow[r, equal] & z_{a+d} \arrow[r, equal] & \cdots \arrow[r, equal] & z_{a+Nd} \arrow[r, equal] & j
\end{tikzcd}$$

for a $j \in \{1, \cdots, k\}$. Hence $a, a+d, \cdots, a+Nd \in B_j$, which are as desired.

□

2 Symbolic Dynamics

Recall that the full shift on a finite alphabet $\mathcal{A} = \{1, \cdots, k\}$ is defined on $X_{full} = \mathcal{A}^{\mathbb{Z}}$ by $(\sigma(x))_n = x_{n+1}$. This defines a homeomorphism $\sigma \colon X_{full} \to X_{full}$ on a compact metric space. A shift space is a σ-invariant closed subset $X \subset X_{full}$ together with $\sigma = \sigma|_X \colon X \to X$.

Definition 2.1. *A cylinder set in X_{full} or X is defined by*

$$[w]_{m,n} = \{x \in X : w_m = x_m, \cdots, w_n = x_n\}$$

where $w \in X$ and $m \leqslant n$.

Proposition 2.2. *Cylinder sets are compact and open.*

Proof. For compactness, note that $[w]_{m,n} \cong \mathcal{A}^{((-\infty, m-1] \sqcup [n+1, +\infty)) \cap \mathbb{Z}}$ and Tychonoff's theorem works. For openness, note that if $m = -n$, then $[w]_{-n,n}$ is an open ball in X. Thus in general case, $[w]_{m,n} = \bigcup_{v \in [w]_{m,n}} [v]_{-N,N}$, where $N = \max\{|m|, |n|\}$, is open. □

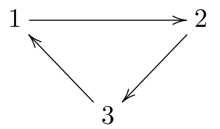
Example 2.3. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an oriented finite graph. We can define the vertex shift by

$$X_{\mathcal{G}} = \left\{x \in \mathcal{V}^{\mathbb{Z}} : \text{for any } n,\ x_n \text{ connects to } x_{n+1} \text{ by an edge}\right\}.$$

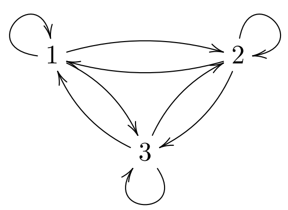
- Possibly $X_{\mathcal{G}} = \varnothing$, e.g.

$$\mathcal{G} = \ 1 \longrightarrow 2 \longrightarrow 3$$

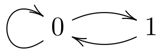
- Possibly $X_{\mathcal{G}}$ is finite, e.g.

$\mathcal{G} =$ 

- Full shift is a vertex shift, .e.g.

$\mathcal{G} =$ 

- Golden mean shift:

$\mathcal{G} =$ 

The adjacency matrix $A = A_{\mathcal{G}}$ is defined by

$$A_{ij} = \begin{cases} 1, & i \text{ connectes to } i; \\ 0, & \text{otherwise} \end{cases}.$$

Lemma 2.4. *1. $(A^n)_{ij}$ is the number of paths from i to j. $\mathrm{Tr}(A^n)$ is the number of periodic points in $X_{\mathcal{G}}$ of period n or a divisor of n.*

2. If $\mathcal{G}$ is connected, i.e. $\forall i, j$, $\exists n \geqslant 1$ such that $(A^n)_{ij} > 0$, then $X_{\mathcal{G}}$ is topological transitive.

3. $\mathcal{G}$ is connected and aperiodic, i.e. there exists n with $(A^n)_{ij} > 0$ for all i, j, if and only if $X_{\mathcal{G}}$ is topological mixing.

Proof. (1): We use induction. For $n = 1$, the proof is by definition. Suppose the claim holds for n. Then

$$(A^{n+1})_{ij} = \sum_l A_{il}(A^n)_{lj}$$

where A_{il} is depending on whether l connects i and $(A^n)_{lj}$ is the number of paths from j to l. The claim follows.

Now $\mathrm{Tr}(A^n) = \sum_i (A^n)_{ii}$ is the number of closed paths with identified starting point, which equals to the number of periodic points with period n or a divisor of n.

(2): Suppose $U, V \subset X_{\mathcal{G}}$ be non-empty open subsets of the vertex shift. We wish to find some n such that $\sigma^{-n}(U) \cap V \neq \varnothing$. Then by Proposition 1.5, the claim follows.

