Definition 1. Suppose that $X$ is a metric space. Let $f : X \to X$ and $x \in X$. Iterate $x$ to obtain the sequence
\[ x, f(x), f(f(x)), f(f(f(x))), \ldots. \]
A shorthand notation for these iterations will be $f^k$, where $k \in \{0, 1, 2, \ldots \}$. Note that $f^0$ is the identity map. The forward orbit of the point $x$ under $f$ is
\[ O^+_f(x) = O^+(x) = \{ x, f(x), f^2(x), \ldots \}. \]
In other words, the orbit is the range of the sequence $\{f^n(x)\}_{n=0}^\infty$. If $f$ is invertible (a bijection) then $f^{-1}$ exists and we can define
\[ f^{-n} = f^{-\leftarrow} \circ f^{-1} \circ \cdots \circ f^{-1} = (f^n)^{-1}. \]
Then $\{x, f^{-1}(x), f^{-2}(x), \ldots \}$ is called the backward orbit. The full orbit is defined by $\{f^n(x) : n \in \mathbb{Z}\}$. Now $(X, f)$ is called a discrete time dynamical system. Alternatively, we have the continuous time dynamical system: \{f^t : t \in \mathbb{R}^+ \} or \{f^t : t \in \mathbb{R} \}. Note that $f^{1+s} = f^1 \circ f^s$ and when $t \in \mathbb{R}$ we call this a flow.

Everything covered in this class will be in discrete time.

Example 2. (Rotations on a Circle $S^1$) We can define a rotation by $T^\lambda(x) = e^{i(\theta + 2\pi \lambda)}$. This map is well defined because we can write any point on the circle in the form of $e^{i\theta}$. We can think of this map as $\theta \mapsto \theta + 2\pi \lambda$ where $\lambda \in [0, 1]$. What do the orbits look like? In this case, we can describe all the iterates. Observe that
\[ T^\lambda(e^{i\theta}) = T^\lambda(e^{i(\theta + 2\pi \lambda)}) = e^{i(\theta + 4\pi \lambda)} = e^{i(\theta + (2\pi)(2\lambda))}. \]
Therefore $T^\lambda$ is the map $\theta \mapsto \theta + 2\pi(2\lambda)$. By induction we can show that $T^n\lambda(e^{i\theta}) = e^{i(\theta + (2\pi)(n\lambda))}$. Two cases now follow:

1. For the first case, assume $\lambda = \frac{p}{q}$, where $p, q$ are positive integers and relatively prime. Then
\[ T^\lambda(e^{i\theta}) = e^{i(\theta + 2\pi(q \frac{p}{q}))} = e^{i\theta} \]
and if $1 \leq k < q$, then $T^\lambda(e^{i\theta}) \neq e^{i\theta}$. We say $x = e^{i\theta}$ is a periodic point with period $q$.

2. For the second case, assume that $\lambda$ is irrational. We claim that the orbits are dense in the circle. (In this case we will have uncountable many infinite orbits.) To prove that each orbit is dense, we will show first that no point is periodic. For a contradiction, suppose there exists a periodic point $x$. Let $k < n$ be such that $T^\lambda(x) = T^n\lambda(x)$. Then $e^{i(\theta + 2\pi n\lambda)} = e^{i(\theta + 2\pi k \lambda)}$ where $x = e^{i\theta}$. Therefore $\theta + 2\pi n\lambda - (\theta + 2\pi k \lambda) = 2\pi j$ for some integer $j$. So $2\pi \lambda(n - k) = 2\pi j$ or $\lambda(n - k) = j$, a contradiction as $\lambda$ is irrational. Thus each orbit is infinite. Now from this, we can show that each orbit is dense in $S^1$.

Claim. Each forward orbit is dense in the circle $S^1$.

Proof. Let $p \in S^1$. We have seen that $O^+(p)$ is an infinite set. Because the circle is compact there exists an accumulation point $w$ of $O^+(p)$. Let $\epsilon > 0$. There exist positive integers $j$ and $k$, with $j < k$, such that $d(T^k\lambda(p), T^j\lambda(p)) < \epsilon$. Set $x = T^k\lambda(p)$ and $n = k - j$. Then
\[ T^n\lambda(x) = T^{k-j}\lambda(p) = T^k\lambda(p). \]
So $d(T^n\lambda(x), x) < \epsilon$. Now, $T\lambda$ has the special property of preserving distance (isometry), so that $T^n\lambda$ has this same property. Consider the points
\[ x, T^n\lambda(x), T^{3n}\lambda(x), \ldots. \]
Every point of $S^1$ is within a distance of $\epsilon$ of one of these points and each of these points is in $O^+(p)$. Therefore $O^+(p)$ is dense in $S^1$. 

So in this example we were able to completely describe the orbits. This will not always be the case as seen in the example below.
Example 3. (We will not prove this now.) Let $f : [0, 1] \to [0, 1]$ and $f(x) = 4x(1-x)$. Then

1. The periodic points are dense.
2. For each $n \geq 1$, $f$ has a periodic point of period $n$.
3. There exists points with dense forward orbit.
4. There exists points with infinite orbits which are not dense.

The following map also has the same properties: $g : S^1 \to S^1$ where $g(x) = x^2$. This map is called the angle doubling map.

Definition 4. Let $f : X \to X$ be continuous where $X$ is a metric space. Let $x \in X$. We define the $\omega$-limit set of $x$, denoted $\omega(x)$ or $\omega_f(x)$ or $\omega(x, f)$, as the set of subsequential limits of the sequence $\{f^n(x)\}$ (i.e. $y \in \omega(x)$ if and only if there exist $N \ni n_k \to \infty$ such that $f^{n_k}(x) \to y$).

Note: $\omega(x) = \bigcap_{k \geq 0} \{f^k(x), f^{k+1}(x), f^{k+2}(x), \ldots\}$. In particular, $\omega(x)$ is closed, and $\omega(f(x)) = \omega(x)$. Also, if $X$ is compact, $\omega(x) \neq \emptyset$, and $f(\omega(x)) = \omega(x)$.

Example 5. Consider a rotation of the circle $T_\lambda : S^1 \to S^1$.

1. For a rational rotation, $\omega(x) = O^+(x)$.
2. For an irrational rotation, $\omega(x) = S^1$.
3. In both cases above $\omega(x) = \overline{O^+(x)}$. It is always true that $\omega(x) \subseteq \overline{O^+(x)}$.

Definition 6. Let $f : X \to X$ be continuous where $X$ is a metric space. Let $x \in X$. We say that $x$ is recurrent if and only if for every open set $V$ with $x \in V$, we have $f^k(x) \in V$ for some positive integer $k$.

Theorem 7. The following are equivalent.

1. $x$ is recurrent.
2. For every open set $V$ with $x \in V$, we have $f^k(x) \in V$ for infinitely many $k$.
3. There exists a sequence $n_k \to \infty$ such that $f^{n_k}(x) \to x$.
4. $x \in \omega(x)$.

Theorem 8. $x$ is recurrent if and only if $\omega(x) = \overline{O^+(x)}$.

Definition 9. Let $f : X \to X$. Let $E \subset X$. We say that $E$ is invariant (or $f$-invariant or forward invariant) if and only if $f(E) \subseteq E$.

Remark 10. Let $X$ be a metric space. Suppose $f : X \to X$ is continuous.

1. If $E$ is invariant then $\bar{E}$ is invariant, $f(\bar{E}) \subseteq \overline{f(E)} \subseteq \bar{E}$.
2. If $E$ is invariant and $x \in E$, then $O^+(x) \subseteq E$.
3. If $E$ and $F$ are invariant, then $E \cap F$ and $E \cup F$ are invariant.
4. If $E$ is invariant $X - E$ need not be invariant.

Definition 11. Let $X$ be a metric space. Suppose $f : X \to X$ is continuous. We say that $f$ is strongly topologically transitive if and only if for every pair of non-empty open subsets $U$ and $V$ of $X$, there exists a positive integer $n$ such that $f^n(U) \cap V \neq \emptyset$.

Definition 12. We say $f$ is forward orbit topologically transitive if and only if there exists $x \in X$ with $\overline{O^+(x)} = X$. If $f$ is a homeomorphism, we say $f$ is full orbit topologically transitive if and only if there exists $x \in X$ with $\overline{O(x)} = X$. 


Proposition 13. Let $X$ be a metric space and suppose $f : X \to X$ is continuous. Then the following are equivalent.

1. $f$ is strongly topologically transitive.

2. For every non-empty open subset $V$ of $X$, $\bigcup_{n=1}^{\infty} f^n(V)$ is dense in $X$.

3. For every non-empty open subset $V$ of $X$, $\bigcup_{n=1}^{\infty} f^{-n}(V)$ is dense in $X$.

4. For every pair $U$ and $V$ of non-empty open subsets of $X$, $f^{-n}(U) \cap V \neq \emptyset$ for some $n > 0$.

5. Every proper closed invariant subset of $X$ has empty interior.

Here $f$ need not be 1-1. Recall, by definition we have $f^{-n}(V) = (f^n)^{-1}(V)$.

Corollary 14. Suppose $f : X \to X$ is a homeomorphism. Then $f$ is strongly topologically transitive if and only if $f^{-1}$ is strongly topologically transitive.

Proposition 15. Let $f : X \to X$, where $X$ is a metric space and $f$ is continuous. Suppose $X$ is compact and $f$ is strongly topologically transitive. Then $f$ is surjective.

Proposition 16. Suppose $X$ is a compact metric space and $f : X \to X$ is continuous. Then $f$ is strongly topologically transitive if and only if there exists $x \in X$ with $\omega(x) = X$. In this case, $f$ is also a forward orbit topologically transitive.

