Dynamical Systems Notes

Lou Yi

April 28, 2025

Contents

1 System With Stable Asymptotic Behaviour **1**

1.1 Contractions . . . 1

1.2 Nondecreasing Interval Map . . . 6

1.3 Quadratic Maps . . . 7

1.4 Eventually Contracting Map . . . 8

1.5 Fractals . . . 10

2 Recurrence and Equidistribution **14**

2.1 Topological Transitivity and Minimality . . . 14

2.2 Introduction to Ergodic Theory . . . 16

2.3 Circle Rotation . . . 20

2.4 Translation on the Torus . . . 25

2.5 Poincaré Classification Theorem for Circle Maps . . . 29

2.5.1 The Denjoy Example . . . 39

3 Hyperbolic Dynamics **41**

3.1 Expanding Maps . . . 41

3.2 Hyperbolic Toral Translation . . . 41

3.3 Topological Mixing . . . 43

4 Symbolic Dynamics **46**

4.1 Coding . . . 46

4.2 Topological Markov Chain . . . 48

5 Entropy **52**

5.1 Capacity . . . 52

5.2 Topological Entropy . . . 53

5.3 Topological Entropy Via Covers and Separated Sets . . . 55

5.4 Examples . . . 58

1 System With Stable Asymptotic Behaviour

1.1 Contractions

Given a linear function $f(x) = kx + \ell$, its corresponding interation is given by

$$x_{i+1} = kx_i + \ell.$$

If we do a change of a variable, by letting $y = x - \frac{\ell}{1-k}$, $k \neq 0$, then

$$y_{i+1} = ky_i$$

Then the asymptotic behaviour of this system can be easily understood.

Definition 1.1 (Differential of Maps). *Let $F : \mathbb{R}^n \to \mathbb{R}^m$, $F = (f_1, \cdots, f_m)$, then*

$$DF = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

If F is regular enough, then

$$F(x_0) = F(x) + DF(x_0 - x) + o(\|x_0 - x\|).$$

In particular, we have

$$F(x_0) = DF(x_0 - x) + O(1)$$

if $|x - x_0|$ is very small.

Definition 1.2 (Lipschitz and Contraction). *A map f of a subset U of a metric space (X, d) is said to be **Lipschitz continuous** with Lipschitz constant λ if*

$$d(f(x), f(y)) \leq \lambda d(x, y).$$

*If $\lambda < 1$, then f is said to be a **contradiction**. If a map f is Lipschitz, then we define its **Lipschitz semi-norm** by*

$$\mathrm{Lip}(f) = \sup_{x \neq y} \frac{d(f(x), f(y))}{d(x, y)}.$$

Definition 1.3 (Orbit). *If f is not invertible, we define its **Orbit***

$$\mathrm{Orb}(f)(x) = \{x, f(x), f^2(x), \cdots, \}$$

*If f is invertible, then we define its **Orbit** by*

$$\mathrm{Orb}(f)(x) = \{\cdots, f^{-1}(x), x, f(x), \cdots\}.$$

*We call the **positive semi-orbit** to be the set of non-negative iterations, and we call the **negative semi-orbit** to* *be the set of non-positive iterations. A* ***fixed point*** *of f is a point x such that*

$$f(x) = x.$$

We denote the set of all fixed points of f to be $\mathrm{Fix}(f)$. *A* ***periodic point*** *of f is a point x such*

$$f^n(x) = x$$

for some $n \in \mathbb{N}_{>0}$. Such n is called the ***periodic*** *of x, and the smallest such n is called the* ***prime period*** *of x.*

Remark 1.3.1. *If x is a periodic point of f, then $x \in \mathrm{Fix}(f^n)$ for some $n \in \mathbb{N}_{>0}$.*

Definition 1.4 (Exponential Convergence)**.** *Let (X, d) be a metric space, we say two sequence (x_n), (y_n)* ***converge exponentially*** *if there exists $C > 0$ and $\alpha \in (0, 1)$ such that*

$$d(x_n, y_n) < C \cdot \alpha^n.$$

Remark 1.4.1. *If two sequences converges exponentially, and one of the sequence converges to a limit, then so does the other sequence.*

Proposition 1.5 (Contraction Principle)**.** *Let (X, d) be a complete metric space under the action of iteration of a contraction $f : X \to X$ all points converge with exponential speed to the unique fixed point of f.*

Proof. Note that

$$d(f^n(x), f^n(y)) \leq \lambda^n d(x, y).$$

Since $\lambda < 1$, $d(x, y)$ is bounded, then

$$d(f^n(x), f^n(y)) \leq C \cdot \lambda^n$$

where $C = d(x, y)$, so the iteration of all points have the same asymptotic behaviour, and in fact converges with exponential speed. Next, let $x \in X$ be arbitrary,WLOG let $n < m$, we note that

$$d(f^n(x), f^m(x)) \leq \sum_{k=0}^{m-n-1} d(f^{n+k+1}(x), f^{n+k}(x)) \leq \sum_{k=0}^{m-n-1} \lambda^{n+k} d(f(x), x).$$

Letting $n \to \infty$, we see that the sequence is a Cauchy sequence, hence for any $x \in X$, the limit $f^{(n)}(x)$ exists as $n \to \infty$. Lastly, we just need to show that this limit is a fixed point, we know it is unique since any two such iterations have the same asymptotic behaviour. Let x_0 be the limit, then

$$d(x_0, f(x_0)) \leq d(x_0, f^n(x)) + d(f^n(x), f^{n+1}(x)) + d(f^{n+1}(x), f(x_0))$$

Letting $n \to \infty$, we see the right hand side can be arbitrarily small, so we must have $d(x_0, f(x_0)) = 0$. □

Definition 1.6 (Convex Set)**.** *A* ***convex subset*** *$C \subseteq \mathbb{R}^n$ is a set such that for all $x, y \in C$, $tx + (1-t)y \in C$ for $t \in [0, 1]$.*

Theorem 1.7. *If $C \subset \mathbb{R}^n$ is convex and open. Suppose $f : C \to \mathbb{R}^n$ is differentiable with*

$$\|Df_x\| \leq M$$

for $x \in \mathbb{C}$. Then for $x, y \in C$, we have

$$\|f(x) - f(y)\| \leq M \cdot \|x - y\|.$$

In particular, if $f : C \to C$ with continuous partial derivative, and

$$\|Df_x\| \leq \lambda < 1, \quad \forall x \in C,$$

then f is a λ-contradiction.

Remark 1.7.1. *The norm of the differential is understood as operator norm.*

Remark 1.7.2. *The condition that C is convex is necessary. For example, there exists a continuously differentiable map $f : U \subset \mathbb{R}^2 \to U$ such that $\|Df_x\| < \lambda < 1$ for all $x \in U$ but f is not a contraction. (Consider the annulus with a sector removed, and continuously deforming it to an annulus with a larger sector removed)*

Proof. Assume that the line segment containing x and y is given by

$$c(t) = x + t(y - x) \quad t \in [0, 1]$$

Then $c(t) \in C$, $\forall t \in [0, 1]$. By the chain rule and mean value theorem, we have

$$\begin{aligned}
\|f(y) - f(x)\| &= \|f(c(1)) - f(c(0))\| \\
&\leq \left\| \frac{d}{dt} f(c(\tilde{t})) \right\| \cdot (1 - 0) \\
&= \left\| Df_{c(\tilde{t})} \cdot \frac{d}{dt} c(t) \right\| \\
&\leq \|Df_{c(\tilde{t})} \cdot (y - x)\| \\
&\leq M \cdot \|y - x\|
\end{aligned}$$

□

Lemma 1.8. *Let $\|A\|$ denote the operator norm of the matrix, then*

$$\|A\| \leq \|A\|_F$$

and

$$\|A\| \geq |\det A|^{1/n}.$$

Moreover, $\|A\|$ is a continuous function of its coefficients.

Proof. Note that if $\|x\| = 1$, then $\|Ax\| \leq \|A\|_F$ by matrix multiplication, so $\|A\| \leq \|A\|_F$.

Next, $\|Ax\|^2 = x^T A^T A x$. Suppose x is any eigenvalue of $A^T A$ corresponding to σ, then $\|Ax\| = \sqrt{\sigma}\|x\|$. Note $\sqrt{\sigma}$ is a singular value of A. Moreover, by singular value decomposition, we know that $|\det A|$ is equal to the product of singular values. Then pick σ to be the biggest eigenvalue, we obtain the second inequality.

Lastly, we show that given $\epsilon > 0$, $\exists \delta > 0$ such that if A, B are matrices such that $\|A - B\|_1 < \delta$ is small, then $|\|A\| - \|B\|| < \epsilon$. A moment of thought tells us that we in fact can take $\delta = \epsilon$. $\square$

Remark 1.8.1. *In fact, from the proof we see that the operator norm of A is equal to the largest singular value of A.*

Proposition 1.9. *Let f be a continuous differentiable function with a fixed point x_0 such that $\|Df_{x_0}\| < 1$. Then there is a closed neighbourhood U of x_0 such that $f(U) \subset U$ and f is a contraction on U.*

Proof. Since Df is continuous, then there exists a small closed ball $U = B(x_0, \eta)$ such that $\|Df_x\| \leq \lambda < 1$ for all $x \in U$. Then $f|_U$ would be a contraction. Lastly, we need to prove that $f(U) \subset U$. Let $y \in B(x_0, \eta)$, since $d(f(y), f(x_0)) < \lambda\eta < \eta$, then $f(y) \in U$ as desired. $\square$

Proposition 1.10 (Perturbation)**.** *Let f be a continuously differentiable map with a fixed point x_0, where $\|Df_{x_0}\| < 1$. Let U be a closed neighbourhood of x_0 such that $f(U) \subset U$. Then any map g sufficiently close to f is a contraction. To be more precisely, if $\epsilon > 0$, then there is $\delta > 0$, s.t., any map g with*

$$\|g(x) - f(x)\| \leq \delta, \ \ \|Dg_x - Df_x\| \leq \delta, \ \ \forall x \in U$$

satisfies $g(U) \subset U$, is a contraction with its unique fixed point $y_0 \in B(x_0, \epsilon)$.

Proof. For simplicity, assume $U = \overline{B(x_0, \eta)}$, $\eta, \epsilon < 1$. Take

$$\delta = \frac{\epsilon\eta(1-\lambda)}{2}$$

Here λ satisfies $\|Df_x\| \leq \lambda < 1$, $\forall x \in U$. Then

$$\begin{aligned}
\|Dg_x\| &\leq \|Dg_x - Df_x\| + \|Df_x\| \\
&\leq \delta + \lambda < \lambda + \frac{(1-\lambda)}{2} < 1.
\end{aligned}$$

If $x \in U$,

$$\begin{aligned}
d(g(x), x_0) &\leq d(g(x), g(x_0)) + d(g(x_0), f(x_0)) + d(f(x_0), x_0) \\
&\leq \frac{1+\lambda}{2} d(x, x_0) + \delta \\
&\leq \frac{1+\lambda}{2} \cdot \eta + \delta \leq \eta.
\end{aligned}$$

So $g(x) \in U$. Now $g^n(x_0) \to y_0$ as $n \to \infty$ (in particular, y_0 is the fixed point). Then

$$\begin{aligned}
d(x_0, y_0) &\leq \sum_{n=0}^{\infty} d(g^n(x_0), g^{n+1}(x_0)) \\
&\leq d(g(x_0), x_0) \cdot \sum_{n=0}^{\infty} \left(\frac{1+\lambda}{2}\right)^n \\
&\leq \frac{2\delta}{1-\lambda} = \epsilon.
\end{aligned}$$

$\square$

Definition 1.11 (Poisson Stable)**.** *A fixed point p is said to be **Poisson stable** if $\forall \epsilon > 0$, $\exists \delta > 0$ such that if a point within δ-neighbourhood of p, then its positive semiorbit is within ϵ-neighbourhood of p.*

Definition 1.12 (Asymptotically Stable)**.** *A fixed point p is said to be **asymptotically stable** or **attracting fixed point**, if it is Poisson stable and $\exists \alpha > 0$ such that every point within α-neighbourhood of p is asymptotic to p.*

Example: (Attracting Fixed Points)
Consider the map

$$f(x) = x + \frac{\sin^2 x}{4}, \quad x \in [0, \pi].$$

Netwon's Iteration:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)},$$

$$F(x) = x - \frac{f(x)}{f'(x)}.$$

Definition 1.13 (Super-attracting)**.** *A fixed point of a differentiable map F is said to be **super-attracting** if $F'(x) = 0$.*

Proposition 1.14. *If $|f'(x)| > \delta$ and $|f''(x)| < M$ on a neighbourhood of the root r of f. Then r is a super-attracting fixed point of F.*

Proof. Firstly $F(r) = r - \frac{f(r)}{f'(r)} = r$ is a fixed point, since $f(r) = 0$. And we also ahve

$$F'(x) = 1 - \frac{(f'(x))^2 - f(x) \cdot f''(x)}{(f'(x))^2} = \frac{f(x) \cdot f''(x)}{(f'(x))^2} = 0$$

□

Theorem 1.15. *Suppose (X, d) is a compact metric space, with $f : X \to X$ such that*

$$d(f(x), f(y)) < d(x, y)$$

for all $x, y \in X$. Then f has a fixed point.

Remark 1.15.1. *The result is false when X is not compact. For example, consider $X = [0, \infty)$, and let $f(x) = x + e^{-x}$, and extend f to be an even function on $\mathbb{R}$. Then we can see this map is a weak contraction, yet has no fixed point. Moreover, $d(f^n(x), f^n(y))$ does not converge to zero for some x, y. Or on $\mathbb{R}$ consider $f(x) = \frac{1}{2}x + \sqrt{\frac{1}{4}x^2 + \frac{1}{2}}$.*

Proposition 1.16. *Suppose X is a complete metric space such that the distance function is at most 1, and $f : X \to X$ is such that $d(f(x), f(y)) \leq d(x, y) - \frac{1}{2}(d(f(x), f(y)))^2$. Then f has a unique fixed point $x_0 \in I$ and $\lim_{n\to\infty} f^n(x) = x_0$ for any $x \in X$.*

Proof. Uniqueness of the fixed point is clear. We show existence.
Since

$$d(fx, fy) \leq d(x, y) - \frac{1}{2} d(fx, fy)^2$$

then

$$d(fx, fy) \leq -1 + \sqrt{1 + 2d(x, y)} = \frac{-1 + \sqrt{1 + 2d(x, y)}}{d(x, y)} d(x, y).$$

This tells us that for each $\delta > 0$ there is a θ with $0 < \theta < 1$ such that if $d(x, y) \geq \delta$, then $d(fx, fy) \leq \theta d(x, y)$.

Now let $x \in X$ be arbitrary, let $Y_n = \{f^n(x), f^{n+1}(x), \cdots\}$, we see $Y_{n+1} \subset Y_n$. Let $a_n = \operatorname{diam}(Y_n)$, then $\{a_n\}$ is a monotonically decreasing sequence. Let $\alpha = \lim_{n\to\infty} a_n$, we claim that $\alpha = 0$.

Suppose not, then $\alpha > 0$, then there exists N such that $\alpha \leq a_n \leq \alpha + \epsilon$ for all $n \geq N$. We know there exists $\theta \in (0, 1)$ such that if $d(x, y) \geq \delta$, then $d(fx, fy) \leq \theta d(x, y)$. Then Choose ϵ small enough such that

$$\epsilon < \delta \left(\frac{1-\theta}{\theta} \right).$$

Then if $i, j \geq N + 1$ is such that

$$\delta \leq d(f^i x, f^j x) < \delta + \epsilon,$$

As d does not increase distance, we must have

$$\delta + \epsilon > d(f^{i-1}x, f^{j-1}x) \geq \delta.$$

However, this is a contradiction, since

$$d(f^i x, f^j x) < \theta d(f^{i-1}x, f^{j-1}x) < \delta.$$

Thus we must have that $\delta = 0$. This in particular shows that $\{f^n(x)\}$ is a Cauchy sequence, hence converges to some point $\tilde{x}$, which must clearly be a fixed point of f.

$\square$

1.2 Nondecreasing Interval Map

In this section, we assume all maps are continuous.

Definition 1.17. *If $I \subseteq \mathbb{R}$ is an interval, then $f : I \to \mathbb{R}$ is **increasing** if $x > y \implies f(x) > f(y)$; it is **decreasing** if $x > y \implies f(x) < f(y)$; it is **non-decreasing** if $x \geq y \implies f(x) \geq f(y)$; it is **non-increasing** if $x \geq y \implies f(x) \leq f(y)$.*

Lemma 1.18. *If $I = [\alpha, \beta] \subseteq \mathbb{R}$ is a closed bounded interval and $f : I \to I$ a nondecreasing map without fixed points in (α, β). Then one end point of I is fixed and all orbits converge to this point, except for the other point if it is fixed as well. If f is invertible, then both endpoints are fixed, and all orbits of points in (α, β) are positively asymptotic to the one endpoint and negatively asymptotic to the other.*

Proof. First notice $f(I) \subset I$, so $f(\alpha) \geq \alpha$ and $f(\beta) \leq \beta$. Then $(f - Id)(\alpha) \geq 0$ and $(f - Id)(\beta) \leq 0$, so there exists a zero of $f - Id$ on $[\alpha, \beta]$. By assumption, it can only happen at α or β.

If $f(x) > x$ on (α, β), then $f(\beta) \geq \beta$, so $f(\beta) = \beta$. Let $x_n = f^n(x)$, $x_{n+1} > x_n$. If $x_0 = \lim\limits_{n\to\infty} f^n(x)$, then $f(x_0) = \lim\limits_{n\to\infty} f^{n+1}(x) = x_0$, so $x_0 = \beta$. Similarly, we can consider the case $f(x) < x$. $\square$

Definition 1.19 (Hetero and Homoclinic). *If $f : X \to X$ is an invertible map and $x \in X$ is a point such that*

$$\lim_{n\to\infty} f^{-n}(x) = a \quad \lim_{n\to\infty} f^{n}(x) = b$$

*then x is said to be **heteroclinic** to a and b. If $a = b$, then x is said to be **homoclinic** to a.*

Proposition 1.20. *If $I \subseteq \mathbb{R}$ is a closed bounded interval and $f : I \to I$ is a nondecreasing map. Then all $x \in I$ are either fixed or asymptotic to a fixed point of f. If f is increasing, then all $x \in I$ are either fixed or heteroclinic to adjacent fixed points.*

Proof. The direction of motion is decided by $f - Id$:

- If $(f - Id)(x) < 0$, then $f(x) < x$, and x moves left;
- If $(f - Id)(x) > 0$, then $f(x) > x$, and x moves right.

Let $I = [\alpha, \beta]$. We have $f(I) \subset I$, so $f(\alpha) \geq \alpha$ and $f(\beta) \leq \beta$. Then there exists a fixed point of f. Let $\text{Fix}(f)$ be the collection of fixed points of f, then this collection is closed.

If $\text{Fix}(f) = I$, then there is nothing to prove. Otherwise, $I \setminus \text{Fix}(f) \neq$ can be written as the disjoint union of open intervals. Then on each subinterval, apply Lemma 1.18 (Do need to consider the endpoints of I). ☐

Corollary 1.20.1. *The only possible prime periods of a nondecreasing interval map is* 1.

Proof. Since every non-periodic point is asymptotic to a periodic point. ☐

Definition 1.21 (Repelling Fixed Point). *A fixed point x is said to be a **repelling fixed point** or a repeller if ϵ is sufficiently small, $\forall y$ within ϵ-neighbourhood of x, there is $n \in \mathbb{N}$ such that the positive semiorbit of $f^n(y)$ has no points within the ϵ neighbourhood of x.*

Lemma 1.22. *Let $f : [0, 1] \to [0, 1]$ be a nonincreasing continuous map. Then the only possible prime period of the periodic points of f is* 1 *and* 2.

Proof. Firstly note that 1 and 2 can be periods. Next we note that f^2 is a non-decreasing interval map. ☐

Proposition 1.23. *Let $f : I \to I$ be a continuous interval map. If $\tilde{x}$ is a fixed point of f such that all points are asymptotic to f, then $\tilde{x}$ is an attracting fixed point.*

1.3 Quadratic Maps

In this section we study a specific type of quadratc maps:

$$f_\lambda : [0, 1] \to [0, 1] \quad f_\lambda(x) = \lambda x(1 - x)$$

where $0 < \lambda < 3$.

Proposition 1.24. *For $0 \leq \lambda \leq 1$, all orbits of f_λ on $[0, 1]$ is asymptotic to* 0.

Proof. If $x \neq 0$, $f_\lambda(x) = \lambda x(1 - x) < x$. Hence $\{f_\lambda^n(x)\}$ is a decreasing sequence that is bounded below. Clearly, we also have f_λ maps $[0, 1]$ to $[0, 1]$. The limit must be a fixed point, which is 0. ☐

Proposition 1.25. *For $1 < \lambda \leq 3$, all orbits of f_λ on $[0,1]$ except for 0 and 1 are asymptotic to the fixed point*

$$x_\lambda = 1 - \lambda^{-1}.$$

Proof. We have $f_\lambda(x) = x$ if and only if $x = 0$ or $x = 1 - \lambda^{-1}$. next we consider two cases:

Case 1: $1 < \lambda \leq 2$, $x_\lambda < \frac{1}{2}$. In this case f_λ is increasing and $f_\lambda(x) > x$ for $x \in (0, x_\lambda)$. So every point $x \in (0, x_\lambda)$ is positively asymptotic to x_λ. Next, recall that $f_\lambda(1-x) = f_\lambda(x)$, then $f_\lambda([1-x_\lambda, 1]) \subset [0, x_\lambda] = f([0, x_\lambda])$. So after one iteration, the points in $[1-x_\lambda, 1]$ will be mapped to the interval $[0, x_\lambda]$. This shows every point $x \in [1-x_\lambda, 1)$ is also positively asymptotic to x_λ. Lastly, we consider points $x \in (x_\lambda, 1-x_\lambda)$. We observe

$$f([x_\lambda, 1-x_\lambda]) \subset [x_\lambda, f(1/2)] = [x_\lambda, \lambda/4] \subset [x_\lambda, 1-x_\lambda].$$

Also we have that on $[x_\lambda, 1-x_\lambda]$, $\max|f'_\lambda(x)|$ happens when $x = x_\lambda$, which is ≤ 1 (equal to 1 iff $\lambda = 2$, in this case $x_\lambda = 1 - x_\lambda = 1/2$ and nothing need to be proved). Then by the contraction principle on $[x_\lambda, 1-x_\lambda]$, $\forall x \in [x_\lambda, 1-x_\lambda]$ is asymptotic to x_λ.