We may assume $[w]_{-N,N} \subset U$ and $[v]_{-N,N} \subset V$. Denote $j = w_N \in \{1, \cdots, k\}$ and $i = v_{-N}$. Then by connectedness there exists some $l \geqslant 1$ such that $(A^l)_{ij} > 0$. This means it's possible to go from j to i in l steps. Denote $u = (u_0 = j, u_1, \cdots, u_{l-1}, u_l = i)$. Then we define

$$\begin{aligned} x = (\cdots, x_{-3N-l-1} = w_{-N-1}, x_{-3N-l} = w_{-N}, \cdots, x_{-N-l} = w_N = j = u_0,& \\ x_{-N-l+1} = u_1, \cdots, x_{-N} = u_l = i = v_{-N},& \\ x_{-N+1} = v_{-N+1}, \cdots).& \end{aligned}$$

In fact, we first go along w until $x_{-N-l} = w_N = u_0$, then we go along u, finally we go along v from $x_{-N} = u_l = v_{-N}$. It's easy to check that $x \in \sigma^{2N+l}([w]_{-N,N}) \cap [v]_{-N,N} \subset \sigma^{2N+l}(U) \cap V \neq \varnothing$. Thus $X_{\mathcal{G}}$ is topological transitive.

(3): ($\Rightarrow$): The proof is similar to (2).

Claim. For any $m \geqslant n$, $(A^m)_{ij} > 0$ for any i, j.

Proof of Claim. Suppose $(A^{m+1})_{ij} = 0$ and $(A^m)_{ij} > 0$ for any i, j. Then $(A^{m+1})_{ij} = \sum_l A_{il}(A^m)_{lj}$ shows that $A_{il} = 0$ for any l. Hence $(A^m)_{il} = 0$ for any m and i by induction. We get a contradiction. □

For any two open sets U, V, we may assume $[w]_{-N,N} \subset U$ and $[v]_{-N,N} \subset V$. For any $m \geqslant n$, we can go from w_N to v_{-N} in m steps along some path u. We first go along w until $x_{-N-l} = w_N = u_0$, then we go along u, finally we go along v from $x_{-N} = u_m = v_{-N}$. It's easy to check that $x \in \sigma^{2N+m}([w]_{-N,N}) \cap [v]_{-N,N} \subset \sigma^{2N+m}(U) \cap V$. It follows that for any $k > 2N + n$, $\sigma^k(U) \cap V \neq \varnothing$. By altering the roles of U, V, we can prove $\sigma^{-k}(U) \cap V = \sigma^{-k}(U \cap \sigma^k(V)) \neq \sigma^{-k}(\varnothing) = \varnothing$ for any $k > 2N + n$. Above all, for any $|k| > 2N + n$, $\sigma^k(U) \cap V \neq \varnothing$. Thus $X_{\mathcal{G}}$ is topological mixing.

($\Leftarrow$): Note that $[(\cdots, i, i, i, \cdots)]_{0,0}$, $i \in \mathcal{V}$ are open balls. Then because $\mathcal{V}$ is finite, there is a universal constant N such that for any $|n| > N$, $\sigma^n([(\cdots, i, \cdots)]_{0,0}) \cap [(\cdots, j, \cdots)]_{0,0} \neq \varnothing$ for any $i, j \in \mathcal{V}$. By above construction, this means for any $|n| > N$, any $i, j \in \mathcal{V}$ can be connected by a path in n steps, i.e. $(A^n)_{ij} > 0$. Thus $\mathcal{G}$ is connected and aperiodic. □

Definition 2.5. *A shift of finite type (sft) is a (closed shift-invariant) subset $X \subset X_{full} = \mathcal{A}^{\mathbb{Z}}$ defined by a finite list of forbidden finite words. More precisely, there should exist N and a finite set $\mathcal{F} \subset \mathcal{A}^{\{1,\cdots,N\}}$ so that $X = \{x \in \mathcal{A}^{\mathbb{Z}} : \forall n, (x_{n+1}, \cdots, x_{n+N}) \notin \mathcal{F}\}$.*

Definition 2.6. *$X \subset \mathcal{A}^{\mathbb{Z}}$ is called a sofic system if there exists a shift of finite type $Y \subset \mathcal{B}^{\mathbb{Z}}$ and a continuous map $\Phi \colon Y \to X$ with $\Phi(Y) = X$ and $\Phi \circ \sigma_Y = \sigma_X \circ \Phi$.*

Lemma 2.7. *For any shift of finite type X, there exists a vertex shift $X_{\mathcal{G}}$ such that $X_{\mathcal{G}}$ is sofic to X.*

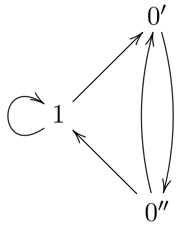
Proof. Suppose X is defined by $\mathcal{F} \subset \mathcal{A}^{\{1,\cdots,N+1\}}$. We define $\mathcal{V} = \mathcal{A}^{\{1,\cdots,N\}}$. We connect $v, v' \in \mathcal{V}$ if there exists $w \in \mathcal{A}^{\{1,\cdots,N+1\}}$ so that $w \notin \mathcal{F}$ and $w = (v, w_{N+1}) = (w_1, v')$. This defines a graph $\mathcal{G}$ and so also a vertex shift $X_{\mathcal{G}}$.