Proposition 17. Suppose $X$ is a compact metric space and $f : X \to X$ is continuous. Suppose that $f$ is surjective. Assume that $x \in X$ satisfies $\overline{O^{+}(x)} = X$. Then

1. $x$ is recurrent, so that $\omega(x) = X = \overline{O^{+}(x)}$.

2. If $y \in O^{+}(x)$, then $y$ is recurrent and $\omega(y) = \overline{O^{+}(y)} = X$.

3. $f$ is strongly topologically transitive.

Proposition 18. Suppose $X$ is a compact metric space and $f : X \to X$ is continuous. Suppose each $x \in X$ is an accumulation point of $X$. Assume that $x \in X$ satisfies $\overline{O^{+}(x)} = X$. Then $\omega(x) = X$. Also, $f$ is surjective.

Proposition 19. Suppose $X$ is a compact metric space and $f : X \to X$ is a homeomorphism. Suppose each $x \in X$ is an accumulation point of $X$. Assume that there exists $x \in X$ with $\overline{O(x)} = X$ (i.e. $f$ is full-orbit topologically transitive). Then $f$ is strongly topologically transitive.

Example 20. In the previous Proposition, it is necessary that $x$ is an accumulation point. Why? Consider the map $f : [0, 1] \to [0, 1]$ given by $f(x) = x^2$. Suppose $X = O(f(\frac{1}{2}))$ and $g : X \to X$ is given by $g = f|_X$. Now there are two accumulation points: 0 and 1. Observe that $g$ is full orbit topologically transitive, but not forward orbit topologically transitive.

Example 21. Consider $f : \{0, 1\} \to \{0, 1\}$ defined by $f(1) = 1$ and $f(0) = 1$. Then $f$ is forward orbit topologically transitive, but not strongly topologically transitive.

Proposition 22. Suppose $X$ is a compact metric space and $f : X \to X$ is continuous. Then $f$ is strongly topologically transitive if and only if there exists $x \in X$ such that $\overline{O^{+}(f(x))} = X$.

Definition 23. Let $X$ be a compact metric space and suppose $M \subseteq X$. Let $f : X \to X$ be continuous. We say that $M$ is a minimal set for $f$ if and only if $M$ is nonempty, closed, invariant and no proper subset of $M$ has these 3 properties.

It is easy to see that this is equivalent to saying the $M$ is non-empty and $M = \overline{O^{+}(x)}$ for all $x \in M$ or $M$ is nonempty and $M = \omega(x)$ for all $x \in M$. We also say that $f|_M$ is minimal and in this case $f$ is strongly topologically transitive.

Definition 24. Suppose that $f : X \to X$ is continuous, where $X$ compact metric space. Let $x \in X$. We say that $x$ is almost periodic (or strongly recurrent or uniformly recurrent) if and only if for any open set $W$ with $x \in W \{n \geq 1 : f^n(x) \in W\}$ (the return times) is infinite and has bounded gaps (or relatively dense) (i.e. if $\{n \geq 1 : f^n(x) \in W\} = \{n_1 < n_2 < n_3 < \ldots\}$ then there exists a positive integer $k$ such that $n_j - n_{j-1} \leq k$ for all $j$).
Theorem 25. Suppose that \( f : X \to X \) is continuous, and \( X \) is a compact metric space. Assume that \( X \) is a minimal set for \( f \). Then each point \( x \in X \) is almost periodic.

Theorem 26. Suppose that \( X \) is a compact metric space, and let \( f : X \to X \) be continuous. Let \( x \in X \), and suppose that \( x \) is almost periodic. Then \( \overline{O^+(x)} \) is a minimal set.

Corollary 27. Suppose that \( X \) is a compact metric space, and \( f : X \to X \) is continuous. Let \( x \in X \), and suppose that \( x \) is almost periodic. Then \( O^+(x) \) is a minimal set.

Corollary 28. A collection \( D = \{ D_i : i \in I \} \) of subsets \( X \) has the finite intersection property if and only if for any finite subset \( F \) of \( D \) the intersection of the sets in \( F \) is nonempty.

Theorem 29. \( X \) is compact if and only if the following condition holds: If \( D = \{ D_i : i \in I \} \) is a collection of closed subsets of \( X \) and \( D \) has the finite intersection property then \( \cap_{i \in I} D_i \neq \emptyset \).

Theorem 30. Suppose that \( X \) is a nonempty compact metric space. Let \( f : X \to X \) be continuous. Then there exists a minimal set for \( f \).

Proof. Let \( C \) denote the collection of all nonempty, closed, invariant subsets of \( X \). Then \( C \neq \emptyset \), as \( X \in C \). Partially order \( C \) by inclusion (i.e. \( A \subset B \)). Let \( D = \{ D_i : i \in I \} \) be a totally ordered subset of \( C \).

We claim that \( D \) as the finite intersection property. Let \( F = \{ D_{i_1}, D_{i_2}, \ldots, D_{i_k} \} \) be a finite subset. Some \( D_{i_j} \) is a subset of each element in \( F \). So \( D_{i_1} \cap D_{i_2} \cap \cdots \cap D_{i_k} \neq \emptyset \). This proves the claim.

Then \( E = \cap_{i \in I} D_i \neq \emptyset \). Thus \( E \) is a lower bound for \( D \). By Zorn’s lemma there exists a minimal element \( C \) (with respect to \( \subset \)).

Corollary 31. Suppose that \( X \) is compact and \( f : X \to X \) is continuous. Then every nonempty closed invariant subset of \( X \) contains a minimal set.

Proof. Let \( K \) be a nonempty closed invariant subset. Consider \( f|_K : K \to K \). This is a continuous map of a compact metric space. The previous theorem then applies.

Definition 32. Suppose \( X \) and \( Y \) are topological spaces. Let \( f : X \to X \) and \( g : Y \to Y \) be continuous. We say that \( f \) and \( g \) are topologically conjugate if and only if there exists a homeomorphism \( h : X \to Y \) such that the diagram

\[
\begin{array}{ccc}
X & \xrightarrow{f} & X \\
\downarrow h & & \downarrow h \\
Y & \xrightarrow{g} & Y
\end{array}
\]

commutes. In other words, \( h(f(x)) = g(h(x)) \) for all \( x \in X \) or \( h \circ f = g \circ h \). We say that \( h \) is a topological conjugacy from \( f \) to \( g \).

We can repeat this diagram, seen below, to see that \( h \) is also a topological conjugacy from \( f^2 \) to \( g^2 \).

\[
\begin{array}{ccc}
X & \xrightarrow{f} & X & \xrightarrow{f} & X \\
\downarrow h & & \downarrow h & & \downarrow h \\
Y & \xrightarrow{g} & Y & \xrightarrow{g} & Y
\end{array}
\]

Proposition 33. Suppose \( X \) and \( Y \) are homeomorphic, and \( f : X \to X \) is continuous. Then there exists \( g : Y \to Y \) such that \( f \) and \( g \) are topologically conjugate.
Definition 34. Let \( f : X \to X \) be continuous. A point \( x \in X \) is \textit{wandering} if and only if there exists an open set \( V \) containing \( x \) such that \( f^k(V) \cap V = \emptyset \) for \( k = 1, 2, 3, \ldots \). If \( x \) is not wandering, we say that \( x \) is \textit{nonwandering}. The set of nonwandering points is called the \textit{nonwandering set} and is denoted \( NW(f) \) or \( \Omega(f) \).

Proposition 35. The set of nonwandering points, \( NW(f) \), has the following properties.

1. \( NW(f) \) is a closed, invariant set.
2. If \( M \) is a minimal set \( M \subset NW(f) \).
3. If \( x \in X, \omega(x) \subset NW(f) \).
4. If \( X \) is compact, \( NW(f) \neq \emptyset \).
5. Let \( \text{Per}(f), \text{AP}(f) \) and \( R(f) \) denote the set of periodic points of \( f \), the almost periodic points of \( f \) and the recurrent points of \( f \), respectively. We have \( \text{Per}(f) \subset \text{AP}(f) \subset R(f) \subset R(f) \subset NW(f) \).

Example 36. Consider the map \( f : [0, 1] \to [0, 1] \) defined by \( f(x) = x^2 \). We claim that all \( x \in (0, 1) \) are wandering. Since \( f(x) < x \) we can find an open interval \( V \) containing \( x \) such that \( f(V) \cap V = \emptyset \). Because \( f(y) < y \) for all \( y \in (0, 1) \) we can repeat this argument and see that all points in \( (0, 1) \) are wandering. Therefore \( x = 0 \) and \( x = 1 \) are the only nonwandering points.

Definition 37. A continuous map \( f : X \to X \) is \textit{topologically mixing} if and only if for every pair \( V, W \) of nonempty open subsets of \( X \), there exists a positive integer \( K \) such that for all \( k \geq K \) we have \( f^k(V) \cap W \neq \emptyset \).

Proposition 38. If \( f \) is topologically mixing, then \( f^j \) is strongly topologically transitive for \( j = 1, 2, 3, \ldots \).

Example 39. Consider an irrational rotation \( f \) on \( S^1 \). Then \( f^j \) is strongly topologically transitive for each positive integer \( j \), but \( f \) is not topologically mixing.