Case 2: $2 < \lambda \leq 3$. Let $I = [1-x_\lambda, f_\lambda(1/2)] = [\lambda^{-1}, \lambda/4]$. One can check that $f_\lambda(I) \subset I$. Since $\forall x \in [0,1]$, $f_\lambda(x) \leq f_\lambda(1/2) = \lambda/4$. And $f(\lambda^{-1}) = 1 - \lambda^{-1} > \lambda^{-1}$, so the $f(\lambda^{-1}0$ and $f(\lambda/4)$ all lie in I.

Next, we show that every orbit except that of 0 and 1 eventually enters I. If $x \in I$, then clearly $f_\lambda(I) \subset I$. Next if $x \in (0, \lambda^{-1})$, $x_n = f_\lambda^n(x)$. $f_\lambda([0, \lambda^{-1}) = [0, x_\lambda]$. Suppose $\{x_n\}_{n\in\mathbb{N}} \cap I = \emptyset$, since $f_\lambda(x)$ is increasing on $[0, \lambda^{-1}]$. And $f_\lambda(x) > x$ implies $x_{n+1} > x_n$. So $\lim\limits_{n\to\infty} x_n = x_0 \in [0, \lambda^{-1}]$, where x_0 is a fixed point of f_λ, this is a contradiction. Now if $x \in (\lambda/4, 1)$, it is easy to show that $f_\lambda([\lambda/4, 1]) \subset (0, \lambda/4)$.

Lastly, we show that $\forall x \in I$, $f^n(x) \to x_\lambda$. $f_\lambda([\lambda^{-1}, x_\lambda]) \subset [x_\lambda, \lambda/4]$ and $f_\lambda([x_\lambda, \lambda/4]) \subset [\lambda^{-1}, x_\lambda]$. So $f_\lambda^2([\lambda^{-1}, x_\lambda]) \subset [\lambda^{-1}, x_\lambda]$. Let $J = [1/2, x_\lambda]$, wish to show $f_\lambda^2(J) \subset [1/2, x_\lambda]$. Since f_λ^2 is strictly increasing on J, then if $x \in J$, $f_\lambda^{2n}(x) \to x_\lambda \in J$. Also $f_\lambda([\lambda^{-1}, 1/2]) \subset J$, which shows that $x \in [\lambda^{-1}, 1/2]$, then $f_\lambda^{2n+1}(x) \to x_\lambda$. Also at $x = x_\lambda$, f_λ is a local contraction by Proposition (1.9), hence the desired result follows. $\square$

1.4 Eventually Contracting Map

Definition 1.26 (Eventually Contracting)**.** *A map of metric space (X, d) is said to be* ***eventually contracting*** *if there are constants c, $\lambda \in (0,1)$ such that $\forall n \in \mathbb{N}$, we have*

$$d(f^n(x), f^n(y)) \leq c\lambda^n d(x,y).$$

Proposition 1.27. *If $f : X \to X$ is a map of a metric space and there are $c, \lambda > 0$ such that for all $x, y \in X$ and $\forall n \in \mathbb{N}$, then*

$$d(f^n(x), f^n(y)) \leq c\lambda^n d(x,y).$$

Then for every $\mu > \lambda$, there exists a metric d_μ uniform equivalent (bilipschitz equivalent) to d such that for all $x, y \in X$,

$$d_\mu(f(x), f(y)) \leq \mu d_\mu(x,y).$$

Proof. Let $n_0 \in \mathbb{N}$ such that

$$c \cdot \left(\frac{\lambda}{\mu}\right)^{n_0} < 1$$

Define d_μ as

$$d_\mu(x,y) := \sum_{i=0}^{n_0-1} \frac{1}{\mu^i} d(f^i(x), f^i(y))$$

Then d_μ is a metric, which is called **Lyapunov metric / adapted metric**.

It is clear that

$$d(x,y) \le d_\mu(x,y) \le \sum_{i=0}^{n-1} c\left(\frac{\lambda}{\mu}\right)^i d(x,y) \le \frac{c}{1-\lambda/\mu} d(x,y).$$

Lastly,

$$\begin{aligned} d_\mu(f(x), f(y)) &= \sum_{i=1}^{n_0} \frac{d(f^i(x), f^i(y))}{\mu^{i-1}} \\ &= \mu\left(\sum_{i=0}^{n_0-1} \frac{d(f^i(x), f^i(y))}{\mu^i}\right) + \frac{d(f^{n_0}(x), f^{n_0}(y))}{\mu^{n_0}} - d(x,y) \\ &\le \mu \cdot (d_\mu(x,y) - (1 - c(\lambda/\mu)^{n_0}) d(x,y) \\ &\le \mu \cdot d_\mu(x,y). \end{aligned}$$

□

Corollary 1.27.1. *Let (X,d) be a complete metric space and let $f : X \to X$ be an eventually contracting map. Then under the iterates of f, all points converges to the uniform fixed points of f with exponential speed.*

Proof. If $\lambda \in (0,1)$. Apply Proposition (1.27) with $\mu \in (\lambda, 1)$. We obtain that under the metric d_μ, $d_\mu(f^n(x), f^n(y)) < \mu^n d_\mu(x,y)$. Then Since d_μ and d are bilipschitiz, exists $c > 0$ such that

$$d(f^n(x), f^n(y)) \le c \cdot \mu^n.$$

□

Proposition 1.28. *Suppose $m\rho(x,y) \le d(x,y) \le M\rho(x,y)$, i.e., d and ρ are bilipschitz. If f is eventually contracting map on (X,ρ), then so is it on (X,d) with the same $\lambda \in (0,1)$.*

Proof.

$$\begin{aligned} d(f^n(x), f^n(y)) &\le M\rho(f^n(x), f^n(y)) \\ &\le Mc \cdot \lambda^n \rho(x,y) \\ &\le \frac{Mc}{m} \cdot \lambda^n \cdot d(x,y). \end{aligned}$$

□

Proposition 1.29. *If X and Y are metric spaces. X is complete, and $f : X \times Y \to X$ is a continuous map such* *that*

$$f_y = f(\cdot, y) \;\; \textit{is a } \lambda\textit{-contraction for all } y \in Y.$$

Then the fixed point $g(y)$ of f_y depends continuously on y.

Proof. By triangle inequality, we have

$$\begin{aligned} d(x, g(y)) &\leq \sum_{i=0}^{\infty} d(f_y^i(x), f_y^{i+1}(x)) \\ &\leq \frac{1}{1-\lambda} d(x, f_y(x)). \end{aligned}$$

Let $x = g(y') = f(g(y'), y') = f_{y'}(g(y'))$. Then

$$d(g(y'), g(y)) \leq \frac{1}{1-\lambda} d(f(g(y'), y'), f(g(y'), y)) \leq C_y d(y, y').$$

□

1.5 Fractals

Lemma 1.30. *The Cantor Set C is the collection of numbers in $[0,1]$ that can be written in ternary expansion without using 1 as a digit expansion. That is $x \in C$ if and only if*

$$x = \sum_{i=0}^{\infty} \frac{c_i}{3^i}$$

where $c_i \neq 1$ for all $i \in \mathbb{N}$.

Proof. Attention is all you need. □

Lemma 1.31. *The ternary cantor set C is totally disconnected.*

Proof. This is because the cantor set contains no intervals. □

Lemma 1.32. *The ternary Cantor set C is uncountable.*

Proof. Attention is all you need. □

Definition 1.33 (Cantor Space)**.** *A set homeomorphic to the ternary Cantor set is referred as a* ***Cantor Space****.*

Lemma 1.34. *Assume X is a compact space, $U \subset X$ is an open subset and $\mathcal{C}$ is a collection of closed subsets of X such that*

$$\bigcap_{C \in \mathcal{C}} C \subset U.$$

Then there exists $C_1, \cdots, C_n \in \mathcal{C}$ such that $C_1 \cap \cdots \cap C_n \subset U$.

Proof. Suppose not, then $\mathcal{C} \cup \{X \setminus U\}$ is a collection of closed subsets of X with finite intersection property. But then the intersection of all sets in this collection is nonempty by compactness of X, which contradicts the assumption $\bigcap_{C \in \mathcal{C}} C \subset U$. Then $C = C_1 \cap \cdots \cap C_n$ is clopen, $x \in C$ and $C \subset U \cup V$ □

Theorem 1.35. *If X is a compact Hausdorff space, then $C(x) = Q(x)$ for all $x \in X$, where $C(x)$ is the connected component that contains x and $Q(x)$ is the quasi-component of x defined by*

$$Q(x) := \bigcap_{\substack{C \subset X \, clopen \\ x \in C}} C.$$

Proof. It is clear that $C(x) \subset Q(x)$. Since if Y is connected, C is clopen and $x \in Y \cap C$, then it is easy to see that $Y \subset C$ by connectedness of Y.

Converse, suppose $Q(x) = A \sqcup B$ for some sets A and B that are clopen in $Q(x)$. We assume that $x \in A$ and show that $B = \emptyset$. Since by definition $Q(x)$ is closed, then A and B are closed in X. Therefore they are disjoint compact subsets of a T_2 space. Hence there exists disjoint open sets $U, V \subset X$ such that $A \subset U$ and $B \subset V$. Then there exists clopen sets $C_1, \cdots, C_n$ containing x such that $C_1 \cap \cdots \cap C_n \subset U \cup V$ by lemma (1.34), since $Q(x) \subset U \cap V$.

Consider the set $D = C \cap U$, it contains x, as $x \in A \subset U$. The set D is clearly open, but it is also closed since U and V are disjoint, we can write $D = C \cap (X \setminus V)$. It follows that $Q(x) \subset D$. Since $B = Q(x) \cap V$ and D does not intersect V, we then conclude that $B = \emptyset$. □

Lemma 1.36. *A compact metric space X is totally disconnected if and only if it has a basis of topology consisting of closed open sets.*

Proof. Suppose X has a basis consisting of compact open sets. Then let V be any set that contains more than two points, namely x, y. Since X is Hausdorff, then exists a neighbourhood N of x such that $y \notin X$. Since N is open, then there exists a clopen set W in the basis, such that $x \in W \subset N$. Then $X \setminus W$ is open. Now clearly, $W \cap X \setminus W = \emptyset$, $W \cup X \setminus W \supset V$, both sets are open. Hence V is not connected.

For the converse. Suppose X is compact metrizable that is also totally disconnected. We claim that the collection of all clopen sets form a basis for X. It is clear that this set covers X and is closed under countable union and finite intersection. Now, suppose $x \in U \subset X$ and U is open. Then we find a clopen set V such that $x \in V \subset U$.

Since X is totally disconnected, for each $y \in X \setminus U$, there exists a clopen set W_y such that $x \in W_y$ and $y \notin W_y$. Then the collection $\{X \setminus W_y \,|\, y \in X \setminus U\}$ is an open cover of $X \setminus U$, which admits a finite subcover $\{F_1, \cdots, F_n\}$ since $X \setminus U$ is closed hence compact. Then take $V = \bigcap_{i=1}^n F_i$, we have V is clopen and $x \in V \subset U$. □

Theorem 1.37 (Brouwer's Theorem)**.** *Let C and C' be any two second topological spaces that are non-empty, compact metrizable, perfect and totally disconnected. Then C and C' are homeomorphic.*

Proof. In fact we show that if C satisfies the condition mentioned in the theorem, then C is homeomorphic to $\{0,1\}^{\mathbb{N}}$ with the product topology.

We first show that X has a countable basis of topology consisting of clopen sets. Let $\{U_n\}_{n=1}^{\infty}$ be an arbitrary countable basis (exists, since C is compact metrizable, hence second countable). For every $n \in \mathbb{N}$ consider the set

$$J_n := \{m \in \mathbb{N} \,|\, \bar{U}_m \subset U_n\}$$

For every $m \in J_n$, choose a clopen set V_{nm} such that $\bar{U}_m \subset V_{nm} \subset U_n$, this is possible since the space X has a basis of topology consisting of clopen sets and we need only finitely many such sets to cover $\bar{U}_m$ as it is compact. Then $\{V_{nm} : n \in \mathbb{N}, m \in J_n\}$ is a countable basis of topology, since if $x \in U_n$, then $x \in \bar{U_m} \subset U_n$ for some m, and then $x \in V_{nm} \subset U_n$.

Let $\{W_n\}_{n=1}^{\infty}$ be a countable basis consisting of clopen sets. We will construct nonempty clopen sets $C_{i_1,\cdots,i_n}$, $(n \in \mathbb{N},\ i_k \in \{0,1\})$ satisfying the following properties:

1. $X = C_0 \sqcup C_1$;
2. $C_{i_1,\cdots,i_n} = C_{i_1,\cdots,i_n,0} \sqcup C_{i_1,\cdots,i_n,1}$ for all $n \in \mathbb{N}$ and $i_1, \cdots, i_n \in \{0,1\}$;
3. each W_n is the union of some of the sets $C_{i_1,\cdots,i_n}$, $i_1, \cdots, i_n \in \{0,1\}$.

We do the construction inductively on n.

For $n = 1$, if W_1 is neither empty nor the entire space X, we take $C_0 = W_1$ and $C_1 = X \setminus W_1$. Otherwise, we take any nonempty proper clopen subset of X as C_0 and let $C_1 = X \setminus C_0$.
For the induction step, assume the sets $C_{i_1,\cdots,i_k}$ are defined for all $k \leq n$ and $i_1, \cdots, i_k \in \{0,1\}$. For every n-tuple $(i_1, \cdots, i_n)$ of zeros and ones, if $W_{n+1} \cap C_{i_1,\cdots,i_n}$ is neither empty nor the entire set $C_{i_1,\cdots,i_n}$, we take $C_{i_1,\cdots,i_n,0} = W_{n+1} \cap C_{i_1,\cdots,i_n}$ and $C_{i_1,\cdots,i_n,1} = C_{i_1,\cdots,i_n} \setminus W_{n+1}$. Otherwise, we take any nonempty proper clopen subset of $C_{i_1,\cdots,i_n}$ as $C_{i_1,\cdots,i_n,0}$ and let $C_{i_1,\cdots,i_n,1} = C_{i_1,\cdots,i_n} \setminus C_{i_1,\cdots,i_n,0}$. Note that such a set exists, since by assumption X has no isolated points, so $C_{i_1,\cdots,i_n}$ cannot consists of a single point.

For every point $I = (i_n)_{n=1}^{\infty} \in \{0,1\}^{\mathbb{N}}$, the sets $C_{i_1,\cdots,i_n}$ form a decreasing sequence of nonempty closed sets, hence they have a nontrivial intersection by compactness of X. Let X_I be a point in this intersection, we claim that there are no other points in $\bigcap_{n=1}^{\infty} C_{i_1,\cdots,i_n}$. This is because for any $\tilde{x} \neq x \in X$, there exists n such that $x \in W_n$ and $\tilde{x} \notin W_n$. But then $C_{i_1,\cdots,i_n} \subset W_n$, hence $\tilde{x} \notin C_{i_1,\cdots,i_n}$.

Furthermore, since for every $n \in \mathbb{N}$, the sets $C_{i_1,\cdots,i_n}$ form a partition of X. Then for every point $x \in X$, there exists a unique $I = (i_n)_{n=1}^{\infty} \in \{0,1\}^{\mathbb{N}}$ such that $x \in \bigcap\limits_{n=1}^{\infty} C_{i_1,\cdots,i_n}$.

We define $f : \{0,1\}^{\mathbb{N}} \to X$ by $f(I) = x_I$, where x_I is the unique element in $\bigcap C_{i_1,\cdots,i_n}$. Then this is a bijection by previous analysis. It remains to check that f is a homeomorphism. For this we observe that the sets $C_{i_1,\cdots,i_n}$ form a basis of topology on X, since there unions W_n do. By construction we have

$$f^{-1}(C_{i_1,\cdots,i_n}) = \{(j_m)_{m=1}^{\infty} \in \{0,1\}^{\mathbb{N}} \,|\, j_1 = i_1, \cdots, j_n = i_n\}.$$

The latter sets form a basis of topology on $\{0,1\}^{\mathbb{N}}$. Therefore f defines a bijection between bases of topologies on X and $\{0,1\}^{\mathbb{N}}$. Hence f is a homeomorphism. $\square$

Corollary 1.37.1. *A topological space is a Cantor space if and only if its is non-empty, perfect, compact, totally disconnected (or have a base for their topologies consisting of clopen sets), and metrizable.*

Proposition 1.38. *The ternary Cantor set C is homeomorphic to the cartesian double $C \times C$.*

Proof. Let $x \in C$ be represented by

$$\sum_{n=1}^{\infty} \frac{a_n}{3^n}$$

where $a_n \in \{0, 2\}$. Then define the homeomorphism by

$$x \mapsto \left(\sum_{k=1}^{\infty} \frac{a_{2k-1}}{3^k}, \sum_{k=1}^{\infty} \frac{a_{2k}}{3^k} \right).$$

$\square$

Theorem 1.39 (Space Filling Curves)**.** *There exists a curve $t \mapsto f(t)$ from the unit interval $[0,1]$ to the unit square $[0,1] \times [0,1]$ with the following properties:*

1. *f is continuous and surjective.*
2. *f satisfies a Lipschitz condition of exponent $1/2$, that is,*

$$|f(t) - f(s)| \leq M|t - s|^{1/2}.$$

3. *The image under f of any sub-interval $[a, b]$ in a compact subset of the square of (two-dimensional) Lebesgue measure exactly $b - a$.*

Proof. Let h denote the Cantor-Lebesgue function on $[0,1]$ and T denote the homeomorphism from C to $C \times C$. Let $H : C \times C \to [0,1] \times [0,1]$ be given by

$$H(x, y) = (h(x), h(y)).$$

Since h is continuous and surjective, then H is continuous and surjective. So the map $H \circ T : C \to [0,1] \times [0,1]$ is a continuous surjective map. Lastly, we extend $f = H \circ T$ "linearly". That is, one each of the deleted open interval (a, b), in the construction of the Cantor set, we define the extension part of F on (a, b) to be the line segment within the unit square joining the values $f(a)$ and $f(b)$. In this way, $f : [0,1] \to [0,1] \times [0,1]$ is continuous and surjective. One can also verify that f satisfies other properties stated in the Theorem. $\square$

2 Recurrence and Equidistribution

2.1 Topological Transitivity and Minimality

Definition 2.1 (Invariance Set)**.** *Let X be a topological space, and f be a map from X to X. Then $A \subset X$ is **invariant** if*

$$f^{-1}(A) = A.$$

Definition 2.2 (Topologically Transitive)**.** *A homeomorphism $f : X \to X$ is said to be **topologically transitive** if there exists a point $x \in X$ such that $\overline{\mathrm{Orb}(x)} = X$.*

*If $f : X \to X$ is non-invertible map, then f is said to be **topologically transitive**, if $\overline{\mathrm{Orb}^+(x)} = X$ for some $x \in X$.*

Lemma 2.3.

1. *There exists a complete metric space and a homeomorphism to itself that has a dense orbit but no dense semiorbit.*

2. *There exists a compact metric space and a homeomorphism to itself that has a dense orbit but no dense semiorbit.*

Proof.

1. Consider $X = \mathbb{Z}$ and $x \mapsto x + 1$.

2. Consider X to be the one point compactification of $\mathbb{Z}$, then X is T_3, second countable so metrizable. The homeomorphism $x \mapsto x + 1$, $\infty \mapsto \infty$ has a dense orbit but no dense semiorbit.

□

Theorem 2.4. *Let $f : X \to X$ be homeomorphism of a locally compact separable metric space X onto itself. Then f is topologically transitive if and only if for any two nonempty open sets $U, V \subset X$, there exists an integer $N = N(U, V)$ such that $f^N(U) \cap V$ is nonempty.*

Proof. $\Rightarrow$: Suppose f is topologically transitive we show for any two nonempty open sets $U, V \subset X$, there exists an integer $N = N(U, V)$ such that $f^N(U) \cap V$ is nonempty:

Let $x \in X$ be such that $\mathrm{Orb}(x)$ is dense in X. Then as U and V are nonempty open sets, $\mathrm{Orb}(x)$ has nonempty intersects U and V respectively. This implies there exists $f^n(x) \in U$, $f^m(x) \in V$. Let $N = m - n \in \mathbb{Z}$. Since f is a homeomorphism, in particular its inverse is well-defined. So no matter what is the sign of N, we have $f^N(f^n(x)) = f^{(m-n)}(f^n(x)) = f^m(x)$. Hence

$$f^N(f^n(x)) \in f^N(U) \cap V$$

which proves that the intersection is nonempty.
$\Leftarrow$: On the other hand, we suppose that or any two nonempty open sets $U, V \subset X$, there exists an integer $N = N(U, V)$ such that $f^N(U) \cap V$ is nonempty, we show that f is topologically transitive:

Since X is separable, then it is second countable. Let $\{U_n\}_{n=1}^{\infty}$ be a countable basis for X. Moreover, since X is locally compact, hausdorff, we may assume that U_1 is such that $\overline{U_1}$ is compact. This is possible since by locally compactness, Hausdorff, we know for any $x \in U_1$ there exists open set $x \in O \subset U_1$ such that $O \subset \overline{O} \subset U_1$ and $\overline{O}$ is compact. Then we may add O to our collection of open basis, and reindex to set O be our new U_1. The new set we get is still a base. It clearly covers X, and for any element U_k in the original basis, $O \cap U_k$ is open, hence contains an element in the base $\{U_n\}$. So WLOG, we may assume our countable base $\{U_n\}_{n=1}^{\infty}$ is such that $\overline{U_1}$ is compact.

Now for any $x \in X$ and any open set U containing x, there exists an $n \in \mathbb{N}$ such that $x \in U_n \subset U$. In order to prove f is topologically transitive, suffices to prove that there exists an $p \in X$, such that there exists $n_k \in \mathbb{Z}$ with $f^{n_k}(p) \in U_n$ for any $k > 0$. This would show $\text{Orb}(p)$ is dense in X.