We define a map $\Phi \colon X \to X_{\mathcal{G}}$ by $x \mapsto (\Phi(x))_n = (x_n, \cdots, x_{n+N-1})$. By construction, x doesn't have forbidden words in $\mathcal{F}$. Thus $(\Phi(x))_n = (x_n, \cdots, x_{n+N-1})$ connects $(\Phi(x))_{n+1} = (x_{n+1}, \cdots, x_{n+N})$. This shows that Φ is well-defined. Moreover, $\Phi(X) = X_{\mathcal{G}}$. To see this, let $v \in X_{\mathcal{G}}$. By definition, there exists $x \in X_{full}$ so that $v_n = (x_n, \cdots, x_{n+N-1})$. It's easy to check that $x \in X$ by definition. Moreover, it's also easy to check that Φ is a homeomorphism and $\Phi \circ \sigma_X = \sigma_{X_{\mathcal{G}}} \circ \Phi$. □

Example 2.8. The even shift $X_{even} \subset \{0,1\}^{\mathbb{Z}}$ is sofic but bot of finite type, where

$$X_{even} = \left\{x \in \{0,1\}^{\mathbb{Z}} : \text{any two 1's in the sequence are separated by an even number of 0's}\right\}.$$

Let Y be the vertex shift defined by



and Φ forget the primes over 0's. Hence we have defined $\Phi\colon Y \to X_{even}$.

Definition 2.9 (Complexity)**.** *Let $X \subset X_{full} = \mathcal{A}^{\mathbb{Z}}$ be a shift. We define the complexity function*

$$p_X(n) := \left|\pi_{\{1,\cdots,n\}}(X)\right| := \text{the number of words of length } n \text{ appearing in any } x \in X.$$

Lemma 2.10. *If X is a shift of finite type, then $p_X(n)$ either grows polynomially (is constant in case X is transitive) or grows exponentially. If X is a topological mixing vertex shift, then either $|X| = 1$ or $p_X(n)$ grows exponentially.*

Proof. By Lemma 2.7, there exists a vertex shift $x_{\mathcal{G}}$ such that $X_{\mathcal{G}}$ is sofic to X. It's easy to check that p_X and $p_{X_{\mathcal{G}}}$ have the same growth type. Thus we can assume that $X = X_{\mathcal{G}}$ for simplicity. Moreover, we can remove all sinks and sources of $\mathcal{G}$ because $X = X_{\mathcal{G}}$ consists of bi-infinite paths. We also assume that any finite path in $\mathcal{G}$ can be extended to a bi-infinite path, i.e. a point in $X = X_{\mathcal{G}}$. Let A be the adjacent matrix of $\mathcal{G}$. By Lemma 2.4 (1),

$$p_X(n) = \sum_{i,j} (A^n)_{ij} \asymp \sum_{i,j} \left(TA^nT^{-1}\right)_{ij},$$

where T takes A to its Jordan normal form by conjugation. For a dimension k Jordan block $\lambda I + N$, we can calculate that $\sum_{i,j}((\lambda I + N)^n)_{ij} = \sum_{d=0}^{k-1}(k-d)\binom{n}{d}\lambda^{n-d}$. Thus if the eigenvalues of A are all 0 or 1, p_X grows polynomially, otherwise p_X grows exponentially.

Suppose $X = X_{\mathcal{G}}$ is topological mixing. By 2.4 (3), $\mathcal{G}$ is connected and aperiodic. Then there exists an n such that $(A^n)_{ij} > 0$ for any $i, j \in \mathcal{V}$. Unless A is the 1-all matrix, p_X grows exponentially. (The professor didn't give an explicit proof. But I think the Perron-Frobenius Theorem works.) □

Theorem 2.11 (Morse-Hedlund)**.** *Let X be a shift space. Then $|X| < \infty$ if and only if there exists some $n \in \mathbb{N}$ so that $p_X(n) \leqslant n$.*

Proof. ($\Rightarrow$): We first claim that any $x \in X$ is periodic. If not, $\infty = |\mathcal{O}(x)| \leqslant |X|$. We get a contradiction. Suppose x has period N. For any $n > N$, words with length n must contain some copies of the period words and be added by some finite choices of symbols at the front or back. In fact, we have $p_x(n) = p_x(n+N) = p_x(n+2N) = \cdots$. Thus there exists C_x such that $p_x \leqslant C_x$. Because $|X| < \infty$, there exists an universal constant C such that $p_X \leqslant \sum_{x \in X} p_x \leqslant C$. Then the proof follows.