Example 40. Define a map \( f : [0, 1] \to [0, 1] \) by

\[
f(x) = \begin{cases} 
 3x, & \text{if } 0 \leq x \leq \frac{1}{3} \\
 1, & \text{if } \frac{1}{3} \leq x \leq \frac{2}{3} \\
 -3(x - 1), & \text{if } \frac{2}{3} \leq x \leq 1 
\end{cases}
\]

Then the graph of \( f \) is as follows

```
1

\( \frac{1}{3} \quad \frac{2}{3} \quad 1 \)
```

Set

\[
W = \bigcup_{n=0}^{\infty} f^{-n} \left( \frac{1}{3}, \frac{2}{3} \right).
\]

We can observe that if \( f(x) \in W \), then \( x \in W \). So

\[
W = \left( \frac{1}{3}, \frac{2}{3} \right) \cup \left( \frac{1}{9}, \frac{2}{9} \right) \cup \left( \frac{7}{9}, \frac{8}{9} \right) \cup \left( \frac{1}{27}, \frac{2}{27} \right) \cup \ldots
\]
and \( W \) is the complement of \( C \), where \( C \) is the middle-third Cantor set. It follows that \( f(C) \subset C \). In fact, \( f(C) = C \). Moreover \( f|_C : C \to C \) is two-to-one. Let’s describe the dynamics of this two-to-one map. We will do this by encoding the intervals. Let \( [0, \frac{1}{3}] \) be interval 1 and \([\frac{2}{3}, 1]\) be interval 2. To each \( x \in C \) we associate a sequence of ones and twos given as follows. We associate the sequence \( a_0, a_1, a_2, \ldots \) where
\[
 a_k = \begin{cases} 
 1, & \text{if } f^k(x) \in [0, \frac{1}{3}] \\
 2, & \text{if } f^k(x) \in [\frac{2}{3}, 1]
\end{cases}
\]
We will let \( \Sigma^+ \) denote the set of all such sequences. We have a map \( h : C \to \Sigma^+ \). If can be verified that \( h \) is a bijection. In fact, \( h \) will be a homeomorphism (with product topology on \( \Sigma^+ \)).

\[
\begin{array}{c}
 C \\
\hline
 h \\
\Sigma^+ \quad \text{-----} \quad \Sigma^+
\end{array}
\]

Then the map of \( \Sigma^+_2 \to \Sigma^+_2 \) given by \( h \circ f|_C \circ h^{-1} \) is topologically conjugate to \( f|_C \). If \( h \circ f|_C \circ h^{-1} = \sigma \), then \( \sigma : \Sigma^+_2 \to \Sigma^+_2 \) is given by \( \sigma(a_0, a_1, a_2, \ldots) = (a_1, a_2, \ldots) \). This is the map we will study in symbolic dynamics.

**Definition 41.** Let \( \Sigma^+_n \) denote the set of all sequences \( \alpha = (a_0, a_1, a_2, \ldots) \) where each \( a_i \in \{1, 2, \ldots, n\} \). (The topology on this set of symbols is the discrete topology). One metric \( d \) on \( \Sigma^+_n \) is as follows: If \( \alpha = \beta \) then \( d(\alpha, \beta) = 0 \). If \( \alpha \neq \beta \), let \( m \) denote the smallest nonnegative integer such that \( a_m \neq b_m \) where \( \alpha = (a_0, a_1, a_2, \ldots) \) and \( \beta = (b_0, b_1, b_2, \ldots) \). Set \( d(\alpha, \beta) = \frac{1}{2^m} \).

Another metric is given by
\[
d(\alpha, \beta) = \sum_{i=0}^{\infty} \frac{\delta(a_i, b_i)}{2^i}
\]
where
\[
\delta(a_i, b_i) = \begin{cases} 
 0, & \text{if } a_i = b_i \\
 1, & \text{if } a_i \neq b_i
\end{cases}
\]

It can be verified that both of these are metrics. These metrics give the same topology, the product topology. We have a map \( \sigma : \Sigma^+_n \to \Sigma^+_n \) where \( \sigma(a_0, a_1, a_2, \ldots) = (a_1, a_2, a_3, \ldots) \). An alternate description of this map is if \( x = (a_0, a_1, a_2, \ldots) \), we may write \( (\sigma(x))_i = x_{i+1} \). The pair \((\Sigma^+_n, \sigma)\) is called the **full one-sided shift** on \( n \) symbols.

Let \( \Sigma_n \) denote the set of doubly infinite sequences \( (\ldots, a_{-2}, a_{-1}, a_0, a_1, a_2, \ldots) \), where each \( a_i \in \{1, 2, \ldots, n\} \). One metric we could place on this space is \( d(\alpha, \beta) = \frac{1}{2^m} \), where \( m \) is the smallest nonnegative integer such that either \( a_m \neq b_m \) or \( a_{-m} \neq b_{-m} \). We have a map \( \sigma : \Sigma_n \to \Sigma_n \) where \( \sigma(\ldots, a_{-2}, a_{-1}, a_0, a_1, a_2, \ldots) = (\ldots, a_{-1}, a_0, a_1, a_2, \ldots) \). An alternate description of this map is if \( x = (\ldots, a_{-2}, a_{-1}, a_0, a_1, a_2, \ldots) \), we may write \( (\sigma(x))_i = x_{i+1} \). The pair \((\Sigma_n, \sigma)\) is called the **full two-sided shift** on \( n \) symbols.

**Proposition 42.**
1. Let \( \epsilon > 0 \). There exists \( k \in \mathbb{Z}^+ \) such that if \( \alpha, \beta \in \Sigma^+_n \), \( \alpha = (a_0, a_1, a_2, \ldots) \), \( \beta = (b_0, b_1, b_2, \ldots) \), and \( a_i = b_i \) for \( i = 0, 1, \ldots, k \), then \( d(\alpha, \beta) < \epsilon \).
2. Let \( k \in \mathbb{Z}^+ \). There exists \( \epsilon > 0 \) such that if \( \alpha, \beta \in \Sigma^+_n \) and \( \alpha = (a_0, a_1, a_2, \ldots) \), \( \beta = (b_0, b_1, b_2, \ldots) \), and \( d(\alpha, \beta) < \epsilon \), then \( a_i = b_i \) for \( i = 0, 1, \ldots, k \).

**Definition 43.** Let \( x_0, x_1, \ldots, x_k \) be a finite sequence of elements of \( \{1, 2, \ldots, n\} \) of length \( k + 1 \). We call such a sequence a **word** in \( \Sigma^+_n \) of length \( k + 1 \). We define a **cylinder set** corresponding to this word by
\[
 C_{x_0, x_1, \ldots, x_k} = \{ \alpha \in \Sigma^+_n : \text{if } \alpha = (a_0, a_1, \ldots) \text{ we have } a_0 = x_0, a_1 = x_1, \ldots, a_k = x_k \}.
\]

**Proposition 44.** The following hold:
1. Each cylinder set is clopen.
2. For a fixed $k$, there exists $n^{k+1}$ distinct cylinder sets corresponding to words of length $k+1$. These sets partition $\Sigma^+_n$.

3. The collection of all cylinder sets forms a basis for the topology.

4. $\Sigma^+_n$ is compact.

**Definition 45.** Let $w = x_0, x_1, \ldots, x_k$ be a word in $\Sigma^+_n$. Let $\alpha = (a_0, a_1, \ldots) \in \Sigma^+_n$. We say that $w$ appears in $\alpha$ if and only if there exists $j \geq 0$ such that $a_j = x_0, a_{j+1} = x_1, \ldots, a_{j+k} = x_k$ (i.e. for some $j \geq 0$, $\sigma^j(\alpha) \in C_{x_0, x_1, \ldots, x_k}$.) We say $w$ is an initial word in $\alpha$ if and only if the above occurs with $j = 0$.

**Proposition 46.** Let $\alpha, \beta \in \Sigma^+_n$. Then $\alpha \in O^+(\beta)$ if and only if every word which appears in $\alpha$ also appears in $\beta$. (Here, the map is understood to be the shift map.)

**Proposition 47.** The set of periodic points of $\sigma$, the one-sided shift map, are dense in $\Sigma^+_n$.

**Proposition 48.** The map $\sigma : \Sigma^+_n \to \Sigma^+_n$ is strongly topologically transitive.

**Proposition 49.** Consider the map $\sigma : \Sigma^+_n \to \Sigma^+_n$. For any positive integer $k$, the number of fixed points of $\sigma^k$ is precisely $n^k$.

**Proposition 50.** Let $\alpha, \beta \in \Sigma^+_n$. Then $\alpha \in \omega(\beta)$ if and only if the following holds: Let $w = x_0, x_1, \ldots, x_k$ be a word which appears in $\alpha$. Then for any positive integer $j$, $w$ appears in $\sigma^j(\beta)$. Informally, $\alpha \in \omega(\beta)$ if and only if every word which appears in $\alpha$ also appears infinitely many times in $\beta$.

**Definition 51.** Let $A$ be an $n \times n$ matrix with all entries 0 or 1. We will assume that in each row, there is at least one 1 (i.e. for each fixed $i$, $\sum_{j=1}^n A_{ij} \geq 1$). We will call such a matrix a transition matrix. Associated to $A$ we have a transition graph, a directed graph with vertices $1, 2, \ldots, n$ and an edge from $i$ to $j$ if and only if $A_{ij} = 1$. If $e$ is an edge in the graph, then $e$ has an initial vertex and a terminal vertex, denoted $I(e)$ and $T(e)$ respectively. If $x$ and $y$ are vertices, by a path from $x$ to $y$ we mean a sequence of edges $e_1, e_2, \ldots, e_k$ where $I(e_1) = x$, $T(e_1) = I(e_2)$, $T(e_2) = I(e_3)$, ..., $T(e_{k-1}) = I(e_k)$, $T(e_k) = y$.

**Definition 52.** We let $\Sigma_A$ denote the subset of $\Sigma^+_n$ consisting of all $\alpha = (s_0, s_1, s_2, \ldots) \in \Sigma^+_n$ such that $A_{s_i, s_{i+1}} = 1$ for each $i = 0, 1, 2, \ldots$ (i.e. all $\alpha = (s_0, s_1, \ldots) \in \Sigma^+_n$ such that for each $i = 0, 1, 2, \ldots$ there is an edge from vertex $s_i$ to the vertex $s_{i+1}$ in the transition graph.)