By assumption, there exists an integer N_1 such that $f^{N_1}(U_1) \cap U_2$ is nonempty. Then $U_1 \cap f^{-N_1}(U_2)$ is the nonempty intersection of two open sets (since f is a homeomorphism, it will map open sets to open sets and preimage of open sets is also an open set), so by locally compactness, we can pick $V_1 \subset \overline{V_1} \subset U_1 \cap f^{-N_1}(U_2)$. Since $\overline{V_1} \subset \overline{U_1}$, then $\overline{V_1}$ is compact. Next, since V_1 is nonempty open, then there exists N_2 such that $f^{N_2}(V_1) \cap U_3$ is nonempty. Again we can pick nonempty open set V_2 such that $\overline{V_2} \subset V_1 \cap f^{-N_2}(U_3)$. By induction we construct a nested sequence of open sets V_n such that $\overline{V_{n+1}} \subset V_n \cap f^{-N_{n+1}}(U_{n+2})$. There for we have $\bigcap\limits_{n=1}^{\infty} \overline{V_n} = \bigcap\limits_{n=1}^{\infty} V_n$. However, we have a nested intersection of compact sets, so it must be nonempty. Let p be a point in the intersection, then by assumption on each V_n, we have $f^{N_{n-1}}(p) \in U_n$ for every $n \in \mathbb{N}$. This shows that $\overline{\text{Orb}(p)} = X$ which concludes the proof of the Theorem. □

Remark 2.4.1. *Suppose f is not an invertible map, then the forward direction is still true, that is. If f is topologically transitive, then for any two nonempty open sets $U, V \subset X$, there exists an integer $N = N(U, V)$ such that $f^N(U) \cap V$ is nonempty. However, we need to interpret f^{-N} for $N \in \mathbb{N}$ as taking preimage N times. For the converse, we need a slightly stronger assumption, that is for any two nonempty open sets $U, V \subset X$, there exists a natural number $N = N(U, V)$ such that $f^N(U) \cap V$ is nonempty, then f is topologically transitive. The proof of this is the same as that of the Theorem. Yet in this case, we will always be taking the intersection of U_n and the preimage $f^{-N_{n-1}(U_n)}$ which is open as f is continuous. Then we could conclude that the positive semiorbit of p is dense in X.*

Corollary 2.4.1. *A self-homeomorphism of a locally compact separable metric space is topologically transitive if and only if there are no two disjoint open non-empty f-invariant sets.*

Proof. Suppose f is topologically transitive, then for any two open sets U and V, there exists N such that $f^N(U) \cap V \neq \emptyset$. But if U and V are f-invariant, then $f^N(U) = U$, hence $U \cap V \neq \emptyset$.

For the other direction, if there are no two disjoint open non-empty f-invariant sets. Let $U, V \subset X$ be open, then the invariant open sets $\tilde{U} := \bigcup\limits_{n \in \mathbb{Z}} f^n(U)$ and $\bigcup\limits_{V} := \bigcup\limits_{n \in \mathbb{Z}} f^n(V)$ are not disjoint, so that $f^n(U) \cap f^m(V) \neq \emptyset$ for some $n, m \in \mathbb{Z}$ and $f^{n-m}(U) \cap V \neq \emptyset$. Hence by Theorem 2.4, f is topologically transitive. □

Definition 2.5 (Minimal Maps)**.** *A homeomorphism $f : X \to X$ is said to be **minimal** if the orbit of every point $\alpha \in X$ is dense in X, equivalent, f is minimal if it has no proper closed invariant subset.*

A closed invariant set is said to be ***minimal*** *if it contains no proper closed invariant subsets or, equivalently, if it is the orbit closure of any of its points.*

Lemma 2.6. *Two minimal sets are either disjoint or the same.*

Proof. If $F \cap G \neq \emptyset$, then $F \cap G$ is a closed invariant subset of F, so $F \cap G = F$. $\square$

2.2 Introduction to Ergodic Theory

Definition 2.7 (Measure-Preserving Transformation)**.** *Let (X, Σ, μ) be a σ-finite measure space with a mapping $\tau : X \to X$ such that whenever E is a measurable subset of X, then so is $\tau^{-1}(E)$ and $\mu(\tau^{-1}(E)) = \mu(E)$. In this case, we say that τ is a* ***measure-preserving transformation*** *and (X, Σ, μ, τ) is called a* ***measure-preserving system****. If in addition τ is bijective and τ^{-1} is also a measure-preserving transformation, then τ is referred to as a* ***measure-preserving isomorphism****.*

Lemma 2.8. *If τ is a measure-preserving transformation, then $f(\tau(x))$ is measurable if $f(x)$ is measurable, in fact, we have*

$$\mu(\{x : f(x) > \alpha\}) = \mu(\{x : f(\tau(x)) > \alpha\}).$$

Moreover, $f(\tau(x))$ is integrable if f is integrable, and

$$\int_X f(\tau(x))d\mu(x) = \int_X f(x)d\mu(x).$$

Proof. The first assertion is clear from the definition of τ being measure-preserving. The second statement follows from the identity, which is by definition true for characteristic functions, hence true for all integrable functions. $\square$

Examples:

1. Consider $(\mathbb{Z}, \mathcal{P}(Z), \mu)$, where μ is a counting measure. Then the map $tau : \mathbb{Z} \to \mathbb{Z}$, $n \mapsto n + 1$ is a measure-preserving isomorphism of $\mathbb{Z}$.

2. The translation on $\mathbb{R}^n$ by h is a measure preserving isomorphism. Moreover, the map $x \mapsto x+\alpha$ on $S^1 = \mathbb{R}/\mathbb{Z}$ with the measure induced from Lebesgue measure $\mathbb{R}$ is a measure preserving isomorphism.

3. The map $\tau(x) = 2x \mod 1$ on $\mathbb{R}/\mathbb{Z}$ is a measure preserving transformation.

Let (X, Σ, μ, τ) be a measure preserving system. Then the map $T : H = L^2(X) \to H$, given by

$$T(f)(x) = f(\tau(x))$$

preserves inner product, since we have

$$\|Tf\| = \|f\|.$$

(The identity is clearly true on characteristic functions, hence applies to all L^2 functions). Now by a moment of thought, this map is also injective, since if two L^2 functions f and g are equal a.e., then so is $f(\tau(x))$ and $g(\tau(x))$. Suppose now τ is a measure-preserving isomorphism, then T is in fact bijective, hence becomes an isometry.

Next, one may consider the subspace S of **invariant vectors**, that is

$$S := \{f \in H \,:\, T(f) = f\}.$$

Clearly, this is a closed subspace, since $T - id$ is bounded. Let P denote the orthogonal projection on this subspace, then we have the following result which is stated in a more general form:

Proposition 2.9. *Suppose T is an isometric embedding of a Hilbert space H to itself, and let P be the orthogonal projection on the subspace of the invariant vectors of T. Let*

$$A_n = \frac{1}{n}(Id + T + T^2 + \cdots + T^{n-1}).$$

Then for each $f \in H$, $A_n(f)$ converges to $P(f)$ in norm, as $n \to \infty$.

Proof. We define

$$S_* = \{f \in H \,:\, T^*(f) = f\}$$

where T^* is the transpose of T and

$$S_1 := \{f \in H \,:\, f = g - Tg,\, g \in H\}.$$

Notice that S_* is closed but S_1 is not necessarily closed. We denote its closure by $\bar{S}_1$. Then we claim the following:

- $S = S_*$;
- The orthogonal complement of $\bar{S}_1$ is S.

Since T is an isometric embedding, we have $T^*T = Id$, so if $Tf = f$ then $T^*Tf = T^*f$, which means that $f = T^*f$. To prove the converse inclusion, assume $T^*f = f$, as a consequence $(f, T^*f - f) = 0$, and thus $(f, T^*f) - (f, f) = 0$, that is $(Tf, f) = \|f\|^2$. However, $\|Tf\| = \|f\|$, so the equality for Cauchy-Schwarz inequality occurs, chich shows $Tf = cf$ with $c \in \mathbb{C}$, which shows $Tf = f$, since $(Tf, f) = \|f\|^2$.

Next, we observe that f is in the orthogonal complement of $\bar{S}_1$ exactly when $(f, g - Tg) = 0$, for all $g \in H$. However, this means that $(f - T^*f, g) = 0$ for all g, and hence $f = T^*f$ which tells us that $f \in S$.

Next, to prove the proposition, let $f \in H$, we may write $f = f_0 + f_1$, where $f_0 \in S$ and $f_1 \in \bar{S}_1$. We also fix $\epsilon > 0$ and pick $f_1' \in S_1$ such that $\|f_1 - f_1'\| < \epsilon$. We then write

$$A_n(f) = A_n(f_0) + A_n(f_1') + A_n(f_1 - f_1').$$

For the first term, we recall that P is the orthogonal projection on S, so $P(f) = f_0$, and since $Tf_0 = f_0$ we deduce

$$A_n(f_0) = \frac{1}{n}\sum_{k=0}^{n-1} T^k(f_0) = f_0 = P(f) \quad \text{for every } n \geq 1.$$

For the second term, we recall the definition of S_1 and pick a $g \in \mathcal{H}$ with $f_1' = g - Tg$. Thus

$$A_n(f_1') = \frac{1}{n}\sum_{k=0}^{n-1} T^k(1 - T)(g) = \frac{1}{n}\sum_{k=0}^{n-1} T^k(g) - T^{k+1}(g)$$

$$= \frac{1}{n}(g - T^n(g)).$$

Since T is an isometric embedding, the above identity shows that $A_n(f_1')$ converges to 0 in the norm as $n \to \infty$. For the last term, we use once again the fact that each T^k is an isometry to obtain

$$\left\|A_n(f_1 - f_1')\right\| \leq \frac{1}{n}\sum_{k=0}^{n-1}\left\|T^k(f_1 - f_1')\right\| \leq \|f_1 - f_1'\| < \epsilon.$$

Hence we deduce that $\limsup_{n\to\infty}\|A_n(f) - P(f)\| \leq \epsilon$, and this concludes the proof of the theorem. $\square$

Definition 2.10 (Maximal function)**.** *Let (X, Σ, μ, τ) be a measure preserving system. For any measurable function $f : X \to \mathbb{R}$, we define the corresponding **maximal function** $f^* : X \to \mathbb{R}$ by*

$$f^*(x) := \sup_{1\leq m<\infty} \frac{1}{m}\sum_{=0}^{m-1} |f(\tau^k(x))|.$$

Remark 2.10.1. *The function f^* is automatically measurable since it is the supremum of a countable number of measurable functions.*

Regarding to the maximal function, we have the following weak-type estimate. However, we will omit the proof here (To see an proof, see Stein's Real Analysis Book, Chapter 6 Section 5.2).

Theorem 2.11. *Whenever $f \in L^1(X, \mu)$, the maximal function f^* is finite for almost every x. Moreover, there is a universal constant A so that*

$$\mu(\{x : f^*(x) > \alpha\}) \leq \frac{A}{\alpha}\|f\|_{L^1(X,\mu)} \quad \textit{for all } \alpha > 0.$$

Theorem 2.12. *Suppose f is integrable over a finite measure space (X, μ) and τ is a measure-preserving transformation. Then for almost every $x \in X$, the average $A_m(f) = \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converge to a limit as $m \to \infty$. If further, we denote this limit by $P'(f)$, we have*

$$\int_X |P'(f)(x)| d\mu(x) \leq \int_X |f(x)| d\mu(x).$$

Moreover, we have $P'(f) = P(f)$, where $f \in L^2(X, \mu)$ (recall $P(f)$ is the projection of f onto S).

Remark 2.12.1. *WLOG, in the proof, we will assume that $\mu(X) = 1$, since $\mu(X)$ is finite, we can always normalize the measure. Then in this case, we have $\|f\|_{L^1} \leq \|f\|_{L^2}$ and then L^2 is dense in L^1.*

Proof. We first show that $A_m(f)$ converges to a limit almost everywhere for a set of functions f that is dense in $L^1(X, \mu)$. We then use the maximal theorem to show that this implies the conclusion for all integrable functions.

Now starting with an integrable f and any $\epsilon > 0$, we can write $f = F + H$, where $\|H\|_{L^1} < \epsilon$, and $F = F_0 + (1-T)G$, where both F_0 and G belong to L^2, and $T(F_0) = F_0$, with $T(F_0) = F_0(\tau(x))$. This is possible since L^2 is dense in L^1 under this situation.

Now $A_m(F) = A_m(F_0) + A_m((1-T)G) = F_0 + \frac{1}{m}(1 - T^m(G))$. Note that $\frac{1}{m}T^m(G) = \frac{1}{m}G(\tau^m(x))$ converges to zero as $m \to \infty$ for almost every $x \in X$. Indeed, the series $\sum_{m=1}^{\infty} \frac{1}{m^2}(G(\tau^m(x)))^2$ converges almost everywhere by the monotone convergence theorem, since its integral over X is

$$\sum_{m=1}^{\infty}\frac{1}{m^2}\|T^m G\|_{L^2}^2 = \|G\|_{L^2}^2\sum_{m=1}^{\infty}\frac{1}{m^2},$$

which is finite. This shows that $A_m(F)(x)$ converges for almost every $x \in X$.

Next, since $A_n(f) = A_n(F) + A_n(H)$, and $A_n(H)$ is bounded by the $\|H\|_{L^1} = \epsilon$. As ϵ is arbitrary, then we can in fact show that $A_n(f)(x)$ converges for almost every x. Hence the first half of the theorem is proved.

Now note that if $f \in L^2(X)$, then we know $A_m(f)$ converges to $P(f)$ in the L^2-norm by Proposition (2.9), and hence a subsequence converges almost everywhere to that limit, which proves $P(f) = P'(f)$ in that case. Now if f is just in L^1, then we have

$$\int_X |A_m(f)|dx \le \frac{1}{m}\sum_{k=0}^{m-1}\int_X |f(\tau^k(x))|d\mu(x) = \int_X |f(x)|d\mu(x),$$

and thus since $A_m(f) \to P'(f)$ almost everywhere, we get by Fatou's lemma that $\int_X |P'(f)(x)|d\mu(x) \le \int_X |f(x)|d\mu(x)$. This concludes the proof of the theorem. □

Definition 2.13 (Ergodic Transformations)**.** *We say that a measure-preserving transformation τ of X is* ***ergodic*** *if whenever E is a measurable set that is invariant, that is $\mu(E\Delta\tau^{-1}(E)) = 0$, then either E or E^c has measure zero.*

Remark 2.13.1. *We say that a measurable function f is* ***invariant*** *under τ if $f(x) = f(\tau(x))$ for a.e. $x \in X$. Then we see that τ is ergodic exactly when the only invariant functions are equivalent to constants a.e. To see this, let f be a real-valued invariant function. Then each set $E_\alpha = \{x : f(x) > \alpha\}$ is invariant, hence $\mu(E_\alpha) = 0$ or $\mu(E_\alpha^c) = 0$ for each α.*

Corollary 2.13.1. *Suppose τ is an ergodic measure-preserving transformation. For any integrable function f, we have*

$$\frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x)) \quad \text{converges to} \quad \int_X f d\mu \quad \text{for a.e. } x \in X \text{ as } m \to \infty.$$

Remark 2.13.2. *The interpretation is that the "time average" of f equals to its space average.*

Proof. By Proposition (2.9) we know that the averages $A_m(f)$ converge to $P(f)$, whenever $f \in L^2$, where P is the orthogonal projection on the subspace of invariant vectors. Since in this case the invariant vectors form a one - dimensional space spanned by the constant functions, we observe that $P(f) = 1(f,1) = \int_X f\ d\mu$, where 1 designates the function identically equal to 1 on X. To verify this, note that P is the identity on constants and annihilates all functions orthogonal to constants. Next we write any $f \in L^1$ as $g+h$, where $g \in L^2$ and $\|h\|_{L^1} < \epsilon$. Then $P'(f) = P'(g) + P'(h)$. However, we also know that $P'(g) = P(g)$, and $\|P'(h)\| \le \|h\|_{L^1} < \epsilon$ by the Theorem (2.12). Thus

$$P'(f) - \int_X f\ d\mu = \int_X (g-f)\ d\mu + P'(h)$$

□

Examples:

1. The map $R_\alpha : S^1 \to S^1$, $x \mapsto x + \alpha$ is ergodic, if and only if $\alpha \in \mathbb{R} \setminus \mathbb{Q}$.

2. For $m \geq 2$, and m and integer, $T_m : S^1 \to S^1$, $x \mapsto mx \mod 1$ is ergodic.

Definition 2.14 (Uniquely Ergodic). *If X is a compact metric space and $f : X \to X$ a continuous map, then f is said to be* ***uniquely ergodic*** *if*

$$\frac{1}{n}\sum_{k=0}^{n-1}\varphi(f^k(x))$$

converges to a constant C_φ uniformly (in x) independently of $x \in X$, for every continuous function φ.

Proposition 2.15 (Equivalent Definition of Uniquely Ergodic). *The following are equivalent:*

1. *T is uniquely ergodic;*

2. *For every $\varphi \in C(X)$,*

$$\frac{1}{n}\sum_{k=0}^{n-1}\varphi(f^k(x))$$

converges to a constant C_φ independent of $x \in X$.

3. *The convergence holds for every φ in a dense subset of $C(X)$.*

Under any of these assumptions, the constant C_φ is $\int_X f d\mu$, where μ is the unique T-invariant measure.

2.3 Circle Rotation

We use S^1 to denote the following set:

$$\{z \in \mathbb{C} : |z| = 1\} = \{e^{2\pi i\phi} : \phi \in \mathbb{R}\}.$$

This can be considered as a multiplicative model for the circle. We can alternatively define $S^1 = \mathbb{R}/\mathbb{Z}$, which is an additive model. There is an isomorphism between the two groups given $z \mapsto e^{2\pi i z}$.

Definition 2.16 (Rotation Operator). *Let $\theta \in \mathbb{R}$, then we define the* ***rotation operator*** *in the multiplicative model of S^1, by*

$$R_\alpha(z) = e^{2\pi i\theta}z.$$

Let $\alpha \in \mathbb{R}$, we define the ***rotation operator*** *in the additive model of S^1, by*

$$R_\alpha(z) = z + \alpha \mod 1.$$

Lemma 2.17. *If $\alpha = \frac{p}{q}$ is a rational number, then the orbit of any $z \in \mathbb{R}$ under R_α over the additive model is finite. In particular, every point $z \in \mathbb{R}$ is periodic under R_α.*

Proof. Since $R_\alpha^q(z) = z + p \mod 1 = z$. □

Lemma 2.18. *If $\alpha \in \mathbb{R}$ is irrational, then for any $z \in \mathbb{R}$, z is not periodic under R_α over the additive model.*

Proposition 2.19 (Density of Orbits). *If $\alpha \in \mathbb{R}$ is irrational, then every positive semi-orbit of R_α (over the multiplicative group) is dense in S^1.*

Proof. Proof 1: Suppose that $x, z \in S^1$. WTS $z \in \overline{\mathrm{Orb}^+(x)}$. Let $\epsilon > 0$, we show $z \in O_\epsilon(\mathrm{Orb}^+(x))$. There exists $k \in \mathbb{N}$ such that $k > \frac{1}{\epsilon^2}$. Then consider a subset of the positive orbit

$$\{x, x+\alpha, \cdots, x+k\alpha\}.$$

Since all points are distinct, then by the Pigeon Hole Principle, there exists $m, n \in \mathbb{N}$, $0 \le m, n \le k$ such that

$$d(R_\alpha^m(x), R_\alpha^n x) \le \frac{1}{k} < \frac{1}{\epsilon}.$$

So $d(x, R_\alpha^{n-m}(x)) < \epsilon$. Let $\beta = (n-m)\alpha$, then

$$\{x, x+\beta, x+2\beta, \cdots\}$$

is a subset of $\mathrm{Orb}^+(x)$, and z is within ϵ neighbourhood of this set, since this set forms what is called a **ϵ-net on** S^1.

Proof 2: Let A be an invariant closed subset of S^1. Then $S^1 \setminus A$ is open so is the union of disjoint open intervals. Let I be the largest of those intervals. We claim that $R_\alpha^n(I)$ are disjoint with each other. If $R_\alpha^n(I) \cap I \neq \emptyset$, then for all $x \in (R_\alpha^N(I) \cup I)$, $x \notin A$, which contradicts the maximality of length of I since $\alpha \notin \mathbb{Q}$. Nevertheless, this is not possible, since I has a positive length. This shows that the invariance closed set is either the empty set or all of S^1. Lastly, notice that $\overline{\mathrm{Orb}^+(x)}$ is an invariance closed subset of S^1. □

Corollary 2.19.1. *If $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, then every negative semi-orbit of R_α is dense in S^1.*

Notation: we denote

$$F_\Delta(x, n) = \mathrm{Card}\,\{k \in \mathbb{Z} \,:\, 0 \le k < n,\ R_\alpha^k x \in \Delta\}$$

where $\Delta \subset S^1$. We have seen that if Δ is an arc on S^1, then $F_\Delta(x, n) \to \infty$ as $n \to \infty$. We will show equidistribution

$$\lim_{n\to\infty} \frac{F_\Delta(x,n)}{n} = \ell(\Delta)$$

when Δ is an arc, and ℓ is the length function.