($\Leftarrow$): Let n be the minimal number such that $p_X(n) \leqslant n$. If $n = 1$, $p_X(1) = 1$ and so $|X| = 1$. The theorem follows. Suppose $n > 1$. We have

$$n - 1 < p_X(n-1) \leqslant p_X(n) \leqslant n.$$

This shows that $p_X(n-1) = p_X(n) = n$. Denote $\mathcal{L}_n = \pi_{\{1,\cdots,n\}}(X) = \{w_1, \cdots, w_n\}$. We define $L\colon \mathcal{L}_n \to \mathcal{L}_{n-1}$ by forgetting the last symbol and $R\colon \mathcal{L}_n \to \mathcal{L}_{n-1}$ by forgetting the first symbol. Note that any word in $\mathcal{L}_{n-1}$ is the last or first of a length n word. Thus L, R are surjective. Note that $|\mathcal{L}_n| = \mathcal{L}_{n-1} = n$, L, R are also injective.

This shows that for any $w \in \mathcal{L}_{n-1}$, there is only one way of adding a symbol to the right or left to obtain a word in $\mathcal{L}_n$. Explicitly, for $w = (w_1, \cdots, w_{n-1})$, there exists unique symbol w_n such that $(w_1, \cdots, w_n) \in \mathcal{L}_n$. Then considering $(w_2, \cdots, w_n)$, we get unique w_{n+1} such that $(w_2, \cdots, w_{n+1}) \in \mathcal{L}_n$. Iterating this process and doing the same things on the left, we get an unique $x \in X$ such that $(x_1, \cdots, x_{n-1}) = w = (w_1, \cdots, w_{n-1})$. This shows that $|X| = |\mathcal{L}_{n-1}| = n$. To see this, for any $x \in X$, $(x_1, \cdots, x_{n-1}) \in \mathcal{L}_{n-1}$. And by uniqueness, if $x, x' \in X$ are with $(x_1, \cdots, x_{n-1}) = (x'_1, \cdots, x'_{n-1})$, then $x = x'$. □

Definition 2.12. *X is Sturmian if $p_X(n) = n + 1$ for all $n \in \mathbb{N}$.*

Example 2.13. $X_{\mathcal{G}}$ is Sturmian where

$$\mathcal{G} = \begin{tikzcd} 0 \arrow[loop left] \arrow[r] & 1 \arrow[loop right] \end{tikzcd}$$

Example 2.14. Let $\alpha \notin \mathbb{Q}$. We define $R_\alpha\colon \mathbb{T} \to \mathbb{T}$ by $x \mapsto x + \alpha$. Consider two intervals $J_1 = [0, 1-\alpha)$ and $J_2 = [1-\alpha, 1)$ as subsets of $\mathbb{T}$. We called a word $\underline{w} \in \{1,2\}^n$ allowed if $J_{\underline{w}} := J_{w_1} \cap R_\alpha^{-1}(J_{w_2}) \cap \cdots \cap R_\alpha^{-(n-1)}(J_{w_n}) \neq \varnothing$.

Claim. There are precisely $n+1$ allowed words of length n, the corresponding sets $J_{\underline{w}}$ are half-open intervals, the end points of these intervals are precisely $\{0, -\alpha, -2\alpha, \cdots, -n\alpha\}$.

Proof of Claim. For $n = 1$, the claim holds. We now assume that the claim holds for n. Let $\underline{w}$ be an allowed word with $J_{\underline{w}} = [a, b) \subset \mathbb{T}$. Note that

$$J_{\underline{w}} = \left\{x \in \mathbb{T} \text{ for which } \underline{w} \text{ describes the locations of } R_\alpha^j(x) \text{ for } j = 0, \cdots, n-1 \text{ with respect to } J_1 \text{ or } J_2\right\}.$$

The question of how the allowed word $\underline{w}$ extends corresponds to the question if $J_{\underline{w}} \supsetneqq J_{\underline{w}a}$ for $a \in \{1,2\}$.

- If $(n+1)\alpha \notin J_{\underline{w}}$, then $J_{\underline{w}} = J_{\underline{w}a}$ for some $a \in \{1,2\}$.
- If $(n+1)\alpha \in J_{\underline{w}}$, then $\underline{w}$ extends in two ways to allowed words.

$\square$

Hence the space of allowed words with the shift map σ is a Sturmian system.