**Definition 53.** The subset $\Sigma_A$ is a closed, invariant subset of $\Sigma^+_n$.

**Definition 54.** The map $\sigma |_{\Sigma_A} : \Sigma_A \to \Sigma_A$ is called the subshift of finite type associated to $A$, and is denoted by $\sigma_A$.

**Example 55.** Consider the following examples.

1. Suppose that

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

Then $\Sigma_A$ has 2 elements $(1, 1, \ldots)$ and $(2, 2, \ldots)$. These are fixed by $\sigma_A$.

2. Suppose that

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

Then $\Sigma_A$ has 2 elements, $(1, 2, 1, 2, \ldots)$ and $(2, 1, 2, 1, \ldots)$ of orbit of period 2.

3. Suppose that

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

This is a full one-sided shift.
4. Suppose that

\[ A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \]

Then 1 can go to 1 or 2, but 2 can only go to 2. The elements of \( \Sigma_A \) \((1, 1, 1, \ldots) \) and \((2, 2, 2, \ldots) \) are fixed points. We could also have

\[ (1, 1, 1, \ldots, 1, 2, 2, 2, \ldots). \]

So \( \Sigma_A \) is countably infinite, and has two fixed points \( x, y \). For all other \( z \in \Sigma_A \), \( \sigma_A^k(z) = y \) for some \( k \geq 1 \).

5. Suppose that

\[ A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \]

Then 1 can go to 1 or 2, but 2 can only go to 1. It can be shown that \( \Sigma_A \) is uncountable and that each element of \( \Sigma_A \) is an accumulation point of \( \Sigma_A \). The point \((1, 1, \ldots)\) is a fixed point and the point \((1, 2, 1, 2, \ldots)\) has period 2. The point \((1, 1, 2, 1, 1, 2, 1, 2, \ldots)\) has period 3. For \( k \geq 3 \), we can construct a point of period \( k \) by considering

\[ (1, 1, 1, \ldots, 1, 2, 2, 2, \ldots). \]

**Definition 56.** Let \( A \) be a transition matrix. We say that \( A \) is **irreducible** if and only if for all \( i, j \in \{1, 2, \ldots, n\} \), there exists \( k \geq 1 \) such that \( (A^k)_{i,j} > 0 \).

**Proposition 57.** Let \( A \) be a transition matrix. \( A \) is irreducible if and only if for all \( i, j \in \{1, 2, \ldots, n\} \) there exists a path in the transition matrix graph from vertex \( i \) to vertex \( j \). In fact, \( (A^k)_{i,j} \) equals the number of paths from \( i \) to \( j \) of length \( k \).

**Proposition 58.** Let \( A \) be a transition matrix. The trace of \( A^k \) is the number of fixed points of \( (\sigma_A)^k \).

**Definition 59.** Let \( w = x_0, x_1, \ldots, x_k \) be a word in \( \Sigma_n^+ \). Let \( A \) be a transition matrix. We say that the word \( w \) is **A-allowable** if and only if \( a_{x_i, x_{i+1}} = 1 \) for all \( i = 1, \ldots, k - 1 \).

**Proposition 60.** Let \( A \) be a transition matrix. The number of \( A \)-allowable words \( x_0, x_1, \ldots, x_k \) of length \( k + 1 \) is equal to the number of paths in the transition graph of length \( k \).

**Theorem 61.** Let \( A \) be a transition matrix. Then \( A \) is irreducible if and only if \( \sigma_A \) is strongly topologically transitive.

**Example 62.** Let’s us consider \( \sigma : \Sigma_2^+ \rightarrow \Sigma_2^+ \), where the symbol set is \( \{0, 1\} \). We will construct a particular element \( \alpha \in \Sigma_2^+ \). This sequence is sometimes called the Morse-Thue sequence. The sequence \( \alpha = (a_0, a_1, a_2, \ldots) \) may be defined in two ways:

1. Let \( n = \sum_{i=0}^{\infty} n_i 2^i \), where each \( n_i \) is 0 or 1. Set \( a_n = \sum_i n_i \mod 2 \).
2. If \( A_k \) is the initial word of \( \alpha \), consisting of the first \( 2^k \) symbols, then set \( A_0 = 0 \) and \( A_{k+1} = A_k \bar{A}_k \), where \( \bar{A}_k \) is obtained from \( A_k \) by replacing 1 by 0 and 0 by 1. From this we can write down any finite portion of \( \alpha \). Thus \( \alpha = (0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, \ldots) \).

**Proposition 63.** Under the shift map, \( \alpha \) is almost periodic, but not periodic.

**Definition 64.** The closure of the orbit of \( \alpha \) is known as the **Morse minimal set**.

**Notation 65.** Suppose that \( X \) is a compact topological space and \( f : X \rightarrow X \) is continuous. We will associate to \( f \) a non-negative extended real number called the **topological entropy** of \( f \). Let \( \alpha, \beta \) denote open covers of \( X \).

1. \( \alpha \vee \beta \) denotes \( \{ A \cap B : A \in \alpha, B \in \beta \} \). Observe that \( \alpha \vee \beta \) is an open cover.
2. \( N(\alpha) \) is the minimum cardinality of a subcover of \( \alpha \).
3. \( H(\alpha) = \log (N(\alpha)) \).
4. We say that $\beta$ is a refinement of $\alpha$, denoted $\alpha < \beta$, if and only if every open set in $\beta$ is contained in some open set in $\alpha$.

5. $f^{-1}(\alpha)$ is the open cover consisting of all sets $f^{-1}(A)$ where $A \in \alpha$.

**Proposition 66.** With the notation above, the following are true.

1. If $\alpha < \beta$, then $H(\alpha) \leq H(\beta)$ and $H(\alpha \vee \beta) = H(\beta)$.
2. $H(\alpha \vee \beta) \leq H(\alpha) + H(\beta)$.
3. If $\alpha < \beta$, then $f^{-1}(\alpha) < f^{-1}(\beta)$.
4. $f^{-1}(\alpha \vee \beta) = f^{-1}(\alpha) \vee f^{-1}(\beta)$.
5. $H(f^{-1}(\alpha)) \leq H(\alpha)$ with equality if $f$ is surjective.

**Lemma 67.** Let $\{a_n\}$ be a sequence of real numbers which is subadditive (i.e. $a_{n+k} \leq a_n + a_k$ for all positive integers $n$ and $k$). Then $\lim_{n \to \infty} \frac{a_n}{n}$ exists and equals $\inf \{\frac{a_n}{n}\}$ (this limit may be $-\infty$).

**Lemma 68.** If $\alpha$ is an open cover of the compact topological space $X$, then

$$
\lim_{n \to \infty} \frac{1}{n} H(\alpha \vee f^{-1}(\alpha) \vee \cdots \vee f^{-n+1}(\alpha))
$$

exists and is finite. Note that $f^{-n+1}(\alpha) = f^{-(n+1)}(\alpha) = (f^n)^{-1}(\alpha)$.

**Definition 69.** Set

$$
h(f, \alpha) = \lim_{n \to \infty} \frac{1}{n} H(\alpha \vee f^{-1}(\alpha) \vee \cdots \vee f^{-n+1}(\alpha)).
$$

We define the **topological entropy** of $f$ by $h(f) = \sup h(f, \alpha)$ over all open covers.

**Proposition 70.** If $\alpha < \beta$, then $h(f, \alpha) \leq h(f, \beta)$.

**Proposition 71.** Suppose $\alpha_n$ is a sequence of open covers such that

1. $\alpha_n < \alpha_{n+1}$ for each $n$;
2. For any open cover $\beta$, there exists $n$ such that $\beta < \alpha_n$;

Then $h(f) = \lim_{n \to \infty} h(f, \alpha_n)$.

**Corollary 72.** Suppose $(X, d)$ is a compact metric space. Suppose $\alpha_n$ is a sequence of open covers such that

1. $\alpha_n < \alpha_{n+1}$ for each $n$;
2. $d(\alpha_n) \to 0$ as $n \to \infty$, where $d(\alpha_n) = \sup \{d(A) : A \in \alpha_n\}$. Sometimes $d(\alpha)$ is called the mesh of $\alpha_n$;

Then $h(f) = \lim_{n \to \infty} h(f, \alpha_n)$.

**Example 73.** Let $\sigma : \Sigma_2^+ \to \Sigma_2^+$ be the full shift map. Then $h(\sigma) = \log 2$.

**Theorem 74.** Let $f : X \to X$ be a shift (i.e. $X$ is a subset of $\Sigma_n^+$, and $f$ is the restriction of the full one-sided shift on $n$ symbols $\sigma : \Sigma_n^+ \to \Sigma_n^+$, and $X$ is closed and invariant under $\sigma$). Let $M_k$ denote the number of words of length $k$ that appear as the initial word of an element of $X$. Then $h(f) = \lim_{k \to \infty} \frac{1}{k} \log M_k$.

**Definition 75.** Recall that that eigenvalues of a square matrix $A$ are the (real or complex) solutions of the equation $\det(A - \lambda I) = 0$.

**Theorem 76.** (Perron-Frobenius) Let $A$ be a square matrix with entries non-negative real numbers. There exists an eigenvalue $\lambda$ of $A$ such that

1. $\lambda$ is real and $\lambda \geq 0$;
2. For every eigenvalue $\mu$ of $A$, $|\mu| \leq \lambda$.

Furthermore, $\lambda = \lim_{n \to \infty} |A^n|^{\frac{1}{n}} = \lim_{n \to \infty} \left( \text{Trace}(A^n) \right)^{\frac{1}{n}}$, where $|A| = \sum a_{i,j}$ and $\text{Trace}(A) = \sum a_{i,i}$.