Proposition 2.20. *Suppose α is irrational and $R_\alpha : S^1 \to S^1$. Let Δ and Δ' be arcs such that $\ell(\Delta) < \ell(\Delta')$. Then there exists an $N_0 \in \mathbb{N}$, such that if $x \in S^1$ and $n \in \mathbb{N}$, then*

$$F_{\Delta'}(x, n+N_0) \ge F_\Delta(x, n).$$

Proof. By the density proposition, $\exists N_0$ such that $R_\alpha^{N_0}(\Delta) \subset \Delta'$. Then if $R_\alpha^n(x) \in \Delta$, $R_\alpha^{n+N_0} \in \Delta'$. So the proposition follows. □

Notation: Let

$$\overline{f_x}(A) := \limsup_{n\to\infty} \frac{F_A(x,n)}{n}$$
$$\underline{f_x}(A) := \liminf_{n\to\infty} \frac{F_A(x,n)}{n}.$$

Corollary 2.20.1. *If $\ell(\Delta) < \ell(\Delta')$. Then $\overline{f_x}(\Delta) \le \overline{f_x}(\Delta')$.*

Lemma 2.21. *Let Δ_1, Δ_2 be arcs that is closed on starting end and open on the end point. Then if the end point of Δ_1 coincides with the starting point of Δ_2, then $\Delta_1 \cap \Delta_2 = \emptyset$ and $\Delta_1 \cup \Delta_2$ is an arc, moreover,*

$$F_{\Delta_1}(x,n) + F_{\Delta_2}(x,n) = F_{\Delta_1 \cup \Delta_2}(x,n).$$

Remark 2.21.1. *In fact, this is true as long as $\Delta_1 \cap \Delta_2 = \emptyset$, even if they are not arcs.*

Proof. Clearly. □

Lemma 2.22. *If Δ is an arc and $\ell(\Delta) = \frac{1}{k}$, then $\overline{f_x}(\Delta) \leq \frac{1}{k-1}$.*

Proof. Let $\Delta_1, \cdots, \Delta_{k-1}$ be $k-1$ disjoint arcs with length $\frac{1}{k-1}$. For every $i \in \{1, \cdots, k-1\}$, there exists N_i such that

$$F_{\Delta_i}(x, n+N_i) \geq F_{\Delta}(x,n).$$

Let $N = \max_i N_i$. Then

$$F_{\Delta_i}(x, n+N) \geq F_{\Delta}(x,n) \quad i = 1, \cdots, k-1.$$

So

$$\sum_{i=1}^{k-1} F_{\Delta_i}(x, n+N) \geq (k-1) F_{\Delta}(x,n).$$

Hence

$$F_{S^1}(x, n+N) \geq (k-1) \cdot F_{\Delta}(x,n).$$

So

$$\frac{F_{\Delta}(x,n)}{n} \leq \frac{n+N}{n(k-1)}.$$

Taking limit sup on both sides, we get

$$\overline{f_x}(\Delta) \leq \frac{1}{k-1}.$$

□

Corollary 2.22.1. *If Δ is an arc and $\ell(\Delta) = \frac{p}{k}$, $p \leq k$, then $\overline{f_x}(\Delta) \leq \frac{p}{k-1}$*

Theorem 2.23 (Equidistribution)**.** *For any arc $\Delta \subset S^1$, and any $x \in S^1$.*

$$f(\Delta) = \lim_{n \to \infty} \frac{F_{\Delta}(x,n)}{n} = \ell(\Delta)$$

and the limit is uniform in x.

Proof. Let $\Delta \subset S^1$ be an arc. $\forall \epsilon > 0$, there exists natural number p, k such that there exists an arc δ' satisfying

$$\Delta \subset \Delta', \;\; \ell(\Delta') = \frac{p}{k}, \;\; \frac{p}{k} < \ell(\Delta) + \epsilon.$$

Then

$$\overline{f_x(\Delta)} \leq \overline{f_x(\Delta')} \leq \frac{p}{k-1} < \frac{\ell(\Delta)+\epsilon}{k-1} \cdot k.$$

Let $\epsilon \to 0$, then $k \to \infty$,

$$\overline{f_x}(\Delta) \leq \ell(\Delta).$$

Lastly, let $A = \Delta^C$, we conclude $\overline{f_x}(\Delta^c) \leq \ell(\Delta^c)$ which implies $\underline{f_x}(\Delta) \geq \ell(\Delta)$. □

Notation: Let A be the union of arcs, then we define

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}.$$

Then

$$F_A(x,n) = \sum_{i=0}^{n-1} \chi_A(R_\alpha^i x).$$

Moreover,

$$\lambda(A) = \int_{S^1} \chi_A(x)dx.$$

Then Theorem (2.23) can be rewritten as follows: Let $\Delta \subset S^1$ be an arc, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \chi_\Delta(R_\alpha^i x) = \int_{S^1} \chi_\Delta(x)dx.$$

Definition 2.24 (Birkhoff Averaging Operator)**.** *We define the* ***Birkhoff averaging operator*** $\mathcal{B}_n$ *by*

$$\mathcal{B}_n(\phi) = \frac{1}{n} \sum_{k=0}^{n-1} \phi \circ (R_\alpha^k).$$

Then

$$\mathcal{B}_n(\phi)(x) = \frac{1}{n} \sum_{k=0}^{n-1} \phi(R_\alpha^k x).$$

Remark 2.24.1. *We would only be considering the case where α is irrational, as the case where $\alpha \in \mathbb{Q}$ is non-interesting since $R_\alpha^k = R_\alpha$ for some $k \in \mathbb{N}$.*

The following properties are easy to verify for the operator $\mathcal{B}_n$:

Proposition 2.25.

1. *$\mathcal{B}_n$ is linear: $\mathcal{B}_n(a\phi + b\psi) = a\mathcal{B}_n(\phi) + b\mathcal{B}_n(\psi)$.*

2. *$\mathcal{B}_n$ is nonnegative: if $\phi \geq 0$, then $\mathcal{B}_n(\phi) \geq 0$.*

3. *$\mathcal{B}_n$ is positive: if $\phi > 0$, then $\mathcal{B}_n(\phi) > 0$.*

4. *$\mathcal{B}_n$ is nonexpanding: $\sup\limits_{x\in S^1} \mathcal{B}_n(\phi)(x) \leq \sup\limits_{x\in S^1} \phi(x)$.*

5. *$\mathcal{B}_n$ preserves the average: $\int_{S^1} \mathcal{B}_n(\phi)(x)dx = \int_{S^1} \phi(x)dx$, since*

$$\int_{S^1} \phi(R_\alpha^k x)dx = \int_{S^1} \phi(x)dx.$$

Proposition 2.26. *For any step function ϕ that is a linear combination of characteristic functions of arcs, then*

$$\lim_{n\to\infty} \mathcal{B}_n(\phi)(x) = \int_{S^1} \phi(x)dx, \quad \forall x \in S^1.$$

Proof. This follows from Linearity of $\mathcal{B}_n$ and Theorem 2.23. $\square$

Proposition 2.27. *For any function ϕ that is uniformly limit of step functions we have*

$$\lim_{n\to\infty} B_n(\phi)(x) = \int_{S^1} \phi(x)dx \quad \forall x \in S^1.$$

Proof. By Uniform Convergence and the nonexpanding property of $\mathcal{B}_n$. $\square$

Proposition 2.28. *Every continuous function is the uniform limit of step functions. The same holds for functions with finitely many discontinuity points and with one-side limit at these points.*

Proof. For any continuous function on S^1. It is uniformly continuous. $\forall \epsilon > 0$, $\exists n \in \mathbb{N}$ such that for every arc I of length less than $\frac{1}{n}$, then $\forall x, y \in I$, $|f(x) - f(y)| < \epsilon$. Then approximate each arc. $\square$

Corollary 2.28.1. *If f is continuous or has finitely many discontinuity points and with one-side limit at these points. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(R_\alpha^i x) = \int_{S^1} f(y)dy \quad \forall x \in S^1.$$

Proposition 2.29. *If $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ and f is Riemann integrable then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(R_\alpha^i x) = \int_{S^1} f(y)dy.$$

Proof. Since we can approximate the Riemann integrable with step function uniformly by considering the upper and lower Riemann sums. $\square$

We can also prove the same result using Fourier Series.

Lemma 2.30. *If f is continuous function on S^1, then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(R_\alpha^i(x)) \to \int_{S^1} f(x)dx.$$

Proof. First check the statement holds for functions f being $1, e^{2\pi i x}, \cdots, e^{2\pi i m x}, \cdots$. Then approximate continuous functions with trigonometric polynomials. $\square$

Proposition 2.31. *If α is irrational and f is Riemann integrable, then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(R_\alpha^k(x)) = \int_{S^1} f(x)dx$$

uniformly in x.

Theorem 2.32. *If $f \in L^1(S^1)$, then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} f(R_\alpha^i(x)) \to \int_0^1 f(x)dx$$

for almost every $x \in S^1$.

Proof. Since $L^2(S^1)$ is dense in $L^1(S^1)$ because $C_c^\infty(S^1)$ is dense in both spaces and the trigonometric polynomials is dense in $L^2(S^1)$, then we conclude that the trigonometric polynomials is dense in $L^1(S^1)$. Then by approximation L^1 functions using trigonometric polynomials, we have the desired result. □

Proposition 2.33. *Let $k, p \in \mathbb{N}$ and k is not a power of 10. Then there exists an $n \in \mathbb{N}$ such that p gives the initial digits of the decimal expansion of k^n.*

Proof. We want to show that there exists an $\ell \in \mathbb{N}$ such that

$$10^\ell \cdot p \leq k^n < 10^\ell(p+1).$$

This is equivalent to showing

$$\ell + \log p \leq n \log k < \ell + \log(p+1)$$

where the log is taken with base 10. Since k is not a power of 10, then $\log k$ is irrational. Let $m = \lfloor \log p \rfloor + 1$, then

$$1 - m \log p \leq 1 - m - l + \log k < 1 - m + \log(p+1).$$

Since we know there exists n such that

$$\log(p/10^{m-1}) \leq \{n \log k\} \leq \log((p+1)/10^{m-1})$$

the Proposition follows. □

Proposition 2.34. *For $k \in \mathbb{N}$, k not a power of 10, and $p \in \mathbb{N}$. Let $F_p^k(n)$ denote the number of integers i between 0 and $n-1$ such that p gives the initial digits of the decimal expansion of k^i. Then*

$$\lim_{n\to\infty} \frac{F_p^k(n)}{n} = \log(p+1) - \log p.$$

Proof. Since the length of the interval given by

$$\log(p/10^{m-1}) \leq \{n \log k\} \leq \log((p+1)/10^{m-1})$$

is $\log(p+1) - \log p$. Then by Theorem 2.23, the proposition follows. □

2.4 Translation on the Torus

Definition 2.35 (Topological Group)**.** *Let X be a topological space and also a group with respect to operation $\cdot$. Then $(X, \cdot)$ is a* ***topological group*** *if the following map are continuous:*

$$\cdot : X \times X, \quad (x, y) \mapsto x \cdot y$$

$$inv : X \times X, \quad x \mapsto x^{-1}$$

Proposition 2.36. *If the translation map $L_{x_0} : X \to X$, $L_{x_0}(x) = x_0 \cdot x$ on a topological group X is topologically transitive, then it is minimal.*

Proof. For $x, x' \in G$, denote $A, A' \subset X$ the closures of the orbits of x and x', respectively. Now $x_0^n x' = x_0^n x(x^{-1}x')$, so $A' = Ag^{-1}g'$ and $A' = X$ if and only if $A = X$. $\square$

Lemma 2.37. *If the self-homeomorphism $f : X \to X$ is topologically transitive, then there is no f-invariant nonconstant continuous function $\varphi : X \to \mathbb{R}$.*

Proof. Let $\varphi : X \to \mathbb{R}$ be f-invariant, that is $\varphi(f(x)) = \varphi(x)$ for all $x \in X$. Since it is not a constant, there exists $t \in \mathbb{R}$ such that both sets $\{x \in X \,|\, \varphi(x) > t\}$ and $\{x \in X \,|\, \varphi(x) < t\}$ are nonempty. Since φ is invariant, these sets are also invariant. Since φ is continuous, they are open. This contradicts Corollary 2.4.1. $\square$

Lemma 2.38. *If $f : \mathbb{T}^n \to \mathbb{R}$ is L^2, then*

$$f(x) = \sum_{(k_1,\cdots,k_n)\in\mathbb{Z}^n} c_{k_1,\cdots,k_n} \exp\left(2\pi i \sum_{j=1}^{n} k_j x_j\right), \quad a.e.$$

Proposition 2.39. *Let $\gamma = (\gamma_1, \cdots, \gamma_n) \in \mathbb{T}^n$ and define the translation*

$$T_\gamma(x_1, \cdots, x_n) = (x_1 + \gamma_1, x_2 + \gamma_2, \cdots, x_n + \gamma_n) \mod 1.$$

Then the translation T_γ is minimal if and only if the numbers $\gamma_1, \cdots, \gamma_n$ and 1 are rationally independent.

Proof. Suppose $\gamma_1, \cdots, \gamma_n$ and 1 are rationally independent. That is $\sum\limits_{i=1}^{n} k_i\gamma_i = k$ have non-trivial solutions. Then we construct a continuous T_γ-invariant function to show that T_γ is not topologically transitive. Define $\varphi(x) = \sin 2\pi \left(\sum\limits_{i=1}^{n} k_i x_i\right)$. It is a nonconstant function on $\mathbb{T}^n$ and invariant since

$$\begin{aligned}
\varphi(T_\gamma x) &= \sin(2\pi \sum k_i(x_i + \gamma_i)) \\
&= \sin(2\pi \sum k_i x_i + 2\pi k) = \varphi(x).
\end{aligned}$$

For the converse, let U and V be two open nonempty f-invariant sets. We show that $U \cap V$ is nonempty. Let χ be the characteristic function of U, then

$$\chi(T_\gamma x) = \chi(x)$$

since U is invariant under T_γ. Taking the Fourier expansion, we have

$$\chi(x_1, \cdots, x_n) = \sum_{k_1,\cdots,k_n)\in\mathbb{Z}^n} c_{k_1,\cdots,k_n} \exp\left(2\pi i \sum_{j=1}^{n} k_j x_j\right).$$

Then

$$\begin{aligned}
\chi(T_\gamma x) &= \chi(x_1 + \gamma_1, \cdots, x_n + \gamma_n) \\
&= \sum_{(k_1,\cdots,k_n)\in\mathbb{Z}^n} \chi_{k_1,\cdots,k_n} \exp\left(2\pi i \sum_{j=1}^{n} k_j(\chi_j + \gamma_j)\right) \\
&= \sum_{(k_1,\cdots,k_n)\in\mathbb{Z}^n} \chi_{k_1,\cdots,k_n} \exp\left(2\pi i \sum_{j=1}^{n} k_j \gamma_j\right) \exp\left(2\pi i \sum_{j=1}^{n} k_j x_j\right)
\end{aligned}$$

$$= \sum_{(k_1,\cdots,k_n)\in\mathbb{Z}^n} c_{k_1,\cdots,k_n} \exp\left(2\pi i \sum_{j=1}^{n} k_j x_j\right)$$

By the uniqueness of Fourier expansion, we must have that

$$c_{(k_1,\cdots,k_n)}\left(1-\exp\left(2\pi i\sum_{j=1}^{n} k_j\gamma_j\right)\right)=0.$$

By the linear independence of $\gamma_1, \cdots, \gamma_n$ and 1 this shows that $c_{k_1,\cdots,k_n} = 0$ except possibly when $k_1 = \cdots = k_n = 0$. This shows $\chi \equiv 1$ outside on a set of Lebesgue measure zero. Thus we conclude that $T^n \setminus V$ has Lebesgue measure zero, which proves $V \cap U$. Then by corollary (2.4.1), T_γ is topologically transitive. Hence minimal by Proposition (2.36).

□

Corollary 2.39.1. *$\{n^2\alpha\}$ is equidistributed on S^1 if α is irrational.*

Proposition 2.40. *For any translation T_γ and any $x \in \mathbb{T}^n$, the closure $C(x)$ of the orbit of x is a finite union of tori of dimension k, $0 \le k \le n$, and that the restriction of T_γ to $C(x)$ is minimal.*

Proof. Suffices to show for the case where $x = 0 \in \mathbb{T}^n$, since we have translation invariance. Let k be the dimension of $\gamma_1, \cdots, \gamma_n, 1$ when $\mathbb{R}$ is considered as a $\mathbb{Q}$ vector space. WLOG, we may assume $\{1, \gamma_1, \cdots, \gamma_k\}$ is a basis for the subspace spanned by $\{\gamma_1, \cdots, \gamma_n, 1\}$, otherwise we can just permute the coordinates (this is a homeomorphism). Now since for $k < i \le n$, γ_i is in the span of $\{1, \gamma_1, \cdots, \gamma_k\}$, then there exists integers such that

$$c_{i,1}\gamma_1 + c_{i,2}\gamma_2 \cdots + c_{i,k}\gamma_k + c_{i,i}\gamma_i \in \mathbb{Z}.$$

In particular, we can minimize $c_{i,i}$ to be the smallest such possible positive integer.

We define the map from $\tau : \mathbb{T}^n \to \mathbb{T}^n$ by

$$(x_1,\cdots,x_k,\cdots,x_n) \mapsto (x_1,\cdots,x_k, x_{k+1} - \frac{1}{c_{k+1,k+1}}\sum_{j=1}^{k} c_{k+1,j}x_j, \cdots, x_n - \frac{1}{c_{n,n}}\sum_{j=1}^{k} c_{n,j}x_j).$$

This is clearly a homeomorphism, hence in order to study $C(0)$, we just need to study $\tau(C(0))$. Note that

$$\tau(\mathrm{Orb}(0)) \subset \mathbb{T}^k \times G$$

where G is a discrete finite abelian subgroup of $\mathbb{T}^{n-k}$. The first k coordinates clearly lies in $\mathbb{T}^k$. Now we show that for $k < i \le n$, $\pi_i(\tau(\mathrm{Orb}(0))$ can only take finitely many values. Since

$$c_{i,1}\gamma_1 + c_{i,2}\gamma_2 \cdots + c_{i,k}\gamma_k + c_{i,i}\gamma_i \in \mathbb{Z}.$$

Then $\pi_i(\tau(T_\gamma^{j_1}(0)))$ and $\pi_i(\tau((T_\gamma^{j_2}(0)))$ are equal iff $c_{i,i}|j_1 - j_2$. This shows that the each $\pi_i(\tau(\mathrm{Orb}(0)))$ is a discrete subspace of $\mathbb{T}^1$, in fact, we can show it is isomorphic to $\mathbb{Z}/c_{i,i}\mathbb{Z}$. Then by taking closure, we conclude that $\tau(C(0)) \subset \mathbb{T}^n \times G$. To show the other direction. Let M be the lcm of $\{c_{i,i}\}_{i>k}$. Then we note the set $\{M\gamma_1, \cdots, M\gamma_k, 1\}$ is still $\mathbb{Q}$ linearly independent. Then for $1 \le j \le M$, we have that projection of the $Orb_{T'}((j\gamma_1, \cdots, j\gamma_k, \cdots, j\gamma_n))$ is dense in $\mathbb{T}^k$ where T' is the translation on $\mathbb{T}^n$ by the element $M\gamma = (M\gamma_1, \cdots, M\gamma_k, \cdots, M\gamma_n)$. Note that

$$\mathrm{Orb}_{T'}((j\gamma_1, \cdots, j\gamma_n)) \subset \mathrm{Orb}_{T_\gamma}(0).$$

And for each j, the projection of the $n-k+1$th coordinates up till nth coordinates on to $\mathbb{T}^{n_k}$ is a constant. Hence we conclude that $\mathrm{Orb}_{T_\gamma}(0)$ is dense in $T^n \times G$. This completes the proof of the problem. □

Proposition 2.41. *The map $A_\alpha : \mathbb{T}^2 \to \mathbb{T}^2$, $A_\alpha(x, y) = (x + \alpha, y + x) \mod 1$ is topologically transitive if and only if α is irrational.*

Proof. If α is rational, then fix any $(x, y) \in \mathbb{T}^2$. Let x_0 be such that $B(x_0, \epsilon) \notin \mathrm{Orb}_{R_\alpha}(x)$. Then note that the strip

$$H := \{(p, q) \in \mathbb{T}^2 \,|\, p \in (x_0 - \epsilon, x_0 + \epsilon),\, q \in \mathbb{T}^1\}$$

do not intersect the orbit of (x, y) under A_α.

On the other hand, suppose α is irrational. Let U, V be two nonempty invariant open set. And consider the characteristic function of U, which is invariant under A_α, so

$$\chi(x, y) = \chi(A_\alpha(x, y)).$$

Taking Fourier series on both side, we get

$$\sum_{(k_1,k_2)\in\mathbb{Z}^2} c_{k_1,k_2} \exp[2\pi i(k_1 x + k_2 y)] = \sum_{(k_1,k_2)\in\mathbb{Z}^2} c_{k_1,k_2} \exp(2\pi i k_1 \alpha) \exp[2\pi i(k_1 x + k_2 x + k_2 y)]$$

By uniqueness of Fourier series, we must have that

$$c_{k_1+k_2,k_2} = c_{k_1,k_2} \exp(2\pi i k_1 \alpha).$$

Now suppose there exists $(k_1, k_2) \in \mathbb{Z}^2 \setminus \{0\}$ such that $c_{k_1,k_2} \neq 0$, then $c_{k_1+mk_2,k_2} \neq 0$ for all $m \in \mathbb{N}$. Moreover, $|c_{k_1+mk_2,k_2}|^2 = |c_{k_1,k_2}|^2$. However, this contradicts Parseval's Identity, since we must have

$$\sum_{(k_1,k_2)\in\mathbb{Z}^2} |c_{k_1,k_2}|^2 = \int_{\mathbb{T}^2} \chi^2(x, y) dx dy.$$

The right hand side is finite. This concludes that χ is a constant function except on a measure zero set. So $\chi = 1$ a.e., which concludes the proof of the proposition. □

Definition 2.42 (n-parallelepiped)**.** *Let $\Delta_1, \cdots, \Delta_n$ be arcs in $\mathbb{T}^1$, then $\Delta := \Delta_1 \times \cdots \times \Delta_n$ is called an **n-parallelepiped**. The **volume** of Δ, $\mathrm{vol}(\Delta)$, is defined as the product of the lengths of the arcs $\Delta_1, \cdots, \Delta_n$.*

Notation: Let Δ be a parallelepiped, and $x \in \mathbb{T}^k$, $\gamma \in \mathbb{T}^k$, then we use

$$F_\Delta(x, n) := \mathrm{card}\{k \in \mathbb{Z} \,:\, 0 \leq k \leq n - 1,\, T_\gamma^k(x) \in \Delta\}.$$

Definition 2.43 (Uniformly distributed). *A sequence $(x_m)_{m\in\mathbb{N}}$ in $\mathbb{T}^k$ is said to be uniformly distributed if*

$$\lim_{m\to\infty}\frac{\operatorname{card}\{k\in\{1,\cdots,m\}\,|\,x_k\in\Delta\}}{m}=\operatorname{vol}(\Delta)$$

for every k-parallelepiped $\Delta\subset\mathbb{T}^n$.

Similar to the One Dimensional Torus Translation. We can prove the following results. We assume that the components and γ together with 1 is rationally independent.

Proposition 2.44. *Let $\Delta=\Delta_1\times\cdots\times\Delta_k$ and $\Delta'=\Delta_1'\times\cdots\times\Delta_k'$ such that $\ell(\Delta_i)<\ell(\Delta_i')$ for $i\in\{1,\cdots,k\}$. Then there is an $N_0\in\mathbb{N}$ which depends on Δ,Δ', and γ, such that if $x\in\mathbb{T}^k$ and $N\geq N_0$ and $n\in\mathbb{N}$, then*

$$F_{\Delta'}(x,n+N)\geq F_{\Delta}(x,n).$$

Theorem 2.45. *Let φ be any Riemann-integrable function on T^k. Then*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\varphi(T_\gamma^i(x))=\int_{\mathbb{T}^k}\varphi(z)dz.$$

uniformly $\forall x\in\mathbb{T}^k$.

Corollary 2.45.1. *The semiorbit of the translation $T_\gamma(x)$ is equidistributed if and only if $\{\gamma_1,\cdots,\gamma_k,1\}$ is rationally independent.*

2.5 Poincaré Classification Theorem for Circle Maps

Let $\pi:\mathbb{R}^1\to S^1$ be the canonical projection,that is $\pi(x)=\{x\}$.

Proposition 2.46. *If $f:S^1\to S^1$ is continuous, then there exists a continuous $F:\mathbb{R}\to\mathbb{R}$ called as a lift of f to $\mathbb{R}$ such that*

$$f\circ\pi=\pi\circ F.$$

Such a lift is unique up to additive integer. The number $F(x+1)-F(x)$ is an integer independent $x\in\mathbb{R}$ and the lift F.

Proof. Let $p\in S^1$ be arbitrary. Then $p=\{x_0\}$ for some $x_0\in\mathbb{R}$. Define $f(p)=\{y_0\}$ for some $y_0\in\mathbb{R}$. We want to define $F:\mathbb{R}\to\mathbb{R}$ such that

- $F(x_0)=y_0$;
- F is continuous;
- $f(\{z\})=\{F(z)\}$, $\forall z\in\mathbb{R}$.

Since f is uniformly continuous, there exists $\delta>0$ such that if $d(\{x\},\{x'\})<\delta$, then $d(f(\{x\}),f(\{x\}))<\frac{1}{2}$. We define F on $(x_0-\delta,x_0+\delta)$. Since $z\in(x_0-\delta,x_0+\delta)$, then there exist a unique $y\in(y_0-\frac{1}{2},y_0+\frac{1}{2})$ such that $\{y\}=f(\{x\})$. We define $F(x)=y$. Then we can extend F to the entire real line.

Next, we show the uniqueness of F up to additive constant. Suppose $\tilde{F} : \mathbb{R} \to \mathbb{R}$ is another lift, then

$$\{\tilde{F}(x)\} = f(\{x\}) = \{F(x)\},$$

so $F - \tilde{F} \in \mathbb{Z}$. Hence $F - \tilde{F}$ is a constant as they are continuous. Then it is also clear that $F(x+1) - F(x)$ is independent of the lifting F. Suppose $x \in \mathbb{R}$ is arbitrary, then

$$\{F(x+1)\} = f(\{x+1\}) = f(\{x\}) = \{F(x)\}.$$

So $F(x+1) - F(x) \in \mathbb{Z}$. Since F is continuous, then this is a constant. So $F(x+1) - F(x)$ is independent of F and x. $\square$

Definition 2.47 (Degree)**.** *We define the degree of a continuous map $f : S^1 \to S^1$, by $\deg(f) = F(x+1) - F(x)$, where F is any lift of f to $\mathbb{R}$, $x \in \mathbb{R}$ is arbitrary.*

Remark 2.47.1. *If $\deg(f) = 0$, then $F(x+1) - F(x) = 0$, so F is not monotone so f is not monotone. This implies that f is not invertible (not injective).*

Remark 2.47.2. *We note that if f is a homeomorphism, then $\deg(f) = \pm 1$. Since if $|\deg(f)| > 1$, $|F(x+1) - F(x)| > 1$, so there exists $y \in [x, x+1]$ such that $F(y) = F(x) + 1$ by the IVT. Then $f(\{y\}) = f(\{x\})$ which contradicts the injectivity of f. In particular, we conclude that if $|\deg(f)| \neq 1$, then f is not injective.*

Definition 2.48 (Orientation Preserving)**.** *Suppose f is invertible, if $\deg(f) = 1$, then we say that f is **orientation preserving**. If $\deg(f) = -1$, then f is said to **reverse orientation**.*

Lemma 2.49. *If f is an orientation preserving circle homeomorphism and F is a lift, then for all $x, y \in \mathbb{R}$,*

$$F(y) - y \leq F(x) - x + 1.$$

Proof. Let $k = \lfloor y - x \rfloor$, then

$$\begin{aligned} F(y) - y &= F(y) + F(x+k) - F(x+k) + (x+k) - (x+k) - y \\ &= (F(x+k) - (x+k)) + (F(y) - F(x+k)) - (y - (x+k)) \end{aligned}$$

Since $\deg(f) = 1$, then $F(x+k) - (x+k) = F(x) - x$. Since $0 \leq y - (x+k) < 1$, then $F(y) - F(x+k) \leq 1$ (otherwise, it is not a homeomorphism). So we conclude that $F(y) - y \leq F(x) - x + 1$. $\square$

Lemma 2.50. *If a sequence $\{a_n\}$ satisfies $a_{n+m} \leq a_n + a_m + L$ for all $m, n \in \mathbb{N}$ and some uniform constant L. Then*

$$\lim_{n\to\infty} \frac{a_n}{n}$$

exists and

$$\lim_{n\to\infty} \frac{a_n}{n} \in \{-\infty\} \cup \mathbb{R}.$$

Proof. Since $a_{n+m} \leq a_n + a_m + L$, then by induction, we have $a_n \leq n \cdot a_1 + n \cdot L$. So

$$\liminf_{n\to\infty} \frac{a_n}{n} \leq a_1 + L$$

Let $b, c \in \mathbb{R}$ be arbitrary such that $\liminf \frac{a_n}{n} < b < c$. We show that $\limsup \frac{a_n}{n} < c$. Notice

$$\frac{a_\ell}{\ell} \le \frac{ka_n + a_r + k\ell}{\ell} = \frac{ka_n}{\ell} + \frac{a_r}{\ell} + \frac{kL}{\ell}$$

if $\ell = nk + r$ for $0 \le r < n$. Then

$$\frac{a_\ell}{\ell} \le \frac{a_n}{n} + \frac{a_r}{\ell} + \frac{L}{n}. \tag{2.1}$$

Let $n > \frac{2L}{c-b}$ such that $\frac{a_n}{n} < b$. Then for any $\ell \ge n$, with $\ell(c-b) \ge 2 \max_{r<n} a_r$. By (2.1), we have $\limsup \frac{a_n}{n} < c$. □

Proposition 2.51. *Let $f : S^1 \to S^1$ be an orientation-preserving homeomorphism and $F : \mathbb{R} \to \mathbb{R}$ a lift of f. Then*

$$\rho(F) := \lim_{|n|\to\infty} \frac{1}{n}(F^n(x) - x)$$

exists for all $x \in \mathbb{R}$, and $\rho(F)$ is independent of x and well-defined up to an integer. That is, if $\tilde{F}$ is another lift, then $\rho(F) - \rho(\tilde{F}) = F - \tilde{F} \in \mathbb{Z}$. $\rho(F)$ is rational if and only if f has a periodic point.

Remark 2.51.1. *This also holds if f is only monotone and admits a lift F such that $F(x+1) = F(x) + 1$ for all $x \in S^1$.*

Proof. We first prove the independence of x if the limit exists. By assumption, $F(x+1) = F(x) + 1$. Then if $x, y \in [0, 1)$, then $|F(x) - F(y)| < 1$, since we have a homeomorphism. Then

$$\left|\frac{1}{n}(F^n(x) - x) - \frac{1}{n}(F^n(y) - y)\right| \le \frac{1}{n}|F^n(x) - F^n(y)| + \frac{1}{n}|x - y| \le \frac{2}{n}.$$

Existence of Limit: Let $a_n = F^n(x) - x$. By definition,

$$a_{n+m} = F^{n+m}(x) - x = F^m(F^n(x)) - F^n(x) + F^n(x) - x = F^m(F^n(x)) - F^n(x) + a_n.$$

Since f is orientation preserving homeomorphism, then so is f^m. Since F^m lifts f^m, then

$$F^m(F^n(x)) - F^n(x) \le F^m(x) - x + 1 = a_m + 1,$$

So we have

$$a_{n+m} \le a_m + a_n + 1.$$

Then by Lemma (2.50),

$$\lim_{n\to\infty} \frac{1}{n}(F^n(x) - x)$$

exits and belongs to $\{-\infty\} \cup \mathbb{R}$. Notice that

$$\frac{1}{n}(F^n(x) - x) = \frac{1}{n}\sum_{i=0}^{n-1}[F^{i+1}(x) - F^i(x)] = \frac{1}{n}\sum_{i=0}^{n-1}(F(F^i(x)) - F^i(x)) \ge \min(F(y) - y).$$

This shows the limit is a real number.

Next we show that $\rho(F) - \rho(\tilde{F}) = F - \tilde{F} \in \mathbb{Z}$. Notice that

$$\rho(F + m) = \lim_{n\to\infty} \frac{1}{n}(F^n(x) + mn - x) = \rho(F) + m$$

Since any for $\tilde{F}$ that also lifts f, we have $\tilde{F} = F + m$ for some m. Then $\rho(\tilde{F}) = \rho(F) + \tilde{F} - F$.

Finally we proceed to show the last assertion in the Proposition. If f has a periodic point with minimum period q. Then $F^q(x) = x + p$ for some $p \in \mathbb{Z}$. So for $m \in \mathbb{N}_+$,

$$\frac{1}{mq}(F^{mq}(x) - x) = \frac{1}{mq} \cdot mp = \frac{p}{q}.$$

Since a subsequence of the convergent sequence $\{\frac{1}{n}(F^n(x) - x)\}$ converges to p/q, we conclude that the limit of the whole sequence if $p/q \in \mathbb{Q}$.

On the other hand if $\rho(F)$ is rational, then f has a periodic point. $\forall m \in \mathbb{N}$,

$$\rho(F^m) = \lim_{n\to\infty} \frac{1}{n}(F^{mn}(x) - x) = \lim_{n\to\infty} \frac{mn}{m}(F^{mn}(x) - x) = m\rho(F).$$

Then if $\rho(F) = p/q$, $\rho(F^q) = p$, which implies $\rho(f^q) = 0$. We claim that if $\rho(g) = 0$, then g has a fixed point. This would imply that f has a periodic point with period q.

Suppose g does not have a fixed point and let G be a lift of G with $G(0) \in [0, 1)$. Since $g(x) \neq x$, then $G(x) - x \notin \mathbb{Z}$. So for all $x \in \mathbb{R}$, $G(x) - x \in (0, 1)$ by the continuity of G. Moreover by compactness of S^1, there exists $\delta > 0$ such that

$$0 < \delta \leq G(x) - x \leq 1 - \delta < 1 \quad \forall x \in \mathbb{R}.$$

Now

$$\begin{aligned} G^n(0) = G^n(0) - 0 &= \sum_{i=0}^{n-1} [G^{i+1}(0) - G^i(0)] \\ &= \sum_{i=0}^{n-1} [G(G^i(0) - G^i(0)] \end{aligned}$$

So we conclude that $n\delta \leq G^n(0) \leq n(1 - \delta)$. But this gives $\rho(G) \in [\delta, 1 - \delta]$, which contradicts $\rho(g) = 0$. Hence we conclude that g must have a fixed point. □

Definition 2.52 (Rotation Number)**.** *We define the* ***rotation number*** *of a homeomorphism f by $\rho(f) = \{\rho(F)\}$, where $\{\cdot\}$ denotes the fractional part of the $\rho(F)$. By the previous Proposition, this is independent of the representative F and x we choose.*

Proposition 2.53. *Let $f : S^1 \to S^1$ be an orientation-preserving homeomorphism. Then all periodic points have the same period. That is if $\rho(f) = \{\frac{p}{q}\}$, $p, q \in \mathbb{Z}$ with $\gcd(p, q) = 1$, then the lift F of f such that $\rho(F) = \frac{p}{q}$ satisfies*

$$F^q(x) = x + p$$

whenever $\{x\}$ is a periodic point. In this case q is the prime period of all periodic point of f.

Remark 2.53.1. *This also holds if f is only monotone and admits a lift F such that $F(x+1) = F(x)+1$ for all $x \in S^1$.*

Proof. If $\{x\}$ is a periodic point. Then $F^r(x) = x + s$ for some $r, s \in \mathbb{Z}$. We know

$$\frac{p}{q} = \rho(F) = \lim_{n\to\infty} \frac{F^{nr}(x) - x}{nr} = \lim_{n\to\infty} \frac{ns}{nr} = \frac{s}{r}.$$

Hence there exists $m \in \mathbb{Z}$ such that $s = mp$, $r = mq$. So $F^{mq}(x) = x + mp$. We claim that $F^q(x) = x + p$. Suppose this is not the case, WLOG, assume $F^q(x) - p > x$. Then, since F is periodic and monotone, we have

$$F^{2q}(x) - 2p = F^q(F^q(x) - p) - p \geq F^q(x) - p > x.$$

where we have used the fact that $F^q(x) + 1 = F^q(x+1)$. Then we would get $F^{mq}(x) - mp > x$, similarly, we have $F^{mq}(x) - mp < x$ if $F^q(x) - p < x$. Thus we must have $F^q(x) = x + p$, so x is a periodic point with period q. □

Proposition 2.54. *$\rho(\cdot)$ is continuous in the C^0 topology for orientation preserving circle homeomorphism.*

Proof. Let $\rho(f) = \rho$, and $\frac{p'}{q'}, \frac{p}{q} \in \mathbb{Q}$ such that $\frac{p'}{q'} < \rho < \frac{p}{q}$. Pick the lift F of f for which $-1 < F^q(x) - x - p \leq 0$ for some $x \in \mathbb{R}$. Then $F^q(x) < x + p$ for all $x \in \mathbb{R}$ since otherwise $F^q(x) = x + p$ for some $x \in \mathbb{R}$ and then $\rho = \frac{p}{q}$. Since the function $F^q = id$ is periodic and continuous, it attains its maximum. Thus there exists $\delta > 0$ such that $F^q(x) < x + p - \delta$ for all $x \in \mathbb{R}$. This implies that every sufficiently small perturbation $\bar{F}$ of F in the uniform topology satisfies $\bar{F}^q(x) < x + p$ for all $x \in \mathbb{R}$. Therefore, $\rho(\bar{f}) < \frac{p}{q}$, where $\bar{f}$ is the circle homeomorphism lifted by $\bar{F}$. Similarly, we can show for the upper bound p'/q'. □

Definition 2.55 (Topological Conjugacy)**.** *Suppose $g : X \to X$, $f : Y \to Y$ are maps of metric space X, Y. Suppose there is a continuous surjective map $h : X \to Y$ such that*

$$h \circ g = f \circ h.$$

*Then f is said to be a **factor** of g via the factor map h. If h is a homeomorphism, then f and g are said to be **conjugate** and h is **conjugacy**. In this case, we have*

$$h \circ g \circ h^{-1} = f.$$

Proposition 2.56. *If $f, h : S^1 \to S^1$ are orientation preserving homeomorphisms. Then*

$$\rho(h^{-1} \circ f \circ h) = \rho(f).$$

Remark 2.56.1. *This also holds if f is only monotone and admits a lift F such that $F(x+1) = F(x)+1$ for all $x \in S^1$.*

Proof. Define F and H to be the lift of f and h, then

$$\pi \circ F = f \circ \pi \quad \text{and} \quad \pi \circ H = h \circ \pi.$$

Then

$$\pi \circ H^{-1} = h^{-1} \circ h \circ \pi \circ H^{-1} = h^{-1} \circ \pi \circ H \circ H^{-1} = h^{-1} \circ \pi.$$

This shows that H^{-1} lifts h^{-1}. Next,

$$\pi \circ H^{-1} \circ F \circ H = h^{-1} \circ \pi \circ F \circ H = h^{-1} \circ f \circ \pi \circ H = h^{-1} \circ f \circ h \circ \pi.$$

This shows that $H^{-1} \circ F \circ H$ lifts $h^{-1} \circ f \circ h$. Hence, the proof would be complete if we can show

$$|(H^{-1} \circ F \circ H)^n(x) - F^n(x)| = |H^{-1} \circ F^n \circ H(x) - F^n(x)|$$

is $o(n)$.

If $x \in [0, 1)$, we can choose $H(0) \in (-1, 0]$, hence

$$-2 < H(x) - x < H(x) < H(1) < 2.$$

Recall $H(x) - x$ is a periodic function in x, so $H(x+1) - (x+1) = H(x) - x$, then $|H(x) - x| < 2$ for all $x \in \mathbb{R}$. Similarly, we can show $|H^{-1}(x) - x| < 2$ for all $x \in \mathbb{R}$. We also claim that if $|y - x| < 2$, then $|F^n(y) - F^n(x)| \leq 3$. Since f is orientation-preserving homeomorphism, so is f^n. Since F^n lifts f^n, then $\forall x, y \in \mathbb{R}$, we have

$$F^n(y) - y \leq F^n(x) - x + 1.$$

So

$$-3 \leq y - x - 1 \leq F^n(y) - F^n(x) \leq y - x + 1 \leq 3$$

when $|y - x| < 2$.

Lastly,

$$\begin{aligned}
&|H^{-1} \circ F^n \circ H(x) - F^n(x)| \\
&\leq |H^{-1} \circ (F^n \circ H(x)) - (F^n \circ H(x))| + |F^n \circ H(x) - F^n(x)| \\
&\leq 2 + 3 = 5
\end{aligned}$$

by the previous inequalities we have established. Thus

$$\frac{1}{n}|H^{-1} \circ F^n \circ H(x) - F^n(x)| \leq \frac{5}{n},$$

this shows $|\rho(f) - \rho(h^{-1} \circ f \circ h)| = 0$. □

Remark 2.56.2. *What happens if h is order reversing, we would get $\rho(h^{-1} \circ f \circ h) = -\rho(f)$. This comes from substituting the following into the original proof of the theorem:*

$$\begin{aligned}
&|H^{-1} \circ F^n \circ H(x) + F^n(x)| \\
&= |H^{-1} \circ (F^n \circ H(x)) + (F^n \circ H(x))| + |F^n \circ H(x) - F^n(x)| \\
&\leq 2 + 3
\end{aligned}$$

Since h^{-1} is order reversing, then $H^{-1}(y) + y \leq 2$ for all $y \in \mathbb{R}$.

Definition 2.57 (Ordering on the Unit Circle)**.** *Given $x_0, x_1, \cdots, x_{n-1} \in S^1$, take $\tilde{x}_0$ as a lift of x_0 and $\tilde{x}_1, \cdots, \tilde{x}_{n-1} \in [\tilde{x}_0, \tilde{x}_0 + 1) \subset \mathbb{R}$ such that*

$$\{\tilde{x}_i\} = x_0,$$

*then the **ordering** of $(x_0, x_1, \cdots, x_{n-1})$ on S^1 is the permutation σ of $\{0, 1, \cdots, n-1\}$ such that*

$$\tilde{x}_0 < \tilde{x}_{\sigma(1)} < \tilde{x}_{\sigma(2)} < \cdots < \tilde{x}_{\sigma(n-1)} < \tilde{x}_0 + 1.$$

Example: Suppose $p, q \in \mathbb{N}$ such that p, q are coprime. Define

$$E := \left\{0, \frac{p}{q}, \frac{2p}{q}, \cdots, \frac{(q-1)p}{q}\right\}.$$

Let $\pi(E) \subset S^1$. Then we have the following observation: since $(p, q) = 1$, by Bezout's Lemma, there exists a unique $k \in \mathbb{N}$ such that $0 < k < q$ such that $kp \equiv 1 \mod q$. Then we define $\sigma(1) = k$. Then we can define $\sigma(i) = k_i$, where $k_i p \equiv i \mod q$, so $k_i = k \cdot i \mod q$.

Proposition 2.58. *Let $f : S^1 \to S^1$ be an orientation preserving homeomorphism. Suppose $p, q > 1 \in \mathbb{N}$ are copime, and $\rho(f) = \{\frac{p}{q}\}$. Then we know there is an $x \in S^1$ such that $f^q(x) = x$. Moreover, the ordering of*

$$\{x, f(x), \cdots, f^{q-1}(x)\}$$

on S^1 is given by $\sigma(i) = k \cdot i \mod q$ where $kp = 1 \mod q$.

Proof. We can fix $\tilde{x} \in \pi^{-1}(\{x\})$ and lift F of f such that

$$F^q(\tilde{x}) = \tilde{x} + p.$$

Let $A := \pi^{-1}(\{x, f(x), \cdots, f^{q-1}(x)\})$ and $I = [\tilde{x}, \tilde{x} + p]$ .Then the $q - 1$ points in $\{F^i(\tilde{x})\}$ partitions I:

$$I = \bigcup_{i=0}^{q-1} [F^i(\tilde{x}), F^{i+1}(\tilde{x})]$$

into q partitions. Similarly, $A \cap I$ partitions I into pq partitions (since there are $pq - 1$ points of A lies in the interior of I).
We have the following observations:

- F is bijective from $I_i := [F^i(\tilde{x}), F^{i+1}(\tilde{x})]$ to $[F^{i+1}(\tilde{x}), F^{i+2}(\tilde{x})]$.
- F preserves A, so each I_i contains exactly $p + 1$ points of A.