We will call $\lambda$ the maximal eigenvalue of $A$.

**Theorem 77.** Let $A$ be a transition matrix, and let $\sigma_A$ denote the corresponding (one-sided) subshift of finite type associated to $A$. Let $\lambda$ denote the maximal eigenvalue of $A$. Then $h(\sigma_A) = \log \lambda$.

**Proposition 78.** For any positive integer $k$, $h(f^k) = k \cdot h(f)$.

**Proposition 79.** Suppose that $f : X \to X$ is continuous and $X$ is compact. Suppose $K$ is a closed invariant subset of $X$. Then $h(f) \geq h(f|_K)$.

**Proposition 80.** Suppose that $X$ is compact. If $f : X \to X$ is homeomorphism, then $h(f) = h(f^{-1})$.

**Theorem 81.** Let $X$ and $Y$ be compact spaces. Let $f : X \to X$ and $g : Y \to Y$ be continuous. Suppose that there exist a continuous, surjective map $\varphi : X \to Y$ such that the diagram

\[
\begin{array}{ccc}
X & \xrightarrow{f} & X \\
\varphi \downarrow & & \varphi \\
Y & \xrightarrow{g} & Y
\end{array}
\]

commutes (i.e. $\varphi(f(x)) = g(\varphi(x))$ for all $x \in X$). Then $h(f) \geq h(g)$. Moreover, if $\varphi$ is a homeomorphism, then $h(f) = h(g)$.

**Remark 82.** The map $\varphi$ in Theorem 81 is called a **semi-conjugacy**. From Theorem 81, we see that topological entropy is an invariant of topological conjugacy.

**Example 83.** There exists a continuous map $f : [0, 1] \to [0, 1]$ with $h(f) = \infty$.

**Definition 84.** Let $(X, d)$ be a compact metric space and let $f : X \to X$ be continuous.

1. For $n \geq 1$ and $\epsilon > 0$, we call a finite subset $A$ of $X$ an **$(n, \epsilon)$-separated set** if and only if for any pair of distinct points $x, y \in A$ we have $d(f^k(x), f^k(y)) > \epsilon$ for some integer $k$ with $0 \leq k \leq n - 1$. Let $s(n, \epsilon)$ denote the maximum cardinality of any $(n, \epsilon)$-separated set.

2. For $n \geq 1$ and $\epsilon > 0$, we call a finite subset $B$ of $X$ an **$(n, \epsilon)$-spanning set** if and only if for all $x \in X$, there exists $y \in B$ such that $d(f^k(x), f^k(y)) \leq \epsilon$ for all $k$ with $0 \leq k \leq n - 1$. Let $r(n, \epsilon)$ denote the minimal cardinality of any $(n, \epsilon)$-spanning set.

**Lemma 85.** $s(n, \epsilon)$ is finite.

**Proof.** Because $f$ is uniformly continuous, there exists $\gamma > 0$ such that if $d(x, y) < \gamma$, then $d(f^k(x), f^k(y)) < \epsilon$ for $0 \leq k \leq n - 1$. There exists a finite cover of $X$ by open sets $\{V_1, V_2, \ldots, V_j\}$ with diameter less than $\gamma$. We claim that $s(n, \epsilon) \leq j$. Let $A$ be a finite subset of $X$ with cardinality greater than $j$. There exists $x, y \in A$ with $x \neq y$ and $x, y \in V_i$ for some $i = 1, 2, \ldots, j$. Hence, $d(x, y) < \gamma$. This implies that $d(f^k(x), f^k(y)) < \epsilon$ for $0 \leq k \leq n - 1$. Thus $A$ is not $(n, \epsilon)$-separated. This proves the claim. \[\blacksquare\]
Lemma 86. \( r(n, \epsilon) \leq s(n, \epsilon). \)

**Proof.** Let \( s(n, j) = j \). Let \( A = \{y_1, y_2, \ldots, y_j\} \) be an \((n, \epsilon)\)-separated set with cardinality \( j \). We claim that \( A \) is an \((n, \epsilon)\)-spanning set. Let \( x \in X \). Two cases follow:

1. If \( x \in A \), choose \( y = x \). Then \( d(f^k(x), f^k(y)) = 0 \leq \epsilon \), for \( 0 \leq k \leq n - 1 \).

2. Suppose that \( x \not\in A \). Then the set \( \{y_1, y_2, \ldots, y_j, x\} \) is not an \((n, \epsilon)\)-separated set. So there exists \( y_i \in \{y_1, y_2, \ldots, y_j\} \) such that \( d(f^k(x), f^k(y_i)) \leq \epsilon \), for \( 0 \leq k \leq n - 1 \).

This proves the claim. Thus \( r(n, \epsilon) \leq j = s(n, \epsilon). \)

Lemma 87. \( s(n, 2\epsilon) \leq r(n, \epsilon). \)

**Proof.** Let \( r(n, \epsilon) = j \). There exists an \((n, \epsilon)\)-spanning set \( B = \{y_1, y_2, \ldots, y_j\} \) with cardinality \( j \). For each \( i \in \{1, 2, \ldots, j\} \) set

\[
D_i = \{x \in X : d(f^k(x), f^k(y_i)) \leq \epsilon \text{ for each } k = 0, 1, \ldots, n - 1\}.
\]

Then \( X = D_1 \cup D_2 \cup \ldots \cup D_j \). Let \( A \) be an \((n, 2\epsilon)\)-separated set.

We claim that \( \text{card}(d(A)) \leq j \). Each point of \( A \) is in some \( D_i \). Suppose distinct points \( v, w \in A \) are in the same \( D_i \), say \( D_{i_0} \). Then for each \( k = 0, 1, \ldots, n - 1 \) we have

\[
d(f^k(v), f^k(w)) \leq d(f^k(v), f^k(y_{i_0})) + d(f^k(y_{i_0}), f^k(w)) \leq \epsilon + \epsilon = 2\epsilon.
\]

However, this is a contradiction as \( A \) is an \((n, 2\epsilon)\)-separated. This proves the claim.

It follows that \( s(n, 2\epsilon) \leq j = r(n, \epsilon). \)

Lemma 88. Let \((X, d)\) be a compact metric space and let \( f : X \to X \) be continuous. Suppose \( n \geq 1 \) and \( \epsilon > 0 \). Let \( \alpha \) be an open cover with \( d(\alpha) < \epsilon \). Then \( s(n, \epsilon) \leq N(\alpha \vee f^{-1}(\alpha) \vee \ldots \vee f^{-(n-1)}(\alpha)). \)

**Proof.** Let \( s(n, \epsilon) = j \) and \( S \) be an \((n, \epsilon)\)-separated set with cardinality \( j \). We claim that each element \( A_{i_0} \cap f^{-1}(A_{i_1}) \cap \ldots \cap f^{-(n-1)}(A_{i_{n-1}}) \) of the cover \( \alpha \vee f^{-1}(\alpha) \vee \ldots \vee f^{-(n-1)}(\alpha) \) contains at most one point of \( S \). By way of contradiction, suppose that some element \( A_{i_0} \cap f^{-1}(A_{i_1}) \cap \ldots \cap f^{-(n-1)}(A_{i_{n-1}}) \) contains distinct points \( x, y \in S \). Then \( x, y \in A_{i_0} \) which implies \( d(x, y) < \epsilon \). Also \( f(x), f(y) \in A_{i_1} \) which implies \( d(f(x), f(y)) < \epsilon \). Similarly, \( d(f^k(x), f^k(y)) < \epsilon \) for \( k = 0, 1, \ldots, n - 1 \). This contradicts that \( S \) is \((n, \epsilon)\)-separated, proving the claim. So we need \( j \) sets in \( \alpha \vee f^{-1}(\alpha) \vee \ldots \vee f^{-(n-1)}(\alpha) \) to cover \( S \). Therefore \( N(\alpha \vee f^{-1}(\alpha) \vee \ldots \vee f^{-(n-1)}(\alpha)) \geq j \).

Lemma 89. Let \((X, d)\) be a compact metric space and let \( f : X \to X \) be continuous. Let \( n \geq 1 \). Suppose that \( \alpha \) is an open cover of \( X \) with Lebesgue number \( \delta > 0 \) (this means that if \( E \subset X \) with \( d(E) < \delta \), then there exists \( A \in \alpha \) with \( E \subset A \)). Set \( \epsilon = \frac{\delta}{2} \). Then \( N(\alpha \vee f^{-1}(\alpha) \vee \ldots \vee f^{-(n-1)}(\alpha)) \leq r(n, \epsilon) \).

**Proof.** Observe that if \( w \in X \) and \( CB(w, \epsilon) = \{x \in X : d(x, w) \leq \epsilon\} \), then \( d(CB(w, \epsilon)) \leq 2\epsilon < \delta \). Set \( r(n, \epsilon) = j \). Let \( B = \{y_1, y_2, \ldots, y_j\} \) be an \((n, \epsilon)\)-spanning set. For each \( i = 1, 2, \ldots, j \), set

\[
D_i = \{x \in X : d(f^k(x), f^k(y_i)) \leq \epsilon \text{ for each } k = 0, 1, \ldots, n - 1\}.
\]

Then \( X = D_1 \cup D_2 \cup \ldots \cup D_j \) because \( B \) is an \((n, \epsilon)\)-spanning set.Fix some \( i = 1, 2, \ldots, j \). Since \( d(CB(y_i, \epsilon)) < \delta \), there exists \( A_{i_0} \in \alpha \) such that \( CB(y_i, \epsilon) \subset A_{i_0} \). Similarly, there exists \( A_{i_1} \in \alpha \) such that \( CB(f(y_i), \epsilon) \subset A_{i_1} \), and for each \( k = 0, 1, \ldots, n - 1 \) there exists \( A_{i_k} \in \alpha \) such that \( CB(f^k(y_i), \epsilon) \subset A_{i_k} \).