Take $k, r \in \mathbb{Z}$ such that the closest right neighbour of $\tilde{x}$ in A is $\tilde{x}_1 = F^k(\tilde{x}) - r$ (note every $p \in A$ can be expressed in this form). In particular, since $q > 1$, then $k \neq q$. Let $\bar{F} = F^k - r$, which is an increasing map since F^k is increasing. Since $k \neq 1$, then it is clear that $\bar{F}$ also preserves A. Then by the increasing property of $\bar{F}(\tilde{x})$, $\tilde{x}_1 = \bar{F}(\tilde{x})$ is the nearest right neighbour of $\tilde{x}$ in A, and similarly we can show $\tilde{x}_n = \bar{F}^n(\tilde{x})$ is the n^{th} nearest right neighbour of $\tilde{x}$. Since $[\tilde{x}, F(\tilde{x})]$ is divided into p subintervals by A. Then $\bar{F}^p(\tilde{x}) = F(\tilde{x})$, so $f^{kp}(x) = f(x)$, which shows $kp \equiv 1 \mod q$. Lastly, the general expression for $\sigma(i)$ follows from induction. □

Proposition 2.59. *Let $f : S^1 \to S^1$ be orientation preserving homeomorphism with $\rho(f) = \{\frac{p}{q}\}$, $(p,q) = 1$. Then there are two types of nonperiodic orbits of f:*

1. *If f has exactly one periodic orbit, then every other point is heteroclinic under f^q to two points on the periodic orbit. These points are different if the period is greater than one. (If the period is one, then all orbits are homoclinic to the fixed point).*

2. *If f has more than one periodic orbit. Then each nonperiodic orbit is heteroclinic under f^q to two points on different periodic orbits.*

Proof. Let F be a lift of f and z a fixed point of f^q. We identify f^q with the homeomorphism $F^q - p$ on $[z, z+1]$. Then $F^q - p$ is a non-decreasing interval map, thus every point on $[z, z+1]$ is either fixed or heteroclinic to adjacent fixed points. So everything except for the last part of (2) follows.

Now suppose that there is an $I = [a,b] \subset \mathbb{R}$ such that a, b are adjacent zeros of $F^q - id - p$ and a, b projects onto the same periodic orbit. Then we show f only has one periodic point. Since we must have $f^k(\pi(a)) = \pi(b)$, for some $k \in \mathbb{N}$ with $(k,q) = 1$. Since $f^q(\pi(a)) = \pi(a)$, then

$$S^1 \setminus \{f^n(\pi(a))\}_{n=0}^{q-1} = \bigcup_{n=0}^{q-1} f^{nk}(\pi(a,b)).$$

The right hand side has no periodic point, by invariance, and the fact that a, b are adjacent zeros. The set $\{f^n(\pi(a))\}_{n=0}^{q-1}$ corresponds to the periodic point $x = \pi(a)$. Hence, we conclude that f has only one periodic point $x = \pi(a)$. □

Lemma 2.60. *If $I \subset \mathbb{R}$ is an interval whose endpoints are adjacent zeros of $F^q - id - p$, where F is the lift of some orientation preserving homeomorphism. Then $F^q - id - p$ has the same sign on the intervals of I and $F(I)$.*

Proof. If $F^q - id - p > 0$ on I, then $F^q(x) > x + p$ for all $x \in I$. $F^q(F(x)) = F(F^q(x)) > F(x+p) = F(x) + p$. So $F^q - id - p > 0$ on $F(I)$. □

Proposition 2.61 (Irrational Rotation)**.** *Let $F : \mathbb{R} \to \mathbb{R}$ be a lift of an orientation-preserving homeomorphism $f : S^1 \to S^1$ with $\rho = \rho(F) \notin \mathbb{Q}$. Then for $n_1, n_2, m_1, m_2 \in \mathbb{Z}$, $n_1 > n_2$, and $x \in \mathbb{R}$,*

$$n_1\rho + m_1 < n_2\rho + m_2$$

if and only if

$$F^{n_1}(x) + m_1 < F^{n_2}(x) + m_2.$$

Proof. Since $\rho(F) \notin \mathbb{Q}$, there is no periodic point of f, and this happens iff $F^{n_1}(x) - F^{n_2}(x) \notin \mathbb{Z}$ for all $x \in \mathbb{R}$, $n_1, n_2 \in \mathbb{Z}$. Given $n_1, n_2, m_1, m_2 \in \mathbb{Z}$, $n_1 > n_2$, then

$$F^{n_1}(x) + m_1 - F^{n_2}(x) - m_2$$

never changes sign.

WLOG, let $F^{n_1}(x) + m_1 < F^{n_2}(x) + m_2$ for all $x \in \mathbb{R}$. Let $y = F^{n_2}(x)$, then

$$F^{n_1-n_2}(y) - y < m_2 - m_1$$

for all $y \in \mathbb{R}$. If $y = 0$, then $F^{n_1-n_2}(0) < m_2 - m_1$. Then

$$F^{2(n_1-n_2)}(0) = F^{(n_1-n_2)}(F^{n_1-n_2}(0)) < F^{n_1-n_2}(0) + m_2 - m_1 < 2(m_2 - m_1).$$

Doing this inductively, we have $F^{n(n_1-n_2)}(0) < n(m_2 - m_1)$.

Recall

$$\rho = \lim_{n\to\infty} \frac{F^{n(n_1-n_2)}(0)}{n(n_1-n_2)} \leq \lim_{n\to\infty} \frac{n(m_2-m_1)}{n(n_1-n_2)} = \frac{m_2-m_1}{n_1-n_2}.$$

Then $(n_1 - n_2)\rho < m_2 - m_1$ if and only if $n_1\rho + m_1 < n_2\rho + m_2$. □

Lemma 2.62. *Let $f : S^1 \to S^1$ be an orientation-preserving homeomorphism without periodic points. Suppose $m, n \in \mathbb{Z}$, $m \neq n$, $x \in S^1$ and $I \subset S^1$ is a closed interval with end points $f^m(x)$ and $f^n(x)$. Then every semiorbits meets I.*

Proof. We prove the statement for positive semi-orbit, the case for negative semiorbit is proved analogously.

Suffices to show $S^1 \subset \bigcup_{k\in\mathbb{N}} f^{-k}(I)$. Define

$$I_k := f^{-k(n-m)}(I),$$

then if $k \in \mathbb{N}$, then I_k and I_{k-1} have one common ending point. $I_k = f^{-k(n-m)}(I) = [f^{-k(n-m)}f^m(x), f^{-k(n-m)}f^n(x)]$, and $I_{k-1} = f^{-(k-1)(n-m)}[I] = [f^{-(k-1)(n-m)}f^n(x), f^{-(k-1)(n-m)}f^n(x)]$. Since f is an order preserving homeomorphism, then $I_k \cap I_{k-1} = \{f^{-k(n-m)}f^m(x)\}$. If $S^1 \neq \bigcup_{k\in\mathbb{N}} I_k$, this shows the endpoint sequence converges to some $z \in S^1$. Then

$$\begin{aligned} z &= \lim_{k\to\infty} f^{-k(n-m)}(f^m(x)) \\ &= \lim_{k\to\infty} f^{-(k+1)(n-m)}(f^m(x)) \\ &= \lim_{n\to\infty} f^{-(n-m)} \circ f^{-k(n-m)}(f^{,}(x)) \\ &= f^{-(n-m)} \lim_{k\to\infty} f^{-k(n-m)}(f^{,}(x)) \\ &= f^{-(n-m)}(z) \end{aligned}$$

which is a periodic point of f, hence we have a contradiction. □

Theorem 2.63 (Poincaré Classification Theorem)**.** *Let $f : S^1 \to S^1$ be an orientation-preserving homeomorphism with irrational rotation number ρ. Then there is a continuous monotone map $h : S^1 \to S^1$ with the property that*

$$h \circ f = R_\rho \circ h.$$

1. *If f is transitive, then h is a homeomorphism.*
2. *If f is not transitive, then h is not invertible.*

Proof. Pick a lift $F : \mathbb{R} \to \mathbb{R}$ of x and for $x \in \mathbb{R}$ define

$$B_x := \{F^n(x) + m \ : \ n, m \in \mathbb{Z}\}$$

then B_x is the total lift of $\{f^n(x)\}$. Define $H : B_x \to \mathbb{R}$ by

$$H(F^n(x) + m) = n\rho + m.$$

By Proposition (2.61) H is monotone. Since the set $\{n\rho + m\}$, $\rho \in \mathbb{Q}$ is dense in $\mathbb{R}$, then $H(B_x)$ is dense in $\mathbb{R}$. Let $\tilde{R}_p = \mathbb{R} \to \mathbb{R}$ as $\tilde{R}_p(x) = x + \rho$, then on B_x, we check $H \circ F = \tilde{R}_\rho \circ H$:

$$H \circ F(F^n(x) + m) = H(F^{n+1}(x) + m) = (n+1)\rho + m = \tilde{R}_\rho \circ H(F^n(x) + m).$$

We claim that H has a continuous extension to the closure $\bar{B}_x$ of B_x. Extend the value of domain of H to $\bar{B}_x$ by limits, since H is monotone, then the left and right hand side limit exits. Since $H(B_x)$ is dense in $\mathbb{R}$, then the left and right limit must coincide.

We define H to be the piecewise constant on the complementary of $\bar{B}_x$, where on each open interval, the constant is equal to the values of H at the endpoints. This gives $H : \mathbb{R} \to \mathbb{R}$ such that

$$H \circ F = \tilde{R}_p \circ H.$$

For all $z \in \bar{B}_x$, $H(z+1) = H(F^n(x) + m + 1) = np + m + 1 = H(z) + 1$. Then by extension, this holds for all $z \in \mathbb{R}$.

Suppose f is topologically transitive, then there exists x such that $\mathrm{Orb}(x)$ is dense in S^1. Then B_x is dense in $\mathbb{R}$, in this case $h = \pi(H)$ is invertible. Now if f is not topologically transitive, then H is not invertible since it is not injective as $\mathbb{R} \setminus \overline{B_x}$ is always nonempty for any $x \in S^1$. Hence h is not invertible.

□

Definition 2.64 (ω-limit Set). *We define the set*

$$\omega(x) := \bigcap_{n \in \mathbb{N}} \overline{\{f^i(x) \ : \ i \geq n\}}$$

*this is set of the accumulation points of the positive semiorbits of x is called the **ω-limit set** of x.*

Definition 2.65 (Nowhere Dense). *A subset of a topological space is called **nowhere dense** or **rare** if its closure has empty interior. Equivalent, a subset S of a topological space X is said to be nowhere dense in X if S is not dense in any nonempty open subset U of X.*

Lemma 2.66. *The boundary of a set is closed, and the boundary of an open set is nowhere dense. Hence the boundary of a boundary of a set is nowhere dense.*

Proposition 2.67. *Let $f : S^1 \to S^1$ be an orientation-preserving circle homeomorphism without periodic points. Then $\omega(x)$ is independent of x and $E = \omega(x)$ is perfect and is either S^1 or nowhere dense.*

Proof. We first show $\omega(x)$ is independent of x. Let $x, y \in S^1$, we show $\omega(x) = \omega(y)$ by showing for any $z \in \omega(x)$, $z \in \omega(y)$. Then there exists a strictly increasing sequence $\{\ell_n\} \subset \mathbb{N}$, such that $f^{\ell_n}(x) \to z$ as $n \to \infty$. Recall Lemma (2.62), by this lemma, we know there exists $k_m \in \mathbb{N}$ such that $f^{k_m}(y) \in [f^{\ell_m}(x), f^{\ell_{m+1}}(x)]$. Then this shows that

$$\lim_{n\to\infty} f^{k_m}(y) = z$$

so $z \in \omega(y)$. By symmetry, we have $\omega(x) = \omega(y)$.

Next we consider S^1 endowed with the subspace topology from $\mathbb{R}^2$. We show E is either S^1 or nowhere dense. We claim that E is the smallest non-empty closed f-invariant set. Firstly notice that E is closed and f-invariant. Then by nested compact sets, we get E is non-empty. If $\emptyset \neq A \subset S^1$ is closed and f-invariant. Let $x \in A$, then $\{f^k(x)\} \subset A$. This in particular shows that

$$\omega(x) \subset \overline{\{f^k(x)\}_{k\in\mathbb{Z}}} = A$$

since A is closed.

This also shows that the only f-invariant set of E can only be E and $\emptyset$. Since $\partial E = E \setminus \mathring{E}$. If $x \in \partial E$ and $x \in U$ is a neighbourhood, then $U \cap E = \emptyset$ and $U \setminus E \neq \emptyset$. Since f is a homeomorphism, then $f(U) \cap f(E) \neq \emptyset$ and $f(U) \setminus f(E) \neq \emptyset$. In particular, $f(E) = E$, this shows that ∂E is a closed f-invariant subset of E. So $\partial E = E$ (E is nowhere dense) or $\partial E = \emptyset$ (E is S^1).

Lastly, we show that E is perfect. Let $x \in E$, we show x is a limit point of E. Since $E = \omega(x)$, then there exists $\{k_n\} \subset \mathbb{N}$ such that

$$f^{k_n}(x) \to x \quad \text{as} \quad n \to \infty.$$

As there is no periodic points, $f^{k_n}(x) \neq x$, for all $n \in \mathbb{N}$. So x is an accumulation point of E, since E is f-invariant, then $f^{k_n}(x) \in E$. □

2.5.1 The Denjoy Example

Definition 2.68 (Wondering Point). *A point is said to be* ***wondering*** *if it has a neighbourhood all of whose images and preimages are pairwise disjoint.*

Remark 2.68.1. *We take the images and preimages infinitely many times. I.e., the forward orbit of the neighbourhood and the backward orbit are all pairwise disjoint.*

Proposition 2.69 (Denjoy Example). *For $\rho \in \mathbb{R} \setminus \mathbb{Q}$, there is a nontransitive C^1-diffeomorphism $f : S^1 \to S^1$ with $\rho(f) = \rho$.*

Proof. Let $\ell_n = (|n| + 3)^{-2}$, $c_n = 2\left(\frac{\ell_{n+1}}{\ell_{n+}} - 1\right) \geq -1$. Then

$$\sum_{n\in\mathbb{Z}} \ell_n < 2\sum_{n=0}^{\infty} \ell_n = 2\sum_{n=3}^{\infty} \frac{1}{n^2} < 1.$$

To blow up the orbit $x_n = R_\rho^n x$, we do the following. We replace x_n by the interval I_n with length ℓ_n, such that they are ordered in the same way as x_n and the space between any I_m and I_n is

$$(1 - \sum_{n \in \mathbb{Z}} \ell_n) d(x_m, x_n) + \sum_{x_k \in (x_m, x_n)} \ell_k.$$

We want to define a map $f(I_n) = I_{n+1}$, and $f|_{S^1 \setminus \bigcup I_n}$ is semiconjugate to R_ρ. On $[a, a+\ell]$, we set

$$h(a, \ell, x) = 1 - \frac{1}{\ell}|2(x-a) - \ell|.$$

Then $h(a, \ell, a + \frac{\ell}{2}) = 1$ and $\int_a^{a+\ell} h(a, \ell, x) dx = \frac{\ell}{2}$. Define

$$f'(x) = \begin{cases} 1 & \text{for } x \in S^1 \setminus \bigcup_{n \in \mathbb{Z}} I_n \\ 1 + c_n h(a_n, \ell_n, x) & \text{if } x \in I_n \end{cases}.$$

We note that

$$\int_{I_n} f'(x) dx = \ell_n + \frac{\ell_n}{2} c_n = \ell_{n+1},$$

so indeed $f(I_n) = I_{n+1}$. $\square$

3 Hyperbolic Dynamics

3.1 Expanding Maps

Definition 3.1 (Expanding Map). *A continuously differentiable map $f : S^1 \to S^1$ is said to be an **expanding map** if $|f'(x)| > 1$ for all $x \in S^1$.*

Lemma 3.2. *If $f, g : S^1 \to S^1$ continuous. Then*

$$\deg(g \circ f) = \deg(f)\deg(g).$$

In particular, $\deg(f^n) = \deg^n(f)$.

Proof. Let F and G be lifts of f and g respectively. Then for any $k \in \mathbb{Z}$, $G(s+k) = G(s) + k\deg(g)$. Then $G \circ F(s+1) = G(F(s) + \deg(f)) = G(F(s)) + \deg(f)\deg(g)$. □

Proposition 3.3. *If $f : S^1 \to S^1$ is an expanding map, then $|\deg(f)| > 1$ and $P_n(f) = \mathrm{Fix}(f^n) = |\deg^n(f) - 1|$.*

Proof. $|\deg(f)| = |F(s+1)-F(s)| > |F'(t)|(s+1-s) > 1$, where $t \in (s, s+1)$. Next we show $\mathrm{Fix}(f^n) = |\deg^n(f)-1|$. We know $\deg(f^n) = \deg^n(f)$ and $|\deg(f^n)| > 1$. Let $g = f^n$, then we are reduced to the case where $|\deg(g)| > 1$, we want to show $\mathrm{Fix}(g) = |\deg(g) - 1|$.

Note that x is a fixed point of g only if $G(x) - x \in \mathbb{Z}$. Let $h(x) = G(x) - x$, then $h(1) = h(0) + \deg(f) - 1$. This shows G has at least $\deg(g) - 1$ fixed points by the intermediate value theorem, since $[h(0), h(1))$ contains at least $\deg(g) - 1$ integers. Since $h'(x) = G'(x) - 1$, then we note that $h'(x)$ is always greater than 0 or less than 0. This shows h is strictly monotone, hence we cannot have more than $\deg(g) - 1$ fixed point. □

Remark 3.3.1. *From the proof, we can see that for any continuous map f, $P_n(f) \geq |\deg(f)^n - 1|$.*

3.2 Hyperbolic Toral Translation

Notation: Let

$$L := \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Then

$$L(x, y) = (2x + y, x + y) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

We know that (x, y) and $(x', y') \in \mathbb{R}^2$ represent the same point on $\mathbb{T}^2$ if and only if $(x - x', y - y') \in \mathbb{Z}^2$. Then $L(x, y)$ represent the same point $L(x', y')$. Since L is invertible over $\mathbb{Z}$, then the map

$$F_L(x, y) = L(x, y) = (2x + y, x + y) \mod 1$$

is an automorphism, in particular, the inverse is given by

$$L^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}.$$

Moreover, we note that the eigenvalues of L is the follows:

$$\lambda_1 = \frac{3+\sqrt{5}}{2} > 1, \quad \lambda_2 = \frac{3-\sqrt{5}}{2} < 1.$$

Proposition 3.4. *Periodic points of F_L are dense and*

$$P_n(F_L) = \lambda_1^n + \lambda_1^{-n} - 2 = \lambda_1^n + \lambda_2^n - 2.$$

Proof. We show that all rational points are periodic points. Let take $(x, y) = \left(\frac{s}{q}, \frac{t}{q}\right)$, then

$$F_L\left(\frac{s}{q}, \frac{t}{q}\right) = \left(\frac{2s+t}{q}, \frac{s+t}{q}\right) \mod 1.$$

Since F_L is an automorphism and the orbit is finite, we conclude that $\left(\frac{s}{q}, \frac{t}{q}\right)$ is a periodic point.

Next we show that if (x, y) is a periodic point, then x, y is a rational point. Notice that $F_L^n(x, y) = (ax+by, cx+dy) \mod 1$ for some integer a, b, c, d such that $ad - bc = 1$, $a + d > 2$. Suppose (x, y) is a periodic point, then

$$ax + by = x + k \quad cx + dy = y + \ell$$

for some $k, \ell \in \mathbb{Z}$. Then by directly solving for x, y, we see that

$$x = \frac{(d-1)k - b\ell}{(a-1)(d-1) - bc} \quad y = \frac{(a-1)\ell - ck}{(a-1)(d-1) - bc}.$$

So $(x, y) \in \mathbb{Q}^2$.

Lastly, we compute $P_n(F_L) = Fix(F_L^n)$. Note that map $G = F_L^n - Id$ is of the form $(x, y) \mapsto ((a-1)x + by, cx + (d-1)y) \mod 1$. So we want to find the cardinality of preimage of 0 under G, which is

$$\text{card}\{(L^n - Id)([0,1) \times [0,1)) \cap \mathbb{Z}^2\}.$$

We claim that

$$|\det(L^n - Id)| = \lambda_1^n + \lambda_1^{-n} - 2,$$

in particular, $L^n - Id$ is invertible. This is true since L can be diagonalized into

$$P \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} P^{-1}.$$

Then we get

$$L^n - Id = P \begin{pmatrix} \lambda_1^n - 1 & 0 \\ 0 & \lambda_2^n - 1 \end{pmatrix} P^{-1}$$

which has determinant $\lambda_1^n + \lambda_1^{-n} - 2$. Then the result follows directly from the following lemma. $\square$

Lemma 3.5. *The area of a parallelogram with integer vertices is the number of lattice points it contains, where points on the edges are counted as half, and all vertices counted as a single point.*

Proof. Denote the area of the parallelogram as A, and the number of lattice points it contains as N. Then we get a canonical tiling of $\mathbb{R}^2$. Denote L as the longest diagonal of the parallelogram. The area of the tiles can be determined by determining how many tiles lies inside the square $[0,n) \times [0,n)$ and covers $[\ell, n-\ell) \times [\ell, n-\ell)$, $n >> \ell$. Note the number of tiles is at least

$$\frac{(n-2\ell)^2}{A} \geq \frac{n^2}{A}\left(1 - \frac{4\ell}{n}\right)$$

since the area of $[\ell, n-\ell) \times [\ell, n-\ell)$ is $(n-2\ell)^2$ and they are all covered by parallelograms.

Since any tile that intersects the square $[0,n) \times [0,n)$ is contained in $[-ell, n+ell) \times [-\ell, n+\ell)$, so the maximum number of tiles is

$$\frac{(n+2\ell)^2}{A} = \frac{n^2}{A}\left(1 + \frac{4\ell}{n}\left(1 + \frac{\ell}{n}\right)\right) < \frac{n^2}{A}\left(1 + \frac{6\ell}{n}\right).$$

Now note the number of integers points in $[0,n) \times [0,n)$ is n^2, and each parallelogram contains the same number of integer points. If we use the counting method stated in the lemma, we have the number of lattice point in the parallelogram, denoted by N, satisfies

$$\frac{(n-2\ell)^2}{A} \cdot N \leq n^2 \leq \frac{(n+2\ell)^2}{A} \cdot N.$$

Letting $n \to \infty$, we get $N = A$. □

Definition 3.6 (Inverse Limit)**.** *If X is a metric space and $f : X \to X$ is continuous and surjective. Then the **inverse limit** is defined on the space*

$$X' = \{(x_n)_{n\in\mathbb{Z}} : x_n \in X \text{ and } f(x_n) = x_{n+1} \, \forall n \in \mathbb{Z}\}.$$

The corresponding shift map $F((x_n)_{n\in\mathbb{Z}}) = (x_{n+1})_{n\in\mathbb{Z}}$

Example: (Solenoid). Let $f = E_2$, $x = S^1$, then the inverse limit is the space

$$S = \{(x_n)_{n\in\mathbb{Z}} : E_2(x_n) = x_{n+1} \, \forall n \in \mathbb{Z}\}.$$

This space is called the solenoid.

3.3 Topological Mixing

Definition 3.7 (Topological Mixing)**.** *A continuous map $f : X \to X$ is said to be **topological maxing** if for every two non-empty open sets U and V, there exists $n \in \mathbb{N}$ such that if $n \geq N$, then $f^n(U) \cap V \neq \emptyset$.*

Proposition 3.8. *Isometries are not topological mixing.*

Proof. WLOG assume $|X|$ is infinite (otherwise it is a discrete space, and we can easily find a counter example, expect $X =$ or X is singleton). Let $f : X \to X$ be an isometry. If $x, y, z \in X$ are three distinct points. Let $\delta = \frac{1}{4}\min(d(x,y), d(x,z), d(y,z))$. Let U, V_1, V_2 be δ-balls around x, y, z respectively. Since f preserves the metric, then $f^n(U)$ is a ball with radius δ. If $f^n(U) \cap V_1 \neq \emptyset$, then $f^n(U) \cap V_2 = \emptyset$; if $f^n(U) \cap V_2 \neq \emptyset$, then $f^n(U) \cap V_1 = \emptyset$. □

Proposition 3.9. *Expanding maps of S^1 are topologically mixing.*

Proof. Let $f : S^1 \to S^1$ such that $|f'(x)| \geq \lambda > 1$ for all $x \in X$. Consider a lift F of f. If $[a, b] \subset \mathbb{R}$ is an interval, then by mean value theorem,

$$|F(b) - F(a)| = |F'(c)| \cdot |b - a| \geq \lambda |b - a|.$$

So the length of any interval is increased by a factor at least λ^n under F^n. Then for every interval I, there exists n such that $|F^n(I)| \geq 1$, so $f^n(I) = S^1$. $\square$

Definition 3.10 (Minimality of Flow)**.** *A **flow** on a set X is a group acton of the additive group of real numbers on X. More explictly, a flow is a family of mapping $\{T_t : X \to X\}_{t\in\mathbb{R}}$ such that for all $x \in X$ and real numbers s and t,*

$$T_0(x) = x$$
$$T_s(T_t(x)) = T_{s+t}(x).$$

*Given a flow T_t on a second countable separable complete metric space space (X, d). We say T_t is **minimal** if every orbit of the flow is dense in X. I.e., $\mathrm{Orb}(x) = \{T_t(x) : t \in \mathbb{R}\}$ is dense in X for every $x \in X$.*

Lemma 3.11. *If a flow is minimal, then so are its time τ map, T_τ, for all but countably many $\tau \in \mathbb{R}$.*

Lemma 3.12. *Let $T_t^\omega : \mathbb{T}^n \to \mathbb{T}^n$ be defined as*

$$T_t^\omega(x_1, \cdots, x_n) = (x_1 + t\omega_1, x_2 + t\omega_2, \cdots, x_n + t\omega_n) \mod 1.$$

Then the flow T_t^ω is minimal if and only if the numbers $\omega_1, \cdots, \omega_n$ are rationally independent.

Proof. Suppose $\omega_1, \cdots, \omega_n$ are rationally independent. We just need to find one $t_0 \in \mathbb{R}$ such that $t_0\omega_1, \cdots, t_0\omega_n, 1$ are rationally independent. Then we know the set $\{T_{t_0}^n(x) : n \in \mathbb{Z}\}$ is dense in $\mathbb{T}^n$ by Proposition (2.40), soT_t^ω is minimal. In fact, we can show that there exists uncountably many t_0 such that the map $T_{t_0}^\omega$ is minimal. Since the equations

$$\sum_{i=1}^n t_0 k_i \omega_i \notin \mathbb{Z} \quad \text{unless } k_1 = k_2 = \cdots = k_n = 0,\ k_i \in \mathbb{Q}$$

has uncountably many solutions in t_0. We can see this in the following way, if $k_i \neq 0$ for all i, then $\sum_{i=1}^n k_i\omega_i \neq 0$, so $\sum_{i=1}^n t k_i \omega_i \in \mathbb{Z}$ if and only if

$$t = \frac{z}{\sum_{i=1}^n k_i \omega_i},\ z \in \mathbb{Z}.$$

Since $k_i \in \mathbb{Q}$ and $z \in \mathbb{Z}$, there are only countable many such t.

Conversely, suppose $\sum_{i=1}^n k_i\omega_i = 0$ for some $k_1, \cdots, k_n \in \mathbb{Z}$, such that $k_1^2 + \cdots + k_n^2 \neq 0$. Then

$$\varphi(x) = \sin\left(2\pi \sum_{i=1}^n k_i x_i\right)$$

is a T_t^ω invariant set. Then we note $\phi^{-1}([0,1]))$ is a closed invariant set of T_t^ω. We note $\phi^{-1}([0,1])$ is not empty nor the entire space $\mathbb{T}^n$. This shows T_t^ω is not minimal. $\square$

Proposition 3.13. *The automorphism F_L is a topological mixing.*

Proof. Let $U, V \subset \mathbb{T}^2$ be open sets. We consider the linear flow T_t^ω, $\omega = \left(1, \frac{\sqrt{5}-1}{2}\right)$. Then T_t^ω is minimal by Lemma (3.12). Hence the projection of the line ℓ given by

$$y = \frac{\sqrt{5}-1}{2}x$$

is dense on the torus $\mathbb{T}^2$. So U must contains a piece of this line ℓ. Next notice $\left(1, \frac{\sqrt{5}-1}{2}\right)$ is an eigenvalue of L with corresponding eigenvalue $\lambda = 3 + \sqrt{5}$, hence the ℓ is F_L-invariant and F_L act on the line as an expanding map. Furthermore, for any $\epsilon > 0$, there exists $T = T(\epsilon)$ such that the projection of any line segment on the line $y = \frac{\sqrt{5}-1}{2}x + c$ of length at least T intersects any ϵ ball on the torus. By translation we just need to show it is possible to find $T(\epsilon)$ for the line ℓ. However, since the projection of ℓ is dense on $\mathbb{T}^2$ and by compactness of $\mathbb{T}^2$, we can find a finite cover of $\mathbb{T}^2$ using ϵ-balls, which allows us to find $T(\epsilon)$ such that the line segment starting from 0 with length $T(\epsilon)$ satisfies the given property. However, the projection of any other line segment on the line is just a translation, hence its holds or any line segment of ℓ with length at least T. Lastly, let ϵ be such that V contains an ϵ ball. Then take N such that for any $n \geq N$, the expanding line segment contained in U has length more than $T(\epsilon)$ under the map F_L^n, then by previous analysis, we have $f^n(U) \cap V \neq \emptyset$. $\square$

4 Symbolic Dynamics

4.1 Coding

We start by considering the following example. Let $E_2(x) = 2x \mod 1$ be the expanding map. We can encode this map by the binary expansion of x. Let

$$\Delta_n^k = \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right]$$

for $n = 1, 2, \cdots$, and $k = 1, \cdots, 2^{n-1}$. Then observe that if $x, y \in \Delta_n^k$, then they share the first n digits in binary expansion. We can use this to prove E_2 is topological transitive.

Suppose $x = 0.x_1x_2 \cdots x_n \cdots$ in binary expansion, then $E_2(x) = x_1.x_2x_3 \cdots \cdots x_n \cdots \mod 1 = 0.x_2x_3 \cdots$. Let $k \in \mathbb{N}$, and its binary expansion as $k_0 \cdots k_{n-1}$, then $x \in \Delta_n^k$ if and only if $x_i = k_i$, $i = 0, \cdots, n-1$. Then let w_n^k be the length n binary sequence such that $0.w_n^k$ is in Δ_n^k. Then consider w_n being the concatenation of $w_n^0 \cdots w_n^{n-1}$. Then there exists $m_k \in \mathbb{N}$ such that $0.w$ lies in the interval Δ_n^k. Lastly, consider

$$w = 0.w_1w_2 \cdots$$

which is the concatenation of $w_1, \cdots, w_n, \cdots$. Then it is clear that the positive iterates of w is dense in $[0, 1)$. Note the similar proof can be used for E_m by considering expansion in base m.

Proposition 4.1. *There exists a point $x \in S^1$ such that the closure of its orbit with respect to E_3 in additive notation coincides with the standard middle three cantor set C. In particular C is invariant under E_3 and contains a dense orbit.*

Proof. Clearly C is E_3 invariant, so suffices to show there exists $x \in C$ such that the positive orbit of x under E_3 is dense in C.

Every point in C has a unique representation in base 3 without 1's. Let $x \in C$ and

$$0.x_1x_2x_3 \cdots$$

be such a representation. Let $h(x)$ be the numbers whose representation in base 2 is

$$0.\frac{x_1}{2}\frac{x_2}{2}\frac{x_3}{2} \cdots$$

Thus we have constructed a map $h : C \to [0, 1]$ that is continuous, nondecreasing and one-to-one, except for the fact that binary rationals have two preimages. Furthermore, $h \circ E_3 = E_2 \circ h$. Let $D \subset [0, 1]$ be a dense set of points that does not contain binary rationals. Then $h^{-1}(D)$ is dense in K because, if Δ is an open interval such that $\Delta \cap K \neq \emptyset$, then $h(\Delta)$ is a nonempty interval, hence contains points of D. Now take any $x \in [0, 1]$ whose E_2-orbit is dense; the E_3-orbit of $h^{-1}(x) \in K$ is dense in K. $\square$

Definition 4.2 (Symbolic Space)**.** *We define*

$$\Omega_N := \{\omega = (\cdots, \omega_{-1}, \omega_0, \omega_1, \cdots\} : \omega_i \in \{0, 1, \cdots, N-1\}\}$$

*which is called the **bi-infinite symbolic space**, similarly, we define*

$$\Omega_N^R := \{\omega = (\omega_0, \omega_1, \omega_2, \cdots\} : \omega_i \in \{0, 1, \cdots, N-1\}\}$$

Remark 4.2.1. *We can define the topology on the symbolic space in several different ways. For example, on Ω_n, we can see it as the direct product of $\mathbb{Z}$ copies of the finite set $\{0, 1, \cdots, N-1\}$ each equipped with discrete topology. Then we can just equipped Ω_N with the product topology.*

*Alternatively, for integers $n_1 < n_2 < \cdots < n_k$ and $\alpha_1, \cdots, \alpha_n \in \{0, 1, \cdots, N-1\}$, then the **cylinder set with rank** k is givenby*

$$C_{\alpha_1,\cdots,\alpha_n}^{n_1,\cdots,n_k} := \{\omega \in \Omega_N : \omega_{n_i} = \alpha_i,\, i = 1, \cdots, k\}$$

Then the set of cylinder sets forms a basis of the product topology. Note that cylinder sets are also closed.

Moreover, we know the product topology with countable sets is metrizable. In fact, for the case of Ω_n, given any $\lambda > 1$,

$$d_\lambda(\omega, \omega') = \sum_{n=-\infty}^{\infty} \frac{|\omega_n - \omega_n'|}{\lambda^{|n|}}$$

induces the product topology on Ω_n.

Remark 4.2.2. *With respect to the product topology, Ω_N is a perfect, disconnected, compact, Hausdorff space. Hence by Brouwer's theorem, it is homeomorphic to the Cantor set.*

Definition 4.3 (Exponential Type)**.** *Let ϕ be a continuous complex valued function defined on Ω_N or any of its closed subsets. We write*

$$\omega = (\cdots, \omega_{-1}, \omega_0, \omega_1, \cdots)$$
$$\omega' = (\cdots, \omega_{-1}', \omega_0', \omega_1', \cdots).$$

We denote

$$V_n(\phi) := \max\{|\phi(\omega) - \phi(\omega')| : \omega_k = \omega_k' \text{ for } |k| \leq n\}.$$

*Since Ω_N is compact, then ϕ is uniformly continuous. Therefore $|V_n(\phi)| \to 0$ as $n \to \infty$. Then we say ϕ has **exponential type** if for some $\alpha > 0$ and $c > 0$,*

$$V_n(\phi) \leq ce^{-an}.$$

Definition 4.4 (Left Shift Operator)**.** *The **left shift operator** or more formally the **topological Bernoulli shift** on Ω_N, $\sigma_N : \Omega_N \to \Omega_N$ given by*

$$\sigma_N(\omega) = \omega' = (\cdots, \omega_{-1}', \omega_0', \omega_1', \cdots)$$

*such that $\omega_n' = \omega_{n+1}$, $n \in \mathbb{Z}$. The **left shift operator** on Ω_N^R, σ_N^R is given by*

$$\sigma_N^R(\omega_0, \cdots, \omega_n, \cdots) = (\omega_1, \cdots, \omega_n, \cdots).$$

Remark 4.4.1. *We note that σ_N is clearly bijective. It is also bicontinuous, since σ_N and its inverse clearly preserves Cylinders.*

Proposition 4.5. *Periodic points for the shift σ_N and σ_N^R is dense in Ω_N and Ω_N^R respectively. Both transformations σ_N and σ_N^R are topological mixing and*

$$P_n(\sigma_N) = P_n(\sigma_N^R) = N^n.$$

Proof. We only show for the case of σ_N. We first note that

$$\sigma_N^m \omega = \omega$$

if and only if $\omega_{n+m} = \omega_n$ for all $n \in \mathbb{Z}$. Thus $P_n(\sigma_N) = N^n$.

Next to show density of periodic point, it is enough to find a periodic point in any cylinder. But then this is clear, since if

$$C_{\alpha_1,\cdots,\alpha_k}^{n_1,\cdots,n_k}$$

is a given cylinder, then ω' such that $\omega_n = \alpha_i$ if $n \equiv \alpha_i \mod (\alpha_k - \alpha_1 + 1)$, and $\omega_n = 0$ otherwise. Then we see that ω is periodic and lies in the given cylinder.

Lastly we show topological mixing. It is enough to show that for any $\alpha = \alpha_{-m}, \cdots, \alpha_m$ and $\beta = \beta_{-m}, \cdots, \beta_m$, and n sufficiently large

$$\sigma_N^n(C_\alpha^m) \cap C_\beta^m \neq \emptyset$$

But then we have the freedom to choose ω_n when $\omega > 2m + 1$. Then by construction, we have

$$\sigma_N^n(C_\alpha^m) \cap C_\beta^m \neq \emptyset$$

□

4.2 Topological Markov Chain

Let $A = (a_{ij})_{i,j=0}^{N-1}$ be a $N \times N$ matrix whose entries a_{ij} are either zeros or ones. Let

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

Definition 4.6 (Topological Markov Chain)**.** *The restriction*

$$\sigma_A = \sigma_N|_{\Omega_A}$$

is called the ***topological Markov chain*** *determined by the matrix A. σ_A is called a subshift of finite type.*

Lemma 4.7. *We identify the symbols $0, 1, \cdots, N_1$ with $x_0, x_1, \cdots, x_{N-1}$ and connected x_i with x_j with an arrow if $a_{ij} = 1$. Then for every $i, j \in \{0, 1, \cdots, N-1\}$, the number N_{ij}^m of admissible paths of lengths m that begins at x_i and end at x_j is equal to the entry $a_{ij}^{(m)}$ of A^m.*

Proof. If $m = 1$, this follows from the definition. Wish to show $N_{ij}^{m+1} = \sum_{i=0}^{N-1} N_{ik}^m a_{kj}$ for every $k \in \{0, \cdots, N-1\}$.

This is true since every admissible path of length m connecting x_i and x_k produces exactly one admissible path with length $m+1$ connecting x_i and x_j. Then by induction, we have the desired result. □

Corollary 4.7.1. $P_n(\sigma_A) = \sum_{i=0}^{N-1} a_{ii}^{(n)} = tr(A^n)$.

Definition 4.8 (Transitive Matrix)**.** *A $0-1$ matrix A is called* ***transitive*** *if for some positive m all entries of the matrix A^m are positive. We call a* ***topological Markov Chain*** *σ_A transitive if A is a transitive matrix.*

Lemma 4.9. *If all entries of A^m are positive, then for any $n \geq m$, all entries of A^n are positive.*

Proof. Assume $a_{ij}^{(m)} > 0$ for all i, j. Then for each j, there is a k such that $a_{kj} = 1$. Otherwise $a_{ij}^{n} = 0$ for every n and i. Then

$$a_{ij}^{(n+1)} = \sum_{k=0}^{N-1} a_{ik}^{(n)} \cdot a_{kj} > 0.$$

□

Lemma 4.10. *If A is transitive and $\alpha = (\alpha_{-k}, \cdots, \alpha_k)$ is admissible, i.e., $a_{\alpha_i \alpha_{i+1}} = 1$ for $i = -k, \cdots, k-1$. Then the intersection*

$$C_{\alpha,A}^k = \Omega_A \cap C_\alpha^k$$

is nonempty and contains a periodic point.

Proof. Since A is transitive, then there exists $m > 0$ such that $a_{\alpha_i \alpha_{i+1}}^{(m)} > 0$ for all i, j. Then we can extend α to an admissible sequence with length $2k+1+m$ such that begins and ends with $\alpha - k$. □

Proposition 4.11. *If A is a transitive matrix, then the topological Markov chain is topological mixing and its periodic orbits are dense in Ω_A.*

Proof. By Lemma (4.10), the periodic points are dense in Ω_A. So wish to show that for any two sequences $\alpha = (\alpha_{-k}, \cdots, \alpha_k)$ and $\beta = (\beta_{-k}, \cdots, \beta_k)$, then for any sufficiently large n, the intersection of two nonempty cylinders

$$\sigma_A^n(C_{\alpha,A}^k) \cap C_{\beta,A}^k \neq \emptyset.$$

Let $n = 2k+1+m+\ell$, $\ell \geq 0$, m being from the transitive definition of A. By Lemma (4.9),

$$a_{\alpha_k \beta_{-k}}^{(m+\ell)} > 0$$

This tells us that we can construct a sequence with length $4k+2+m+\ell$ whose first $2k+1$ symbols are α and last $2k+1$ symbols are β. By Lemma (4.10), this word can be extended bi-infinitely. □

Theorem 4.12 (Coding of the Toral Automorphism)**.** *For the map*

$$F(x,y) = (2x+y, x+y) \mod 1$$

of the torus, there is a semi-conjugacy $h : \Omega_A \to \mathbb{T}^2$ such that

$$F \circ h = h \circ \sigma_5|_{\Omega_A}$$

where

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

Proof. Let ω be a bi-infinite sequence in Ω_A. If $\omega = (\omega_n)_{n\in\mathbb{Z}}$, then there exists at most one point in the intersection

$$\bigcap_{n\in\mathbb{Z}} F^{-n}(\Delta_{\omega_n}).$$

Consider $\tilde{\omega} = \omega_{-\ell}\omega_{-\ell+1}\cdots\omega_k$, $\ell > 0$, $k > 0$. Since

$$\bigcap_{n=-\ell}^{k} F^{-n}(\Delta_{\omega_n}) = \left(\bigcap_{n=-\ell}^{0} F^{-n}(\Delta_{\omega_n})\right) \cap \left(\bigcap_{n=1}^{k}(F^{-n}(\Delta_{\omega_n})\right).$$

Let $k, \ell \to \infty$, we see this intersection has at most one point.

We also see that if $\omega \in \Omega_5$ and $F^{-1}(\Delta_{\omega_n}) \cap \Delta_{\omega_{n+1}} \neq \emptyset$ for all $n \in \mathbb{Z}$ (i.e., $\omega \in \Omega_A$), then $\bigcap_{n\in\mathbb{Z}} F^{-n}(\Delta_{\omega_n})$ is non-empty. Then we define $h(\omega) = \bigcap_{n\in\mathbb{Z}} F^{-n}(\Delta_{\omega_n})$ for $\omega \in \Omega_A$. It is clear that h is the factor map between F and $\sigma_5|_{\Omega_A}$. □

Definition 4.13 (Markov Partition)**.** *A* ***Markov partition*** *is a finite cover $R = \{R_0, R_1, \cdots, R_{n-1}\}$ of $\mathbb{T}^2$ by proper rectangles such that the following are true:*

1. $\mathrm{Int}(R_i) \cap \mathrm{Int}(R_j) = \emptyset$ *if* $i \neq j$.
2. *Whenever* $x \in \mathrm{Int}(R_i)$ *and* $f(x) \in \mathrm{Int}(R_i)$, *then*

$$W^u_{R_j}(f(x)) \subset F(W^u_{R_i}(x))$$

and

$$f(W^s_{R_i}(x)) \subset W^s_{R_j}(f(x)).$$

Proposition 4.14. *The semi-conjugacy h between σ_A and F is one to one on periodic points except for the fixed point. The number of preimages of any point not negatively asymptotic to the fixed point is bounded.*

Proof. The fixed points for σ_A are the constant sequences 0's $1'$ and 4's which all maps to $(0,0)$ under h. By Jordan Normal Form, we have

$$P_n(\sigma_A) = \lambda_1^n + \lambda_1^{-n}$$

where λ_1 is the largest eigenvalue of

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

Note

$$P_n(\sigma_A) = P_n(F) + 2.$$

The fixed point for F is $(0,0)$. So except for the fixed point, all the other periodic points are identified with each other.

If $q \in \mathbb{T}^2$ and its positive and negative iterations avoid boundaries $\partial R^{(1)}$ and $\partial R^{(2)}$. Then by Markov partition property, it has a unique preimage under h, and vice versa. In particular, periodic points other than the origin fall into this category as they have rational coordinates. The points of Ω_A on those boundaries or their iterates under F fall into three categories corresponding to the three segments of stable and unstable manifolds through 0 that define parts of the boundary. Thus sequences are identified in the following case: they have a constant infinite right (future) tail consisting of 0's or 4's and agree otherwise- this correspods to a stale boundary piece; or else an infinite left (past) tail of (0's and 1's, or of 4's) and agree otherwise-this corresponds to an unstable boundary piece. $\square$

5 Entropy

5.1 Capacity

Definition 5.1 (Capacity)**.** *Suppose X is a compact space with metric d.*

- *A set $E \subset X$ is said to be r-**dense** if*
$$X \subset \bigcup_{x \in E} B_d(x, r).$$
- *We define the r-**capacity** of (X, d) to be the minimal cardinality $S_d(r)$ of an r-dense set.*

Example: Let $X = [0, 1]$ with the usual metric, i.e., $d(x, y) = |x - y|$. Then $S_d(r) = \lfloor \frac{1}{2r} \rfloor + 2$.