We claim that \( D_i \subset (A_{i_0} \cap f^{-1}(A_{i_1}) \cap \ldots \cap f^{-(n-1)}(A_{i_{n-1}})) \). To prove this, observe that \( x \in D_i \) implies that \( d(x, y) \leq \epsilon \) which implies \( x \in A_{i_0} \). Also \( x \in D_i \) implies \( d(f(x), f(y_i)) \leq \epsilon \) which implies \( f(x) \in A_{i_1} \) and \( x \in f^{-1}(A_{i_1}) \). Similarly, \( x \in D_i \) implies \( d(f^k(x), f^k(y_i)) \leq \epsilon \) which implies \( f(x) \in A_{i_k} \) and \( x \in f^{-1}(A_{i_k}) \) for all \( k = 0, 1, \ldots, n - 1 \). This proves the claim. Now we can cover \( X = D_1 \cup D_2 \cup \ldots \cup D_j \) by at most \( j \) sets in \( \alpha \vee f^{-1}(\alpha) \vee \ldots \vee f^{-(n-1)}(\alpha) \). Therefore \( N(\alpha \vee f^{-1}(\alpha) \vee \ldots \vee f^{-(n-1)}(\alpha)) \leq r(n, \epsilon) \).

Lemma 90. Let \((X, d)\) be a compact metric space and let \( f : X \to X \) be continuous. Let \( n \geq 1 \) and suppose that \( 0 < \epsilon < \epsilon' \). Then

1. \( s(n, \epsilon) \geq s(n, \epsilon') \)
2. \( r(n, \epsilon) \geq r(n, \epsilon') \).

**Proof.** (1) If \( A \) is \((n, \epsilon')\)-separated, then \( A \) is \((n, \epsilon)\)-separated because in Definition 84 we have \( > \epsilon \). If \( s(n, \epsilon') = j \), then there exists an \((n, \epsilon)\)-separated set with cardinality \( j \). This set is also \((n, \epsilon)\)-separated. This implies that \( s(n, \epsilon) \), which is the maximum cardinality of an \((n, \epsilon)\)-separated set is greater than or equal to \( j = s(n, \epsilon') \).

(2) If \( B \) is an \((n, \epsilon)\)-spanning set, then \( B \) is an \((n, \epsilon')\)-spanning set. Set \( r(n, \epsilon) = j \). Then there exists an \((n, \epsilon)\)-spanning set with cardinality \( j \). This set is also a \((n, \epsilon')\)-spanning set. So \( r(n, \epsilon') \) which is the minimum cardinality of an \((n, \epsilon')\)-spanning set is less that or equal to \( r(n, \epsilon) \).

**Theorem 91.** Let \((X, d)\) be a compact metric space and let \( f : X \to X \) be continuous. Then

\[
  h(f) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log s(n, \epsilon) = \liminf_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log s(n, \epsilon) = \limsup_{\epsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \log r(n, \epsilon).
\]

**Proof.** The limits exist (may be \( \infty \)) by the monotonicity in Lemma 90 and in fact \( \liminf \) may be replaced by \( \sup \). We will prove two claims.

1. \( h(f) \leq \liminf_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log s(n, \epsilon) = \liminf_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log r(n, \epsilon) \)

2. \( h(f) \geq \limsup_{\epsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \log s(n, \epsilon) = \limsup_{\epsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \log r(n, \epsilon) \)

The conclusion follows. To prove the first claim, let \( \alpha \) be an open cover of \( X \). Also, let \( \delta > 0 \) be a Lebesgue number for \( \alpha \). Then

\[
  h(f, \alpha) = \liminf_{n \to \infty} \frac{1}{n} \log N(\alpha \cup f^{-1}(\alpha) \cup \ldots \cup f^{-(n-1)}(\alpha)) \leq \liminf_{n \to \infty} \frac{1}{n} \log r(n, \epsilon) \leq \lim_{\epsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \log r(n, \epsilon) \leq \limsup_{\epsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \log s(n, \epsilon).
\]

So

\[
  h(f) = \sup_{\alpha} h(f, \alpha) \leq \liminf_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log r(n, \epsilon) = \liminf_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log s(n, \epsilon).
\]

To prove the second claim, let \( \epsilon > 0 \). There exists an open cover \( \alpha \) of \( X \) with \( d(\alpha) < \epsilon \). We have

\[
  h(f) \geq h(f, \alpha) = \limsup_{n \to \infty} \frac{1}{n} \log N(\alpha \cup f^{-1}(\alpha) \cup \ldots \cup f^{-(n-1)}(\alpha)) \geq \limsup_{n \to \infty} \frac{1}{n} \log s(n, \epsilon) \quad \text{(Lemma 88)}.
\]

Since \( \epsilon > 0 \) was arbitrary, we have

\[
  h(f) \geq \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log s(n, \epsilon) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log r(n, \epsilon),
\]

as desired.
**Remark 92.** We will now look at some results on maps of the interval to itself. The setting is $I = [0, 1]$ and $f : I \to I$ is continuous. We will use the terms interval or subinterval to mean a subset of $I$ of one of the forms $[a, b], (a, b), [a, b), (a, b]$ where $a < b$.

**Lemma 93.** Let $J$ be a closed subinterval of $I$. If $f(J) \supset J$, then $f$ has a fixed point in $J$.

**Lemma 94.** Suppose $J$ and $K$ are closed subintervals of $I$ with $f(J) \supset K$. Then there exists a closed subinterval $L$ of $J$ with $f(L) = K$.

**Proposition 95.** Let $J_0, J_1, \ldots, J_n$ be closed subintervals of $I$ with $J_0 = J_n$ and $f(J_0) \supset J_1$, $f(J_1) \supset J_2$, $\ldots$, $f(J_{n-1}) \supset J_n$. Then there exists $x \in J_0$ with $f^n(x) = x$ and $f(x) \in J_1$, $f^2(x) \in J_2$, $\ldots$, $f^{n-1}(x) \in J_{n-1}$.

**Definition 96.** Let $P$ be a periodic orbit of $f$, $P = \{x_1, x_2, \ldots, x_n\}$ with $x_1 < x_2 < \ldots < x_n$ (notation: $P = \{x_1 < x_2 < \ldots < x_n\}$). The P-graph associated to $P$ is defined as follows: Let $I_1 = [x_1, x_2]$, $I_2 = [x_2, x_3]$, $\ldots$, $I_{n-1} = [x_{n-1}, x_n]$ be the vertices of the graph with an arrow (or edge) from $I_i$ to $I_j$ if and only if $I_j \subset (f(x_i), f(x_{i+1}))$ where $[c, d]$ denotes the closed interval with endpoints $c$ and $d$ (either $[c, d]$ or $[d, c]$). The P-graph is sometimes called the Markov graph associated to $P$.

**Remark 97.** Let $P$ be a periodic orbit of $f$, with $P = \{x_1 < x_2 < \ldots < x_n\}$. We associate to $P$ a piecewise linear map $L$ (called the linearization or connect-the-dots map) defined by

1. The domain of $L$ is $[x_1, x_n]$;
2. $L(x_i) = f(x_i)$ for all $i$;
3. $L_{[x_i, x_{i+1}]}$ is linear for all $i$.

Note that the P-graph may be determined from the map $L$.

**Remark 98.** We defined P-graph for a periodic orbit $P$. More generally, we could consider any finite subset $P$ of $I$ with $f(P) \subset P$. We could form a P-graph in the same way. Here we might have a vertex with no arrows (i.e. we might have $f(x_i) = f(x_{i+1})$ where $x_i < x_{i+1}$ are adjacent points of $P$).

**Proposition 99.** Suppose $f : I \to I$ is continuous. Suppose $f$ has a periodic orbit of period 3. Then for each positive integer $k$, $f$ has a periodic point of period $k$.

**Example 100.** Let $P = \{x_1 < x_2 < x_3\}$ where $f(x_1) = x_1$, $f(x_2) = x_3$ and $f(x_3) = x_1$. Show that for each positive integer $k$, $f$ has a periodic point of period $k$.

**Example 101.** Let $P = \{x_1 < x_2 < x_3 < x_4 < x_5\}$ where $f(x_1) = x_2$, $f(x_2) = x_3$, $f(x_3) = x_4$, $f(x_4) = x_5$ and $f(x_5) = x_1$. Show that for each positive integer $k$, $f$ has a periodic point of period $k$.

**Example 102.** Let $P = \{x_1 < x_2 < x_3 < x_4 < x_5\}$ where $f(x_1) = x_4$, $f(x_4) = x_2$, $f(x_2) = x_5$, $f(x_5) = x_1$ and $f(x_1) = x_3$. Show that for each positive integer $k \neq 3$, $f$ has a periodic point of period $k$. Must $f$ have a periodic point of period 3?
Proposition 103. Let \( P = \{x_1 < x_2 < \ldots < x_n\} \) be a periodic orbit of period \( n \geq 3 \). Suppose \( K = [x_k, x_{k+1}] \) is a vertex in the \( P \)-graph with \( K \rightarrow K \). Let \( J = [x_j, x_{j+1}] \) be an arbitrary vertex. Then there exists a path from \( K \) to \( J \).

Definition 104. Let \( P \) be a periodic orbit. Let \( J_1 \rightarrow J_2 \rightarrow \ldots \rightarrow J_k \rightarrow J_1 \) be a path in the \( P \)-graph from some vertex \( J_1 \) to itself of length \( k \). We say the path is repetitive if and only if there exists positive integers \( s, t \) with \( t > 1 \) and \( k = st \) such that for each \( i = 1, 2, \ldots, s \) we have \( J_i = J_{i+t} \). In other words the given path is just \( J_1 \rightarrow J_2 \rightarrow \ldots \rightarrow J_s \rightarrow J_1 \rightarrow J_2 \rightarrow \ldots \rightarrow J_s \rightarrow \ldots \rightarrow J_1 \) repeated \( t \) times.

Theorem 105. Suppose that \( P \) is a periodic orbit of \( f \). Suppose that \( J_1 \rightarrow J_2 \rightarrow \ldots \rightarrow J_k \rightarrow J_1 \) is a path in the \( P \)-graph of length \( k \) from some vertex \( J_1 \) to itself. Suppose also that this path is not repetitive. Then there exists \( y \in J_1 \) with \( f(y) \in J_2, \ldots, f^{k-1}(y) \in J_k \) such that \( y \) is a periodic point of \( f \) with period \( k \).

Definition 106. Let \( P = \{x_1 < x_2 < \ldots < x_n\} \) be a periodic orbit of period \( n \), where \( n \geq 3 \) is odd. Let \( c \) be the midpoint of the orbit (i.e. \( c = \frac{x_{n+1}}{2} \)). We say that \( P \) is a Stefan orbit if and only if either
\[
f^{n-1}(c) < \ldots < f^4(c) < f^2(c) < c < f(c) < f^3(c) < \ldots < f^{n-2}(c)
\]
or
\[
f^{n-2}(c) < \ldots < f^3(c) < f(c) < f^2(c) < f^4(c) < \ldots < f^{n-1}(c).
\]

Example 107. Suppose that \( n = 7 \). Let \( P = \{x_1 < x_2 < \ldots < x_7\} \) be a Stefan orbit. Suppose we have the following
\[
\begin{array}{cccc}
  f^6(c) & f^5(c) & f^2(c) & c \\
  f(c) & f^3(c) & f^5(c) & f^4(c)
\end{array}
\]

Let \( I_1 = [c, f(c)], I_2 = [f^2(c), c], I_3 = [f(c), f^3(c)], I_4 = [f^4(c), f^2(c)], I_5 = [f^3(c), f^5(c)] \) and \( I_6 = [f^6(c), f^4(c)] \). Then the \( P \)-graph is

\[\text{Diagram}\]

Theorem 108. Let \( f : I \rightarrow I \) be continuous. Let \( n \geq 5 \) be an odd positive integer. Suppose that \( P = \{x_1 < x_2 < \ldots < x_n\} \) is a periodic orbit of \( f \) with period \( n \). Suppose also that \( f \) has no periodic point of period \( k \) for any \( k \) with \( 1 < k < n \) and \( k \) odd. Then \( P \) is a Stefan orbit. Moreover, the vertices in the \( P \)-graph may be renumbered so that the \( P \)-graph has the form

\[\text{Diagram}\]
with arrows from $I_{n-1}$ to $I_j$ if and only if $j$ is odd.

**Definition 109.** The Sharkovskoy order of the positive integers is as follows:

$$3 \triangleright 5 \triangleright 7 \triangleright 9 \triangleright \ldots \triangleright 3 \cdot 2 \triangleright 5 \cdot 2 \triangleright 7 \cdot 2 \triangleright 9 \cdot 2 \triangleright \ldots \triangleright 3^2 \triangleright 5^2 \triangleright 7^2 \triangleright 9^2 \triangleright \ldots \triangleright 2^4 \triangleright 3^2 \triangleright 2^2 \triangleright 2 \triangleright 1$$

**Theorem 110.** Suppose that $f : I \to I$ is continuous and has a periodic point of period $2j + 1$ for some $j \geq 1$. Suppose that $2j + 1 \triangleright k$. Then $f$ also has a periodic point of period $k$.

**Theorem 111.** Suppose that $k$ is a positive integer and set $n = 2k$. Suppose that $P = \{x_1 < x_2 < \ldots < x_n\}$ is a periodic orbit of $f$, where $f$ is a continuous map from the interval to itself. Suppose that for each odd $j \geq 3$, $f$ has no periodic point of period $j$. Set $L = \{x_1, x_2, \ldots, x_k\}$ and $R = \{x_{k+1}, x_{k+2}, \ldots, x_n\}$. Then $f(L) = R$ and $f(R) = L$.

**Corollary 112.** Suppose that $f$ has a periodic orbit of period $n$ where $n \triangleright 2$ (i.e. $n > 2$). Then $f$ also has a periodic point of period 2.

**Remark 113.** Suppose $s \geq 1$ and suppose $x$ is a periodic point of $f$ with period $k$. Let $j$ denote the period of $x$ under the map $f^{(s)}$. Then

1. If $k$ is odd, then $j = k$.
2. If $k = 2^t \cdot n$ where $1 \leq t \leq s$ and $n \geq 1$ is odd, then $j = n$.
3. If $k = 2^t \cdot n$ where $t > s$ and $n \geq 1$ is odd, then $j = (2^{(t-s)}) \cdot n$

**Lemma 114.** Suppose $s \geq 1$ and $f$ has a periodic point of period $2^s$. Then also $f$ has a periodic point of period $2^{s-1}$.

**Theorem 115.** (Sharkovsky’s Theorem) Suppose $n$ and $k$ are positive integers with $n \triangleright k$. Suppose $f : I \to I$ is continuous and $f$ has a periodic point of period $n$. Then $f$ also has a periodic point of period $k$.

**Definition 116.** A continuous map $f : X \to X$ is said to be a factor of the continuous map $g : Y \to Y$ if and only if there exists a continuous surjective map $\pi : Y \to X$ such that the diagram

$$
\begin{array}{ccc}
Y & \xrightarrow{g} & Y \\
\pi & & \pi \\
X & \xrightarrow{f} & X
\end{array}
$$

commutes. The map $\pi$ is called a semi-conjugacy.

**Definition 117.** Let $f : I \to I$ be continuous, and let $C = \{J_1, \ldots, J_n\}$ be a finite collection of closed intervals with pairwise disjoint interiors. The matrix $A$ associated to $f$ and $C$ (called the adjacency matrix) is the $n \times n$ matrix $A = (a_{ij})$ where

$$a_{ij} = \begin{cases} 
1, & \text{if } f(J_i) \supset J_j \\
0, & \text{otherwise}
\end{cases}.$$

**Theorem 118.** Suppose $f$, $C$ and $A$ are as in Definition 117. Suppose that each row of $A$ has at least one 1. Suppose in addition, that the intervals are pairwise disjoint. Then there is a closed, invariant subset $X$ of $I$ such that $\sigma_A : \Sigma_A \to \Sigma_A$ is a factor of $f \big|_X : X \to X$. In particular, $h(f) \geq \log \lambda$, where $\lambda$ is the maximal eigenvalue of $A$.

**Corollary 119.** Suppose $f : I \to I$ is continuous. Suppose $J_1, \ldots, J_p$ are closed intervals, which are pairwise disjoint. Suppose for some positive integer $n$, we have

$$f^n(J_i) \supset (J_1 \cup \ldots \cup J_p)$$

for each $i = 1, 2, \ldots, p$. Then $h(f) \geq \frac{1}{n} \log p$. 

(1)
Proof. Consider $f^n$ and the collection $C = \{J_1, \ldots, J_p\}$. The corresponding matrix is the $p \times p$ matrix with all entries 1. So $\sigma_A$ is the full one-sided shift and the maximal eigenvalue is $p$. Thus

$$h(f) = \frac{1}{n} h(f^n) \geq \frac{1}{n} \log p.$$ 

Remark 120. The following is a true theorem. Let $f : I \rightarrow I$ be continuous. Suppose $\alpha$ is a real number with $0 < \alpha < h(f)$. Then there exists a positive integers $n$ and $p$, and a collection $J_1, \ldots, J_p$ of pairwise disjoint intervals such that (1) holds and $\frac{1}{n} \log p > \alpha$.

Theorem 121. Suppose $f$, $C$ and $A$ are as in Definition 117. Let $\lambda$ denote the maximal eigenvalue of $A$. Then $h(f) \geq \log \lambda$.

Corollary 122. Let $f : I \rightarrow I$ be continuous. Suppose that $J_1, \ldots, J_p$ are closed subintervals of $I$ with pairwise disjoint interiors. Suppose that for some positive integer $n$, $f^n(J_i) \supset (J_1 \cup \ldots \cup J_p)$ for all $i = 1, 2, \ldots, p$. Then $h(f) \geq \frac{1}{n} \log p$.

Definition 123. Let $f : I \rightarrow I$ be continuous. Suppose $P$ is a finite subset of $I$ with $0 \in P$ and $1 \in P$ such that $f(P) \subset P$. Let $P = \{x_1 < x_2 < \ldots < x_n\}$. We say that $f$ is $P$-monotone if and only if for each $i = 1, \ldots, n - 1$, $f|_{[x_i, x_{i+1}]}$ is monotone with $f(x_i) \neq f(x_{i+1})$. In this case, the collection of intervals $\{[x_1, x_2], \ldots, [x_{n-1}, x_n]\}$ is called a Markov partition, $f$ is called a Markov map and the $P$-graph $G$ is called the Markov graph. The subshift of finite type $\sigma_A : \Sigma_A \rightarrow \Sigma_A$, where $A$ is the matrix corresponding to $G$ is called the subshift of finite type corresponding to $f$. Finally, if in addition, $f|_{[x_i, x_{i+1}]}$ is linear for all $i = 1, 2, \ldots, n - 1$ then we say $f$ is $P$-linear and $f$ is called a linear Markov map.