Example: Let X be the Cantor set with usual metric. To compute the order of $S_d(\frac{1}{3^i})$, We first consider $S_d((3 - \frac{1}{i})^{-i})$ which is just 2^i. Then $S_d(\frac{1}{3^i})$ is of order 2^i.

Definition 5.2 (Box Dimension)**.** *If X is a totally bounded metric space, then the quantity*

$$\mathrm{bdim}(x) = \lim_{r \to 0} -\frac{S_d(r)}{\log r}$$

is called the box dimension of X.

Example: The box dimension of $X = [0, 1]$ is

$$\lim_{r \to 0} -\frac{\log S_d(r)}{\log r} = \frac{\log 2r}{\log r} = 1.$$

Example: The box dimension of the Cantor set is $\frac{\log 2}{\log 3}$.

Example: We consider the space (Ω_N, d_λ), $\lambda > 1$. Let $\alpha = \alpha_{1-n} \cdots \alpha_{n-1}$, then $B_{d_\lambda}(\alpha, \lambda^{1-n}) = C_{\alpha_{1-n} \cdots \alpha_{n-1}}$. Hence $S_{d_\lambda}(r) = N^{2n-1}$. So

$$\mathrm{bdim}(\Omega_N, d_\lambda) = \lim_{r \to 0} -\frac{\log S_d(r)}{\log r} = \lim_{n \to \infty} -\frac{\log N^{2n-1}}{\log \lambda^{1-n}} = 2 \cdot \frac{\log N}{\log \lambda}.$$

Proposition 5.3. *Given a totally bounded metric space (X, d). Then for any $a > 0$,*

$$\mathrm{bdim}(X, d) = \mathrm{bdim}(X, ad).$$

Proof. It is clear that $S_{ad}(ar) = S_d(r)$ Then

$$\mathrm{bdim}(X, ad) = \lim_{r \to 0} -\frac{\log S_{ad}(r)}{\log r} = \lim_{r \to 0} -\frac{\log S_{ad}(ar)}{\log ar} = \lim_{r \to 0} -\frac{\log S_d(r)}{\log a + \log r} = \lim_{r \to 0} -\frac{\log S_d(r)}{\log r} = \mathrm{bdim}(X, d).$$

□

5.2 Topological Entropy

Definition 5.4 (Bowen Metric)**.** *Suppose $f : X \to X$ is a continuous map of a compact metric space X with distance function d. We define the **Bowen metric** d_n^f as*

$$d_n^f(x,y) = \max_{0 \leq i \leq n-1} d(f^i(x), f^i(y)).$$

Definition 5.5 (Topological Entropy)**.** *A set $E \subset X$ is r-dense with respect to d_n^f or (n,r)-dense if $X \subset \bigcup_{x \in E} B_f(x,r,n)$, where*

$$B_f(x,r,n) = \{y \in X \,:\, d_n^f(x,y) < r\}.$$

$S_d(f,r,n)$ is the minimal cardinality of an (n,r)-dense set. The exponential growth rate of $S_d(f,r,n)$ is defined as

$$h_d(f,r) = \limsup_{n \to \infty} \frac{1}{n} \log S_d(f,r,n).$$

*We define the **topological entropy** $h_d(f)$ of f to be*

$$h_d(f) = \lim_{r \to 0} h_d(f,r).$$

Proposition 5.6. *Let $f : [0,1] \to [0,1]$ be a homeomorphism, then $h_{\text{top}}(f) = 0$.*

Proof. WLOG, let $f(0) = 0$ and $f(1) = 1$ and so that f is monotonically increasing. Fix $n \in \mathbb{Z}_{\geq 1}$ and $r \in \mathbb{R}_{>0}$.

Consider the set $S' = \left\{\frac{1}{p}, \frac{2}{p}, \cdots, \frac{p-1}{p}\right\}$ with $p \in \mathbb{Z}_{>1}$ be the smallest such that $\frac{1}{p} < r$, and define $S := \bigcup_{k=0}^{n-1} f^{-k}(S')$.

We claim that S is an (n,r) - spanning set for f. Indeed, let $x \in [0,1]$. Then there are at most two elements $s_1, s_2 \in S$ that are the closest in S to x, i.e. $d(x,S) = d(x,s_1) = d(x,s_2)$. Let us choose one of them and write instead $s_x \in S$. S being an (n,r) - spanning set for f is equivalent to $d_n^f(x,s_x) < r$. Suppose otherwise. Then there is some $k \in \{0,1,\ldots,n-1\}$ such that $d(f^k(x), f^k(s_x)) \geq r$ (in particular $x \neq s_x$; WLOG let's assume $x \leq s_x$). Then for some $s'_x \in S'$, we have $f^k(x) \leq s'_x \leq f^k(s_x)$. Since f^k is a homeomorphism, its restriction is also a homeomorphism $[x, s_x] \to [f^k(x), f^k(s_x)]$; consequently we have $x \leq f^{-k}(s'_x) \leq s_x$, a contradiction since $f^{-k}(s'_x) \in S$ but S has no points closer to x than s_x. Thus S is an (n,r) - spanning set for f.

Therefore we have

$$S_d(f,n,r) \leq \#(S) = n(p-1) \leq \frac{n}{r},$$

Then by direct computation we find out that $h_{\text{top}}(f) = 0$. □

Proposition 5.7. *If d' is a metric on X equivalent to d, then*

$$h_{d'}(f) = h_d(f).$$

Proof. We consider the identity map

$$\text{id} : (X,d) \to (X,d')$$

which is an homeomorphism, since d, d' are equivalent. We also know that id is uniformly continuous.

Given $r > 0$, there exists $\delta = \delta(r) > 0$ such that if

$$d'(x_1, x_2) < \delta$$

then $d(x_1, x_2) < r$. So $S_{d'}(f, \delta, n) \geq S_d(f, r, n)$. Then by the definition of topological entropy, we have

$$h_{d'}(f) \geq h_d(f).$$

Similarly, we can show that

$$h_d(f) \geq h_{d'}(f).$$

Thus completing the proof. □

Proposition 5.8. *Suppose $h : X \to Y$ is an isometry, and $f : X \to X$. Then*

$$h_{top}(h \circ f \circ h^{-1}) = h_{top}(f).$$

Proof. Let the metric on X and Y be d and d' respectively and let $g = h \circ f \circ h^{-1}$. We show that if $\mathcal{S}$ is a (n, r)-dense set of X under d_f^n, then $h(\mathcal{S})$ is an (n, r)-dense set of Y under $(d')_g^n$. This is the case since if $y \in Y$, then let $x = h^{-1}(y) \in X$. By definition, there exists $\tilde{x} \in \mathcal{S}$ such that

$$d_f^n(\tilde{x}, x) < r.$$

Then

$$\begin{aligned}
d_g^n(h(\tilde{x}), y) &= \max_{0 \leq n \leq n-1} d'(g^n(h(\tilde{x}), g^n(h(x))) \\
&= \max_{0 \leq n \leq n-1} d'(h \circ f^n(\tilde{x}), h \circ f^n(x)) \\
&= d(f^n(\tilde{x}), f^n(x)) \\
&< r.
\end{aligned}$$

Conversely, we can show that if $\mathcal{S}$ is a (n, r)-dense set of Y under $(d')_g^n$, then $h^{-1}(\mathcal{S})$ is an (n, r)-dense set of X under d_f^n. Hence we conclude that

$$h_{top}(g) = h_{top}(f).$$

□

Corollary 5.8.1. *Topological Entropy is an invariant of topological conjugacy.*

Proof. Let $f : X \to X$ and $g : Y \to Y$ be conjugate via a homeomorphism $h : X \to Y$, i.e.,

$$h \circ f = g \circ h.$$

Let d be a metric on X. Then define the metric d' on Y by

$$d'(y_1, y_2) = d(h^{-1}(y_1), h^{-1}(y_2))$$

then h is an isometry from (X, d) to (Y, d'), hence by Proposition (5.8), $h_d(f) = h_{d'}(g)$. It remains to show that

d' is equivalent to the original metric ρ of Y, but this is the case since the identity map $(Y, \rho) \to (Y, d')$ is the composition of h_1 and h_2^{-1}, where $h_1 = h : (X, d) \to (Y, d')$ and $h_2^{-1} = h^{-1} : (Y, \rho') \to (X, d)$, both of which are homeomorphism. Hence by Proposition (5.7), we conclude that proof. $\square$

Proposition 5.9. *The topological entropy of contractions and isometries are zero.*

Proof. Since f is a contraction or isometry, then

$$d_n^f(x, y) = d(x, y).$$

So $S_d(f, r, n)$ is independent with n, then $h_d(f, r) = 0$. $\square$

5.3 Topological Entropy Via Covers and Separated Sets

Definition 5.10. *Let $D_d(f, r, n)$ be the minimal number of sets whose diameter in the metric d_n^f is less than r and whose union covers X.*

Lemma 5.11. $\tilde{h}(f, r) = \lim_{n\to\infty} \frac{1}{n} \log D_d(f, r, n)$ *exists for any $r > 0$.*

Proof. If A is a set of d_n^f-diameter less than r, and B is a set of d_m^f-diameter less than r, then $A \cap f^{-n}(B)$ has d_{n+m}^f-diameter less than r. This is because, if x, y are two points in the intersection, then by $x, y \in A$, we know $d(f^i(x), f^i(y)) < r$ for $i = 0, \cdots, n-1$ and by $x, y \in f^{-n}(B)$, we know $d(f^i(x), f^i(y)) < r$ for $i = n, \cdots, n+m-1$.

If $\mathcal{U}$ is a cover of X by $D_d(f, r, n)$ sets of d_n^f-diameter less than r. $\mathcal{B}$ is a cover of X by $D_d(f, r, m)$ sets of d_m^f-diameter of less than r. Then the cover $A \cap f^{-n}(B)$, $A \in \mathcal{U}$, $B \in \mathcal{B}$ contains at most $D_d(f, r, n) \cdot D_d(f, r, m)$ sets with d_{n+m}^f-diameter less than r. Moreover, since $\mathcal{B}$ covers X, then $\bigcup f^{-n}(B)$ covers X, so $\bigcup A \cap f^{-n}(B)$ covers X. Hence we conclude that

$$D_d(f, r, n+m) \leq D_d(f, r, n) \cdot D_d(f, r, m)$$

which implies

$$\log D_d(f, r, n+m) \leq \log D_d(f, r, n) + \log D_d(f, r, m).$$

Then by the sub-additivity lemma, we concludes the proof of the lemma. $\square$

Proposition 5.12. *If $\underline{h}_d(f, r) = \liminf_{n\to\infty} \frac{1}{n} \log S_d(f, r, n)$, then*

$$\lim_{r\to 0} \tilde{h}_d(f, r) = \lim_{r\to 0} \underline{h}_d(f, r) = \lim_{r\to 0} h_d(f, r) = h_{top}(f).$$

Proof. We note that the diameter of an r-ball is at most $2r$, so every covering by r-ball is a covering of diameter less than $2r$. So

$$D_d(f, 2r, n) \leq S_d(f, r, n).$$

Any set of diameter less than r is contained in the r-ball centered at any points inside the set, so

$$S_d(f, r, n) \leq D_d(f, r, n).$$

So

$$\tilde{h}_d(f, 2r) \le \underline{h}_d(f, r) \le h_d(f, r) \le \tilde{h}_d(f, r).$$

$\square$

Definition 5.13. *Let $N_d(f, r, n)$ be the maximal number of points in X with pairwise d_n^f distance at least r. Then we call a set with cardinality $N_d(f, r, n)$ such that the pairwise d_n^f distance at least r to be an (n, r)-separated sets.*

Proposition 5.14.

$$h_{top}(f) = \lim_{r \to 0} \limsup_{n \to \infty} \frac{1}{n} \log N_d(f, r, n) = \lim_{r \to 0} \liminf_{n \to \infty} \frac{1}{n} \log N_d(f, r, n).$$

Proof. We first observe that a maximal (n, r)-separated sets is (n, r)-dense. Then $S_d(f, r, n) \le N_d(f, r, n)$. On the other hand, we note that each $\frac{r}{2}$-ball cannot contain two points r apart, hence $N_d(f, r, n) \le S_d(f, \frac{r}{2}, n)$. This concludes the proof of the proposition. $\square$

Proposition 5.15.

1. *If Λ is a closed f-invariant set, then*
$$h_{top}(f|_\Lambda) \le h_{top}(f).$$

2. *If $X = \bigcup_{i=1}^{m} \Lambda_i$, where Λ_i are closed f-invariant sets, then*
$$h_{top}(f) = \max_{1 \le i \le m} h_{top}(f|_{\Lambda_i}).$$

3. $h_{top}(f^m) = |m| h_{top}(f)$.

4. *If g is a factor of f, then $h_{top}(g) \le h_{top}(f)$.*

5. $h_{top}(f \times g) = h_{top}(f) + h_{top}(g)$, *where $f : X \to X$, $g : Y \to Y$, $(f \times g)(x, y) = (f(x), g(y))$.*

Proof.

1. Take any (n, r)-cover S of X. Then $S|_\Lambda := \{A \cap \Lambda : A \in S\}$ is an (n, r)-cover of Λ since Λ is closed and f-invariant.

2. Since the union of covers of $\Lambda_1, \cdots, \Lambda_m$ by sets of diameter less than r is a cover of X. Hence
$$D_d(f, r, n) \le \sum_{i=1}^{m} D_d(f|_{\Lambda_i}, r, n).$$
Then there exists at least one i such that for infinitely many $n \in \mathbb{N}$,
$$\frac{1}{m} D_d(f, r, n) \le D_d(f|_{\Lambda_i}, r, n)$$
This shows that for each r, there exists i such that
$$\max_{1 \le i \le m} \limsup_{n \to \infty} \frac{\log D_d(f|_{\Lambda_i}, r, n)}{n} \ge \limsup_{n \to \infty} \frac{(\log D_d(f, r, n)) - \log m}{n},$$

hence

$$h_{top}(f) \le \max_{1\le i\le m} h_{top}(f|_{\Lambda_i}).$$

It is clear that

$$h_{top}(f) \ge \max_{1\le i\le m} h_{top}(f|_{\Lambda_i})$$

by part (1).

3. Let $m > 0$. Then

$$d_n^{f^m}(x,y) = \max_{0\le i\le n-1} d(f^{im}(x), f^{im}(y)) \le \max_{0\le i\le mn-1} d(f^i(x), f^i(y)) = d_{mn}^f(x,y).$$

This shows any $d_n^{f^m}$-r-ball contains a d_{mn}^f-r-ball. Hence $S_d(f^m, r, n) \le S^d(f, r, mn)$. Then

$$h_{top}(f^m) \le m \cdot h_{top}(f).$$

On the other hand, $\forall r > 0$, there exists $\delta(r) > 0$ (independent of n) such that

$$B(x, \delta(r)) \subset B_f(x, r, m) \quad \forall x \in X.$$

Then

$$\begin{aligned} B_{f^m}(x, \delta(r), n) &= \bigcap_{i=0}^{n-1} f^{-im}((B(f^{im(x)}, \delta(r))) \\ &\subset \bigcap_{i=0}^{n-1} f^{-im}(B_f(f^{im}(x), r, m)) \\ &= B_f(x, r, mn) \end{aligned}$$

So

$$S_d(f, r, mn) \le S_d(f^m, \delta(r), n)$$

Hence

$$m \cdot h_{top}(f) \le h_{top}(f^m).$$

Next we consider the case $m < 0$, we assume f is invertible. Then

$$B_f(x, r, n) = B_{f^{-1}}(f^{n-1}(x), r, n)$$

So

$$S_d(f, r, n) = S_d(f^{-1}, r, n)$$

since f is invertible. Thus $h_{top}(f) = h_{top}(f^{-1})$.

4. If g is a factor of f. There exists $h : X \to Y$ such that $h \circ f = g \circ h$ and moreover $h(X) = Y$. Let d_X, d_Y be the corresponding metric on X and Y. Since X, Y are compact, then h is uniformly continuous. Then for

any $r > 0$, there exists $\delta(r) > 0$ such that if $d_X(x_1, x_2) < \delta(r)$, then $d_Y(h(x_1), h(x_2)) < r$. The image of any $(d_X)_n^f$-ball of radius $\delta(r)$ lies inside a $(d_Y)_n^g$-ball of radius r. Hence

$$S_{d_X}(f, \delta(r), n) \geq S_{d_Y}(g, r, n)$$

Hence

$$h_{top}(f) \geq h_{top}(g).$$

5. We consider the product metric given b $d((x_1, y_1), (x_2, y_2)) = \max(d(x_1, x_2), d(y_1, y_2))$ The balls in the product metric are products of balls on X and Y. This is also true on the Bowen metric. The above is also true for $d_n^{f \times g}$. Hence

$$S_d(f \times g, r, n) \leq S_{d_X}(f, r, n) \cdot S_{d_Y}(g, r, n).$$

This shows

$$h_{top}(f \times g) \leq h_{top}(f) + h_{top}(g).$$

The product of any (n, r)-separated set in X and any (n, r)-separated set in Y for g is an (n, r)-separated set for $f \times g$. Then

$$N_d(f \times g, r, n) \geq N_{d_X}(f, r, n) \cdot N_{d_Y}(g, r, n).$$

Hence

$$h_{top}(f \times g) \geq h_{top}(f) + h_{top}(g).$$

□

5.4 Examples

Theorem 5.16. *Let $E_m : S^1 \to S^1$ be the map given by $E_m(x) = mx \mod 1$. If $m \in \mathbb{N}$ with $|m| \geq 2$. Then*

$$h_{top}(E_m) = \log |m|.$$

Proof. For the map E_m, the distance between iterates any two points grows until it becomes greater than a certain constant depending on m (specifically $\frac{1}{2m}$).

Assume $m > 0$. If $d(x, y) < \frac{m^{-n}}{2}$, we have

$$d_n^{E_m}(x, y) = d(E_m^{(n-1)}x, E_m^{(n-1)}y).$$

Also note that $d_n^{E_m}(x, y) \geq r$ if and only if $d(x, y) \geq r \cdot m^{-n}$. Then

$$\{i \cdot m^{-n-k} \; : \; i = 0, 1, \cdots, m^{n+k} - 1\}$$

is a maximal set of points whose pairwise distance $d_n^{E_m}$-distance are at least $r = m^{-k}$ for some $k \in \mathbb{N}$. This shows

$$N_d(E_n, m^{-k}, n) = m^{n+k}$$

so

$$h(E_m) = \lim_{R\to\infty} \limsup_{n\to\infty} \frac{\log N_d(E_m, m^{-k}, n)}{n} = \log m.$$

Similarly, we can do the same for $m < 0$. $\square$

Proposition 5.17. *For any transitive topological Markov chain, we have*

$$h_{top}(\sigma_A) = \log|\lambda_A^{\max}|$$

where $\lambda_A^{\max}$ is the largest eigenvalue of A (in terms of norm).

Proof. Recall

$$C_{\alpha_{-m},\cdots,\alpha_{m+n}}^{-m,\cdots,m+n} = \{\omega \in \Omega_N \,:\, \omega_i = \alpha_i \text{ for } -m \le i \le m+n\}.$$

We observe that this is the ball of radius $r_m = \frac{\lambda^{-m}}{2}$ around each of C_α^m points with respect to the metric $d_n^{\sigma_N}$, here λ is from d_λ , which the the metric associated to σ_N (for simplicity sake, we can assume $\lambda > 3$).

Then we notice that for any two $d_n^{\sigma_N}$ balls of radius r_m are either identical or they are disjoint. Hence

$$S_{d_\lambda}(\sigma_N, r_m, n) = N^{2m+1+n}.$$

This shows $h(\sigma_N) = \log N$. If σ_A is a topological Markov chain, then, then $S_d(\lambda_A, r_m, r_n)$ is the number of those cylinders that have nonempty intersection with Ω_A. Assume that each row of A contains at least one 1. Since the number of admissible paths of length n that begin with i and end with j is the entry a_{ij}^n of A^n. Then the number of nonempty cylinders of rank n in Ω_A is $\sum_{i,j=0}^{N-1} a_{ij}^n < C \cdot \|A^n\|$ for some constant C. On the other hand, $\sum\limits_{i,j=0}^{N-1} a_{ij}^n > c\|A^n\|$ for another constant $c > 0$. We need transitivity here to guarantee that this cylinders are non-empty. Because all numbers a_{ij}^n are nonnegative and hence the left-hand side is the norm $\sum_{i,j=0}^{N-1} a_{ij}^n$ of A^n, which is equivalent to the usual norm because all norms on $\mathbb{R}^{N^2}$ are equivalent. Thus, we have

$$S_{d_\lambda}(\sigma_A, r_m, n) = \sum_{i,j=0}^{N-1} a_{ij}^{n+2m}$$

and

$$h_{top}(\sigma_A) = \lim_{m\to\infty} \lim_{n\to\infty} \frac{1}{n} \log\|A^{n+2m}\| = \lim_{n\to\infty} \frac{1}{n} \log\|A^n\| = \log|\lambda_A^{\max}|.$$

$\square$

Proposition 5.18. *Let $F_L : \mathbb{T}^2 \to \mathbb{T}^2$ be given by*

$$F_L(x, y) = (2x + y, x + y) \mod 1$$

Then $h_{top}(F_L) = \log \frac{3+\sqrt{5}}{2}$.

Proof. We note that $h(F_L) \leq h(\sigma_A)$ where

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

We note $\lambda_A^{\max} = \frac{3+\sqrt{5}}{2}$. Hence

$$h_{top}(F_L) \leq \log \frac{3+\sqrt{5}}{2}.$$

Next we claim that these of n-periodic points of F_L is $(n, \frac{1}{4})$ separated set for any $n \in \mathbb{N}$. Then since the cardinality of n-periodic points of F_L is

$$P_n(F_L) = \left(\frac{3+\sqrt{5}}{2}\right)^n + \left(\frac{3+\sqrt{5}}{2}\right)^{-n} - 2.$$

Then this would show that

$$h_{top}(F_L) \geq \log \frac{3+\sqrt{5}}{2}.$$

□