Remark 124. Suppose $f$ and $P$ are as in Definition 123, and $f$ is $P$-monotone. We may consider itineraries, but itineraries are need not be unique. The point $y \in I$ has a unique itinerary if and only if $O^+(y) \cap P = \emptyset$. We claim the following: Let $\alpha = (s_0, s_1, \ldots) \in \Sigma_A$. The set of $y \in I$ such that $\alpha$ is an itinerary of $y$ is either a closed interval or a point.

Remark 125. If $f$ is Markov, then $h(f) = h(\sigma_A) = \log \lambda$, where $\lambda$ is the maximal eigenvalue of $A$. markov term
**Example 126.** (The Horseshoe Map, S. Smale)

Part 1: The ultimate objective is to obtain a homeomorphism $F : S^2 \to S^2$. First, we construct a map $F$ from a disk into itself. Consider the unit square $S$, with endpoints $(0,0), (1,0), (0,1)$ and $(1,1)$, in the plane. Divide the unit square into 5 vertical strips, labeled from left to right $A, V_1, C, V_2$ and $E$. We will adjoin regions $D_1$ and $D_2$ to $S$ as in the picture below:

Let $S \cup D_1 \cup D_2 = D$. Define $F : D \to D$ as seen in the picture below:

$F$ maps $D$ homeomorphically onto $F(D)$. $F$ maps $V_1$ onto $F(V_1)$ linearly, contracting by $\frac{1}{5}$ in vertical direction and stretching by a factor of 5 in the horizontal direction:

$$F(x,y) = \left(5\left(x - \frac{1}{5}\right), \frac{1}{5}y + \frac{3}{5}\right)$$

for $(x,y) \in V_1$. Also, $F$ maps $V_2$ onto $F(V_2)$ similarly, say

$$F(x,y) = \left(1 - 5(x - \frac{3}{5}), \frac{2}{5} - \frac{1}{5}y\right)$$

for $(x,y) \in V_2$. We construct $F$ on $D_1$, so that $F\big|_{D_1}$ is a contraction mapping of $D_1$ into itself (i.e. there exists $c$ with $0 < c < 1$ such that $d(F(x), F(y)) \leq c \cdot d(x, y)$ for all $x, y \in D$). Then the Contraction Mapping Theorem applies.

Part 2: We extend $F$ to a homeomorphism $F : S^2 \to S^2$. First, let $\hat{D}$ be a disk containing $D$ in its interior.
We can extend $F$ to a homeomorphism $F : \hat{D} \to D$ which maps the boundary of $\hat{D}$ to the boundary of $D$ and maps $\hat{D}$ onto $D$. Let $\hat{D}$ be the bottom hemisphere of $S^2$. Draw horizontal rings approaching the north pole $q$. A sequence $R_n$ of such rings $n = 1, 2, \ldots$ lets map the ring $R_n$ onto the ring $R_{n-1}$ for $n = 2, 3, \ldots$ and fill in the area between the rings. Set $F(q) = q$. Note that $q$ is called a source or expanding fixed point. We have a well defined homeomorphism $F : S^2 \to S^2$. The point $p$ (the unique fixed point in $D_1$) is a sink or contracting fixed point.

Set $X = \{x \in S : F^k(x) \in S$ for all $k \in \mathbb{Z}\}$. We can show that $NW(F) \subset (X \cup \{p, q\})$.

Part 3: We give a geometric description of the set $X$. We see that $X = C \times C$ where $C$ is the “middle two-fifths” Cantor set.

Part 4: We see that $F|_X$ and $\sigma : \Sigma_2 \to \Sigma_2$ (full two-sided shift on two symbols) are topologically conjugate. Define $h : X \to \Sigma_2$ by $h(x) = (\ldots s_{-2}s_{-1}s_0s_1 \ldots)$ where

$$s_n = \begin{cases} 1, & \text{if } F^n(x) \in V_1 \\ 2, & \text{if } F^n(x) \in V_2 \end{cases}.$$

Automatically, the diagram

$$\begin{array}{ccc} X & \xrightarrow{F|_X} & X \\
\downarrow{h} & & \downarrow{h} \\
\Sigma_2 & \xrightarrow{\sigma} & \Sigma_2 
\end{array}$$

commutes. Also, $h$ is continuous because if $x \in X$, $y \in Y$ are close, we can get $F^k(x), F^k(y)$ in same strip $V_1, V_2$ for $-N \leq k \leq N$.

**Definition 127.** Let $F : S^2 \to S^2$ be the homeomorphism in Example 126. Let $x \in NW(F)$. The stable manifold of $x$ is given by

$$W^s(x) = \{y \in S^2 : d(F^n(y), F^n(x)) \to 0 \text{ as } n \to \infty\}.$$

The unstable manifold of $x$ is given by

$$W^u(x) = \{y \in S^2 : d(F^{-n}(y), F^{-n}(x)) \to 0 \text{ as } n \to \infty\}.$$

**Proposition 128.** Let $F : S^2 \to S^2$ be the homeomorphism in Example 126. Let $p$ denote the unique fixed point of $F$ in $D_1$. Then $W^s(p)$ is an open dense subset of $S^2$ and $W^s(q) = \{q\}$.

**Definition 129.** Let $F : S^2 \to S^2$ be the homeomorphism in Example 126. Let $x \in X$. The local stable manifold of $x$ is the set of points in the square $S$ on the vertical line through $x$. We let $W^s_l(x)$ denote the local stable manifold of $x$. The local unstable manifold of $x$ is the set of points in the square $S$ on the horizontal line through $x$. We let $W^u_l(x)$ denote the local unstable manifold of $x$.

**Proposition 130.** Let $F : S^2 \to S^2$ be the homeomorphism in Example 126. Let $x \in X$. We have

$$W^s(x) = \bigcup_{n=0}^{\infty} F^{-n}(W^s_l(F^n(x))).$$

**Proposition 131.** Let $F : S^2 \to S^2$ be the homeomorphism in Example 126. The collection of stable manifolds of points in $NW(F)$ forms a partition of $S^2$. 
**Example 132.** (Hyperbolic Toral Automorphism) Let $L : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear map obtained from the matrix

$$
\begin{pmatrix}
2 & 1 \\
1 & 1 \\
\end{pmatrix}.
$$

The eigenvalues are $\lambda = \frac{3 + \sqrt{5}}{2} > 1$ and $\frac{1}{\lambda} < 1$. We have eigenvectors, $v_\lambda = (\frac{1 + \sqrt{5}}{2}, 1)$ and $v_{\frac{1}{\lambda}} = (\frac{1 - \sqrt{5}}{2}, 1)$.

Define an equivalence relation on $\mathbb{R}^2$ by $(x_1, y_1) \sim (x_2, y_2)$ if and only if $x_2 - x_1 \in \mathbb{Z}$ and $y_2 - y_1 \in \mathbb{Z}$. Set $T^2 = \mathbb{R}^2 / \sim$. We have a projection $\pi : \mathbb{R}^2 \to T^2$. There exists a unique continuous $\hat{L}$ such that the following diagram commutes:

$$
\begin{array}{ccc}
\mathbb{R}^2 & \xrightarrow{L} & \mathbb{R}^2 \\
\pi \downarrow & & \downarrow \pi \\
T^2 & \xrightarrow{\hat{L}} & T^2
\end{array}
$$

This is because if $(x_1, y_1) \sim (x_2, y_2)$, then $L(x_1, y_1) \sim L(x_2, y_2)$. In fact, $\hat{L}$ is a homeomorphism.

**Theorem 133.** The periodic points of $\hat{L}$ form a dense subset of $T^2$.

**Definition 134.** Let $x \in T^2$. We define the stable manifold of $x$ by

$$
W^s(x) = \left\{ y \in T^2 : d(\hat{L}^n(y), \hat{L}^n(x)) \to 0 \text{ as } n \to \infty \right\}.
$$

We define the unstable manifold of $x$ by

$$
W^u(x) = \left\{ y \in T^2 : d(\hat{L}^n(y), \hat{L}^n(x)) \to 0 \text{ as } n \to -\infty \right\}.
$$

**Proposition 135.** Let $\hat{L} : T^2 \to T^2$ be the homeomorphism in Example 132. The collection of stable manifolds of points in $T^2$ forms a partition of $T^2$.

**Proposition 136.** Let $E$ be a line in $\mathbb{R}^2$ which is not vertical and has irrational slope. Let $p : \mathbb{R}^2 \to T^2$ be defined by $p(x, y) = (e^{2\pi ix}, e^{2\pi iy})$. Then $p(E)$ is a dense subset of $T^2$.

**Proposition 137.** For any $x \in T^2$, both $W^s(x)$ and $W^u(x)$ are dense in $T^2$.

**Theorem 138.** The homeomorphism $\hat{L}$ is topologically mixing.

**Definition 139.** Let $\hat{L} : T^2 \to T^2$ be the homeomorphism in Example 132. A rectangle is a subset of $T^2$ of the form $I_x \times I_y$ where $I_x$ is a small closed interval on a stable manifold, and $I_y$ is a small closed interval on an unstable manifold. A finite cover of $T^2$ by rectangles is a rectangle partition if and only if the intersection of any two distinct rectangles is contained in the boundaries of each of the rectangles. A rectangle partition $\{A_1, \ldots, A_r\}$ is a Markov partition if and only if whenever

$$
x \in \text{int}A_i \cap \hat{L}^{-1}\text{int}A_j
$$

we have

$$
\hat{L}(W^u_{\text{loc}}(x) \cap A_i) \supset (W^u_{\text{loc}}(\hat{L}(x)) \cap A_j)
$$

and also

$$
\hat{L}(W^s_{\text{loc}}(x) \cap A_i) \subset (W^s_{\text{loc}}(\hat{L}(x)) \cap A_j).
$$

**Theorem 140.** There exists a Markov partition for $\hat{L}$. 